

Kinetic-MHD stability of virtually collisionless plasmas

Présentée le 27 août 2020

à la Faculté des sciences de base
SPC - Théorie
Programme doctoral en physique

pour l'obtention du grade de Docteur ès Sciences

par

Samuel LANTHALER

Acceptée sur proposition du jury

Prof. F. Mila, président du jury
Dr J. Graves, directeur de thèse
Prof. F. Zonca, rapporteur
Dr X. Garbet, rapporteur
Dr H. Reimerdes, rapporteur

Abstract

The stability of pressure driven modes such as the 1/1 internal kink is known to depend sensitively on a multitude of physical effects such as toroidal rotation, kinetic effects due to thermal and suprathermal particle species and finite Larmor radius effects. Presently available models do not take into account these combined effects in a consistent way. This thesis presents the derivation of a novel kinetic magnetohydrodynamic (MHD) model utilizing a kinetic pressure closure which incorporates all of these physical mechanisms and can in particular be used to study the interplay of important centrifugal and kinetic effects in strongly rotating plasmas. The kinetic-MHD model is based on an original derivation of a consistent set of guiding-centre equations allowing for sonic flow. Important higher-order Larmor radius corrections to the guiding-centre coordinates, which are conventionally discarded, are discussed in detail for two applications: The first application concerns neutral beam injection (NBI) heating. It is shown that higher-order (Baños drift) corrections affect the expected resonances of particles with resonant magnetic perturbations (RMP), as well as the estimated NBI driven current in slowing-down simulations in a MAST-like equilibrium by up to 8%. As a second application, the full expression for the gyroviscous contribution to the pressure tensor is obtained from guiding-centre theory. Higher-order guiding-centre corrections are shown to lead to a non-circular Larmor motion of the particle around its guiding-centre which result in off-diagonal components of the pressure tensor. The derived expression for the pressure tensor in terms of the guiding-centre distribution function is used to formulate a consistent linear kinetic-MHD model with kinetic closure for the pressure. The proposed kinetic-MHD model allows for strong flows and includes centrifugal as well as diamagnetic flows. The model also includes a drift-kinetic form of the quasi-neutrality equation, and allows the effects of a parallel electric field on global MHD modes to be studied self-consistently. Pressure closure of the kinetic-MHD model is obtained from a solution of the guiding-centre equations, thus taking into account finite orbit-width effects and particle-wave interactions such as precession resonance. The benefits of the pressure closure approach over an approach following current-closure are discussed. It is shown that due to several convenient cancellations, the pressure closure approach can be based on first-order guiding-centre equations while an equivalent model formulated in terms of current closure would require second-order corrections to be retained. Thus, the benefits and the efficiency of a formulation of kinetic-MHD models with pressure closure over alternative models based on current closure are demonstrated.

Keywords: MHD stability, internal kink, guiding-centre theory, strong rotation, kinetic-MHD, finite Larmor-radius, Baños drift, diamagnetic flow

Zusammenfassung

Es ist bekannt, dass die Stabilität von druckgetriebenen Modi wie dem 1/1 “Internal Kink” empfindlich von mehreren physikalischen Mechanismen abhängt, einschliesslich der Toroidalrotation, kinetischer Effekte bedingt durch thermische und superthermische Partikelarten und endliche Larmor-Radius-Effekte. Die verfügbaren Modelle berücksichtigen diese kombinierten Effekte nicht in konsistenter Art und Weise. Diese Dissertation präsentiert die Herleitung eines neuartigen kinetischen Magnetohydrodynamik (MHD)-Modells unter Verwendung einer kinetischen Druck-Schliessung der Impulsgleichungen, welches all diese physikalischen Effekte berücksichtigt und zur Untersuchung des Zusammenspiels wichtiger zentrifugaler und kinetischer Effekte in rotierenden Plasmen verwendet werden kann. Das kinetische MHD-Modell basiert auf einer neuen Herleitung konsistenter Gleichungen für die Leitzentren (guiding-centres) mit Flussgeschwindigkeiten in Höhe der Schallgeschwindigkeit. Wichtige Larmor-Radius-Korrekturen höherer Ordnung werden im Detail in zwei Anwendungen diskutiert: Die erste Anwendung betrifft die Neutrale-Partikelstrahl-Injektion (NBI). Es wird gezeigt, dass (Baños-Drift) Korrekturen die erwarteten Resonanzen von Partikeln mit resonanten magnetischen Störungen (RMP) beeinflussen, sowie den geschätzten NBI-induzierten Strom in einem MAST-ähnlichen Gleichgewicht um bis zu 8% ändern können. Als zweite Anwendung wird der gyroviskose Beitrag zum Drucktensor aus der Leitzentrumstheorie hergeleitet. Korrekturen höherer Ordnung führen zu Abweichungen von einer kreisförmigen Larmor-Bewegung des Partikels um sein Leitzentrum, welche in nicht-diagonalen Komponenten des Drucktensors resultieren. Der hergeleitete Ausdruck für den Drucktensor wird verwendet, um ein konsistentes, lineares kinetisches MHD-Modell mit kinetischer Druckschliessung zu formulieren. Das Modell berücksichtigt sowohl zentrifugale als auch diamagnetische Effekte. Eine drift-kinetische Quasi-Neutralitätsgleichung ermöglicht es, die Auswirkungen eines parallelen elektrischen Feldes auf globale MHD-Modi zu untersuchen. Druckschliessung der Impulsgleichung wird durch die Lösung der Leitzentren-Gleichungen unter Berücksichtigung endlicher Bahnbreiteneffekte (finite orbit width) und Partikel-Wellen-Wechselwirkungen, wie etwa Präzessionsresonanzen, erreicht. Die Vorteile des Druckschliessungsansatzes gegenüber einer Stromschliessung werden diskutiert. Es wird gezeigt, dass unter dem Druckschliessungsansatz Leitzentren-Gleichungen erster Ordnung ausreichend sind, wohingegen ein vergleichbares Modell basierend auf einer Stromschliessung Korrekturen zweiter Ordnung erfordern würde. Dies zeigt die Vorteile und die Effizienz einer Formulierung des kinetischen MHD-Modells mit Druckschliessung der Impulsgleichung gegenüber alternativen Modellen basierend auf einer Stromschliessung auf.

Keywords: MHD-Stabilität, Internal-Kink, Leitzentrentheorie, Toroidalrotation, kinetische MHD, finiter Larmor-Radius, Baños-Drift, diamagnetische Strömung

Contents

Abstract (English/Deutsch)	iii
List of figures	ix
List of tables	xiii
1 Ideal MHD and beyond	1
1.1 Introduction	1
1.2 Ideal MHD	4
1.2.1 Rotating equilibrium	6
1.2.2 Straight field-line coordinates	8
1.2.3 Linear stability – Plasma displacement	9
1.2.4 Parallel and perpendicular dynamics	11
1.3 Basic formulation of kinetic-MHD	13
1.4 The 1/1 internal kink	15
1.4.1 Centrifugal effects	18
1.4.2 Kinetic effects	19
1.5 Summary and contributions of this thesis	22
2 Guiding-centre theory	25
2.1 Introduction	25
2.2 Lagrangian mechanics and Lie perturbation theory	28
2.2.1 Variational principles and Lagrangian one-form	29
2.3 Guiding-centre theory in co-moving frame	34
2.3.1 Phase-space coordinates	34
2.3.2 Summary of guiding-centre transformation	36
2.4 Derivation of guiding-centre transformation	41
2.4.1 Phase-space transformations induced by a vector field	41
2.4.2 Order-by-order analysis	45
2.4.3 Collection of useful identities	55
2.5 Summary	57

Contents

3 Application of higher-order guiding-centre corrections to full-<i>f</i> calculation	59
3.1 Introduction	59
3.2 Guiding-centre theory	61
3.2.1 Introduction	61
3.2.2 Equations of motion	62
3.2.3 Choice of guiding-centre variables	63
3.2.4 Guiding-centre push-forward	65
3.2.5 Discussion	66
3.2.6 Practical implications	68
3.3 Single particle dynamics	70
3.4 Application to slowing-down simulations	72
3.4.1 Corrections for strong anisotropy (NBI)	73
3.4.2 Unbalanced NBI in MAST	75
3.4.3 Unbalanced NBI losses due to RMP	76
3.4.4 Balanced NBI with RMP	77
3.4.5 Fusion α particle loss in DEMO	78
3.5 Discussion	80
3.6 Summary and conclusions	81
4 Kinetic-MHD	83
4.1 Introduction	83
4.2 Guiding-centre equations in co-moving frame	87
4.3 Pressure in guiding-centre coordinates	89
4.3.1 Intuitive picture in slab geometry	89
4.3.2 General expression for pressure	92
4.4 A Kinetic-MHD model for strong flows	97
4.4.1 Nonlinear formulation including diamagnetic effects	97
4.4.2 Equilibrium	100
4.4.3 Linear Perturbation	103
4.4.4 Efficiency of kinetic-MHD pressure closure	111
4.4.5 Linear kinetic-MHD model incl. centrifugal and diamagnetic effects . .	112
4.5 Summary	117
5 Final remarks and perspectives	119
5.1 Summary and conclusions	119
5.2 Future perspectives	122
5.2.1 Analytic considerations	122
5.2.2 Numerical implementation	123
A Banos drift in purely sheared background	133
A.1 Introduction	133
A.2 Derivation from Lie perturbation methods	134

B Detailed derivation of gyroviscous tensor components	137
B.1 FLR correction (I) – evaluation of eq. (4.13)	137
B.2 FLR correction (II) – evaluation of eq. (4.14)	137
B.3 FLR correction (III) – evaluation of eq. (4.15)	142
B.4 FLR correction (IV) – evaluation of eq. (4.16)	143
B.5 Identities involving moving frame	144
Bibliography	157
Acknowledgements	159
Curriculum Vitae	161

List of Figures

1.1	A schematic illustration of a tokamak. <i>Courtesy of EUROfusion</i>	2
1.2	Domain of validity of ideal MHD (blue) and conditions found in fusion relevant plasmas (orange). Based on the derivation of ideal MHD from the kinetic Boltzmann-Maxwell equations in [Fre14]. The relevant quantities for the validity of the ideal MHD model are the particle density n , and the temperature T	6
1.3	Plasma profiles $p(\rho)$ and $q(\rho)$ required to compute an equilibrium, in the absence of rotation $\Omega \equiv 0$.	8
1.4	Flux aligned equilibria, (a) obtained from the VMEC code and (b) the corresponding transformed straight field-line coordinates. Computed from the plasma profiles shown in Figure 1.3.	9
1.5	Internal kink mode structure: Shown are the dominant (m, n) -Fourier components $\rho\xi_m^\rho$ of the radial displacement ξ^ρ , obtained for the straight field-line equilibrium of Figure 1.4, with toroidal mode number $n = 1$. Superimposed is the q -profile (safety factor). Visible is the dominant $m = n = 1$ mode structure with a sharp gradient when $q = m/n = 1$, where the field-line bending stabilisation is weakest. Also visible is the $m = 2, n = 1$ side-band, which has a sharp change in gradient only at the $q = 2$ surface, consistent with $m/n = 2/1$ field-line bending stabilization.	16
1.6	Internal kink: Comparison of growth rates vs. poloidal β , obtained from VENUS-MHD vs. analytic form due to Bussac, for equilibrium with profiles shown in Figure 1.3 and scan over different scaled pressure profiles. The internal kink mode structure of Figure 1.5 corresponds to the first unstable value of β_p .	17
3.1	$B_{ }^*$ for a H^+ ion at energy $E = 10\text{ keV}$, parallel pitch $\lambda = 0.5$. Shown are approximations expanded at different orders in ϵ_B for the axisymmetric MAST-like equilibrium considered in section 3.4.1, along the midplane ($Z = 0$)	66
3.2	Physical particles sharing the same guiding-centre are initialized according to conventional (left) and higher-order (right) expressions (initialization near $y = z = 0$). Dots indicate snapshots of the position of these markers at several fixed points in time during their evolution. The lack of first-order corrections in the parallel velocity at the initialization of full Lorentzian markers (left) causes distortions, as well as an average lag behind the guiding-centre (missing Baños drift correction).	69

List of Figures

3.3 Poloidal projection of representative guiding-centre orbits initialized at different energies and evolved using higher-order equations of motion ('high') and conventional equations of motion ('low'). The difference between the curves 'high' and 'low' are due to the $O(\epsilon_B^2)$ drifts, which arise from the additional term $-\epsilon_B m \mu / q^2 \bar{R}$ in the higher-order Lagrangian (3.4).	71
3.4 Energy dependence ($E = 10\text{ keV}$ to $E = 120\text{ keV}$) of characteristic orbit quantities obtained from higher-order ('high') and conventional ('low') guiding-centre equations of motion. Shown are the bounce-averaged toroidal drift precession $\langle \dot{\phi} \rangle$ (left) and the bounce-averaged value of the normalized poloidal flux $\langle \Psi_N \rangle = \langle \Psi \rangle / \Psi_{\text{edge}}$ along the orbit. The difference between the curves 'high' and 'low' are due to the $O(\epsilon_B^2)$ drifts, which arise from the additional \bar{R} -term in the higher-order Lagrangian (3.4).	72
3.5 Deposition rate of NBI markers in terms of parallel guiding-centre pitch variable λ . A comparison is made between the conventional model and the model including the Baños drift correction to the parallel guiding-centre velocity (3.23).	74
3.6 Profiles determining the MAST-like equilibrium considered in this section.	75
3.7 Comparison of current profiles as computed from the slowing-down simulation for the NBI injection models considered in this section for a MAST-like equilibrium (bifurcated state).	75
3.8 Model of RMP perturbation coils in MAST (5.6 kA). The colors in (a) indicate the direction of current in the coils, corresponding to the sign of the normal component of the magnetic perturbation ($\delta B_r > 0$ (red), $\delta B_r < 0$ (blue), no current in coil (black)).	76
3.9 Comparison of predicted NBI heat fluxes through the LCFS computed in presence of RMP. Compared are two NBI models, one neglecting and one including the Baños drift correction in the initialisation of guiding-centre distributions. For the loss distribution in ϕ the results are plotted for $\phi \in [0, 2\pi/3]$, taking into account the $n = 3$ symmetry of the resonant magnetic field perturbations. Subfigure (a) shows the temporal evolution of the lost / injected power over the simulation. A steady state is reached after about 0.06 s. Subfigure (b) shows the local heat flux at steady state as a function of toroidal angle ϕ . The solid lines show the heat fluxes obtained for the two considered models in presence of RMP. For comparison, the pale lines correspond to a simulation in the absence of RMP, i.e. axisymmetric background.	77
3.10 Comparison of models including/neglecting the Baños drift correction (3.23) for (hypothetical) balanced NBI in a MAST-like equilibrium, in the presence of RMP.	79
4.1 Assumed slab geometry.	89
4.2 Particle motion in slab geometry with $\nabla E_2 \neq 0$: motion in lab frame with $E \times B$ drift clearly visible (left). Motion in frame co-moving with the guiding-centre (right) clearly shows deviation from circular motion due to ∇E_2 . In this figure, epsilon stands for $\epsilon \equiv \nabla \mathbf{u}_E / \Omega_c = \nabla E_2 / B \Omega_c$	91

4.3 Trajectories intersecting $\mathbf{r} = 0$ and corresponding anisotropic velocity distribution.	91
4.4 Terms related to off-diagonal components Π_{12} and $\Pi_{11} - \Pi_{22}$ as function of gyroangle ζ at fixed \mathbf{r} (solid), and their gyroaverages (dashed). Here: $v_x \equiv v_1$, $v_y \equiv v_2$.	92
4.5 Two-step transformation from physical particle (phy) to gyro-centre (gy) coordinates via guiding-centre (gc) coordinates.	104
5.1 Flow chart of an implementation of the linear kinetic-MHD model.	124
5.2 Work flow of abstract implementation of Galerkin discretization. The choice of model determines the functional form of the operator coefficients, based upon which the discretization can be carried out uniformly. This allows new models to be added to the code with relative ease.	129
5.3 Radial mesh-packing based on q-profile: Specification of the list of values $q = 1, 2$ leads to more densely packed radial mesh points near the corresponding q -values, and enables higher resolution of the singular layers in ideal MHD. Higher-order radial basis function (3rd order B-splines) (top), lower-order radial basis functions (2nd order B-splines) (middle) and corresponding q -profile (bottom).	130
A.1 Temporal evolution of x , y , z coordinates for particle with initialization (A.1). The parameters were set as $E = 10\text{keV}$, $\tau = 1$, $B_0 = 0.5\text{T}$. The Baños drift in the z -direction is clearly visible.	134
A.2 Temporal evolution of $v_{ }$ and the z -coordinate for the same particle of figure A.1. Indicated in red are relevant expressions obtained from the Lie perturbation theory.	135

List of Tables

3.1 Comparison of computed total heat flux to LCFS, as predicted by guiding-centre approximation at different orders in ϵ_B for fusion α -particles in DEMO. (A) was based on Lagrangian \mathcal{L}_0 , marker weights $w \propto B$, (B) based on Lagrangian \mathcal{L}_0 , marker weight $w \propto B_{\parallel}^*$, (C) based on Lagrangian \mathcal{L} , weight $w \propto B_{\parallel}^*$. Shown are the heat fluxes to the LCFS after saturation. The uncertainty in these values is estimated as the standard deviation after saturation of the fluctuations about the mean value (numerical/Monte-Carlo noise).	80
--	----

1 Ideal MHD and beyond

1.1 Introduction

The central goal of the continued international research effort in fusion-oriented plasma physics is the stable confinement of plasmas at temperatures and pressures at which fusion reactions can take place [Fre07]. A plasma is an electrically conducting, neutral gas of electrons and ions, with a very low electrical resistivity. The best candidate for a fusion reaction achievable on Earth is the deuterium-tritium (DT) reaction, which occurs at temperatures of around 150 million degrees Celsius [OKWZ16]. One of the most promising candidate devices for the confinement of plasmas at these temperatures is the tokamak: A device in which external coils generate strong magnetic fields to keep the plasma magnetically confined, schematically shown in Figure 5.1.

It has been proven experimentally that the temperatures and pressures required for fusion reactions to take place can be created and sustained in tokamak devices. In 1991, a large-scale DT experiment conducted at the Joint European Torus (JET) experiment in Culham, UK, achieved up to 1.7 MW of fusion power production in a short pulse of around 2 s [Reb92]. A later DT campaign in JET, in 1997, achieved over 16 MW in fusion power, for about 1 s [KGG⁺99]. A longer duration of sustained fusion power of over 4 MW over a duration of over 5 s in JET are reported in [J⁺99]. Despite these experimental achievements, the ultimate goal of a commercial use of fusion for electricity generation has not been attained at present.

The future success of this research programme will depend in part on a better understanding of the complicated physical mechanisms present in magnetically confined plasmas at such high temperatures [SCM⁺07, FBC⁺16]. New physics insights may result in more efficient heating [GSL⁺07], better confinement of the plasma with reduced transport of heat away from the hot core [PMP⁺07], and at the same time allow stable operation over long time-scales and free of disruptive instabilities [HWB⁺07]. While it has been shown that, in principle, conditions suitable for fusion can be achieved experimentally, the exciting next question is whether this can be done in a sufficiently efficient manner such that the energy required for the operation of the tokamak is outweighed by the energy set free in the fusion reactions taking place in the

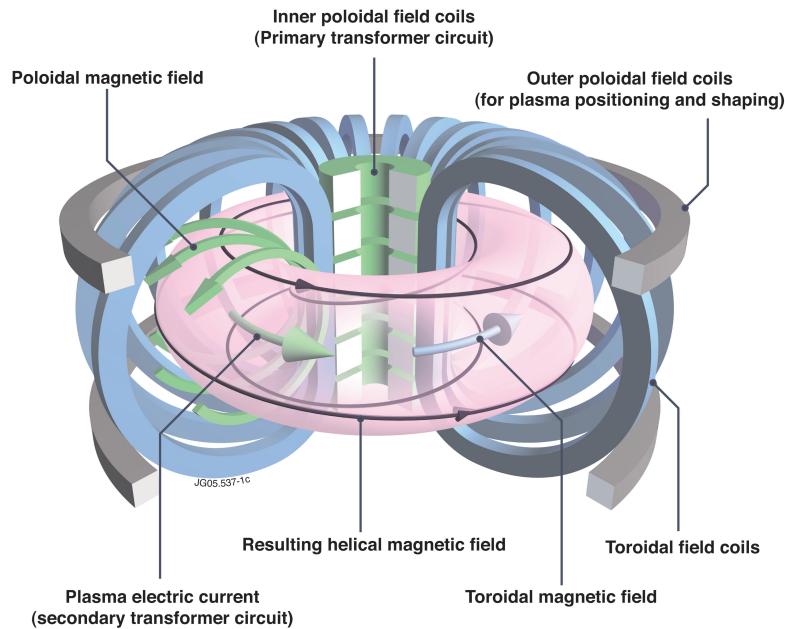


Figure 1.1 – A schematic illustration of a tokamak. *Courtesy of EUROfusion*

plasma core.

As already noted above, one of the most successful approaches to the confinement of high-temperature plasmas is the tokamak configuration. In a tokamak (cp. Figure 1.1), a strong toroidal field is generated by currents running through external coils. Additional transformer coils (the central solenoid) are used to induce a toroidal current in the plasma. This toroidal current serves two objectives. Firstly, it is used to heat the plasma through Ohmic dissipation. Secondly, this plasma current induces a poloidal magnetic field in addition to the externally applied toroidal field. This means that the magnetic field in a tokamak is not purely toroidal, but instead winds helically around the machine, and is designed to fill out well-defined torus-shaped surfaces, called flux surfaces.

The enforced axisymmetry of the plasma in a tokamak configuration makes such devices not only easier to build, but also simplifies the theoretical understanding of the physics relevant for tokamaks, compared to other magnetic confinement configurations without such symmetry. The toroidal symmetry in a tokamak implies that the canonical toroidal momentum is a constant of charged particle motion. The presence of such a conserved quantity has far-ranging consequences for the confinement of particles in a tokamak configuration. An important aspect in understanding and predicting the enhanced particle transport in real-world experiments are thus deviations from axisymmetry, allowing the conservation of toroidal momentum to be broken. One source of the breaking of axisymmetry is due the finite number of field coils, which in reality achieve to produce only an approximately axisymmetric toroidal field. Another source of axisymmetry breaking are fluctuations in the magnetic field, which occur in the presence of plasma instabilities. Both of these effects are sources of symmetry

breaking, and are found to markedly increase particle transport, and therefore deteriorate plasma confinement.

The fact that the magnetic field lines wrap around the device helically to form flux surfaces in a tokamak has many implications for the plasma dynamics: Since charged particles approximately follow magnetic field lines, the transport of energy and momentum in the direction parallel to the field lines is very large compared to the transport perpendicular to the flux surfaces. This means that temperature gradients along the field lines are very quickly damped out by diffusion. As the magnetic field lines densely fill out most flux surfaces, this has as a consequence that the temperature at equilibrium must be constant on each flux surface to a good approximation, and can therefore depend only on a radial variable perpendicular to the flux surfaces. By this argument, large gradients are permitted only across flux surfaces.

To obtain the best performance in such a plasma configuration, one seeks to maximize the pressure and temperature at the core of the plasma. For plasma heating of the core in high-performance plasmas, Ohmic heating due to the externally induced current is not sufficient, because the electrical resistivity of plasmas decreases rapidly as the temperature is increased [Fre07]. Additional heating mechanisms are used to heat the plasma core. Such heating schemes include the heating of ions by the generation of externally injected electromagnetic waves which resonate with the ions cyclotron motion (ion cyclotron resonance heating, or ICRH) and the injection of neutral beam ions (NBI). In both cases, this heating results in a fraction of highly energetic ions in the plasma. The objective is then to confine this suprathermal species of ions over time-scales that are sufficiently long for their energy to be transferred to the background ions and electrons of the plasma via collisions.

Invariably, the achievable performance even with such auxiliary heating mechanisms is limited by plasma instabilities, which typically occur when the large gradients perpendicular to the flux surfaces exceed a threshold value [Fre14]. In the worst case, such instabilities can lead to plasma disruptions which result in the complete loss of the plasma. These types of catastrophic instabilities set hard limits on the available parameter space at which safe operation of a tokamak is possible. A second type of instability does not result in the complete termination of the plasma, but rather results in a degradation of its performance [CPG⁺07]. Examples of such instabilities are internal instabilities such as the sawtooth and the fishbone instabilities. The effect of these instabilities is typically the sudden expulsion of hot ions in the plasma core, resulting in a collapse of the peaked pressure and density distribution at the centre, and consequently a less well performing plasma.

Various plasma models have been proposed to describe the physical mechanisms that play a role in such plasma instabilities. Such models are not only required to understand and avoid plasma instabilities. They can also be used to explore possible mechanisms to suppress these instabilities and potentially allow operation of high-performance plasma beyond the apparent operational limits, e.g. by a judicious choice of external heating sources to control the distribution of suprathermal particles inside the plasma, thereby exerting a stabilizing

influence on the global plasma dynamics [CGL⁺15, CIG⁺09, GCC⁺09].

In the following, we will review some of the most successful models which have been used to describe the stability of plasmas. Our review will focus in particular on the so-called 1/1 internal kink mode, which describes the linear stage of the previously mentioned nonlinear sawteeth and fishbone modes, which are very commonly observed instabilities in plasma experiments with auxiliary heating. Due to this fact, the understanding of the 1/1 internal kink instability is of central importance, and takes a prominent role in the research literature on global macroscopic instabilities. This magnetohydrodynamics (MHD)-mode is particularly interesting, because it is in fact only very weakly unstable, and thus its destabilization in a given experiment can depend sensitively on various physical effects; the plasma shaping, the plasma conductivity, toroidal plasma flow or the resonant interaction with suprathermal particles have all been shown to significantly influence the stability of the 1/1 internal kink.

It is hoped that the discussion in the remaining sections of this introduction will motivate the need for the derivation of a new model, which retains several physical effects that have previously only been studied separately. As will be seen, the complex interplay of various physical effects influencing the internal kink stability will require us to stretch a vast range of plasma physics – from ideal MHD and extended fluid models, to higher-order guiding-centre theory and even touching on gyrokinetics. This effort will result in the derivation of a kinetic-MHD model that is suitable for a numerical implementation and can be used to describe the stability of plasmas to internal kink modes, including the effects of strong toroidal flows, kinetic effects due to the resonant interaction with thermal and supra-thermal (virtually) collisionless ions and the effects of a parallel electric field in a single, unified model based on guiding-centre theory. In the process, we will gain a better understanding of the differences and links between kinetic-MHD and a purely kinetic description of the plasma, which would e.g. be obtained from a standard gyrokinetic approach, meaning the usual physics models behind gyrokinetic codes. In particular, we will argue based on a rigorous analysis from modern guiding-centre theory that under suitable approximations the kinetic-MHD formulation achieves a more efficient description of the kinetic effects for global macroscopic instabilities, compared to a fully kinetic formulation; as will be explained in detail in the following chapters, a fully kinetic description would require higher-order in Larmor radius effects to be retained to obtain a suitable description of macroscopic instabilities. To circumvent the need for a higher-order kinetic description, kinetic-MHD instead makes instead use of the additional fluid variables (bulk flow, current, magnetic field) as well as several convenient cancellations, thus obviating the need for a kinetic theory expanded to very high orders. The derivation of a novel kinetic-MHD model has been the main goal and is the central contribution of the present thesis.

1.2 Ideal MHD

The ideal MHD equations model the plasma as an electrically conducting, inviscid fluid with zero electric resistivity, according to the following equations:

$$\left\{ \begin{array}{ll} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, & \text{(conservation of mass)} \\ \rho \frac{d\mathbf{u}}{dt} + \nabla p = \mathbf{j} \times \mathbf{B}, & \text{(momentum equation)} \\ \frac{d}{dt}(p\rho^{-\gamma}) = 0, & \text{(adiabatic equation of state)} \\ \frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0, & \text{(Faraday's law)} \\ \mathbf{E} + \mathbf{u} \times \mathbf{B} = 0, & \text{(ideal Ohm's law)} \\ \nabla \cdot \mathbf{B} = 0, & \text{(1.6)} \\ \nabla \times \mathbf{B} = \mu_0 \mathbf{j}, & \text{(Ampère's law)} \end{array} \right. \quad \begin{array}{l} (1.1) \\ (1.2) \\ (1.3) \\ (1.4) \\ (1.5) \\ (1.6) \\ (1.7) \end{array}$$

where $d/dt := (\partial/\partial t + \mathbf{u} \cdot \nabla)$ denotes the convective derivative along the bulk velocity \mathbf{u} . These equations are expressed in terms of the plasma mass density ρ , the (isotropic) pressure p , the electric and magnetic fields \mathbf{E} , \mathbf{B} , and the current density \mathbf{j} . The constants μ_0 and γ denote the vacuum permeability and the adiabatic index (ideal gas), respectively.

The ideal MHD equations are used extensively in plasma physics, and are often successful in predicting the behaviour even of high-temperature plasmas. This is despite the fact that the derivation of the ideal MHD equations is strictly speaking invalid at fusion-relevant temperatures [Fre14]. Despite the difficulty in rigorously justifying their application to fusion plasmas, many breakthroughs in plasma physics have been achieved relying on the ideal MHD equations. Their practical success must therefore be attributed to the fact that in many situations of practical interest, the ideal MHD equations capture the essential aspects of the physics even of high-temperature plasmas. To quote Freidberg [Fre14, p.7, chapter 2]

[...] one of the basic assumptions used in the derivation [of ideal MHD], i.e., that the plasma is collision dominated, is *never* satisfied in plasmas of fusion interest. Even so, there is overwhelming empirical evidence that MHD provides an accurate description of macroscopic plasma behavior. This apparent good fortune is not a lucky coincidence but the consequence of some subtle physics; namely, those parts of the MHD model that are not valid because of violation of the collision dominated assumption are not directly involved in many if not most phenomena of interest. [...]

Freidberg's claim could be regarded as a claim of an "unreasonable effectiveness" of ideal MHD in plasma physics. While this is certainly true in general and for many strongly growing instabilities, it should be pointed out that there are important conditions which are frequently encountered in fusion experiments, for which ideal MHD is insufficient; This is in particular

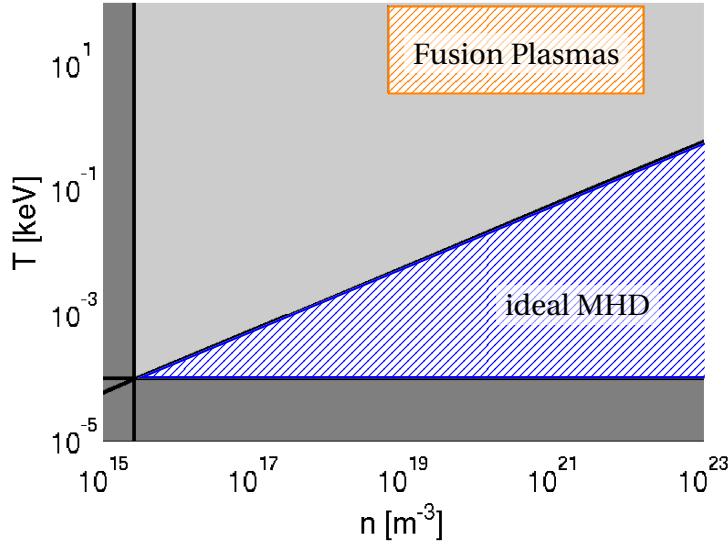


Figure 1.2 – Domain of validity of ideal MHD (blue) and conditions found in fusion relevant plasmas (orange). Based on the derivation of ideal MHD from the kinetic Boltzmann-Maxwell equations in [Fre14]. The relevant quantities for the validity of the ideal MHD model are the particle density n , and the temperature T .

true for the 1/1 internal kink mode. The main objective of this thesis is to rigorously derive a model that can be used to analyse the stability of plasmas under such conditions. It is important to emphasize at this point that among the set of equations (1.1)-(1.7), mainly the adiabatic equation of state and ideal Ohm's law are truly problematic from the physical point of view, in the context of macroscopic instabilities of plasmas [Fre14]. In particular, the assumption of an adiabatic equation of state is only valid at high collisionality, and therefore never applies to fusion relevant plasmas (cp. Figure 1.2).

1.2.1 Rotating equilibrium

The relevant equations describing ideal MHD equilibria are obtained by setting $\partial/\partial t = 0$ in the ideal MHD equations (1.1)-(1.7). We will usually refer to equilibrium quantities with a subscript 0. At equilibrium, the momentum equation (1.2) is written

$$\rho_0(\mathbf{v}_0 \cdot \nabla) \mathbf{v}_0 = -\nabla p_0 + \mathbf{j}_0 \times \mathbf{B}_0.$$

In an axisymmetric configuration, the B-field can be expressed in the form $\mathbf{B}_0 = \nabla\phi \times \nabla\psi + F(\psi)\nabla\phi$, in terms of the poloidal flux function ψ and a function $F = F(\psi)$, which is related to the toroidal component (upper indices) of the magnetic field by $B_0^\phi \equiv \mathbf{B}_0 \cdot \nabla\phi = F/R^2$, where R is the major radius. If ψ has a single extremum inside the plasma, then the flux surfaces are nested. In this situation, we can introduce coordinates (ψ, θ, ϕ) , where the poloidal magnetic flux ψ , measured from the core of the plasma outward, serves as a radial coordinate, ϕ is the

geometric toroidal angle, and θ is a choice of poloidal angle. The angles θ and ϕ parametrize the $\psi = \text{const}$ flux surfaces (for a more detailed description of flux coordinates, we refer e.g. to the thorough discussion in [HM03, Chapter 3.1-3.4]).

From an analysis of the particle motion, particles following equilibrium trajectories are confined to flux surfaces to leading order. In fact, it follows from the detailed derivation in chapter 2, that to leading order, the canonical toroidal momentum P_ϕ is given by $P_\phi \approx q\psi$ with q the particle charge; Hence the confinement of particles to a flux surface is strongly linked to the conservation of toroidal momentum in axisymmetry. In particular, we expect the equilibrium velocity to be tangent to flux surfaces, to good approximation. Following the model studied by Maschke and Perrin [MP80], and consistent with neoclassical theory [CBT87] from which it is expected that poloidal flows are strongly damped in axisymmetry, we furthermore assume that there are no poloidal components of the flow at equilibrium, and that the velocity can be written in the form

$$\mathbf{v}_0 = \Omega(\psi)R^2\nabla\phi.$$

Here R denotes the major radius, and $\Omega(\psi)$ is the toroidal rotation frequency, which may vary across different flux surfaces. If we furthermore assume that the temperature $T = T(\psi)$ is constant on flux surfaces, then under these assumptions, the following form of the Grad-Shafranov equation [GR58, LS57, Sha58] for ψ can be obtained, allowing for toroidal rotation (see e.g. [HdBK12]):

$$R^2\nabla\cdot\left(\frac{1}{R^2}\nabla\psi\right) = \frac{\partial P}{\partial\psi}\Big|_R - F\frac{dF}{d\psi}. \quad (1.8)$$

Here $P(\psi, R) = \bar{p}_0(\psi)\exp\left(\frac{m_i\Omega(\psi)^2R^2}{4T(\psi)}\right)$ is the pressure, provided the ideal gas law is assumed [MP80, CBF⁺15], and it is noted that due to centrifugal effects, the pressure varies on a flux surface. We recall that the function $F(\psi)$ is related to the toroidal component of the magnetic field via $\mathbf{B}_0 \cdot \nabla\phi = F/R^2$. As compared to the non-rotating case ($\Omega \equiv 0$), to determine the rotating equilibrium, one thus needs to specify in addition to the usual profiles, such as e.g. the pressure profile $\bar{p}_0(\psi)$ and the $F(\psi)$ -profile (which in practice is usually determined indirectly from either a specified current or ‘safety factor’ profile, see below), also the rotation profile $\Omega(\psi)^2/T(\psi)$.

In chapter 4, we will revisit the above form of the Grad-Shafranov equation from the kinetic point of view, based on guiding-centre theory. In particular, we will see that this form of the equilibrium relations, which here was obtained from ideal MHD, can also be derived to leading order from guiding-centre theory.

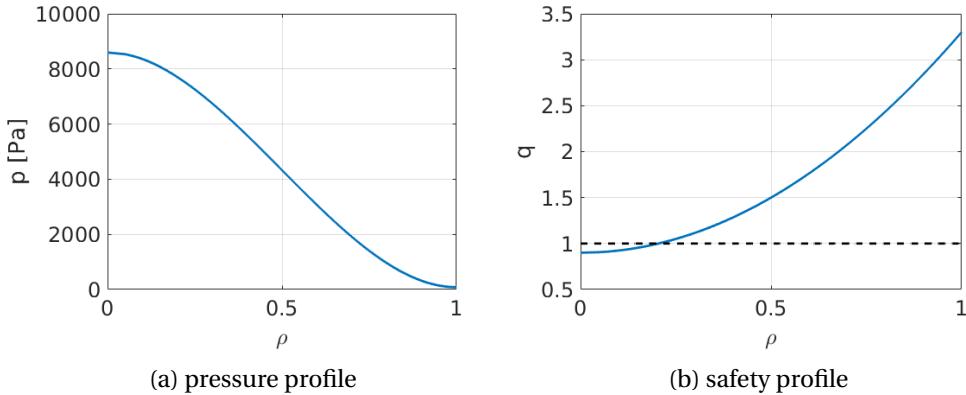


Figure 1.3 – Plasma profiles $p(\rho)$ and $q(\rho)$ required to compute an equilibrium, in the absence of rotation $\Omega \equiv 0$.

1.2.2 Straight field-line coordinates

In the description above, we have not specified the poloidal angle. For stability analysis, the preferred poloidal angle is the so-called straight-field line (SFL) angle θ_{SFL} , which is chosen such that the safety factor $q \equiv (\mathbf{B} \cdot \nabla \phi) / (\mathbf{B} \cdot \nabla \theta_{\text{SFL}})$ is a flux surface quantity, i.e. $q = q(\psi)$. As the name indicates, with this choice of poloidal angle, the field lines are straight, describing curves of the form $(\theta, \phi)(\tau) = (\theta_0 + \tau, \phi_0 + q\tau)$ when following a field line for time τ , on a fixed flux surface ψ . Given an arbitrary choice of a poloidal angle θ and the SFL angle θ_{SFL} on the flux surface, we clearly have

$$\mathbf{B} \cdot \nabla \theta = \frac{d\theta}{d\theta_{\text{SFL}}} (\mathbf{B} \cdot \nabla \theta_{\text{SFL}}).$$

Therefore, given an arbitrary angle θ , we can find the corresponding SFL angle, by solving

$$\frac{d\theta_{\text{SFL}}}{d\theta} = \frac{\mathbf{B} \cdot \nabla \theta_{\text{SFL}}}{\mathbf{B} \cdot \nabla \theta} = \frac{1}{q(\psi)} \frac{\mathbf{B} \cdot \nabla \phi}{\mathbf{B} \cdot \nabla \theta}, \quad (1.9)$$

where

$$q(\psi) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\mathbf{B} \cdot \nabla \phi}{\mathbf{B} \cdot \nabla \theta} d\theta,$$

is the flux surface average of the “local” $q_{\text{loc}}(\psi, \theta) \equiv (\mathbf{B} \cdot \nabla \phi) / (\mathbf{B} \cdot \nabla \theta)$. Clearly, θ_{SFL} is unique up to the arbitrary choice of the reference point at which $\theta_{\text{SFL}} = 0$. Given an equilibrium in flux coordinates generated by a MHD equilibrium code, such as VMEC [CH87], the first step in the numerical stability analysis is therefore often to map the equilibrium to straight field-line coordinates, which, as mentioned earlier, are more suitable for the stability analysis. Conversely, the equilibrium problem is not usually conveniently defined in terms of θ_{SFL} . For example, the VMEC equilibrium code employs a Fourier expansion of R and Z (vertical coordinate) in terms of θ , with a fixed number of Fourier modes. The freedom in the choice

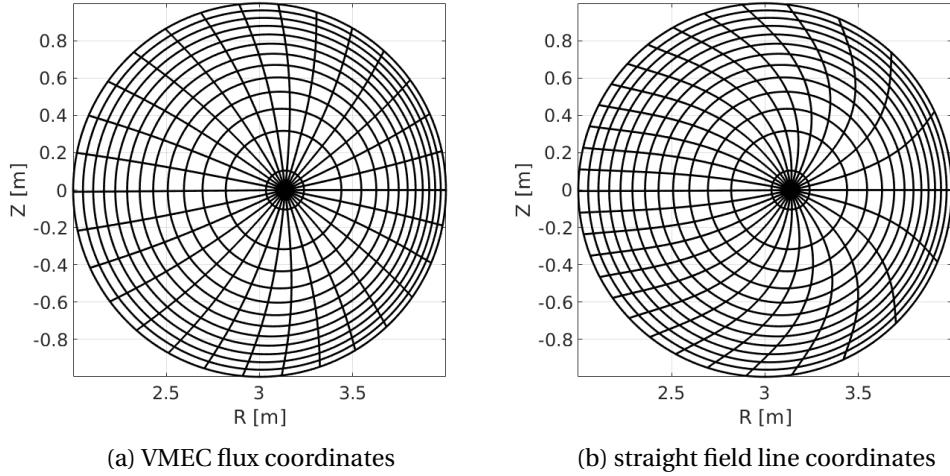


Figure 1.4 – Flux aligned equilibria, (a) obtained from the VMEC code and (b) the corresponding transformed straight field-line coordinates. Computed from the plasma profiles shown in Figure 1.3.

of θ is used to improve the accuracy of the Fourier representation of R and Z . The resulting VMEC poloidal angle θ differs significantly from θ_{SFL} , as is shown in Figure 1.3. It shows an example of an equilibrium generated by VMEC, corresponding to the profiles displayed in Figure 1.3, and the corresponding equilibrium mapped to SFL coordinates.

1.2.3 Linear stability – Plasma displacement

In ideal MHD, the linearization of the set of equations (1.1)-(1.7) can be very conveniently described in terms of a single vector field ξ , called the plasma displacement. In the non-rotating case ($\mathbf{v}_0 = 0$), the plasma displacement is simply given by $\partial \xi / \partial t = \delta \mathbf{v}$, where $\delta \mathbf{v}$ denotes the perturbed ideal MHD velocity [Fre14] in the Eulerian frame. The correct generalization of the displacement to the rotating case goes back to Frieman and Rotenberg [FR60], in which case one defines the plasma displacement via

$$\frac{\partial \xi}{\partial t} + \mathbf{v}_0 \cdot \nabla \xi = \delta \mathbf{v}.$$

It turns out that with this definition of ξ , all perturbed ideal MHD quantities can be entirely written in terms of ξ in a very compact manner. As a consequence, the linearisation of the ideal MHD equations can be fully expressed in terms of the relevant momentum equation, describing the motion the plasma displacement ξ . The resulting momentum equation for the evolution of ξ takes the form (cp. [GKP10, Chapter 12] for a thorough discussion)

$$\rho_0 \frac{\partial^2 \xi}{\partial t^2} + 2\rho_0 (\mathbf{v}_0 \cdot \nabla) \frac{\partial \xi}{\partial t} = \delta G(\xi), \quad (1.10)$$

Chapter 1. Ideal MHD and beyond

where

$$\delta \mathbf{G}(\boldsymbol{\xi}) = \delta \mathbf{F}(\boldsymbol{\xi}) + \nabla \cdot (\boldsymbol{\xi} \rho_0 \mathbf{v}_0 \cdot \nabla \mathbf{v}_0 - \rho_0 \mathbf{v}_0 \mathbf{v}_0 \cdot \nabla \boldsymbol{\xi}),$$

is the extended force operator, and $\delta \mathbf{F}$ is the conventional force operator for static ($\mathbf{v}_0 = 0$) plasmas [GP04, Fre14]:

$$\delta \mathbf{F}(\boldsymbol{\xi}) = \nabla(\Gamma p_0 \nabla \cdot \boldsymbol{\xi}) - \mathbf{B}_0 \times (\nabla \times \delta \mathbf{B}) + \nabla(\boldsymbol{\xi} \cdot \nabla p_0) + \mathbf{j}_0 \times \delta \mathbf{B}.$$

In these expressions, ρ_0 is the equilibrium mass density, p_0 is the equilibrium pressure, \mathbf{v}_0 is the background velocity, $\mathbf{j}_0 = \mu_0^{-1} \nabla \times \mathbf{B}_0$ is the current density corresponding to the equilibrium magnetic field \mathbf{B}_0 . The perturbed magnetic field $\delta \mathbf{B}$ is given by

$$\delta \mathbf{B} = \nabla \times (\boldsymbol{\xi} \times \mathbf{B}_0),$$

so that $\delta \mathbf{A} \equiv \boldsymbol{\xi} \times \mathbf{B}_0$ can be identified as the perturbed vector potential.

The formulation in terms of a plasma displacement is not only convenient because it reduces the number of equations to be analysed from the original 8 equations to 3 equations; **In the non-rotating case**, this formulation also makes the self-adjoint character of the equations manifest, which has important consequences on the spectra of ideal MHD equilibria. Indeed, it can be shown that the MHD force operator $\delta \mathbf{F}(\boldsymbol{\xi})$ is self-adjoint in the sense that for any two vector fields $\boldsymbol{\xi}, \boldsymbol{\eta}$, we have

$$\int \boldsymbol{\eta}^* \cdot \delta \mathbf{F}(\boldsymbol{\xi}) d^3x = \int \delta \mathbf{F}(\boldsymbol{\eta})^* \cdot \boldsymbol{\xi} d^3x,$$

where $\boldsymbol{\eta}^*$ denotes the complex conjugate of $\boldsymbol{\eta}$. Furthermore, in the non-rotating case, the left-hand side of the momentum equation simplifies to $\rho_0 \partial^2 \boldsymbol{\xi} / \partial t^2$. Making the ansatz $\boldsymbol{\xi}(x, t) = \hat{\boldsymbol{\xi}}(x) e^{i\omega t}$, the problem of ideal MHD stability in the non-rotating case thus leads to an eigenvalue problem for self-adjoint operators:

$$-\omega^2 \rho_0 \boldsymbol{\xi} = \delta \mathbf{F}(\boldsymbol{\xi}). \quad (1.11)$$

Since the spectrum of a self-adjoint operator is purely real, we can distinguish a unstable plasma from an stable one, by considering whether the spectrum of $\delta \mathbf{F}$ contains positive eigenvalues, or not. If $\delta \mathbf{F}(\boldsymbol{\xi})$ possesses a positive eigenvalue $\Lambda > 0$, then equation (1.11) has a solution with $-\omega^2 = \Lambda > 0$, implying that $\omega = \pm \sqrt{-\Lambda} \equiv \pm i\lambda$ is a complex number (here $\lambda \equiv \sqrt{\Lambda} > 0$), and hence there exists a normal mode solution $\boldsymbol{\xi} = \hat{\boldsymbol{\xi}} e^{i\omega t} = \hat{\boldsymbol{\xi}} e^{\lambda t}$ of (1.11) which grows exponentially like $\sim e^{\lambda t}$.

Due to the self-adjoint nature of $\delta \mathbf{F}$, the most unstable eigenvalue can be determined by a variational principle [Fre14, Chapter 6.4]:

$$\omega^2 = \min_{\boldsymbol{\xi}} \frac{\delta W(\boldsymbol{\xi})}{\delta K(\boldsymbol{\xi})}, \quad \delta W(\boldsymbol{\xi}) = \int \boldsymbol{\xi}^* \cdot \delta \mathbf{F}(\boldsymbol{\xi}) d^3x, \quad \delta K(\boldsymbol{\xi}) = \int \rho_0 (\boldsymbol{\xi}^* \cdot \boldsymbol{\xi}) d^3x, \quad (1.12)$$

leading to the following “**energy principle**”, which roughly states the following (for a more precise statement, see [Fre14, p.262]):

A static equilibrium ($\mathbf{v}_0 = 0$) is stable if and only if $\delta W(\xi) > 0$ for all plasma displacements ξ .

An equilibrium for which there exists a displacement ξ , for which $\delta W(\xi) = 0$, is often referred to as *marginally stable*. It is noted here that the rotating case unfortunately does not exhibit the convenient self-adjoint properties described above.

1.2.4 Parallel and perpendicular dynamics

In non-rotating plasmas the linear momentum equation for ξ can be effectively split into parallel and perpendicular parts, as will presently be discussed. It turns out to be convenient to introduce $\beta \equiv B^{-2} \mathbf{B} \cdot \xi = B^{-1} \xi_{||}$, such that $\nabla \cdot \xi_{||} = \nabla \cdot (\beta \mathbf{B}_0) = \mathbf{B}_0 \cdot \nabla \beta$. With this definition, we find from the momentum equation (1.10) (the growth rate $\lambda \equiv i\omega$ is considered below, rather than the frequency ω):

$$\lambda^2 \rho_0 \xi_{\perp} = \nabla_{\perp} \cdot (\xi_{\perp} \nabla p_0) + \nabla_{\perp} \cdot (\Gamma p_0 \nabla \cdot \xi_{\perp}) + \nabla_{\perp} \cdot (\Gamma p_0 \mathbf{B}_0 \cdot \nabla \beta) + [\delta(\mathbf{j} \times \mathbf{B})]_{\perp},$$

perpendicular to the field \mathbf{B}_0 , and

$$\lambda^2 \rho_0 \beta = B_0^{-2} \mathbf{B}_0 \cdot \nabla (\xi_{\perp} \nabla p_0 + \Gamma p_0 \nabla \cdot \xi_{\perp}) + \Gamma B_0^{-2} p_0 \mathbf{B}_0 \cdot \nabla (\mathbf{B}_0 \cdot \nabla \beta) + B_0^{-2} [\delta(\mathbf{j} \times \mathbf{B})]_{||},$$

parallel to the field lines. It can be shown that the first and last terms on the right-hand side of the parallel equation cancel,¹ leaving us with

$$\lambda^2 \rho_0 \beta = \Gamma B_0^{-2} p_0 \mathbf{B}_0 \cdot \nabla (\nabla \cdot \xi_{\perp} + \mathbf{B}_0 \cdot \nabla \beta). \quad (1.13)$$

Comparing terms on the left- and right-hand sides of this equality, we now note that they are of relative order of magnitude:

$$\frac{\lambda^2}{\Gamma p_0 / R_0^2 \rho_0} = \frac{\lambda^2}{\omega_s^2},$$

where $\omega_s^2 = \Gamma p_0 / R_0^2 \rho_0$ is the sound frequency. For many ideal MHD modes of interest, such as the 1/1 internal kink, we have $\lambda^2 \ll \omega_s^2$, so that for such modes, the left-hand side term of (1.14) is much smaller than the terms on the right-hand side. This means that we must have

$$B^{-2} \mathbf{B} \cdot \nabla (\nabla \cdot \xi_{\perp} + \mathbf{B} \cdot \nabla \beta) \approx 0, \quad (1.14)$$

and therefore to good accuracy (corrections of order λ^2 / ω_s^2), we have $\mathbf{B} \cdot \nabla \beta \approx -\nabla \cdot \xi_{\perp}$, or equivalently $\nabla \cdot \xi \approx 0$. Under this incompressibility approximation, the perpendicular dynamics

¹This can be seen easily by dotting the un-linearized momentum equation with the total field $\mathbf{B} = \mathbf{B}_0 + \delta \mathbf{B}$, noting that $(\mathbf{j} \times \mathbf{B}) \cdot \mathbf{B} = 0$, then linearizing.

is then governed by the “incompressible MHD equations”:

$$\lambda^2 \rho_0 \xi_{\perp} = \nabla_{\perp} (\xi_{\perp} \cdot \nabla p_0) + [\delta(\mathbf{j} \times \mathbf{B})]_{\perp}, \quad (1.15)$$

while the parallel component $\xi_{\parallel} = B_0^{-1} \beta$ can be solved for from $\mathbf{B} \cdot \nabla \beta \approx -\nabla \cdot \xi_{\perp}$.

The above analysis shows that the threshold for stability within the ideal MHD framework, and crucially *in the absence of toroidal rotation*, is entirely determined by the incompressible MHD equations (1.15) for ξ_{\perp} , which could (formally) also have been obtained from our momentum equation by formally taking $\Gamma = 0$, $\xi_{\parallel} = 0$. The important difference between the incompressible model (neglecting parallel dynamics altogether) and the full ideal MHD momentum equation is that even under the approximation $\nabla \cdot \xi = 0$, the compressibility term gives us additional information about the parallel component $\xi_{\parallel} \neq 0$. It is clear that under this approximation ξ_{\parallel} will not contribute to δW . Instead, the parallel displacement will contribute to the inertia δK . In fact, for the internal kink mode, it can be shown that ξ_{\parallel} leads to an inertial enhancement $\delta K(\xi) \approx \sqrt{1+2q_s^2}\delta K(\xi_{\perp}) = \sqrt{3}\delta K(\xi_{\perp})$, where $q_s = 1 = 1/1$ is the q -value at the rational surface, where the majority of the inertia originates [GGJ75, GW17]. Thus, this discussion shows that while the determination of stability or instability agree between the compressible and incompressible models, it is expected that the incompressible model overestimates the growth rate $\gamma_{\text{incomp}} \approx \sqrt{3}\gamma$ by a factor of $\sqrt{3} \approx 1.4$ for the internal kink mode compared to the growth rate γ determined by the full ideal MHD model. If one instead takes projections of the momentum equation for solving for the eigenvalue problem, ξ_{\parallel} enters the replaced momentum equation through the $\nabla \cdot \xi$ contribution to δp . To recover the correct growth rate one must eliminate ξ_{\parallel} in $\nabla \cdot \xi$ via the solution to the parallel component of the momentum equation, and thus allow for weak contributions to $\nabla \cdot \xi$.

The fact that the determination of instability for the compressible and incompressible MHD models agree may be seen as a more rigorous justification for Freidberg’s quote at the beginning of this chapter. As has been pointed out above, the essential reason is that the perpendicular dynamics, represented by ξ_{\perp} , are only very weakly coupled to the parallel dynamics ξ_{\parallel} in the ideal MHD model without toroidal flow. And the dynamics perpendicular to the field lines is well modeled by ideal MHD. This weak coupling is good news, since the parallel dynamics are not well represented in the ideal MHD model. Indeed, in the collisionless limit, particles are free-streaming along the field-lines, and the fluid assumption (which requires high-collisionality) completely breaks down parallel to the field lines. Thus, in high-temperature, low-collisional plasmas the mean-free path along field lines can be very large, allowing complicated wave-particle interactions to take place, and we cannot expect any fluid model to accurately represent such effects parallel to the field lines. Perpendicular to the field lines, the magnetic field serves to restrict the mean-free path of particles, which gives some justification for a fluid approach. If the perpendicular and parallel dynamics are only weakly coupled, then the inaccurate representation of the dynamics along the field lines may be of minor concern.

This argument of weak coupling of perpendicular and parallel dynamics breaks down in the strongly rotating case. As is clear from the momentum equation (1.10), the non-vanishing fluid velocity $\mathbf{v}_0 \neq 0$ introduces an additional, complicated coupling between parallel and perpendicular dynamics, both due to the non-vanishing advective term, and through the centrifugal and coriolis forces (additional terms in the extended force operator $\delta\mathbf{G}(\xi)$ as opposed to the conventional ideal MHD force operator $\delta\mathbf{F}(\xi)$). Due to this fact, it is expected that an improved treatment of the parallel dynamics will be particularly important for rotating plasmas.

1.3 Basic formulation of kinetic-MHD

In the collisionless limit, the dynamics of a (non-relativistic) plasma are well described by the (kinetic) Vlasov equations, describing the evolution of the particle distribution function $f_s(\mathbf{x}, \mathbf{v}, t)$ for each particle species s :

$$\frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \nabla f_s + \frac{q_s}{m_s} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \frac{\partial \mathbf{f}_s}{\partial \mathbf{v}} = 0, \quad (1.16)$$

coupled to the Maxwell equations to determine the electromagnetic fields \mathbf{E} , \mathbf{B} . To obtain the corresponding moment equations, we multiply by m_s , $m_s \mathbf{v}$ and integrate against $d\mathbf{v}$. This leads to the *exact moment equations* for each species s ,

$$\frac{\partial \rho_s}{\partial t} + \nabla \cdot (\rho_s \mathbf{u}_s) = 0, \quad (1.17)$$

$$\rho_s \left(\frac{\partial \mathbf{u}_s}{\partial t} + \mathbf{u}_s \cdot \nabla \mathbf{u}_s \right) = -\nabla \cdot \mathbf{P}_s + q_s n_s (\mathbf{E} + \mathbf{u}_s \times \mathbf{B}). \quad (1.18)$$

Here ρ_s is the mass density [kg/m^3], \mathbf{u}_s is the macroscopic fluid velocity [m/s], n_s is the number density [$1/\text{m}^3$], and \mathbf{P}_s is the pressure tensor [N/m^2], defined by

$$\mathbf{P}_s \equiv m_s \int (\mathbf{v} - \mathbf{u}_s) \otimes (\mathbf{v} - \mathbf{u}_s) f_s d\mathbf{v}. \quad (1.19)$$

These equations are exact, but they are not closed, since there is no equation to determine the pressure \mathbf{P}_s in terms of the other moments ρ_s and \mathbf{u}_s .

Summing equation (1.18) over all species, assuming quasi-neutrality $\sum_s q_s n_s = 0$ and introducing the mass density $\rho \equiv \sum_s m_s n_s$, and mass velocity \mathbf{u} by $\rho \mathbf{u} \equiv \sum_s m_s n_s \mathbf{u}_s$, we find the following basic form of a kinetic-MHD model:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (1.20)$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla \cdot \mathbf{P} + \mathbf{j} \times \mathbf{B} - \nabla \cdot \boldsymbol{\sigma}. \quad (1.21)$$

where the pressure is now given by $\mathbf{P} \equiv \sum_s \mathbf{P}_s$, and for each species, \mathbf{P}_s is obtained from

(1.19). Summation of the Lorentz force terms in (1.18) yields the familiar $\mathbf{j} \times \mathbf{B}$ force, since $\sum_s q_s n_s \mathbf{u}_s = \mathbf{j}$. Note that the summation over all species eliminates the explicit appearance of the electric field \mathbf{E} , thanks to the quasi-neutrality relation $\sum_s q_s n_s = 0$. This cancellation is good news, since the electromagnetic forces are formally large terms in the exact momentum equations (1.18). The last term in (1.21) represents a stress which arises due to the difference between the flow velocity \mathbf{u}_s of species s and the mass velocity \mathbf{u} :

$$\boldsymbol{\sigma} \equiv \sum_s m_s n_s (\mathbf{u}_s - \mathbf{u}) \otimes (\mathbf{u}_s - \mathbf{u}).$$

The ideal MHD model is obtained from (1.20) and (1.21), assuming that the inertia is dominated by a single bulk ion species i and making an adiabatic closure assumption, so that the total pressure $\sum_s \mathbf{P}_s$ can be written in the isotropic form $p \mathbf{I}$, in terms of a scalar pressure p . Note that in this case, we have $\boldsymbol{\sigma} \approx 0$, since $\rho \approx \rho_i$, $\mathbf{u}_i \approx \mathbf{u}$ and $m_e \ll m_i$. We will assume that the plasma inertia is dominated by a single species i , in the following. As already pointed out above, the main problem with the ideal MHD model as applied to fusion relevant plasmas lies in the adiabatic closure assumption. Without this assumption, we find instead the following basic form of a kinetic-MHD model (here $\rho = \rho_i$, $\mathbf{u} = \mathbf{u}_i$):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (1.22)$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla \cdot \mathbf{P} + \mathbf{j} \times \mathbf{B}. \quad (1.23)$$

where the pressure is given by $\mathbf{P} \equiv \sum_s \mathbf{P}_s$, and for each species, \mathbf{P}_s is obtained from (1.19). Note that even though the electric field has been eliminated from the momentum equation, its effects will still be felt in the evolution equation for the particle distributions f_s (1.16), and hence will have an indirect influence on the momentum equation via \mathbf{P}_s , which is determined from the kinetic closure relation (1.19). To leading order, an adequate closure for the current \mathbf{j} is obtained as in the ideal MHD model from ideal Ohm's law and the Maxwell equations (without displacement current), equations (1.4)-(1.7). In this approximation, the perpendicular velocity \mathbf{u}_\perp is given by the familiar $\mathbf{E} \times \mathbf{B}$ -velocity,

$$\mathbf{u}_\perp = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (1.24)$$

As a result of this discussion, we obtain the basic form of a kinetic-MHD model, comprising of the continuity equation (1.1), the momentum equation (1.23) (with kinetic closure), and the field equations (1.4), (1.5), (1.6) and (1.7). The adiabatic equation of state of the ideal MHD model (1.3) is now replaced by the kinetic closure relation (1.19). The latter requires solution of the kinetic equation (1.16).

In contrast to the ideal MHD model, a kinetic-MHD model allows for the tensorial nature of the pressure tensor. It goes back to the early work of Chew, Low and Goldberger [CGL56], that the pressure in plasmas at low collisionality should be of the so-called CGL form $\mathbf{P} \approx$

$p_{\parallel} \mathbf{b} \mathbf{b} + p_{\perp} (\mathbf{I} - \mathbf{b} \mathbf{b})$, though the focus of the work [CGL56] was not on a kinetic closure. This leading order form of the kinetic pressure tensor is an immediate consequence of guiding-centre theory (cp. chapter 4, where also higher-order, off-diagonal corrections will be derived). The first investigation of linear stability in a kinetic-MHD model dates back at least to the work of Rosenbluth and Rostocker [RR59], where the linearly perturbed kinetic equation was solved along unperturbed trajectories. The work [RR59] also first pointed out the benefits, in terms of a formal ordering in the Larmor radius, of a pressure closure in the momentum equation over a kinetic current closure, or indeed a direct kinetic formulation. We will return to this topic in chapter 4, where we derive a suitable kinetic-MHD model with pressure closure based on a rigorous expansion of the guiding-centre equations in the Larmor radius. A kinetic-MHD model based on a reduced, leading-order (“beads-on-a-string”) drift-kinetic equation to provide kinetic pressure closure was discussed by Kulsrud [Kul83]. The present work follows a very similar approach, but aims to consistently include higher-order effects such as guiding-centre drifts and finite-Larmor radius effects. The derivation of a suitable set of higher-order guiding-centre equations is carried out in detail in chapter 2. Kinetic-MHD models have in particular made an important and essential contribution to the understanding of the $m = 1/n = 1$ internal kink mode, which is a linear precursor for the often-observed sawtooth (and fishbone) instabilities in present day experimental tokamak devices.

The kinetic-MHD model described above differs from the ideal MHD model only in the way in which pressure closure is obtained. Due to this fact, the linearisation of this kinetic-MHD model can again be described in terms of a plasma displacement ξ , leading (in the static case $\mathbf{v}_0 = 0$) to a momentum equation of the form

$$-\omega^2 \rho_0 \xi = -\nabla \cdot \delta \mathbf{P} + \delta \mathbf{j} \times \mathbf{B}_0 + \mathbf{j}_0 \times \delta \mathbf{B}, \quad (1.25)$$

where in contrast to the momentum equation (1.11) in ideal MHD, the perturbed pressure $\delta \mathbf{P} = \delta \mathbf{P}(\xi; \omega)$ is now a tensor which depends on the mode frequency ω in addition to the displacement ξ (as well, potentially, as the perturbed parallel electric field). $\delta \mathbf{P}$ has to be solved for in terms of the perturbed distribution functions δf_s , of all kinetic species s , and is given by $\delta \mathbf{P} = \sum_s \delta \mathbf{P}_s$, where $\delta \mathbf{P}_s = m_s \int (\mathbf{v} \otimes \mathbf{v}) \delta f_s d\mathbf{v}$. One of the central problems of kinetic-MHD is the derivation of suitable (simplified) equations to determine the perturbed distribution functions δf_s , and an analysis of the complicated dependence of $\delta \mathbf{P}$ on ω .

1.4 The 1/1 internal kink

A particularly important instability, which describes the linear stage of sawteeth and fishbones, is a perturbation that can be described by a plasma displacement ξ with a dominant $m = n = 1$ mode structure $\sim \exp(im\theta + in\phi)$. A first analysis of the internal kink instability based on the ideal MHD model was undertaken by Shafranov [Sha70]. Employing an expansion of the linearized stability equations in the inverse aspect ratio ($\epsilon_a \sim a/R_0 \ll 1$), a minimisation of $\delta W = \delta W_0 + \epsilon_a^2 \delta W_2 + \dots$ was carried out for a straight cylindrical plasma

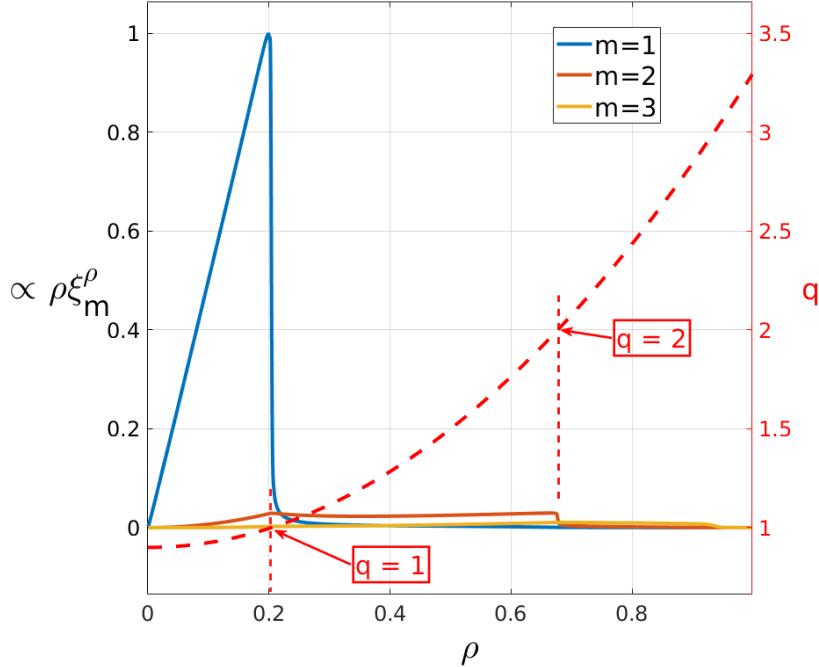


Figure 1.5 – Internal kink mode structure: Shown are the dominant (m, n) -Fourier components $\rho \xi_m^\rho$ of the radial displacement ξ^ρ , obtained for the straight field-line equilibrium of Figure 1.4, with toroidal mode number $n = 1$. Superimposed is the q -profile (safety factor). Visible is the dominant $m = n = 1$ mode structure with a sharp gradient when $q = m/n = 1$, where the field-line bending stabilisation is weakest. Also visible is the $m = 2, n = 1$ side-band, which has a sharp change in gradient only at the $q = 2$ surface, consistent with $m/n = 2/1$ field-line bending stabilization.

column (neglecting effects due to the toroidal bending of the torus). From such a derivation (see e.g. [Gra99, Chapter 2.4] for a more thorough discussion), one concludes that to lowest order the displacement must be incompressible $\nabla \cdot \boldsymbol{\xi} \approx 0$, and that to the next order, a top-hat displacement achieves the minimisation of δW . In fact, it can be shown that the dominant contribution for a perturbation with poloidal and toroidal m/n mode structure is of the form

$$\delta W_2 \sim \int \left[\left(r \frac{d\xi_r}{dr} \right)^2 + (m^2 - 1)(\xi_r)^2 \right] \left(\frac{n}{m} - \frac{1}{q(r)} \right)^2 r dr,$$

which physically corresponds to the energy required for the mode to bend the field lines. Clearly, this field-line bending is a stabilizing ($\delta W_2 \geq 0$) effect. For a perturbation to be unstable, it is therefore necessary that it minimize such field-line bending, resulting for $m = n = 1$ in a perturbation with $d\xi_r/dr = 0$, except at any radial position r where $q(r) = m/n = 1$; at those

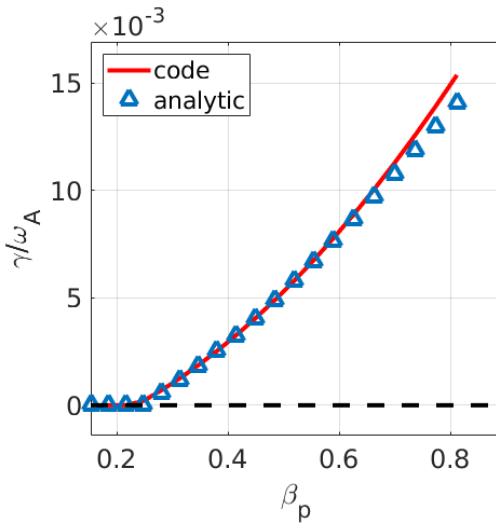


Figure 1.6 – Internal kink: Comparison of growth rates vs. poloidal β , obtained from VENUS-MHD vs. analytic form due to Bussac, for equilibrium with profiles shown in Figure 1.3 and scan over different scaled pressure profiles. The internal kink mode structure of Figure 1.5 corresponds to the first unstable value of β_p .

radial positions, $d\xi_r/dr$ may be non-zero since it is cancelled by the factor $(n/m - 1/q)^2$. With the appropriate boundary conditions, this implies that the radial component ξ^r have a top-hat shape, as can be inferred from Figure 1.5, where $\rho\xi^{\rho} \approx r\xi_r$ is plotted for a numerically obtained 1/1 internal kink. Since such a top-hat displacement is marginally stable to leading order in the inverse aspect ratio ϵ_a , it is necessary to go to higher order in ϵ_a to determine the stability of the 1/1 internal kink. After the initial work by Shafranov [Sha70], a treatment up to fourth order in ϵ_a was later achieved by Rosenbluth et al. [RDR73], who again considered a straight cylindrical plasma column, thus neglecting toroidal effects. A perhaps surprising result of the analysis of Rosenbluth et al. is the realization that by a careful treatment of the “layer region” (radial region close to $q = 1$, where field line bending stabilization is weak), the growth rate $\gamma = i\omega$ turns out to be linearly related to the minimised potential energy δW_e in the “external region”, i.e. $\gamma \sim \delta W_e$. In contrast, a naive consideration of the ideal MHD variational principle would instead suggest a quadratic relationship $\gamma^2 \sim \delta W$.

The results of Rosenbluth et al. [RDR73] were later revisited by Bussac et al. [BPES75], who included corrections due to the toroidal bending of the plasma into the shape of a torus. Hinting at the delicate nature of the 1/1 internal kink instability, Bussac et al. showed that the cylindrical corrections to δW derived by Rosenbluth et al. are exactly cancelled by certain toroidal corrections, so that the remaining terms arise purely due to toroidal effects. Employing the singular layer analysis of [RDR73], Bussac et al. show that for a parabolic q -profile, the internal kink mode is unstable provided that $\beta_p \gtrsim 0.2$ to 0.3, where β_p is the poloidal beta, a measure for the ratio of the plasma internal energy (pressure) vs. the poloidal magnetic energy, inside the $q = 1$ surface, with an approximately quadratic growth as a function of β_p (cp. Figure 1.6, where a scan in β_p is shown computed by the stability code VENUS-MHD; for

a description of this code, see chapter 5).

As already mentioned in the previous section, the plasma compressibility (parallel dynamics) leads to inertial enhancement and a corresponding adjustment of the growth rate by a factor $\sqrt{3}$ [GGJ75]. Ara et al. [ABC⁺78] have later taken into account the diamagnetic motion of ions and electrons in the layer: Introducing the ion diamagnetic frequency $\omega_{i,p} = -p'_i/(Zen_i B_0 r)$, the most important modification due to such diamagnetic effects is the replacement of the growth rate $\gamma = i\omega \rightarrow i\sqrt{\omega(\omega - \omega_{i,p})}$, resulting in a dispersion relation $i\sqrt{\omega(\omega - \omega_{i,p})} \sim \delta W_e$, where δW_e is the minimized potential energy of Bussac et al. [BPES75]. Note that the introduction of diamagnetic effects breaks the symmetry present for ideal MHD modes (which lie symmetric about the origin). Further references and a more extended review can be found in [Gra99].

1.4.1 Centrifugal effects

The stability of the 1/1 internal kink in tokamaks and in the absence of toroidal rotation, has received considerable attention since the result due to Bussac et al. [BPES75] and is reasonably well understood. The effect of strong toroidal rotation on this instability has first been considered by Walbroeck [Wae96] only two decades after the work by Bussac et al. [BPES75]. In [Wae96] it was shown that, in addition the trivial doppler shift of the mode frequency in the lab frame, the centrifugal forces cause a “gyroscopic” stabilization mechanism even in rigidly rotating plasmas, i.e. in the absence of any shear in the rotation frequency $\Omega(\psi) = \text{const.}$. This gyroscopic mechanism is especially important for tokamaks with a large fraction of neutral beam injection (NBI) heating, where it can lead to significant stabilising effects. The rotation velocities can in fact exceed the sound velocity in low aspect-ratio tokamaks such as MAST [LAA⁺07] and NSTX [MBF⁺05]. Therefore centrifugal effects can not be neglected in these devices.

The observations of a stabilizing mechanism by Waelbroeck [Wae96] has later been confirmed in work by Wahlberg and Bondeson [WB97, WB00] (see also [WB01] for similar effects on interchange modes in an axisymmetric torus). The latter work was assisted by a computer-algebra program to carry out the inverse aspect ratio expansion of the Frieman-Rotenberg equations [FR60] to the required order in inverse aspect ratio $\epsilon_a \sim a/R_0$. Wahlberg and Bondeson relate the stabilising effects due to toroidal rotation to the created nonuniform plasma density and pressure on the $q = 1$ surface due to the centrifugal force. The stabilisation in the work [Wae96, WB00] is thus related a finite continuum frequency in the $q \approx 1$ layer, the so-called “Brunt-Väisälä” (BV) frequency,

$$\omega_{\text{BV}}^2 \equiv \frac{\mathcal{M}^2 \Omega^2}{3} \left(1 - \frac{1}{\Gamma}\right),$$

where Γ is the adiabatic index, \mathcal{M} is the sonic Mach number, given by $\mathcal{M} = \Omega/\omega_s = \sqrt{\rho\Omega^2 R_0^2/2p}$, and p the plasma pressure. This continuum frequency is partly responsible for the stabilization

of the internal kink mode in the analysis of [Wae96]. To leading order, it is found in [WB00] that the stability criterion can be written as $\omega_{BV} > \gamma_B$, where γ_B is the growth rate computed in the absence of any flow. As already noted in [WB00], the explicit dependency of this stability criterion on the value of the adiabatic index Γ shows that this result depends crucially on the fluid closure in the ideal MHD model and thus the parallel dynamics. This is consistent with our discussion of the coupling of perpendicular and parallel dynamics due to centrifugal effects in section 1.2.4. Since ideal MHD is a bad description of the parallel dynamics in realistic high-temperature tokamak plasmas, it is expected that the *kinetic behavior* along the field lines will influence this stabilization mechanism considerably.

Since the stabilizing mechanism discovered in the analysis of [WB00] crucially depends on the non-uniform distribution of pressure and density in the equilibrium on the $q = 1$ flux surface due to centrifugal forces, the work [WB00] also suggests that it is crucial to consistently include flow effects not only in the stability model, but also in the equilibrium. The importance of centrifugal effects for internal kink mode stability for numerical stability analyses in toroidally rotating plasmas with realistic tokamak geometry has been further investigated by Wahlberg, Chapman and Graves in [WCG09]. In this work, analytical results of [WB00] were compared to results obtained from two numerical codes: CASTOR-FLOW) [SGM⁺05], with a self-consistent treatment of centrifugal effects in both the equilibrium as well as the linearized equations of motion, and another code, MISHKA-F [CSHM06], which includes flow effects correctly in the linearized stability equations, but neglects centrifugal effects on the equilibrium. Centrifugal effects have been found to be important for the stability of the internal kink mode, even at modest flow rates. In particular, guided by the analytical results, the authors of [WCG09] find that while the assessment of stability from the two codes [SGM⁺05, CSHM06] can be very similar for certain choices of plasma profiles, for other choices, they make completely opposite predictions; In the most extreme case reported in [WCG09], the consistent model predicts complete stabilization of the internal kink mode, while the inconsistent model predicts a destabilization due to rotation. The marked differences between a self-consistent and non-self-consistent treatment of centrifugal effects on mode stability are traced back to the stabilizing influence of a rotation-induced geodesic acoustic mode (GAM), which exists in the ideal MHD model only when centrifugal effects on rotating equilibria are taken into account. Centrifugal effects lead to pressure corrections, enhancing the Shafranov shift of flux surfaces, and induce a nonuniform plasma density on flux surfaces. A recent review on GAMs in rotating plasmas, and their correction to instabilities, can be seen in [GW17].

1.4.2 Kinetic effects

Going beyond fluid theory, it has been found that additional important effects on the internal kink mode stability can come from interactions of this macroscopic mode with both an external species of (suprathermal) kinetic ions, as well as additional kinetic effects due to a kinetic description the bulk plasma. As already indicated in section 1.3, a suitable kinetic-MHD model to describe such kinetic effects consists of two ingredients: A fluid model to determine

the evolution of the electro-magnetic fields, as well as macroscopic fluid moments, such as the bulk velocity and current density on the one hand, and on the other hand a kinetic model which accounts for kinetic effects such as particle-wave interactions and Landau damping, and which provides a more accurate moment closure (pressure) for the moment equations of the fluid model.

At this level, we can distinguish between self-consistent and perturbative approaches. Whereas in a self-consistent approach, the kinetic-MHD model that is used to assess the stability includes the full interaction between the kinetic and the fluid components by solving for the evolution of the full coupled system of fluid and kinetic equations, in a perturbative approach, one first solves the fluid equations in the absence of kinetic effects, and then adds kinetic corrections in a second step. Typically in such a perturbative approach, the mode structure is fixed from ideal MHD, and only corrections to the mode frequency ω are solved for based on a modified dispersion relation. A suitable dispersion relation is obtained from the momentum equation (1.25), by multiplication with ξ^* , and integration against $d\mathbf{x}$. This general procedure results in an equation of the form

$$i\omega \sim \delta W_{\text{fluid}} + \delta W_{\text{kin}}(\omega),$$

where, as a rule of thumb, ω is replaced by $\sqrt{\omega(\omega - \omega_{i,p})}$, when diamagnetic effects due to ions are included. As indicated by this equation, two contributions to δW are now taken into account; the first one corresponds to the fluid potential energy contributions δW_{fluid} , while the second term corresponds to

$$\delta W_{\text{kin}}(\omega) \sim \int \xi^* \cdot (\nabla \cdot \delta \mathbf{P}(\xi; \omega)) d\mathbf{x},$$

and is typically a complicated function of ω . The additional term δW_{kin} accounts for kinetic effects such as particle-wave interactions. Different versions of such a dispersion relation have been derived under various approximations. Notably, this procedure has been employed by Chen et al. [CWR84] to investigate the excitation of internal kink modes by trapped energetic beam ions. The model of Chen et al. is based on an ideal MHD description of the bulk plasma species (adiabatic closure), and includes an additional kinetic species of suprathermal particles which contributes a hot particle pressure tensor of CGL form $\delta P_h = p_{||,h} \mathbf{b}\mathbf{b} + p_{\perp,h} (\mathbf{I} - \mathbf{b}\mathbf{b})$ with parallel and perpendicular pressure components $p_{||,h}$, $p_{\perp,h}$, respectively. A bounce-averaged form of the kinetic equation is solved, taking into account the toroidal drift frequency, while neglecting finite orbit-width effects (drifts off the flux surface). The effects of a parallel electric field are neglected (i.e. the model assumes $\delta E_{||} = 0$). Based on an expansion in inverse aspect ratio, the authors find that the inclusion of trapped supra-thermal particles induces an additional trapped-particle branch in the dispersion relation, with mode frequency $\omega \sim \omega_{dh}$, where ω_{dh} is the drift frequency of the suprathermal trapped particles. It is proposed that this trapped-particle induced branch is related to the fishbone instability.

In addition to such a minority species of supra-thermal ions, Kruskal and Oberman [KO58]

have also shown that thermal ions can stabilize MHD instabilities such as the internal kink [HH88]. Kruskal Oberman theory predicts that thermal trapped ions contribute to δW via an additional term

$$\delta W_{\text{KO}} \sim \int \boldsymbol{\xi}_\perp^* \cdot (\nabla \cdot \delta p_{k,i}) d\mathbf{x},$$

where the perturbed pressure tensor is given by $\delta p_{k,i} = -\boldsymbol{\xi}_\perp \cdot \nabla p_{i0}$. This stabilizing term is analogous to the adiabatic compressibility term in ideal MHD, and arises due to the fact that thermal trapped particles resist compression. A generalisation of the inertial enhancement first derived by Glasser et al. [GGJ75] in a kinetic treatment of the thermal ions, has been derived by Graves et al. in [GHH00]. The derivation of [GHH00], which takes into account radial drift motions across the rational surface, shows that the inertial enhancement factor $\sqrt{1 + \Delta}$, where $\Delta_{\text{MHD}} = 1 + 2q^2$ in ideal MHD, as derived by Glasser et al., is replaced by

$$\Delta \approx \left[\frac{1.6}{\sqrt{\epsilon_{a,1}}} + 0.5 \right] q^2,$$

when kinetic effects in the singular layer (where $q \approx 1$) due to thermal ions are taken into account. Here $\epsilon_{a,1} = r_1/R_0$ is the inverse aspect ratio at the $q = 1$ surface. Thus, kinetic effects due to thermal ions are expected to significantly affect the growth rate due to corrections to the inertia. It is noted that analogous inertial enhancement factors have been evaluated for higher frequency modes [ZC14]. Besides such inertial corrections, the work by Graves et al. [GHH00] investigates the effects of sheared toroidal plasma rotation on the internal kink mode. To this end, a kinetic-MHD model is derived which includes toroidal rotation in the toroidal drift motion of the kinetic ions. While this model is suitable to investigate the effects of velocity shear, centrifugal effects are not taken into account. Therefore, to investigate the combined centrifugal and kinetic effects in a rotating plasma, the model of Graves et al. [GHH00] should be extended to allow for centrifugal effects on the thermal ions, as well as centrifugal and coriolis forces in the equation for the plasma displacement. Also in [GHH00], the parallel electric field was assumed to vanish ($\delta E_{||} = 0$).

The inclusion of an electrostatic contribution (parallel electric field) in the computation of the perturbed potential energy δW has been discussed by Antonsen and Lee [Ant82], who employ Lagrangian coordinates to solve the equation of the kinetic species, and include a quasi-neutrality relation. Their model is bounce-averaged and takes into account the toroidal drift motion of trapped particles, but neglects drifts off the flux surface. Though the work of Antonsen and Lee does not focus on the internal kink, their model has later been employed by Antonsen and Bondeson [AB93] who numerically investigate the effects of thermal trapped particles on the $m = n = 1$ internal kink, with a particular emphasis on the effect of unequal electron- and ion-temperatures, and the coupling of this mode to electrostatic trapped particle modes.

An extension of the model of Antonsen and Lee [Ant82] to include full orbit width effects has been proposed by Porcelli et al. [PSBZ92, PSK94], though their model was applied to

cases where $\delta E_{||}$ was also set to zero (no attempt was made to eliminate $\delta E_{||}$ in favor of ξ). In [PSK94], starting from the guiding-centre equation derived by Littlejohn [Lit83], the authors obtain a suitable linearisation of the kinetic equation for δf by direct perturbation of the guiding-centre equations of motion. Their method based on the Lagrangian formulation of guiding-centre theory allows for the full finite-orbit width effects to be retained. Although a perturbed electrostatic potential $\delta\phi$ is included in the derivation of the kinetic equation as mentioned above, the corresponding quasi-neutrality relation that would be needed to self-consistently calculate $\delta\phi$ is not discussed in the work [PSK94]. It is shown that results from models with kinetic closure in the zero-orbit width limit overestimate the fast particle stabilization of the internal kink in the potato orbit limit, where the bounce-frequencies become of the same of magnitude as the drift frequency, and the orbit width is of the order of the local minor radius. The formulation by Porcelli is in fact implemented in the numerical code CASTOR-K [BK99, NBC⁺15], to compute the influence of a population of suprathermal particles on a fluid bulk plasma mode. A model similar to that of Porcelli et al. [PSK94] has also been obtained by Helander et al. [HGHM97], who present a different derivation of the form of the perturbed distribution function δf , based on physical intuition. The authors of [HGHM97] apply their kinetic-MHD model to “double-kink” modes, which are similar to the conventional internal kink modes, except that the q -profile is no longer assumed to be monotonic, allowing it to have multiple $q = 1$ surfaces inside the plasma. A similar role of the fast ions on the stability is observed also for such double-kink modes.

We emphasize that none of the kinetic-MHD models above include centrifugal effects.

1.5 Summary and contributions of this thesis

Basic elements of the ideal MHD model have been reviewed in the context of tokamak equilibria and as a model for macroscopic plasma instabilities. A particular emphasis has been given on the inadequacies of the simple adiabatic pressure closure of this model for plasmas of fusion relevance, at low collisionality. It has been pointed out that for a variety of instabilities, in particular in the absence of toroidal rotation, the stability analysis based on the ideal MHD model depends mostly on the perpendicular dynamics, and hence the incorrect treatment of the parallel dynamics in this model does not have any dramatic consequences. In contrast, for pressure-driven instabilities such as the $m = n = 1$ internal kink, which are marginally stable to high order in the inverse aspect ratio expansion, corrections to the adiabatic pressure closure assumption can have a significant influence on the stability analysis. As illustrated by our short survey, the internal kink instability depends on many effects, such as toroidal rotation, kinetic effects due to a suprathermal species (generated e.g. by external heating mechanisms such as ICRH, NBI), kinetic effects due to the thermal bulk species and finite Larmor-radius, or diamagnetic effects. A considerable research effort has illuminated many aspects of the stability of instabilities such as the internal kink, and several models have been proposed to understand various aspects of the underlying plasma dynamics. However, no model is available at present, which incorporates important centrifugal as well as kinetic effects. As

explained in this introduction, centrifugal effects cause a stronger coupling of the perpendicular and parallel dynamics, even within the ideal MHD model, and therefore a kinetic treatment should be expected to be of particular importance for plasmas with strong toroidal rotation. The extension of existing kinetic-MHD models to consistently include centrifugal as well as kinetic effects is the main goal of this thesis. This requires a better understanding of an appropriate simplified kinetic description (a guiding-centre description) of the dynamics of charged particles in a tokamak plasma, and the relation of this microscopic guiding-centre theory to the macroscopic fluid description which describes the gross motion of the plasma in response to the kinetic effects.

The remainder of this thesis is organised as follows. In chapter 2, we will review guiding-centre theory as developed originally by Littlejohn [Lit83]. A full derivation of the higher-order guiding-centre Lagrangian allowing for flow velocities on the order of the sound velocity will be presented. A complete and detailed derivation is given in this work. It is hoped that the level of detail may serve as a reference for anyone wishing to learn about the derivation of these equations, which are not only implemented in most modern particle pushing codes, but also form the basis of modern gyrokinetic theory [BH07]. It will be shown how the kinetic-MHD model of Porcelli et al. [PSK94] can be derived from gyrokinetics. The non-adiabatic and adiabatic contributions to δf identified in [PSK94] will be shown to correspond to the perturbed distribution function in gyrocentre coordinates and the contribution arising due to the pull-back to guiding-centre coordinates, respectively. Moreover, our derivation allows for the inclusion of centrifugal effects as well as a parallel component of the perturbed electric field. Diamagnetic effects will be discussed at length in chapter 4. From the point of view of guiding-centre theory, it will be shown that diamagnetic effects arise due to higher-order corrections in the Larmor motion of particles about their guiding centre. These higher-order corrections stem from gradients in the electric field, due to which the particles describe an elliptic rather than a circular motion perpendicular to the field lines. In particular, the deviation from a circular Larmor motion is shown to cause off-diagonal components of the pressure tensor, resulting in a so-called “gyroviscous” contribution to the pressure tensor. Based on well-known results from fluid theory together with this novel derivation of the gyroviscous pressure tensor from guiding-centre theory, the pressure closure approach is then followed to derive a consistent kinetic-MHD model that is suitable for a numerical implementation and can be used to describe the stability of plasmas to internal kink modes, including the effects of strong toroidal flows, kinetic effects due to the resonant interaction with thermal and supra-thermal (virtually) collisionless ions and the effects of a parallel electric field in a single, unified model based on guiding-centre theory. It will be argued in chapter 4 that under suitable approximations the kinetic-MHD formulation with pressure-closure achieves a more efficient description of the kinetic effects for global macroscopic instabilities, compared to a fully kinetic formulation based on a current-closure approach. Indeed, a fully kinetic description appears to require higher-order $O(\epsilon_B^2)$ corrections (in Larmor radius) to be retained to achieve a suitable description of macroscopic instabilities. In contrast, within the proposed kinetic-MHD approach, $O(\epsilon_B)$ corrections appear to be sufficient. Thus, we argue that the kinetic-

Chapter 1. Ideal MHD and beyond

MHD approach relying on pressure-closure may achieve a more efficient description of the plasma dynamics, as compared to an alternative approach based on current-closure.

2 Guiding-centre theory

2.1 Introduction

As has been pointed out in the previous chapter, it is commonly agreed that, in principle, the dynamics of plasmas of interest to fusion research are well-described by the Vlasov-Maxwell system of equations. These equations are able to describe most aspects of the plasma dynamics; ranging from macroscopic (fluid) instabilities to the enhanced transport due to small amplitude microscopic instabilities associated to turbulence, the interaction of electromagnetic waves with the gyromotion of particles (used for plasma heating) and even describing the propagation of high-frequency waves through the plasma. With this power of describing a wide range of effects comes an inherent complexity, which makes these equations not only difficult to understand and analyse analytically, but also extremely difficult to solve numerically in all but the simplest of settings. It is therefore not astonishing that the greatest progress in plasma physics has been achieved by the development of simplified models, which for a given problem, retain the essential physical effects required to understand the plasma dynamics, while at the same time discarding superfluous information, such as effects taking place at length or time scales which are not relevant to the physics processes of interest. If a suitable model can be found, this allows both analytical as well as numerical progress to be made, resulting in a deeper understanding of the physical processes taking place.

One important example of such a reduction in complexity occurs when passing from kinetic to fluid theory. Fluid theory, and in particular ideal MHD has led to a better understanding of macroscopic instabilities, as has already been discussed in the previous chapter. Another example of such a reduced model will be discussed in the present chapter. Its goal is to give a description of the motion of individual charged particles in strong background magnetic field, whose exact description under most conditions of interest is assumed to be given by the Lorentz equations:

$$m\ddot{\mathbf{x}} = q[\mathbf{E}(\mathbf{x}) + \mathbf{v} \times \mathbf{B}(\mathbf{x})]. \quad (2.1)$$

Here, \mathbf{x} denotes the particle position, $\mathbf{v} = \dot{\mathbf{x}}$ the particle velocity, m is the particle mass, q the

Chapter 2. Guiding-centre theory

particle charge, and \mathbf{E} , \mathbf{B} denote the background electric and magnetic fields, respectively.

As is well known, such particles carry out a helical motion around field lines. The frequency of the gyration increasing with higher magnetic field strength, and when lowering the particle mass. Assuming a homogeneous background B-field $\mathbf{B}(\mathbf{x}) = \mathbf{B}_0$ and $\mathbf{E}(\mathbf{x}) = \mathbf{0}$, and choosing a cartesian frame $\mathbf{e}_1, \mathbf{e}_2, \mathbf{b}$, such that $\mathbf{B}_0 = B_0 \mathbf{b}$, the solution of the Lorentz equations, can be written in the form

$$\mathbf{x}(t) = \mathbf{X}_0 + \mathbf{v}_{0,\parallel} t + \boldsymbol{\rho}(\zeta(t)), \quad \boldsymbol{\rho}(\zeta) = \rho_0 \cos(\zeta) \mathbf{e}_1 - \rho_0 \sin(\zeta) \mathbf{e}_2, \quad \zeta(t) = \zeta_0 + \Omega_0 t. \quad (2.2)$$

The six constants \mathbf{X}_0 , ρ_0 , $\mathbf{v}_{0,\parallel} = v_{0,\parallel} \mathbf{b}$, ζ_0 are to be determined from the initial conditions, i.e. form the specification of the six (3+3) components of the initial particle position and velocity (\mathbf{x}_0 , \mathbf{v}_0). The constant Ω_0 depends on the particle mass m , the charge q and the background field strength B_0 : $\Omega_0 = qB_0/m$. According to (2.2), the particle describes a helical motion around its “guiding-centre” \mathbf{X}_0 , with a constant parallel velocity $v_{0,\parallel}$ along the field lines, and with gyration frequency, or *gyrofrequency*, given by Ω_0 . Note in particular that Ω_0 depends only on the particle species (characterised by charge and mass), as well as the field strength of the background field, but it is independent of e.g. the particle energy. It turns out that for most typical macroscopic instabilities in magnetically confined plasmas, the characteristic time scale of interest is much longer than the gyrofrequency of the particle species. This means that even during a potentially disruptive plasma instability, all particles involved will typically undergo a large number of gyrations before the fields have changed appreciably. This suggests that to describe the interaction of the particles with such plasma instabilities, only the “gyro-averaged” particle motion (average over one gyroperiod), is usually of relevance.

In addition to this difference of time scales between the particles’ fast gyromotion and typical plasma instabilities, the magnetic confinement in tokamaks is, *by design*, such that the applied magnetic field binds the particles as tightly as technically feasible. This makes the typical particle gyroradius $\rho_0 = v_\perp/\Omega_0$ of most particles ($v_\perp \sim v_{\text{th}}$) much smaller than the typical length scales $L_B \sim [|\nabla B|/B]^{-1}$ of the plasma. Macroscopic instabilities, which are the most dangerous to the confinement of the plasma, have typical scale lengths on the order of L_B . To obtain a reduced description of the particle dynamics, it is therefore reasonable to assume $\rho_0 \ll L_B$, and $\Omega_0 \gg \omega$, where ω is the typical frequency of the plasma instability (inversely proportional to the time-scale on which the instability changes appreciably).

The reduced description of particle motion under these conditions on the time- and length-scales of interest is usually referred to as *guiding-centre theory*. The central aim of guiding-centre theory is to find a reduced description of the average particle motion, or more precisely, to describe the motion of the particle’s guiding-centre. To derive reduced equations of motion for the particles, one introduces the formal guiding-centre ordering parameter

$$\epsilon_B \sim \frac{\rho_0}{L_B} \sim \frac{\omega}{\Omega} \ll 1.$$

Historically, two approaches to deriving suitable guiding-centre equations of motion have

been followed. The first approach, pioneered in the 1950's and 60's by Alfvén, Kruskal, Northrop, led to the first insights into the properties of guiding-centre motion. This approach is conceptually straight-forward, but becomes somewhat computationally involved, in particular when attempting to derive higher than first-order corrections to the guiding-centre motion in ϵ_B . Using this approach, one starts from the ordered (and normalised) Lorentz equations

$$\dot{\mathbf{v}} = \epsilon_B^{-1} [\mathbf{E}(\mathbf{x}) + \mathbf{v} \times \mathbf{B}(\mathbf{x})], \quad \dot{\mathbf{x}} = \mathbf{v},$$

and proceeds to solve these equations for the average particle motion, order by order in ϵ_B . Using this approach, it has been observed that the magnetic moment $\mu = \frac{1}{2}mv_\perp^2/B$ is a conserved quantity to order ϵ_B . The existence of such an (approximate) conserved quantity is a fundamental insight that underlies much of guiding-centre theory. It also provides a first clue that a different description and derivation of the guiding-centre equations of motion should be possible.

As is well-known, it is a mantra of classical mechanics that classical systems should be described by a Lagrangian, and the equations of motion obtained from variational principles. In particular, there should be a Lagrangian description of guiding-centre motion. The second approach to deriving guiding-centre motion, comes from an attempt to find this guiding-centre Lagrangian. In accordance with Noether's theorem, the conserved quantities of the particle motion correspond to the symmetries of the Lagrangian. Therefore, if μ is a (approximate) constant of guiding-centre motion, then there should correspond to it a (approximate) symmetry, and as the experience from Hamiltonian mechanics indicates, the variable conjugate to μ should be a periodic (or approximately periodic) variable. One obvious candidate for such a variable is the gyroangle ζ , that we have already encountered in equation (2.2). It turns out that the derivation of such a guiding-centre Lagrangian requires mathematical techniques beyond those that one can typically find in textbooks on Lagrangian or Hamiltonian mechanics, which are mostly formulated with respect to canonical variables. After several attempts, an elegant derivation of the guiding-centre Lagrangian in a natural (non-canonical) set of coordinates was finally achieved in the celebrated work of Littlejohn [Lit83], whose methods underpin most of "modern" guiding-centre theory (as presented e.g. in [CB09]) and natural extensions such as gyrokinetics, not only in theory but also in their numerical implementation.

One crucial advantage of the Lagrangian formulation of guiding-centre theory over the equations obtained by direct expansion of the Lorentzian equations of motion, is that the Lagrangian guiding-centre equations automatically ensure that additional conservation laws are satisfied precisely; in a static background the guiding-centre energy is exactly conserved, and in an axisymmetric background field, the canonical toroidal momentum P_ϕ is a constant of guiding-centre motion. Since the conservation of toroidal momentum is intimately related to the confinement of particles in tokamaks, physically realistic simulations are usually based on the Lagrangian guiding-centre equations, as only they ensure that particles will not experience an unphysical drifting motion which might cause them to escape the plasma even in a perfectly axisymmetric configuration.

The general mathematical technique that has been developed and introduced to guiding-centre theory by Littlejohn [Lit83], will be used in section 2.4 to provide a full and detailed derivation of the guiding-centre equations in strongly flowing plasmas. While this derivation is not entirely new, it is here given in unprecedented detail, including computations that are usually omitted in the published literature (probably for the sake of brevity). The intention of including these calculations here is two-fold: Firstly, our results disagree in some details with expressions found in the literature. These differences turn out to be important for our subsequent discussion of finite Larmor-radius effects in kinetic MHD. It is hoped that including the detailed derivation of the equations here will make it possible for anyone interested to check the correctness of the results. Secondly, a major hurdle to entering the field of guiding-centre theory (and gyrokinetics) is the understanding of fundamental concepts such as the push-forward and pull-back relations. A proper understanding of these concepts, in particular to higher than leading-order, turns out to be rather subtle and requires a closer look at the meaning of guiding-centre variables, and their relation to the particle position and velocity. As pointed out by Brizard, modern guiding-centre theory is not fundamentally a description of the “average” particle motion, but rather a description of particle motion in the most convenient coordinates, making the conserved magnetic moment μ and its conjugate ignorable variable ζ , appear explicitly. In this formulation, the reduction is not achieved by any averaging, as is often erroneously claimed, but by uncovering a hidden symmetry of particle motion, the so-called *gyro-symmetry* [Lit84]. It is the author’s hope that the detailed derivation given below will allow more researchers to appreciate not only the usefulness, but also the beauty of Littlejohn’s Lagrangian formalism of guiding-centre motion, without the pain of having to derive most details on their own.

2.2 Lagrangian mechanics and Lie perturbation theory

The differential-geometric point of view for deriving reduced equations describing the motion of charged particles in strongly magnetized background has first been introduced by Littlejohn [Lit82, Lit83, Lit84], who bases his analysis on earlier methods introduced in Cary and Littlejohn [CL83]. An overview of previous perturbation methods at that time is given by Cary in [Car81]. In the Lie perturbation approach, the equations for the motion of the particle’s guiding-centre [Lit83] are derived by a perturbative method based on the geometric theory of classical mechanics, as e.g. developed in the classic textbooks [AM87, Arn78].

In the following sections, we will first remind the reader of some basic notions of classical mechanics, giving mostly references to existing excellent accounts of the background theory in the literature, and then derive the starting equations of Lie perturbation theory. One of the intentions of this section is to collect material which had to be assembled from various sources in the literature. This section is therefore intended to serve as a convenient reference for the reader, as well as building the basis for later sections. These further developments will require a detailed understanding of not only the guiding-centre motion itself, but also the precise meaning of the guiding-centre coordinates. By providing a more detailed discussion of the

mathematical basis of guiding-centre theory, we hope to clarify the sometimes subtle issues surrounding the guiding-centre transformation and to emphasize the fact that the guiding-centre equations of motion are meaningless without knowledge of the chosen guiding-centre coordinates, and how these coordinates relate to the phase-space coordinates of the physical particle which they describe.

2.2.1 Variational principles and Lagrangian one-form

Informative discussions of the Lagrangian formulation of guiding-centre motion, including the details of the variational principle can be found by some of the founders of this field, in the reviews by Cary, Brizard [CB09, section II] and by Brizard, Hahm [BH07, section IV]. We here discuss a few key elements which we shall make use of in the following.

It is well-known that given a physical system with N degrees of freedom, and set of coordinates $\mathbf{q} = (q^1, q^2, \dots, q^N)$, the dynamics of this physical system can often be described by a Lagrangian $L(\mathbf{q}, \dot{\mathbf{q}}, t)$. The Lagrangian is a function of the coordinates and their time derivatives. The variational principle states that the equations of motion can be obtained from the Lagrangian, by finding the curve $t \mapsto \mathbf{q}(t)$ which minimizes the action

$$A[\mathbf{q}(t)] = \int L(\mathbf{q}, \dot{\mathbf{q}}, t) dt.$$

An equivalent condition for $t \mapsto \mathbf{q}(t)$ to minimize the action A is that \mathbf{q} solves the system of Euler-Lagrange equations

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\mathbf{q}}} \right) - \frac{\partial L}{\partial \mathbf{q}} = 0.$$

For the motion of a charged particle of mass m and charge q in a background electromagnetic field, the Lagrangian is given by

$$L(\mathbf{x}, \dot{\mathbf{x}}, t) = \frac{m}{2} \dot{\mathbf{x}}^2 + q \mathbf{A}(\mathbf{x}, t) \cdot \dot{\mathbf{x}} - q \Phi(\mathbf{x}, t). \quad (2.3)$$

The Lagrangian is expressed in terms of the time-dependent vector potential \mathbf{A} , and the electrostatic potential Φ , which are related to the electric and magnetic fields \mathbf{E} , \mathbf{B} as usual via $\mathbf{E} = -\nabla\Phi - \partial_t \mathbf{A}$, $\mathbf{B} = \nabla \times \mathbf{A}$. A short calculation confirms that the Euler-Lagrange equations for (2.3) lead to the well-known Lorentz equations (2.1).

Unfortunately, this conventional Lagrangian formulation of the particle motion is not immediately suitable for a perturbation analysis. The goal of perturbation theory is to find a transformation to new coordinates to a given order in an expansion parameter, in which the Lagrangian takes a simpler form, e.g. because the transformed variables make approximate symmetries of the original Lagrangian manifestly apparent. The possible transformations for a Lagrangian of the form (2.3) are transformations of the particle position coordinates $\mathbf{x} = (x^1, x^2, x^3)$, only; whereas our experience with the particle motion in homogeneous fields

(cp. equations (2.1), (2.2)) clearly indicates that suitable guiding-centre coordinates should depend on the particle velocity in addition to the particle position. This motivates an approach, where not only the particle position \mathbf{x} , but also the particle velocity \mathbf{v} are treated as independent coordinates.

Early attempts at deriving guiding-centre equations have therefore focused on classical Hamiltonian perturbation techniques, where one first introduces the canonical momentum $\mathbf{p} = \partial L / \partial \dot{\mathbf{q}}(\mathbf{q}, \dot{\mathbf{q}}, t)$, and uses this to interpret the velocity $\dot{\mathbf{q}} = \dot{\mathbf{q}}(\mathbf{q}, \mathbf{p}, t)$ as a (time-dependent) function of the phase-space variables (\mathbf{p}, \mathbf{q}) . In canonical coordinates (\mathbf{p}, \mathbf{q}) , the dynamics can then be expressed in terms of a scalar function, the Hamiltonian $H(\mathbf{p}, \mathbf{q}, t)$, by Hamilton's equations $\dot{\mathbf{q}} = \partial H / \partial \mathbf{p}$ and $\dot{\mathbf{p}} = -\partial H / \partial \mathbf{q}$. With this canonical formulation in place, the perturbation analysis can then proceed via canonical transformations.

That such canonical methods are not favoured in the context of guiding-centre theory can be attributed in large part to the fact that the canonical momentum of a charged particle is in this case given by $\mathbf{p} = \partial L / \partial \dot{\mathbf{x}} = q\mathbf{A}(\mathbf{x}, t) + m\mathbf{v}$, and is therefore not a simple function of the velocity \mathbf{v} , as would be the case in other systems. Furthermore, since $q\mathbf{A}$ is formally ϵ_B^{-1} larger than $m\mathbf{v}$ in the perturbative ordering, the transformation to canonical variables mixes terms of different orders. This makes both the perturbative analysis, as well as the physical interpretation of the results more difficult, compared to the viewpoint introduced by Littlejohn, which we will describe next.

Instead of working in canonical coordinates, Littlejohn's approach uses arbitrary phase-space coordinates, which one is therefore free to choose. The two main benefits of this approach are that it is possible to choose suitable coordinates with a simple physical interpretation, and that there is more freedom in the choice of transformations in the perturbation analysis. This comes at the expense of having to transform not only the scalar Hamiltonian, but instead all components of the so-called phase-space Lagrangian.

Since general phase-space coordinates are considered, it is most natural (or maybe inevitable) to formulate this approach in the language of differential geometry. In such a geometric formulation of Hamiltonian mechanics, the phase-space Lagrangian is now interpreted as a one-form Γ on phase-space. We will assume the reader to be familiar with this formulation of mechanics, for reviews of the mathematical elements we refer to [CB09, BH07, BSQ13]. In canonical coordinates, this one-form can be expressed simply as

$$\Gamma = \mathbf{p} \cdot d\mathbf{q} - H(\mathbf{p}, \mathbf{q}, t) dt \equiv \left(\sum_{i=1}^N p_i dq^i \right) - H(\mathbf{p}, \mathbf{q}, t) dt.$$

To avoid writing out sums as above, we will make use of Einstein's summation rule whereby repeated indices are summed over, in the following. One may think of Γ as an infinitesimal action, since the action along a path $t \mapsto \alpha(t) := (\mathbf{p}(t), \mathbf{q}(t))$ (say, defined for $t \in [t_0, t_1]$) in

phase-space can be simply obtained by integration of Γ along the path α :

$$\mathcal{A}(\alpha(t)) = \int_{\alpha} \Gamma = \int_{t_0}^{t_1} \dot{\alpha}(t) \cdot \Gamma|_{\alpha(t)} = \int_{t_0}^{t_1} [\mathbf{p}(t) \cdot \dot{\mathbf{q}}(t) - H(\mathbf{p}(t), \mathbf{q}(t), t)] dt.$$

Minimizing the action \mathcal{A} is equivalent to Hamilton's equations. In contrast to the Hamiltonian formulation, the equivalent formulation in terms of the one-form Γ also indicates the correct phase-space Lagrangian formulation with respect to arbitrary phase-space coordinates: If $\mathcal{Z} = (\mathcal{Z}^1, \dots, \mathcal{Z}^N)$ are arbitrary phase-space coordinates, then we can express $\mathbf{p} = \mathbf{p}(\mathcal{Z})$, $\mathbf{q} = \mathbf{q}(\mathcal{Z})$ and $H = H(\mathcal{Z}, t)$. The phase-space Lagrangian Γ is then given with respect to \mathcal{Z} , by

$$\Gamma = \Gamma_i(\mathcal{Z}, t) d\mathcal{Z}^i - H dt, \quad \Gamma_i(\mathcal{Z}, t) \equiv \mathbf{p}(\mathcal{Z}, t) \cdot \frac{\partial \mathbf{q}(\mathcal{Z})}{\partial \mathcal{Z}^i}. \quad (2.4)$$

Example 2.2.1 (Charged particle phase-space Lagrangian). To give a concrete example of the above abstract formulation, let us compute the phase-space Lagrangian of a charged particle in a background field. In this case the canonical momentum is given by $\mathbf{p} = q\mathbf{A}(\mathbf{x}, t) + m\dot{\mathbf{x}}$, so that

$$\dot{\mathbf{x}}(\mathbf{x}, \mathbf{p}, t) = \mathbf{p}/m - q\mathbf{A}(\mathbf{x}, t)/m.$$

The Hamiltonian is

$$H(\mathbf{x}, \mathbf{p}, t) = \mathbf{p} \cdot \dot{\mathbf{x}}(\mathbf{x}, \mathbf{p}, t) - L(\mathbf{x}, \dot{\mathbf{x}}(\mathbf{x}, \mathbf{p}, t), t) = \frac{1}{2m} |\mathbf{p} - q\mathbf{A}(\mathbf{x}, t)|^2 - q\Phi(\mathbf{x}, t). \quad (2.5)$$

We now choose as our phase-space coordinates $\mathcal{Z} = (\mathbf{x}, \mathbf{v})$, the position and velocity of the particle. From the general transformation rule (2.4), we then find: $\mathbf{q} = \mathbf{x}$, $\mathbf{p} = q\mathbf{A}(\mathbf{x}, t) + m\mathbf{v}$ and $H(\mathcal{Z}) = \frac{1}{2}m\mathbf{v}^2 + q\Phi(\mathbf{x}, t)$. Therefore

$$\left(\mathbf{p} \cdot \frac{\partial \mathbf{q}}{\partial \mathcal{Z}^i} \right) d\mathcal{Z}^i = p_j \delta_i^j dx^i = \mathbf{p} \cdot d\mathbf{x} = [q\mathbf{A}(\mathbf{x}, t) + m\mathbf{v}] \cdot d\mathbf{x},$$

where δ_i^j is the Kronecker delta. Hence it follows that the phase-space Lagrangian in coordinates (\mathbf{x}, \mathbf{v}) is given by

$$\Gamma = [q\mathbf{A}(\mathbf{x}, t) + m\mathbf{v}] \cdot d\mathbf{x} - \left[\frac{1}{2}m\mathbf{v}^2 + q\Phi(\mathbf{x}, t) \right] dt. \quad (2.6)$$

Given an expression for Γ in a chosen set of phase-space coordinates \mathcal{Z} , the equations of motion are obtained by minimizing the action along the path $\mathcal{Z}(t)$,

$$\mathcal{A}(\mathcal{Z}(t)) = \int_{t_0}^{t_1} [\Gamma_i(\mathcal{Z}(t), t) \dot{\mathcal{Z}}^i(t) - H(\mathcal{Z}, t)] dt.$$

Chapter 2. Guiding-centre theory

The Euler-Lagrange equations corresponding to this variational problem take the form

$$\left(\frac{\partial \Gamma_j(\mathcal{Z}, t)}{\partial \mathcal{Z}^i} - \frac{\partial \Gamma_i(\mathcal{Z}, t)}{\partial \mathcal{Z}^j} \right) \cdot \dot{\mathcal{Z}}^j = \frac{\partial \Gamma_i(\mathcal{Z}, t)}{\partial t} + \frac{\partial H(\mathcal{Z}, t)}{\partial \mathcal{Z}^i}, \quad \text{for } i = 1, \dots, N. \quad (2.7)$$

We note that for any function $S(\mathcal{Z}, t)$, adding the exact one-form $dS = \partial_i S d\mathcal{Z}^i + \partial_t S dt$ to the Lagrangian, we obtain a new Lagrangian $\hat{\Gamma}$ given by

$$\hat{\Gamma} \equiv \Gamma + dS = \left(\Gamma_i(\mathcal{Z}, t) + \frac{\partial S}{\partial \mathcal{Z}^i} \right) d\mathcal{Z}^i - \left(H - \frac{\partial S}{\partial t} \right) dt.$$

Writing $\hat{\Gamma} = \hat{\Gamma}_i d\mathcal{Z}^i - \hat{H} dt$, with

$$\hat{\Gamma}_i \equiv \Gamma_i + \frac{\partial S}{\partial \mathcal{Z}^i}, \quad \hat{H} \equiv H - \frac{\partial S}{\partial t},$$

we observe that

$$\frac{\partial \hat{\Gamma}_j}{\partial \mathcal{Z}^i} - \frac{\partial \hat{\Gamma}_i}{\partial \mathcal{Z}^j} = \left(\frac{\partial \Gamma_j}{\partial \mathcal{Z}^i} + \frac{\partial^2 S}{\partial \mathcal{Z}^i \partial \mathcal{Z}^j} \right) - \left(\frac{\partial \Gamma_i}{\partial \mathcal{Z}^j} + \frac{\partial^2 S}{\partial \mathcal{Z}^j \partial \mathcal{Z}^i} \right) = \frac{\partial \Gamma_j}{\partial \mathcal{Z}^i} - \frac{\partial \Gamma_i}{\partial \mathcal{Z}^j},$$

and

$$\frac{\partial \hat{\Gamma}_i}{\partial t} + \frac{\partial \hat{H}}{\partial \mathcal{Z}^i} = \left(\frac{\partial \Gamma_i}{\partial t} + \frac{\partial^2 S}{\partial t \partial \mathcal{Z}^i} \right) + \left(\frac{\partial H}{\partial \mathcal{Z}^i} - \frac{\partial^2 S}{\partial \mathcal{Z}^i \partial t} \right) = \frac{\partial \Gamma_i}{\partial t} + \frac{\partial H}{\partial \mathcal{Z}^i}.$$

Hence, we find the well-known fact that the equations of motion (2.7) for $\dot{\mathcal{Z}}^j$, derived from Γ and from $\hat{\Gamma} = \Gamma + dS$ are the same and hence the physics described by the Lagrangian one-form Γ is invariant with respect to the addition of a “total derivative” dS .

Case of a time-invariant symplectic part $\Gamma_i(\mathcal{Z}) d\mathcal{Z}^i$

Let us assume now that in fact $\Gamma_i = \Gamma_i(\mathcal{Z})$, does not depend explicitly on time. The Hamiltonian is still allowed to explicitly depend on time, $H = H(\mathcal{Z}, t)$. Introducing the notation $\omega_{ij} \equiv \partial \Gamma_j / \partial \mathcal{Z}^i - \partial \Gamma_i / \partial \mathcal{Z}^j$ for the coefficients in front of $\dot{\mathcal{Z}}^i$ in (2.7), we can then write the Euler-Lagrange equations more succinctly in the form

$$\dot{\mathcal{Z}}^i = X_H^i(\mathcal{Z}), \quad (2.8)$$

where X_H is the Hamiltonian vector field, which is defined as the solution of

$$\omega_{ij} X_H^j = \frac{\partial H}{\partial \mathcal{Z}^i}, \quad \text{for } i = 1, \dots, N. \quad (2.9)$$

Clearly, if $\mathcal{J}^{ij}(\mathcal{Z})$ denote the matrix coefficients of the inverse matrix of $\omega_{ij}(\mathcal{Z})$, then this can be equivalently written in the form $X_H^i(\mathcal{Z}, t) = \mathcal{J}^{ij}(\mathcal{Z}) \partial_j H(\mathcal{Z}, t)$. This suggests another common formalism; to express the equations of motion in terms of the *Poisson bracket* $\{\cdot, \cdot\}$.

The Poisson bracket acts on functions $f(\mathcal{Z}), g(\mathcal{Z})$, and is defined by

$$\{f(\mathcal{Z}), g(\mathcal{Z})\} \equiv \frac{\partial f(\mathcal{Z})}{\partial \mathcal{Z}^i} \left\{ \mathcal{Z}^i, \mathcal{Z}^j \right\} \frac{\partial g(\mathcal{Z})}{\partial \mathcal{Z}^j} \equiv \frac{\partial f(\mathcal{Z})}{\partial \mathcal{Z}^i} \mathcal{J}^{ij}(\mathcal{Z}) \frac{\partial g(\mathcal{Z})}{\partial \mathcal{Z}^j}. \quad (2.10)$$

In terms of the Poisson bracket, we can now very concisely combine (2.8), (2.9) and write the time evolution equation (2.7) (under the assumption that $\partial_t \Gamma_i = 0$), as

$$\dot{\mathcal{Z}}^i = \left\{ \mathcal{Z}^i, H \right\}, \quad \text{for } i = 1, \dots, N. \quad (2.11)$$

We note that the coefficients \mathcal{J}^{ij} of the Poisson matrix are given by inverting the matrix with coefficients $\omega_{ij} = \partial_i \Gamma_j - \partial_j \Gamma_i$, and hence the expression of the Poisson bracket in a given set of coordinates \mathcal{Z} , depend only on the so-called *symplectic part* $\Gamma_i d\mathcal{Z}^i$ of the phase-space Lagrangian Γ , but not on the Hamiltonian $H(\mathcal{Z}, t)$. More precisely, the coefficients depend on the symplectic (two-)form

$$\omega \equiv d\Gamma_i \wedge d\mathcal{Z}^i = \left(\frac{\partial \Gamma_i}{\partial \mathcal{Z}^j} - \frac{\partial \Gamma_j}{\partial \mathcal{Z}^i} \right) d\mathcal{Z}^i \wedge d\mathcal{Z}^j. \quad (2.12)$$

Note that in this time-independent case, for any function $S = S(\mathcal{Z})$, adding the closed form dS (“total derivative”) to the phase-space Lagrangian $\Gamma \rightarrow \Gamma + dS$ results in the same symplectic structure, since $d^2 S = 0$ implies that $d(\Gamma + dS) = d\Gamma = \omega$. Furthermore, the addition of dS also leaves the Hamiltonian invariant. Therefore, with this Poisson bracket formulation, the resulting equations of motion are independent of the addition of dS for time-independent S . We will employ this Poisson formalism in chapter 4, to extend the ad hoc derivation of the linearised kinetic-MHD equation for the perturbed distribution function δf of Porcelli et al. [PSK94], to allow for strong toroidal rotation. While the work [PSK94] achieved this by carrying out a direct linearization of the guiding-centre equations of motion, the abstract approach in terms of Poisson which we will present, results in a shorter argument and clarifies the meaning and origin of the adiabatic and non-adiabatic contributions to δf derived by Porcelli et al. in the non-rotating case.

As has been explained in the present section, the general phase-space Lagrangian formulation of mechanics requires knowledge of the symplectic form ω , in addition to the Hamiltonian H . Canonical coordinates are precisely those phase-space coordinates $\mathcal{Z}_c \equiv (\mathbf{p}, \mathbf{q})$, in which the symplectic form ω has a coordinate representation of the canonical form $\omega = d\mathbf{p} \wedge d\mathbf{q}$, corresponding to a phase-space Lagrangian of the form $\Gamma = \mathbf{p} \cdot d\mathbf{q} - H dt$. Compared to an approach based on canonical coordinates, the phase-space formulation allows considerably more flexibility in the choice of phase-space coordinates, and therefore enables us to choose the most suitable form for a given problem. This observation is particularly important for the Lagrangian formulation of guiding-centre theory.

2.3 Guiding-centre theory in co-moving frame

In this section, we apply the Lie perturbation theory explained in the previous section to obtain a reduced set of guiding-centre equations of motion describing the motion of a charged particle. For later applications, it will be useful to develop a set of guiding-centre equations based on the velocity variable decomposition $\mathbf{v} = \mathbf{u}(\mathbf{x}) + \mathbf{w}$, where \mathbf{u} is the (leading-order in ϵ_B) fluid velocity and \mathbf{w} can be viewed as a thermal fluctuation. This has been achieved in [Bri95] for the case of a time-independent equilibrium flow \mathbf{u} . A similar approach has been followed in [Mad10] in the case of time-varying fields, but a slightly different decomposition $\mathbf{v} = \mathbf{u}_E(\mathbf{x}, t) + \mathbf{v}_{||}$ was employed, with \mathbf{u}_E the $E \times B$ velocity so that the parallel velocity corresponds to the parallel velocity as measured in the lab frame. Here, we extend the results in [Bri95] to allow for time-varying fields. It turns out that the calculation is an almost verbatim repetition of [Mad10, appendix], with the only difference being that \mathbf{u}_E is now replaced by \mathbf{u} and the parallel guiding-centre velocity $\mathbf{v}_{||}$ is replaced by the fluctuating parallel part $w\mathbf{b}$. Along the way we have, however, found two errors in [Mad10, appendix], which do not affect the results in the main text of that work but which will be crucial for the present work. This section serves two purposes. Firstly, we explain our precise choice of guiding-centre coordinates, including higher-order corrections. Secondly, since our results disagree with published work [Mad10], we would like to present the perturbation analysis in full detail to enable the reader to convince himself of the correctness of our derivation. Furthermore, we hope that the detailed calculations included in this section may serve as a convenient reference for anyone wishing to learn about guiding-centre theory from the Lie perturbation approach.

2.3.1 Phase-space coordinates

We start our discussion of the guiding-centre transformation by choosing a more convenient set of coordinates on phase-space. To this end, we first fix unit vector fields $\mathbf{e}_1(\mathbf{x}, t), \mathbf{e}_2(\mathbf{x}, t)$ perpendicular to $\mathbf{b}(\mathbf{x}, t)$, so that $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$ forms a right-handed orthonormal basis at any point \mathbf{x} and time t . Instead of phase-space coordinates (\mathbf{x}, \mathbf{v}) , it will be more convenient to write the particle velocity relative to a given background flow velocity $\mathbf{u}(\mathbf{x}, t)$ in the following form

$$\mathbf{v} = \mathbf{u}(\mathbf{x}, t) + w_0 \mathbf{b}(\mathbf{x}, t) + \mathbf{c}_\perp(\mathbf{x}, \mu_0, \zeta_0, t), \quad (2.13)$$

in terms of a new set of particle coordinates $(\mathbf{x}, w_0, \mu_0, \zeta_0)$. Then, clearly

$$w_0 \equiv [\mathbf{v} - \mathbf{u}(\mathbf{x}, t)] \cdot \mathbf{b}(\mathbf{x}, t) \quad (2.14)$$

is the parallel particle velocity measured in a frame relative to \mathbf{u} , and \mathbf{c}_\perp is the perpendicular velocity relative to \mathbf{u} . Following [Bri95, Mad10], we have introduced a variable μ_0 , defined by

$$\mu_0 = \frac{mc_\perp^2}{2B(\mathbf{x}, t)}. \quad (2.15)$$

2.3. Guiding-centre theory in co-moving frame

Finally, the gyro-angle variable ζ_0 is defined relative to $\mathbf{e}_1, \mathbf{e}_2$ by the relationship $\mathbf{c}_\perp = -c_\perp [\sin(\zeta_0)\mathbf{e}_1 + \cos(\zeta_0)\mathbf{e}_2]$. More explicitly, from (2.15) we find that

$$\mathbf{c}_\perp(\mathbf{x}, \mu_0, \zeta_0, t) = -\sqrt{\frac{2\mu_0 B(\mathbf{x}, t)}{m}} [\sin(\zeta_0)\mathbf{e}_1(\mathbf{x}, t) + \cos(\zeta_0)\mathbf{e}_2(\mathbf{x}, t)].$$

We shall assume that \mathbf{u} is the leading-order mass flow of the plasma, whose perpendicular component is given by the $E \times B$ flow, i.e. we assume $\mathbf{u} \times \mathbf{B} = -\mathbf{E}_\perp$. The parallel component is left undefined at the moment and can be chosen depending on the particular physical application without any change to the derivation presented in this section. With this choice of particle coordinates $(\mathbf{x}, w_0, \mu_0, \zeta_0)$, the phase-space Lagrangian for charged particle motion, giving rise to the exact Lorentz equations of motion (which we will denote by γ in the following), becomes

$$\gamma = [q\mathbf{A} + m\mathbf{w} + mc_\perp] \cdot d\mathbf{x} - \left[q\Phi + \frac{1}{2}m\mathbf{w}^2 + m\mathbf{w} \cdot \mathbf{c}_\perp + \mu_0 B \right] dt. \quad (2.16)$$

Here we have denoted $\mathbf{w} \equiv \mathbf{u}(\mathbf{x}, t) + w_0 \mathbf{b}(\mathbf{x}, t)$. We have the following ordering, where the background fields vary on a typical length scale L_B :

$$w \sim u \sim c_\perp \sim v_{\text{th}}, \quad A \sim L_B B, \quad E \sim |\mathbf{u} \times \mathbf{B}| \sim v_{\text{th}} B, \quad \Phi \sim L_B E, \quad (2.17)$$

and, with $\Omega = qB/m$ the gyro-frequency, ρ_{th} the gyro-radius, it is assumed that

$$v_{\text{th}} = \rho_{\text{th}}\Omega, \quad \rho_{\text{th}}/L_B \sim \epsilon_B \ll 1, \quad \partial_t/\Omega \sim \omega/\Omega \sim \epsilon_B \ll 1, \quad (2.18)$$

where it is reminded that ϵ_B is the formal guiding-centre ordering parameter. From the above, we find

$$\frac{mw + mc_\perp}{qA} \sim \frac{mv_{\text{th}}}{qBL_B} \sim \frac{\rho_{\text{th}}}{L_B} \sim \epsilon_B \ll 1, \quad (2.19)$$

and

$$\frac{\frac{1}{2}m\mathbf{w}^2 + m\mathbf{w} \cdot \mathbf{c}_\perp + \mu_0 B}{q\Phi} \sim \frac{mv_{\text{th}}^2}{q\Phi} \sim \frac{mv_{\text{th}}^2}{qBL_B v_{\text{th}}} \sim \frac{v_{\text{th}}}{\Omega L_B} \sim \frac{\rho_{\text{th}}}{L_B} \sim \epsilon_B \ll 1. \quad (2.20)$$

Based on this ordering, we can write $\gamma = \gamma_0 + \epsilon_B \gamma_1$ where

$$\begin{cases} \gamma_0 = q\mathbf{A} \cdot d\mathbf{x} - q\Phi dt, \\ \gamma_1 = [m\mathbf{w} + mc_\perp] \cdot d\mathbf{x} - \left[\frac{1}{2}m\mathbf{w}^2 + m\mathbf{w} \cdot \mathbf{c}_\perp + \mu_0 B \right] dt. \end{cases} \quad (2.21)$$

It now becomes evident that the Lagrangian (2.16) possesses an approximate symmetry with respect to ζ_0 . More precisely, if it were not for the terms involving \mathbf{c}_\perp , all coefficients of the Lagrangian would be independent of ζ_0 which would imply the existence of a conserved quantity by Noether's theorem. In reality, there is no such exact invariant. However, since

the ζ_0 -dependent contributions occur in the higher-order ‘‘correction’’ γ_1 , there does exist an approximately conserved quantity, an *adiabatic invariant*, which is referred to as the magnetic moment μ . To lowest order, it turns out that $\mu \approx \mu_0$ with μ_0 given by (2.15). The Lie perturbation approach allows us to compute an expression for μ to arbitrary order in ϵ_B . In the present work, we will only explain the derivation of the first-order correction. For higher-order expressions and an algorithm suitable for implementation in computer algebra systems we refer the interested reader to the work of Burby et al. [BSQ13].

2.3.2 Summary of guiding-centre transformation

In section 2.4 below, we will present a detailed derivation of the guiding-centre transformation in a frame co-moving with a background flow \mathbf{u} , based on a Lie perturbation analysis up to third order in ϵ_B . As will be seen, an expansion to ϵ_B^3 is necessary (at least of certain terms) in order to uniquely determine all the components of the second-order guiding-centre Lagrangian $\Gamma = \Gamma_0 + \epsilon_B \Gamma_1 + \epsilon_B^2 \Gamma_2$ (cp. [Bri95, Mad10, BSQ13]). Here, we summarize our results.

Rotating frame

Given the right-handed orthonormal frame $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$, which was used to define the gyro-angle ζ_0 above, we define a rotating frame $(\hat{\rho}, \mathbf{b}, \hat{\perp})$ by

$$\begin{aligned}\hat{\perp}(\mathbf{X}, \zeta, t) &\equiv -\sin(\zeta)\mathbf{e}_1(\mathbf{X}, t) - \cos(\zeta)\mathbf{e}_2(\mathbf{X}, t), \\ \hat{\rho}(\mathbf{X}, \zeta, t) &\equiv \cos(\zeta)\mathbf{e}_1(\mathbf{X}, t) - \sin(\zeta)\mathbf{e}_2(\mathbf{X}, t),\end{aligned}\tag{2.22}$$

Following [CB09], we also introduce the dyadic tensors

$$\mathbf{a}_1 = -\frac{1}{2}(\hat{\rho}\hat{\perp} + \hat{\perp}\hat{\rho}), \quad \mathbf{a}_2 = \frac{1}{4}(\hat{\perp}\hat{\perp} - \hat{\rho}\hat{\rho}).\tag{2.23}$$

Particle coordinates

The guiding-centre transformation maps physical particle coordinates $(\mathbf{x}, w_0, \mu_0, \zeta_0)$ to guiding-centre coordinates $(\mathbf{X}, w, \mu, \zeta)$. The particle coordinates are defined with respect to a reference flow $\mathbf{u}(\mathbf{x}, t)$, in terms of the particle position \mathbf{x} and velocity variables (w_0, μ_0, ζ_0) . The velocity variables are defined via the particle velocity \mathbf{v} in the lab frame, as follows: First, we introduce the particle velocity, as measured in the frame co-moving with $\mathbf{u}(\mathbf{x}, t)$ by $\mathbf{w} \equiv \mathbf{v} - \mathbf{u}(\mathbf{x}, t)$. Then we set

$$w_0 \equiv \mathbf{b}(\mathbf{x}, t) \cdot \mathbf{w}, \quad \mu_0 \equiv \frac{\frac{1}{2}m|\mathbf{w}_\perp|^2}{B(\mathbf{x}, t)}, \quad \zeta_0 \equiv \arctan\left(\frac{\mathbf{w} \cdot \mathbf{e}_1(\mathbf{x}, t)}{\mathbf{w} \cdot \mathbf{e}_2(\mathbf{x}, t)}\right),\tag{2.24}$$

so that the particle velocity in the lab frame can be written as

$$\mathbf{v} = \mathbf{u}(\mathbf{x}, t) + w_0 \mathbf{b}(\mathbf{x}, t) + \sqrt{\frac{2B(\mathbf{x}, t)\mu_0}{m}} \hat{\mathbf{L}}(\mathbf{x}, t, \zeta_0), \quad (2.25)$$

with $\hat{\mathbf{L}}$ given by (2.22).

Guiding-centre Lagrangian in co-moving frame

The second-order guiding-centre Lagrangian corresponding to this guiding-centre transformation is given by (cp. (2.55), (2.59), (2.66), (2.99), (2.100)):

$$\Gamma = q\mathbf{A}^* \cdot d\mathbf{X} + \epsilon_B^2 \frac{m\mu}{q} d\zeta - H dt, \quad (2.26)$$

where

$$\mathbf{A}^* \equiv \mathbf{A} + \epsilon_B \frac{m}{q} \mathbf{W} - \epsilon_B^2 \frac{m\mu}{q^2} \left(\mathbf{R} + \frac{1}{2} [\mathbf{b} \cdot \nabla \times \mathbf{b}] \mathbf{b} \right), \quad (2.27)$$

and

$$H = q\Phi + \epsilon_B \left(\frac{1}{2} m W^2 + \mu B \right) + \epsilon_B^2 \frac{m\mu}{q} \left(\frac{1}{2} [\mathbf{b} \cdot (\nabla \times \mathbf{u}_\perp)] + S \right). \quad (2.28)$$

Here, we have defined the leading-order guiding-centre velocity $\mathbf{W} \equiv \mathbf{u}(\mathbf{X}, t) + w\mathbf{b}(\mathbf{X}, t)$ and the gyrogauge fields are given by

$$\mathbf{R} \equiv (\nabla \mathbf{e}_1) \cdot \mathbf{e}_2, \quad S \equiv \left(\frac{\partial \mathbf{e}_1}{\partial t} \right) \cdot \mathbf{e}_2. \quad (2.29)$$

Guiding-centre equations of motion

The guiding-centre equations of motion are derived from (2.26) by the variational principle (2.7). Following [CB09], we introduce

$$\mathbf{E}^* \equiv -\frac{1}{q} \nabla H - \frac{\partial \mathbf{A}^*}{\partial t}, \quad \mathbf{B}^* \equiv \nabla \times \mathbf{A}^*. \quad (2.30)$$

Then, the equations of motion for \mathbf{X} and w can be expressed in the form

$$\begin{aligned} \dot{\mathbf{X}} &= ((\mathbf{b} \cdot \mathbf{u}) + w) \frac{\mathbf{B}^*}{B_{||}^*} + \frac{\mathbf{E}^* \times \mathbf{b}}{B_{||}^*}, \\ \dot{w} &= \frac{q}{m} \frac{\mathbf{B}^* \cdot \mathbf{E}^*}{B_{||}^*}, \end{aligned} \quad (2.31)$$

with $B_{||}^* \equiv \mathbf{b} \cdot \mathbf{B}^*$.

The equations of motion (2.31) including higher-order corrections at order ϵ_B^2 are considered in more detail in chapter 3, in the limit of vanishing background flow $\mathbf{u} \equiv 0$, in the context of numerical slowing-down simulations.

Expanding the equation of motion for \mathbf{X} in (2.31) to first order in ϵ_B , it can be shown [Mad10, eq.(49)] that (2.31) can be written in the more familiar form

$$\dot{\mathbf{X}} = \mathbf{W} + \epsilon_B \left[\frac{\mu}{qB} \mathbf{b} \times \nabla B + \frac{1}{\Omega} \mathbf{b} \times \left(\frac{\partial}{\partial t} + \mathbf{W} \cdot \nabla \right) \mathbf{W} \right] + O(\epsilon_B^2), \quad (2.32)$$

with $\Omega = \Omega(\mathbf{X}, t) \equiv \frac{m}{qB(\mathbf{X}, t)}$ the gyrofrequency at the guiding-centre position \mathbf{X} . In (2.32), we can clearly identify the leading-order guiding-centre velocity $\mathbf{W} = \mathbf{u} + w\mathbf{b}$ (which includes the $E \times B$ drift perpendicular to the magnetic field via $\mathbf{u}_\perp = E \times \mathbf{B}/B^2$), as well as the well-known ∇B -drift. In the absence of a strong background flow ($\mathbf{u} \equiv 0$), the last term is given by $\Omega^{-1} w^2 \mathbf{b} \times (\mathbf{b} \cdot \nabla \mathbf{b})$, which is therefore identified as the curvature drift in limit $\mathbf{u} \approx 0$.

The approximate form (2.32) of the guiding-centre equations of motion with strong background flow will be one ingredient in our derivation of higher-order (gyroviscous) corrections to the pressure tensor in chapter 4. It turns out that the correct form of these gyroviscous corrections requires not only the equations of motion, but crucially depends on higher-order Larmor-radius corrections in the particle position (expressing deviations from circular Larmor motion about its guiding-centre). We summarize the relation between particle coordinates and the coordinates of the corresponding guiding-centre next.

Relation between guiding-centre and particle coordinates

To the accuracy that is required to obtain the second-order guiding-centre Lagrangian, the guiding-centre transformation is given in terms of phase-space generating vector fields G_1 and G_2 according to (2.54), below, as follows:

$$\begin{aligned} \mathbf{X} &= \mathbf{x} + \epsilon_B G_1^\mathbf{X} + \epsilon_B^2 \left(G_2^\mathbf{X} + \frac{1}{2} G_1 \cdot dG_1^\mathbf{X} \right) + O(\epsilon_B^3), \\ w &= w_0 + \epsilon_B G_1^w + O(\epsilon_B^2), \\ \mu &= \mu_0 + \epsilon_B G_1^\mu + O(\epsilon_B^2), \\ \zeta &= \zeta_0 + \epsilon_B G_1^\zeta + O(\epsilon_B^2), \end{aligned} \quad (2.33)$$

We recall that the terms on the right hand side of (2.33) need to be evaluated at the *particle coordinates*, i.e. by formally substituting $(\mathbf{X}, w, \mu, \zeta) \rightarrow (\mathbf{x}, w_0, \mu_0, \zeta_0)$ in the expressions (2.37)-(2.42).

The inverse mapping from guiding-centre to particle coordinates is obtained according to

(2.53) as follows:

$$\begin{aligned}\mathbf{x} &= \mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 + \epsilon_B^2 \boldsymbol{\rho}_1 + O(\epsilon_B^3), \\ w_0 &= w - \epsilon_B G_1^w + O(\epsilon_B^2), \\ \mu_0 &= \mu - \epsilon_B G_1^\mu + O(\epsilon_B^2), \\ \zeta_0 &= \zeta - \epsilon_B G_1^\zeta + O(\epsilon_B^2),\end{aligned}\tag{2.34}$$

where the right-hand side coefficients (2.37)-(2.42) are now evaluated at the *guiding-centre coordinates* $(\mathbf{X}, w, \mu, \zeta)$, and where we have used the notation $\boldsymbol{\rho}_0 = -G_1^X$ and

$$\boldsymbol{\rho}_1 \equiv -G_2^X + \frac{1}{2} G_1 \cdot dG_1^X.\tag{2.35}$$

After some algebraic manipulation detailed in the appendix, the higher-order displacement (2.35) can be written more conveniently in the form

$$\boldsymbol{\rho}_1 = -\left(g_\mu \frac{\partial}{\partial \mu} + g_\zeta \frac{\partial}{\partial \zeta}\right) \boldsymbol{\rho}_0 - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \boldsymbol{\rho}_0 - \left(G_{2,\parallel}^X + \frac{1}{2} \boldsymbol{\rho}_0 \cdot \nabla \mathbf{b} \cdot \boldsymbol{\rho}_0\right) \mathbf{b}.\tag{2.36}$$

Coefficients of generating vector fields

We now collect the required components of G_1 , G_2 . The spatial component of G_1 is given by (cp. (2.67))

$$G_1^X = -\boldsymbol{\rho}_0 = -\sqrt{\frac{2\mu}{q\Omega}} \hat{\boldsymbol{\rho}}(\mathbf{X}, t, \zeta).\tag{2.37}$$

We recall that $\hat{\boldsymbol{\rho}}$ is the rotating unit vector in direction of $\boldsymbol{\rho}_0$, defined by (2.22). From (2.74), (2.75) and (2.98), we obtain

$$G_1^w = \frac{\mu}{q} [\mathbf{b} \cdot (\nabla \times \mathbf{b}) + \mathbf{a}_1 : \nabla \mathbf{b}] - w \boldsymbol{\rho}_0 \cdot \boldsymbol{\kappa} + \boldsymbol{\rho}_0 \times [\mathbf{b} \times (\nabla \times \mathbf{u})].\tag{2.38}$$

This expression involves the dyadic tensor \mathbf{a}_1 defined in (2.23). The first-order correction to the parallel guiding-centre velocity G_1^w is expressed in terms of the field line curvature $\boldsymbol{\kappa} \equiv \mathbf{b} \cdot \nabla \mathbf{b}$, the field strength gradient ∇B and the magnetic field line twist $\tau \equiv \mathbf{b} \cdot \nabla \times \mathbf{b}$, as well as the curl $\nabla \times \mathbf{u}$ of the background velocity \mathbf{u} . We note that, given two dyadic tensors \mathbf{C} and \mathbf{D} , we denote by

$$\mathbf{C} : \mathbf{D} \equiv \sum_{i,j=1}^3 C_{ij} D_{ij},$$

the contraction on both indices, where C_{ij} , D_{ij} are the components of \mathbf{C} and \mathbf{D} with respect to an orthonormal basis.

Chapter 2. Guiding-centre theory

Collecting (2.72), (2.84) and (2.93), we find

$$G_1^\mu = -\frac{\mu}{\Omega} (\mathbf{b} \cdot (\nabla \times \mathbf{W}) + \mathbf{a}_1 : \nabla \mathbf{W}) + \frac{m}{B} \boldsymbol{\rho}_0 \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} \right) + \mu \boldsymbol{\rho}_0 \cdot \nabla \log B. \quad (2.39)$$

This expression for G_1^μ corrects an error in [Mad10, eq. (A24)], which mistakenly contains an additional factor of $\frac{1}{2}$ in front of the $\partial \mathbf{W} / \partial t$ term.

Equations (2.73) and (B.6) yield

$$G_1^\zeta = -\boldsymbol{\rho}_0 \cdot \mathbf{R} - \frac{1}{\Omega} \mathbf{a}_2 : \nabla \mathbf{W} - \frac{q}{2\mu\Omega} (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} + \frac{2\mu}{3m} \nabla B \right). \quad (2.40)$$

Finally, we write down the spatial component $G_2^X = G_{2,\perp}^X + G_{2,\parallel}^X \mathbf{b}$. The perpendicular contribution is given by (2.76)

$$G_{2,\perp}^X = \frac{1}{\Omega} [\mathbf{b} \cdot (\nabla \times \mathbf{W})] \boldsymbol{\rho}_0 + \frac{1}{2\Omega} \mathbf{b} \times \left[g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right], \quad (2.41)$$

where

$$g_\mu = G_1^\mu - \mu \boldsymbol{\rho}_0 \cdot \nabla \log B,$$

$$g_\zeta = G_1^\zeta + \boldsymbol{\rho}_0 \cdot \mathbf{R}.$$

According to (2.96), the parallel component $G_{2,\parallel}^X$ is

$$G_{2,\parallel}^X = -\frac{1}{\Omega} \boldsymbol{\rho}_0 \cdot \left\{ \mathbf{b} \times \left[\mathbf{b} \cdot \nabla \mathbf{W} + \mathbf{W} \cdot \nabla \mathbf{b} + \frac{\partial \mathbf{b}}{\partial t} \right] \right\} + \frac{\mu}{q\Omega} \mathbf{a}_2 : \nabla \mathbf{b}. \quad (2.42)$$

Here the dyadic tensor \mathbf{a}_2 is defined by (2.23), above.

This expression for the second order components G_2^X corrects two errors in [Mad10, eq. (A35)], where the $\frac{\partial \mathbf{b}}{\partial t}$ contribution to $G_{2,\parallel}^X$ contains an erroneous additional factor of $\frac{1}{2}$, and the last term in [Mad10, eq. (A35)] does not gyro-average to zero.

The expressions for G_1^μ and G_2^X presented here will be crucial for the derivation of the correct form of the gyroviscous components of the pressure tensor in chapter 4. Only with these corrections does the derivation from guiding-centre theory recover the corresponding result based on fluid theory [Mac65].

2.4 Derivation of guiding-centre transformation

2.4.1 Phase-space transformations induced by a vector field

Given a vector field G on phase-space M , we associate to it a flow $\phi_s : M \rightarrow M$, defined by following the vector field G for a time s , i.e. ϕ_s is defined as the solution of the system of ODEs

$$\frac{d\phi_s(\mathcal{Z})}{ds} = G(\phi_s(\mathcal{Z})), \quad \phi_0(\mathcal{Z}) = \mathcal{Z}, \quad \forall \mathcal{Z} \in M.$$

For fixed s , the mapping ϕ_s is a bijective phase-space transformation (diffeomorphism) and we can use it to define new phase-space coordinates $\bar{\mathcal{Z}}$ by the relation $\bar{\mathcal{Z}} \equiv \phi_1(\mathcal{Z})$, i.e. following G up to time $s = 1$.

If $G = O(\epsilon)$ is formally small in an expansion parameter ϵ , then the difference between $\bar{\mathcal{Z}}$ and \mathcal{Z} will be small as well, in the sense that the distance between $\bar{\mathcal{Z}}$ and \mathcal{Z} is $O(\epsilon)$. In this case, we introduce a formal ordering parameter (or tag) ϵ , which in reality is set to $\epsilon = 1$, but is used to keep track of formally small quantities in our perturbation analysis. To emphasize that G is considered small, we would then write ϵG rather than G , and denote

$$\bar{\mathcal{Z}} = \phi_\epsilon(\mathcal{Z}).$$

In this case, we call ϕ_ϵ a *near-identity transformation*, since to lowest order in ϵ we find $\bar{\mathcal{Z}}^\alpha \approx \mathcal{Z}^\alpha + \epsilon G^\alpha(\mathcal{Z}) + O(\epsilon^2)$, and ϕ_ϵ differs from the identity only by terms of order ϵ and higher.

The basic idea of Lie perturbation theory is to choose the vector field ϵG in a suitable way so that the dynamics in terms of the new coordinates $\bar{\mathcal{Z}}$ is simpler than the original dynamics in coordinates \mathcal{Z} . If such a reduction in complexity can be achieved and a suitable ϵG can be found, then one can solve the simpler equations of motion for $\bar{\mathcal{Z}}$ and obtain the corresponding evolution of \mathcal{Z} from the inverse relationship $\mathcal{Z} = \phi_{-\epsilon}(\bar{\mathcal{Z}})$. To find equations that can be used to determine a suitable ϵG , we first need to consider how geometric quantities such as functions, vector fields and differential forms transform under the near-identity transformation induced by ϵG . We begin by explaining this in detail for a given phase-space function $f(\mathcal{Z})$. To such a function f , we can associate the corresponding function $\mathcal{T}^\epsilon f$ (it's pull-back), which is written in new variables $\bar{\mathcal{Z}}$, by requiring that $[\mathcal{T}^\epsilon f](\bar{\mathcal{Z}}) \equiv f(\mathcal{Z})$, or more explicitly

$$[\mathcal{T}^\epsilon f](\bar{\mathcal{Z}}) = [(\phi_\epsilon)^* f](\bar{\mathcal{Z}}) \equiv f(\phi_{-\epsilon}(\bar{\mathcal{Z}}))$$

is the pull-back of f from \mathcal{Z} to $\bar{\mathcal{Z}}$. Then, to leading order in ϵ ,

$$\begin{aligned} [\mathcal{T}^\epsilon f](\bar{\mathcal{Z}}) &\approx f(\phi_{-s}(\bar{\mathcal{Z}}))|_{s=0} + \epsilon \frac{d\phi_{-s}}{ds}\Big|_{s=0} \cdot df(\phi_{-s}(\bar{\mathcal{Z}}))|_{s=0} + O(\epsilon^2) \\ &= f(\bar{\mathcal{Z}}) - \epsilon(G \cdot df)|_{\bar{\mathcal{Z}}} + O(\epsilon^2), \end{aligned}$$

which is usually expressed more simply, yet slightly ambiguously, in the form $\mathcal{T}^\epsilon f = f - \epsilon(G \cdot$

$\mathrm{d}f) + O(\epsilon^2)$. To go beyond first order, we consider any $0 \leq s \leq \epsilon$:

$$\partial_s \mathcal{T}^s f = \partial_s \left[(\phi_s^{-1})^* f \right] = \lim_{\Delta s \rightarrow 0} \frac{(\phi_{s+\Delta s}^{-1})^* f - (\phi_s^{-1})^* f}{\Delta s} = (\phi_s^{-1})^* \left(\lim_{\Delta s \rightarrow 0} \frac{(\phi_{\Delta s}^{-1})^* f - f}{\Delta s} \right),$$

where we have used that the flow and its inverse satisfy $\phi_{s+\Delta s}^{-1} = \phi_s^{-1} \circ \phi_{\Delta s}^{-1}$, so that $(\phi_{s+\Delta s}^{-1})^* f = (\phi_s^{-1})^* ((\phi_{\Delta s}^{-1})^* f)$. The limit on the right-hand side can be expressed in terms of G as a Lie derivative

$$-\epsilon \mathcal{L}_G f = \lim_{\Delta s \rightarrow 0} \frac{(\phi_{\Delta s}^{-1})^* f - f}{\Delta s},$$

so that \mathcal{T}^s is found to satisfy the following relation

$$\partial_s \mathcal{T}^s f = -\epsilon \mathcal{T}^s \mathcal{L}_G f. \quad (2.43)$$

Since, \mathcal{T}^s and \mathcal{L}_G commute, we find the formal solution for \mathcal{T}^s in terms of G as

$$\mathcal{T}^\epsilon = \exp(-\epsilon \mathcal{L}_G) = \sum_{k=0}^{\infty} \frac{(-\epsilon \mathcal{L}_G)^k}{k!}. \quad (2.44)$$

Since \mathcal{T}^s and \mathcal{L}_G also commute with the exterior derivative d , one can show that formula (2.44) remains true also for differential forms, and therefore in particular the Lagrangian 1-form [BSQ13].

This gives us the required recipe for determining the form of $\mathcal{T}^s \gamma$ for a given function, or more generally for a differential form γ , in new coordinates $\bar{\mathcal{Z}}$ to any desired order. Given an explicit functional form for γ : We first find $\mathcal{L}_G^k \gamma$ for all required k , and then evaluate

$$(\mathcal{T}^\epsilon \gamma)|_{\bar{\mathcal{Z}}} = \sum_k \frac{(-\epsilon \mathcal{L}_G)^k \gamma}{k!} \Big|_{\bar{\mathcal{Z}}}.$$

With formula (2.44) at hand, we can now proceed to write down the **transformation rules** required by Lie perturbation theory. We consider a Lagrangian one-form γ which we assume to be written in the form $\gamma = \gamma_0 + \epsilon \gamma_1$, where γ_0 is a leading-order term and $\epsilon \gamma_1$ is a formally small first-order correction. Rather than using a near-identity transformation induced by a single generating vector field ϵG , we consider a family of vector fields at different orders $\epsilon G_1, \epsilon^2 G_2, \epsilon^3 G_3, \dots$ with associated near-identity transformations $\phi_{\ell, \epsilon}$ pull-back operators $\mathcal{T}_\ell^\epsilon \equiv \exp(-\epsilon^\ell \mathcal{L}_{G_\ell})$, where $\mathcal{L}_\ell \equiv \mathcal{L}_{G_\ell}$, for $\ell = 1, 2, 3, \dots$. In this way, the transformation can be carried out order-by-order and we define the total transformation to be given by

$$\bar{\mathcal{Z}} \equiv \phi_\epsilon(\mathcal{Z}) := [\cdots \circ \phi_{3,\epsilon} \circ \phi_{2,\epsilon} \circ \phi_{1,\epsilon}] (\mathcal{Z}),$$

i.e. to pass from \mathcal{Z} to $\bar{\mathcal{Z}}$, we start at \mathcal{Z} and first follow G_1 , then G_2 , followed by G_3 and so forth.

The corresponding transformation operator \mathcal{T}^ϵ is given by

$$\begin{aligned}\mathcal{T}^\epsilon &= [\dots \circ \exp(-\epsilon^3 \mathcal{L}_3) \circ \exp(-\epsilon^2 \mathcal{L}_2) \circ \exp(-\epsilon \mathcal{L}_1)] \\ &= 1 - \epsilon \mathcal{L}_1 - \epsilon^2 \left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) - \epsilon^3 \left(\mathcal{L}_3 - \mathcal{L}_2 \mathcal{L}_1 + \frac{1}{6} \mathcal{L}_1^3 \right) + O(\epsilon^4),\end{aligned}\tag{2.45}$$

where the last expression is obtained from straight-forward expansion of the exponentials.

The inverse transformation from $\bar{\mathcal{Z}}$ to \mathcal{Z} coordinates and to a given order N , is obtained by starting at \bar{Z} and following first $-\epsilon^N G_N$, then $-\epsilon^{N-1} G_{N-1}$, and so forth until finally we follow $-\epsilon G_1$ to find \bar{Z} , i.e. $\mathcal{Z} \approx [\phi_{1,-\epsilon} \circ \dots \circ \phi_{N-1,-\epsilon} \circ \phi_{N,-\epsilon}] (\bar{\mathcal{Z}})$. The associated transformation operator is correspondingly given by

$$\begin{aligned}[\mathcal{T}^\epsilon]^{-1} &= [\exp(\epsilon \mathcal{L}_1) \circ \exp(\epsilon^2 \mathcal{L}_2) \circ \exp(\epsilon^3 \mathcal{L}_3) \circ \dots] \\ &= 1 + \epsilon \mathcal{L}_1 + \epsilon^2 \left(\mathcal{L}_2 + \frac{1}{2} \mathcal{L}_1^2 \right) + \epsilon^3 \left(\mathcal{L}_3 + \mathcal{L}_1 \mathcal{L}_2 + \frac{1}{6} \mathcal{L}_1^3 \right) + O(\epsilon_B^4).\end{aligned}\tag{2.46}$$

We thus find that the inverse relation (2.46) is not simply (2.45) with reversed signs of the corrections terms, as one might think. This perhaps subtle difference between the (forward and backward) transformation operators will turn out to be crucial in the correct evaluation of the gyroviscous corrections of the pressure moment from guiding-centre theory in chapter 4 below.

Transformation rules for Lagrangian

We return to the Lagrangian one-form $\gamma = \gamma_0 + \epsilon \gamma_1$ in coordinates \mathcal{Z} . Note that this is precisely the form taken by the Lagrangian in physical coordinates (2.21), which gives rise to the Lorentzian equations of motion. Following the above discussion, we can now write the corresponding expression in the transformed coordinates $\bar{\mathcal{Z}}$. As discussed previously, the physical Lagrangian $\gamma = \gamma_0 + \epsilon_B \gamma_1$ exhibits a gyro-symmetry only to 0-th order. Our goal is ultimately to find a suitable set of coordinates $\bar{\mathcal{Z}}$ (corresponding to guiding-centre coordinates), in which the transformed Lagrangian $\Gamma = \Gamma_0 + \epsilon_B \Gamma_1 + \epsilon_B^2 \Gamma_2 + \dots$ is gyro-symmetric to higher order in ϵ_B . To this end, we first write down the general expression for Γ in terms of arbitrary generating vector fields G_1, G_2, \dots and then seek to determine G_1, G_2, \dots for which $\Gamma_1, \Gamma_2, \dots$ are gyro-symmetric, i.e. do not depend on the (transformed) gyro-angle ζ .

Using (2.45), we find

$$\Gamma = \mathcal{T}^\epsilon \gamma = \Gamma_0 + \epsilon \Gamma_1 + \epsilon^2 \Gamma_2 + \epsilon^3 \Gamma_3 + \dots$$

with

$$\Gamma_0 = \gamma_0 + dS_0, \quad (2.47)$$

$$\Gamma_1 = \gamma_1 - \mathcal{L}_1 \gamma_0 + dS_1, \quad (2.48)$$

$$\Gamma_2 = -\mathcal{L}_1 \gamma_1 - \left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) \gamma_0 + dS_2, \quad (2.49)$$

$$\Gamma_3 = -\left(\mathcal{L}_2 - \frac{1}{2} \mathcal{L}_1^2 \right) \gamma_1 - \left(\mathcal{L}_3 - \mathcal{L}_2 \mathcal{L}_1 + \frac{1}{6} \mathcal{L}_1^3 \right) \gamma_0 + dS_3. \quad (2.50)$$

⋮

Here, we have also introduced closed forms dS_0, dS_1, \dots , which represent additional gauge terms that can be used to simplify the resulting form of the transformed Lagrangian but which will not affect the physics (see the discussion following (2.7)). The expressions (2.49) and (2.50) can be simplified somewhat by noting that from (2.48) we have, up to the physically irrelevant closed forms, $\mathcal{L}_1 \gamma_0 = \gamma_1 - \Gamma_1$. Inserting in (2.49) yields

$$\Gamma_2 = -\mathcal{L}_2 \gamma_0 - \frac{1}{2} \mathcal{L}_1 (\gamma_1 + \Gamma_1) + dS_2. \quad (2.51)$$

Similarly, we find that (2.50) can be rewritten, neglecting contributions due to exact forms, using

$$\begin{cases} -\mathcal{L}_2 \gamma_1 + \mathcal{L}_2 \mathcal{L}_1 \gamma_0 = -\mathcal{L}_2 \gamma_1 + \mathcal{L}_2 (\gamma_1 - \Gamma_1) = -\mathcal{L}_2 \Gamma_1, \\ \frac{1}{2} \mathcal{L}_1^2 \gamma_1 - \frac{1}{6} \mathcal{L}_1^3 \gamma_0 = \frac{1}{2} \mathcal{L}_1^2 \gamma_1 - \frac{1}{6} \mathcal{L}_1^2 (\gamma_1 - \Gamma_1) = \frac{1}{3} \mathcal{L}_1^2 (\gamma_1 + \frac{1}{2} \Gamma_1). \end{cases}$$

From which it follows that

$$\Gamma_3 = -\mathcal{L}_3 \gamma_0 - \mathcal{L}_2 \Gamma_1 + \frac{1}{3} \mathcal{L}_1^2 \left(\gamma_1 + \frac{1}{2} \Gamma_1 \right) + dS_3. \quad (2.52)$$

Equations (2.47), (2.48), (2.51), (2.52) express the transformed Lagrangian up to third-order in ϵ , in terms of the generating vector fields G_1, G_2, G_3 and the phase-space gauge functions S_0, \dots, S_3 . The Lie perturbation approach now consists of fixing a desired form for the transformed Lagrangian Γ , and solving for G_ℓ, S_ℓ order-by-order in ϵ . Usually, the form of Γ is restricted by solvability conditions for the generating vector fields and gauge functions. However, even after these solvability conditions are taken into account, there will generally be many different forms of the resulting Lagrangian, corresponding to different choices of the transformed variables $\bar{\mathcal{Z}}$. This non-uniqueness has led to some confusion in the past, but any apparent contradictions between different formulations are resolved when carefully taking into account the associated phase-space transformations, as has been shown in [BSQ13].

Applying (2.45) to the coordinate functions $\mathcal{Z}^\alpha = [\phi_{-\epsilon}(\bar{\mathcal{Z}})]^\alpha = [\mathcal{T}^\epsilon] \bar{\mathcal{Z}}^\alpha$, we find that

$$\mathcal{Z}^\alpha = \bar{\mathcal{Z}}^\alpha - \epsilon G_1^\alpha - \epsilon^2 \left(G_2^\alpha - \frac{1}{2} G_1 \cdot dG_1^\alpha \right) + O(\epsilon^3), \quad (2.53)$$

where the right-hand side terms are evaluated at $\bar{\mathcal{Z}}$. Employing the inverse transformation (2.54) to $\bar{\mathcal{Z}}^\alpha = [\phi_\epsilon(\mathcal{Z})]^\alpha = [\mathcal{T}^\epsilon]^{-1} \mathcal{Z}^\alpha$, we find

$$\bar{\mathcal{Z}}^\alpha = \mathcal{Z}^\alpha + \epsilon G_1^\alpha + \epsilon^2 \left(G_2^\alpha + \frac{1}{2} G_1 \cdot dG_1^\alpha \right) + O(\epsilon_B^3), \quad (2.54)$$

and the right-hand side terms are evaluated at \mathcal{Z} .

2.4.2 Order-by-order analysis

The guiding-centre transformation is based on the approximately gyro-symmetric Lagrangian $\gamma = \gamma_0 + \epsilon_B \gamma_1$ given by equation (2.21). The idea is to find a coordinate transformation such that the transformed Lagrangian will be gyro-symmetric to a higher order in ϵ_B . Then, the transformed Lagrangian is truncated at the chosen order, which yields an approximate Lagrangian with an exact gyro-symmetry. The existence of this symmetry implies that a reduced description of the dynamics can be obtained, which despite the reduction in complexity still is formally correct to a given order in ϵ_B . The mathematical procedure is an application of the Lie perturbation theory approach outlined in section 2.4.1.

Two basic formulae (equations (α) and (β))

Before beginning our order-by-order analysis, we establish the following equations for phase-space functions $\alpha, \beta = (\beta_1, \beta_2, \beta_3)$:

$$i_G d(\alpha dt) = \left[G^X \cdot \nabla \alpha + G^w \frac{\partial \alpha}{\partial w} + G^\mu \frac{\partial \alpha}{\partial \mu} + G^\zeta \frac{\partial \alpha}{\partial \zeta} \right] dt, \quad (\alpha)$$

and

$$\begin{aligned} i_G d(\beta \cdot dX) &= \left[- [G^X \times (\nabla \times \beta)] + G^w \frac{\partial \beta}{\partial w} + G^\mu \frac{\partial \beta}{\partial \mu} + G^\zeta \frac{\partial \beta}{\partial \zeta} \right] \cdot dX \\ &\quad - G^X \cdot \frac{\partial \beta}{\partial w} dw - G^X \cdot \frac{\partial \beta}{\partial \mu} d\mu - G^X \cdot \frac{\partial \beta}{\partial \zeta} d\zeta - G^X \cdot \frac{\partial \beta}{\partial t} dt. \end{aligned} \quad (\beta)$$

These formulae will be used multiple times in the following.

The derivation follows directly from the definition of the exterior derivative d and the inner product i_G . To see (α), we first note that $d(\alpha dt) = d\alpha \wedge dt$ and $i_G d(\alpha dt) = (i_G d\alpha) \wedge dt - d\alpha \wedge (i_G dt) = (i_G d\alpha) \wedge dt$, since $G^t = 0$. Now,

$$\begin{aligned} (i_G d\alpha) \wedge dt &= \left(i_G \left[\frac{\partial \alpha}{\partial X^i} dX^i + \frac{\partial \alpha}{\partial w} dw + \frac{\partial \alpha}{\partial \mu} d\mu + \frac{\partial \alpha}{\partial \zeta} d\zeta + \frac{\partial \alpha}{\partial t} dt \right] \right) \wedge dt \\ &= \left[\frac{\partial \alpha}{\partial X^i} G^{X^i} + \frac{\partial \alpha}{\partial w} G^w + \frac{\partial \alpha}{\partial \mu} G^\mu + \frac{\partial \alpha}{\partial \zeta} G^\zeta + \frac{\partial \alpha}{\partial t} G^t \right] dt, \end{aligned}$$

implies (α) . To see (β) , we note that

$$\begin{aligned} d(\boldsymbol{\beta} \cdot d\mathbf{X}) &= d\beta_i \wedge dX^i \\ &= \frac{\partial \beta_i}{\partial X^j} dX^j \wedge dX^i + \frac{\partial \beta_i}{\partial w} dw \wedge dX^i + \frac{\partial \beta_i}{\partial \mu} d\mu \wedge dX^i \\ &\quad + \frac{\partial \beta_i}{\partial \zeta} d\zeta \wedge dX^i + \frac{\partial \beta_i}{\partial t} dt \wedge dX^i. \end{aligned}$$

Taking the interior product of the individual terms with G , first yields

$$\begin{aligned} i_G \left(\frac{\partial \beta_i}{\partial X^j} dX^j \wedge dX^i \right) &= \frac{\partial \beta_i}{\partial X^j} (i_G dX^j) \wedge dX^i - \frac{\partial \beta_i}{\partial X^j} dX^j \wedge (i_G dX^i) \\ &= \frac{\partial \beta_i}{\partial X^j} G^{X^j} dX^i - \frac{\partial \beta_i}{\partial X^j} G^{X^i} dX^j \\ &= \left(\frac{\partial \beta_i}{\partial X^j} - \frac{\partial \beta_j}{\partial X^i} \right) G^{X^j} dX^i. \end{aligned}$$

Employing the vector identity $G^X \times (\nabla \times \mathbf{A}) = (\nabla \mathbf{A}) \cdot G^X - G^X \cdot (\nabla \mathbf{A})$, we find

$$i_G \left(\frac{\partial \beta_i}{\partial X^j} dX^j \wedge dX^i \right) = -[G^X \times (\nabla \times \boldsymbol{\beta})] \cdot d\mathbf{X}.$$

Similarly, the w -components can be obtained from

$$\begin{aligned} i_G \left(\frac{\partial \beta_i}{\partial w} dw \wedge dX^i \right) &= \frac{\partial \beta_i}{\partial w} (i_G dw) \wedge dX^i - \frac{\partial \beta_i}{\partial w} dw \wedge (i_G dX^i) \\ &= \frac{\partial \beta_i}{\partial w} G^w dX^i - \frac{\partial \beta_i}{\partial w} G^{X^i} dw. \end{aligned}$$

The terms involving μ, ζ are obtained by substituting $w \rightarrow \mu, \zeta$. The temporal component is found by substituting $w \rightarrow t$ and noting that $G^t = 0$, i.e.

$$i_G \left(\frac{\partial \beta_i}{\partial t} dt \wedge dX^i \right) = -\frac{\partial \beta_i}{\partial t} G^{X^i} dt.$$

Combining these expressions yields (β) . With (α) and (β) at our disposal we now give a detailed derivation of the guiding-centre transformation via an order-by-order analysis.

0th-order analysis

The lowest-order Lagrangian is already gyro-symmetric so that we can simply choose $\Gamma_0 = \gamma_0$ with gauge function $S_0 = 0$, so that

$$\Gamma_0 = q\mathbf{A} \cdot d\mathbf{X} - q\Phi dt. \tag{2.55}$$

Before proceeding to higher orders, we point out that the transformed Lagrangian is only required *up to closed forms*. Since the transformation equations (2.48)-(2.52) are expressed in

terms of Lie derivatives \mathcal{L}_G and since we have Cartan's formula for differential forms

$$\mathcal{L}_G \gamma = i_G d\gamma + d(i_G \gamma),$$

which expresses \mathcal{L}_G as a sum of a differential form $i_G d\gamma$ and a closed form $d(i_G \gamma)$, we can in practice substitute $\mathcal{L}_G \doteq i_G d$ [Lit82], which will simplify the following analysis considerably.

1st-order analysis

From equation (2.48), we find that the first-order transformed Lagrangian is given by $\Gamma_1 = -\mathcal{L}_{G_1} \gamma_0 + \gamma_1$. A very detailed calculation of the first term on the right-hand side is now given. First, we recall that $\mathcal{L}_G \gamma_0 \doteq i_G d\gamma_0$. Then, we apply (α) and (β) to find

$$i_{G_1} d\gamma_0 = i_{G_1} d(q\mathbf{A} \cdot d\mathbf{X}) - i_{G_1} d(q\Phi dt) = [-G_1^X \times (\nabla \times \mathbf{A})] \cdot d\mathbf{X} + G_1^X \cdot \left(-\frac{\partial \mathbf{A}}{\partial t} - \nabla \Phi \right) dt.$$

Recalling that $\mathbf{B} = \nabla \times \mathbf{A}$ and $\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla \Phi$, we can write this in the compact form

$$i_G d\gamma_0 = -q[G^X \times \mathbf{B}] \cdot d\mathbf{X} + qG^X \cdot \mathbf{E} dt. \quad (2.56)$$

From (2.59), (2.21) and (2.56), we find

$$\Gamma_1 = [m\mathbf{W} + m\mathbf{c}_\perp + qG_1^X \times \mathbf{B}] \cdot d\mathbf{X} - \left[\frac{1}{2}mW^2 + m\mathbf{W} \cdot \mathbf{c}_\perp + \mu B + qG^X \cdot \mathbf{E} \right] dt,$$

in the transformed coordinates $(\mathbf{X}, w, \mu, \zeta)$. To arrive at this form, we have chosen $S_1 = 0$. In these expressions, the terms involving \mathbf{c}_\perp are clearly dependent on ζ . To make the $d\mathbf{X}$ components gyro-symmetric, we thus require that $qG_1^X \times \mathbf{B} = -m\mathbf{c}_\perp$. We can clearly identify

$$G_{1,\perp}^X \equiv \frac{\mathbf{c}_\perp \times \mathbf{b}}{\Omega} = -\boldsymbol{\rho}_0, \quad (2.57)$$

as the negative Larmor radius vector $\boldsymbol{\rho}_0$, which we already found in the case of a homogeneous B-field. The dt coefficient of Γ_1 now contains the a priori gyro-angle dependent combination

$$m\mathbf{W} \cdot \mathbf{c}_\perp + q\mathbf{E} \cdot G_1^X = m \left[\mathbf{u}_\perp \cdot \mathbf{c}_\perp + \mathbf{E} \cdot \left(\mathbf{c}_\perp \times \frac{\mathbf{b}}{B} \right) \right] = m \left[\mathbf{u}_\perp + \frac{\mathbf{E} \times \mathbf{b}}{B} \right] \cdot \mathbf{c}_\perp.$$

This term will be gyro-angle independent if the term in square brackets vanishes, i.e. provided that

$$\mathbf{E}_\perp + \mathbf{u} \times \mathbf{B} = 0 \quad \Leftrightarrow \quad \mathbf{u}_\perp = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (2.58)$$

We will henceforth assume \mathbf{u} to satisfy (2.58). In contrast to [Mad10], we will allow for $\mathbf{u}_{||} \neq 0$.

With $G_{1,\perp}^X$ chosen according to (2.57) and assuming (2.58), we obtain

$$\Gamma_1 = m\mathbf{W} \cdot d\mathbf{X} - \left[\frac{1}{2}mW^2 + \mu B \right] dt. \quad (2.59)$$

The other components $G_{1,\parallel}^X$, G_1^w , G_1^μ and G_1^ζ are not determined at this order, so that knowledge of the higher-order corrections in the guiding-centre coordinates require consideration of higher-order contributions to Γ .

2nd-order analysis

To the next order, we recall that by equation (2.51):

$$\Gamma_2 = -\mathcal{L}_{G_2}\gamma_0 - \frac{1}{2}\mathcal{L}_{G_1}(\gamma_1 + \Gamma_1) + dS_2. \quad (2.60)$$

From (2.56), we immediately find

$$-\mathcal{L}_{G_2}\gamma_0 \doteq q[G_2^X \times \mathbf{B}] \cdot d\mathbf{X} - qG_2^X \cdot \mathbf{E} dt. \quad (2.61)$$

Next, we note that $\gamma_1 = \Gamma_1 + mc_\perp \cdot d\mathbf{X} - m\mathbf{W} \cdot \mathbf{c}_\perp dt$ and $\Gamma_1 = m\mathbf{W} \cdot d\mathbf{X} - [\frac{1}{2}mW^2 + \mu B] dt$, so that

$$-\frac{1}{2}\mathcal{L}_{G_1}(\gamma_1 + \Gamma_1) \doteq -i_{G_1} d(m\mathbf{W} \cdot d\mathbf{X}) - \frac{1}{2}i_{G_1} d(mc_\perp \cdot d\mathbf{X}) \quad (2.62)$$

$$+ i_{G_1} d\left(\left[\frac{1}{2}mW^2 + \mu B + \frac{1}{2}m\mathbf{W} \cdot \mathbf{c}_\perp\right] dt\right). \quad (2.63)$$

To compute the first term we use that $\mathbf{W} = \mathbf{u} + w\mathbf{b}$ depends on \mathbf{X} , w and t . Formula (β) implies that

$$i_{G_1} d(\mathbf{W} \cdot d\mathbf{X}) = [-G_1^X \times (\nabla \times \mathbf{W}) + G_1^w \mathbf{b}] \cdot d\mathbf{X} - (\mathbf{b} \cdot G_1^X) dw - G_1^X \cdot \frac{\partial \mathbf{W}}{\partial t} dt. \quad (2.64)$$

The computation of the term $i_{G_1} d(\mathbf{c}_\perp \cdot d\mathbf{X})$ is similar to (2.64), but one needs to take into account that \mathbf{c}_\perp depends on $\mathbf{X}, \mu, \zeta, t$, instead of \mathbf{X}, w, t . We find from (β):

$$i_{G_1} d(\mathbf{c}_\perp \cdot d\mathbf{X}) = \left[-G_1^X \times (\nabla \times \mathbf{c}_\perp) + G_1^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + G_1^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot d\mathbf{X} - \frac{2\mu}{q} d\zeta + \frac{2\mu}{q} S dt, \quad (2.65)$$

The third term in (2.62) will only contribute to the dt -component $\Gamma_{2,t}$ and will be computed at the end of this section. Combining equations (2.60)-(2.65), we conclude that the dw , $d\mu$ and $d\zeta$ -components of Γ_2 are given by

$$\Gamma_{2,w} = m\mathbf{b} \cdot G_1^X + \frac{\partial S_2}{\partial w}, \quad \Gamma_{2,\mu} = \frac{\partial S_2}{\partial \mu}, \quad \Gamma_{2,\zeta} = \frac{m}{q}\mu + \frac{\partial S_2}{\partial \zeta}. \quad (2.66)$$

2.4. Derivation of guiding-centre transformation

To enforce gyro-symmetry of the transformed Lagrangian, we wish to ensure that $\Gamma_k = \langle \Gamma_k \rangle$ for all components $k = X, w, \mu, \zeta, t$. Furthermore, the gyrogauge function will be chosen so that its gyro-average vanishes, $\langle S_2 \rangle = 0$. Comparing with (2.66), we find that this is possible if $S_2 = 0$ and $\Gamma_{2,\mu} = 0$, $\Gamma_{2,\zeta} = \frac{m}{q}\mu$. We also choose $\mathbf{b} \cdot G_1^X = 0$, so that $\Gamma_{2,w} = 0$ and

$$G_1^X \equiv G_{1,\perp}^X = -\boldsymbol{\rho}_0, \quad (2.67)$$

where $G_{1,\perp}^X$ is given by (2.57). The more difficult task is to remove the gyro-angle dependency in $\Gamma_{2,X}$ and $\Gamma_{2,t}$. We will first consider $\Gamma_{2,X}$, which will determine $G_{2,\perp}^X$ and G_1^w . Then, we will consider $\Gamma_{2,t}$.

To analyse $\Gamma_{2,X}$, we first write

$$-G_1^X \times (\nabla \times \mathbf{W}) = -[(\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot (\nabla \times \mathbf{W})] \mathbf{b} + [\mathbf{b} \cdot (\nabla \times \mathbf{W})] (\boldsymbol{\rho}_0 \times \mathbf{b}). \quad (2.68)$$

It can be shown that

$$\begin{aligned} -G_1^X \times (\nabla \times \mathbf{c}_\perp) &= \frac{2\mu}{q} \mathbf{R} - \frac{\mu}{q} [\mathbf{b} \cdot (\nabla \times \mathbf{b})] \mathbf{b} \\ &\quad - \frac{2\mu}{q} (\mathbf{a}_1 : \nabla \mathbf{b}) \mathbf{b} - \mu (\boldsymbol{\rho}_0 \cdot \nabla \log B) \frac{\partial \mathbf{c}_\perp}{\partial \mu} + (\boldsymbol{\rho}_0 \cdot \mathbf{R}) \frac{\partial \mathbf{c}_\perp}{\partial \zeta}. \end{aligned} \quad (2.69)$$

In this expression, the first line is gyro-angle independent, the second line is oscillatory. We can now rewrite (2.64) and (2.65) as follows:

$$\begin{aligned} i_{G_1} d(\mathbf{W} \cdot d\mathbf{X}) &= [(\mathbf{b} \cdot (\nabla \times \mathbf{W})) (\boldsymbol{\rho}_0 \times \mathbf{b}) + (G_1^w - (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot (\nabla \times \mathbf{W})) \mathbf{b}] \cdot d\mathbf{X} \\ &\quad + \boldsymbol{\rho}_0 \cdot \frac{\partial \mathbf{W}}{\partial t} dt. \end{aligned} \quad (2.70)$$

and

$$\begin{aligned} i_{G_1} d(\mathbf{c}_\perp \cdot d\mathbf{X}) &= \frac{2\mu}{q} \mathbf{R} \cdot d\mathbf{X} - \frac{2\mu}{q} \left(\frac{1}{2} [\mathbf{b} \cdot (\nabla \times \mathbf{b})] + \mathbf{a}_1 : \nabla \mathbf{b} \right) \mathbf{b} \cdot d\mathbf{X} \\ &\quad + \left[g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot d\mathbf{X} - \frac{2\mu}{q} d\zeta + \frac{2\mu}{q} S dt. \end{aligned} \quad (2.71)$$

where, similar to [Bri95, Mad10], we introduce the following convenient notation

$$g^\mu \equiv G_1^\mu - \mu (\boldsymbol{\rho}_0 \cdot \nabla \log B), \quad (2.72)$$

$$g^\zeta \equiv G_1^\zeta + (\boldsymbol{\rho}_0 \cdot \mathbf{R}). \quad (2.73)$$

Similarly, it will be convenient to use

$$g^w \equiv G_1^w - \frac{\mu}{q} \mathbf{a}_1 : \nabla \mathbf{b} - (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot (\nabla \times \mathbf{W}). \quad (2.74)$$

Combining the expression for Γ_2 (2.60), (2.61), (2.62) with the above (2.70) and (2.71), we

obtain

$$\begin{aligned}\Gamma_{2,X} = & -\frac{m\mu}{q}\mathbf{R} - m\left[g^w - \frac{\mu}{2q}[\mathbf{b} \cdot (\nabla \times \mathbf{b})]\right]\mathbf{b} \\ & + \left[q\left(G_2^X - \frac{1}{\Omega}[\mathbf{b} \cdot (\nabla \times \mathbf{W})]\boldsymbol{\rho}_0\right) \times \mathbf{B} - \frac{m}{2}\left(g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta}\right)\right].\end{aligned}$$

To make $\Gamma_{2,X}$ gyro-angle independent, we require that $\widetilde{g^w} = 0$ and we require the second line to vanish. We remind the reader that for any physical quantity X , we denote by \tilde{X} the gyro-angle oscillatory part,

$$\tilde{X} \equiv X - \langle X \rangle, \quad \langle X \rangle \equiv \frac{1}{2\pi} \oint X d\zeta.$$

Requiring that $\widetilde{g^w} = 0$ and that the second line in the above expression for $\Gamma_{2,X}$ vanish, yields

$$\widetilde{G_1^w} = \frac{\mu}{q} \mathbf{a}_1 : \nabla \mathbf{b} + \boldsymbol{\rho}_0 \cdot (\mathbf{b} \times (\nabla \times \mathbf{W})) = \frac{\mu}{q} \mathbf{a}_1 : \nabla \mathbf{b} - w \boldsymbol{\rho}_0 \cdot \boldsymbol{\kappa} + \boldsymbol{\rho}_0 \times [\mathbf{b} \times (\nabla \times \mathbf{u})], \quad (2.75)$$

and

$$G_{2,\perp}^X = \frac{1}{\Omega}[\mathbf{b} \cdot (\nabla \times \mathbf{W})]\boldsymbol{\rho}_0 + \frac{1}{2\Omega}\mathbf{b} \times \left[g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta}\right]. \quad (2.76)$$

And we find

$$\Gamma_{2,X} = -\frac{m\mu}{q}\mathbf{R} - m\left[\langle g^w \rangle - \frac{\mu}{2q}[\mathbf{b} \cdot (\nabla \times \mathbf{b})]\right]\mathbf{b}. \quad (2.77)$$

In addition, we need to consider the dt -component $\Gamma_{2,t}$. To evaluate this component, we first consider the third term in (2.62). Using (a), we find from formula (a)

$$i_{G_1} d\left(\left[\frac{1}{2}mW^2 + \mu B\right] dt\right) = \left[\frac{1}{2}mG_1^X \cdot \nabla W^2 + mG_1^w(\mathbf{b} \cdot \mathbf{W}) + BG_1^\mu + \mu G_1^X \cdot \nabla B\right] dt.$$

From (2.72) and (2.67), we can identify $BG_1^\mu + \mu G_1^X \cdot \nabla B = Bg^\mu$. We can furthermore show that

$$\frac{m}{2}G_1^X \cdot \nabla W^2 = -m(\mathbf{b} \cdot \mathbf{W})[(\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot (\nabla \times \mathbf{W})] + \frac{m}{\Omega}(\mathbf{c}_\perp \cdot \mathbf{W})[\mathbf{b} \cdot (\nabla \times \mathbf{W})] - m\boldsymbol{\rho}_0 \cdot (\mathbf{W} \cdot \nabla \mathbf{W}).$$

Thus,

$$\begin{aligned}i_{G_1} d\left(\left[\frac{1}{2}mW^2 + \mu B\right] dt\right) = & \left[m\left(G_1^w - (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot (\nabla \times \mathbf{W})\right)(\mathbf{b} \cdot \mathbf{W}) + Bg^\mu\right. \\ & \left.- m\boldsymbol{\rho}_0 \cdot (\mathbf{W} \cdot \nabla \mathbf{W}) + \frac{m}{\Omega}(\mathbf{c}_\perp \cdot \mathbf{W})[\mathbf{b} \cdot (\nabla \times \mathbf{W})]\right] dt.\end{aligned} \quad (2.78)$$

2.4. Derivation of guiding-centre transformation

Finally, we need to evaluate $i_{G_1} d(\mathbf{W} \cdot \mathbf{c}_\perp dt)$. Using (α) once more, we find

$$i_{G_1} d(\mathbf{W} \cdot \mathbf{c}_\perp dt) = \left[G_1^X \cdot (\nabla \mathbf{W}) \cdot \mathbf{c}_\perp + G_1^w \frac{\partial \mathbf{W}}{\partial w} \cdot \overset{0}{\mathbf{c}_\perp} + G_1^X \cdot (\nabla \mathbf{c}_\perp) \cdot \mathbf{W} \right. \\ \left. + G_1^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} \cdot \mathbf{W} + G_1^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot \mathbf{W} \right] dt.$$

having taken into account that $\partial_w \mathbf{W} = \mathbf{b} \perp \mathbf{c}_\perp$ to cancel one term. So that

$$i_{G_1} d\left(\frac{1}{2} m \mathbf{W} \cdot \mathbf{c}_\perp dt\right) = \frac{m}{2} \left[G_1^X \cdot (\nabla \mathbf{W}) \cdot \mathbf{c}_\perp + G_1^X \cdot (\nabla \mathbf{c}_\perp) \cdot \mathbf{W} \right. \\ \left. + G_1^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} \cdot \mathbf{W} + G_1^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot \mathbf{W} \right] dt, \quad (2.79)$$

The first two terms can be simplified using

$$G_1^X \cdot (\nabla \mathbf{W}) \cdot \mathbf{c}_\perp = \frac{2\mu}{q} \left(\frac{1}{2} \mathbf{b} \cdot \nabla \times \mathbf{W} + \mathbf{a}_1 : \nabla \mathbf{W} \right) \\ = \frac{\mu}{q} [\mathbf{b} \cdot \nabla \times \mathbf{b}] (\mathbf{b} \cdot \mathbf{W}) + \frac{\mu}{q} [\mathbf{b} \cdot \nabla \times \mathbf{u}_\perp] + \frac{2\mu}{q} \mathbf{a}_1 : \nabla \mathbf{W},$$

and

$$G_1^X \cdot \nabla \mathbf{c}_\perp \cdot \mathbf{W} = -\mu (\boldsymbol{\rho}_0 \cdot \nabla \log B) \frac{\partial \mathbf{c}_\perp}{\partial \mu} \cdot \mathbf{W} + (\boldsymbol{\rho}_0 \cdot \mathbf{R}) \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot \mathbf{W} \\ - \left[\frac{\mu}{q} \mathbf{b} \cdot \nabla \times \mathbf{b} + \frac{2\mu}{q} \mathbf{a}_1 : \nabla \mathbf{b} \right] (\mathbf{b} \cdot \mathbf{W}). \quad (2.80)$$

We note a cancellation between the terms involving $[\mathbf{b} \cdot \nabla \times \mathbf{b}] (\mathbf{b} \cdot \mathbf{W})$. Recalling the definitions of g^μ, g^ζ, g^w (2.72)-(2.74), the last three terms in (2.80) can also be combined to obtain

$$i_{G_1} d\left(\frac{1}{2} m \mathbf{W} \cdot \mathbf{c}_\perp dt\right) = \left[\frac{m\mu}{2q} [\mathbf{b} \cdot \nabla \times \mathbf{u}_\perp] + \frac{m\mu}{q} \mathbf{a}_1 : \nabla \mathbf{W} - \frac{m\mu}{q} [\mathbf{a}_1 : \nabla \mathbf{b}] (\mathbf{b} \cdot \mathbf{W}) \right. \\ \left. + \frac{m}{2} \left(g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} \cdot \mathbf{W} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot \mathbf{W} \right) \right] dt, \quad (2.81)$$

Thus, combining (2.61), (2.70), (2.71), (2.78) and (2.81), the component $\Gamma_{2,t}$ is found to be given by

$$\Gamma_{2,t} = -q G_2^X \cdot \mathbf{E} + \frac{m}{2} \left[g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot \mathbf{W} + \frac{m}{\Omega} (\mathbf{c}_\perp \cdot \mathbf{W}) (\mathbf{b} \cdot \nabla \times \mathbf{W}) - \frac{m\mu}{q} S \\ + B g^\mu + \frac{m\mu}{2q} \mathbf{b} \cdot (\nabla \times \mathbf{u}_\perp) + m g^w (\mathbf{b} \cdot \mathbf{W}) + \frac{m\mu}{q} \mathbf{a}_1 : \nabla \mathbf{W} - m \boldsymbol{\rho}_0 \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} \right). \quad (2.82)$$

Chapter 2. Guiding-centre theory

Now we note that $\mathbf{E} = \mathbf{E}_\perp + O(\epsilon_B)$, and $\mathbf{E}_\perp = -\mathbf{W} \times \mathbf{B}$, so that

$$-qG_2^X \cdot \mathbf{E} = (-qG_2^X \times \mathbf{B}) \cdot \mathbf{W} + O(\epsilon_B). \quad (2.83)$$

We thus find from (2.76)

$$\begin{aligned} -qG_2^X \cdot \mathbf{E} + \frac{m}{2} \left[g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot \mathbf{W} &= \left[-q(G_2^X \times \mathbf{B}) + \frac{m}{2} g^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + \frac{m}{2} g^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot \mathbf{W} + O(\epsilon_B) \\ &= -m[\mathbf{b} \cdot (\nabla \times \mathbf{W})](\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot \mathbf{W} + O(\epsilon_B) \\ &= -\frac{m}{\Omega} [\mathbf{b} \cdot (\nabla \times \mathbf{W})](\mathbf{c}_\perp \cdot \mathbf{W}) + O(\epsilon_B). \end{aligned}$$

This term will evidently cancel the third term in (2.82). Recalling also that $g^w = \langle g^w \rangle$, we obtain

$$\begin{aligned} \Gamma_{2,t} &= -\frac{m\mu}{q} S + \frac{m\mu}{2q} \mathbf{b} \cdot (\nabla \times \mathbf{u}_\perp) + m\langle g^w \rangle (\mathbf{b} \cdot \mathbf{W}) \\ &\quad + Bg^\mu + \frac{m\mu}{q} \mathbf{a}_1 : \nabla \mathbf{W} - m\boldsymbol{\rho}_0 \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} \right). \end{aligned}$$

To eliminate the oscillatory contributions, we require

$$\widetilde{g^\mu} = -\frac{\mu}{\Omega} \mathbf{a}_1 : \nabla \mathbf{W} + \frac{m}{B} \boldsymbol{\rho}_0 \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} \right). \quad (2.84)$$

The component $\langle g^\mu \rangle$ will be determined at the next order. For now, we simply note that

$$\Gamma_{2,t} = -\frac{m\mu}{q} S + \frac{m\mu}{2q} \mathbf{b} \cdot (\nabla \times \mathbf{u}_\perp) + m\langle g^w \rangle (\mathbf{b} \cdot \mathbf{W}) + B\langle g^\mu \rangle. \quad (2.85)$$

3rd-order analysis

The general formula for the third-order transformed Lagrangian is given according to (2.52), by

$$\Gamma_3 = -\mathcal{L}_{G_3} \gamma_0 - \mathcal{L}_{G_2} \Gamma_1 + \frac{1}{3} \mathcal{L}_{G_1}^2 \left(\gamma_1 + \frac{1}{2} \Gamma_1 \right) + dS_3.$$

By (2.56), the third-order generating vector field will only contribute to the perpendicular component of $\Gamma_{3,X}$ and to $\Gamma_{3,t}$. The parallel component of $\Gamma_{3,X}$ will give an equation for G_2^w . Therefore, in the same way that the second-order Lagrangian allowed us to solve for $G_{2,\perp}^X$ and G_2^w , the choice of the spatial component $\Gamma_{3,X}$ will now determine G_2^w and $G_{3,\perp}^X$. Since we will neither require knowledge of G_2^w nor of $G_{3,\perp}^X$ in the following, we will not compute the spatial component $\Gamma_{3,X}$. Similarly, the dt -component could be used to determine (the gyro-angle oscillatory part of) G_2^μ . We will not require this component of G_2 , either. Thus, it will suffice to compute only the three components $\Gamma_{3,w}$, $\Gamma_{3,\mu}$ and $\Gamma_{3,\zeta}$ in order to obtain three equations for the unknowns $\langle g^\mu \rangle$, $G_{2,\parallel}^X$ and g^ζ .

Because we will not be interested in the dX , dt components, we introduce the following

notation $\stackrel{(v)}{=}$ for equality *up to terms of the form* $(\dots) dX + (\dots) dt$. From (2.64), we thus find

$$-\mathcal{L}_{G_2} \Gamma_1 \doteq -i_{G_2} d\Gamma_1 \stackrel{(v)}{=} m(\mathbf{b} \cdot G_2^X) dw. \quad (2.86)$$

Let us also point out that for any phase-space function G , we have

$$i_{G_1} d(G\mathbf{b} \cdot d\mathbf{X}) = (i_{G_1} dG)(\mathbf{b} \cdot d\mathbf{X}) - (\mathbf{b} \cdot G_1^X) dG \stackrel{(v)}{=} 0, \quad (2.87)$$

since $(\dots)(\mathbf{b} \cdot d\mathbf{X}) \stackrel{(v)}{=} 0$ by definition of $\stackrel{(v)}{=}$.

More involved calculations are necessary to evaluate

$$\frac{1}{3} \mathcal{L}_{G_1}^2 \left(\gamma_1 + \frac{1}{2} \Gamma_1 \right) = \frac{1}{2} \mathcal{L}_{G_1}^2 \Gamma_1 + \frac{1}{3} \mathcal{L}_{G_1}^2 (m\mathbf{c}_\perp \cdot d\mathbf{X} - m\mathbf{W} \cdot \mathbf{c}_\perp dt).$$

From (2.70) and our observation (2.87), we obtain

$$\mathcal{L}_{G_1}^2 \Gamma_1 \doteq i_{G_1} d(i_{G_1} d\Gamma_1) \stackrel{(v)}{=} i_{G_1} d\left(\frac{m}{\Omega}(\mathbf{b} \cdot (\nabla \times \mathbf{W})) \mathbf{c}_\perp \cdot d\mathbf{X}\right).$$

From formula (β) , and the fact that $\mathbf{c}_\perp \perp G_1^X$, it follows that

$$\frac{1}{2} \mathcal{L}_{G_1}^2 \Gamma_1 \stackrel{(v)}{=} -\frac{m}{2\Omega}(\mathbf{b} \cdot (\nabla \times \mathbf{W})) \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot G_1^X d\zeta = -\frac{m\mu}{q\Omega} [\mathbf{b} \cdot (\nabla \times \mathbf{W})] d\zeta. \quad (2.88)$$

For the second term, we observe that

$$\mathcal{L}_{G_1}^2 (m\mathbf{c}_\perp \cdot d\mathbf{X} - m\mathbf{W} \cdot \mathbf{c}_\perp dt) \stackrel{(v)}{=} i_{G_1} d(i_{G_1} d(m\mathbf{c}_\perp \cdot d\mathbf{X})).$$

Using (2.71), the last term can be written – again not writing out terms that will contribute to dt , only, or terms of the form $i_{G_1} d((\dots)\mathbf{b} \cdot \mathbf{X})$ which vanish due to (2.87) – as

$$i_{G_1} d\left(\frac{2m\mu}{q} \mathbf{R} \cdot d\mathbf{X} + \left[mg^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + mg^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta}\right] \cdot d\mathbf{X} - \frac{2m\mu}{q} d\zeta\right).$$

The four contributing terms evaluate as follows

$$\begin{aligned} i_{G_1} d\left(\frac{2m\mu}{q} \mathbf{R} \cdot d\mathbf{X}\right) &\stackrel{(v)}{=} \frac{2m}{q} (\boldsymbol{\rho}_0 \cdot \mathbf{R}) d\mu, \\ i_{G_1} d\left(mg^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} \cdot d\mathbf{X}\right) &\stackrel{(v)}{=} -\frac{m}{q} g^\mu d\zeta, \\ i_{G_1} d\left(mg^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot d\mathbf{X}\right) &\stackrel{(v)}{=} -\frac{\partial}{\partial \zeta} \left(\frac{2m\mu}{q} g^\zeta\right) d\zeta + \left[\frac{m}{q} g^\zeta - \frac{\partial}{\partial \mu} \left(\frac{2m\mu}{q} g^\zeta\right)\right] d\mu - \frac{\partial}{\partial w} \left(\frac{2m\mu}{q} g^\zeta\right) dw, \\ i_{G_1} d\left(-\frac{2m\mu}{q} d\zeta\right) &\stackrel{(v)}{=} -\frac{2m}{q} [g^\mu + \mu(\boldsymbol{\rho}_0 \cdot \nabla \log B)] d\zeta + \frac{2m}{q} [g^\zeta - (\boldsymbol{\rho}_0 \cdot \mathbf{R})] d\mu. \end{aligned}$$

Chapter 2. Guiding-centre theory

Taking the sum and denoting $s \equiv \frac{2m\mu}{3q} g^\zeta$, we find

$$\begin{aligned} \frac{1}{3} \mathcal{L}_{G_1}^2 (m\mathbf{c}_\perp \cdot d\mathbf{X} - m\mathbf{W} \cdot \mathbf{c}_\perp dt) &\stackrel{(v)}{=} \left[-\frac{m}{q} g^\mu - \frac{2m\mu}{3q} (\boldsymbol{\rho}_0 \cdot \nabla B) - \frac{\partial s}{\partial \zeta} \right] d\zeta \\ &+ \left[\frac{m}{q} g^\zeta - \frac{\partial s}{\partial \mu} \right] d\mu - \frac{\partial s}{\partial w} dw \end{aligned} \quad (2.89)$$

We note that s contributes to these components of Γ_3 in the form of a closed term ds , which can be absorbed in the gyrogauge function S_3 . From (2.86), (2.88) and (2.89), we conclude that

$$\Gamma_{3,w} = m(\mathbf{b} \cdot G_2^X) + \frac{\partial S_3}{\partial w}, \quad (2.90)$$

$$\Gamma_{3,\mu} = \frac{m}{q} g^\zeta + \frac{\partial S_3}{\partial \mu}, \quad (2.91)$$

$$\Gamma_{3,\zeta} = -\frac{m}{q} g^\mu - \frac{2m\mu}{3q} (\boldsymbol{\rho}_0 \cdot \nabla \log B) - \frac{m\mu}{q\Omega} [\mathbf{b} \cdot (\nabla \times \mathbf{W})] + \frac{\partial S_3}{\partial \zeta}. \quad (2.92)$$

For the component $\Gamma_{3,\zeta}$ to vanish, we require

$$\langle g^\mu \rangle = -\frac{\mu}{\Omega} [\mathbf{b} \cdot (\nabla \times \mathbf{W})], \quad (2.93)$$

and

$$\frac{\partial S_3}{\partial \zeta} = \frac{m}{q} \widetilde{g^\mu} + \frac{2m\mu}{3q} (\boldsymbol{\rho}_0 \cdot \nabla \log B).$$

Recalling (2.84), and observing that $\mathbf{a}_1 = \partial_\zeta \mathbf{a}_2$, $\boldsymbol{\rho}_0 = -\partial_\zeta (\boldsymbol{\rho}_0 \times \mathbf{b})$, we can integrate in ζ to find

$$S_3 = -\frac{m\mu}{q\Omega} \mathbf{a}_2 : \nabla \mathbf{W} + \frac{m}{\Omega} (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} - \frac{2\mu}{3m} \nabla B \right). \quad (2.94)$$

Note that the integration in ζ is uniquely determined, provided $\langle S_3 \rangle = 0$. Having found an explicit expression for S_3 , we can now compute g^ζ from (2.91) as

$$g^\zeta = -\frac{1}{\Omega} \mathbf{a}_2 : \nabla \mathbf{W} - \frac{q}{2\mu\Omega} (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot \left(\frac{\partial \mathbf{W}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{W} + \frac{2\mu}{3m} \nabla B \right). \quad (2.95)$$

Similarly, we find from (2.90)

$$\mathbf{b} \cdot G_2^X = \frac{\mu}{q\Omega} \mathbf{a}_2 : \nabla \mathbf{b} - \frac{1}{\Omega} (\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot \left(\frac{\partial \mathbf{b}}{\partial t} + \mathbf{W} \cdot \nabla \mathbf{b} + \mathbf{b} \cdot \nabla \mathbf{W} \right). \quad (2.96)$$

We have now completed our derivation of G_1 and G_2^X . Returning to (2.85), and substituting

$$B\langle g^\mu \rangle = -\frac{m\mu}{q} [\mathbf{b} \cdot (\nabla \times \mathbf{W})] = -\frac{m\mu}{q} [\mathbf{b} \cdot (\nabla \times \mathbf{u}_\perp)] - \frac{m\mu}{q} [\mathbf{b} \cdot (\nabla \times \mathbf{b})](\mathbf{b} \cdot \mathbf{W}),$$

we find

$$\Gamma_{2,t} = -\frac{m\mu}{q}S - \frac{m\mu}{2q}\mathbf{b} \cdot \nabla \times \mathbf{u}_\perp + m\left(\langle g^w \rangle - \frac{\mu}{q}\mathbf{b} \cdot \nabla \times \mathbf{b}\right)(\mathbf{b} \cdot \mathbf{W}). \quad (2.97)$$

To simplify $\Gamma_{2,t}$, we choose

$$\langle g^w \rangle = \frac{\mu}{q}\mathbf{b} \cdot (\nabla \times \mathbf{b}). \quad (2.98)$$

So that, finally, the missing components of Γ_2 are found to be given by

$$\Gamma_{2,X} = -\frac{m\mu}{q}\left(\mathbf{R} + \frac{1}{2}[\mathbf{b} \cdot (\nabla \times \mathbf{b})]\mathbf{b}\right), \quad (2.99)$$

$$\Gamma_{2,t} = -\frac{m\mu}{q}S - \frac{m\mu}{2q}[\mathbf{b} \cdot (\nabla \times \mathbf{u}_\perp)]. \quad (2.100)$$

We recall that in the above derivation, we have imposed that

$$\mathbf{u}_\perp = \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (2.101)$$

2.4.3 Collection of useful identities

With the notation as above, $G_1^X = -\boldsymbol{\rho}_0$.

Since $(\hat{\boldsymbol{\rho}}, \mathbf{b}, \hat{\perp})$ is right-handed:

$$\hat{\boldsymbol{\rho}} \times \mathbf{b} = \hat{\perp}, \quad \mathbf{b} \times \hat{\perp} = \hat{\boldsymbol{\rho}}, \quad \hat{\perp} \times \hat{\boldsymbol{\rho}} = \mathbf{b}, \quad (2.102)$$

Using that they are mutually orthogonal unit vectors, we find from the decomposition $\nabla \hat{\perp} = (\nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}}) \otimes \hat{\boldsymbol{\rho}} + (\nabla \hat{\perp} \cdot \mathbf{b}) \otimes \mathbf{b}$ that

$$\nabla \hat{\perp} = \mathbf{R} \otimes \hat{\boldsymbol{\rho}} - (\nabla \mathbf{b} \cdot \hat{\perp}) \otimes \mathbf{b}, \quad (2.103)$$

$$\nabla \hat{\boldsymbol{\rho}} = -\mathbf{R} \otimes \hat{\perp} - (\nabla \mathbf{b} \cdot \hat{\boldsymbol{\rho}}) \otimes \mathbf{b}. \quad (2.104)$$

where $\mathbf{R} = \nabla \mathbf{e}_1 \cdot \mathbf{e}_2 = \nabla \hat{\perp} \cdot \hat{\boldsymbol{\rho}}$ is the gyrogauge vector field. We also note that

$$\frac{\partial \hat{\boldsymbol{\rho}}}{\partial \zeta} = \hat{\perp}, \quad (2.105)$$

$$\frac{\partial \hat{\perp}}{\partial \zeta} = -\hat{\boldsymbol{\rho}}. \quad (2.106)$$

Chapter 2. Guiding-centre theory

Based on the fact that $\mathbf{c}_\perp = \sqrt{2\mu B/m} \hat{\mathbf{L}} = \Omega \rho_0 \hat{\mathbf{L}}$, we find

$$\frac{\partial \mathbf{c}_\perp}{\partial \zeta} = -\Omega \boldsymbol{\rho}_0 = \mathbf{c}_\perp \times \mathbf{b}, \quad (2.107)$$

$$\frac{\partial \mathbf{c}_\perp}{\partial \mu} = \frac{1}{2\mu} \mathbf{c}_\perp, \quad (2.108)$$

$$\nabla \mathbf{c}_\perp = \frac{1}{2} \nabla \log B \otimes \mathbf{c}_\perp + \mathbf{R} \otimes \Omega \boldsymbol{\rho}_0 - (\nabla \mathbf{b} \cdot \mathbf{c}_\perp) \otimes \mathbf{b} \quad (2.109)$$

$$= \mu \nabla \log B \otimes \frac{\partial \mathbf{c}_\perp}{\partial \mu} - \mathbf{R} \otimes \frac{\partial \mathbf{c}_\perp}{\partial \zeta} - (\nabla \mathbf{b} \cdot \mathbf{c}_\perp) \otimes \mathbf{b}. \quad (2.110)$$

Using the vector calculus identity $\mathbf{A} \times (\nabla \times \mathbf{B}) = \nabla \mathbf{B} \cdot \mathbf{A} - \mathbf{A} \cdot \nabla \mathbf{B}$, it follows that

$$-G_1^X \times (\nabla \times \mathbf{c}_\perp) = \boldsymbol{\rho}_0 \times (\nabla \times \mathbf{c}_\perp) \quad (2.111)$$

$$= (\nabla \mathbf{c}_\perp) \cdot \boldsymbol{\rho}_0 - \boldsymbol{\rho}_0 \cdot \nabla \mathbf{c}_\perp \quad (2.112)$$

$$= \Omega \rho_0^2 \mathbf{R} - \mu (\boldsymbol{\rho}_0 \cdot \nabla \log B) \frac{\partial \mathbf{c}_\perp}{\partial \mu} + (\boldsymbol{\rho}_0 \cdot \mathbf{R}) \frac{\partial \mathbf{c}_\perp}{\partial \zeta} + (\boldsymbol{\rho}_0 \cdot \nabla \mathbf{b} \cdot \mathbf{c}_\perp) \mathbf{b}. \quad (2.113)$$

Note that

$$\Omega \rho_0^2 = \frac{\mathbf{c}_\perp^2}{\Omega} = \frac{2\mu B}{m\Omega} = \frac{2\mu}{q}. \quad (2.114)$$

Furthermore, for any vector field \mathbf{A} , we have

$$\boldsymbol{\rho}_0 \cdot \nabla \mathbf{A} \cdot \mathbf{c}_\perp = \frac{2\mu}{q} \hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{A} \cdot \hat{\mathbf{L}} \quad (2.115)$$

$$= \frac{\mu}{q} [\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{A} \cdot \hat{\mathbf{L}} - \hat{\mathbf{L}} \cdot \nabla \mathbf{A} \cdot \hat{\boldsymbol{\rho}}] \quad (2.116)$$

$$+ \frac{\mu}{q} [\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{A} \cdot \hat{\mathbf{L}} + \hat{\mathbf{L}} \cdot \nabla \mathbf{A} \cdot \hat{\boldsymbol{\rho}}] \quad (2.117)$$

To simplify the last expression, we used the identity

$$\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{A} \cdot \hat{\mathbf{L}} - \hat{\mathbf{L}} \cdot \nabla \mathbf{A} \cdot \hat{\boldsymbol{\rho}} = \mathbf{e}_2 \cdot \nabla \mathbf{A} \cdot \mathbf{e}_1 - \mathbf{e}_1 \cdot \nabla \mathbf{A} \cdot \mathbf{e}_2 = \nabla_2 \mathbf{A}_1 - \nabla_1 \mathbf{A}_2,$$

and observe that $(\mathbf{e}_1, \mathbf{e}_2, \mathbf{b})$ is a right-handed orthonormal basis, so that

$$\nabla_2 \mathbf{A}_1 - \nabla_1 \mathbf{A}_2 = -\epsilon^{3ij} \nabla_i \mathbf{A}_j = -\mathbf{b} \cdot \nabla \times \mathbf{A}.$$

We find

$$\boldsymbol{\rho}_0 \cdot \nabla \mathbf{A} \cdot \mathbf{c}_\perp = -\frac{\mu}{q} [\mathbf{b} \cdot (\nabla \times \mathbf{A})] - \frac{2\mu}{q} \mathbf{a}_1 : \nabla \mathbf{A}, \quad (2.118)$$

where $\mathbf{a}_1 \equiv \frac{1}{2} (\hat{\boldsymbol{\rho}} \otimes \hat{\mathbf{L}} + \hat{\mathbf{L}} \otimes \hat{\boldsymbol{\rho}})$.

We can now write

$$-G_1^X \times (\nabla \times \mathbf{c}_\perp) = \frac{2\mu}{q} \mathbf{R} - \mu(\boldsymbol{\rho}_0 \cdot \nabla \log B) \frac{\partial \mathbf{c}_\perp}{\partial \mu} + (\boldsymbol{\rho}_0 \cdot \mathbf{R}) \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \quad (2.119)$$

$$- \frac{2\mu}{q} \left[\frac{1}{2} \mathbf{b} \cdot (\nabla \times \mathbf{b}) + \mathbf{a}_1 : \nabla \mathbf{A} \right]. \quad (2.120)$$

We can also decompose

$$\begin{aligned} -G_1^X \times (\nabla \times \mathbf{W}) &= \hat{\mathbf{L}} \cdot \{\boldsymbol{\rho}_0 \times (\nabla \times \mathbf{W})\} \hat{\mathbf{L}} + \mathbf{b} \cdot \{\boldsymbol{\rho}_0 \times (\nabla \times \mathbf{W})\} \mathbf{b} \\ &= (\hat{\mathbf{L}} \times \boldsymbol{\rho}_0) \cdot (\nabla \times \mathbf{W}) \hat{\mathbf{L}} + (\mathbf{b} \times \boldsymbol{\rho}_0) \cdot (\nabla \times \mathbf{W}) \mathbf{b} \end{aligned}$$

to find

$$-G_1^X \times (\nabla \times \mathbf{W}) = [\mathbf{b} \cdot (\nabla \times \mathbf{W})] (\boldsymbol{\rho}_0 \times \mathbf{b}) - [(\boldsymbol{\rho}_0 \times \mathbf{b}) \cdot (\nabla \times \mathbf{W})] \mathbf{b} \quad (2.121)$$

Another important equality will be the following:

$$G_1^X \cdot \frac{\partial \mathbf{c}_\perp}{\partial \zeta} = \Omega \rho_0^2 = \frac{2\mu}{q}. \quad (2.122)$$

It follows that

$$\begin{aligned} i_{G_1} d(\mathbf{c}_\perp \cdot d\mathbf{X}) &= \left[-G_1^X \times (\nabla \times \mathbf{c}_\perp) + G_1^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + G_1^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot d\mathbf{X} \\ &\quad - \overbrace{\frac{\partial \mathbf{c}_\perp}{\partial \mu} \cdot G_1^X d\mu}^0 - \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \cdot G_1^X d\zeta - \frac{\partial \mathbf{c}_\perp}{\partial t} \cdot G_1^X dt \\ &= \left[-G_1^X \times (\nabla \times \mathbf{c}_\perp) + G_1^\mu \frac{\partial \mathbf{c}_\perp}{\partial \mu} + G_1^\zeta \frac{\partial \mathbf{c}_\perp}{\partial \zeta} \right] \cdot d\mathbf{X} - \frac{2\mu}{q} d\zeta + \frac{2\mu}{q} S dt. \end{aligned}$$

2.5 Summary

In this chapter, a detailed discussion of the Lagrangian formulation of guiding-centre theory in a rotating plasma has been given. After a short review of the historical development of guiding-centre theory, the guiding-centre Lagrangian has been presented in a suitable set of guiding-centre coordinates. As pointed out in the introduction to this chapter, the approach to guiding-centre theory followed in the present work focuses on (gyro-)symmetry as the core mechanism by which a reduction in complexity can be achieved. This is in contrast to an alternative approach based on gyro-averaging of the equations. While the two approaches can be used to derive guiding-centre equations of motion which are formally equivalent, up to higher-order correction terms in ϵ_B , the main advantage of the Lagrangian approach is the additional preservation of the underlying physical structure encoded in the Lagrangian formulation. In particular, using the present Lagrangian approach, if quantities such as the energy and toroidal momentum are conserved by the particle motion, then the corresponding

guiding-centre motion will also ensure exact conservation of these quantities. Such an exact conservation of physical constants of motion in the reduced description is especially important for numerical simulations over long time-scales, and is difficult to achieve when the equations of motion are derived by gyroaveraging. A full derivation of the guiding-centre Lagrangian including centrifugal effects in a time-varying background is given, following an order-by-order expansion in Larmor radius ϵ_B . Throughout, a particular emphasis is given on the meaning of the guiding-centre coordinates and their relation to the physical particle coordinates. The guiding-centre equations and the higher-order Larmor radius corrections derived in the present chapter form the basis of the original results presented in chapters 3 and 4. Chapter 3 focuses on the possible implications of higher-order ϵ_B corrections which are conventionally neglected in the context of slowing-down simulation of neutral beam injection in a MAST-like equilibrium, in the absence of strong background rotation ($\mathbf{u} \approx 0$). It will be seen that the neglect of higher-order corrections for the magnetic moment μ and the parallel guiding-centre velocity w corresponds to a neglect of a guiding-centre drift parallel to the field lines, the so-called Baños drift. This often overlooked drift motion along the field lines, induced by magnetic field line shear, is shown to affect the expected resonances of particles with external magnetic perturbations, as well as the estimated NBI driven current in a MAST-like equilibrium by up to 8% [LPGC17]. In chapter 4, higher-order corrections, such as the higher-order displacement ρ_1 (2.36) (which expresses a deviation of the particles Larmor motion from a circular motion due to gradients in the background fields), will be shown to give rise to off-diagonal components of the pressure tensor corresponding to a given distribution of guiding-centres. Several errors in the available expressions for the required higher-order corrections in the literature [Mad10, appendix] have been uncovered in the process of the detailed derivation given in the present chapter. While these errors do not affect the results in the main text of [Mad10], they prove crucial for the correct derivation of the gyroviscous pressure components in chapter 4. Based on the results of the present chapter, first-order (gyroviscous) corrections to the pressure tensor will be evaluated from guiding-centre theory without any simplifying assumptions on the background geometry, for the first time. This derivation will form one of the main ingredients to show the consistency of the kinetic-MHD model presented in chapter 4 [LGPC19].

3 Application of higher-order guiding-centre corrections to full- f calculation

3.1 Introduction

Much effort has recently been expended on calculating realistic, consistent fast ion distributions in present and future devices [AGB⁺15, SKSS10, PMA⁺04, PMCG15, BGC03]. Some research focuses on special ion orbits that may hit sensitive plasma-facing components [SKSS10], other efforts attempt to estimate driven currents and pressure gradients associated with fast ions, which ultimately affect plasma equilibria [BGC03, PMA⁺04]. Most numerical studies involving fast particles apply low-order guiding-centre equations to track their orbits and evolve their distribution function. There exist numerous papers on the derivation and theoretical discussion of higher-order guiding-centre equations (see e.g. [Lit83, Bri95, CB09, BSQ13] and references therein). But in the context of guiding-centre simulations, higher-order effects seem to be largely ignored.

The consideration of guiding-centre equations at one order higher in Larmor radius than is conventionally done is motivated by the findings of previous work [PGC15]. There, it was shown from a simple example that the Baños drift, a drift parallel to the field lines induced by the magnetic field line twist $\tau \equiv \mathbf{b} \cdot \nabla \times \mathbf{b} = \hat{\mu}_0 j_{||}/B$ (where $\hat{\mu}_0$ is the vacuum permeability) is not correctly accounted for using the conventional guiding-centre expressions. To correctly account for the Baños drift, we derive a set of higher-order guiding-centre equations, similar to Littlejohn [Lit83] but following the formulation of Brizard [Bri95] and Cary and Brizard [CB09]. We then show that these higher-order guiding-centre equations of motion can actually be used in orbit-following codes instead of solving the more costly full-orbit equations. The code VENUS-LEVIS [PCGM14] has indeed been upgraded in the publication [LPGC17] (which forms the basis of the present chapter) to include these higher-order guiding-centre equations and is applied to revisit three cases involving energetic ions: 1) neutral beam injection in a MAST-like low aspect-ratio spherical equilibrium where the fast ion driven current is significantly larger with respect to previous calculations [PGC⁺14] because of the large Baños drift due to the combination of large equilibrium currents, low magnetic field strength and comparatively high injection energies, 2) fast ion losses due to resonant magnetic perturbations in a MAST-like

equilibrium where a lower lost fraction and a better confinement than previous calculations [PMCG15] is confirmed, 3) alpha particles in the ripple field of the European DEMO [PCFG16, LPGC17] where the effect is found to be marginal.

Fast ion distributions obtained via neutral beam heating are naturally anisotropic in pressure, and exhibit strong gradients in velocity parallel to the field lines. Higher-order parallel dynamics, such as the parallel Baños drift, dismissed by Hazeltine and Meiss [HM03, p.136, ch.4.2] as having no known application, can become important for studies in the exotic properties and consequences of auxiliary generated fast ion populations. Denoting E the particle energy, I_p the plasma current and B_0 the magnetic field strength, the Baños drift scales as $\sim \sqrt{E}I_p/B_0^2$. The higher-order corrections considered are expected to be important at high particle energies, large plasma current and/or small magnetic field strength. To our knowledge, the implications of higher-order corrections for guiding-centre following codes and the modelling of energetic ion source distribution functions has not been discussed before.

Higher-order corrections to the guiding-centre's magnetic moment and toroidal momentum have previously been considered to compute consistent hybrid kinetic-MHD equilibria by Belova et al. [BGC03] with neutral beam injection (NBI). The focus of their work was to find appropriate variables to be used in the modelled energetic ion equilibrium distribution function [BGC03, equations (15)–(17)] in order to include kinetic effects in the computed MHD equilibrium. Their fast ion distribution functions were however not computed based on guiding-centre slowing-down simulations that would be consistent with such higher-order corrections in the equations of motion. Instead, their results are obtained with an ad hoc analytic expression resembling a slowing-down distribution, expressed as a function of the corrected (adiabatic) invariants. Their ad hoc distribution contains free parameters that are tuned to match typical NSTX parameters as well as the profiles of the beam ion density calculated by the TRANSP code. To compare the conservation properties of guiding-centre invariants expanded at various orders in Larmor radius, Belova et al. [BGC03] employed full-orbit following with the HYM code (see e.g. [BJJ⁺00]). From their work, they concluded that good conservation of the guiding-centre invariants could only be achieved via the inclusion of higher-order terms [BGC03, sec. III.A, III.B]. A second important question addressed by Belova et al. [BGC03] was to correctly evaluate the fast ion contribution to the current density, when the spatial gradient length scale of the slowing-down distribution is small compared to the Larmor radius [BGC03, sec. IV]. They found that strong gradients in the final slowing-down NBI distribution impacted the fast ion driven current (evaluated after the NBI slowing-down distribution has been computed). In the present work, the corrections provided in the context of NBI modelling are complementary to the work of Belova et al. in the sense that the focus here is on higher-order corrections related to strong *velocity* gradients from NBI deposition and on the question of correctly initialising the position and velocity variables of guiding-centre tracers. From a computational point of view, the present work therefore applies the results of chapter 2 to address how to initialize guiding-centre distributions for slowing-down simulations and what equations to solve in order to obtain consistent saturated distribution functions, while the work by Belova et al. [BGC03] discusses the properties of the saturated

solution in terms of the corrected invariants.

The work is organised as follows. In section 3.2, we summarise some elements of modern guiding-centre theory and explain our choice of guiding-centre variables as implemented in the code. To this end, we will recall several results from chapter 2, in the appropriate limit. It will be argued that the form of the Lagrangian derived in chapter 2 (see also Brizard [Bri95]) may have practical advantages over the original (equivalent) expression by Littlejohn [Lit83]. In section 3.3, the higher-order corrections are investigated numerically. In section 3.4, higher-order guiding-centre equations are applied in slowing-down simulations of NBI in a MAST-like equilibrium and α -particles in DEMO. Higher-order corrections are found to lead to an increase of the fast ion driven current by up to 8% for NBI in the MAST-like equilibrium [LPGC17]. When applying resonant magnetic perturbations (RMP), these corrections are seen to affect the lost fraction of NBI ions. Better confinement is found for co-current injection in a MAST-like equilibrium as compared to previous calculations. The impact on fusion α -particle losses in DEMO is found to be marginal. Section 3.5 is dedicated to a discussion on the surprising fact that the Baños drift finds an application in current drive corrections for parallel anisotropic distributions. Conclusive remarks are reserved for recitation in section 3.6.

3.2 Guiding-centre theory

3.2.1 Introduction

The Lagrangian formulation of guiding-centre theory as outlined in chapter 2 (see [Lit83, Bri95, CB09]) is based on the removal of the gyroangle dependence of the particle phase-space Lagrangian

$$\mathcal{L}_{\text{phys}} = (\epsilon_B^{-1} q \mathbf{A} + m \mathbf{v}) \cdot \dot{\mathbf{x}} - \frac{1}{2} m v^2, \quad (3.1)$$

by the use of a near-identity phase-space transformation. We shall assume a time-independent background equilibrium without electric field throughout this current chapter. The more general formulation, considered in chapter 2, with strong flows will be used in chapter 4, below. We recall that following the Lie perturbation procedure explained in chapter 2, one introduces the expansion parameter $\epsilon_B \sim \rho/L_B \ll 1$, with the Larmor radius $\rho \sim mv/qB$, and characteristic scale-length $L_B \sim B|\nabla \mathbf{B}|^{-1}$. Employing the right-handed orthonormal frame $\mathbf{e}_1, \mathbf{e}_2, \mathbf{b}$ with $\mathbf{b} = \mathbf{B}/B$, the gyroangle ζ_0 (in the absence of background flow $\mathbf{u} \equiv 0$) is defined through (cp. (2.25))

$$\mathbf{v} = v_{||} \mathbf{b}(\mathbf{x}) - v_{\perp} (\sin(\zeta_0) \mathbf{e}_1(\mathbf{x}) + \cos(\zeta_0) \mathbf{e}_2(\mathbf{x})). \quad (3.2)$$

In the context of this perturbative treatment, it is convenient to introduce physical particle coordinates $\mathcal{Z}_0 = (\mathbf{x}, v_{||}, \mu_0, \zeta_0)$, with μ_0 the lowest-order expression for the magnetic moment $\mu_0 \equiv \frac{1}{2} m v_{\perp}^2 / B(\mathbf{x})$. The Lie transformation method provides a framework for the asymptotic

removal of the gyroangle-dependence in (3.1), moving gyroangle-dependent terms to higher order in ϵ_B . As explained in Section 2.4.1, the method results in a phase-space transformation

$$\mathcal{T}_\epsilon : (\mathbf{x}, v_{||}, \mu_0, \zeta_0) \rightarrow (\mathbf{X}, U, \mu, \zeta) \quad (3.3)$$

from physical particle variables to guiding-centre variables $\mathcal{Z}_{\text{gc}} = (\mathbf{X}, U, \mu, \zeta)$. The Lagrangian $\mathcal{L}_{\text{phys}}$ expressed in these guiding-centre variables, and neglecting terms of order $O(\epsilon_B^2)$ and higher, is given by (cp. (2.26) in the limit $\mathbf{u} = 0$)

$$\mathcal{L} = q\mathbf{A}^* \cdot \dot{\mathbf{X}} + \epsilon_B m\mu/q\dot{\zeta} - H, \quad (3.4)$$

where the symplectic part of the Lagrangian \mathbf{A}^* is given by

$$\mathbf{A}^* \equiv \epsilon_B^{-1} \mathbf{A}(\mathbf{X}) + m/qU\mathbf{b}(\mathbf{X}) - \epsilon_B m\mu/q^2 \bar{\mathbf{R}}(\mathbf{X}) + O(\epsilon_B^2), \quad (3.5)$$

and the Hamiltonian is

$$H = \frac{1}{2}mU^2 + \mu B(\mathbf{X}) + O(\epsilon_B^2). \quad (3.6)$$

Note that in contrast to chapter 2, where the parallel guiding-centre velocity was denoted by w measured in a frame relative to the background velocity \mathbf{u} , we denote the parallel guiding-centre velocity in this chapter by U , thinking of it as the guiding-centre velocity in a Eulerian frame unrelated to possible macroscopic background flows. The vector field $\bar{\mathbf{R}} = \mathbf{R} + \frac{1}{2}(\mathbf{b} \cdot \nabla \times \mathbf{b})\mathbf{b}$ is expressed in terms of the gyrogauge vector field $\mathbf{R} \equiv (\nabla \mathbf{e}_1) \cdot \mathbf{e}_2$ (2.29) (first introduced by Littlejohn [Lit83, Lit88]) and the magnetic field line twist $\tau \equiv (\mathbf{b} \cdot \nabla \times \mathbf{b})$ [CB09]. We emphasize that with the present choice of guiding-centre coordinates, terms at order ϵ_B in the Hamiltonian vanish. A detailed discussion of different choices of guiding-centre coordinates as well as an algorithm for their calculation to arbitrary order can be found in [BSQ13]. It turns out that the formulation considered in the present work is particularly advantageous for numerical simulation, as explained in the next section.

3.2.2 Equations of motion

The equations of motion have been derived from the Euler-Lagrange equations $\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\mathcal{Z}}^\alpha} \right) = \frac{\partial \mathcal{L}}{\partial \mathcal{Z}^\alpha}$ in chapter 2, eq. (2.31). As shown in section 2.3.2, assuming time-invariant background fields, they are conveniently expressed in terms of $\mathbf{B}^* \equiv \nabla \times \mathbf{A}^*$ and $\mathbf{E}^* \equiv -\nabla H/q$. Setting $\mathcal{Z}^\alpha = X$ and $\mathcal{Z}^\alpha = U$, one obtains

$$\dot{\mathbf{X}} = U \frac{\mathbf{B}^*}{B_{||}^*} + \frac{\mathbf{E}^* \times \mathbf{b}}{B_{||}^*}, \quad \dot{U} = \frac{e}{m} \frac{\mathbf{B}^* \cdot \mathbf{E}^*}{B_{||}^*}, \quad (3.7)$$

with $B_{||}^* \equiv \mathbf{b} \cdot \mathbf{B}^*$. The Euler-Lagrange equations corresponding to $\mathcal{Z}^\alpha = \zeta$ yields $\dot{\mu} = 0$. For $\mathcal{Z}^\alpha = \mu$, we obtain $\dot{\zeta} = \Omega + \mathbf{R}(\mathbf{X}) \cdot \dot{\mathbf{X}}$, where $\Omega \equiv qB(\mathbf{X})/m$ is the gyrofrequency evaluated at the guiding-centre position. The lower-order correction term $\mathbf{R}(\mathbf{X}) \cdot \dot{\mathbf{X}}$ accounts for any rotation

along the guiding-centre trajectory of the frame $\mathbf{e}_1, \mathbf{e}_2, \mathbf{b}$ which is used to define ζ . It guarantees that the resulting guiding-centre dynamics is independent of the particular choice of $\mathbf{e}_1, \mathbf{e}_2$, a property called gyrogauge-invariance [Lit88, Lit84].

3.2.3 Choice of guiding-centre variables

The form of the higher-order guiding-centre Lagrangian is not unique and depends on the choice of guiding-centre variables \mathcal{Z}_{gc} . The choice used to arrive at (3.4) differs from the one made in the original work by Littlejohn [Lit83]. In particular, the different choices between chapter 2 (cp. also [CB09]) and [Lit83] can be tracked down to equation (2.98). Littlejohn's Lagrangian is obtained by setting the gyroaverage $\langle g_1^w \rangle = (\mu/2q)\mathbf{b} \cdot \nabla \times \mathbf{b}$, whereas the form considered in the present work corresponds to the choice $\langle g_1^w \rangle = (\mu/q)\mathbf{b} \cdot \nabla \times \mathbf{b}$. Therefore, the higher-order guiding-centre Lagrangian should not be viewed in isolation of the associated guiding-centre phase-space transformation.

The relevant relation between full-Lorentzian and guiding-centre variables for the present work, neglecting terms of order $O(\epsilon_B^2)$ and higher, is expressed in terms of the first-order generating vector field G_1 (2.37)-(2.40) (see also [CB09, eq. (5.41)]). Formally, the relation is given by $\mathcal{Z}_{\text{gc}}^\alpha = \mathcal{Z}_0^\alpha + \epsilon_B G_1^\alpha$, which yields (see (2.33) and (2.37)-(2.40))

$$X = \mathbf{x} - \boldsymbol{\rho}_0, \quad (3.8)$$

$$U = v_{||} + \mu_0/q(\mathbf{a}_1 : \nabla \mathbf{b} + \tau) - v_{||}\boldsymbol{\rho}_0 \cdot \boldsymbol{\kappa}, \quad (3.9)$$

$$\begin{aligned} \mu &= \mu_0 + \boldsymbol{\rho}_0 \cdot \left(\mu_0 \nabla \log B + \frac{mv_{||}^2}{B} \boldsymbol{\kappa} \right) \\ &\quad - \mu_0 \frac{v_{||}}{\Omega} (\mathbf{a}_1 : \nabla \mathbf{b} + \tau), \end{aligned} \quad (3.10)$$

$$\begin{aligned} \zeta &= \zeta_0 - \boldsymbol{\rho}_0 \cdot \mathbf{R} + \frac{\partial \boldsymbol{\rho}_0}{\partial \zeta} \cdot \left(\nabla \log B + \frac{mv_{||}^2}{2\mu_0 B} \right) \\ &\quad + \frac{v_{||}}{\Omega} \mathbf{a}_2 : \nabla \mathbf{b}. \end{aligned} \quad (3.11)$$

These expressions involve dyadic tensors $\mathbf{a}_1, \mathbf{a}_2$, which are defined (cp. (2.23) in chapter 2) in terms of the rotating (right-handed) orthonormal frame $(\mathbf{b}, \hat{\mathbf{l}}, \hat{\boldsymbol{\rho}})$, given by (cp. (2.22))

$$\hat{\mathbf{l}} = -\sin(\zeta)\mathbf{e}_1 - \cos(\zeta)\mathbf{e}_2, \quad (3.12)$$

$$\hat{\boldsymbol{\rho}} = \cos(\zeta)\mathbf{e}_1 - \sin(\zeta)\mathbf{e}_2, \quad (3.13)$$

and $\boldsymbol{\rho}_0 \equiv \rho_0 \hat{\boldsymbol{\rho}} \equiv -\mathbf{b} \times m\mathbf{v}/qB$ is the usual expression for the Larmor vector. We remind the reader that the dyadic tensors (cp. (2.23) in chapter 2) are given by

$$\mathbf{a}_1 = -\frac{1}{2}(\hat{\boldsymbol{\rho}}\hat{\mathbf{l}} + \hat{\mathbf{l}}\hat{\boldsymbol{\rho}}), \quad \mathbf{a}_2 = \frac{1}{4}(\hat{\mathbf{l}}\hat{\mathbf{l}} - \hat{\boldsymbol{\rho}}\hat{\boldsymbol{\rho}}).$$

The terms also involve the field line curvature $\kappa \equiv \mathbf{b} \cdot \nabla \mathbf{b}$, the field strength gradient ∇B and the magnetic field line twist $\tau \equiv \mathbf{b} \cdot \nabla \times \mathbf{b}$.

The appearance of terms involving

$$\mathbf{a}_1 : \nabla \mathbf{b} \equiv -\frac{1}{2} (\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{b} \cdot \hat{\mathbf{L}} + \hat{\mathbf{L}} \cdot \nabla \mathbf{b} \cdot \hat{\boldsymbol{\rho}}) \quad (3.14)$$

may be more intuitively understood when considering the projection of the perpendicular particle velocity component \mathbf{v}_\perp onto $\mathbf{b}(X)$, which is one of the contributions to the parallel guiding-centre velocity

$$U = \mathbf{b}(X) \cdot \dot{X} = \mathbf{b}(X) \cdot \mathbf{v} - \mathbf{b}(X) \cdot \dot{\boldsymbol{\rho}}. \quad (3.15)$$

We obtain

$$\begin{aligned} \mathbf{b}(X) \cdot \mathbf{v}_\perp &= v_\perp \mathbf{b}(X) \cdot \hat{\mathbf{L}}(X + \boldsymbol{\rho}_0) \\ &\approx v_\perp \mathbf{b}(X) \cdot (\hat{\mathbf{L}}(X) + (\boldsymbol{\rho}_0 \cdot \nabla) \hat{\mathbf{L}}|_X) \\ &= (mv_\perp^2/qB) \mathbf{b} \cdot (\hat{\boldsymbol{\rho}} \cdot \nabla \hat{\mathbf{L}}) \\ &\approx -2\mu/q (\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{b} \cdot \hat{\mathbf{L}}), \end{aligned} \quad (3.16)$$

and we observe that

$$\begin{aligned} -\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{b} \cdot \hat{\mathbf{L}} &= -\frac{1}{2} (\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{b} \cdot \hat{\mathbf{L}} + \hat{\mathbf{L}} \cdot \nabla \mathbf{b} \cdot \hat{\boldsymbol{\rho}}) \\ &\quad -\frac{1}{2} (\hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{b} \cdot \hat{\mathbf{L}} - \hat{\mathbf{L}} \cdot \nabla \mathbf{b} \cdot \hat{\boldsymbol{\rho}}) \\ &= \mathbf{a}_1 : \nabla \mathbf{b} - \frac{1}{2} \tau. \end{aligned} \quad (3.17)$$

Calculation of $\mathbf{b}(X) \cdot \dot{\boldsymbol{\rho}}$ to the required order would involve a higher-order correction to $\boldsymbol{\rho} \equiv \mathbf{x} - X = \boldsymbol{\rho}_0 + \epsilon_B \boldsymbol{\rho}_1 + \dots$, see [Bri95, appendix] for details. According to eq. (3.17), the appearance of \mathbf{a}_1 in (3.9) and (3.10) is a direct consequence of the distinction $\mathbf{b}(X) \neq \mathbf{b}(\mathbf{x})$, a first-order finite Larmor radius effect. More explicit expressions for these corrections in a simple background field may be found in appendix A.

In a conventional guiding-centre treatment as used in most guiding-centre following codes, the additional first order correction terms in equations (3.4), (3.8), (3.9), (3.10), (3.11) are neglected, relying instead on the leading order relations

$$X = \mathbf{x} - \boldsymbol{\rho}_0, \quad (3.18)$$

$$U = v_{||} \equiv \mathbf{v} \cdot \mathbf{b}(\mathbf{x}), \quad (3.19)$$

$$\mu = \mu_0 \equiv \frac{mv_\perp^2}{2B(\mathbf{x})}. \quad (3.20)$$

And the conventional guiding-centre Lagrangian

$$\mathcal{L}_0 = (\epsilon_B^{-1} q \mathbf{A} + m \mathbf{U} \mathbf{b}) \cdot \dot{\mathbf{X}} - H_0 + O(\epsilon_B), \quad (3.21)$$

$$H_0 = \frac{1}{2} m U^2 + \mu B(\mathbf{X}) + O(\epsilon_B). \quad (3.22)$$

is considered on reduced phase-space (\mathbf{X}, U, μ) , with $\mu = \text{const.}$ treated as a parameter. The conventional Lagrangian (3.21) is correct to order $O(1)$. Additional $O(\epsilon_B)$ terms in the Lagrangian are then neglected.

3.2.4 Guiding-centre push-forward

To consistently include all first-order effects in numerical guiding-centre following, guiding-centres have to be initialized correctly according to a given physical distribution function f . This requires the relation between full-Lorentzian distributions and guiding-centre distributions to be established including terms at order $O(\epsilon_B)$. Assuming $f = f(\mathbf{x}, v_{||}, \mu)$ with $\rho_0 |\nabla \log f| \sim \epsilon_B$, expanding to first order in gradient length scale, the guiding-centre distribution F_{gc} is found to be given by

$$F_{\text{gc}} = f - \epsilon_B G_1^X \cdot \nabla f - \epsilon_f G_1^{v_{||}} \frac{\partial f}{\partial v_{||}} - \epsilon_f G_1^\mu \frac{\partial f}{\partial \mu} + O(\epsilon_B^2, \epsilon_f^2, \epsilon_B \epsilon_f). \quad (3.23)$$

Where an additional expansion parameter $\epsilon_f \sim \frac{\mu \tau}{q} \left| \frac{\partial \log f}{\partial v_{||}} \right|$ has been introduced. While ϵ_B is related to spatial gradients of the background equilibrium, ϵ_f measures the effect of velocity space gradients of the particle distribution. The parameter ϵ_f depends both on the equilibrium (via the magnetic twist τ) as well as the anisotropy of the distribution function. For a fusion α -particle birth distribution considered in section 3.4.5, we would e.g. have $\epsilon_f \ll \epsilon_B$, while for the NBI case considered in section 3.4.1, it is found that $\epsilon_f > \epsilon_B$.

For a correct initialization of guiding-centre markers, the first order phase-space density $d^3 \mathbf{x} d^3 \mathbf{v}$ is also required in guiding-centre coordinates. As explained in [CB09, p.720], the Jacobian of the phase-space transformation $(\mathbf{x}, \mathbf{v}) \rightarrow (\mathbf{X}, U, \mu, \zeta)$ (3.8)-(3.11) is given by $\mathcal{J} = m^{-1} B_{||}^* \equiv \mathbf{b} \cdot \mathbf{B}^*$ in terms of $B_{||}^* \equiv \mathbf{b} \cdot \mathbf{B}^*$. Thus

$$d^3 \mathbf{x} d^3 \mathbf{v} = m^{-1} B d^3 \mathbf{x} d\nu_{||} d\mu_0 d\zeta_0 \quad (3.24)$$

$$= m^{-1} B_{||}^* d^3 \mathbf{X} dU d\mu d\zeta. \quad (3.25)$$

One can arrive at the same result by consideration of the canonical phase-space volume element as computed from the Lagrangian (3.4) [CB09, equation (3.43)].

Neglecting first-order corrections associated to the magnetic field line twist, we find $B_{||}^* = B$ and no distinction needs to be made between the particle and guiding-centre volume elements. Figure 3.1 shows a comparison between approximations to $B_{||}^*$ at different orders in ϵ_B for a

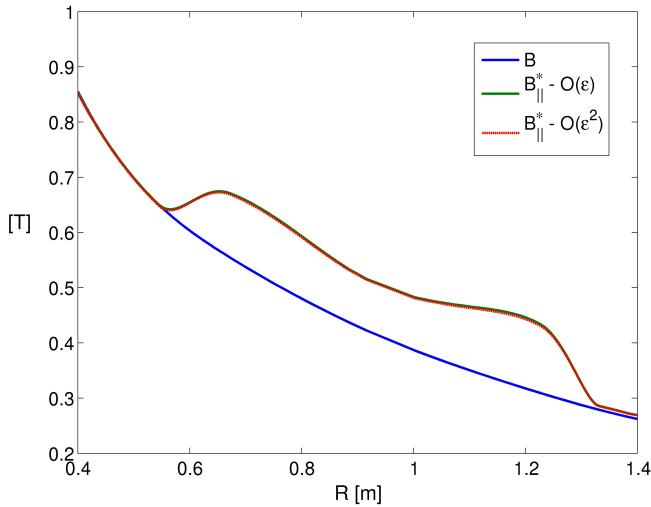


Figure 3.1 – $B_{||}^*$ for a H^+ ion at energy $E = 10\text{keV}$, parallel pitch $\lambda = 0.5$. Shown are approximations expanded at different orders in ϵ_B for the axisymmetric MAST-like equilibrium considered in section 3.4.1, along the midplane ($Z = 0$).

H^+ hydrogen ion guiding-centre with $E = 10\text{keV}$ and parallel pitch $\lambda = 0.5$ for the MAST-like equilibrium of section 3.4.1. The first order term associated with the magnetic field line twist τ is found to make a noticeable difference, while the second-order term related to $\bar{\mathbf{R}}$ makes only a negligible (< 1% throughout the domain) contribution to $B_{||}^*$ in this case.

3.2.5 Discussion

As pointed out in [CB09, eq. (3.48) and discussion], $O(\epsilon_B)$ corrections to the Hamiltonian H lead to $O(\epsilon_B)$ drifts, while $O(\epsilon_B)$ corrections to the symplectic part A^* (3.5) only lead to additional $O(\epsilon_B^2)$ drifts in the equations of motion. Owing to this observation, the form of the higher-order Hamiltonian H in equation (3.6) is preferable for a numerical treatment over the choice considered in [Lit83]. Indeed, the observation implies that guiding-centre simulations including all drift corrections at $O(\epsilon_B)$ (perpendicular *and* parallel to the field lines) can actually be achieved based on the expression for the conventional guiding-centre Lagrangian (i.e. neglecting the $\bar{\mathbf{R}}$ -term in A^*). Any drift terms that are neglected are then seen to be of order $O(\epsilon_B^2)$ or higher. To consistently include the $O(\epsilon_B)$ drift-correction parallel to the field lines, higher-order terms in the correspondence between guiding-centre and physical particle variables must however be retained.

To provide a concrete example where the above observation applies, we consider the problem of evaluating the NBI driven current. As discussed e.g. in [Spi52, QTRL00], there are two

contributions to the physical current density $\mathbf{j} = \mathbf{j}_{\text{gc}} + \mathbf{j}_{\text{mag}} + O(\epsilon_B^2)$, where

$$\mathbf{j}_{\text{gc}} = q \int \dot{\mathbf{X}} F_{\text{gc}} d^3 v, \quad (3.26)$$

$$\mathbf{j}_{\text{mag}} = \epsilon_B \nabla \times \left(\int -\mu \mathbf{b} F_{\text{gc}} d^3 v \right). \quad (3.27)$$

In these expressions, F_{gc} denotes the guiding-centre distribution. To obtain \mathbf{j}_{gc} including all corrections at $O(\epsilon_B)$, we need to take into account guiding-centre drifts at $O(\epsilon_B)$ perpendicular *and* parallel to the field lines. Perpendicular to the field lines, these drifts are accounted for when relying on the conventional Lagrangian. Parallel to the field lines, the inclusion of the relevant corrections relies on the correct definition of the parallel guiding-centre velocity U according to (3.9).

It is important to note that the higher-order terms in (3.9), (3.10) do not gyroaverage to zero. Gyroaveraging (3.9) $\langle \dots \rangle = \frac{1}{2\pi} \oint d\zeta \dots$, yields

$$U = \langle v_{||} \rangle + \epsilon_B \frac{\mu}{q} \tau + O(\epsilon_B^2) \quad (3.28)$$

where $\tau(\mathbf{x}) \equiv \mathbf{b} \cdot \nabla \times \mathbf{b}$ is the field-line twist. Equation (3.28) shows that there is a difference between the gyroaverage of the local parallel particle velocity $v_{||}$ and the parallel guiding-centre velocity $U = \mathbf{b}(\mathbf{X}) \cdot \dot{\mathbf{X}}$. The difference between the two is the Baños drift parallel to the field lines, which is induced by the magnetic field line twist [Ban67] (see (A.4) and surrounding discussion). In general, this drift correction therefore would be expected to give a finite contribution to the guiding-centre current \mathbf{j}_{gc} at $O(\epsilon_B)$.

As will be seen below, the computation of a slowing-down distribution for NBI represents a special case, because the source distribution is given in physical coordinates, while the slowing-down distribution is to be computed in guiding-centre coordinates. The use of a consistent pushforward operation is thus required to obtain the correct source distribution in guiding-centre coordinates. This is in contrast to e.g. the computation of the distribution for ion cyclotron resonance heating (ICRH) or ion cyclotron current drive (ICCD) where particles are heated from a thermal background distribution that is usually expressed a priori as a function of the guiding-centre constants of motion. In the latter case, the push-forward relation (3.23) is not needed for the correct guiding-centre initialization. In particular, the Baños drift is then already implicitly included in the parallel guiding-centre velocity and will not appear explicitly. In a similar way, the Baños drift would be implicitly accounted for in extensions of guiding-centre theory such as gyrokinetics, if the guiding-centre Lagrangian and coordinates considered in the present work are used. Our discussion of higher-order effects in section 3.4.1 focuses on higher-order corrections related to strong anisotropy in the source distribution and does not apply in cases that do not involve the push-forward operation.

To make contact with the approach to drift-kinetic theory based on direct gyroaveraging of the

Vlasov equation, as discussed in detail in [HM03], we also consider the gyroaverage of (3.10)

$$\mu = \langle \mu_0 \rangle - \epsilon_B \frac{\mu U}{\Omega} \tau + O(\epsilon_B^2). \quad (3.29)$$

Combining this expression with $d\mu/dt = 0$ yields at leading order

$$\frac{d\langle \mu_0 \rangle}{dt} = \epsilon_B \frac{d}{dt} \left(\frac{\mu U}{\Omega} \tau \right) + O(\epsilon_B^2) \quad (3.30)$$

$$= \epsilon_B U \mu \mathbf{b} \cdot \nabla \left(\frac{U}{\Omega} \tau \right) + O(\epsilon_B^2). \quad (3.31)$$

This expression recovers the corresponding result obtained in [HM03, eq. (4.69)], under the assumption of a time-independent background field. We note that the $O(\epsilon_B)$ -term on the right-hand side of (3.30) appears to be subsequently neglected in [HM03, eq. (4.44) and discussion] to arrive at a simpler form of the gyro-averaged Vlasov equation. In contrast, the approach based on Lie perturbation theory discussed in the current chapter retains a simple form of the Vlasov equation while accounting for the variation in $\langle \mu_0 \rangle$. This is achieved by employing μ instead of $\langle \mu_0 \rangle$ as a guiding-centre variable. Neglect of the right hand side of (3.30) as proposed in [HM03] is only justified under additional assumptions on either the ordering of the parallel current $j_{||} \propto \tau \equiv \mathbf{b} \cdot \nabla \times \mathbf{b}$ (cp. eq. (3.43) below), or on the distribution of guiding-centres (approximate isotropy). Neither of these assumptions hold for the case of NBI in a MAST-like equilibrium, which is studied in section 3.4.1 below.

In accordance with [PGC15], we conclude that relying on the leading order terms in the definition of the guiding-centre variables results in a neglect of the Baños drift. In appendix A, it is shown how the results of [Pfe15, PGC15], which have been derived by direct inspection of the full particle dynamics, are recovered by Lie perturbation methods to first order in ϵ_B . The discussion of appendix A will be based on the Lagrangian (3.21) and relations (3.8)–(3.11), which as already indicated, lead to guiding-centre treatment neglecting only drift terms of order $O(\epsilon_B^2)$.

3.2.6 Practical implications

For numerical applications, an increase in computational speed can only be achieved if the short time-scales of the gyromotion do not have to be resolved, thereby allowing the use of a much larger timestep. In practice, simulations are carried out on the reduced phase-space (X, U, μ) , with μ treated as a parameter. As the mapping $(x, v) \rightarrow (X, U, \mu)$ is no longer invertible, this implies a loss of information. This is a trade-off we are willing to make for an increase in computational speed. Since the push-forward relation (3.23) might in general contain gyroangle dependent information, we proceed to eliminate this information by gyroaveraging. Using the relations (cp. [BH07, equation (B13)])

$$\langle \rho_0 \rangle = 0, \langle G_1^{v_{||}} \rangle = \mu/q\tau, \langle G_1^{\mu} \rangle = -\mu v_{||}\tau/\Omega, \quad (3.32)$$

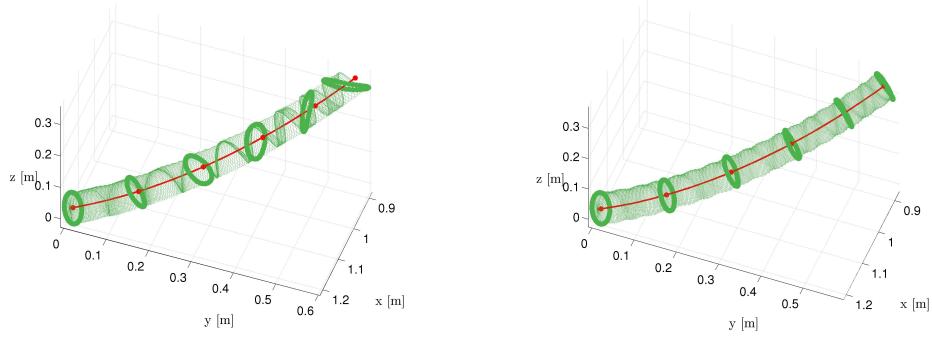


Figure 3.2 – Physical particles sharing the same guiding-centre are initialized according to conventional (left) and higher-order (right) expressions (initialization near $y = z = 0$). Dots indicate snapshots of the position of these markers at several fixed points in time during their evolution. The lack of first-order corrections in the parallel velocity at the initialization of full Lorentzian markers (left) causes distortions, as well as an average lag behind the guiding-centre (missing Baños drift correction).

assuming $f = f(\mathbf{x}, v_{||}, \mu)$ and defining the gyro-averaged part of the distribution function $F(\mathcal{Z}_{gc}) \equiv \langle F_{gc}(\mathcal{Z}_{gc}) \rangle = \langle f(\mathcal{Z}_0) \rangle = \langle f(\mathcal{Z}_{gc}) \rangle - \epsilon_f \langle G_1 \cdot df|_{\mathcal{Z}_{gc}} \rangle + O(\epsilon_B^2, \epsilon_f^2)$, we obtain

$$F = f - \epsilon_f \frac{\mu}{q} \tau \left(\frac{\partial f}{\partial v_{||}} - \frac{v_{||}}{\Omega} \frac{\partial f}{\partial \mu/q} \right) + O(\epsilon_B^2, \epsilon_f^2, \epsilon_B \epsilon_f). \quad (3.33)$$

In a conventional leading order guiding-centre treatment, the first order contributions involving $\tau \equiv \mathbf{b} \cdot \nabla \times \mathbf{b}$ would be neglected.

For numerical applications, it is convenient to express these relations in terms of $(\mathbf{x}, \lambda_{||}, E)$ and $(\mathbf{X}, \lambda, E_{gc})$, where $\lambda_{||} \equiv v_{||}/v$ and $\lambda = U/\sqrt{2E/m}$ denote the particle and guiding-centre parallel pitch variables, respectively. We evaluate the particle energy $E \equiv \frac{1}{2}mv_{||}^2 + \mu_0 B(\mathbf{x})$ in terms of the guiding-centre coordinates \mathcal{Z}_{gc} , given by (3.8)-(3.11):

$$E_{gc}(\mathcal{Z}_{gc}) \equiv E(\mathcal{Z}_0) = E(\mathcal{Z}_{gc}) - \epsilon_B G_1^E|_{\mathcal{Z}_{gc}} + O(\epsilon_B^2). \quad (3.34)$$

Thanks to the convenient cancellation in the energy component G_1^E of the generating vector field [CB09]:

$$G_1^E \equiv G_1^{v_{||}} \frac{\partial E}{\partial v_{||}} + G_1^\mu \frac{\partial E}{\partial \mu} + G_1^x \cdot \nabla_x E = 0, \quad (3.35)$$

we obtain $E_{gc} = E + O(\epsilon_B^2) = \frac{1}{2}mU^2 + \mu B(\mathbf{X}) + O(\epsilon_B^2)$. No distinction therefore needs to be made between the particle energy and the energy of its associated guiding-centre, to first order in ϵ_B .

The phase-space density in terms of the variables $(\mathbf{x}, \lambda_{||}, E)$ and (\mathbf{X}, λ, E) is expressed as

$$d^3\mathbf{x} d^3\mathbf{v} = \sqrt{\frac{2E}{m^3}} d^3\mathbf{x} d\lambda_{||} dE d\zeta_0 \quad (3.36)$$

$$= \sqrt{\frac{2E}{m^3}} \frac{B_{||}^*}{B} d^3\mathbf{X} d\lambda dE d\zeta. \quad (3.37)$$

In terms of the coordinates (\mathbf{X}, λ, E) , we obtain the following expression for the numerical distribution function

$$F = f - \epsilon_f \frac{\mu\tau}{qv} \frac{\partial f}{\partial \lambda_{||}} \Big|_E + O(\epsilon_B^2, \epsilon_f^2, \epsilon_B \epsilon_f). \quad (3.38)$$

The terms on the right-hand side are evaluated at the guiding-centre coordinates, e.g.

$$f = f(\mathbf{x}, E, \lambda_{||})|_{\mathbf{x} \rightarrow \mathbf{X}, E \rightarrow E_{gc}, \lambda_{||} \rightarrow \lambda}, \quad (3.39)$$

a clear distinction is made between $\lambda_{||}$ and λ at the order considered. Neglect of the correction term in (3.38) is questionable unless we may assume an additional ordering $\epsilon_f \ll \epsilon_B$. This is manifestly not the case for NBI in a MAST-like equilibrium, though the correction would clearly be zero for a fusion α -particle birth distribution. A comparison between models including and neglecting corrections related to ϵ_f for NBI in a MAST-like equilibrium is given in section 3.4.1.

3.3 Single particle dynamics

The theory described in the last section has been implemented in the guiding-centre following code VENUS-LEVIS [PCGM14]. This includes a consistent full-Lorentz/guiding-centre switching taking into account higher-order corrections, as well as higher-order guiding-centre dynamics obtained from the higher-order Lagrangian (3.4). Numerical investigations of the effect of higher-order corrections will be presented in this section.

An algorithm for switching between full-Lorentzian and guiding-centre following based on leading-order relations, as well as a field variation estimator have been presented in [PGC15]. In the present work, the switching algorithm is refined by inclusion of higher-order corrections.

The meaning of the higher-order terms in (3.8)-(3.11) are illustrated in eq. (3.9). We identify two sets of terms

$$U = v_{||} + \underbrace{\frac{\mu}{q} \mathbf{a}_1 : \nabla \mathbf{b} - v_{||} \boldsymbol{\rho}_0 \cdot \boldsymbol{\kappa}}_{(I)} + \underbrace{\frac{\mu}{q} \mathbf{b} \cdot \nabla \times \mathbf{b}}_{(II)}. \quad (3.40)$$

The terms (I) are gyroangle-dependent and gyroaverage to zero. These terms account for the variations of the parallel particle velocity on the gyration time scale; they carry out a

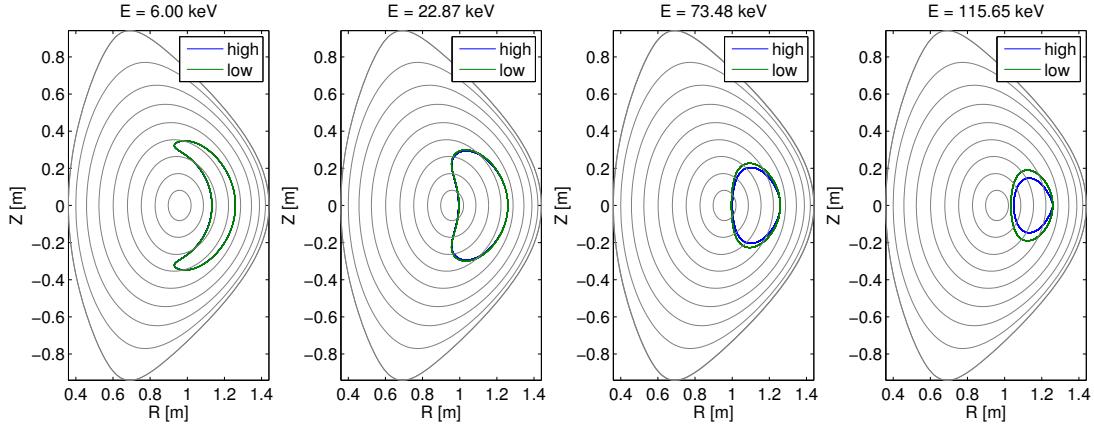


Figure 3.3 – Poloidal projection of representative guiding-centre orbits initialized at different energies and evolved using higher-order equations of motion ('high') and conventional equations of motion ('low'). The difference between the curves 'high' and 'low' are due to the $O(\epsilon_B^2)$ drifts, which arise from the additional term $-\epsilon_B m \mu / q^2 \bar{R}$ in the higher-order Lagrangian (3.4).

gyroaverage of $v_{||}$, so that $v_{||} + (I) = \langle v_{||} \rangle + O(\epsilon_B^2)$ is invariant on the gyromotion time-scale. The term (II) on the other hand is a drift term that varies on the slower time-scale of guiding-centre motion. The difference between $v_{||}$, $\langle v_{||} \rangle$, and U in a purely sheared background is clearly visible in figure A.2 of appendix A. Note that the term (I) serves to generalize that of (A.4) for more general magnetic fields.

In figure 3.2, we show the evolution of full Lorentzian particles initialized to correspond to the same guiding-centre. Compared are the results obtained from initialization based on the conventional relations (3.18)–(3.20) and the higher-order relations (3.8)–(3.11). It is observed that the conventional leading-order expressions quickly lead to distortions from the full-Lorentz/guiding-centre correspondence. The difference is due to the lack of first-order corrections in the conventional relations, and in particular the neglect of corrections in the correspondence of particle and guiding-centre parallel velocities (eq. (3.40)). The consequences of the inclusion of such correction terms for models of NBI injection will be discussed in section 3.4.1, a detailed analytical derivation in a simple background field is presented in appendix A.

To investigate the effect of the additional higher-order terms in (3.4), scans in energy of characteristic orbit quantities have been performed in a MAST-like equilibrium. A representative guiding-centre was initialized at a fixed initial position $\rho_T = 0.55$, $\theta = 0$ in the poloidal plane (here $\rho_T = \sqrt{\Phi_N}$, where Φ_N = normalized toroidal flux), and the parallel pitch was fixed at $\lambda = -0.4$. The energy was varied from 10 keV to 116 keV. These energies can be compared with the NBI model considered in section 3.4.1, where guiding-centres are injected at a maximal energy $E_0 = 56$ keV. The results from conventional guiding-centre equations (neglecting the higher-order terms related to \bar{R} in (3.4)) and the higher-order equations are compared in figure 3.3.

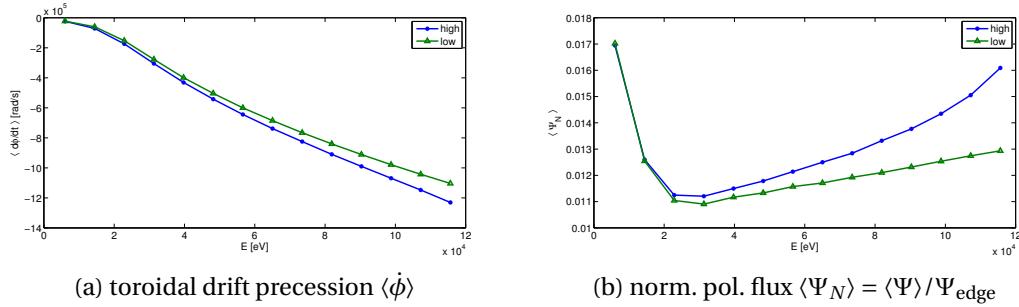


Figure 3.4 – Energy dependence ($E = 10\text{ keV}$ to $E = 120\text{ keV}$) of characteristic orbit quantities obtained from higher-order ('high') and conventional ('low') guiding-centre equations of motion. Shown are the bounce-averaged toroidal drift precession $\langle \dot{\phi} \rangle$ (left) and the bounce-averaged value of the normalized poloidal flux $\langle \Psi_N \rangle = \langle \Psi \rangle / \Psi_{\text{edge}}$ along the orbit. The difference between the curves 'high' and 'low' are due to the $O(\epsilon_B^2)$ drifts, which arise from the additional \bar{R} -term in the higher-order Lagrangian (3.4).

The additional higher-order corrections in A^* scale as $\epsilon_B \sim \sqrt{E}$ as compared to the next lower-order terms. As expected, the results obtained from the conventional and higher-order guiding-centre Lagrangian agree to good accuracy in the limit of low energies, but differ as the energy is increased. Going even further up in energy, the correction terms have been observed to locally dominate at energies $E > 130\text{ keV}$, indicating that the guiding-centre approximation completely breaks down at such high energies in a MAST-like equilibrium. For this reason, the poloidal projections of orbits in figure 3.3 are restricted to a maximum energy of 116 keV . Shown in figure 3.4 are results for the toroidal drift precession $\langle \dot{\phi} \rangle$ and the average (normalized) poloidal flux value $\langle \Psi_N \rangle$, with $\Psi_N \equiv \Psi / \Psi_{\text{edge}}$.

The results of figures 3.3 and 3.4 indicate that only at very high energies with associated large Larmor radii and when approaching the limits of the guiding-centre approximation, should one see appreciable differences in the guiding-centre dynamics expanded at different orders.

3.4 Application to slowing-down simulations

In 3.4.1, we discuss the implications of higher-order terms for the modeling of NBI injection in a MAST-like equilibrium. In sections 3.4.2 and 3.4.3, we compare results obtained with the inclusion of higher-order correction (3.23) to the results obtained from a conventional approach. In section 3.4.5, slowing-down simulations based on the higher-order equations presented in section 3.2 are applied to the computation of fusion α -particle loss in the European DEMO design, and compared to results obtained from conventional guiding-centre equations.

3.4.1 Corrections for strong anisotropy (NBI)

The focus of this section is on the implications of the higher-order corrections for models of NBI. The deposition distribution of NBI ions in MAST is strongly anisotropic, so that the first-order correction in equation (3.38) is non-zero. Furthermore, NBI is usually focused in the core region, where the magnetic field-line twist $\tau \equiv \mathbf{b} \cdot \nabla \times \mathbf{b}$ is maximal. Therefore, the ϵ_f -term in (3.38) is not expected to be negligible in this case. Focusing on the effects of strong anisotropy, we shall assume the ordering $\epsilon_f > \epsilon_B$ in this section. In a first approximation, we will neglect the corrections associated to the $\bar{\mathbf{R}}$ -term in (3.4), and only retain the correction associated with ϵ_f . A more formal justification for the ordering may be given by considering

$$\epsilon_f \sim \frac{v_{Baños}}{v} \left| \frac{\partial \log f}{\partial \lambda_{||}} \right| \sim \epsilon_B \left| \frac{\partial \log f}{\partial \lambda_{||}} \right|. \quad (3.41)$$

The relation between ϵ_B and ϵ_f is therefore uniquely determined by the scale length of the distribution in the parallel pitch $\lambda_{||}$. In this section, we consider the limit of strong anisotropy, expecting the corrections to the source distribution which are induced by the push-forward to dominate higher-order corrections. Figure 3.5 shows the general shape of the NBI source distributions obtained in MAST, which indicates that we may expect a value of this scale length of around $\Delta\lambda = 0.3 - 0.5$. Hence, we expect $\epsilon_f \sim 2\epsilon_B$ to $3\epsilon_B$. We acknowledge that higher-order drift terms in the guiding-centre dynamics may give additional corrections in the case of MAST by e.g. slightly altering the toroidal drift precession as shown in figure 3.4, which may in turn e.g. influence the resonant interaction with field perturbations and hence the induced losses. The detailed investigation of such effects has been found to be associated with numerical difficulties due to the coordinate singularity near the axis in the case of MAST, and is left for future work.

The simplified NBI model used in VENUS-LEVIS was described in detail in [Alb11]. This model accounts for ion deposition computed based on ionisation rates depending on the plasma background profiles and takes into account beam spread. Markers are distributed uniformly along the beamline and given a weight proportional to the ionisation rate. The beam spread is accounted for by the introduction of a random Gaussian displacement. The position, energy and the parallel component of the velocity along the field line $v_{||} = \mathbf{b}(\mathbf{x}) \cdot \mathbf{v}$ are used to initialize the guiding-centre marker at ionization. Finite Larmor radius corrections, in particular the Baños drift correction to $v_{||}$ (expressed by the first order term in (3.23)), were previously not taken into account in [Alb11]. While there are NBI models which take into account more detailed information on beam geometry [PMA⁺04, AGB⁺15], and even finite Larmor radius effects associated to the ρ_0 correction, the corrections due to the higher-order effects considered in the present work have, to the best of our knowledge, not been discussed before.

The finite Larmor-radius corrections considered in this section are expressed by the gyroaveraged push-forward (3.38). We consider two models. The first model (A) employs the simplified VENUS-LEVIS NBI model where $F(\mathbf{X}, U, E) = \langle f(\mathbf{X}, U, E, \zeta) \rangle$ neglects the Baños drift correction.

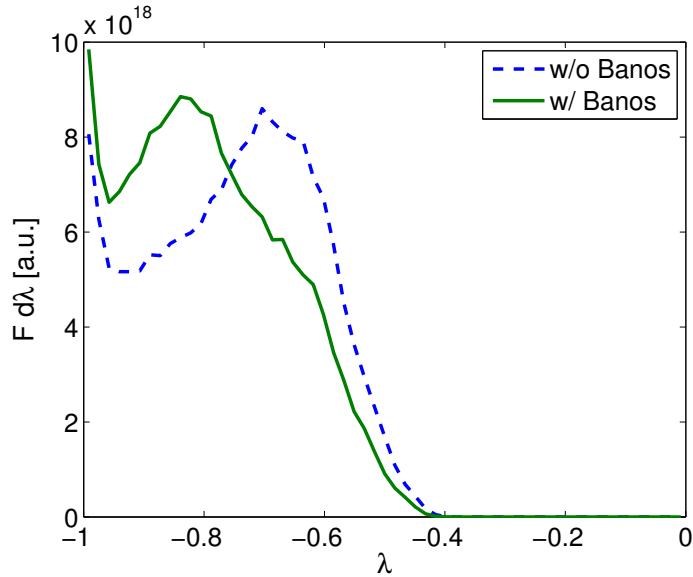


Figure 3.5 – Deposition rate of NBI markers in terms of parallel guiding-centre pitch variable λ . A comparison is made between the conventional model and the model including the Baños drift correction to the parallel guiding-centre velocity (3.23).

The second model (B) on the other hand includes this parallel correction. The simulations were set up as follows: an ensemble of markers with variables $X, \lambda_{||}, E$ was initialized according to the physical distribution provided by the model in VENUS-LEVIS. This furnished the NBI deposition distribution for model (A). To obtain the source distribution for (B), the same distribution was taken, but the relation $\lambda = \lambda_{||} + v_{Baños}/v$ was used to initialize guiding-centres including the Baños drift correction. This corresponds to a marker initialization in accordance with the gyro-averaged push-forward relation (3.38). Additional complications with a possible gyroangle dependence of the physical source distribution are here neglected. As a result, the markers of models (A) and (B) are deposited identically in the poloidal plane, as well as toroidally. The only difference is the missing Baños correction of model (A), which is taken into account in model (B). This allows us to isolate and quantify effects related to the consistent inclusion of parallel drifts, independently of other finite Larmor-radius (FLR) corrections, associated to e.g. ρ_0 or the gyroangle dependent terms in (3.9). The source distributions for models (A) and (B) as a function of parallel guiding-centre pitch λ are depicted in figure 3.5 for the unbalanced NBI beam application of MAST that is considered in the next subsection. All slowing-down simulations are performed until a steady state is reached, so that the NBI source is balanced by the fast particle sinks. The sinks of our fast ion model are given by losses to the last closed flux surface (LCFS), as well as thermalization with the background. A guiding-centre is considered thermalized, and removed from the simulation, if its energy falls below a constant multiple of the background thermal energy. As in earlier studies [PMCG15], a thermalization threshold factor of 3 is chosen for all simulations.

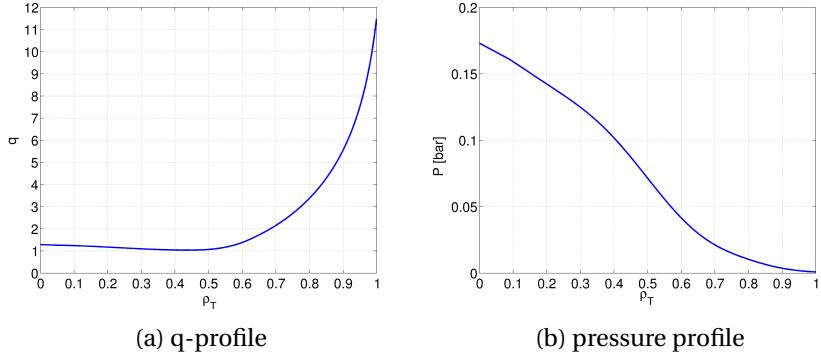


Figure 3.6 – Profiles determining the MAST-like equilibrium considered in this section.

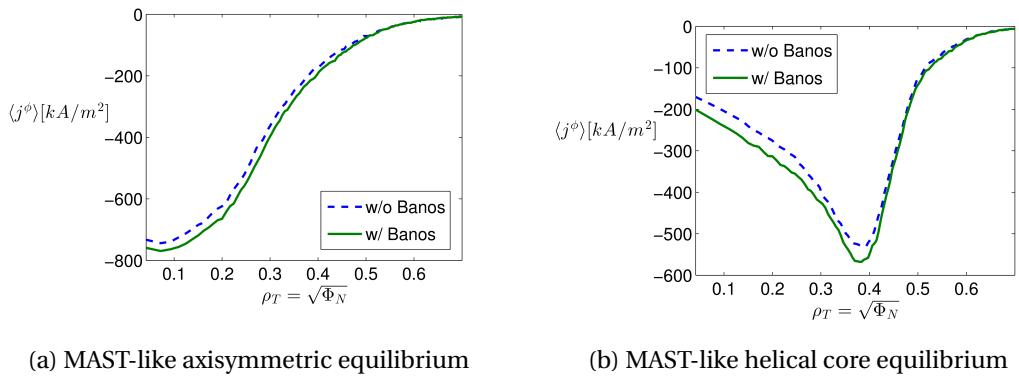


Figure 3.7 – Comparison of current profiles as computed from the slowing-down simulation for the NBI injection models considered in this section for a MAST-like equilibrium (bifurcated state).

3.4.2 Unbalanced NBI in MAST

Slowing-down simulations have been performed for models (A) and (B), and for a MAST-like equilibrium with non-monotonic q -profile and q_{\min} very close to 1 (cp. fig. 3.6). The MAST magnetic equilibrium was modelled with the fixed boundary VMEC code. This equilibrium bifurcates into two sister states; an axisymmetric equilibrium and a helical core sister state, modelling a saturated internal kink [CGP⁺10]. We choose first to compare the models for the axisymmetric branch. NBI markers are injected at energies E_0 , $E_0/2$, $E_0/3$ with $E_0 = 56\text{keV}$. To compare models (A) and (B) the NBI current density associated to the slowed-down distribution is computed and shown in figure 3.7a. The inclusion of the FLR correction related to the Baños drift is found to yield an increase of $\approx 4\%$ to the computed NBI driven current from models (B) as compared to (A) after slowing-down, reflecting the shift in the source distribution seen in figure 3.5.

Models (A) and (B) have similarly been compared for the helical branch with helical displacement $\delta_h = 0.23$ [CGP⁺10, PGC15]. Again, inclusion of the Baños drift correction has been found to lead to an increased NBI-driven current in model (B) as compared to (A). In this case, the

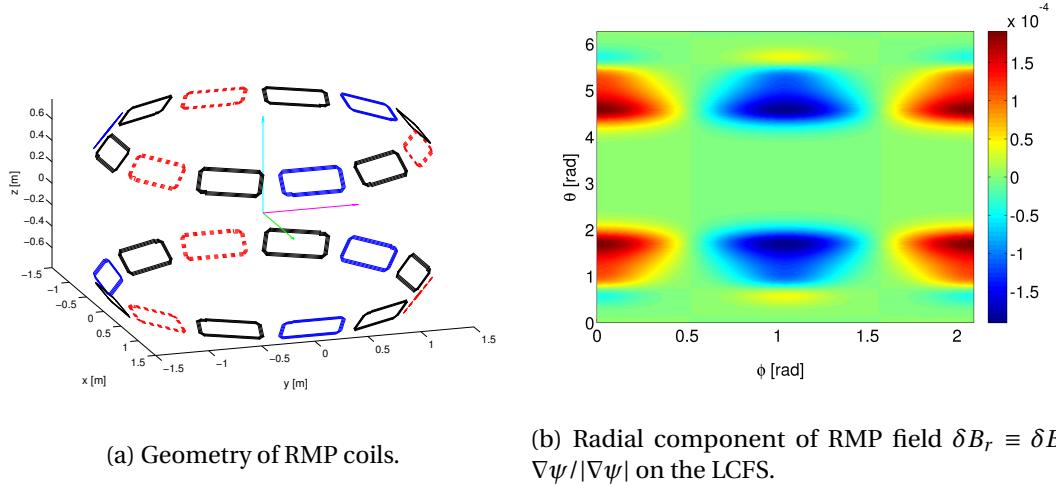
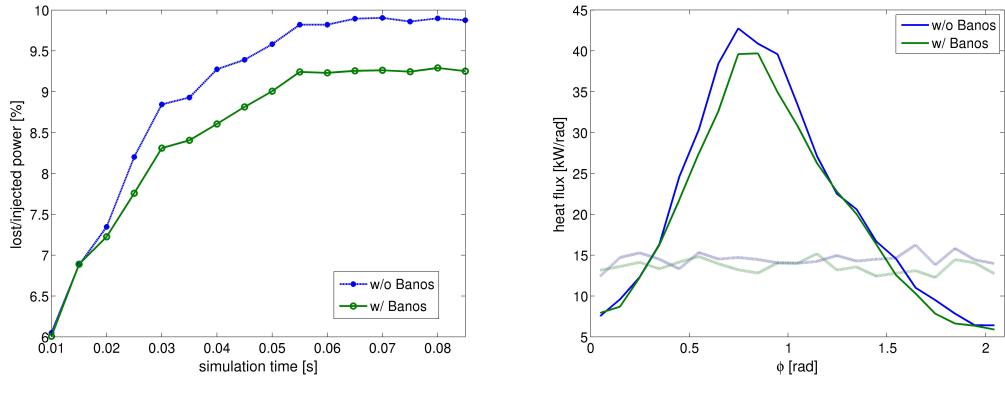


Figure 3.8 – Model of RMP perturbation coils in MAST (5.6 kA). The colors in (a) indicate the direction of current in the coils, corresponding to the sign of the normal component of the magnetic perturbation ($\delta B_r > 0$ (red), $\delta B_r < 0$ (blue), no current in coil (black)).

difference between the computed currents after slowing-down is found to be higher by $\approx 8\%$, when comparing (B) to (A). The resulting current density profiles are compared in figure 3.7b. It is observed that the largest difference in the current profiles occur in the high-current region near the axis.

3.4.3 Unbalanced NBI losses due to RMP

We consider the axisymmetric branch of the equilibrium of the last section, but in addition, add RMP in the vacuum approximation (neglecting the plasma response). The assumed RMP coil configuration has a $n = 3$ symmetry, with 0° -phase [Pfe15, PMCG15], as shown in figs. 3.8a and 3.8b. The RMP coil currents are set to 5.6 kA. Again, we compare two models for NBI with and without inclusion of the Baños drift. In the case of RMP, we expect the shift in the deposition distribution in λ , which reflects the inclusion of the Baños drift, to affect the overall particle losses, which are due to resonances with the magnetic field perturbations [PMCG15]. As deeply passing particles are less affected by such perturbations, we would expect the Baños drift visible in figure 3.5, to increase NBI particle confinement, i.e. to reduce the fast ion heat flux to the LCFS. The results of slowing-down simulations for the two models are shown in figure 3.9. The saturation of the total computed heat flux over the simulation time is shown in figure 3.9a, the toroidal distribution of the heat flux is shown in 3.9b, where we take into account the $n = 3$ symmetry of the applied perturbation, and show only the heat flux for $\phi \in [0, 2\pi/3]$. The results of these simulations agree qualitatively with the expected results. At a neutral beam injected power of 1500 kW, the total heat fluxes computed with inclusion of the Baños drift are 138.8 kW as compared to 148.2 kW without this drift. The inclusion of the Baños drift is thus seen to reduce the expected losses by 6.3% in the present case. The inclusion of the Baños drift is therefore seen to affect the expected resonant behaviour of ions with MHD



(a) Evolution of total heat flux through the LCFS over simulation time.
 (b) Local heat flux through the LCFS as function of toroidal variable ϕ .

Figure 3.9 – Comparison of predicted NBI heat fluxes through the LCFS computed in presence of RMP. Compared are two NBI models, one neglecting and one including the Baños drift correction in the initialisation of guiding-centre distributions. For the loss distribution in ϕ the results are plotted for $\phi \in [0, 2\pi/3]$, taking into account the $n = 3$ symmetry of the resonant magnetic field perturbations. Subfigure (a) shows the temporal evolution of the lost / injected power over the simulation. A steady state is reached after about 0.06 s. Subfigure (b) shows the local heat flux at steady state as a function of toroidal angle ϕ . The solid lines show the heat fluxes obtained for the two considered models in presence of RMP. For comparison, the pale lines correspond to a simulation in the absence of RMP, i.e. axisymmetric background.

perturbations.

3.4.4 Balanced NBI with RMP

The previous applications demonstrate that the Baños drift correction can be non-negligible for unbalanced NBI in a MAST-like equilibrium. To show that the corrections discussed in the present work may also have implications for balanced NBI, we consider a (hypothetical) balanced NBI model for the axisymmetric MAST-like equilibrium. This model is obtained by doubling the number of beam lines in comparison to the unbalanced model, adding the mirror image of each beam-line with respect to the vertical plane running parallel to the beam-line, and containing the $R = Z = 0$ axis. Again, models (A) and (B) are obtained in the same way as before, with model (B) accounting for the Baños drift correction. The same external magnetic (RMP) perturbations are applied (see figure 3.9). Figure 3.10 collects results from this simulation. For this application, the deposition depicted in figure 3.10a replaces the unbalanced deposition shown in fig. 3.5. It is seen that the balanced model without Baños drift correction has a distribution symmetric in λ . This symmetry is broken by the Baños drift correction (3.23). Overall, the losses for both models are clearly dominated by the injection with $\lambda > 0$, due to the well-known finite orbit-width effects. The Baños drift (FLR) correction leads to improved confinement for $\lambda < 0$, while it increases the losses for $\lambda > 0$ (cp. figure 3.10b). For a NBI power of 1500 kW, the losses for model (B) with inclusion of the Baños drift are found

to be 585.0 kW, as compared to 559.8 kW for model (A). This corresponds to an increase of 4.6%. The saturated currents are shown in figure 3.10c. The fast ion current is of course weak for balanced neutral beam injection. The Baños drift is found to lead to a relative increase of more than 28% over the weak but finite current obtained for balanced injection without Baños drift. It is noted that although the current is weak for balanced injection, the distribution function is nevertheless distorted for the Baños case. One may expect such modifications to be important for pressure driven MHD instabilities, where e.g. it has been mentioned in chapter 1 that such modes are modified by weak corrections to the equilibrium.

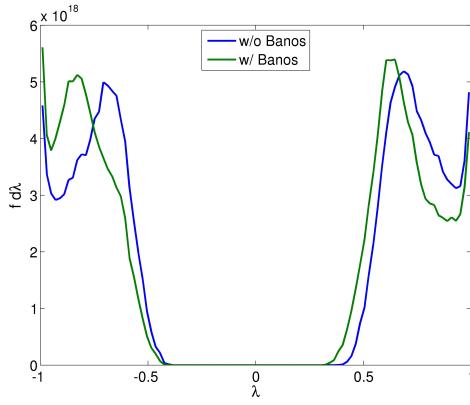
3.4.5 Fusion α particle loss in DEMO

We investigate the sufficiency of the conventional guiding-centre equations for the current European DEMO design, focusing in particular on toroidal field ripple induced losses. The study is based on an 18 toroidal field coil design. A detailed description of the equilibrium used for this study, as well as a discussion of the expected toroidal field ripple in DEMO, can be found in [PCFG16] and [WAA⁺17]. For our slowing-down simulations, the (axisymmetric) background is computed using the free-boundary version of the VMEC equilibrium code in realistic coil geometry. The ripple field perturbation is computed in the vacuum approximation and algebraically added to the axisymmetric background (termed “2D+vacuum” approximation in [PCFG16]).

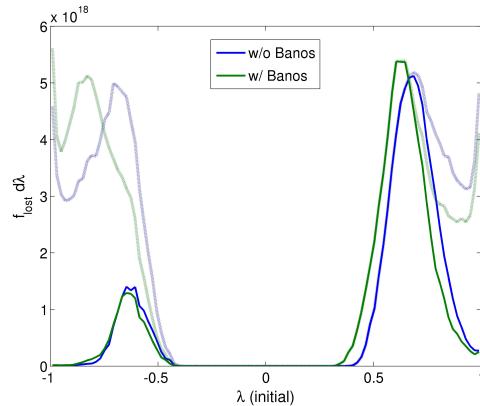
The physical fusion α birth distribution density is assumed isotropic, i.e. of the form $f = f(\mathbf{x}, E)$. As derived in (3.23), the corresponding guiding-center distribution F is then functionally the same $F(\mathbf{X}, E) = f(\mathbf{X}, E)$ to first order in Larmor radius (3.23), which is to say that $\epsilon_f = 0$ in this case. The higher-order corrections considered in this section are given by (3.36) and the $\bar{\mathbf{R}}$ -term in (3.5). Note that the anisotropic push-forward corrections of the last section are not relevant for this application, due to the intrinsic isotropy of the alpha distribution (see earlier discussion).

As discussed in [Pfe15, sec. 4.7.2] the correct inclusion of the guiding-centre Jacobian (3.36) can be represented by a weight proportional to $B_{||}^*$. A marker weight $\propto B_{||}^*$ has already been employed in earlier studies using the VENUS-LEVIS code, where the correct factor was found by inspection of the canonical volume element associated to (3.21). Therefore, even though the consistent switching based on Lie transform methods was not available previously, the results presented in [PCFG16] are consistent with the *higher-order* relations (3.8)-(3.11). In the following, we consider the possible effect of higher ($O(\epsilon_B^2)$) order drifts arising from the $\bar{\mathbf{R}}$ -terms in equation (3.5). To this end, two slowing-down simulations have been performed first neglecting higher-order drifts (simulation B, based on (3.21)) and then including higher-order drifts (simulation C, based on (3.4)) for identical birth distributions. The results of these simulations are shown in table 3.1. In addition, we also evaluated the losses obtained when neglecting the higher-order correction in the marker weights (simulation A, based on (3.21)), i.e. under the approximation $B_{||}^* \approx B$.

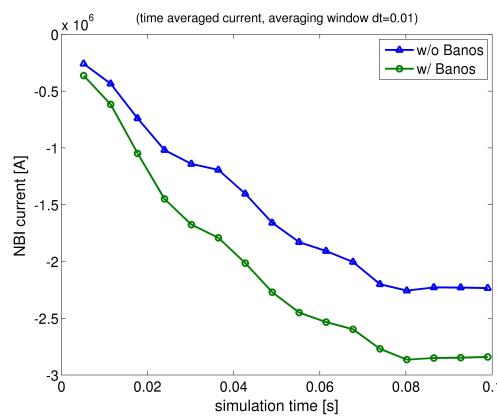
3.4. Application to slowing-down simulations



(a) Deposition rate (initial source) of NBI markers in terms of parallel guiding-centre pitch variable λ for the balanced injection model. The sign of the Baños drift correction is independent of the sign of λ , leading to an overall shift of the guiding-centre distribution.



(b) Loss rate of NBI markers in terms of initial parallel guiding-centre pitch variable λ (solid lines). Also indicated is the deposition rate (pale lines). Losses are dominated by counter-current injection due to finite orbit width effects ($\lambda > 0$).



(c) Evolution of the fast ion guiding-centre current over simulation time.

Figure 3.10 – Comparison of models including/neglecting the Baños drift correction (3.23) for (hypothetical) balanced NBI in a MAST-like equilibrium, in the presence of RMP. ⁷⁹

	P_{LCFS}	$P_{\text{LCFS}}/P_{\text{fus}}$	$\Delta P/P_{(B)}$
(A)	$736 \pm 6 \text{ kW}$	0.1223 %	$0.11 \pm 0.96 \text{ %}$
(B)	$736 \pm 6 \text{ kW}$	0.1222 %	$0.0 \pm 0.84 \text{ %}$
(C)	$762 \pm 7 \text{ kW}$	0.1255 %	$3.65 \pm 0.84 \text{ %}$

Table 3.1 – Comparison of computed total heat flux to LCFS, as predicted by guiding-centre approximation at different orders in ϵ_B for fusion α -particles in DEMO. (A) was based on Lagrangian \mathcal{L}_0 , marker weights $w \propto B$, (B) based on Lagrangian \mathcal{L}_0 , marker weight $w \propto B_{||}^*$, (C) based on Lagrangian \mathcal{L} , weight $w \propto B_{||}^*$. Shown are the heat fluxes to the LCFS after saturation. The uncertainty in these values is estimated as the standard deviation after saturation of the fluctuations about the mean value (numerical/Monte-Carlo noise).

Due to the large machine size and strong B-field, no essential differences were observed when comparing the higher-order and conventional guiding-centre models for the case of fusion α particles in the European DEMO design. Differences are within other modelling uncertainties as discussed in detail in [PCFG16]. It is interesting to note that, for these simulations, the difference between simulations (A) and (B), which is due to the neglect of first-order terms in $B_{||}^*$, is in fact smaller than the difference in simulations (C) and (B), which formally differ in the addition of second order drift terms. This can be explained in part by the fact that α -particle losses in DEMO occur dominantly close to the last closed flux surface (LCFS) on the outboard side, where $\mathbf{b} \cdot \nabla \times \mathbf{b} = \hat{\mu}_0 j_{||}/B = 0$. Therefore, we have $B_{||}^* \approx B$ to high accuracy for almost all guiding-centres that are lost to the LCFS. The difference between simulations (B) and (C) on the other hand is to be attributed to second-order drifts. These drifts are non-zero (though still essentially negligible) even close to LCFS, as indicated by the finite value of the $O(\epsilon_B)$ -correction ($\approx -\mu\tau b_z/q$) to P_ϕ on the LCFS.

3.5 Discussion

In this section, the importance of the push-forward operation is discussed and summarized for cases where source distributions are anisotropic. In particular, auxiliary heating systems such as NBI give rise to highly anisotropic energetic particle populations (cp. figure 3.5). For such distributions, the first-order correction term in the push-forward (3.23) is finite due to the non-trivial dependence of the distribution function on the parallel particle pitch $\lambda_{||}$, and can become non-negligible due to the particles' large Larmor radii.

The neglect of the correction term of equation (3.23) is expected to lead to an error

$$\Delta \mathbf{j} \sim \epsilon_B \frac{\tau(p_\perp - p_{||})}{B} \mathbf{b} + O(\epsilon_B^2) \quad (3.42)$$

in the estimated current density (cp. the difference between the current profiles in figure 3.7). This error evidently only vanishes if the distribution function is isotropic. Furthermore, $\Delta \mathbf{j}$ is expected to be biggest in the region with highest field-line twist τ , usually near the magnetic

axis (cp. also with eq. (3.43)). Formally, we have $\Delta \mathbf{j} \sim \epsilon_B \mathbf{j}$, in agreement with the size of the correction to the current profiles visible in figure 3.7.

In the case of the estimation of NBI losses due to RMP, the correction to the guiding-centre distribution given by eq. (3.23), is seen to predict an decreased fraction of trapped particles and, as a consequence, lower particle losses are obtained. This effect clearly depends on the sign of the equilibrium parallel current, since from Ampère's law:

$$\hat{\mu}_0 j_{||,\text{eq}} \equiv \hat{\mu}_0 \mathbf{b} \cdot \mathbf{j}_{\text{eq}} = \mathbf{b} \cdot \nabla \times \mathbf{B} = B\tau, \quad (3.43)$$

where $\hat{\mu}_0$ denotes the vacuum permeability. The hat serves to distinguish $\hat{\mu}_0$ from the magnetic moment at particle position μ_0 introduced earlier. Equation (3.38) may now equivalently be written as

$$F \approx f - \hat{\mu}_0 j_{||,\text{eq}} \frac{\mu}{qB} \frac{\partial f}{\partial v_{||}} \Big|_E, \quad (3.44)$$

with F the guiding-centre distribution function, f the physical particle distribution function, $\hat{\mu}_0$ the vacuum permeability, $j_{||,\text{eq}}$ the parallel component of the equilibrium current density, $\mu \approx mv_{\perp}^2/B$ the magnetic moment and q the charge of the species under consideration. Equation (3.44) reflects the gyroaveraged relation

$$U = \mathbf{b}(\mathbf{X}) \cdot \dot{\mathbf{X}} = \langle v_{||} \rangle + \hat{\mu}_0 j_{||,\text{eq}} \frac{\mu}{qB}. \quad (3.45)$$

Equation (3.45) indicates that in the case of co-current injection, the Baños drift correction expresses a shift of the NBI guiding-centre distribution towards the deeply passing region, which yields a reduction of losses. For counter-current injection, the Baños drift causes a shift towards the barely passing or trapped regions of phase-space, which tends to enhance losses. It is emphasized that the Baños drift correction is purely a finite-Larmor radius effect, which is not to be confused with finite *orbit width* effects which affect losses and similarly depend on co-/counter-current injection.

3.6 Summary and conclusions

Guiding-centre equations based on the Lie transformation approach, including a switching algorithm consistent with the near-identity transformation to $O(\epsilon_B^2)$ and retaining higher-order drift terms have been implemented in the VENUS-LEVIS code for the first time. The resulting higher-order guiding-centre model has been compared with the conventional model used in most guiding-centre following codes.

Higher-order corrections, in particular the first-order Baños drift parallel to the field lines expressed by the push-forward relation (3.23), have been discussed in detail. It has been shown that first-order drift parallel to the field lines can be accounted for even when relying on the conventional guiding-centre Lagrangian. This necessitates an appropriate choice of

guiding-centre variables and the inclusion of higher-order corrections in the relation between particle and guiding-centre, distinguishing between the physical particle parallel velocity and that of its associated guiding-centre. It has been shown that these corrections can affect the expected resonances of particles with RMP, as well as the estimated NBI driven current in a MAST-like equilibrium by up to 8%.

Additional higher-order drift terms arising from the higher-order correction to the symplectic part of the Lagrangian (3.5) have been considered in combination with 3D effects for the study of toroidal field ripple induced losses in the European DEMO design. A clear need for the inclusion of such terms has not been found. Our results suggest that conventional guiding-centre following is sufficient for the purpose of estimating ripple induced fast ion losses in DEMO. The difference between conventional and higher-order models has been found to be within other modelling uncertainties. In general, the fast ion heat load due to toroidal field ripple has been found to be of minor concern for the European DEMO design, with a value well below the expected wall-load limit [WAA⁺17].

Guiding-centre simulations including all drift terms at order ϵ_B can only be assured if higher-order corrections in the full-Lorentz/guiding-centre correspondence are implemented either at the single-particle level, as expressed by equations (3.8)-(3.11), or at the level of distribution functions, given by the push-forward operation (3.38), (3.36), which is directly derived from (3.8)-(3.11). Higher-order corrections to the Lagrangian (3.21) are however not necessary when drift terms of $O(\epsilon_B^2)$ and higher can be neglected, as explained in section 3.2.2. The relevance of corrections in the push-forward correspondence between full-Lorentz/guiding-centre has been discussed in detail for models of neutral beam injection in a MAST-like equilibrium. Existing guiding-centre codes should benefit from the improved accuracy and consistency provided by the inclusion of the higher-order order corrections discussed in the present work.

4 Kinetic-MHD

4.1 Introduction

Plasmas of relevance to fusion research operate at high temperatures and low collisionality. The conditions found in such plasmas clearly place them outside the domain of validity of fluid models such as ideal magnetohydrodynamics (MHD), which are derived based on an assumption of high collisionality. Despite this fact, MHD has proven to be an indispensable and robust tool for the analysis of experiments and remains one of the main sources of understanding of macroscopic instabilities. This ‘unreasonable effectiveness’ of MHD even at high temperatures and corresponding long mean-free path lengths of ions, is commonly understood to arise as a consequence of the presence of a strong magnetic field; Perpendicular to the field lines, a fluid description may be justified, because the short mean-free path that is found in collisional regimes (allowing fluid closure in MHD) is effectively replaced by the small Larmor radius of particles even at low collisionality. However, parallel to the field lines the particles are free-streaming in strongly magnetized near-collisionless plasmas and kinetic effects such as resonant wave-particle interaction can become important.

The study of collisionless, strongly magnetized plasmas was pioneered by Chew, Goldberger and Low [CGL56]. In [CGL56], an attempt was made to derive fluid equations based on an expansion in the parameter $\epsilon_B \sim \rho_i/L \ll 1$, where the Larmor radius of ions ρ_i is assumed to be small relative to the characteristic length scales of the (macroscopic) plasma motion. As a result of the asymptotic expansion of moments of the Vlasov equation in ϵ_B , it was found that the leading-order form of the pressure tensor is no longer isotropic, but instead is given in the so-called Chew-Goldberger-Low (CGL) form $\mathbf{P} = p_{\parallel} \mathbf{bb} + p_{\perp} (\mathbf{I} - \mathbf{bb})$. Here \mathbf{b} is the unit vector pointing along the magnetic field, and p_{\parallel} , p_{\perp} denote the parallel and perpendicular pressure components, respectively. The derivation in [CGL56] does not lead to a closed set of fluid equations because the equations which determine p_{\parallel} , p_{\perp} depend on heat fluxes which are not provided by the model. If these heat fluxes are simply neglected, the CGL double adiabatic fluid model is obtained, according to which the evolution of p_{\parallel} and p_{\perp} is determined

by [CGL56]:

$$\frac{d}{dt} \left(\frac{p_{\parallel} B^2}{\rho^3} \right) = 0, \quad \frac{d}{dt} \left(\frac{p_{\perp}}{\rho B} \right) = 0. \quad (4.1)$$

Here, $d/dt \equiv \partial/\partial t + \mathbf{u} \cdot \nabla$ is the total time derivative with \mathbf{u} the fluid velocity, B the magnetic field strength and ρ denotes the mass density. The derivation of Chew, Goldberger and Low was restricted to lowest order in the ϵ_B -expansion. It was subsequently recognized that finite Larmor-radius (FLR) effects can exert a stabilizing influence on plasmas. Such FLR effects are reflected by higher-order ϵ_B -corrections to the pressure tensor and independent of collisionality. Among the early papers on this subject, Roberts and Taylor [RT62] have pointed out the importance of the additional (higher-order in ϵ_B) components of the pressure tensor which persist in the limit of vanishing collisionality and are closely related to diamagnetic flows. These components form a tensor $\boldsymbol{\pi}_{\wedge}$ that is commonly referred to as the *gyroviscous* tensor. It was shown that the inclusion of the resulting gyroviscous force $\mathbf{F}_{\wedge} \equiv -\nabla \cdot \boldsymbol{\pi}_{\wedge}$ in the momentum equation leads to a *gyroviscous cancellation*, where \mathbf{F}_{\wedge} approximately cancels out the effect of diamagnetic flows in the fluid inertia. A number of authors have subsequently extended the initial results of [RT62] to include additional effects, such as temperature variations and stress tensor drift [CC92], temperature gradients and higher-order moments of the distribution function [Smo98], or the combined effects of collisions and FLR corrections [Kau60, MT71]. Most of these results are obtained in simple slab geometry – a discussion of the assumptions required for the validity of this approximation has for example been given by Hazeltine and Meiss [HM85, Section 4.3.3]. In complex geometry, an expression for the gyroviscous pressure tensor has first been obtained under the fast flow ordering $\partial/\partial t \sim \mathbf{u} \cdot \nabla \sim v_{\text{th}}/L$ by MacMahon [Mac65], whose result remains the state-of-the-art in collisionless fluid theory under these assumptions. The findings of [Mac65] have recently been confirmed by Ramos [Ram05a], who presents a comprehensive discussion of fluid models for collisionless plasmas including alternative orderings. In a later publication [Ram05b], an explicit expression for the gyroviscous force \mathbf{F}_{\wedge} in complex geometry is also provided.

While the fluid moment approach provides useful information on extended fluid effects (diamagnetic flows, heat fluxes, higher-order moments), it does not provide a closed set of equations describing the dynamics of collisionless plasmas, except under very particular circumstances which are not usually found [CGL56]. In general, some form of ad hoc closure remains necessary.

The present chapter is based on the publication [LGPC19]. Following [LGPC19], this chapter focuses on an approach where a closure can be achieved from the solution of a reduced kinetic (guiding-centre or gyrokinetic) equation, taking into account both FLR as well as other kinetic corrections. Our approach shares some similarity with the one recently adapted by several authors in gyrokinetic theory [Bri92, Bel01]. In the gyrokinetic approach, a dynamical reduction is first employed to transform the Vlasov equation to suitable gyro-centre coordinates, thereby eliminating the fast time-scales associated with gyro-motion. The gyrokinetic

reduction allows FLR effects to be retained under the ordering $k_{\perp}\rho_i \sim 1$, where the Larmor radius of ions is allowed to be of the same order of magnitude as the characteristic length scale of the perturbations. This comes at the expense of limiting consideration to small amplitude fluctuations of the electromagnetic fields about a long-wavelength background equilibrium.

The gyrokinetic derivation of reduced fluid equations including FLR corrections has first been proposed by Brizard [Bri92]. In [Bri92], a set of gyrofluid equations has been obtained in gyrocentre coordinates and then transformed to physical particle space. Using this method, gyroviscous cancellations are automatically accounted for and extensions to arbitrary order in $k_{\perp}\rho_i$ can be systematically carried out. Building on these ideas, an expression for the non-linear gyroviscous force has subsequently been obtained by Belova [Bel01]. The results of [Bel01] are in agreement with the fluid results [Smo98], but are only established for electrostatic perturbations and in slab geometry. To the knowledge of the authors, no derivation of FLR corrections to the pressure tensor is available in the existing literature from a reduced kinetic description and in complex geometry.

The present work fills this gap by giving a detailed derivation of the collisionless gyroviscous tensor in complex geometry. In contrast to [Bri92, Bel01], we will follow the *guiding-centre approach* which allows for arbitrary amplitude perturbations, while requiring $k_{\perp}\rho_i \ll 1$. Strong flows are allowed for and full electromagnetic perturbations are retained. In this way, a guiding-centre kinetic-MHD model for strong flows including diamagnetic (FLR) effects is obtained. The pressure coupling approach is followed; closure of the momentum equation is achieved by expressing the pressure moment in terms of the solution of the guiding-centre equations. It is shown that the higher-order corrections to the Larmor motion that describe the deviation of a particle trajectory from circular motion around the magnetic field play a crucial role in determining the off-diagonal components of the pressure tensor. Our derivation naturally leads to a consistent hybrid kinetic-MHD description of collisionless plasmas in which the exact fluid equations are closed by approximate pressure moments obtained from the solution of a reduced kinetic equation. The proposed model accounts for kinetic effects such as Landau damping, includes an exact treatment of finite orbit-width effects and allows the investigation of strong flows and diamagnetic effects based on the consistent framework of guiding-centre theory. Special consideration is given to the kinetic-MHD equilibrium, and a set of equations suitable for the study of linear dynamics within the proposed model is derived. In particular, the model equations derived here generalise the results of Porcelli [PSK94] and AntonSEN, Lee [Ant82] to include centrifugal effects, and FLR corrections in the fluid contributions to the plasma inertia.

The kinetic-MHD approach taken in the present chapter should be contrasted with gyrokinetic [CZ16], gyrofluid [Sco07] and two-fluid approaches [LL10], that have previously been used to investigate diamagnetic effects. Gyrofluid models are based on taking moments of the gyrokinetic equation, and therefore both gyrokinetic and gyrofluid models intrinsically rely on a splitting between background fields and small-scale fluctuations. Instead, we follow a more classical kinetic-MHD approach which does not require such a splitting. Our approach

is instead restricted to the study of macroscopic, long-wavelength perturbations. Furthermore, our objective is not the derivation of a self-consistent set of fluid equations including FLR corrections for the study of small-scale turbulence. Rather, our goal is to discuss how several important, missing kinetic effects can be added to fluid descriptions of macroscopic instabilities. In particular, the present work concerns the interplay between the fluid and guiding-centre descriptions of plasmas, and shows how they can be combined in a consistent framework.

The present approach may more naturally be viewed as being complementary to a (two-)fluid approach, such as is for example implemented in the XTOR-2F code [LL08, LL10]. Where instead of employing an ad hoc fluid closure, a kinetic closure is achieved, thus taking into account kinetic effects such as wave-particle resonances. As mentioned at the beginning of this introduction, a kinetic closure for the pressure is required at weak collisionality. Our work presents a model including both fluid effects and a reduced kinetic equation. It is explicitly shown that the kinetic description is consistent with two-fluid effects such as diamagnetic drifts. Our equations do however not form a superset of the two-fluid model employed by XTOR-2F. In particular, collisions and related diffusive processes are not included in our collisionless kinetic description. Nevertheless, we do give attention to quasi-neutrality, and corrections to the parallel electric field, which some kinetic-fluid codes and treatments simplify.

This chapter is organized as follows. In section 4.2, a set of guiding-centre equations is developed in a local frame moving with a time-dependent background flow $\mathbf{u}(x, t)$. Important higher-order corrections due to gradients of the background flow are considered in detail in section 4.3.1. Before advancing to the general calculation of the pressure tensor including gyroviscous corrections, the discussion of section 4.3.1 focuses on a special case in simple geometry; this provides an intuitive picture of the meaning and origin of the gyroviscous stresses, and clearly relates them to deviations of particle trajectories from circular motion. Section 4.3.2 then presents our derivation of the general form of the gyroviscous pressure tensor from guiding-centre theory, making use of the results presented in chapter 2. The derivation is shown to recover MacMahon's results [Mac65]. The full set of non-linear kinetic-MHD equations is discussed in Section 4.4.1. The kinetic-MHD equilibrium is considered in section 4.4.2. Equations suitable for the computation of the linear kinetic response are derived in section 4.4.3. Combining the results of the current work, a linear kinetic-MHD model is finally obtained by extending the Frieman-Rotenberg equation [FR60] to include the kinetic equations of section 4.4.5. An argument for the efficiency of the kinetic-MHD pressure closure as opposed to an alternative purely kinetic approach based on current closure is given in Section 4.4.4, where we point out that the present kinetic-MHD approach appears to require only $O(\epsilon_B)$ corrections to the guiding-centre equations to be retained, whereas a fully kinetic approach would require $O(\epsilon_B^2)$ corrections to be considered. Further details on the calculations of the gyroviscous pressure components as well as several mathematical identities that may be of interest in other contexts are provided in the appendix.

4.2 Guiding-centre equations in co-moving frame

For the derivation of the FLR corrections to the pressure tensor, it will be useful to utilize a set of guiding-centre equations based on the velocity variable decomposition $\mathbf{v} = \mathbf{u}(\mathbf{x}, t) + \mathbf{w}$ as developed in chapter 2, where \mathbf{u} is the (leading-order in ϵ_B) fluid velocity and \mathbf{w} can be viewed as a thermal fluctuation. It is interesting to emphasize at this point that the gyroviscous corrections to the pressure are analogous in their origin to the guiding-centre polarization corrections to the density moment and magnetization corrections to the velocity moment, respectively. Therefore, the subtle distinction between the higher-order displacement expressed in guiding-centre \mathcal{Z}_{gc} versus physical particle coordinates \mathcal{Z}_{phy} , which has previously been pointed out and discussed in detail by Brizard [Bri10] for the non-rotating ($\mathbf{u} = 0$) case, is important also in the present context. Even though the important distinction occurs in a second-order (ϵ_B^2) term when expressing the particle position \mathbf{x} in terms of the guiding-centre position \mathbf{X} (in guiding-centre variables) in the form (cp. the first equation of (2.34) in chapter 2)

$$\mathbf{x} = \mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 + \epsilon_B^2 \boldsymbol{\rho}_1 + O(\epsilon_B^3), \quad (4.2)$$

the second-order correction $\boldsymbol{\rho}_1$ will play an important role in calculating the correct *first-order* ϵ_B -correction to the pressure tensor. This perhaps perplexing fact can be understood as follows: The pressure moment is written in terms of the physical particle velocity $\mathbf{v} = \dot{\mathbf{x}} = \dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}_0 + \dot{\boldsymbol{\rho}}_1 + \dots$. While the higher-order guiding-centre displacement $\boldsymbol{\rho}_1$ is ordered at $O(\epsilon_B^2)$, the velocity associated with the second-order displacement is $\dot{\boldsymbol{\rho}}_1 \approx \Omega \partial_\zeta \boldsymbol{\rho}_1$. Because the gyrofrequency $\Omega \sim \epsilon_B^{-1}$ is a formally large term, it causes the velocity contribution due to $\boldsymbol{\rho}_1$ to be of order ϵ_B relative to $\dot{\mathbf{X}} \sim O(1)$; therefore, indicating the ordering in ϵ_B we find that

$$\begin{aligned} \dot{\mathbf{x}} &= \dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}_0 + \epsilon_B \dot{\boldsymbol{\rho}}_1 + O(\epsilon_B^2) \\ &= \mathbf{W} + \epsilon_B \mathbf{V}_{\text{gc}} + \dot{\boldsymbol{\rho}}_0 + \epsilon_B \dot{\boldsymbol{\rho}}_1 + O(\epsilon_B^2), \end{aligned}$$

and $\epsilon_B \dot{\boldsymbol{\rho}}_1$ is of the same order of magnitude as the guiding-centre ∇B - and $\boldsymbol{\kappa}$ -drifts contained in $\epsilon_B \mathbf{V}_{\text{gc}}$ (defined in (4.8), below), and therefore needs to be retained when evaluating the particle velocity including $O(\epsilon_B)$ -corrections.

A detailed derivation of the equations has been presented in chapter 2. In the current section, we will summarize the main results which are of relevance for the computation of the pressure. We use the guiding-centre coordinates $(\mathbf{X}, w, \mu, \zeta)$, of which \mathbf{X} , μ and ζ are summarized in section 2.3.2, and we recall that we retain the leading-order parallel flow component of \mathbf{u} , so that $\mathbf{u} = \mathbf{u}_{||} + \mathbf{u}_E$. We will choose $\mathbf{u}_{||}$ in a way that is suitable for the given application. Correspondingly, in our formulation the parallel velocity variable w is chosen so that $\dot{\mathbf{X}}_{||} = \mathbf{u}_{||}(\mathbf{X}, t) + w \mathbf{b}(\mathbf{X}, t)$. The guiding-centre Lagrangian to the relevant order (neglecting $O(\epsilon_B^2)$

corrections in (2.26)) is here written

$$\begin{aligned}\mathcal{L} &= [q\mathbf{A} + m\mathbf{W}] \cdot d\mathbf{X} + J d\zeta - H dt \\ H &= q\Phi + \frac{1}{2}mW^2 + \mu B.\end{aligned}\tag{4.3}$$

with $\mathbf{W} \equiv w\mathbf{b}(\mathbf{X}, t) + \mathbf{u}(\mathbf{X}, t)$ the leading-order guiding-centre velocity and $J \equiv \frac{m}{q}\mu$ the gyroangle action. The phase-space transformation corresponding to this form of the guiding-centre Lagrangian is generated by the phase-space vector field $G = G_1 + \epsilon_B G_2$ according to $\mathcal{Z}_{\text{gc}}^\alpha = \mathcal{Z}_{\text{phy}}^\alpha + \epsilon_B G_1^\alpha + \epsilon_B^2 (G_2^\alpha + \frac{1}{2}G_1 \cdot dG_1^\alpha)$. Following (2.53), the inverse transformation expressed in guiding-centre coordinates is given by

$$\mathcal{Z}_{\text{phy}}^\alpha = \mathcal{Z}_{\text{gc}}^\alpha - \epsilon_B G_1^\alpha - \epsilon_B^2 \left(G_2^\alpha - \frac{1}{2}G_1 \cdot dG_1^\alpha \right).\tag{4.4}$$

As pointed out above, for the purposes of the present work, the second-order correction to the spatial component of the guiding-centre transformation is required. Comparing (4.2) and (4.4), we can identify $\mathbf{x} = \mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 + \epsilon_B^2 \boldsymbol{\rho}_1$, where (recall equations (2.57) and (2.36))

$$\boldsymbol{\rho}_0 = -G_1^x,\tag{4.5}$$

$$\boldsymbol{\rho}_1 = -G_2^x + \frac{1}{2}G_1 \cdot dG_1^x.\tag{4.6}$$

As in chapter 2, we fix a perpendicular frame $\mathbf{e}_1(\mathbf{X}, t), \mathbf{e}_2(\mathbf{X}, t)$ consisting of unit vectors such that $\mathbf{e}_1, \mathbf{e}_2, \mathbf{b}$ form a right-handed orthonormal basis at each point. We recall from (2.37) that the leading-order displacement is given in terms of the gyroangle-dependent unit vector $\hat{\boldsymbol{\rho}}$ as

$$\boldsymbol{\rho}_0(\mathbf{X}, \mu, \zeta, t) = \sqrt{\frac{2\mu}{mB(\mathbf{X}, t)}} \hat{\boldsymbol{\rho}}(\mathbf{X}, \zeta, t),\tag{4.7}$$

with $\hat{\boldsymbol{\rho}} = \cos(\zeta)\mathbf{e}_1 - \sin(\zeta)\mathbf{e}_2$. The leading order velocity is then along

$$\hat{\mathbf{T}}(\mathbf{X}, \zeta, t) \equiv \partial_\zeta \hat{\boldsymbol{\rho}} = -\sin(\zeta)\mathbf{e}_1(\mathbf{X}, t) - \cos(\zeta)\mathbf{e}_2(\mathbf{X}, t).$$

By (2.36), the higher-order displacement can be written

$$\boldsymbol{\rho}_1 = - (g_\mu \partial_\mu + g_\zeta \partial_\zeta) \boldsymbol{\rho}_0 - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \boldsymbol{\rho}_0 - \left(G_{2,\parallel}^X + \frac{1}{2} \boldsymbol{\rho}_0 \cdot \nabla \mathbf{b} \cdot \boldsymbol{\rho}_0 \right) \mathbf{b}.$$

where (cp. (2.72), (2.73))

$$g_\mu = G_1^\mu - \mu \boldsymbol{\rho}_0 \cdot \nabla \log B,$$

$$g_\zeta = G_1^\zeta + \boldsymbol{\rho}_0 \cdot \mathbf{R},$$

with $\mathbf{R} \equiv (\nabla \mathbf{e}_1) \cdot \mathbf{e}_2$. Explicit expressions for the components g_μ, g_ζ are given by (B.5),(B.6) in appendix B. Correcting an error in [Mad10], the parallel component $G_{2,\parallel}^X$ is found to be given by (2.42). The guiding-centre equations of motion are obtained from variations of the Lagrangian

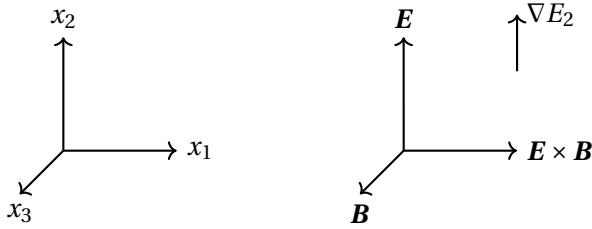


Figure 4.1 – Assumed slab geometry.

(4.3). For the computation of the pressure, the following explicit expressions will be used (see also (2.32) in chapter 2)

$$\begin{aligned}\dot{\mathbf{X}} &= \mathbf{W} + \epsilon_B \mathbf{V}_{\text{gc}} + O(\epsilon_B^2), \\ \mathbf{V}_{\text{gc}} &= \frac{\mu}{qB} \mathbf{b} \times \nabla B + \frac{\mathbf{b}}{\Omega} \times \left(\frac{\partial}{\partial t} + \mathbf{W} \cdot \nabla \right) \mathbf{W}, \\ \dot{\zeta} &= \epsilon_B^{-1} \Omega + \epsilon_B^0 \left[\mathbf{R} \cdot \mathbf{W} + S + \frac{1}{2} \mathbf{b} \cdot (\nabla \times \mathbf{b}) \right] + O(\epsilon_B).\end{aligned}\tag{4.8}$$

Here \$S\$ is the “time-like” analogue of \$\mathbf{R}\$, namely \$S = (\partial_t \mathbf{e}_1) \cdot \mathbf{e}_2\$.

4.3 Pressure in guiding-centre coordinates

4.3.1 Intuitive picture in slab geometry

To provide an intuitive account of the guiding-centre corrections described in the last section, we focus here on the particle motion and implications of these corrections in a simplified magnetic geometry. Following [Kau60], we assume a cartesian slab geometry described by coordinates \$(x_1, x_2, x_3)\$ and constant magnetic field \$\mathbf{B} = B\mathbf{b}\$, pointing in the \$x_3\$-direction. We assume the \$E \times B\$ velocity \$\mathbf{u}\$ to be orientated in direction \$\mathbf{e}_1\$, with amplitude linear in \$x_2\$, such that

$$\mathbf{u}(\mathbf{x}) = u_{1,2} x_2 \mathbf{e}_1,$$

corresponding to an electric field \$\mathbf{E} = E_2(x_2) \mathbf{e}_2 = u_{1,2} B x_2 \mathbf{e}_2\$ (cf. Figure 4.1). Here, we denote by \$u_{1,2} \equiv \partial u_1 / \partial x_2\$ the partial derivative of the \$x_1\$-component of \$\mathbf{u}\$.

Particle trajectory

We first derive the motion of a charged particle with mass m , charge q as predicted by guiding-centre theory. We find

$$\begin{aligned}\boldsymbol{\rho}_1 \cdot \mathbf{b} &= 0, \\ \frac{1}{2\mu} g_\mu &= -\frac{1}{2\Omega} \mathbf{b} \cdot \nabla \times \mathbf{u} + \frac{1}{4\Omega} (\hat{\boldsymbol{\rho}} \hat{\mathbf{L}} + \hat{\mathbf{L}} \hat{\boldsymbol{\rho}}) : \nabla \mathbf{u}, \\ g_\zeta &= -\frac{1}{4\Omega} (\hat{\boldsymbol{\rho}} \hat{\boldsymbol{\rho}} - \hat{\mathbf{L}} \hat{\mathbf{L}}) : \nabla \mathbf{u}\end{aligned}$$

where $\mathbf{b} \cdot \nabla \times \mathbf{u} = -u_{1,2}$, and $\nabla \mathbf{u} = u_{1,2} \mathbf{e}_2 \mathbf{e}_1$. Also $\Omega \equiv qB/m$ is the gyrofrequency. So that

$$\begin{aligned}(\hat{\boldsymbol{\rho}} \hat{\mathbf{L}} + \hat{\mathbf{L}} \hat{\boldsymbol{\rho}}) : \nabla \mathbf{u} &= u_{1,2} [\sin^2(\zeta) - \cos^2(\zeta)], \\ (\hat{\boldsymbol{\rho}} \hat{\boldsymbol{\rho}} - \hat{\mathbf{L}} \hat{\mathbf{L}}) : \nabla \mathbf{u} &= u_{1,2} [-2 \cos(\zeta) \sin(\zeta)].\end{aligned}$$

We then find from equations (4.5)-(4.7): $\boldsymbol{\rho}_0 = \rho_0 (\cos(\zeta) \mathbf{e}_1 - \sin(\zeta) \mathbf{e}_2)$, and

$$\boldsymbol{\rho}_1 = \frac{3\rho_0 u_{1,2}}{4\Omega} \boldsymbol{\rho}_0 + \frac{\rho_0 u_{1,2}}{2\Omega} \sin(\zeta) \mathbf{e}_2.$$

Introducing $a \equiv \rho_0 \left(1 + \frac{3u_{1,2}}{4\Omega}\right)$, we can finally write

$$\boldsymbol{\rho} \approx \boldsymbol{\rho}_0 + \epsilon_B \boldsymbol{\rho}_1 = a \cos(\zeta) \mathbf{e}_1 - a \left(\frac{\Omega - \epsilon_B \frac{1}{2} u_{1,2}}{\Omega} \right) \sin(\zeta) \mathbf{e}_2.$$

Observing also that $\dot{\mathbf{X}} = \mathbf{u}(\mathbf{X})$ with $\mathbf{X}(t=0) = X_1^{(0)} \mathbf{e}_1 + X_2^{(0)} \mathbf{e}_2$ implies $\mathbf{X} = \mathbf{X}(t=0) + u_{1,2} X_2^{(0)} t \mathbf{e}_1$, it follows from $\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}$ that

$$\begin{cases} x_1(t) = X_1^{(0)} + u_{1,2} X_2^{(0)} t + a \cos(\zeta(t)), \\ x_2(t) = X_2^{(0)} - a \left(\frac{\Omega - \epsilon_B \frac{1}{2} u_{1,2}}{\Omega} \right) \sin(\zeta(t)). \end{cases} \quad (4.9)$$

In addition, we note that the guiding-centre equations of motion yield

$$\dot{\zeta} = \Omega + \epsilon_B \frac{1}{2} \mathbf{b} \cdot (\nabla \times \mathbf{u}) = \Omega - \epsilon_B \frac{1}{2} u_{1,2}. \quad (4.10)$$

We can compare these guiding-centre results with the results obtained from the Lorentzian equations of motion [Kau60]. According to [Kau60, equation (20)], the general solution of the resulting Lorentzian equations of motion is given by

$$\begin{cases} x_1(t) = x_1^{(0)} + u_{1,2} x_2^{(0)} t + a \cos \left(\sqrt{\Omega(\Omega - \epsilon_B u_{1,2})} t + \alpha \right), \\ x_2(t) = x_2^{(0)} - a \left(\frac{\Omega - \epsilon_B u_{1,2}}{\Omega} \right)^{1/2} \sin \left(\sqrt{\Omega(\Omega - \epsilon_B u_{1,2})} t + \alpha \right). \end{cases} \quad (4.11)$$

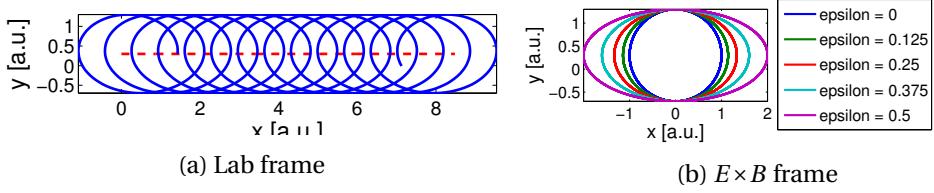


Figure 4.2 – Particle motion in slab geometry with $\nabla E_2 \neq 0$: motion in lab frame with $E \times B$ drift clearly visible (left). Motion in frame co-moving with the guiding-centre (right) clearly shows deviation from circular motion due to ∇E_2 . In this figure, epsilon stands for $\epsilon \equiv |\nabla \mathbf{u}_E|/\Omega_c = |\nabla E_2|/B\Omega_c$.

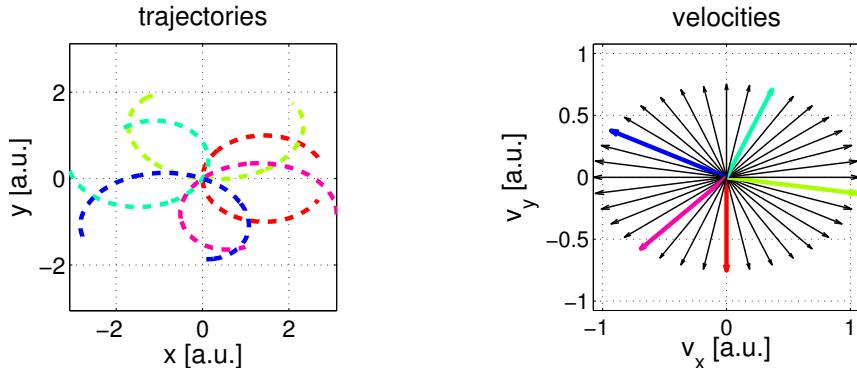


Figure 4.3 – Trajectories intersecting $\mathbf{r} = 0$ and corresponding anisotropic velocity distribution.

where $x_1^{(0)}, x_2^{(0)}, a, \alpha$ are constants to be determined from the initial conditions. By inspection of (4.9) and (4.11), we can immediately identify $\mathbf{x}^{(0)} = \mathbf{X}^{(0)}$, and we see that the leading-order approximation

$$\zeta = \sqrt{\Omega(\Omega - \epsilon_B u_{1,2})} t + \alpha \Rightarrow \dot{\zeta} \approx \Omega - \epsilon_B \frac{1}{2} u_{1,2} = \Omega - \epsilon_B \frac{1}{2} \mathbf{b} \cdot (\nabla \times \mathbf{u}),$$

is consistent with the guiding-centre picture (cp. equation (4.10)).

Effects due to higher-order corrections

The fact that the particle trajectory has an elliptic shape, rather than a circular one, turns out to cause off-diagonal terms in the pressure tensor. To see this, we fix a physical position $\mathbf{r} = 0$ and consider all particle trajectories that intersect \mathbf{r} , as depicted in Figure 4.3. The fact that the trajectories are elliptical rather than circular translates into an anisotropy in the velocity distribution at \mathbf{r} (cf. Figure 4.3 (right)).

The off-diagonal components $\mathbf{\Pi}_{12} \sim \langle v_1 v_2 \rangle$, $\mathbf{\Pi}_{11} - \mathbf{\Pi}_{22} \sim \langle v_1^2 - v_2^2 \rangle$ of the pressure tensor are non-zero because this anisotropy corresponds to correlations between v_1, v_2 as depicted in Figure 4.4, leading to $\mathbf{\Pi} \neq 0$. Here we have denoted the gyro-average by $\langle \dots \rangle \equiv (2\pi)^{-1} \int d\zeta(\dots)$.

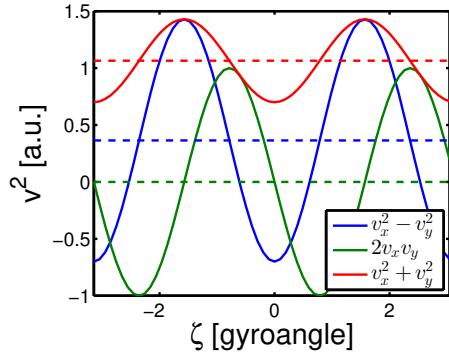


Figure 4.4 – Terms related to off-diagonal components Π_{12} and $\Pi_{11} - \Pi_{22}$ as function of gyroangle ζ at fixed \mathbf{r} (solid), and their gyroaverages (dashed). Here: $v_x \equiv v_1$, $v_y \equiv v_2$.

4.3.2 General expression for pressure

Approach

Let $\bar{\mathbf{u}} = \mathbf{u} + \epsilon_B \mathbf{u}_* + O(\epsilon_B^2)$ denote the (macroscopic) mass fluid velocity, including a leading-order term \mathbf{u} with respect to which the guiding-centre equations are defined. The exact form of \mathbf{u} in our context will be specified later on, for now we only require that the difference between \mathbf{u} and $\bar{\mathbf{u}}$ is of higher order in ϵ_B . We wish to evaluate the pressure tensor, retaining effects related to gradients in \mathbf{u} to first order in ϵ_B . The first step is to recognize that the distinction between \mathbf{u} and $\bar{\mathbf{u}}$ will be unimportant for the evaluation of the pressure. The difference between these velocities, $\epsilon_B \mathbf{u}_*$, is a diamagnetic flow, which itself is of collective (distribution) origin. To see why the contribution due to $\epsilon_B \mathbf{u}_*$ is unimportant, let us first write the pressure tensor \mathbf{P} in guiding-centre coordinates.

$$\begin{aligned}\mathbf{P}(\mathbf{r}) &= \int d^3\mathbf{v} m [\mathbf{v} - \bar{\mathbf{u}}(\mathbf{r})] [\mathbf{v} - \bar{\mathbf{u}}(\mathbf{r})] f(\mathbf{r}, \mathbf{v}) \\ &= \int d^6\mathcal{Z}_{\text{phy}} \delta(\mathbf{x} - \mathbf{r}) m [\mathbf{v} - \bar{\mathbf{u}}(\mathbf{r})] [\mathbf{v} - \bar{\mathbf{u}}(\mathbf{r})] f(\mathcal{Z}_{\text{phy}}) \\ &= \int d^6\mathcal{Z}_{\text{gc}} \delta(\mathbf{X} + \boldsymbol{\rho} - \mathbf{r}) m [\mathbf{v} - \bar{\mathbf{u}}(\mathbf{r})] [\mathbf{v} - \bar{\mathbf{u}}(\mathbf{r})] F(\mathcal{Z}_{\text{gc}}).\end{aligned}$$

We have denoted $\mathcal{Z}_{\text{phy}} = (\mathbf{x}, \mathbf{v})$ the particle phase-space coordinates and \mathcal{Z}_{gc} the guiding-centre phase-space coordinates, and f, F are the distributions functions expressed in particle and guiding-centre coordinates, respectively.

Expanding now $\bar{\mathbf{u}} = \mathbf{u} + \epsilon_B \mathbf{u}_*$, and recognizing that the leading order particle velocity $\mathbf{v} =$

$\mathbf{u} + w\mathbf{b} + \dot{\boldsymbol{\rho}}_0 + O(\epsilon_B)$, we find

$$\begin{aligned}\mathbf{P}(\mathbf{r}) &= \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 - \mathbf{r}) m [\mathbf{v} - \mathbf{u}(\mathbf{r})] [\mathbf{v} - \mathbf{u}(\mathbf{r})] F(\mathcal{Z}_{gc}) \\ &\quad + \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 - \mathbf{r}) m [\mathbf{v} - \mathbf{u}(\mathbf{r})] [-\epsilon_B \mathbf{u}_*(\mathbf{r})] F(\mathcal{Z}_{gc}) + (T) + O(\epsilon_B^2) \\ &= \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 - \mathbf{r}) m [\mathbf{v} - \mathbf{u}(\mathbf{r})] [\mathbf{v} - \mathbf{u}(\mathbf{r})] F(\mathcal{Z}_{gc}) \\ &\quad - \epsilon_B \mathbf{u}_*(\mathbf{r}) \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} - \mathbf{r}) m [w\mathbf{b} + \dot{\boldsymbol{\rho}}_0] F(\mathcal{Z}_{gc}) + (T) + O(\epsilon_B^2).\end{aligned}$$

We denote by (T) the transpose of the second term. Now, we find that the second term in the last equation is of order ϵ_B^2 , because

$$\int d^3 \mathbf{v}_{gc} m [w\mathbf{b} + \dot{\boldsymbol{\rho}}_0] F(\mathcal{Z}_{gc}) = O(\epsilon_B),$$

by our choice of parallel velocity variable w as an (approximate) thermal fluctuation. We conclude that

$$\mathbf{P}(\mathbf{r}) = \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 - \mathbf{r}) m [\mathbf{v} - \mathbf{u}(\mathbf{r})] [\mathbf{v} - \mathbf{u}(\mathbf{r})] F(\mathcal{Z}_{gc}) + O(\epsilon_B^2).$$

The goal of this section is to expand this expression for \mathbf{P} , retaining all corrections at order ϵ_B . In guiding-centre coordinates \mathcal{Z}_{gc} , we find

$$\mathbf{v} = \dot{\mathbf{x}} = \dot{\mathbf{X}} + \dot{\boldsymbol{\rho}}.$$

The guiding centre velocity $\dot{\mathbf{X}}$ can be written in the form (cp. eq. (4.8))

$$\dot{\mathbf{X}} = w\mathbf{b}(\mathbf{X}) + \mathbf{u}(\mathbf{X}) + \epsilon_B \mathbf{V}_{gc} + O(\epsilon_B^2)$$

It is important to note that $\mathbf{u}(\mathbf{r}) \neq \mathbf{u}(\mathbf{X})$. To the required accuracy,

$$\dot{\mathbf{X}} - \mathbf{u}(\mathbf{r}) = \dot{\mathbf{X}} - \mathbf{u}(\mathbf{X}) - \epsilon_B \boldsymbol{\rho}_0 \cdot \nabla \mathbf{u}(\mathbf{X}) + O(\epsilon_B^2),$$

so that

$$\mathbf{P}(\mathbf{r}) = \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 - \mathbf{r}) m [w\mathbf{b} + \dot{\boldsymbol{\rho}}_{gc} + \epsilon_B \mathbf{V}_{gc} - \epsilon_B \boldsymbol{\rho}_0 \cdot \nabla \mathbf{u}]^{\otimes 2} F(\mathcal{Z}_{gc})$$

Expanding $\delta(\mathbf{X} + \epsilon_B \boldsymbol{\rho}_0 - \mathbf{r}) = [1 + \epsilon_B \boldsymbol{\rho}_0 \cdot \nabla] \delta(\mathbf{X} - \mathbf{r}) + O(\epsilon_B^2)$, integrating by parts, and neglecting terms at order $O(\epsilon_B^2)$ and higher throughout, we arrive at

$$\begin{aligned}\mathbf{P}(\mathbf{r}) &= \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} - \mathbf{r}) m [w\mathbf{b} + \dot{\boldsymbol{\rho}} + \epsilon_B \mathbf{V}_{gc} - \epsilon_B \boldsymbol{\rho}_0 \cdot \nabla \mathbf{u}]^{\otimes 2} F \\ &\quad - \nabla \cdot \int d^6 \mathcal{Z}_{gc} \delta(\mathbf{X} - \mathbf{r}) m \boldsymbol{\rho}_0 [w\mathbf{b} + \dot{\boldsymbol{\rho}}] [w\mathbf{b} + \dot{\boldsymbol{\rho}}] F.\end{aligned}$$

We will furthermore write the (canonical) phase-space volume in the form $d^6\mathcal{Z}_{gc} = dX d\mathbf{v}_{gc}$, where the guiding-centre velocity volume element is $d\mathbf{v}_{gc} \equiv m^{-1}B_{||}^* dw d\mu d\zeta$ and $B_{||}^* \equiv \mathbf{b} \cdot \mathbf{B}^*$ is defined in terms of \mathbf{B}^* defined in chapter 2, equation (2.30).

Let us note that both $\dot{\rho} \equiv \dot{\rho}^{(0)} + \epsilon_B \dot{\rho}^{(1)} + O(\epsilon_B^2)$ and ρ_0 are purely oscillatory in gyroangle. Expanding the above expression for \mathbf{P} and taking into account that several terms gyro-average to zero, we obtain (assuming a single ion species)

$$\mathbf{P}(\mathbf{r}) = \left(\int d^3\mathbf{v}_{gc} mw^2 F \right) \mathbf{b}\mathbf{b} + \left(\int d^3\mathbf{v}_{gc} m\dot{\rho}^{(0)} \dot{\rho}^{(0)} F \right) \quad (4.12)$$

$$+ \epsilon_B \int d^3\mathbf{v}_{gc} mw [\mathbf{b}V_{gc} + V_{gc}\mathbf{b}] F \quad (4.13)$$

$$+ \epsilon_B \int d^3\mathbf{v}_{gc} m [\dot{\rho}^{(0)} \dot{\rho}^{(1)} + \dot{\rho}^{(1)} \dot{\rho}^{(0)}] F \quad (4.14)$$

$$- \epsilon_B \int d^3\mathbf{v}_{gc} m [\dot{\rho}^{(0)} (\rho_0 \cdot \nabla \mathbf{u}) + (\rho_0 \cdot \nabla \mathbf{u}) \dot{\rho}^{(0)}] F \quad (4.15)$$

$$- \epsilon_B \nabla \cdot \int d^3\mathbf{v}_{gc} mw \rho_0 [\mathbf{b}\dot{\rho}^{(0)} + \dot{\rho}^{(0)}\mathbf{b}] F. \quad (4.16)$$

We remind the reader that in these expressions F denotes the guiding-centre distribution. We also note that we are not including $O(\epsilon_B^2)$ -terms, so that we may use

$$d^3\mathbf{v}_{gc} \approx m^{-1}B dw d\mu d\zeta \quad (4.17)$$

to calculate the first-order correction terms (4.13)-(4.16).

Leading-order expression

Observing that

$$\dot{\rho}^{(0)} = \Omega \partial_\zeta \rho_0 = \sqrt{2\mu B/m} \hat{\mathbf{L}} \approx \mathbf{v}_\perp,$$

we find that the gyroaverage $m\langle \dot{\rho}^{(0)} \dot{\rho}^{(0)} \rangle = 2\mu B \langle \hat{\mathbf{L}} \hat{\mathbf{L}} \rangle = \mu B (\mathbf{I} - \mathbf{b}\mathbf{b})$, and the CGL form of the pressure tensor is recovered to lowest order

$$\mathbf{P} = P_\perp (\mathbf{I} - \mathbf{b}\mathbf{b}) + P_{||} \mathbf{b}\mathbf{b} + O(\epsilon_B).$$

In these expressions, we define the perpendicular and parallel guiding-centre pressure moments as

$$P_\perp = \int d^3\mathbf{v}_{gc} \mu B F, \quad P_{||} = \int d^3\mathbf{v}_{gc} mw^2 F.$$

For future reference, we also define the guiding-centre parallel fluxes of parallel heat $q_{||}^{\parallel}$ and transverse heat $q_{||}^{\perp}$ as follows:

$$q_{||}^{\parallel} = \int d^3 v_{gc} m w^3 F, \quad q_{||}^{\perp} = \int d^3 v_{gc} w \mu B F.$$

Higher-order correction

Detailed calculations for the higher-order terms have been included in appendix B. Here we only cite the final results. It is convenient to decompose the pressure tensor as follows:

$$\mathbf{P} = \mathbf{P}^{CGL} + \Pi_{\delta}(\mathbf{e}_2 \mathbf{e}_2 - \mathbf{e}_1 \mathbf{e}_1) + \Pi_{12}(\mathbf{e}_1 \mathbf{e}_2 + \mathbf{e}_2 \mathbf{e}_1) + \mathbf{b} \otimes \Pi_{b\perp} + \Pi_{\perp b} \otimes \mathbf{b},$$

where the last two contributions are defined as

$$\Pi_{b\perp} \equiv \mathbf{b} \cdot \mathbf{P} \cdot (\mathbf{I} - \mathbf{b}\mathbf{b}), \quad \Pi_{\perp b} \equiv (\mathbf{I} - \mathbf{b}\mathbf{b}) \cdot \mathbf{P} \cdot \mathbf{b},$$

and we note that by symmetry $\Pi_{b\perp} = \Pi_{\perp b}$. The three contributions $\Pi_{b\perp} = \Pi_{\perp b}$, Π_{δ} , Π_{12} arise from corrections to \mathbf{P} in each of (4.13), (4.14), (4.15), (4.16). These corrections in each equation are listed below:

Correction (I) (eq. (4.13)): From perpendicular Ohm's law, we have $[d\mathbf{b}/dt]_{\perp} \approx [\mathbf{b} \cdot \nabla \mathbf{u}]_{\perp}$. This term therefore gives the following non-zero contributions

$$\begin{aligned} \Pi_{b\perp} &= \mathbf{b} \otimes \left\{ \frac{1}{\Omega} \mathbf{b} \times \left[2P_{||} \mathbf{b} \cdot \nabla \mathbf{u} + q_{||}^{\perp} \nabla B + 2q_{||}^{\parallel} \boldsymbol{\kappa} \right] \right\} \\ \Pi_{\delta} &= 0 \\ \Pi_{12} &= 0. \end{aligned}$$

Correction (II) (eq. (4.14)): We find

$$\begin{aligned} \Pi_{b\perp} &= \mathbf{b} \otimes \left\{ \frac{1}{\Omega} \mathbf{b} \times \left[-q_{||}^{\perp} \boldsymbol{\kappa} - P_{\perp} \mathbf{b} \cdot \nabla \mathbf{u} \right] \right\} \\ \Pi_{\delta} &= \frac{1}{2\Omega} \mathbf{I}_{\gamma} : \left[P_{\perp} \nabla \mathbf{u} + q_{||}^{\perp} \nabla \mathbf{b} \right] \\ \Pi_{12} &= \frac{1}{2\Omega} \mathbf{I}_{\delta} : \left[P_{\perp} \nabla \mathbf{u} + q_{||}^{\perp} \nabla \mathbf{b} \right]. \end{aligned}$$

Correction (III): (eq. (4.15))

$$\begin{aligned}\boldsymbol{\Pi}_{\mathbf{b}\perp} &= \mathbf{b} \otimes \left\{ \frac{1}{\Omega} \mathbf{b} \times [P_{\perp} \nabla \mathbf{u} \cdot \mathbf{b}] \right\} \\ \Pi_{\delta} &= \frac{1}{\Omega} \mathbf{I}_{\gamma} : [P_{\perp} \nabla \mathbf{u}] \\ \Pi_{12} &= \frac{1}{\Omega} \mathbf{I}_{\delta} : [P_{\perp} \nabla \mathbf{u}].\end{aligned}$$

Correction (IV): (eq. (4.16))

$$\begin{aligned}\boldsymbol{\Pi}_{\mathbf{b}\perp} &= \mathbf{b} \otimes \left\{ \frac{1}{\Omega} \mathbf{b} \times [\nabla q_{\parallel}^{\perp} - q_{\parallel}^{\perp} \nabla B - q_{\parallel}^{\perp} \boldsymbol{\kappa}] \right\} \\ \Pi_{\delta} &= \frac{1}{\Omega} \mathbf{I}_{\gamma} : [q_{\parallel}^{\perp} \nabla b] \\ \Pi_{12} &= \frac{1}{\Omega} \mathbf{I}_{\delta} : [q_{\parallel}^{\perp} \nabla b].\end{aligned}$$

Sum of all terms (I-IV): Based on the above expressions, we finally obtain:

$$\begin{aligned}\boldsymbol{\Pi}_{\mathbf{b}\perp} &= \mathbf{b} \otimes \left\{ \frac{1}{\Omega} \mathbf{b} \times [(2P_{\parallel} - P_{\perp}) \mathbf{b} \cdot \nabla \mathbf{u} + P_{\perp} (\nabla \mathbf{u}) \cdot \mathbf{b} \right. \\ &\quad \left. + 2(q_{\parallel}^{\parallel} - q_{\parallel}^{\perp}) \boldsymbol{\kappa} + \nabla q_{\parallel}^{\perp}] \right\}, \\ \Pi_{\delta} &= \frac{1}{2\Omega} \mathbf{I}_{\gamma} : [P_{\perp} \nabla \mathbf{u} + q_{\parallel}^{\perp} \nabla b], \\ \Pi_{12} &= \frac{1}{2\Omega} \mathbf{I}_{\delta} : [P_{\perp} \nabla \mathbf{u} + q_{\parallel}^{\perp} \nabla b].\end{aligned}\tag{4.18}$$

In addition, using $\boldsymbol{\Pi}_{\perp\mathbf{b}} = (\boldsymbol{\Pi}_{\mathbf{b}\perp})^T$, we have for completion,

$$\begin{aligned}\boldsymbol{\Pi}_{\perp\mathbf{b}} &= \left\{ \frac{1}{\Omega} \mathbf{b} \times [(2P_{\parallel} - P_{\perp}) \mathbf{b} \cdot \nabla \mathbf{u} + P_{\perp} (\nabla \mathbf{u}) \cdot \mathbf{b} \right. \\ &\quad \left. + 2(q_{\parallel}^{\parallel} - q_{\parallel}^{\perp}) \boldsymbol{\kappa} + \nabla q_{\parallel}^{\perp}] \right\} \otimes \mathbf{b}\end{aligned}\tag{4.19}$$

This form of the gyroviscous tensor, as derived from guiding-centre theory, recovers MacMahon's result [Mac65, eq. (11)-(13)]. Macmahon's result remains the most general expression for the gyroviscous tensor for collisionless dynamics, and under the present fast-dynamics ordering, with fluid velocity of the order of the ion thermal velocity $\mathbf{u} \sim v_{\text{th},i}$. Note that the more conventionally cited expression for the gyroviscous tensor due to Braginskii [Bra65] is instead based on the assumption of fast dynamics and *high collisionality*, and hence may be appropriate if fluid equations with an ad hoc closure are discussed, but does not pertain in the present context. A thorough discussion and comparison of several alternate orderings for FLR corrected fluid equations has been given by Ramos [Ram05a, Ram05b], and we refer the interested reader to that work for further information. For the present work, it is

interesting to point out that those contributions to the gyroviscous tensor, which are proportional to gradients in \mathbf{u} , are the same for all the various orderings discussed in [Ram05b]. It is this contribution proportional to $\nabla \mathbf{u}$ which is most commonly considered, and which is responsible for the approximate gyroviscous cancellation. For a discussion of the derivation of the gyroviscous force from the gyroviscous tensor, and the approximate gyroviscous cancellation in the momentum equation that is implied by the gyroviscous force, we refer to [CC92, Smo98, Kau60, MT71, HM85, Ram05b, Bra65]. In particular, our derivation shows that the generalized guiding-centre theory developed in the present work is consistent with the inclusion of finite Larmor-radius effects (diamagnetic flows) in the momentum equation.

4.4 A Kinetic-MHD model for strong flows

4.4.1 Nonlinear formulation including diamagnetic effects

Kinetic-MHD model

Based on the above development of guiding-centre theory in terms of the MHD-like velocity \mathbf{u} , and the demonstration that this formulation is consistent with the gyroviscous cancellation, we can now re-write the exact momentum equation, obtained in the limit $m_e \rightarrow 0$: As already pointed out in Section 1.3, to avoid additional complications due to inertia contributions by multiple ion species, we shall assume that the inertia is dominated by single bulk ion species, in the following discussion.¹ The exact momentum equation in this limit is given in terms of the ion flow velocity (“i” denotes the dominant bulk ion species) $\bar{\mathbf{u}} = \mathbf{u} + \mathbf{u}_{*,i} + O(\epsilon_B^2)$:

$$m_i n \left(\frac{\partial \bar{\mathbf{u}}}{\partial t} + \bar{\mathbf{u}} \cdot \nabla \bar{\mathbf{u}} \right) + \nabla \cdot \boldsymbol{\Pi}_i = \mathbf{j} \times \mathbf{B} - \nabla \cdot \mathbf{P}^{\text{CGL}}.$$

While the derivation in the last section has been based on a guiding-centre calculation, the resulting expression for $\boldsymbol{\Pi}_i$ agrees with known results obtained from fluid theory. Therefore following [Aib16], and employing the approximate gyro-viscous cancellation,² we write

$$m_i n \left(\frac{\partial \bar{\mathbf{u}}}{\partial t} + \bar{\mathbf{u}} \cdot \nabla \bar{\mathbf{u}} \right) + \nabla \cdot \boldsymbol{\Pi}_i \approx m_i n \left(\frac{d \mathbf{u}}{dt} + \mathbf{u}_{*,i} \cdot \nabla \mathbf{u}_\perp \right), \quad (4.20)$$

with $\frac{d}{dt} = \partial_t + \mathbf{u} \cdot \nabla$ the convective derivative along \mathbf{u} . Thus, we find the following form of the momentum equation

$$m_i n \left(\frac{d \mathbf{u}}{dt} + \mathbf{u}_{*,i} \cdot \nabla \mathbf{u}_\perp \right) = \mathbf{j} \times \mathbf{B} - \nabla \cdot \mathbf{P}^{\text{CGL}}, \quad (4.21)$$

¹An extension to contributions to the inertia by multiple ion bulk species would be straight-forward if $\mathbf{u}_s \approx \mathbf{u}$ for all bulk ion species “s”, though. The general case is left for future work.

²We employ the approximate form of the gyroviscous force at this point, rather than the exact expression as e.g. derived in [Ram05b], mainly for simplicity of the resulting equations.

where

$$\boldsymbol{u}_{*,i} = \frac{\boldsymbol{b} \times \nabla P_{\perp,i}}{ZenB}$$

is the ion diamagnetic flow. In this last expression, we have allowed for a general diagonal Chew-Goldberger-Low form \mathbf{P}^{CGL} , instead of the isotropic form $\mathbf{P}^{\text{CGL}} = p\mathbf{I}$, which was assumed in the fluid treatment [Aib16]. The density n in (4.21) satisfies the continuity equation

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\boldsymbol{u}) = 0. \quad (4.22)$$

Assuming $\mathbf{E}_{||} \approx 0$, Ohm's law can be written:

$$\mathbf{E} + \boldsymbol{u} \times \mathbf{B} = 0.$$

This form of Ohm's law then leads to the conventional ideal MHD result

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\boldsymbol{u} \times \mathbf{B}). \quad (4.23)$$

Finally, closure is provided by the solution of the collisionless guiding-centre Vlasov equation

$$\partial_t F_{s,\text{gc}} + \dot{\mathbf{X}} \cdot \nabla F_{s,\text{gc}} + \dot{w} \frac{\partial F_{s,\text{gc}}}{\partial w} = 0. \quad (4.24)$$

for each species $s \in \{i, e\}$. We note that

$$\dot{\mathbf{X}} = w \frac{\mathbf{B}^*}{B_{||}^*} + \frac{\mathbf{E}^* \times \mathbf{b}}{B_{||}}, \quad \dot{w} = \frac{q_s}{m_s} \frac{\mathbf{E}^* \cdot \mathbf{B}^*}{BB_{||}^*}, \quad (4.25)$$

with $\mathbf{E}^* \equiv -\nabla H/q - \frac{\partial \mathbf{A}^*}{\partial t}$, $\mathbf{B}^* \equiv \nabla \times \mathbf{A}^*$ and $\mathbf{A}^* \equiv \mathbf{A} + \frac{m}{q}\mathbf{W}$, are derived from the guiding-centre Lagrangian \mathcal{L} (cp. equation (4.3)). The pressure tensor \mathbf{P}^{CGL} is computed at leading order as

$$\begin{aligned} \mathbf{P}^{\text{CGL}} &= \sum_s \mathbf{P}_s^{\text{CGL}}, \\ \mathbf{P}_s^{\text{CGL}} &= P_{s,||} \mathbf{b} \mathbf{b} + P_{s,\perp} (\mathbf{I} - \mathbf{b} \mathbf{b}), \end{aligned} \quad (4.26)$$

where

$$\begin{aligned} P_{s,||} &\equiv \int m_s w^2 F_{s,\text{gc}} d^3 \boldsymbol{v}_{\text{gc}}, \\ P_{s,\perp} &\equiv \int \mu B F_{s,\text{gc}} d^3 \boldsymbol{v}_{\text{gc}}. \end{aligned}$$

The resulting system of equations is written in term of the variables $(n, \boldsymbol{u}, \mathbf{B}, F_s)$, where n is the ion number density, \boldsymbol{u} is the leading-order fluid velocity, \mathbf{B} is the total magnetic field, and F_s denote the guiding-centre distribution functions for each species s . The numerical solution of the non-linear kinetic-MHD system requires the evolution of n according to equation (4.22),

\mathbf{u} is solved for from equation (4.21) and the solution of Ohm's law (4.23) is needed to advance \mathbf{B} . In addition, the guiding-centre distributions are evolved according to the collisionless guiding-centre Vlasov equation (4.24), from which the CGL-contribution to the pressure are obtained according to (4.26). The resulting system of equations can be used to study the non-linear dynamics of low-frequency modes for which diamagnetic effects are important, i.e. when $\omega \sim n\omega_{*,i}$. Both FLR, as well as particle wave interactions (precession resonance) are included in this non-linear model.

Discussion of the non-linear kinetic-MHD model

The kinetic-MHD approach has a long history going back at least to the pioneering work of Rosenbluth and Rostocker [RR59]. In this section we would like to point out a subtlety that make this approach particularly suitable for the study of kinetic effects on macroscopic instabilities. To explain why, we start out by introducing the formal ordering of the terms appearing in the momentum equation (4.21):

$$m_i n \left(\frac{d\mathbf{u}}{dt} + \epsilon_B \mathbf{u}_{*,i} \cdot \nabla \mathbf{u}_\perp \right) = \epsilon_B^{-1} \mathbf{j} \times \mathbf{B} - \nabla \cdot \mathbf{P}^{\text{CGL}}. \quad (4.27)$$

In order to obtain a consistent set of equations, we need to provide closure relations for \mathbf{j} and \mathbf{P}^{CGL} . The crucial point of kinetic-MHD is that it is beneficial *not* to compute the current density directly from the particle distribution function f_s (cp. the classic work by Rosenbluth, Rostocker [RR59, p.25], or the book by Freidberg [Fre14, p.391]). Indeed, if we were to write $\mathbf{j} = q_s \int \mathbf{v} f_s d^3 v$, we would require knowledge of f_s including corrections at order ϵ_B^2 , in order to compute \mathbf{j} to an order that is consistent with the other terms in (4.27). Instead, we express \mathbf{j} in terms of \mathbf{B} using Ampère's law $\nabla \times \mathbf{B} = \mu_0 \mathbf{j}$, and the evolution of \mathbf{B} is governed by equation (4.23).

Thus, the only variable that requires kinetic closure is now the CGL-part of the pressure tensor \mathbf{P}^{CGL} . And comparing with the ordered momentum equation (4.27), we find that the pressure can be consistently computed from knowledge of the distribution function f_s including FLR correction to order ϵ_B . In this way, the kinetic-MHD approach with pressure closure achieves an efficient use of kinetic information, in the sense that the guiding-centre distribution F_s , computed based on guiding-centre theory including all $O(\epsilon_B)$ -corrections is sufficient. Higher-order corrections at $O(\epsilon_B^2)$ are not required in a kinetic-MHD approach, whereas they would be required to compute all terms in the momentum equation (4.27) from a purely kinetic approach.

Previous versions of the kinetic-MHD model, as e.g. discussed in [Fre14, Chap. 9.5], have focused on leading-order effects, for which only information about the zeroth-order distribution function is necessary. In this work, we also consider first-order corrections in the guiding-centre equations which lead to important particle drift effects, and which allow us to consistently include the effects due to the diamagnetic drift velocity in the momentum

equation (4.27). We emphasize that the above non-linear model is based on the MHD form of Ohm's law (also parallel to the field lines), and hence does not include kinetic effects due to $\mathbf{E}_\parallel \neq 0$. In sections 4.3 and 4.4 below, we will instead consider a *linearized* kinetic-MHD model where the effects due to $\mathbf{E}_\parallel \neq 0$ are included.

In closing our discussion of the non-linear kinetic-MHD model, we remark that a similar pressure-coupling approach is for example followed in the numerical code XTOR-K, which is currently under development. XTOR-K is an extension of the two-fluid code XTOR-2F [LL08, LL10]. In XTOR-K, kinetic closure is achieved by computing the particle distribution function f_s , based on direct solution of the Vlasov equation, rather than a set of reduced guiding-centre equations. Numerically, this puts more stringent demands on the maximum allowed time-step size, since the fast gyro-motion of particles needs to be resolved in XTOR-K. In contrast the fastest motion in the guiding-centre description is on the order of the bounce-frequency of particles. The model proposed in this work could thus serve as a computationally less demanding alternative to provide kinetic closure for XTOR-K, in the collisionless limit. Similarly to our non-linear model, XTOR-K also does not determine \mathbf{E}_\parallel from a self-consistent quasi-neutrality relation.

In comparison to non-linear, global gyrokinetic codes [JBA⁺07, GLB⁺11], we expect that the kinetic-MHD system might be less computationally expensive to solve in practice, due to the fact that the considered modes are long-wavelength, in contrast to the small-scale turbulence considered in gyrokinetics. Hence, we would expect such modes to require less spatial resolution. In addition, the kinetic-MHD treatment requires no (non-local) gyroaveraging operations to be carried out, and thus fewer operations per time-step.

While the guiding-centre equations which we used to derive the expression for the gyroviscous tensor (4.18) have naturally led to a non-linear kinetic-MHD model, the main goal of the present work is the derivation of a *linear* model, allowing for a non-zero parallel electric field. The following sections are devoted to the derivation of such a linear model.

4.4.2 Equilibrium

Disclaimer: In the following, we will consistently write fields as a sum of equilibrium and perturbed parts. Perturbed quantities will be denoted with a δ . To ease notation, we will however *not* indicate all equilibrium quantities with a subscript 0, e.g. we shall replace $\mathbf{B}(\mathbf{x}, t) \rightarrow \mathbf{B}(\mathbf{x}) + \delta \mathbf{B}(\mathbf{x}, t)$. Unless otherwise specified, henceforth any quantity without a δ will refer to an equilibrium quantity.

Before deriving equations to study the linear dynamics, we briefly discuss the magnetohydrodynamic equilibrium from a kinetic point of view. We will assume for simplicity that the equilibrium is isotropic. So that the CGL equilibrium pressure tensor becomes a scalar. Note though that the lowest order perturbed pressure is not isotropic. Following [Bri95], we will

assume an axisymmetric equilibrium with nested flux surfaces, so that we can choose flux coordinates (ψ, θ, ϕ) with ψ the poloidal flux variable, θ the poloidal angle and ϕ the toroidal angle. We furthermore assume the equilibrium flow, defining the co-moving frame, to be purely toroidal and given by

$$\mathbf{u}_0 = \Omega(\psi) R^2 \nabla \phi,$$

where $\Omega(\psi)$ is the leading-order toroidal rotation frequency and $R = R(\psi, \theta)$ is the major radius. This is consistent with lowest-order kinetic theory for all collisionalities in an axisymmetric plasma [CBT87]. We remind the reader that the rotation frequency is related to the leading-order potential $\Phi^{(0)}(\psi)$ by $\Omega = -d\Phi^{(0)}/d\psi$ [Bri95].

Guiding-centre equations allowing for strong toroidal flows have been developed in [Bri95], where the following form of the equilibrium guiding-centre Lagrangian was derived:

$$\begin{aligned} \Gamma_0 &= [q\mathbf{A} + m\mathbf{W}] \cdot d\mathbf{X} + \frac{m}{q}\mu d\zeta - H_0 dt, \\ H_0 &= \frac{1}{2}m\mathbf{W}^2 + \mu B + q\Phi^{(0)} + q\Phi^{(1)}, \end{aligned} \quad (4.28)$$

where $\mathbf{W} = \mathbf{u}_0 + w\mathbf{b}$, $\Phi^{(0)}$ is the leading-order electrostatic potential, \mathbf{A} denotes the equilibrium vector potential, and $\mathbf{b} = \mathbf{B}/B$ with $\mathbf{B} = \nabla\phi \times \nabla\psi + F(\psi)\nabla\phi$ the magnetic field. We have allowed for a higher-order electrostatic correction $\Phi^{(1)}$, which is induced by centrifugal effects and which will be computed from quasi-neutrality below. It is imposed that the flux surface average of $\Phi^{(1)}$ vanish. We will denote the total electrostatic potential by $\Phi = \Phi^{(0)} + \Phi^{(1)}$. As guiding-centre variables, we take the guiding-centre position \mathbf{X} , the guiding-centre parallel velocity w measured in a frame moving with \mathbf{u}_0 , the magnetic moment μ suitably defined in the moving frame [Bri95], and the gyroangle ζ . We point out that (4.28) can be viewed as a special case of the non-linear guiding-centre Lagrangian (4.3) when the fields are at equilibrium.

Since the equilibrium is assumed axisymmetric and stationary, the coefficients of the Lagrangian do not explicitly depend on toroidal angle ϕ and time t , nor are they dependent on the gyroangle ζ . Correspondingly, we have following the three constant of equilibrium guiding-centre motion: toroidal momentum $P_\phi \equiv qA_\phi + mW_\phi$, total energy $E = \frac{1}{2}m\mathbf{W}^2 + \mu B + q\Phi$ and magnetic moment μ .

The equilibrium Hamiltonian H_0 corresponds to the energy E as measured in the lab frame, and contains the formally large contribution $q\Phi^{(0)} \sim \epsilon_B^{-1} \frac{1}{2}m\mathbf{W}^2$. It turns out that it will be useful to introduce a new constant of motion, \mathcal{E} , which does not contain such formally large terms. To this end and following Brizard [Bri95], we define a pseudo-radial variable ψ^* via the toroidal momentum P_ϕ :

$$\psi^* \equiv -P_\phi/q = \psi - \frac{m}{q}W_\phi = \psi + \Delta\psi,$$

where $\Delta\psi = -\frac{m}{q}W_\phi$ is the banana width in units of ψ . Clearly, ψ^* is a constant of motion in

axisymmetric equilibria, so that we can define a new constant of motion as the combination $\mathcal{E} \equiv E - q\Phi^{(0)}(\psi^*)$. Expanding

$$\begin{aligned}\Phi^{(0)}(\psi^*) &\approx \Phi^{(0)}(\psi) - \frac{m}{q} W_\phi \frac{d\Phi^{(0)}}{d\psi} \\ &= \Phi^{(0)}(\psi) + \frac{m}{q} \mathbf{W} \cdot \mathbf{u}_0,\end{aligned}$$

and $\mathbf{W} = w\mathbf{b} + \mathbf{u}_0$, then including order $\Delta\psi$ corrections, we can write

$$\begin{aligned}\mathcal{E} &= E - q\Phi^{(0)}(\psi^*) \\ &\approx \left[\frac{1}{2} mw^2 + mw\mathbf{b} \cdot \mathbf{u}_0 + \frac{1}{2} m\mathbf{u}_0^2 + q\Phi(\psi) \right] \\ &\quad - q \left[\Phi^{(0)}(\psi) + \frac{m}{q} (w\mathbf{b} + \mathbf{u}_0) \cdot \mathbf{u}_0 \right] \\ &= \frac{1}{2} mw^2 + \mu B + q\Phi^{(1)} - \frac{1}{2} m\mathbf{u}_0^2.\end{aligned}$$

Thus, \mathcal{E} corresponds to the energy measured in the rotating frame. Under collisionless dynamics, the equilibrium distribution function $F_0 = F_0(P_\phi, E, \mu)$ of a given species must depend on the constants of motion, only. Instead of P_ϕ and E , we can equivalently use the constants of motion ψ^* and \mathcal{E} . For thermal species ($s = i, e$), we shall impose $F_0 = F_0(\psi^*, \mathcal{E})$ to be Maxwellian in \mathcal{E} , so that

$$F_{0,s}(\psi^*, \mathcal{E}) \equiv \frac{N_s(\psi^*)}{(2\pi T_s(\psi^*)/m_s)^{3/2}} \exp(-\mathcal{E}/T_s(\psi^*)) \quad (4.29)$$

with $N_s(\psi)$, $T_s(\psi)$ given profiles. Notice that this equilibrium distribution is isotropic $p_{||} = p_\perp = p$. That the temperature $T_s(\psi)$ is a flux function is consistent with Fokker-Planck equation solutions of arbitrary collisionalities [CBT87]. Thermal gradients cannot build up along magnetic field lines, where particles can freely stream.

We next show that this natural choice for $F_{0,s}$ can be used to recover the ideal MHD equilibrium relations. The physical ion and electron densities are given by

$$\begin{aligned}n_s(\psi, \theta) &= \int d^3\mathbf{v} F_{0,s} \\ &\approx N_s(\psi) \exp\left(\frac{-q_s\Phi^{(1)} + \frac{1}{2}m_sR^2\Omega^2}{T_s(\psi)}\right).\end{aligned}$$

Assuming quasi-neutrality between electrons and ions, $n_e = Zn_i$, and neglecting centrifugal effects for electrons ($m_e \ll m_i$), we obtain

$$q\Phi^{(1)} = \frac{m_i\Omega^2}{2(1 + T_i/T_e)} (R^2 - \langle\langle R^2 \rangle\rangle),$$

where $\langle\langle \dots \rangle\rangle$ denotes flux-surface averaging. Assuming in addition that $T_i = T_e$, we find that

the mass density and pressure $p = p_i + p_e$ are poloidally dependent and of the form

$$\rho(\psi, \theta) \approx \bar{\rho}(\psi) \exp\left(\frac{m_i(R^2 - \langle\langle R^2 \rangle\rangle) \Omega^2}{4T(\psi)}\right),$$

$$p(\psi, \theta) \approx \bar{p}(\psi) \exp\left(\frac{m_i(R^2 - \langle\langle R^2 \rangle\rangle) \Omega^2}{4T(\psi)}\right).$$

We can now write

$$\nabla p = \frac{\partial p}{\partial \psi} \Big|_R \nabla \psi + \frac{\partial p}{\partial R} \Big|_\psi \nabla R,$$

and recall that $\mathbf{u}_0 \cdot \nabla \mathbf{u}_0 = -\Omega^2 R \nabla R$. From the parallel projection of equilibrium force balance, $\mathbf{b} \cdot (\rho \mathbf{u}_0 \cdot \nabla \mathbf{u}_0 + \nabla p) = 0$, we then find the following relationship between ρ, p, T :

$$\rho = \frac{m_i p}{2T} \iff p = 2\rho T / m_i.$$

Note that $\rho = m_i n_i$ and $p = p_i + p_e = 2p_i$, so that the above is equivalent to $2p_i = 2n_i T$. Recalling that $p = 2p_i$ and $n = n_e + n_i = 2n_i$, we arrive at the equation of state for an ideal gas $p = nT$.

Radial force balance leads to the conventional form of the Grad-Shafranov equation for toroidally rotating plasmas

$$\Delta^* \psi + \mu_0 R^2 \frac{\partial p}{\partial \psi} \Big|_R + \frac{1}{2} \frac{dF^2}{d\psi} = 0, \quad \Delta^* \psi \equiv R \nabla \cdot (R^{-1} \nabla \psi),$$

where we recall that the magnetic field is represented as $\mathbf{B} = \nabla \phi \times \nabla \psi + F(\psi) \nabla \phi$ and ϕ is the geometric toroidal angle. We have thus shown that our assumed Maxwellian form of the distribution function is consistent with ideal MHD equilibria. Note that a similar result was summarized in [GW17]. In the present work, we have shown that the derivation of [GW17] is consistent with the guiding-centre equations in a co-moving frame (cf. section 4.2).

4.4.3 Linear Perturbation

Next, we wish to derive suitable equations for the linear kinetic response. The physical parameters to consider are $\epsilon_B \sim \rho_i / L_B$ with ρ_i the thermal ion Larmor radius and L_B the background equilibrium length scale, $\delta \sim \delta B / B$ the amplitude of the perturbations and $k_{\perp} \rho_i$, with k_{\perp} the perpendicular wave vector of the perturbed quantities. The validity of the guiding-centre transformation requires that $\epsilon_B \ll 1$. The study of linear dynamics imposes $\delta \ll 1$. Different forms of the linearized kinetic equation can then be derived assuming $k_{\perp} \rho_i \ll 1$ (guiding-centre approach), or allowing for $k_{\perp} \rho_i \lesssim 1$ (gyrokinetic approach). Clearly, for the study of the linearized dynamics, the gyrokinetic approach is the more general one, motivating us to follow that approach in the following.

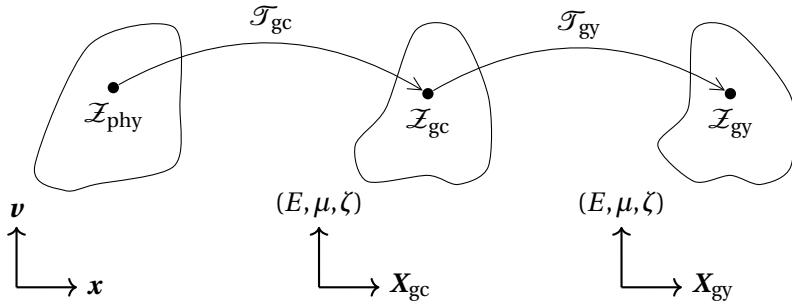


Figure 4.5 – Two-step transformation from physical particle (phy) to gyro-centre (gy) coordinates via guiding-centre (gc) coordinates.

Gyrokinetic transformation

We shall initially allow for $k_{\perp} \rho_i \sim 1$, before discussing the long-wavelength limit $k_{\perp} \rho_i \ll 1$ in section 4.4.3. In section 4.4.3, we show that the results obtained from gyrokinetic theory in the long wavelength limit are consistent with the linearization of the guiding-centre equations of section 4.2.

The gyrokinetic approach, as developed in [BH07, and references therein], is a two-step process: In a first step, a near-identity transformation is carried out to pass from particle coordinates Z_{phy} to guiding-centre coordinates Z_{gc} , defined with respect to the equilibrium fields. In a second step, perturbations are introduced. To obtain a reduced set of kinetic equations taking into account both background inhomogeneities as well as the perturbations, a second transformation to gyrocentre coordinates Z_{gy} is carried out (cf. Figure 4.5).

After the guiding-centre transformation with respect to the equilibrium fields, explained in detail in chapter 2 (see also [Bri95]), the Lagrangian in guiding-centre coordinates now takes the form $\Gamma = \Gamma_0 + \delta \Gamma_1$ with Γ_0 given by equation (4.28), and the perturbed Lagrangian (first order with respect to the perturbation parameter δ),

$$\Gamma_1 = q\delta A(X + \boldsymbol{\rho}) \cdot d[X + \boldsymbol{\rho}] - q\delta\phi(X + \boldsymbol{\rho}) dt.$$

We denote by $\boldsymbol{\rho}$ the guiding-centre displacement, written in guiding-centre coordinates, such that the particle position \mathbf{x} is related to the guiding-centre position \mathbf{X} by $\mathbf{x} = \mathbf{X} + \boldsymbol{\rho}$. Clearly, the perturbed Lagrangian Γ_1 is gyro-angle dependent, because of the gyro-angle dependent terms in $\boldsymbol{\rho}$. To remove this gyroangle-dependence we carry out a near identity transformation to gyrocentre coordinates induced by the phase-space vector field $\bar{\mathbf{G}}_1$. While the final result has previously been obtained in [Bri95], we here supply a detailed discussion of the derivation. Following [BH07], the components of the perturbed Lagrangian in gyrocentre coordinates are related to the components of $\bar{\mathbf{G}}_1$ by

$$\bar{G}_1^a = \{\bar{S}_1, Z^a\}_0 + (\Gamma_{1b} - \bar{\Gamma}_{1b}) J_0^{ab}, \quad (4.30)$$

where \bar{S}_1 is a gauge function, solving the equation

$$\frac{d_0 \bar{S}_1}{dt} = \left(\widetilde{H_1 - \bar{H}_1} \right) - \left(\widetilde{\Gamma_{1a} - \bar{\Gamma}_{1a}} \right) \dot{\mathcal{Z}}_0^a,$$

and $\dot{\mathcal{Z}}_0^a = \{\mathcal{Z}^a, H_0\}$ is the time derivative along an unperturbed orbit. Here, given any physical quantity X , we denote by $\tilde{X} \equiv X - \langle X \rangle$ the gyro-angle oscillatory part of X , and $\langle X \rangle \equiv (2\pi)^{-1} \oint X d\zeta$ is the gyro-average of X . We shall choose a Hamiltonian representation, meaning that only perturbed terms in the Hamiltonian are present, i.e. we choose $\bar{\Gamma}_1 = (\dots) dt$, with vanishing dX -components. It then follows that

$$\frac{d_0 \bar{S}_1}{dt} = q \widetilde{\delta_* \phi} - q [\widetilde{\delta_* A \cdot [\dot{X} + \dot{\rho}]}],$$

where we have introduced $\delta_* \phi \equiv \delta \phi(X + \rho)$, and $\delta_* A \equiv \delta A(X + \rho)$, following the notation in [Bri95].

With this choice of \bar{G}_1 and \bar{S}_1 , the perturbed Hamiltonian in gyrocentre coordinates is given by

$$\begin{aligned} \bar{H}_1 &= \langle H_1 \rangle - \left(\Gamma_{1a} - \bar{\Gamma}_{1a} \right) \dot{\mathcal{Z}}_0^a - \{\bar{S}_1, H_0\}_0 \\ &= q \langle \delta_* \phi \rangle - q \langle \delta_* A \cdot [\dot{X} + \dot{\rho}] \rangle. \end{aligned}$$

where $\langle \dots \rangle$ denotes gyro-averaging.

Collisionless linear gyro-kinetic equation

In Hamiltonian gyrocentre coordinates, the perturbation only affects the Hamiltonian part, so that the Poisson bracket (which is determined by the symplectic part of the Lagrangian) is formally the same as for the guiding-centre equilibrium dynamics $\{\cdot, \cdot\}_0$ with strong toroidal flow \mathbf{u}_0 [Bri95]. We will denote this Poisson bracket simply by $\{\cdot, \cdot\}$ in the following. The collisionless gyrokinetic equation is given by

$$\frac{dF}{dt} = \frac{\partial F}{\partial t} + \{F, H\} = 0.$$

We write the total distribution (in gyrocentre coordinates) in the form $F = F_0 + \delta f$ as a sum of equilibrium and perturbation. Similarly, the Hamiltonian is written in terms of its equilibrium part H_0 and the perturbation \bar{H}_1 . Then, to order $O(\delta^0)$ in δ :

$$\frac{d_0 F_0}{dt} \equiv \frac{\partial F_0}{\partial t} + \{F_0, H_0\} = 0,$$

expressing the fact that the equilibrium distribution function F_0 is a function of the constants of the unperturbed motion, E, P_ϕ, μ (in gyrocentre coordinates). To first order $O(\delta)$, we obtain

$$\frac{d_0 \delta f}{dt} \equiv \frac{\partial \delta f}{\partial t} + \{\delta f, H_0\} = -\{F_0, \bar{H}_1\} = \{\bar{H}_1, F_0\}.$$

Writing now $F_0 = F_0(E, P_\phi, \mu; \sigma)$, where σ is an additional label, allowing to distinguish between co-passing and counter-passing orbits, we find

$$\{\bar{H}_1, F_0\} = \{\bar{H}_1, H_0\} \frac{\partial F_0}{\partial E} + \{\bar{H}_1, P_\phi\} \frac{\partial F_0}{\partial P_\phi} + \{\bar{H}_1, \mu\} \frac{\partial F_0}{\partial \mu}.$$

We can simplify the first term by noting that $d_0 \bar{H}_1 / dt = \partial_t \bar{H}_1 + \{\bar{H}_1, H_0\}$, so that

$$\{\bar{H}_1, H_0\} = \frac{d_0 \bar{H}_1}{dt} - \frac{\partial \bar{H}_1}{\partial t}.$$

Furthermore, it is easy to show from a general property of Lagrangians with symmetries that

$$\{\cdot, P_\phi\} = \frac{\partial}{\partial \phi}, \quad \{\cdot, \mu\} = \frac{q}{m} \frac{\partial}{\partial \zeta},$$

because the toroidal angle ϕ is conjugate to P_ϕ and the gyroangle ζ is conjugate to the gyroaction $J = m\mu/q$. We thus find

$$\frac{d_0 \delta f}{dt} = \left(\frac{d_0 \bar{H}_1}{dt} - \frac{\partial \bar{H}_1}{\partial t} \right) \frac{\partial F_0}{\partial E} + \frac{\partial \bar{H}_1}{\partial \phi} \frac{\partial F_0}{\partial P_\phi},$$

having used that \bar{H}_1 is gyro-angle independent to eliminate the $\partial/\partial\zeta$ term. We may equivalently write this in terms of the non-adiabatic part

$$\delta h \equiv \delta f - \bar{H}_1 \frac{\partial F_0}{\partial E},$$

as

$$\frac{d_0 \delta h}{dt} = \frac{\partial \bar{H}_1}{\partial \phi} \frac{\partial F_0}{\partial P_\phi} - \frac{\partial \bar{H}_1}{\partial t} \frac{\partial F_0}{\partial E}. \quad (4.31)$$

A similar equation for the non-adiabatic part has previously been found by Porcelli et al. [PSK94] in the absence of strong flows. The relative brevity of the present calculation compared to that of [PSK94] is evident.

Pull-back to guiding-centre coordinates

The gyrokinetic distribution, as written in gyrocentre coordinates, has now been shown to be of the form $F_{gy} = F_0 + \delta f$, where $F_0 = F_0(P_\phi, E, \mu; \sigma)$ is a function of the constants of motion. The corresponding distribution in guiding-centre coordinates, is given by $F_{gc} = F_{gy} + \bar{G}_1$.

$dF_{gy} + O(\delta^2)$, or more explicitly

$$F_{gc} = F_0 + \delta f + \bar{G}_1^{P_\phi} \frac{\partial F_0}{\partial P_\phi} + \bar{G}_1^E \frac{\partial F_0}{\partial E} + \bar{G}_1^\mu \frac{\partial F_0}{\partial \mu},$$

with \bar{G}_1^a given by eq. (4.30). The component $\bar{G}_1^{P_\phi}$ is found from

$$\bar{G}_1^{P_\phi} = \bar{\mathbf{G}}_1 \cdot dP_\phi = \mathbf{e}_\phi \cdot (\bar{\mathbf{G}}_1 \cdot d\Gamma_0) = \mathbf{e}_\phi \cdot (\Gamma_1 + d\bar{S}_1) = \Gamma_{1\phi} + \frac{\partial \bar{S}_1}{\partial \phi}.$$

Similarly, we find

$$\bar{G}_1^\mu = \bar{\mathbf{G}}_1 \cdot d\mu = \mathbf{e}_\zeta \cdot (\bar{\mathbf{G}}_1 \cdot d\Gamma_0) = \mathbf{e}_\zeta \cdot (\Gamma_1 + d\bar{S}_1) = \Gamma_{1\zeta} + \frac{\partial \bar{S}_1}{\partial \zeta}.$$

Finally, \bar{G}_1^E is determined from the relation $\bar{H}_1 = H_1 - \bar{G}_1 \cdot dH_0 = H_1 - \bar{G}_1^E$:

$$\bar{G}_1^E = H_1 - \bar{H}_1.$$

Using these expressions and writing $\delta f = \bar{H}_1 \partial F_0 / \partial E + \delta h$, we obtain the general expression (valid to arbitrary order in $k_{\perp} \rho$):

$$F_{gc} = F_0 + \bar{G}_1^{P_\phi} \frac{\partial F_0}{\partial P_\phi} + H_1 \frac{\partial F_0}{\partial E} + \bar{G}_1^\mu \frac{\partial F_0}{\partial \mu} + \delta h. \quad (4.32)$$

The non-adiabatic perturbed distribution function δh is given by equation (4.31).

Note that Helander et al. [HGHM97] have argued that the perturbed distribution δf should be of the form $\delta f = -\delta \mathbf{I} \cdot \frac{\partial F_0}{\partial \mathbf{I}}$, where $\delta \mathbf{I} = (\delta \mathcal{E}, \delta \mathcal{P}_\phi, \delta \mu)$ are the perturbed constants of motion including adiabatic and non-adiabatic contributions. In [HGHM97], it is also claimed that this is true to arbitrary order in ϵ_B . This is in accordance with our results based on modern gyrokinetic theory. In fact, our derivation shows that this is the case even in the presence of a strong electric field at equilibrium (which was assumed to vanish in [HGHM97]).³ Furthermore, from our derivation following a Lie perturbation two-step reduction (to guiding-centre coordinates, then to gyro-centre coordinates), we have an algorithm to compute higher-order corrections in ϵ_B .

³Of course, to obtain an explicit expression for δf from (4.32) to a given order in ϵ_B , one would first need to expand the guiding-centre Lagrangian to that order. What is claimed here is that the abstract form of (4.32) remains the same to arbitrary order in ϵ_B .

Long-wavelength limit $\epsilon_B \sim k_\perp \rho \ll 1$

In the long-wavelength limit, we have $\tilde{F}_{\text{gc}} \ll \langle F_{\text{gc}} \rangle$, where $\tilde{F}_{\text{gc}} \equiv F_{\text{gc}} - \langle F_{\text{gc}} \rangle$. By direct calculation, we find $\langle \bar{G}_1^{\mu} \rangle \approx q\delta A_\phi$, $\langle H_1 \rangle \approx q\delta\phi$, and

$$\langle \bar{G}_1^\mu \rangle \approx \langle q\boldsymbol{\rho} \cdot \nabla \delta \mathbf{A} \cdot \partial_\zeta \boldsymbol{\rho} \rangle = -\mu \frac{\delta B_{||}}{B}.$$

Therefore, in the long-wavelength limit where we are justified to drop oscillatory contributions in gyroangle, one obtains to leading order:

$$F_{\text{gc}} = F_0 + q\delta A_\phi \frac{\partial F_0}{\partial P_\phi} + q\delta\phi \frac{\partial F_0}{\partial E} - \mu \frac{\delta B_{||}}{B} \frac{\partial F_0}{\partial \mu} + \delta h, \quad (4.33)$$

where the non-adiabatic contribution δh is a solution of the drift-kinetic equation

$$\frac{d_0 \delta h}{dt} = \frac{\partial \bar{H}_{\text{dk}}}{\partial \phi} \frac{\partial F_0}{\partial P_\phi} - \frac{\partial \bar{H}_{\text{dk}}}{\partial t} \frac{\partial F_0}{\partial E}. \quad (4.34)$$

with

$$\bar{H}_{\text{dk}} \equiv q\delta\phi - q\delta \mathbf{A} \cdot \dot{\mathbf{X}} + \mu\delta B_{||}. \quad (4.35)$$

These equations allow for strong flows on the order of ion sound velocity, and in this respect generalize the result of Porcelli et al [PSK94], despite the calculation being much more compact.

Consistency with non-linear guiding-centre equations

To show that the linearized kinetic equation derived in the previous section is consistent with the linear limit of the non-linear guiding-centre Lagrangian of section 4.4.1, we briefly consider the linearization of (4.3). To this end, we write \mathcal{L} in the form $\mathcal{L} = \mathcal{L}_0 + \delta\mathcal{L}$, where \mathcal{L}_0 is the equilibrium Lagrangian and $\delta\mathcal{L}$ collects all perturbed terms, and is given by

$$\delta\mathcal{L} = [q\delta \mathbf{A} + m\delta \mathbf{W}] \cdot d\mathbf{X} - [m\delta \mathbf{W} \cdot \mathbf{W}_0 + \mu\delta B_{||} + q\delta\phi] dt.$$

We have introduced the convenient notation $\delta \mathbf{W} \equiv w\delta \mathbf{b} + \delta \mathbf{u}$, which is the perturbed counterpart of $\mathbf{W}_0 = w\mathbf{b} + \mathbf{u}_0$. By a near-identity phase-space transformation to ‘‘Hamiltonian’’ coordinates, we can achieve that the Lagrangian written in these new ‘‘Hamiltonian’’ coordinates takes the form

$$\overline{\delta\mathcal{L}} = \left\{ [q\delta \mathbf{A} + m\delta \mathbf{W}] \cdot \frac{d_0 \mathbf{X}}{dt} - [m\delta \mathbf{W} \cdot \mathbf{W}_0 + \mu\delta B_{||} + q\delta\phi] \right\} dt.$$

We are thus able to move the perturbation from the symplectic part to the Hamiltonian part of the Lagrangian.

Writing $\dot{\mathbf{X}}_0 \equiv d_0 \mathbf{X} / dt$, we can express the perturbed Lagrangian in Hamiltonian coordinates as

$$\begin{aligned}\overline{\delta\mathcal{L}} &= -[q\delta\phi - q\delta\mathbf{A}\cdot\dot{\mathbf{X}}_0 + \mu\delta B_{||}] dt - [m\delta\mathbf{W}\cdot(\dot{\mathbf{X}}_0 - \mathbf{W}_0)] dt \\ &= -[q\delta\phi - q\delta\mathbf{A}\cdot\dot{\mathbf{X}}_0 + \mu\delta B_{||}] dt + O(\epsilon_B^2),\end{aligned}$$

having taken into account that $m\delta\mathbf{W} \sim \epsilon_B q\delta\mathbf{A}\cdot\dot{\mathbf{X}}_0$ and $\dot{\mathbf{X}}_0 = \mathbf{W}_0 + O(\epsilon_B)$ in the last equality. Clearly, we then have $\overline{\delta\mathcal{L}} = -\overline{H}_{dk} dt$ in the long-wavelength limit (4.35), and the kinetic equation (4.34) can be recovered also from guiding-centre theory.

Quasi-neutrality

We now show how to recover the leading-order quasi-neutrality equation first derived by Antonisen, Lee [Ant82]. Note that (4.33) expresses the pull-back in terms of $F_0 = F_0(P_\phi, E, \mu)$. Taking F_0 to be a function of $\psi^* \equiv -P_\phi/q, \mathcal{E}$, instead, we find from $\mathcal{E} = E - q\Phi(\psi^*)$:

$$\begin{aligned}\frac{\partial F_0}{\partial E} \Big|_{P_\phi} &= \frac{\partial F_0}{\partial \mathcal{E}} \Big|_{\psi^*}, \\ q \frac{\partial F_0}{\partial P_\phi} \Big|_E &= -\frac{\partial F_0}{\partial \psi^*} \Big|_E = -\frac{\partial F_0}{\partial \psi^*} \Big|_{\mathcal{E}} - \frac{\partial \mathcal{E}}{\partial \psi^*} \Big|_E \frac{\partial F_0}{\partial \mathcal{E}} \Big|_{\psi^*}.\end{aligned}$$

Using

$$\frac{\partial \mathcal{E}}{\partial \psi^*} \Big|_E = -q \frac{d\Phi^{(0)}}{d\psi},$$

and choosing the ideal MHD gauge $\delta\mathbf{A} = \boldsymbol{\xi} \times \mathbf{B}$, we have

$$\delta A_\phi = [\boldsymbol{\xi} \times \mathbf{B}]_\phi = \xi^\psi,$$

and obtain

$$\delta f = -\xi^\psi \frac{\partial F_0}{\partial \psi^*} \Big|_{\mathcal{E}} + q \left(\delta\phi + \xi^\psi \frac{d\Phi^{(0)}}{d\psi} \right) \frac{\partial F_0}{\partial \mathcal{E}} \Big|_{\psi^*} + \delta h.$$

Writing $\delta\phi = -\xi^\psi \frac{d\Phi^{(0)}}{d\psi} + \delta\phi_L$, where $\delta\phi_L$ is the ‘‘Lagrangian’’ perturbation, we can express this equivalently in the form

$$\delta f = -\xi^\psi \frac{\partial F_0}{\partial \psi^*} \Big|_{\mathcal{E}} + q\delta\phi_L \frac{\partial F_0}{\partial \mathcal{E}} \Big|_{\psi^*} + \delta h.$$

The result of Antonisen and Lee [Ant82] is now recovered by making the approximation $\psi^* \approx \psi$ in the pull-back relation, so that

$$\xi^\psi \frac{\partial F_0}{\partial \psi^*} \approx \boldsymbol{\xi} \cdot \nabla F_0,$$

where $F_0 \approx F_0(\psi, \mathcal{E})$. We can then express the pull-back (4.33) in the form

$$\delta f \approx -\boldsymbol{\xi}_\perp \cdot \nabla F_0 + q\delta\phi_L \frac{\partial F_0}{\partial \mathcal{E}} + \delta h.$$

Integrating over velocity space and summing over species, we find the Lagrangian perturbation

$$0 = \sum_s q_s \delta n_s = \sum_s q_s \int d^3 \mathbf{v}_{gc} \delta h_s + \delta\phi_L \sum_s q_s^2 \int d^3 \mathbf{v}_{gc} \frac{\partial F_{0s}}{\partial \mathcal{E}}, \quad (4.36)$$

where we have assumed quasi-neutrality at equilibrium $\sum_s q_s n_{0s} = 0$. For Maxwellian distributions, $\partial F_0 / \partial \mathcal{E} = -1/T_0$, so that the leading-order quasi-neutrality equation leads to

$$\delta\phi_L = \rho_\Phi^{-1} \sum_s q_s \int d^3 \mathbf{v}_{gc} \delta h_s, \quad (4.37)$$

where $\rho_\Phi \equiv \sum_s q_s^2 n_{0s} / T_{0s}$. As explained below, in section 4.4.5, $\delta\phi_L$ directly determines the parallel electric field (measured in a frame moving with \mathbf{u}_0). Therefore, (4.36) corrects ideal Ohm's law.

Clearly, 4.37 is only the leading-order expression for $\delta\phi_L$. In a current-closing gyrokinetic approach, it is necessary to retain polarization drift effects in the quasi-neutrality equation, in order to correctly reproduce the dispersion relation for drift-waves from gyrokinetic theory [DKOL83]. Let us therefore indicate how such additional FLR corrections may be obtained from the present approach. In taking the long-wavelength limit in section 4.4.3, we have dropped the gyro-angle dependent contribution \tilde{F}_{gc} to the guiding-centre distribution. If we retain this contribution, then we find to the required accuracy

$$\tilde{F}_{gc} \approx q(\boldsymbol{\rho} \cdot \nabla) \xi^\psi \frac{\partial F_0}{\partial P_\phi} + q(\boldsymbol{\rho} \cdot \nabla) \delta\phi \frac{\partial F_0}{\partial E} \approx -(\boldsymbol{\rho} \cdot \nabla) \xi^\psi \frac{\partial F_0}{\partial \psi^*} + q(\boldsymbol{\rho} \cdot \nabla) \delta\phi_L \frac{\partial F_0}{\partial \mathcal{E}},$$

where $F_{gc} = \langle F_{gc} \rangle + \tilde{F}_{gc}$. Correspondingly, instead of the quasi-neutrality relation (4.36), we now obtain an additional term in the general expression for the physical number density of species s ,

$$\begin{aligned} n_s(\mathbf{r}) &= \int d^3 \mathbf{v}_{gc} F_{gc} \delta(\mathbf{X} + \boldsymbol{\rho} - \mathbf{r}) \\ &\approx \int d^3 \mathbf{v}_{gc} \langle F_{gc} \rangle \delta(\mathbf{X} + \boldsymbol{\rho} - \mathbf{r}) + \int d^3 \mathbf{v}_{gc} \tilde{F}_{gc} (\boldsymbol{\rho} \cdot \nabla \delta(\mathbf{X} - \mathbf{r})). \end{aligned}$$

After an integration by parts in the last term on the right, we now have to retain an additional term in the quasi-neutrality relation, which arises due to the $\boldsymbol{\rho} \cdot \nabla \delta\phi_L$ -correction in \tilde{F}_{gc} :

$$\begin{aligned} 0 &= \sum_s q_s \int d^3 \mathbf{v}_{gc} \delta h_s \delta(\mathbf{X} + \boldsymbol{\rho}_0 - \mathbf{r}) + \left(\sum_s q_s^2 \int d^3 \mathbf{v}_{gc} \frac{\partial F_{0s}}{\partial \mathcal{E}} \right) \delta\phi_L \\ &\quad - \nabla \cdot \left[\left(\sum_s q_s^2 \int d^3 \mathbf{v}_{gc} \boldsymbol{\rho}_0 \boldsymbol{\rho}_0 \frac{\partial F_{0s}}{\partial \mathcal{E}} \right) \cdot \nabla \delta\phi_L \right]. \end{aligned}$$

For Maxwellian distribution functions, and neglecting the FLR-correction due to electrons, this can be re-written in the form

$$0 = \sum_s q_s \int d^3 v_{gc} \delta h_s \delta(\mathbf{X} + \boldsymbol{\rho}_0 - \mathbf{r}) - \left(\sum_s \frac{q_s^2 n_{0s}}{T_{0s}} \right) \delta \phi_L \\ + \nabla \cdot \left(\frac{q_i^2 n_{0i}}{T_{0i}} \rho_i^2 \nabla_{\perp} \delta \phi_L \right), \quad (4.38)$$

where $\rho_i \equiv \sqrt{m_i T_{0i}} / (q_i B)$ denotes the thermal gyroradius of the ions. Equation (4.38) is the relevant quasi-neutrality relation including FLR-corrections. As will be seen now, these corrections are irrelevant for a pressure closing kinetic-MHD model. Hence, it will be shown that (4.36) is all that is required to recover e.g. drift waves.

4.4.4 Efficiency of kinetic-MHD pressure closure

As seen from (4.38), the additional FLR corrections to the quasi-neutrality relation appear at order $(k_{\perp} \rho_i)^2 \sim \epsilon_B^2$ in the long-wavelength limit. As we have argued in section 4.4.1, the kinetic-MHD model only requires guiding-centre corrections up to order ϵ_B to consistently provide pressure closure in the ordered momentum equation including diamagnetic effects (4.27). In addition, as explained in the same section 4.4.1, the current density needs to be known to order ϵ_B^2 to consistently close the momentum equation. We recall that $\delta \mathbf{j}$ is obtained from $\mu_0 \delta \mathbf{j} = \nabla \times \delta \mathbf{B}$, and $\delta \mathbf{B}$ is found from Faraday's law

$$-\frac{\partial \delta \mathbf{B}}{\partial t} = \nabla \times \delta \mathbf{E}.$$

Crucially for the present approach, the *electrostatic contribution* $\nabla \delta \phi$ to $\delta \mathbf{E}$ cancels in the above equation. Therefore, the calculation of $\delta \mathbf{B}$ (and $\delta \mathbf{j}$) to the required higher order does also not require knowledge of higher-order corrections to $\delta \phi_L$. Thus, owing to the convenient cancellation of the electric field upon summation of the momentum equation over species (using quasi-neutrality), and thanks to the cancellation of the electrostatic contribution in Faraday's law, the consideration of higher-order corrections in (4.38) does not appear to be necessary within the kinetic-MHD approach followed in the present work.⁴ Consistent with the above observations, the authors of [AB93] have used the leading-order form of the quasi-neutrality relation (4.36) to study the combined effects of diamagnetic flows and a parallel electric field, based on a self-consistent kinetic-MHD model with bounce-averaged kinetic closure for trapped particles. In addition to resolving MHD instabilities, the model of Ref. [AB93] also obtained electrostatic instabilities, and the coupling between them.

⁴We remind the reader that it is assumed that $k_{\perp} \rho_i \sim \epsilon_B \ll 1$ in the derivation of this model. This ordering assumption might potentially break down for certain radially strongly localized modes. For such modes, it may be interesting to compare results obtain with and without the inclusion of the higher-order corrections in (4.38). In general, we however remark that the inclusion of kinetic effects have been observed to tend to broaden the mode structure due to finite orbit-width effects. [PSK94]

The convenient cancellations in a kinetic-MHD approach are thus to be contrasted with a purely kinetic approach, where a consistent treatment would have forced us to *derive from reduced kinetic equations* all the necessary terms in the momentum equation⁵

$$m_s n_s \left(\frac{\partial \bar{\mathbf{u}}_s}{\partial t} + \bar{\mathbf{u}}_s \cdot \nabla \bar{\mathbf{u}}_s \right) = -\nabla \cdot \mathbf{P}_s + \epsilon_B^{-1} q_s n_s (\mathbf{E} + \bar{\mathbf{u}}_s \times \mathbf{B}). \quad (4.39)$$

In this case, we would then clearly have had to retain *second-order corrections* to $\mathbf{E} \approx \mathbf{E}^{(0)} + \epsilon_B \mathbf{E}^{(1)} + \epsilon_B^2 \mathbf{E}^{(2)}$ (and hence higher-order corrections to $\delta\phi_L$), in order to consistently include the required diamagnetic terms (entering at order ϵ_B) in equation (4.39). The inclusion of such effects within a purely guiding-centre kinetic approach would consequently require the derivation (and solution) of the guiding-centre equations including all ϵ_B^2 corrections – a formidable task!

4.4.5 Linear kinetic-MHD model including centrifugal and diamagnetic effects

A non-linear form of the kinetic-MHD model has been presented in section 4.4.1. In this section, we will propose a more refined, *linear* kinetic-MHD model for a two-component plasma consisting of thermal electrons (e) and thermal ions (i), and self-consistently taking into account centrifugal, diamagnetic as well as kinetic effects due to e.g. precession resonance. Furthermore, the model also allows the effects due to a parallel electric field to be taken into account via the solution of a quasi-neutrality equation. We will assume perturbed quantities have the following time- and toroidal dependencies $\delta X \sim \exp(-i\omega t + in\phi)$. In section 4.4.2, we have shown that, under the assumption of Maxwellian distributions for $s = i, e$, the kinetic equilibrium is consistent with MHD equilibria allowing for strong flows. A drift-kinetic equation to determine the linear kinetic response to fluctuating electro-magnetic fields, and including finite-orbit width effects as well as resonances has been presented in section 4.4.3. We next want to combine the kinetic derivation with suitable fluid equations in a linear kinetic-MHD model suitable to study the stability properties of strongly flowing plasmas.

Frieman and Rotenberg [FR60] were the first to consider the hydrodynamic stability of strongly flowing plasmas within ideal MHD. Their derivation of the linearized stability equations is based on a Lagrangian displacement ξ , which measures the difference between fluid elements advected by the perturbed and equilibrium flows. In [FR60], it was found that – in contrast to the stability problem in ideal MHD for a static plasma which is written in terms of a Hermitian operator – the corresponding operator for stationary plasmas with strong flows is non-Hermitian, admitting a much more complex structure of the spectrum. More recently, Aiba [Aib16] has extended the Frieman-Rotenberg approach to include diamagnetic effects in addition to strong flows. The results of [Aib16] are obtained under the assumption of a closure relation for the pressure analogous to the one employed in ideal MHD. In this section, we will combine the kinetic equation with a version of a Frieman-Rotenberg-like equation including

⁵Here written in non-linear form for notational convenience.

diamagnetic effects [Aib16], to obtain a linear MHD model with kinetic closure suitable to study the combined diamagnetic and centrifugal effects.

We now follow largely [Aib16], but we include a Lagrangian perturbation $\delta\phi_L \neq 0$ as well as a kinetic expression for the pressure tensor. We first follow the discussion in Section 1.2.3, where the Frieman-Rotenberg formulation of ideal MHD in terms of a plasma displacement ξ was described. We use perpendicular Ohm's law

$$\mathbf{E}_\perp + \mathbf{u} \times \mathbf{B} = 0,$$

for the lowest-order (MHD-)velocity \mathbf{u} , and impose $\delta\mathbf{B} = \nabla \times (\xi_\perp \times \mathbf{B})$ in terms of a new variable ξ , which will be identified as the displacement. Correspondingly, we choose the gauge $\delta\mathbf{A} = \xi_\perp \times \mathbf{B}$ and we find $\delta\mathbf{E} = -\partial\xi_\perp/\partial t \times \mathbf{B} - \nabla\delta\phi$. In this expression for $\delta\mathbf{E}$, we have allowed for an additional electrostatic contribution $-\nabla\delta\phi$ to $\delta\mathbf{E}$, which was assumed to vanish in the ideal MHD model discussed in Section 1.2.3. With this choice of gauge, we obtain the following expression for the perturbed electric field $\delta\mathbf{E}'$, as measured in a frame co-moving with \mathbf{u}_0 :

$$\delta\mathbf{E}' \equiv \delta\mathbf{E} + \mathbf{u}_0 \times \delta\mathbf{B} = -\nabla\delta\phi_L - \left[\frac{\partial\xi}{\partial t} + \mathbf{u}_0 \cdot \nabla\xi - \xi \cdot \nabla\mathbf{u}_0 \right] \times \mathbf{B}. \quad (4.40)$$

Note in particular that $\delta\mathbf{E}'_\parallel = -\nabla_\parallel\delta\phi_L$ is uniquely determined by $\delta\phi_L$. Thus, with the present choice of gauge, a non-zero $\delta\phi_L$ is directly related to a non-vanishing parallel electric field $\delta\mathbf{E}'_\parallel \neq 0$. This is in stark contrast to a conventional gyrokinetic treatment, which is often based on variables $(\delta\phi, \delta A_\parallel)$, sometimes neglecting δB_\parallel (see recent discussion on comparison between MHD and gyro-kinetic variables in [GZB⁺19]). We note that the precise relation between the present formulation in terms of $(\xi, \delta\phi_L)$ and a formulation in terms of $(\delta\phi, \delta A_\parallel)$ appears to be a challenging open problem. In fact, a comparison of these two approaches is non-trivial even the simplified case where $\delta\phi_L = 0$, as discussed in [GZB⁺19].

Continuing with the derivation of the linearized fluid equations, we can use the relation (4.40) to express the perturbed (Eulerian) MHD-velocity $\delta\mathbf{u}$ by re-writing the perturbed form of perpendicular Ohm's law

$$\delta\mathbf{E}_\perp + \delta\mathbf{u} \times \mathbf{B} + [\mathbf{u}_0 \times \delta\mathbf{B}]_\perp = 0,$$

as follows

$$\delta\mathbf{u}_\perp = \left[\frac{\partial\xi_\perp}{\partial t} + (\mathbf{u}_0 \cdot \nabla)\xi_\perp - (\xi_\perp \cdot \nabla)\mathbf{u}_0 \right]_\perp - \frac{\nabla\delta\phi_L \times \mathbf{B}}{B^2}.$$

This establishes a relation for ξ_\perp , but still leaves ξ_\parallel undefined.⁶ We will *define* the parallel component ξ_\parallel , by requiring without loss of generality that $\delta\mathbf{u}$ be conveniently expressed as

$$\delta\mathbf{u} = \frac{\partial\xi}{\partial t} + (\mathbf{u}_0 \cdot \nabla)\xi - (\xi \cdot \nabla)\mathbf{u}_0 - \frac{\nabla\delta\phi_L \times \mathbf{B}}{B^2}.$$

⁶When centrifugal effects are neglected, the parallel displacement is not required for the kinetic-MHD problem. The potential $\delta\phi$ and the quasi-neutrality equation replaces ξ_\parallel and parallel momentum equation for the MHD problem. When centrifugal effects are important, parallel and perpendicular flows cannot be trivially decoupled.

The corresponding Lagrangian velocity perturbation is then seen to be given by $\delta \mathbf{u}_L = \frac{\partial \boldsymbol{\xi}}{\partial t} + (\mathbf{u}_0 \cdot \nabla) \boldsymbol{\xi} - (\nabla \delta \phi_L \times \mathbf{B}) / B^2$, generalizing the corresponding result in [Aib16, eq. (34)] to include an electrostatic contribution.

Based on the derivation in [Aib16], but allowing for a perturbed pressure tensor with kinetic effects, and parallel electric fields, the following extended Frieman-Rotenberg equation for $\boldsymbol{\xi}$ is derived from this form of the Lagrangian displacement:

$$\begin{aligned} \rho_0 & \left[\frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} + 2(\mathbf{u}_0 \cdot \nabla) \frac{\partial \boldsymbol{\xi}}{\partial t} + (\mathbf{u}_{0,*i} \cdot \nabla) \frac{\partial \boldsymbol{\xi}_\perp}{\partial t} \right] \\ & + \rho_0 \left[\frac{\partial}{\partial t} + (\{\mathbf{u}_0 + \mathbf{u}_{0,*i}\} \cdot \nabla) \right] \left[\frac{-\nabla \delta \phi_L \times \mathbf{B}}{B^2} \right] \\ & + \nabla \cdot \delta \mathbf{P}^{\text{CGL}} - \mathbf{j} \times \delta \mathbf{B} - \delta \mathbf{j} \times \mathbf{B} \\ & = \nabla \otimes [\rho_0 \boldsymbol{\xi} \otimes (\mathbf{u}_0 \cdot \nabla) \mathbf{u}_0 - \rho_0 \mathbf{u}_0 \otimes (\mathbf{u}_0 \cdot \nabla) \boldsymbol{\xi}]. \end{aligned} \quad (4.41)$$

Here, we have defined the ion diamagnetic drift velocity as

$$\rho_0 \mathbf{u}_{0,*i} \equiv \frac{1}{\Omega_{ci}} \mathbf{b} \times \nabla p_{0,i},$$

with $\Omega_{ci} = q_i B / m_i$ the ion cyclotron frequency, and we have retained fluid diamagnetic effects only in the inertia where they are most important. This is justified if the additional diamagnetic force terms [Aib16, e.g. eq. (40)] are dominated by either the fluid force terms, or the kinetic pressure correction arising in $\nabla \cdot \delta \mathbf{P}^{\text{CGL}}$ ⁷. More precisely, we have neglected an additional force term [Aib16, eq. (43)]

$$\rho_0 [\{(\boldsymbol{\xi} \cdot \nabla) \mathbf{u}_{0,*i} - (\mathbf{u}_{0,*i} \cdot \nabla) \boldsymbol{\xi} + \delta \mathbf{u}_{*i}\} \cdot \nabla] \mathbf{u}_{0,\perp},$$

and similarly, we have replaced $\bar{\mathbf{u}}_0 = \mathbf{u}_0 + \mathbf{u}_{0,*i}$ by \mathbf{u}_0 on the right-hand side of equation (4.41).

To close the momentum equation (4.41), we note that $\delta \mathbf{B} = \nabla \times (\boldsymbol{\xi}_\perp \times \mathbf{B})$ and the perturbed current is given by $\delta \mathbf{j} = \nabla \times \delta \mathbf{B} / \mu_0$. Finally, the perturbed CGL-contributions to the pressure are given by

$$\delta \mathbf{P}^{\text{CGL}} = \delta p_{||} \mathbf{b} \mathbf{b} + \delta p_\perp (\mathbf{I} - \mathbf{b} \mathbf{b}), \quad (4.42)$$

with

$$\delta p_{||} \equiv \sum_s \int d^3 \mathbf{v}_{\text{gc}} m_s w^2 \delta f_s, \quad \delta p_\perp \equiv \sum_s \int d^3 \mathbf{v}_{\text{gc}} \mu B \delta f_s. \quad (4.43)$$

Here, the perturbed distribution function

$$\delta f_s = \delta f_{s,a} + \delta h_s \quad (4.44)$$

⁷Incidentally, based on analytic theory in the large aspect-ratio limit, we would indeed expect kinetic corrections to dominate these diamagnetic fluid corrections, as has been shown explicitly for internal 1/1 kink modes [GHH00]

for $s = i, e$ is given in terms of an adiabatic contribution $\delta f_{s,a}$ and a non-adiabatic contribution δh_s . For the adiabatic contribution, we have found

$$\delta f_{s,a} = -\xi^\psi \frac{\partial F_{0,s}}{\partial \psi^*} \Big|_{\mathcal{E}} + q\delta\phi_L \frac{\partial F_{0,s}}{\partial \mathcal{E}} \Big|_{\psi^*} \quad (4.45)$$

To express the non-adiabatic contribution, we first note that [GHH00]

$$\frac{\partial F_0}{\partial E} \Big|_{P_\phi} = \frac{\partial F_0}{\partial \mathcal{E}} \Big|_{\psi^*}, \quad \frac{\partial F_0}{\partial P_\phi} \Big|_E = -\frac{1}{q} \frac{\partial F_0}{\partial \psi^*} \Big|_{\mathcal{E}} - \frac{\partial F_0}{\partial \mathcal{E}} \Big|_{\psi^*} \frac{1}{q} \frac{\partial E}{\partial \psi^*} \Big|_{\mathcal{E}}.$$

The drift-kinetic equation (4.35) can now be expressed in the form

$$\frac{d_0 \delta h_s}{dt} = i(\omega - n\Omega - n\omega_{*,s}) \frac{\partial F_{0,s}}{\partial \mathcal{E}} \overline{H}_{dk}, \quad (4.46)$$

where $\Omega = \Omega(\psi^*) = d\Phi^{(0)}/d\psi|_{\psi=\psi^*}$ is a contribution due to the background flow and the sum of the first two terms $\omega_D \equiv \omega - n\Omega$ represents a Doppler shifted frequency. The third contribution $\omega_{*,s} \equiv \frac{1}{q_s} \frac{\partial F_{0,s}}{\partial \psi^*} / \frac{\partial F_{0,s}}{\partial \mathcal{E}}$ is a diamagnetic rotation frequency. Imposing the ideal MHD gauge $\delta \mathbf{A} = \boldsymbol{\xi}_\perp \times \mathbf{B}$, we can equivalently write the perturbed drift-kinetic Hamiltonian (4.35) in the form

$$\begin{aligned} \overline{H}_{dk} = & q\delta\phi_L + [mw^2 - \mu B] \boldsymbol{\xi}_\perp \cdot \boldsymbol{\kappa} - \mu B (\nabla \cdot \boldsymbol{\xi}_\perp) \\ & - \boldsymbol{\xi}_\perp \cdot [m(\mathbf{u}_0 \cdot \nabla) \mathbf{u}_0 + 2mw(\mathbf{b} \cdot \nabla) \mathbf{u}_0], \end{aligned} \quad (4.47)$$

where we have used several vector identities and $\nabla \cdot \mathbf{u}_0 = 0$. Equations (4.46), (4.47) extend [GHH00, eq. (12)], taking into account centrifugal and Coriolis effects, together with corrections associated with quasi-neutrality ($\delta\phi_L \neq 0$).

The full system is written in terms of $\boldsymbol{\xi}$, δh_s and $\delta\phi_L$. It consists of: the momentum equation (4.41) to determine $\boldsymbol{\xi}$; the closure relations (4.42)-(4.45); the kinetic equation (4.46) for δh_s , which is expressed in terms of the Hamiltonian (4.35); and the quasi-neutrality equation (4.36) for $\delta\phi_L$. Note that we require from the kinetic equation three even moments: perpendicular pressure, parallel pressure and density. The density is required only in the quasi-neutrality relation. Clearly, this model does not require a direct evaluation of the current from kinetic information, which as pointed out above is crucial for the kinetic-MHD approach.

Thus, we have arrived at the linear counterpart to the non-linear model of section 4.4.1. In contrast to the non-linear model, this linear model now includes the effects due to a non-vanishing parallel electric field $\delta E_\parallel \neq 0$. We remind the reader that, even though we have assumed that the plasma inertia is dominated by a single bulk ion species, the kinetic pressure closure includes naturally all collisionless kinetic species ($\delta \mathbf{P} = \sum_s \delta \mathbf{P}_s$).

Linear model with weak flows

In the case of weakly flowing plasmas, centrifugal and Coriolis forces can be neglected. Nevertheless retaining weak flow and $\delta\phi_L$, the momentum equation (4.41) is conveniently written in terms of $(\xi_\perp, \delta h_s)$, because the parallel component $\xi_{||}$ is decoupled from ξ_\perp in this case. Indeed, we then find

$$\rho_0 \left[\left(\frac{\partial}{\partial t} + \mathbf{u}_{0,*i} \cdot \nabla \right) \frac{\partial \xi_\perp}{\partial t} + \left(\frac{\partial}{\partial t} + \mathbf{u}_{0,*i} \cdot \nabla \right) \frac{-\nabla \delta\phi_L \times \mathbf{B}}{B^2} \right] = -[\nabla \cdot \delta \mathbf{P}^{\text{CGL}}]_\perp + [\mathbf{j} \times \delta \mathbf{B}]_\perp + \delta \mathbf{j} \times \mathbf{B}. \quad (4.48)$$

with $\delta \mathbf{j} = \nabla \times \delta \mathbf{B}/\mu_0$ and $\delta \mathbf{B} = \nabla \times (\xi_\perp \times \mathbf{B})$, and $\delta \mathbf{P}^{\text{CGL}}$ is given by (4.42)-(4.45) as a function of $\xi_\perp, \delta h_s, \delta\phi_L$. The latter variables are determined from the solution of the kinetic equation (without centrifugal effects)

$$\frac{d_0 \delta h_s}{dt} = i(\omega - n\Omega - n\omega_{*,s}) \frac{\partial F_{0,s}}{\partial \mathcal{E}} \bar{H}_{dk}, \quad (4.49)$$

where $\omega_{*,s} \equiv \frac{1}{q_s} \frac{\partial F_{0,s}}{\partial \psi^*} / \frac{\partial F_{0,s}}{\partial \mathcal{E}}$. The perturbed drift-kinetic Hamiltonian is given by

$$\bar{H}_{dk} = q\delta\phi_L + [mw^2 - \mu B]\xi_\perp \cdot \boldsymbol{\kappa} - \mu B(\nabla \cdot \xi_\perp). \quad (4.50)$$

The quasi-neutrality equation reads

$$\left(\sum_s \frac{q_s^2 n_{0s}}{T_{0s}} \right) \delta\phi_L = \sum_s q_s \int d^3 \mathbf{v}_{gc} \delta h_s, \quad (4.51)$$

These equations are identical to those of reference [GHH00], except that in [GHH00] the parallel electric field was neglected ($\delta\phi_L = 0$). In the system of equations of [GHH00], it was found that weak flows do affect the kinetic corrections to MHD if the plasma rotation is sheared.

The system of equations (4.48)-(4.51) is expressed in terms of $\xi_\perp, \delta h_s$ and $\delta\phi_L$ and can be used to study electrostatic effects on global MHD instabilities [AB93]. A very similar model has previously been proposed in [Ant82, AB93], but in contrast to [Ant82] the derivation (from gyrokinetics) adopted in this work makes no restriction on the allowed bounce-frequencies and includes full finite-orbit width effects. In addition [Ant82, AB93] neglected diamagnetic corrections in the inertia, and by neglecting finite orbit width effects in solving for the non-adiabatic part δh , the kinetic inertia effects described in [GHH00] were also not included in [Ant82, AB93].

Linear model with strong flows, without parallel electric field and diamagnetic corrections in the inertia

In the presence of strong equilibrium flows, the equations for ξ_{\perp} and $\xi_{||}$ are no longer decoupled and the parallel dynamics can have an important effect on the stability analysis. Neglecting diamagnetic and parallel electric field ($\delta\phi_L = 0$), but retaining kinetic wave-particle interaction, we arrive at the following model: The kinetic-MHD momentum equation (4.41) now becomes a Frieman-Rotenberg-like equation with kinetic closure

$$\begin{aligned} \rho_0 \left[\frac{\partial^2 \boldsymbol{\xi}}{\partial t^2} + 2(\mathbf{u}_0 \cdot \nabla) \frac{\partial \boldsymbol{\xi}}{\partial t} \right] = & \\ - \nabla \cdot \delta \mathbf{P}^{\text{CGL}} + \mathbf{j} \times \delta \mathbf{B} + \delta \mathbf{j} \times \mathbf{B} & \\ + \nabla \otimes [\rho_0 \boldsymbol{\xi} \otimes (\mathbf{u}_0 \cdot \nabla) \mathbf{u}_0 - \rho_0 \mathbf{u}_0 \otimes (\mathbf{u}_0 \cdot \nabla) \boldsymbol{\xi}] . & \end{aligned} \quad (4.52)$$

with $\delta \mathbf{j} = \nabla \times \delta \mathbf{B} / \mu_0$ and $\delta \mathbf{B} = \nabla \times (\boldsymbol{\xi} \times \mathbf{B})$, and $\delta \mathbf{P}^{\text{CGL}}$ is given by (4.42)-(4.45) as a function of $\boldsymbol{\xi}, \delta h_s$ where we set $\delta\phi_L = 0$. The δh_s are determined from the solution of the kinetic equation (4.46). In this limit, the model is expressed in terms of $\boldsymbol{\xi}$ and δh_s , and it allows to study the combined kinetic and centrifugal effects on MHD modes. Equation (4.52) is an exact linearisation of the simplified non-linear model of (4.21) if $\mathbf{u}_{*,i}$ is neglected on the LHS of (4.21). Retaining $\mathbf{u}_{*,i}$ effects on the LHS of (4.52) would render this linear model as the linearisation of equation (4.21)

4.5 Summary

A kinetic-MHD model has been derived from a consistent set of guiding-centre equations. The proposed kinetic-MHD model allows for strong flows and includes centrifugal as well as FLR effects related to diamagnetic flows. Closure of the momentum equation is obtained from a solution of the guiding-centre equations, thus accounting for finite orbit-width effects and particle-wave interactions such as precession resonance.

For the first time, the full expression for the gyroviscous contribution to the pressure tensor has been obtained from kinetic theory, without resorting to any simplifying assumptions on the background geometry. Our detailed calculations demonstrate that the formulation of guiding-centre theory proposed in this work can be used to study the influence of kinetic effects on global MHD modes on time-scales $\omega_D \sim n\omega_{*,i}$ when the diamagnetic drift frequency $\omega_{*,i} = -p'_i/(Zen_i B_0 r)$ is of the order of the Doppler shifted rotation frequency $\omega_D = \omega - n\Omega$, with n the toroidal mode number.

We have explicitly shown that for a two-component plasma ($s = i, e$), and under the assumption of Maxwellian equilibrium distribution functions, the usual ideal MHD equilibrium equations are obtained from guiding-centre theory.

Linearized equations describing the evolution of long-wavelength ($k_{\perp}\rho_i \ll 1$) global (collis-

sionless) kinetic-MHD modes have been obtained from a kinetic extension of the Frieman-Rotenberg equations, allowing for centrifugal effects, diamagnetic fluid drift, as well as a kinetic closure. The model naturally includes a drift-kinetic form of the quasi-neutrality equation, and allows the effects of a parallel electric field on global MHD modes to be studied self-consistently.

As discussed in Section 4.4.4, the kinetic-MHD approach with pressure-closure relies on several convenient cancellations of high-order terms. Kinetic information is only used to provide closure for terms such as the pressure which are already formally small in the momentum equation (in terms a guiding-centre ordering). In particular, the current is *not* computed directly from the kinetic distribution, but is instead determined from the field equations. Based on this, it is argued that the resulting kinetic-MHD model with pressure-closure achieves to consistently include diamagnetic effects based on guiding-centre equations expanded to order ϵ_B , whereas an equivalent kinetic model based on current-closure would have required guiding-centre corrections to order ϵ_B^2 to be retained.

In closing, we mention that while the results of this work have focused on the collisionless kinetic effects due to thermal species in a two-component plasma, an additional population of supra-thermal ions (or electrons) can be added trivially in the pressure coupling scheme, provided that their contribution to the inertia can be neglected and that either the suprathermal distributions are isotropic, or that their contributions to pressure gradients are weak. An extension of the results presented in this chapter to the case of a multi-species plasma, in which multiple ion species contribute significantly to the inertia, is left for future work.

The implications of the centrifugal, diamagnetic and electrostatic effects discussed in the present work can be the subject of future investigations. The proposed model can also serve as a starting point for analytical work, and is suitable for the derivation of a dispersion relation taking into account the centrifugal effects on the kinetic response of thermal ions.

5 Final remarks and perspectives

5.1 Summary and conclusions

As pointed out in the first chapter, the contributions presented in this thesis have been motivated in large part due to the need to consider various physical effects to assess the stability of tokamak plasmas to the 1/1 internal kink instability. Due to the particular importance of kinetic particle-wave interactions, kinetic-MHD models have been widely used to study and analyse the stability of this mode. More recently, the importance of the stabilizing effect of toroidal rotation has been revealed within an ideal MHD approach, both by analytical and numerical work. Since strong toroidal rotation is usually achieved in devices with plasma heating by (unbalanced) neutral beam injection, strong toroidal rotation frequencies are often obtained in conjunction with a kinetic species of suprathermal particles. Under these conditions, a complete analysis of the stability, including kinetic and centrifugal effects as well as their interaction, should be based upon a consistent kinetic-MHD model, which takes into account centrifugal effects in the kinetic equations and the fluid equations. However, past work in this context has focused on either kinetic closure in the absence of rotation, or centrifugal effects within a fluid model, and in particular within ideal MHD. In contrast to non-rotating plasmas, the perpendicular and parallel dynamics in rotating plasmas is strongly coupled. Fluid models do not provide an accurate description of the parallel dynamics of high-temperature plasmas. A realistic model must take into account the long mean-free path of particles in the direction parallel to the field-lines, and hence a consistent kinetic closure is expected to be of particular importance for strongly rotating plasmas. Existing kinetic-MHD models have thus far largely neglected centrifugal effects. Therefore, a gap has remained in the research literature, calling for the need to develop a novel, consistent linear kinetic-MHD model, which includes centrifugal as well as kinetic effects.

This thesis attempts to fill this gap, by providing a detailed derivation of such a linear kinetic-MHD model with a kinetic closure based on guiding-centre theory: After a short review of the main concepts of fusion research and ideal MHD, and the basic formulation of kinetic-MHD models, several aspects of the 1/1 internal kink instability are reviewed in chapter 1.

Chapter 5. Final remarks and perspectives

This first chapter is intended to provide the background and motivation for several original contributions in this thesis, which are presented in chapters 2-4.

Since guiding-centre theory is central to the approach to kinetic-MHD taken in the present thesis, a thorough discussion of the Lagrangian formulation of guiding-centre theory is given in chapter 2. After an introduction including a short overview and several elements of the historical development of guiding-centre theory, a formulation of the guiding-centre equations of motion including centrifugal effects, and with respect to guiding-centre coordinates which are particularly suitable for the development of a kinetic-MHD model, are presented. In contrast to similar presentations in the literature, a particular concern, which runs throughout this thesis, is the precise meaning of the guiding-centre coordinates and their relation to the coordinates of the particle that this guiding-centre describes, including higher-order corrections which are often neglected. In particular, the often overlooked subtlety arising due to this distinction between guiding-centre and particle coordinates is pointed out in appendix A, where an original discussion of the Baños drift parallel to the field lines in a simple background magnetic field is given, based on the results presented in chapter 2. A comparison of the physical particle motion (governed by the Lorentz equations) and the guiding-centre motion is given, suggesting that a careful distinction should be made between the parallel particle velocity and the parallel velocity of the corresponding guiding-centre (Baños drift). The consequences of this higher-order correction parallel to the field lines is investigated further in the context of full- f slowing-down simulations in chapter 3. The main contribution of chapter 2 is the detailed derivation of the guiding-centre Lagrangian based on Lie perturbation methods in section 2.4. While very similar expressions for the guiding-centre Lagrangian have been derived by several authors in the past, a very limited amount of detail (if any) of the mathematical derivation is usually provided. This does not only make such derivations difficult to verify, but also represents an obstacle for students and researchers wishing to learn more about guiding-centre theory. While the reader is assumed to be familiar with several mathematical prerequisites, a thorough discussion of the steps required in the derivation is provided. It turns out that the derivation proved critically important for chapter 4, where certain errors in the literature were discovered, enabling the final correct kinetic-MHD model to be derived.

Chapter 3 presents a concrete application of the higher-order corrections derived in chapter 2. The application is that of slowing-down simulations of NBI injection in a MAST-like device. For this problem, higher-order guiding-centre equations have been implemented in the VENUS-LEVIS guiding-centre following code, as well as an algorithm to switch between particle and guiding-centre coordinates consistent with the theory presented in chapter 2. While higher-order corrections to the guiding-centre Lagrangian are not found to be important for this particular application, inclusion of the Baños drift is found to significantly affect the expected resonances of NBI injected particles in the presence of RMP coils, as well as the NBI driven current in a MAST-like equilibrium.

Finally, chapter 4 discusses the derivation of a novel kinetic-MHD model with pressure closure

based on the guiding-centre theory developed in chapter 2. Besides the derivation of suitable equations for the evolution of the distribution function of guiding-centres of the kinetic species, the other main issue is the evaluation of the pressure tensor from our knowledge about the distribution of guiding-centres. Due to the complicated, spatially non-local nature of the guiding-centre coordinates (all complicated functions of both the physical particle position \mathbf{x} and particle velocity \mathbf{v}), this turns out to be a non-trivial task. While leading-order terms are readily computed and yield the conventional Chew-Goldberg-Lew contributions to the pressure tensor, which are conventionally used in existing models, higher-order FLR corrections are closely related to diamagnetic effects, and the so-called gyroviscous cancellation. These higher-order FLR corrections are known to lead to off-diagonal, gyroviscous contributions to the pressure tensor.

Chapter 4 begins with a short survey of past work on this problem. Before advancing to the general derivation of the pressure tensor including gyroviscous components, simplified geometry is considered in section 4.3.1 to provide an intuitive picture of the origin and meaning of these gyroviscous contributions to the pressure tensor. It is observed by explicit calculation both based on the Lorentz equations and from guiding-centre theory, that small (second-order) corrections to the circular motion of the particles' gyration around the field lines result in elliptically shaped trajectories, which directly causes off-diagonal components in the pressure.

After this introductory discussion, the full expression for the pressure tensor including higher-order corrections is presented in section 4.3.2. This represents the first time that these gyroviscous corrections to the pressure have been obtained directly from guiding-centre theory allowing for a sonic-ordered $\mathbf{E} \times \mathbf{B}$ velocity, and without making any simplifying assumptions on the background geometry. As a direct application of this calculation, a non-linear formulation of a kinetic-MHD model with pressure closure based on guiding-centre theory is obtained in section 4.4.1.

A discussion of the kinetic-MHD equilibrium, and how the conventional rotating Grad-Shafranov equation can be obtained to leading order from our guiding-centre formulation is presented in 4.4.2. This discussion is important from a practical point of view, because it provides a form of the equilibrium distributions, which would need to be considered when solving the linearized kinetic-MHD equations numerically and which is consistent with the background equilibrium which would, at least in a first step, be computed by a Grad-Shafranov solver based on ideal MHD.

In section 4.4.3, linearized kinetic equations are developed. These equations are required for the solution of the perturbed kinetic distribution function δf in the linearized kinetic-MHD model. Our discussion is based on the long-wavelength limit of gyrokinetic theory. An original and concise derivation of the results of Porcelli et al. [PSK94] and an extension of these results to the rotating case is provided, based on the gyrokinetic formalism. The adiabatic contribution to δf , initially found by Porcelli et al. [PSK94], is explicitly identified as that contribution to δf which arises due to the pull-back from gyrocentre to guiding-centre

coordinates. A discussion is also given towards the appropriate quasi-neutrality equation that can be used in our kinetic treatment, if the effects of a non-zero parallel electric field $\delta E_{||} \neq 0$ are to be taken into account.

A main benefit of a kinetic-MHD model over a purely kinetic model is discovered. It is argued that – under the assumed approximations – a kinetic-MHD approach makes more efficient use of the kinetic information (in terms of an expansion in Larmor radius), in comparison to a purely kinetic approach.

Finally, these kinetic equations are combined with a fluid model formulated in terms of a plasma displacement in section 4.4.5. The main original results of this thesis are combined, resulting in a linear collisionless kinetic-MHD model, including centrifugal effects, diamagnetic effects, kinetic effects and providing a quasi-neutrality equation allowing for the consideration of a non-zero perturbed electric field. This model is derived based on a consistent set of guiding-centre equations. Kinetic-MHD models that can be obtained in appropriate limits, such as the limit of weak background rotation and with vanishing parallel electric field, are also presented. The derived kinetic-MHD model allows for strong toroidal rotation on the order of the ion sound velocity and can in particular be used to study the combined effects of rotation and kinetic particle-wave interactions on the 1/1 internal kink. Other modes such as GAMs and Kelvin Helmholtz instabilities in a toroidally rotating plasma [CWGW11, WGC13, WG16, GW17] may also be considered with this kinetic-MHD model.

5.2 Future perspectives

5.2.1 Analytic considerations

A natural direction for future work would be to extend the analytic work of Graves et al. [Gra13] to include centrifugal effects in addition to the effects of sheared flow. To this end, the expressions for the bounce frequencies and toroidal drift precession, as e.g. derived analytically in [Gra13], should be extended to include the additional guiding-centre drifts discussed in chapter 2. We remark that the analytical expressions which are commonly used in the non-rotating case, rely on the conservation of energy

$$E = \frac{1}{2}mw^2 + \mu B,$$

to obtain the expression $w = \pm\sqrt{2(E - \mu B)/m}$ for the parallel guiding-centre velocity w . As both the guiding-centre energy E and the magnetic moment μ are constant of motion, and the poloidal flux along a particle trajectory $\tau \mapsto \psi(\tau) = \langle\psi\rangle + \epsilon_B \Delta\psi(\langle\psi\rangle, \theta(\tau))$ remains approximately constant. This allows the velocity variable to be directly related to the poloidal angle $\tau \mapsto \theta(\tau)$ along the particle trajectory, in addition to the constants of motion E , μ and the bounce-averaged poloidal flux $\langle\psi\rangle$ [Gra13]. This fact allows the average drifts to be computed analytically by suitable averaging in the poloidal angle for large aspect-ratio equilibria.

In the rotating case, the corresponding guiding-centre energy is instead given by the more complicated expression

$$E = q\Phi + \frac{1}{2}m[\mathbf{u} + w\mathbf{b}]^2 + \mu B.$$

This expression apparently exhibits a more complicated dependence of w on the poloidal angle, as w is the solution of a general quadratic equation of the form $w^2 + \alpha w + \beta = 0$, where $\alpha = 2\mathbf{b} \cdot \mathbf{u} \neq 0$. That this complication is only apparent can be inferred from the discussion of section 4.4.2, from which it follows that the “energy in the rotating frame” \mathcal{E} , which we may define as

$$\mathcal{E} \equiv E - q\Phi^{(0)}(\langle\psi\rangle),$$

can be approximately expressed, upon taking into account the guiding-centre's radial drifting motion, in the form

$$\mathcal{E} \approx \frac{1}{2}mw^2 + \mu B + q\Phi^{(1)} - \frac{1}{2}m\mathbf{u}_0^2.$$

This allows the much simpler expression $w = \pm\sqrt{2(\mathcal{E} - \mu B - q\Phi^*)/m}$ for w , where we have introduced the effective potential in the rotating frame $q\Phi^* \equiv q\Phi^{(1)} - \frac{1}{2}m\mathbf{u}_0^2$. A formulation in terms of the constant of motion \mathcal{E} , μ , and $\langle\psi\rangle$, in addition to the poloidal angle θ could therefore be suitable to form the basis of an extension of the analytic results in [Gra13], to include centrifugal effects. Of course, other modes of interest may also be investigated, e.g. Kelvin Helmholtz and resistive wall modes, etc.

5.2.2 Numerical implementation

Another direction for future work is the numerical implementation of the proposed kinetic-MHD model, perhaps neglecting diamagnetic effects and assuming vanishing parallel electric field $\delta E_{||} = 0$ in a first step. Schematically, such a linear kinetic-MHD stability code would have the structure depicted in figure 5.1.

Such an implementation needs two basic ingredients: A MHD solver, which discretizes the fluid equations for the macroscopic fields such as the plasma displacement, and a kinetic solver, which discretizes the kinetic guiding-centre equations to solve for the perturbed distribution function δf . In addition, a coupling code must be implemented to pass information about the electromagnetic fields (e.g. the perturbed electromagnetic potentials δA , $\delta\phi$) to the kinetic solver, and in return outputs fields necessary for the kinetic closure (the perturbed pressure) to the MHD solver.

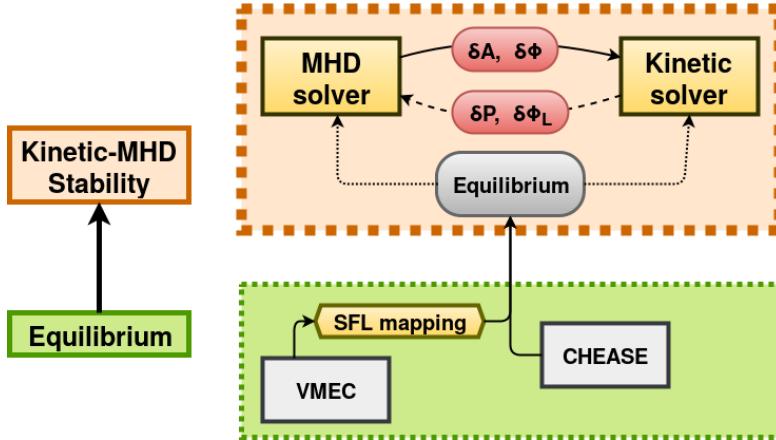


Figure 5.1 – Flow chart of an implementation of the linear kinetic-MHD model.

MHD stability code

A first “beta”-version of an adaptable MHD stability code (tentatively named VENUS-MHD) have been undertaken within the framework of this thesis. VENUS-MHD is written in Fortran. The representation of the MHD equilibrium is taken from the guiding-centre pushing code VENUS-LEVIS [PCGM14,Pfe15]. The first ingredient in this stability code is a program SATIRE2sf1.x to carry out the mapping of the equilibrium produced by the VMEC equilibrium solver (including toroidal rotation [CH87,CBF⁺15]) to straight-field line coordinates. To this end, equation (1.9) of chapter 1 is discretized and solved by a Fourier decomposition on each flux surface. In addition, the VMEC radial variable $s = \Phi/\Phi_N$ (normalized toroidal flux) is remapped to $\rho_{\text{tor}} \equiv \sqrt{s}$. The remapped equilibrium is output in right-handed SFL coordinates $(\rho_{\text{tor}}, \theta_{\text{SFL}}, \phi)$.

Given the remapped SFL equilibrium, the MHD stability equations are discretized using a Galerkin projection approach. To increase flexibility of the resulting code, this Galerkin discretization is first formulated in terms of a general linear system of the form

$$\mathbf{B} \frac{\partial^\sigma \mathbf{X}}{\partial t^\sigma} = \mathbf{A}\mathbf{X}, \quad (\text{where } \sigma = 1 \text{ or } 2). \quad (5.1)$$

Here $\mathbf{X} = (X_1, \dots, X_N)$ are the numerical variables (N depends on the particular model under consideration), and \mathbf{A} and \mathbf{B} are matrices of differential operators, i.e. \mathbf{A} is of the form

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & A_{13} & \dots & A_{1N} \\ A_{21} & A_{22} & A_{23} & \dots & A_{2N} \\ A_{31} & A_{32} & A_{33} & & A_{3N} \\ \vdots & \vdots & & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & & A_{NN} \end{pmatrix},$$

where we assume each A_{ij} to be a differential operator of the form

$$A_{ij}f = a_{ij}^{(0)}f + \sum_{k=1}^3 a_{ij}^{(1);k} \frac{\partial f}{\partial x^k} + \sum_{k,\ell=1}^3 a_{ij}^{(2);k,\ell} \frac{\partial^2 f}{\partial x^k \partial x^\ell} \quad (5.2)$$

Here $(x^1, x^2, x^3) \equiv (\rho, \theta, \phi)$ and we have assumed partial derivatives of at most second-order to be present. The coefficients $a_{ij}^{(0)}, a_{ij}^{(1);k}, a_{ij}^{(2);k,\ell}$ are expressed in terms of equilibrium quantities and have to be derived based on the fluid model under consideration, and the choice of numerical variables X_1, \dots, X_N . As a consequence of the assumed axisymmetric equilibrium, all the coefficients $a_{ij}^{(0)}, a_{ij}^{(1);k}, a_{ij}^{(2);k,\ell}$ are assumed to be functions of ρ, θ , but independent of the toroidal angle ϕ .

To give a concrete example, for the ideal MHD model formulated in terms of a displacement ξ , one could take $N = 3$ with numerical variables $\mathbf{X} = (\xi^\rho, \xi^\theta, \xi^\phi)$, and we would have $\mathbf{B} = \rho_0 \mathbf{I}$, where \mathbf{I} is the 3×3 identity matrix, and \mathbf{A} would need to be determined by identifying the dependency of each component of the ideal MHD force operator

$$\delta F(\xi) = -\nabla \delta p + \delta \mathbf{j} \times \mathbf{B} + \mathbf{j} \times \delta \mathbf{B}.$$

To this end, explicit expressions for $\delta p, \delta \mathbf{j}, \delta \mathbf{B}$ need to be derived and written in terms of $X_1 = \xi^\rho, X_2 = \xi^\theta, X_3 = \xi^\phi$, to identify the coefficients A_{ij} of \mathbf{A} .

For the numerical discretization, each numerical variable X_μ ($\mu = 1, \dots, N$) is expanded in terms of a poloidal and toroidal Fourier series

$$X_\mu(\rho, \theta, \phi, t) = \sum_{m=-\underline{m}}^{\bar{m}} X_{(\mu,m)}(\rho) e^{im\theta + in\phi + \lambda t}. \quad (5.3)$$

The toroidal mode number n , and the range of poloidal modes $m \in [\underline{m}, \bar{m}]$ can be specified by the user. The growth rate λ (which is in general a complex number), is to be determined. To discretize the operators \mathbf{A} and \mathbf{B} , the “weak form” of equation (5.1) is now considered: \mathbf{X} is a solution of the weak form, provided that

$$\lambda^\sigma \mathbf{B}(\mathbf{Y}, \mathbf{X}) = \mathbf{A}(\mathbf{Y}, \mathbf{X}),$$

for all $\mathbf{Y} = (Y_1, \dots, Y_N)$, where each “test-function” Y_μ is itself of the form (5.3). Here $\mathbf{A}(\mathbf{Y}, \mathbf{X})$ is defined by

$$A(\mathbf{Y}, \mathbf{X}) \equiv \int \mathbf{Y}^* \cdot \mathbf{A} \mathbf{X} \, dx.$$

Here \mathbf{Y}^* denotes the complex conjugate of \mathbf{Y} . The Fourier expansion in θ and ϕ is particularly convenient, since it allows us to replace differentiation with respect to θ and ϕ , by a simple

Chapter 5. Final remarks and perspectives

multiplication:

$$\frac{\partial}{\partial \theta} \rightarrow i m, \quad \frac{\partial}{\partial \phi} \rightarrow i n.$$

The weak form of an operator A can now be represented as follows: In general for vectors \mathbf{Y}, \mathbf{X} :

$$A(\mathbf{Y}, \mathbf{X}) = \sum_{\mu', \mu=1}^N A_{\mu' \mu}(Y_{\mu'}, X_{\mu}),$$

with $A_{\mu' \mu}(Y_{\mu'}, X_{\mu})$ representing the weak form of $A_{\mu' \mu}$. assuming at most second order derivatives to occur, we can then write this in the form

$$A_{\mu' \mu}(Y_{\mu'}, X_{\mu}) = \int d\rho d\theta \sum_{r', r} \sum_{p', p} \alpha_{\mu' \mu}^{r', p'; r, p} \left[\frac{\partial^{r'}}{\partial \rho^{r'}} \frac{\partial^{p'}}{\partial \theta^{p'}} Y_{\mu'} \right]^* \left[\frac{\partial^r}{\partial \rho^r} \frac{\partial^p}{\partial \theta^p} X_{\mu} \right] \quad (5.4)$$

Here, radial derivative indices are denoted by r', r , and the poloidal derivatives by p', p . We have already assumed the toroidal dependency $\sim \exp(in\phi)$, so that toroidal derivatives correspond to multiplication with $i n$, which can be incorporated in the coefficients $\alpha_{\mu' \mu}^{r', p'; r, p} = \alpha_{\mu' \mu}^{r', p'; r, p}(\rho, \theta; n)$. Since both \mathbf{Y} and \mathbf{X} have only one toroidal Fourier mode number, and $Y_{\mu'}^* \sim e^{-in\phi}$, whereas $X_{\mu} \sim e^{in\phi}$, these factors cancel in (5.4), and hence an integration over ϕ is not necessary. Due to the poloidal Fourier expansion the integration over θ can be carried out analytically, leading to a coupled system in the radial direction, where the coupling of the m' -th Fourier mode of the μ' -th component $Y_{(\mu', m')}$ of \mathbf{Y} with the m -th Fourier mode of the μ -th component $X_{(\mu, m)}$ of \mathbf{X} is given by

$$\begin{aligned} A(Y_{(\mu', m')} e^{im'\theta}, X_{(\mu, m)} e^{im\theta}) \\ = \int d\rho \sum_{r', r} [\underline{\alpha}]_{(\mu', m'), (\mu, m)}^{r'; r} \left[\left(\frac{d}{d\rho} \right)^{r'} Y_{(\mu', m')}(\rho) \right]^* \left[\left(\frac{d}{d\rho} \right)^r X_{(\mu, m)}(\rho) \right]. \end{aligned}$$

Here, the coefficients $[\underline{\alpha}]_{(\mu', m'), (\mu, m)}^{r'; r} = [\underline{\alpha}]_{(\mu', m'), (\mu, m)}^{r'; r}(\rho)$ are defined by

$$[\underline{\alpha}]_{(\mu', m'), (\mu, m)}^{r'; r}(\rho) \equiv \sum_{p', p} (-im')^{p'} (im)^p \left(\int d\theta e^{i(m'-m)\theta} \alpha_{\mu' \mu}^{r', p'; r, p}(\rho, \theta) \right).$$

Note that for a straight cylinder equilibrium (no toroidal effects), the magnetic background has no dependency on θ , and hence all the coefficients $\alpha_{\mu' \mu}^{r', p'; r, p}$ are functions of ρ , only. In particular, this implies that the coupling between different mode-numbers vanishes in this case. In the above Galerkin formulation, this is equivalent to the vanishing of the coefficients $[\alpha]_{(\mu', m'), (\mu, m)}^{r'; r} = 0$ for $m' \neq m$. When toroidal effects are included, this is evidently no longer the case, and the coupling between different mode numbers needs to be taken into account.

Up to this point, no discretization in the radial direction has been carried out. To discretize in

ρ , we make an additional expansion in terms of radial bases. To this end, for each numerical variable X_μ , we fix a set of radial basis functions $\Lambda_k^{(\mu)}(\rho)$, where $k = 1, \dots, N_r$ and N_r is determined from the number of radial grid points (specified by the user). In our implementation, we have chosen $\Lambda_k^{(\mu)}$ to be a basis of either 2nd- or 3rd-order Bsplines (with degree depending on the numerical variable). As is well-known, a judicious choice of radial expansion functions is required to avoid the phenomenon of “spectral pollution” [GR85]. Given a Fourier mode number m , we expand $X_{(\mu,m)}$ as

$$X_{(\mu,m)}(\rho) = \sum_{k=1}^{N_r} \hat{X}_{(\mu,m,k)} \Lambda_k^{(\mu)}(\rho).$$

Thus, the numerical variables \mathbf{X} can now be determined entirely in terms of the collection of (complex) coefficients $\hat{X}_{(\mu,m,k)}$, where $\mu = 1, \dots, N$, $m = \{\underline{m}, \dots, \bar{m}\}$, and $k = 1, \dots, N_r$. Instead of this triple of indices (μ, m, k) , for the numerical implementation, we introduce a new (logical) index $\mathbf{v} \triangleq (\mu, m, k)$. The numerical discretization is now completed by defining the matrix coefficients

$$\hat{A}_{\mathbf{v}',\mathbf{v}} \equiv \hat{A}_{(\mu',m',k'),(\mu,m,k)} \equiv \sum_{r',r} \int d\rho [\underline{\alpha}]_{(\mu',m'),(\mu,m)}^{r';r} \left[\left(\frac{d}{d\rho} \right)^{r'} \Lambda_{k'}^{(\mu')} \right] \left[\left(\frac{d}{d\rho} \right)^r \Lambda_k^{(\mu)} \right].$$

The above procedure thus associates to an operator \mathbf{A} , the corresponding coefficients $\hat{A}_{\mathbf{v}',\mathbf{v}}$. In terms of these coefficients, we have

$$A(\mathbf{Y}, \mathbf{X}) = \sum_{\mathbf{v}',\mathbf{v}} \hat{Y}_{\mathbf{v}'} \hat{A}_{\mathbf{v}',\mathbf{v}} \hat{X}_{\mathbf{v}} \equiv \sum_{\mu',\mu=1}^N \sum_{m',m=\underline{m}}^{\bar{m}} \sum_{k',k=1}^{N_r} \hat{Y}_{(\mu',m',k')} \hat{A}_{(\mu',m',k'),(\mu,m,k)} \hat{X}_{(\mu,m,k)}.$$

Thus, the original equation

$$\mathbf{B} \frac{\partial^\sigma \mathbf{X}}{\partial t^\sigma} = \mathbf{AX}, \quad (\text{where } \sigma = 1 \text{ or } 2),$$

is discretized as a generalized eigenvalue problem of the form

$$\lambda^\sigma \hat{\mathbf{B}} \hat{\mathbf{X}} = \hat{\mathbf{A}} \hat{\mathbf{X}}, \quad (\text{where } \sigma = 1 \text{ or } 2),$$

which can also be formulated in terms of an initial value problem

$$\hat{\mathbf{B}} \frac{d^\sigma \hat{\mathbf{X}}}{dt^\sigma} = \hat{\mathbf{A}} \hat{\mathbf{X}}, \quad (\text{where } \sigma = 1 \text{ or } 2).$$

The eigenvalue problem is presently solved by loading the discretized matrices in MATLAB and using the `eigs` function (the implemented Stab MATLAB-class provides convenient wrappers to the relevant function calls to carry this out in practice, and to illustrate the results including spectra and eigenfunctions).

In practice, the operators \mathbf{A} and \mathbf{B} are determined by specifying the coupling coefficients of

the m' -th poloidal Fourier mode of the μ' -th numerical variable with the m -th poloidal Fourier mode of the μ -th numerical variable for all μ', μ and m', m . A convenient short-hand notation for this, which is in common use is to write down the explicit expressions for

$$\begin{aligned} A(k, \ell) &\equiv \sum_{p', p} (-im')^{p'} (im)^p \alpha_{k\ell}^{r', p'; r, p} \Big|_{r'=0, r=0}, \\ A(k', \ell) &\equiv \sum_{p', p} (-im')^{p'} (im)^p \alpha_{k\ell}^{r', p'; r, p} \Big|_{r'=1, r=0}, \\ A(k, \ell') &\equiv \sum_{p', p} (-im')^{p'} (im)^p \alpha_{k\ell}^{r', p'; r, p} \Big|_{r'=0, r=1}, \\ A(k', \ell') &\equiv \sum_{p', p} (-im')^{p'} (im)^p \alpha_{k\ell}^{r', p'; r, p} \Big|_{r'=1, r=1}, \\ &\vdots \end{aligned}$$

and so forth, for a given model.

For the numerical implementation, this means that once these coefficients $A(k, \ell)$, $A(k', \ell)$, $A(k, \ell')$, $A(k', \ell')$ are specified, the remainder of the Galerkin discretization of the operator can be carried out abstractly and in general, following the derivation presented above. In the present thesis, a numerical implementation of the Galerkin discretization has been carried out. To increase flexibility of the resulting code, we have chosen to encapsulate the Galerkin discretization of systems in an abstract Fortran type `PDE_op`, which contains procedure pointers `Aop_pt` and `Bop_pt`. A new model can now be specified by creating a concrete instance of this abstract `PDE_op` type, which provides explicit procedures to evaluate the coefficients of the operators \mathbf{A} and \mathbf{B} in weak form. The `PDE_op` module then automatically computes the corresponding matrices $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ for a given equilibrium, based on the specified operator coefficients (cp. figure 5.2).

Once a model is written down, the determination of the operator coefficients is an arduous, but mechanical process. For the currently implemented MHD models in the current version of the code VENUS-MHD, we have made extensive use of `Mathematica`, to carry out the necessary algebraic manipulations, and to write out the resulting operator coefficients which can then be copied into the `Fortran` code.

The radial discretization allows the user to set not only the total number of radial grid points to be used for this discretization, but also to specify values of the q -profile at which more resolution is required. The code will then correspondingly create more densely packed mesh points at the corresponding radial positions. The radial mesh packing is illustrated in Figure 5.3.

Several tests have been carried out on this preliminary implementation of the Galerkin discretization, including a tentative benchmark for a implemented discretization of the ideal

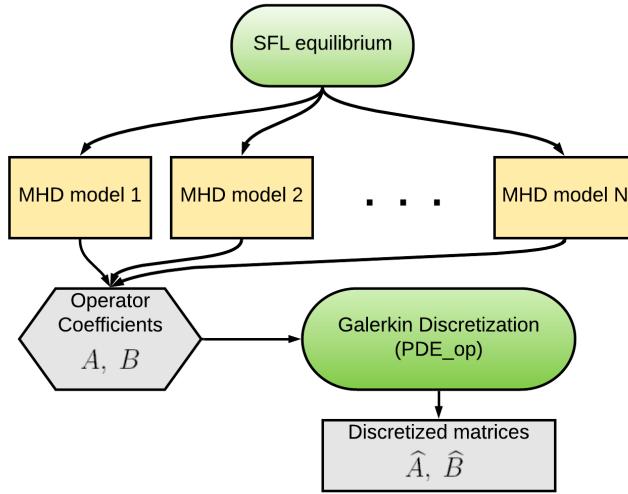


Figure 5.2 – Work flow of abstract implementation of Galerkin discretization. The choice of model determines the functional form of the operator coefficients, based upon which the discretization can be carried out uniformly. This allows new models to be added to the code with relative ease.

MHD momentum equation formulated in terms of a displacement

$$\lambda^2 \rho_0 \xi = \delta F(\xi).$$

for an internal kink, shown in figure 1.6, chapter 1. A more thorough benchmarking of the code and the various models that are currently implemented is however left for future work.¹

Kinetic solver

The following approach to solving the spectral problem for the kinetic-MHD model of chapter 4, in the absence of rotation, diamagnetic effects in the fluid inertia or a quasi-neutrality equation, has also been attempted. However, this project was abandoned due to several practical difficulties, which we will summarize below. In the attempted approach, the goal is to solve the following non-linear eigenvalue problem

$$-\omega^2 \rho_0 \xi = \delta F(\xi) - \nabla \cdot \delta P_h(\xi; \omega),$$

where $\delta F(\xi)$ is the conventional ideal MHD force operator, and an explicit expression for the perturbed (hot) pressure tensor $\delta P_h(\xi; \omega)$ including full-orbit width effects has been given by Porcelli et al. [PSK94]. This system describes the interaction of an ideal MHD mode with a species of supra-thermal ions (e.g. generated by NBI or ICRH). Following the Galerkin

¹Detailed derivations and explanations of the available models can be found by opening the various Mathematica Scripts that come with the code, and which have been used to generate the Fortran code for each of these models.

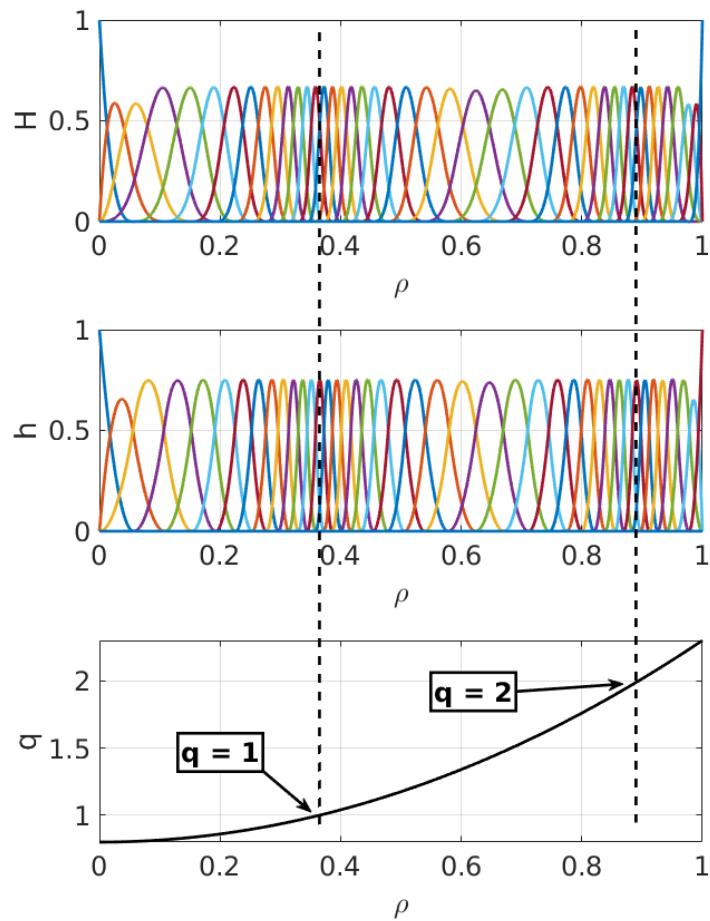


Figure 5.3 – Radial mesh-packing based on q -profile: Specification of the list of values $q = 1, 2$ leads to more densely packed radial mesh points near the corresponding q -values, and enables higher resolution of the singular layers in ideal MHD. Higher-order radial basis function (3rd order B-splines) (top), lower-order radial basis functions (2nd order B-splines) (middle) and corresponding q -profile (bottom).

procedure described above, leads to an eigenvalue problem of the form

$$-\omega^2 \hat{\mathbf{B}}\mathbf{X} = \hat{\mathbf{A}}\mathbf{X} + \hat{\mathbf{C}}(\omega)\mathbf{X},$$

where the additional matrix $\hat{\mathbf{C}}(\omega)$ depends on the eigenvalue, and $\omega \mapsto \hat{\mathbf{C}}(\omega)$ is a holomorphic function of ω in the upper half-plane (corresponding to unstable modes – we did not consider continua). This approach thus suggests to treat the spectral problem of this kinetic-MHD model as a “non-linear” eigenvalue problem, which asks to determine ω and \mathbf{X} , such that $\mathbf{M}(\omega)\mathbf{X} = 0$, with $\mathbf{M}(\omega) \equiv \omega^2 \hat{\mathbf{B}} + \hat{\mathbf{A}} + \hat{\mathbf{C}}(\omega)$.

Major difficulties with this approach include the following:

- While there exist some numerical libraries such as SLEPC, which can in principle tackle non-linear eigenvalue problems of this form, the methods are by far not as well developed as for usual eigenvalue problems.
- The matrix structure of $\hat{\mathbf{M}}(\omega)$ goes from a sparse to a full matrix with increasing radial orbit width of the additional hot particle species. This poses problems with the available memory.

Other codes, which successfully follow a similar strategy (such as MARS-K [LCCH08], LIGKA [LGKP07]) make various simplifications to make this problem more tractable. As a consequence they do not take into account full finite-orbit width (at least not in a completely consistent way). Work by Porcelli et al. [PSBZ92] has in particular shown such effects to be important for the internal kink, and therefore it would be desirable to not make such simplifications. Perhaps even more fundamentally, the problem of kinetic corrections to the inertia (see [GHH00, appendix], or [ZC14]) are captured only on inclusion of the radial drift of a particle across the radial structure of the mode. Inertia corrections apply in particular to thermal ions, so it appears that capturing the orbit width corrections is necessary for all ions, not just minority ions.

Therefore, a fresh start appears necessary for the implementation of the kinetic-MHD system. If only the most unstable mode is to be determined, then an initial value problem would be suitable, which evolves the fluid and kinetic models congruently, suggesting a conceptually simple PIC approach to this problem.

A Baños drift in purely sheared background

A.1 Introduction

It is instructive to see how the inclusion of higher-order terms in equations (3.8)-(3.11) discussed in chapter 3 (based on the detailed derivation presented in chapter 2) can account for the Baños drift for a simple example. The present discussion extends the results of [PGC15, Appendix A] and has been included in the publication [LPGC17, appendix].

We consider the sheared magnetic background $\mathbf{B}(\mathbf{x}) = B_0[\sin(\tau x)\hat{\mathbf{y}} + \cos(\tau x)\hat{\mathbf{z}}]$ in cartesian coordinates (x, y, z) . This field satisfies $\nabla \cdot \mathbf{B} = 0$, $\nabla B = 0$, $\boldsymbol{\kappa} \equiv \mathbf{b} \cdot \nabla \mathbf{b} = 0$ and $\tau = \mathbf{b} \cdot \nabla \times \mathbf{b}$, justifying our choice of the notation for the parameter τ . The analysis detailed in [PGC15] shows that for small enough values of τ , a particle with initial data

$$\begin{aligned} x_0 &= 0 & y_0 &= \rho_0 & z_0 &= 0 \\ \dot{x}_0 &= u_0 & \dot{y}_0 &= 0 & \dot{z}_0 &= v_0 \end{aligned} \tag{A.1}$$

will follow a closed orbit in the (x, y) plane, and that to leading order in τ , we have

$$\langle \dot{z} \rangle - v_0 = \frac{\tau u_0^2}{4\Omega_0}, \tag{A.2}$$

$$\langle v_{||} \rangle - v_0 = -\frac{\tau u_0^2}{4\Omega_0}, \tag{A.3}$$

so that

$$\langle \dot{z} \rangle - \langle v_{||} \rangle = \frac{\tau u_0^2}{2\Omega_0} = \frac{\tau \mu}{q}. \tag{A.4}$$

This result is astonishing, because it implies that even in the absence of a finite $v_{||}$ component, the guiding-centre will drift along \mathbf{B} . This guiding-centre motion is induced by the shear.

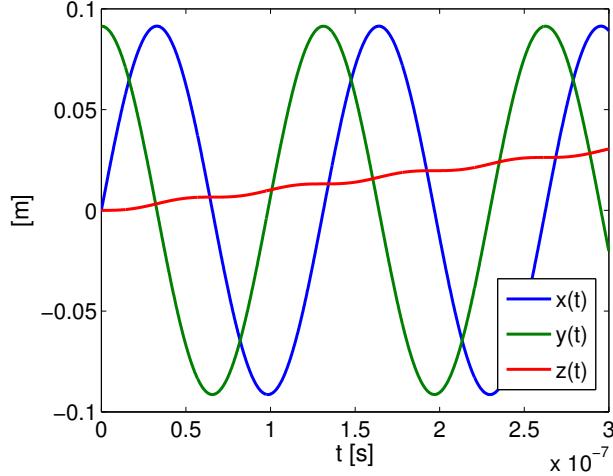


Figure A.1 – Temporal evolution of x , y , z coordinates for particle with initialization (A.1). The parameters were set as $E = 10\text{keV}$, $\tau = 1$, $B_0 = 0.5\text{T}$. The Baños drift in the z -direction is clearly visible.

A.2 Derivation from Lie perturbation methods

We shall now explain how the above result based on the Lorentzian equations of motion can be obtained from the equations presented in section 3.2. We fix $\mathbf{e}_1 = \hat{\mathbf{x}}$, which determines $\mathbf{e}_2 = \mathbf{b} \times \mathbf{e}_1 = -\sin(\tau x)\hat{\mathbf{y}} + \cos(\tau x)\hat{\mathbf{z}}$. One obtains $\nabla \mathbf{b} = \tau \cos(\tau x)\hat{\mathbf{x}}\hat{\mathbf{y}} - \tau \sin(\tau x)\hat{\mathbf{x}}\hat{\mathbf{z}}$, so that

$$\begin{aligned}\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 &= \tau, & \mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 &= 0, \\ \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 &= 0, & \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 &= 0.\end{aligned}$$

Using the definition (cp. equation (2.23)) of \mathbf{a}_1 , we arrive at the simple relation

$$\mathbf{a}_1 : \nabla \mathbf{b} = \tau / 2 (2 \cos^2 \zeta - 1). \quad (\text{A.5})$$

Using also the vanishing of the other terms in (3.9), the general relations evaluated at the guiding-centre position $X = Y = 0$ reduce to first order in gyroradius to the simple form

$$x = \rho_0 \cos(\zeta), \quad (\text{A.6})$$

$$y = -\rho_0 \sin(\zeta), \quad (\text{A.7})$$

$$Z = z, \quad (\text{A.8})$$

$$U = v_{||} + \frac{\mu\tau}{q} \left(\cos^2(\zeta) + \frac{1}{2} \right). \quad (\text{A.9})$$

The initialization considered here is then seen to correspond to $\zeta = -\pi/2$, which yields the following initialization of the guiding-centre: $X_0 = 0$, $U_0 = v_0 + \frac{\mu\tau}{2q}$. Due to the vanishing of ∇B , the guiding-centre velocity satisfies $\dot{U} = 0$ for this particular equilibrium, from which it follows that $U = U_0$ for all times, i.e. $U = v_0 + \frac{\mu\tau}{2q}$. Gyroaveraging equation (A.9) yields $U = \langle v_{||} \rangle + \frac{\mu\tau}{q}$.

Equating the two expression for U , we find

$$\langle v_{||} \rangle - v_0 = \frac{\mu\tau}{2q}. \quad (\text{A.10})$$

Observing that $\dot{X} = 0$, $\dot{Y} = 0$ and $\dot{Z} = U + \eta\mu\tau/q$, we find

$$\dot{Z} - v_0 = \frac{\mu\tau}{2q}. \quad (\text{A.11})$$

Equation (A.10) is equivalent with (A.3), whereas equation (A.11) recovers (A.2).

Substitution of $\zeta(t) \approx \Omega t - \pi/2$ in (A.9) yields

$$U \approx v_{||} + \frac{\mu\tau}{q} \left(\sin^2(\Omega t) + \frac{1}{2} \right), \quad (*)$$

relating the temporal evolution of the full-Lorentzian parallel velocity $v_{||}$ to the guiding-centre parallel velocity U .

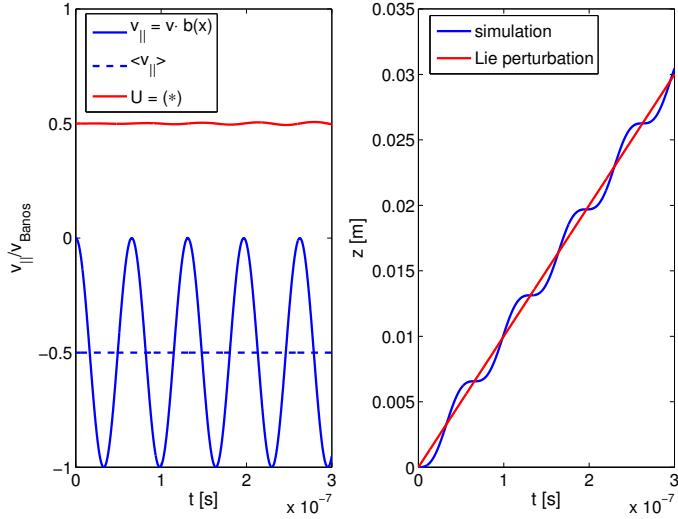


Figure A.2 – Temporal evolution of $v_{||}$ and the z -coordinate for the same particle of figure A.1. Indicated in red are relevant expressions obtained from the Lie perturbation theory.

The temporal evolution of $v_{||} \equiv \mathbf{b}(\mathbf{x}) \cdot \mathbf{v}$ computed from numerical integration of the Lorentzian equations of motion is depicted in figure A.2. It is interesting to note that for the present case under consideration the local particle expression for $v_{||}$ does not even agree in sign with the guiding-centre expression $(*)$ for the parallel guiding-centre velocity U , above. A careful distinction should be made between the guiding-centre and local particle parallel velocities. In figure A.2 on the right, it is shown that the predicted parallel drift recovers the average motion in the z -direction to good accuracy.

We emphasize that these results have been obtained using the conventional guiding-centre

Appendix A. Banos drift in purely sheared background

Lagrangian, without the additional high-order terms considered in (3.4). Consistent inclusion of all first order corrections in (3.9) was necessary. In particular, this indicates that with a judicious choice of guiding-centre variables, guiding-centre following based on the Lagrangian (3.21) is sufficient to account for the particle dynamics induced by first-order variations in \mathbf{B} , also parallel to the field lines.

B Detailed derivation of gyroviscous tensor components

In this appendix, we present the details of the derivation of the gyroviscous pressure tensor components of chapter 4.

B.1 FLR correction (I) – evaluation of eq. (4.13)

We evaluate the contribution of

$$\int d^3v mw [\mathbf{b}V_{gc} + V_{gc}\mathbf{b}] F$$

where we use (2.32)

$$V_{gc} = \mathbf{b} \times \left[\frac{\mu}{qB} \nabla B + \frac{w^2}{\Omega} \mathbf{b} \cdot \nabla \mathbf{b} \right] + \frac{w}{\Omega} \mathbf{b} \times \left[\frac{d\mathbf{b}}{dt} + \mathbf{b} \cdot \nabla \mathbf{u} \right] + \frac{1}{\Omega} \mathbf{b} \times \frac{d\mathbf{u}}{dt}.$$

The relevant contribution of this correction to \mathbf{P} , is therefore

$$\Omega P_{b\perp} = \mathbf{b} \times \left[P_{||} \left(\frac{d\mathbf{b}}{dt} + \mathbf{b} \cdot \nabla \mathbf{u} \right) + q_{||}^\perp \nabla B + 2q_{||}^{||} \boldsymbol{\kappa} \right],$$

where $\boldsymbol{\kappa} \equiv \mathbf{b} \cdot \nabla \mathbf{b}$ denotes the field line curvature.

B.2 FLR correction (II) – evaluation of eq. (4.14)

We evaluate

$$\int d^3v m [\dot{\rho}_{gc}^{(0)} \dot{\rho}_{gc}^{(1)} + \dot{\rho}_{gc}^{(1)} \dot{\rho}_{gc}^{(0)}] F$$

Appendix B. Detailed derivation of gyroviscous tensor components

This term requires the evaluation of the first order correction to $\dot{\rho}_{\text{gc}}$. We have

$$\begin{aligned}\dot{\rho}_{\text{gc}} &= (\partial_t + \dot{X} \cdot \nabla + i\omega \partial_w + \dot{\zeta} \partial_\zeta) \rho_{\text{gc}} \\ &= \dot{\zeta}^{(0)} \partial_\zeta \rho_0 + \epsilon_B \dot{\zeta}^{(1)} \partial_\zeta \rho_0 + \epsilon_B \partial_t \rho_0 + \epsilon_B \dot{X}^{(0)} \cdot \nabla \rho_0 + \epsilon_B \dot{\zeta}^{(0)} \partial_\zeta \rho_1.\end{aligned}$$

We note that

$$\rho_0 = \rho_0 \hat{\rho} = \sqrt{\frac{2\mu B(\mathbf{X}, t)}{m}} \hat{\rho}(\mathbf{X}, \zeta, t),$$

so that

$$\nabla \rho_0 = \frac{1}{2} \nabla \log B \otimes \rho_0 + \rho_0 \nabla \hat{\rho}, \quad (\text{B.1})$$

$$\partial_t \rho_0 = \frac{1}{2} \partial_t \log B \rho_0 + \rho_0 \partial_t \hat{\rho}. \quad (\text{B.2})$$

Furthermore,

$$\dot{\zeta}^{(0)} = \Omega, \quad \dot{\zeta}^{(1)} = \mathbf{R} \cdot \mathbf{W} + S + \frac{1}{2} \mathbf{b} \cdot \nabla \times \mathbf{W},$$

with $\mathbf{R} \equiv (\nabla \hat{\mathbf{L}}) \cdot \hat{\mathbf{p}} = (\nabla \mathbf{e}_1) \cdot \mathbf{e}_2$, $S = \hat{\mathbf{p}} \cdot \partial_t \hat{\mathbf{L}} = \mathbf{e}_2 \cdot \partial_t \mathbf{e}_1$, and $\mathbf{W} \equiv \mathbf{u} + w \mathbf{b}$. We can now write

$$\begin{aligned}\frac{1}{\rho_0} \dot{\rho}_1 &= [\mathbf{R} \cdot \mathbf{W} + S] \hat{\mathbf{L}} + \left[\frac{1}{2} \mathbf{b} \cdot \nabla \times \mathbf{W} \right] \hat{\mathbf{L}} + \Omega \partial_\zeta \rho_1 / \rho_0 \\ &\quad + \frac{1}{2} [(\partial_t + \mathbf{W} \cdot \nabla) \log B] \hat{\rho} + (\partial_t + \mathbf{W} \cdot \nabla) \hat{\rho}.\end{aligned}$$

We can further simplify this expression, by noting that

$$\begin{aligned}[\partial_t + \mathbf{W} \cdot \nabla] \hat{\rho} &= -[(\partial_t + \mathbf{W} \cdot \nabla) \hat{\mathbf{L}} \cdot \hat{\mathbf{p}}] \hat{\mathbf{L}} - [(\partial_t + \mathbf{W} \cdot \nabla) \mathbf{b} \cdot \hat{\mathbf{p}}] \mathbf{b} \\ &= -[S + \mathbf{W} \cdot \mathbf{R}] \hat{\mathbf{L}} - \left(\left[\frac{d\mathbf{b}}{dt} + w \boldsymbol{\kappa} \right] \cdot \hat{\mathbf{p}} \right) \mathbf{b}.\end{aligned} \quad (\text{B.3})$$

Thus, we find that the terms involving \mathbf{R}, S cancel, leaving an expression that is written in terms of derivatives of \mathbf{B} and \mathbf{u} :

$$\begin{aligned}\frac{1}{\rho_0} \dot{\rho}_1 &= \left[\frac{1}{2} \mathbf{b} \cdot \nabla \times \mathbf{W} \right] \hat{\mathbf{L}} + \Omega \partial_\zeta \rho_1 / \rho_0 \\ &\quad + \frac{1}{2} [(\partial_t + \mathbf{W} \cdot \nabla) \log B] \hat{\rho} - \left(\left[\frac{d\mathbf{b}}{dt} + w \boldsymbol{\kappa} \right] \cdot \hat{\mathbf{p}} \right) \mathbf{b}.\end{aligned}$$

We then find, using $\hat{\rho} = \mathbf{b} \times \hat{\perp}$ and hence $\mathbf{A} \cdot \hat{\rho} = -\mathbf{b} \times \mathbf{A} \cdot \hat{\perp}$, that

$$\begin{aligned} m\langle \dot{\rho}_{\text{gc}}^{(0)} \dot{\rho}_{\text{gc}}^{(1)} \rangle &= \frac{2\mu B}{\Omega} \langle \hat{\perp} \dot{\rho}_{\text{gc}}^{(1)} / \rho_0 \rangle \\ &= \frac{\mu B}{\Omega} [\mathbf{b} \cdot \nabla \times \mathbf{W}] \langle \hat{\perp} \hat{\perp} \rangle + (2\mu B) \langle \hat{\rho} \rho_1 / \rho_0 \rangle \\ &\quad + \frac{\mu B}{\Omega} [(\partial_t + \mathbf{W} \cdot \nabla) \log B] \langle \hat{\perp} \hat{\rho} \rangle \\ &\quad + \frac{2\mu B}{\Omega} \mathbf{b} \times \left[\frac{d\mathbf{b}}{dt} + w\boldsymbol{\kappa} \right] \cdot \langle \hat{\perp} \hat{\perp} \rangle \mathbf{b} \\ &= \frac{\mu B}{\Omega} \frac{1}{2} [\mathbf{b} \cdot \nabla \times \mathbf{W}] (\mathbf{I} - \mathbf{b}\mathbf{b}) + (2\mu B) \langle \hat{\rho} \rho_1 / \rho_0 \rangle \\ &\quad + (\text{AS}) + \frac{\mu B}{\Omega} \left(\mathbf{b} \times \left[\frac{d\mathbf{b}}{dt} + w\boldsymbol{\kappa} \right] \right) \mathbf{b}, \end{aligned}$$

where (AS) is an anti-symmetric term, which will cancel in the final expression for the pressure, which is symmetric. To complete this calculation, we still need to evaluate $\langle \hat{\rho} \rho_1 \rangle / \rho_0$. Since this expression is a gyro-average, we only require the contributions to ρ_1 , which can be written as odd polynomials in $\hat{\rho}, \hat{\perp}$.

In general, we have the following expression (cp. (2.36))

$$\rho_1 = -G_2^X + \frac{1}{2} (G_1^\mu \partial_\mu + G_1^\zeta \partial_\zeta) G_1^X + \frac{1}{2} G_1^X \cdot \nabla G_1^X,$$

where $G_1^X = -\rho_0$. The perpendicular component of G_2^X can be written (cp. (2.41))

$$G_{2,\perp}^X = \frac{1}{2} (g_\mu \partial_\mu + g_\zeta \partial_\zeta) \rho_0 + \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \rho_0,$$

where

$$\begin{aligned} G_1^\mu &= g_\mu + \mu \rho_0 \cdot \nabla \log B, \\ G_1^\zeta &= g_\zeta - \rho_0 \cdot \mathbf{R}. \end{aligned}$$

We find that

$$\begin{aligned} \rho_1 &= -G_{2,\parallel}^X \mathbf{b} - (g_\mu \partial_\mu + g_\zeta \partial_\zeta) \rho_0 - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \rho_0 \\ &\quad - \frac{1}{2} \mu (\rho_0 \cdot \nabla \log B) \partial_\mu \rho_0 + \frac{1}{2} (\rho_0 \cdot \mathbf{R}) \partial_\zeta \rho_0 + \frac{1}{2} \rho_0 \cdot \nabla \rho_0 \\ &= -G_{2,\parallel}^X \mathbf{b} - (g_\mu \partial_\mu + g_\zeta \partial_\zeta) \rho_0 - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \rho_0 \\ &\quad - \frac{1}{4} (\rho_0 \cdot \nabla \log B) \rho_0 + \frac{1}{2} (\rho_0 \cdot \mathbf{R}) \partial_\zeta \rho_0 + \frac{1}{2} \rho_0 \cdot \nabla \rho_0. \end{aligned}$$

Appendix B. Detailed derivation of gyroviscous tensor components

By similar manipulations as in equation (B.3), we observe that the last term can be written

$$\frac{\rho_0^2}{2} [(\hat{\boldsymbol{\rho}} \cdot \mathbf{R}) \hat{\mathbf{L}} + \hat{\boldsymbol{\rho}} \cdot \nabla \hat{\boldsymbol{\rho}}] = -\frac{\rho_0^2}{2} \hat{\boldsymbol{\rho}} \cdot \nabla \mathbf{b} \cdot \hat{\boldsymbol{\rho}} = -\frac{1}{2} \boldsymbol{\rho}_0 \cdot \nabla \mathbf{b} \cdot \boldsymbol{\rho}_0,$$

and thus

$$\begin{aligned} \boldsymbol{\rho}_1 &= -(g_\mu \partial_\mu + g_\zeta \partial_\zeta) \boldsymbol{\rho}_0 - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \boldsymbol{\rho}_0 \\ &\quad - \left(G_{2,\parallel}^X + \frac{1}{2} \boldsymbol{\rho}_0 \cdot \nabla \mathbf{b} \cdot \boldsymbol{\rho}_0 \right) \mathbf{b}. \end{aligned}$$

This expression is still completely general, and recovers equation (49) [Bri10] in the limit $\mathbf{u} \rightarrow 0$.

Citing the results of chapter 2, equations (2.96), (2.84), (2.93) and (B.6), we find

$$\begin{aligned} G_{2,\parallel}^X &= -\frac{1}{\Omega} \boldsymbol{\rho}_0 \cdot \mathbf{b} \times [\mathbf{b} \cdot \nabla \mathbf{W} + \mathbf{W} \cdot \nabla \mathbf{b} + \partial_t \mathbf{b}] \\ &\quad - \frac{\mu}{4q\Omega} [\hat{\boldsymbol{\rho}} \hat{\boldsymbol{\rho}} - \hat{\mathbf{L}} \hat{\mathbf{L}}] : \nabla \mathbf{b} \end{aligned} \tag{B.4}$$

$$\begin{aligned} g_\mu &= -\frac{\mu}{\Omega} \mathbf{b} \cdot \nabla \times \mathbf{W} + \frac{m}{B} \boldsymbol{\rho}_0 \cdot [(\mathbf{W} \cdot \nabla) \mathbf{W}] \\ &\quad + \frac{\mu}{2\Omega} (\hat{\boldsymbol{\rho}} \hat{\mathbf{L}} + \hat{\mathbf{L}} \hat{\boldsymbol{\rho}}) : \nabla \mathbf{W} + \frac{m}{B} \boldsymbol{\rho}_0 \cdot \partial_t \mathbf{W}, \end{aligned} \tag{B.5}$$

$$\begin{aligned} g_\zeta &= -\frac{q}{m} \partial_\mu s_3 \\ &= -\frac{1}{4\Omega} (\hat{\boldsymbol{\rho}} \hat{\boldsymbol{\rho}} - \hat{\mathbf{L}} \hat{\mathbf{L}}) : \nabla \mathbf{W} \\ &\quad + \boldsymbol{\rho}_0 \cdot \frac{\mathbf{b} \times \nabla B}{B} - \frac{q}{2\mu} \frac{1}{\Omega} \boldsymbol{\rho}_0 \cdot \mathbf{b} \times [(\partial_t + \mathbf{W} \cdot \nabla) \mathbf{W}]. \end{aligned} \tag{B.6}$$

We only require the contribution to $G_{2,\parallel}^X$ that is linear in $\hat{\boldsymbol{\rho}}, \hat{\mathbf{L}}$ and the contributions to g_μ, g_ζ which are constant or quadratic in $\hat{\boldsymbol{\rho}}, \hat{\mathbf{L}}$. We can therefore take

$$\begin{aligned} G_{2,\parallel}^X &\rightarrow -\frac{1}{\Omega} \boldsymbol{\rho}_0 \cdot \mathbf{b} \times [\mathbf{b} \cdot \nabla \mathbf{W} + \mathbf{W} \cdot \nabla \mathbf{b} + \partial_t \mathbf{b}] \\ g_\mu / (2\mu) &\rightarrow -\frac{1}{2\Omega} \mathbf{b} \cdot \nabla \times \mathbf{W} + \frac{1}{4\Omega} (\hat{\boldsymbol{\rho}} \hat{\mathbf{L}} + \hat{\mathbf{L}} \hat{\boldsymbol{\rho}}) : \nabla \mathbf{W}, \\ g_\zeta &\rightarrow -\frac{1}{4\Omega} (\hat{\boldsymbol{\rho}} \hat{\boldsymbol{\rho}} - \hat{\mathbf{L}} \hat{\mathbf{L}}) : \nabla \mathbf{W}, \end{aligned}$$

when carrying out the gyroaverage $\langle \hat{\rho} \rho_1 / \rho_0 \rangle$.

$$\begin{aligned}
 \langle \hat{\rho} \rho_1 / \rho_0 \rangle &= -\langle g_\mu / (2\mu) \hat{\rho} \hat{\rho} \rangle - \langle g_\zeta \hat{\rho} \hat{\perp} \rangle \\
 &\quad - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \langle \hat{\rho} \hat{\rho} \rangle - \langle G_{2,||}^X \hat{\rho} \rangle \mathbf{b} \\
 &= \frac{1}{2\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \langle \hat{\rho} \hat{\rho} \rangle + \frac{1}{\Omega} \left\langle \hat{\rho} \hat{\rho} \left[-\frac{(\hat{\rho} \hat{\perp} + \hat{\perp} \hat{\rho})}{4} \right] : \nabla \mathbf{W} \right\rangle \\
 &\quad + \frac{1}{\Omega} \left\langle \hat{\rho} \hat{\perp} \left[\frac{(\hat{\rho} \hat{\rho} - \hat{\perp} \hat{\perp})}{4} \right] : \nabla \mathbf{W} \right\rangle - \frac{1}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) \langle \hat{\rho} \hat{\rho} \rangle \\
 &\quad - \frac{1}{\Omega} \langle \hat{\rho} \hat{\rho} \rangle \cdot \left(\mathbf{b} \times \left[\mathbf{b} \cdot \nabla \mathbf{W} + \mathbf{W} \cdot \nabla \mathbf{b} + \frac{1}{2} \partial_t \mathbf{b} \right] \right) \mathbf{b} \\
 &= -\frac{1}{4\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) (\mathbf{I} - \mathbf{b}\mathbf{b}) \\
 &\quad + \frac{1}{\Omega} \left\langle \hat{\rho} \hat{\rho} \left[-\frac{(\hat{\rho} \hat{\perp} + \hat{\perp} \hat{\rho})}{4} \right] : \nabla \mathbf{W} \right\rangle \\
 &\quad + \frac{1}{\Omega} \left\langle \hat{\rho} \hat{\perp} \left[\frac{(\hat{\rho} \hat{\rho} - \hat{\perp} \hat{\perp})}{4} \right] : \nabla \mathbf{W} \right\rangle \\
 &\quad - \frac{1}{2\Omega} \left(\mathbf{b} \times \left[2w\boldsymbol{\kappa} + \mathbf{b} \cdot \nabla \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{b} + \frac{1}{2} \partial_t \mathbf{b} \right] \right) \mathbf{b}.
 \end{aligned}$$

After carrying out the remaining gyro-averages, we find

$$\begin{aligned}
 (2\mu B) \langle \hat{\rho} \rho_1 / \rho_0 \rangle &= -\frac{\mu B}{2\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{W}) (\mathbf{I} - \mathbf{b}\mathbf{b}) \\
 &\quad + \frac{\mu B}{4\Omega} [(\nabla_1 W_2 + \nabla_2 W_1)(\mathbf{e}_1 \mathbf{e}_1 - \mathbf{e}_2 \mathbf{e}_2) \\
 &\quad \quad + (\nabla_2 W_2 - \nabla_1 W_1)(\mathbf{e}_1 \mathbf{e}_1 + \mathbf{e}_2 \mathbf{e}_1)] \\
 &\quad - \frac{\mu B}{\Omega} (\mathbf{b} \times [2w\boldsymbol{\kappa} + \mathbf{b} \cdot \nabla \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{b} + \partial_t \mathbf{b}]) \mathbf{b}.
 \end{aligned}$$

Appendix B. Detailed derivation of gyroviscous tensor components

We therefore finally have

$$\begin{aligned}
m \left[\langle \dot{\boldsymbol{\rho}}_{\text{gc}}^{(0)} \dot{\boldsymbol{\rho}}_{\text{gc}}^{(1)} \rangle + \langle \dot{\boldsymbol{\rho}}_{\text{gc}}^{(0)} \dot{\boldsymbol{\rho}}_{\text{gc}}^{(1)} \rangle \right] &= \frac{\mu B}{\Omega} [\mathbf{b} \cdot \nabla \times \mathbf{W}] (\mathbf{I} - \mathbf{b} \mathbf{b}) \\
&\quad + (2\mu B) \left[\langle \widehat{\boldsymbol{\rho}} \boldsymbol{\rho}_{\text{gc}}^{(1)} / \rho_0 \rangle + \langle \boldsymbol{\rho}_{\text{gc}}^{(1)} \widehat{\boldsymbol{\rho}} / \rho_0 \rangle \right] \\
&\quad + \frac{\mu B}{\Omega} \left[\left(\mathbf{b} \times \left[\frac{d\mathbf{b}}{dt} + w\boldsymbol{\kappa} \right] \right) \mathbf{b} + (T) \right] \\
&= \frac{\mu B}{2\Omega} [(\nabla_1 W_2 + \nabla_2 W_1)(\mathbf{e}_1 \mathbf{e}_1 - \mathbf{e}_2 \mathbf{e}_2) \\
&\quad + (\nabla_2 W_2 - \nabla_1 W_1)(\mathbf{e}_1 \mathbf{e}_1 + \mathbf{e}_2 \mathbf{e}_1)] \\
&\quad + \frac{\mu B}{\Omega} \left[\left(\mathbf{b} \times \left[\frac{d\mathbf{b}}{dt} + w\boldsymbol{\kappa} \right] \right) \mathbf{b} + (T) \right] \\
&\quad - \frac{\mu B}{\Omega} \left[\left(\mathbf{b} \times \left[2w\boldsymbol{\kappa} + \mathbf{b} \cdot \nabla \mathbf{u} + \frac{d\mathbf{b}}{dt} \right] \right) \mathbf{b} + (T) \right] \\
&= \frac{\mu B}{2\Omega} [(\nabla_1 W_2 + \nabla_2 W_1)(\mathbf{e}_1 \mathbf{e}_1 - \mathbf{e}_2 \mathbf{e}_2) \\
&\quad + (\nabla_2 W_2 - \nabla_1 W_1)(\mathbf{e}_1 \mathbf{e}_1 + \mathbf{e}_2 \mathbf{e}_1)] \\
&\quad + \frac{\mu B}{\Omega} [(\mathbf{b} \times [-w\boldsymbol{\kappa} - \mathbf{b} \cdot \nabla \mathbf{u}]) \mathbf{b} + (T)],
\end{aligned}$$

where (T) indicates the transpose of the term to its left.

B.3 FLR correction (III) – evaluation of eq. (4.15)

We evaluate

$$(III) = -\epsilon_B d^3 \mathbf{v} m \left[\dot{\boldsymbol{\rho}}_{\text{gc}}^{(0)} (\boldsymbol{\rho}_0 \cdot \nabla \mathbf{u}) + (\boldsymbol{\rho}_0 \cdot \nabla \mathbf{u}) \dot{\boldsymbol{\rho}}_{\text{gc}}^{(0)} \right] F$$

We find

$$\begin{aligned}
(III) &= -\epsilon_B \int d^3 \mathbf{v} m [\dot{\boldsymbol{\rho}}_{\text{gc}} (\boldsymbol{\rho}_0 \cdot \nabla \mathbf{u}) + (\boldsymbol{\rho}_{\text{gc}} \cdot \nabla \mathbf{u}) \dot{\boldsymbol{\rho}}_0] F \\
&= -\epsilon_B \int d^3 \mathbf{v} \frac{2\mu B}{\Omega} [\widehat{\mathbf{l}} (\widehat{\boldsymbol{\rho}} \cdot \nabla \mathbf{u}) + (\widehat{\boldsymbol{\rho}} \cdot \nabla \mathbf{u}) \widehat{\mathbf{l}}] F \\
&= -\epsilon_B \frac{p_\perp}{\Omega} [\mathbf{e}_1 (\mathbf{e}_2 \cdot \nabla \mathbf{u}) - \mathbf{e}_2 (\mathbf{e}_1 \cdot \nabla \mathbf{u}) \\
&\quad + (\mathbf{e}_2 \cdot \nabla \mathbf{u}) \mathbf{e}_1 - (\mathbf{e}_1 \cdot \nabla \mathbf{u}) \mathbf{e}_2].
\end{aligned}$$

This can be further simplified as

$$\begin{aligned}
 (\text{III}) &= \frac{P_\perp}{\Omega} [(\nabla_2 u_2 - \nabla_1 u_1)(\mathbf{e}_1 \mathbf{e}_2 + \mathbf{e}_2 \mathbf{e}_1) \\
 &\quad + (\nabla_2 u_1 + \nabla_1 u_2)(\mathbf{e}_1 \mathbf{e}_1 - \mathbf{e}_2 \mathbf{e}_2)] \\
 &\quad + \frac{P_\perp}{\Omega} [\mathbf{b} \times (\nabla \mathbf{u} \cdot \mathbf{b}) + (T)] \\
 &\quad - \frac{P_\perp}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{u})(\mathbf{I} - \mathbf{b}\mathbf{b}).
 \end{aligned}$$

B.4 FLR correction (IV) – evaluation of eq. (4.16)

We evaluate

$$\begin{aligned}
 -\nabla \cdot \int d^3v mw\rho_0 [\mathbf{b}\dot{\rho}_{\text{gc}}^{(0)} + \dot{\rho}_{\text{gc}}^{(0)}\mathbf{b}] F &= -\nabla \cdot \int d^3v \frac{w\mu B}{\Omega} F [2\langle \hat{\rho}\mathbf{b}\hat{1} \rangle + 2\langle \hat{\rho}\hat{1}\mathbf{b} \rangle] \\
 &= -\nabla \cdot \left(\frac{q_{||}^\perp}{\Omega} [2\langle \hat{\rho}\mathbf{b}\hat{1} \rangle + 2\langle \hat{\rho}\hat{1}\mathbf{b} \rangle] \right),
 \end{aligned}$$

and

$$2\langle \hat{\rho}\mathbf{b}\hat{1} \rangle + 2\langle \hat{\rho}\hat{1}\mathbf{b} \rangle = \mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) - \mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1).$$

We have

$$\begin{aligned}
 -\nabla \cdot \left(\frac{q_{||}^\perp}{\Omega} [2\langle \hat{\rho}\mathbf{b}\hat{1} \rangle + 2\langle \hat{\rho}\hat{1}\mathbf{b} \rangle] \right) &= -\left(\frac{1}{\Omega} \nabla q_{||}^\perp - \frac{q_{||}^\perp}{\Omega} \nabla \log(B) \right) \cdot [\mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) - \mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1)] \\
 &\quad - \frac{q_{||}^\perp}{\Omega} \nabla \cdot [\mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) - \mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1)] \\
 &= \frac{1}{\Omega} \left(\mathbf{b} \times [\nabla q_{||}^\perp - q_{||}^\perp \nabla \log(B)] \right) \mathbf{b} \\
 &\quad - \frac{q_{||}^\perp}{\Omega} \nabla \cdot [\mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) - \mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1)].
 \end{aligned}$$

Here

$$\begin{aligned}
 \nabla \cdot [\mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) - \mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1)] &= \mathbf{b}(\mathbf{b} \times \boldsymbol{\kappa}) + (\mathbf{b} \times \boldsymbol{\kappa})\mathbf{b} \\
 &\quad - 2(\mathbf{b} \cdot \nabla \times \mathbf{b})\mathbf{b}\mathbf{b} + (\mathbf{b} \cdot \nabla \times \mathbf{b})(\mathbf{I} - \mathbf{b}\mathbf{b}) \\
 &\quad + (\nabla_1 b_1 - \nabla_2 b_2)(\mathbf{e}_1 \mathbf{e}_2 + \mathbf{e}_2 \mathbf{e}_1) \\
 &\quad + (\nabla_1 b_2 + \nabla_2 b_1)(\mathbf{e}_2 \mathbf{e}_2 - \mathbf{e}_1 \mathbf{e}_1),
 \end{aligned}$$

Appendix B. Detailed derivation of gyroviscous tensor components

hence

$$\begin{aligned}
-\nabla \cdot d^3v mw \rho_0 \left[\mathbf{b} \dot{\rho}_{gc}^{(0)} + \dot{\rho}_{gc}^{(0)} \mathbf{b} \right] F = & \frac{1}{\Omega} \mathbf{b} \times \left[\nabla q_{||}^\perp - q_{||}^\perp \nabla \log(B) - q_{||}^\perp \boldsymbol{\kappa} \right] \mathbf{b} \\
& - \frac{q_{||}^\perp}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{b}) (\mathbf{I} - \mathbf{b} \mathbf{b}) \\
& - \frac{2q_{||}^\perp}{\Omega} (\mathbf{b} \cdot \nabla \times \mathbf{b}) \mathbf{b} \mathbf{b} \\
& - \frac{q_{||}^\perp}{\Omega} [(\nabla_1 b_1 - \nabla_2 b_2)(\mathbf{e}_1 \mathbf{e}_2 + \mathbf{e}_2 \mathbf{e}_1) \\
& + (\nabla_1 b_2 + \nabla_2 b_1)(\mathbf{e}_2 \mathbf{e}_2 - \mathbf{e}_1 \mathbf{e}_1)].
\end{aligned}$$

Interestingly, the contribution $\nabla \log(B)$ here exactly cancels a similar ∇B contribution from eq. (4.13) which is obtained as a result of the magnetic ∇B drift, so that the final expression is written in terms of the field line curvature $\boldsymbol{\kappa} \equiv \mathbf{b} \cdot \nabla \mathbf{b}$ and not ∇B .

B.5 Identities involving moving frame

We have a moving frame $\mathbf{e}_1, \mathbf{e}_2, \mathbf{b}$, as well as a rotating frame $\hat{\rho}, \hat{\perp}, \mathbf{b}$ (cp. (2.22)), where

$$\begin{aligned}
\hat{\rho} &= \cos(\zeta) \mathbf{e}_1 - \sin(\zeta) \mathbf{e}_2, \\
\hat{\perp} &= \sin(\zeta) \mathbf{e}_1 + \cos(\zeta) \mathbf{e}_2.
\end{aligned}$$

It follows that

$$\langle \hat{\rho} \hat{\perp} \rangle = \frac{1}{2} (\mathbf{e}_1 \mathbf{e}_2 - \mathbf{e}_2 \mathbf{e}_1).$$

and hence

$$\langle \hat{\rho} \mathbf{b} \hat{\perp} \rangle = \frac{1}{2} (\mathbf{e}_1 \mathbf{b} \mathbf{e}_2 - \mathbf{e}_2 \mathbf{b} \mathbf{e}_1), \quad \langle \hat{\rho} \hat{\perp} \mathbf{b} \rangle = \frac{1}{2} (\mathbf{e}_2 \mathbf{b} \mathbf{e}_1 - \mathbf{e}_1 \mathbf{b} \mathbf{e}_2).$$

Then

$$\langle \hat{\rho} \mathbf{b} \hat{\perp} \rangle + \langle \hat{\rho} \hat{\perp} \mathbf{b} \rangle = \frac{1}{2} [\mathbf{e}_1 (\mathbf{e}_2 \mathbf{b} + \mathbf{b} \mathbf{e}_2) - \mathbf{e}_2 (\mathbf{e}_1 \mathbf{b} + \mathbf{b} \mathbf{e}_1)].$$

We compute

$$\begin{aligned}
 \nabla \cdot [\mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2)] &= (\nabla \cdot \mathbf{e}_1)(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) + [(\mathbf{e}_1 \cdot \nabla \mathbf{e}_2)\mathbf{b} + \mathbf{b}(\mathbf{e}_1 \cdot \nabla \mathbf{e}_2)] \\
 &\quad + [\mathbf{e}_2(\mathbf{e}_1 \cdot \nabla \mathbf{b}) + (\mathbf{e}_1 \cdot \nabla \mathbf{b})\mathbf{e}_2] \\
 &= [\mathbf{e}_1 \cdot \nabla \mathbf{e}_2 \cdot \mathbf{e}_1](\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1) \\
 &\quad + [\nabla \cdot \mathbf{e}_1](\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) \\
 &\quad + 2[\mathbf{e}_1 \cdot \nabla \mathbf{e}_2 \cdot \mathbf{b}]\mathbf{b}\mathbf{b} \\
 &\quad + 2[\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2]\mathbf{e}_2\mathbf{e}_2 \\
 &\quad + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1](\mathbf{e}_1\mathbf{e}_2 + \mathbf{e}_2\mathbf{e}_1),
 \end{aligned}$$

and, permuting $1 \leftrightarrow 2$, we find

$$\begin{aligned}
 \nabla \cdot [\mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1)] &= [\mathbf{e}_2 \cdot \nabla \mathbf{e}_1 \cdot \mathbf{e}_2](\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) \\
 &\quad + [\nabla \cdot \mathbf{e}_2](\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1) \\
 &\quad + 2[\mathbf{e}_2 \cdot \nabla \mathbf{e}_1 \cdot \mathbf{b}]\mathbf{b}\mathbf{b} \\
 &\quad + 2[\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1]\mathbf{e}_1\mathbf{e}_1 \\
 &\quad + [\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2](\mathbf{e}_2\mathbf{e}_1 + \mathbf{e}_1\mathbf{e}_2).
 \end{aligned}$$

Thus

$$\begin{aligned}
 \nabla \cdot \{[\mathbf{e}_1(\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2)] - [\mathbf{e}_2(\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1)]\} &= [\mathbf{e}_1 \cdot \nabla \mathbf{e}_2 \cdot \mathbf{e}_1 - \nabla \cdot \mathbf{e}_2](\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1) \\
 &\quad + [\nabla \cdot \mathbf{e}_1 - \mathbf{e}_2 \cdot \nabla \mathbf{e}_1 \cdot \mathbf{e}_2](\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) \\
 &\quad + 2[\mathbf{e}_1 \cdot \nabla \mathbf{e}_2 \cdot \mathbf{b} - \mathbf{e}_2 \cdot \nabla \mathbf{e}_1 \cdot \mathbf{b}]\mathbf{b}\mathbf{b} \\
 &\quad - 2[\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1]\mathbf{e}_1\mathbf{e}_1 \\
 &\quad + 2[\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2]\mathbf{e}_2\mathbf{e}_2 \\
 &\quad + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 - \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2](\mathbf{e}_1\mathbf{e}_2 + \mathbf{e}_2\mathbf{e}_1) \\
 &= [-\mathbf{b} \cdot \nabla \mathbf{e}_2 \cdot \mathbf{b}](\mathbf{e}_1\mathbf{b} + \mathbf{b}\mathbf{e}_1) \\
 &\quad + [\mathbf{b} \cdot \nabla \mathbf{e}_1 \cdot \mathbf{b}](\mathbf{e}_2\mathbf{b} + \mathbf{b}\mathbf{e}_2) \\
 &\quad - 2[\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 - \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1]\mathbf{b}\mathbf{b} \\
 &\quad - 2[\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1]\mathbf{e}_1\mathbf{e}_1 \\
 &\quad + 2[\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2]\mathbf{e}_2\mathbf{e}_2 \\
 &\quad + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 - \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2](\mathbf{e}_1\mathbf{e}_2 + \mathbf{e}_2\mathbf{e}_1).
 \end{aligned}$$

Appendix B. Detailed derivation of gyroviscous tensor components

We can combine the terms proportional to $\mathbf{e}_1 \mathbf{e}_1$ and $\mathbf{e}_2 \mathbf{e}_2$ by observing that

$$\begin{aligned} 2[\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1] &= [\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 + \mathbf{e}_1 \nabla \mathbf{b} \cdot \mathbf{e}_2] + [\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 - \mathbf{e}_1 \nabla \mathbf{b} \cdot \mathbf{e}_2] \\ &= [\mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 + \mathbf{e}_1 \nabla \mathbf{b} \cdot \mathbf{e}_2] - \mathbf{b} \cdot \nabla \times \mathbf{b}, \\ 2[\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2] &= [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 + \mathbf{e}_2 \nabla \mathbf{b} \cdot \mathbf{e}_1] + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 - \mathbf{e}_2 \nabla \mathbf{b} \cdot \mathbf{e}_1] \\ &= [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 + \mathbf{e}_2 \nabla \mathbf{b} \cdot \mathbf{e}_1] + \mathbf{b} \cdot \nabla \times \mathbf{b}. \end{aligned}$$

Therefore,

$$\begin{aligned} \nabla \cdot \{[\mathbf{e}_1(\mathbf{e}_2 \mathbf{b} + \mathbf{b} \mathbf{e}_2)] - [\mathbf{e}_2(\mathbf{e}_1 \mathbf{b} + \mathbf{b} \mathbf{e}_1)]\} \\ = [\boldsymbol{\kappa} \cdot \mathbf{e}_2](\mathbf{e}_1 \mathbf{b} + \mathbf{b} \mathbf{e}_1) + [-\boldsymbol{\kappa} \cdot \mathbf{e}_1](\mathbf{e}_2 \mathbf{b} + \mathbf{b} \mathbf{e}_2) \\ - 2[\mathbf{b} \cdot \nabla \times \mathbf{b}] \mathbf{b} \mathbf{b} + [\mathbf{b} \cdot \nabla \times \mathbf{b}] (\mathbf{e}_1 \mathbf{e}_1 + \mathbf{e}_2 \mathbf{e}_2) \\ + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 + \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1] (\mathbf{e}_2 \mathbf{e}_2 - \mathbf{e}_1 \mathbf{e}_1) \\ + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 - \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2] (\mathbf{e}_1 \mathbf{e}_2 + \mathbf{e}_2 \mathbf{e}_1). \end{aligned}$$

Finally, we observe that

$$\begin{aligned} [\boldsymbol{\kappa} \cdot \mathbf{e}_2](\mathbf{e}_1 \mathbf{b}) + [-\boldsymbol{\kappa} \cdot \mathbf{e}_1](\mathbf{e}_2 \mathbf{b}) &= [(\boldsymbol{\kappa} \cdot \mathbf{e}_2)\mathbf{e}_1 - (\boldsymbol{\kappa} \cdot \mathbf{e}_1)\mathbf{e}_2] \mathbf{b} \\ &= [(\boldsymbol{\kappa} \times \mathbf{b}) \cdot \mathbf{e}_1 \mathbf{e}_1 + (\boldsymbol{\kappa} \times \mathbf{b}) \cdot \mathbf{e}_2 \mathbf{e}_2] \mathbf{b} \\ &= (\boldsymbol{\kappa} \times \mathbf{b}) \mathbf{b}, \end{aligned}$$

so

$$\begin{aligned} \nabla \cdot \{(\hat{\rho} \mathbf{b} \hat{\perp}) + (\hat{\rho} \hat{\perp} \mathbf{b})\} &= \nabla \cdot \{[\mathbf{e}_1(\mathbf{e}_2 \mathbf{b} + \mathbf{b} \mathbf{e}_2)] - [\mathbf{e}_2(\mathbf{e}_1 \mathbf{b} + \mathbf{b} \mathbf{e}_1)]\} \\ &= [(\boldsymbol{\kappa} \times \mathbf{b}) \mathbf{b} + \mathbf{b} (\boldsymbol{\kappa} \times \mathbf{b})] \\ &\quad - 2[\mathbf{b} \cdot \nabla \times \mathbf{b}] \mathbf{b} \mathbf{b} + [\mathbf{b} \cdot \nabla \times \mathbf{b}] (\mathbf{e}_1 \mathbf{e}_1 + \mathbf{e}_2 \mathbf{e}_2) \\ &\quad + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2 + \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1] (\mathbf{e}_2 \mathbf{e}_2 - \mathbf{e}_1 \mathbf{e}_1) \\ &\quad + [\mathbf{e}_1 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_1 - \mathbf{e}_2 \cdot \nabla \mathbf{b} \cdot \mathbf{e}_2] (\mathbf{e}_1 \mathbf{e}_2 + \mathbf{e}_2 \mathbf{e}_1). \end{aligned}$$

Bibliography

- [AB93] T. M. Antonsen and A. Bondeson. Effects of trapped thermal particles on the n=1 internal kink mode in tokamaks. *Phys. Rev. Lett.*, 71:2046–2049, 1993.
- [ABC⁺78] G Ara, B Basu, B Coppi, G Laval, MN Rosenbluth, and BV Waddell. Magnetic reconnection and m=1 oscillations in current carrying plasmas. *Annals of Physics*, 112(2):443–476, 1978.
- [AGB⁺15] O. Asunta, J. Govenius, R. Budny, M. Gorelenkova, G. Tardini, T. Kurki-Suonio, A. Salmi, and S. Sipilä. Modelling neutral beams in fusion devices: Beamlet-based model for fast particle simulations. *Computer Physics Communications*, 188:33 – 46, 2015.
- [Aib16] N Aiba. Impact of ion diamagnetic drift on ideal ballooning mode stability in rotating tokamak plasmas. *Plasma Physics and Controlled Fusion*, 58(4):045020, 2016.
- [Alb11] M. Albergante. *Interaction between fast ions and microturbulence in thermonuclear devices*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2011.
- [AM87] R. Abraham and J.E. Marsden. *Foundations of mechanics, 2nd ed.* Addison-Wesley, 1987.
- [Ant82] T. M. Antonsen. Electrostatic modification of variational principles for anisotropic plasmas. *Physics of Fluids*, 25(1):132, 1982.
- [Arn78] V.I. Arnold. *Mathematical methods of classical mechanics*. Graduate Texts in Mathematics, Springer, 1978.
- [AT94] M. Artun and W. M. Tang. Nonlinear electromagnetic gyrokinetic equations for rotating axisymmetric plasmas. *Physics of Plasmas*, 1(8):2682, 1994.
- [Ban67] A. Banos. The guiding centre approximation in lowest order. *Journal of Plasma Physics*, 1(8):305 – 3016, 1967.
- [Bel01] E. V. Belova. Nonlinear gyroviscous force in a collisionless plasma. *Physics of Plasmas*, 8(9):3936–3944, 2001.

Bibliography

- [Bel13] E V Belova. Nonlinear gyroviscous force in a collisionless plasma Nonlinear gyroviscous force in a collisionless plasma. 3936(2001), 2013.
- [BGC03] E. V. Belova, N. N. Gorelenkov, and C. Z. Cheng. Self-consistent equilibrium model of low aspect-ratio toroidal plasma with energetic beam ions. *Physics of Plasmas*, 10(8):3240–3251, 2003.
- [BH07] A. J. Brizard and T. S. Hahm. Foundations of nonlinear gyrokinetic theory. *Rev. Mod. Phys.*, 79:421–468, Apr 2007.
- [BJJ⁺00] E. V. Belova, S. C. Jardin, H. Ji, M. Yamada, and R. Kulsrud. Numerical study of tilt stability of prolate field-reversed configurations. *Physics of Plasmas*, 7(12):4996–5006, 2000.
- [BK99] Duarte Borba and Wolfgang Kerner. Castor-k: Stability analysis of alfvén eigenmodes in the presence of energetic ions in tokamaks. *Journal of Computational Physics*, 153(1):101 – 138, 1999.
- [BPES75] M. N. Bussac, R. Pellat, D. Edery, and J. L. Soule. Internal kink modes in toroidal plasmas with circular cross sections. *Phys. Rev. Lett.*, 35:1638–1641, Dec 1975.
- [Bra65] S. I. Braginskii. Transport processes in a plasma. *Reviews of Plasma Physics*, 1, 1965.
- [Bri92] A. J. Brizard. Nonlinear gyrofluid description of turbulent magnetized plasmas. *Physics of Fluids B*, 4(5):1213–1228, 1992.
- [Bri95] Alain J. Brizard. Nonlinear gyrokinetic Vlasov equation for toroidally rotating axisymmetric tokamaks. *Physics of Plasmas*, 2(May 2014):459, 1995.
- [Bri01] Alain J Brizard. A geometric view of hamiltonian perturbation theory. *Physics Letters A*, 291(2):146 – 149, 2001.
- [Bri05] Alain J. Brizard. Energy-conserving Finite-beta Electromagnetic Drift-fluid Equations. pages 1–24, 2005.
- [Bri10] A. J. Brizard. Guiding-center polarization and magnetization effects in gyrokinetic theory. *arXiv e-prints*, page arXiv:1008.3888, Aug 2010.
- [BSQ13] J. W. Burby, J. Squire, and H. Qin. Automation of the guiding center expansion. *Physics of Plasmas*, 20(7):072105, 2013.
- [Car81] John R. Cary. Lie transform perturbation theory for hamiltonian systems. *Physics Reports*, 79(2):129 – 159, 1981.
- [CB09] John R. Cary and Alain J. Brizard. Hamiltonian theory of guiding-center motion. *Rev. Mod. Phys.*, 81:693–738, May 2009.

- [CBF⁺15] W.A. Cooper, D. Brunetti, J.M. Faustin, J.P. Graves, D. Pfefferl , M. Raghunathan, O. Sauter, T.M. Tran, I.T. Chapman, C.J. Ham, N. Aiba, and and. Free boundary equilibrium in 3d tokamaks with toroidal rotation. *Nuclear Fusion*, 55(6):063032, may 2015.
- [CBT87] P. J. Catto, I. B. Bernstein, and M. Tessarotto. Ion transport in toroidally rotating tokamak plasmas. *The Physics of Fluids*, 30(9):2784–2795, 1987.
- [CC92] Z. Chang and J. D. Callen. Generalized gyroviscous force and its effect on the momentum balance equation. *Physics of Fluids B: Plasma Physics*, 4(7):1766–1771, 1992.
- [CG08] C Z Cheng and N N Gorelenkov. Trapped electron stabilization of ballooning modes in low aspect ratio toroidal plasmas. 4784(2004), 2008.
- [CGL56] G.F. Chew, M.L. Goldberger, and F.E. Low. The Boltzmann Equation and the One-Fluid Hydromagnetic Equations in the Absence of Particle Collisions. *Proc. R. Soc. London (London)*, A236:112, 1956.
- [CGL⁺15] I. T. Chapman, J. P. Graves, M. Lennholm, J. Faustin, E. Lerche, T. Johnson, and S. Tholerus. The merits of ion cyclotron resonance heating schemes for sawtooth control in tokamak plasmas. *Journal of Plasma Physics*, 81(6):365810601, 2015.
- [CGP⁺10] W. A. Cooper, J. P. Graves, A. Pochebon, O. Sauter, and L. Villard. Tokamak magnetohydrodynamic equilibrium states with axisymmetric boundary and a 3D helical core. *Phys. Rev. Lett.*, 105:035003, Jul 2010.
- [CH87] W A Cooper and S P Hirshman. Axisymmetric MHD equilibria with isothermal toroidal mass flow by variational steepest descent moments method. *Plasma Physics and Controlled Fusion*, 29(7):933–943, jul 1987.
- [CIG⁺09] I.T. Chapman, V.G. Igochine, J.P. Graves, S.D. Pinches, A. Gude, I. Jenkins, M. Maraschek, G. Tardini, and and. Sawtooth control and the interaction of energetic particles. *Nuclear Fusion*, 49(3):035006, feb 2009.
- [CL83] John R Cary and Robert G Littlejohn. Noncanonical hamiltonian mechanics and its application to magnetic field line flow. *Annals of Physics*, 151(1):1 – 34, 1983.
- [CMP88] B Coppi, Stefano Migliuolo, and F Porcelli. Macroscopic plasma oscillation bursts (fishbones) resulting from high-energy populations. *The Physics of fluids*, 31(6):1630–1648, 1988.
- [CP86] B Coppi and F Porcelli. Theoretical model of fishbone oscillations in magnetically confined plasmas. *Physical review letters*, 57(18):2272, 1986.
- [CPG⁺07] I T Chapman, S D Pinches, J P Graves, R J Akers, L C Appel, R V Budny, S Coda, N J Conway, M de Bock, L-G Eriksson, R J Hastie, T C Hender, G T A Huysmans,

Bibliography

- T Johnson, H R Koslowski, A Krämer-Flecken, M Lennholm, Y Liang, S Saarelma, S E Sharapov, and I Voitsekhovitch and. The physics of sawtooth stabilization. *Plasma Physics and Controlled Fusion*, 49(12B):B385–B394, nov 2007.
- [CSHM06] I. T. Chapman, S. E. Sharapov, G. T. A. Huysmans, and A. B. Mikhailovskii. Modeling the effect of toroidal plasma rotation on drift-magnetohydrodynamic modes in tokamaks. *Physics of Plasmas*, 13(6):062511, 2006.
- [CWGW11] IT Chapman, NR Walkden, JP Graves, and Christer Wahlberg. The effects of sheared toroidal rotation on stability limits in tokamak plasmas. *Plasma Physics and Controlled Fusion*, 53(12):125002, 2011.
- [CWR84] Liu Chen, RB White, and MN Rosenbluth. Excitation of internal kink modes by trapped energetic beam ions. *Physical Review Letters*, 52(13):1122, 1984.
- [CZ16] Liu Chen and Fulvio Zonca. Physics of alfvén waves and energetic particles in burning plasmas. *Reviews of Modern Physics*, 88(1):015008, 2016.
- [DKOL83] D. Dubin, J. A. Krommes, C. Oberman, and W. Lee. Nonlinear gyrokinetic equations. *The Physics of fluids*, 26(12):3524–3535, 1983.
- [FBC⁺16] A Fasoli, S Brunner, WA Cooper, JP Graves, P Ricci, O Sauter, and L Villard. Computational challenges in magnetic-confinement fusion physics. *Nature Physics*, 12(5):411–423, 2016.
- [FR60] E Frieman and Manuel Rotenberg. On hydromagnetic stability of stationary equilibria. *Reviews of Modern Physics*, 32(4):898, 1960.
- [Fre07] Jeffrey P. Freidberg. *Plasma Physics and Fusion Energy*. Cambridge University Press, 2007.
- [Fre14] Jeffrey P. Freidberg. *Ideal MHD*. Cambridge University Press, 2014.
- [GCC⁺09] J. P. Graves, I. Chapman, S. Coda, L.-G. Eriksson, and T. Johnson. Sawtooth-control mechanism using toroidally propagating ion-cyclotron-resonance waves in tokamaks. *Phys. Rev. Lett.*, 102:065005, Feb 2009.
- [GGJ75] AH Glasser, JM Greene, and JL Johnson. Resistive instabilities in general toroidal plasma configurations. *The Physics of Fluids*, 18(7):875–888, 1975.
- [GHH00] J P Graves, R J Hastie, and K I Hopcraft. The effects of sheared toroidal plasma rotation on the internal kink mode in the banana regime. *Plasma Physics and Controlled Fusion*, 42(10):1049–1066, oct 2000.
- [GKP10] J. P. Goedbloed, Rony Keppens, and Stefaan Poedts. *Advanced Magnetohydrodynamics: With Applications to Laboratory and Astrophysical Plasmas*. Cambridge University Press, 2010.

- [GLB⁺11] T. Görler, X. Lapillonne, S. Brunner, T. Dannert, F. Jenko, F. Merz, and D. Told. The global version of the gyrokinetic turbulence code gene. *Journal of Computational Physics*, 230(18):7053 – 7071, 2011.
- [GP04] J. P. Hans Goedbloed and Stefaan Poedts. *Principles of Magnetohydrodynamics: With Applications to Laboratory and Astrophysical Plasmas*. Cambridge University Press, 2004.
- [GR58] Harold Grad and Hanan Rubin. Hydromagnetic equilibria and force-free fields. *Journal of Nuclear Energy (1954)*, 7(3-4):284–285, 1958.
- [GR85] R Gruber and J Rappaz. Finite element methods in linear ideal magnetohydrodynamics. *Berlin and New York, Springer-Verlag*, 1985, 189 p., 1985.
- [Gra99] J. P. Graves. *Kinetic stabilisation of the internal kink mode for fusion plasmas*. PhD thesis, University of Nottingham, School of Mathematical Sciences, 1999.
- [Gra04] Jonathan P. Graves. Influence of asymmetric energetic ion distributions on sawtooth stabilization. *Phys. Rev. Lett.*, 92:185003, May 2004.
- [Gra05] Jonathan P. Graves. Internal kink mode stabilization and the properties of auxiliary heated ions. *Physics of Plasmas*, 12(9):090908, 2005.
- [Gra13] Jonathan P Graves. Toroidal drift precession and wave-particle interaction in shaped tokamaks with finite beta and neoclassical equilibrium effects. *Plasma Physics and Controlled Fusion*, 55(7):074009, jun 2013.
- [GSL⁺07] C Gormezano, A.C.C Sips, T.C Luce, S Ide, A Becoulet, X Litaudon, A Isayama, J Hobirk, M.R Wade, T Oikawa, R Prater, A Zvonkov, B Lloyd, T Suzuki, E Barbato, P Bonoli, C.K Phillips, VVdovin, E Joffrin, T Casper, J Ferron, D Mazon, D Moreau, R Bundy, C Kessel, A Fukuyama, N Hayashi, F Imbeaux, M Murakami, A.R Polevoi, and H.E. St John. Chapter 6: Steady state operation. *Nuclear Fusion*, 47(6):S285–S336, jun 2007.
- [GW17] J P Graves and C Wahlberg. Generalised zonal modes in stationary axisymmetric plasmas. *Plasma Physics and Controlled Fusion*, 59(5):054011, apr 2017.
- [GZB⁺19] J P Graves, D Zullino, D Brunetti, S Lanthaler, and C Wahlberg. Reduced models for parallel magnetic field fluctuations and their impact on pressure gradient driven MHD instabilities in axisymmetric toroidal plasmas. *Plasma Physics and Controlled Fusion*, 61(10):104003, aug 2019.
- [HAK⁺14] E. Hirvijoki, O. Asunta, T. Koskela, T. Kurki-Suonio, J. Miettunen, S. Sipilä, A. Snicker, and S. Äkäslompolo. ASCOT: Solving the kinetic equation of minority particle species in tokamak plasmas. *Computer Physics Communications*, 185(4):1310 – 1321, 2014.

Bibliography

- [HdBK12] JW Haverkort, Hugo J de Blank, and Barry Koren. The brunt–väisälä frequency of rotating tokamak plasmas. *Journal of Computational Physics*, 231(3):981–1001, 2012.
- [HGHM97] P. Helander, C. G. Gimblett, R. J. Hastie, and K. G. McClements. The influence of fast ions on the magnetohydrodynamic stability of negative shear profiles. *Physics of Plasmas*, 4(6):2181–2187, 1997.
- [HH88] R.J. Hastie and T.C. Hender. Toroidal internal kink stability in tokamaks with ultra flat q profiles. *Nuclear Fusion*, 28(4):585–594, apr 1988.
- [HM85] R. D. Hazeltine and J. D. Meiss. Shear-Alfvén dynamics of toroidally confined plasmas. *Physics Reports*, 121(1-2):1–164, 1985.
- [HM03] R. D. Hazeltine and J. D. Meiss. *Plasma confinement*. Dover Publications, Mineola, N.Y, 2003.
- [HWB⁺07] T.C Hender, J.C Wesley, J Bialek, A Bondeson, A.H Boozer, R.J Buttery, A Garofalo, T.P Goodman, R.S Granetz, Y Gribov, O Gruber, M Gryaznevich, G Giruzzi, S Günter, N Hayashi, P Helander, C.C Hegna, D.F Howell, D.A Humphreys, G.T.A Huysmans, A.W Hyatt, A Isayama, S.C Jardin, Y Kawano, A Kellman, C Kessel, H.R Koslowski, R.J. La Haye, E Lazzaro, Y.Q Liu, V Lukash, J Manickam, S Medvedev, V Mertens, S.V Mirnov, Y Nakamura, G Navratil, M Okabayashi, T Ozeki, R Paccagnella, G Pautasso, F Porcelli, V.D Pustovitov, V Riccardo, M Sato, O Sauter, M.J Schaffer, M Shimada, P Sonato, E.J Strait, M Sugihara, M Takechi, A.D Turnbull, E Westerhof, D.G Whyte, R Yoshino, H Zohm, Disruption the ITPA MHD, and Magnet Group. Chapter 3: MHD stability, operational limits and disruptions. *Nuclear Fusion*, 47(6):S128–S202, jun 2007.
- [J⁺99] J Jacquinot et al. Deuterium-tritium operation in magnetic confinement experiments: results and underlying physics. *Plasma physics and controlled fusion*, 41(3A):A13, 1999.
- [JBA⁺07] S. Jolliet, A. Bottino, P. Angelino, R. Hatzky, T.M. Tran, B.F. Mcmillan, O. Sauter, K. Appert, Y. Idomura, and L. Villard. A global collisionless pic code in magnetic coordinates. *Computer Physics Communications*, 177(5):409 – 425, 2007.
- [Kau60] A. N. Kaufman. Plasma Viscosity in a Magnetic Field. *Phys. Fluids*, 3(4):610, 1960.
- [KGG⁺99] M Keilhacker, A Gibson, C Gormezano, PJ Lomas, PR Thomas, ML Watkins, P Andrew, B Balet, D Borba, CD Challis, et al. High fusion performance from deuterium-tritium plasmas in jet. *Nuclear Fusion*, 39(2):209, 1999.
- [KO58] M. D. Kruskal and C. R. Oberman. On the stability of plasma in static equilibrium. *The Physics of Fluids*, 1(4):275–280, 1958.

- [KSAH⁺09] T. Kurki-Suonio, O. Asunta, T. Hellsten, V. Hynönen, T. Johnson, T. Koskela, J. Lönnroth, V. Parail, M. Roccella, G. Saibene, A. Salmi, and S. Sipilä. ASCOT simulations of fast ion power loads to the plasma-facing components in ITER. *Nuclear Fusion*, 49(9):095001, 2009.
- [Kul83] R. M. Kulsrud. volume I, pages 115–146. North-Holland, Amsterdam, 1983.
- [LAA⁺07] B Lloyd, RJ Akers, F Alladio, Y Andrew, LC Appel, D Applegate, KB Axon, N Ben Ayed, C Bunting, RJ Buttery, et al. Overview of physics results from mast. *Nuclear fusion*, 47(10):S658, 2007.
- [LCCH08] Yueqiang Liu, Ming-Sheng Chu, IT Chapman, and TC Hender. Toroidal self-consistent modeling of drift kinetic effects on the resistive wall mode. *Physics of plasmas*, 15(11):112503, 2008.
- [LGKP07] Ph Lauber, Sibylle Günter, Axel Könies, and Simon D Pinches. Ligka: A linear gyrokinetic code for the description of background kinetic and fast particle effects on the mhd stability in tokamaks. *Journal of Computational Physics*, 226(1):447–465, 2007.
- [LGPC19] S Lanthaler, J P Graves, D Pfefferl  , and W A Cooper. Guiding-centre theory for kinetic-magnetohydrodynamic modes in strongly flowing plasmas. *Plasma Physics and Controlled Fusion*, 61(7):074006, may 2019.
- [Lit82] Robert G. Littlejohn. Hamiltonian perturbation theory in noncanonical coordinates. *Journal of Mathematical Physics*, 23(5):742–747, 1982.
- [Lit83] Robert G. Littlejohn. Variational principles of guiding centre motion. *Journal of Plasma Physics*, 29(1):111–125, 1983.
- [Lit84] Robert G. Littlejohn. Geometry and guiding center motion. *Contemporary Mathematics*, 28:151–167, 1984.
- [Lit88] Robert G. Littlejohn. Phase anholonomy in the classical adiabatic motion of charged particles. *Phys. Rev. A*, 38:6034–6045, Dec 1988.
- [LL08] H. Lütjens and J.-F. Luciani. The XTOR code for nonlinear 3D simulations of MHD instabilities in tokamak plasmas. *Journal of Computational Physics*, 227(14):6944 – 6966, 2008.
- [LL10] H. Lütjens and J.-F. Luciani. XTOR-2F: A fully implicit Newton–Krylov solver applied to nonlinear 3D extended MHD in tokamaks. *Journal of Computational Physics*, 229(21):8130 – 8143, 2010.
- [LPGC17] S Lanthaler, D Pfefferl  , J P Graves, and W A Cooper. Higher order larmor radius corrections to guiding-centre equations and application to fast ion equilibrium distributions. *Plasma Physics and Controlled Fusion*, 59(4):044014, mar 2017.

Bibliography

- [LS57] R Lüst and A Schlüter. Axialsymmetrische magnetohydrodynamische gleichgewichtskonfigurationen. *Zeitschrift für Naturforschung A*, 12(10):850–854, 1957.
- [Mac65] A. Macmahon. Finite Gyro-Radius Corrections to the Hydromagnetic Equations for a Vlasov Plasma. *Physics of Fluids*, 8(10):1840, 1965.
- [Mad10] J. Madsen. Second order guiding-center vlasov-maxwell equations. *Physics of Plasmas*, 17(8):082107, 2010.
- [MBF⁺05] J.E Menard, R.E Bell, E.D Fredrickson, D.A Gates, S.M Kaye, B.P LeBlanc, R Maingi, S.S Medley, W Park, S.A Sabbagh, A Sontag, D Stutman, K Tritz, W Zhu, and the NSTX Research Team. Internal kink mode dynamics in high- β NSTX plasmas. *Nuclear Fusion*, 45(7):539–556, jun 2005.
- [MGB⁺83] K McGuire, R Goldston, M Bell, M Bitter, K Bol, K Brau, D Buchenauer, T Crowley, S Davis, F Dylla, et al. Study of high-beta magnetohydrodynamic modes and fast-ion losses in pdx. *Physical Review Letters*, 50(12):891, 1983.
- [Mig93] S Migliuolo. Theory of ideal and resistive m= 1 modes in tokamaks. *Nuclear fusion*, 33(11):1721, 1993.
- [MP80] EK Maschke and H Perrin. Exact solutions of the stationary mhd equations for a rotating toroidal plasma. *Plasma Physics*, 22(6):579, 1980.
- [MS11] N. Miyato and B. D. Scott. Fluid moments in the reduced model for plasmas with large flow velocity. *Plasma and Fusion Research*, 6(2011):1–11, 2011.
- [MT71] A. B. Mikhailovskii and V. S. Tsypin. Transport equations and gradient instabilities in a high pressure collisional plasma. *Plasma Physics*, 13(9):785–798, sep 1971.
- [MYS15] N Miyato, M Yagi, and B D Scott. On push-forward representations in the standard gyrokinetic model On push-forward representations in the standard gyrokinetic model. 012103, 2015.
- [NBC⁺15] F. Nabais, D. Borba, R. Coelho, A. Figueiredo, J. Ferreira, N. Loureiro, and P. Rodrigues. The CASTOR-k code, recent developments and applications. *Plasma Science and Technology*, 17(2):89–96, feb 2015.
- [NCJ⁺91] MFF Nave, DJ Campbell, E Joffrin, FB Marcus, G Sadler, P Smeulders, and K Thomsen. Fishbone activity in jet. *Nuclear Fusion*, 31(4):697, 1991.
- [OKWZ16] J Ongena, R Koch, R Wolf, and H Zohm. Magnetic-confinement fusion. *Nature Physics*, 12(5):398–410, 2016.
- [PCFG16] D. Pfefferlé, W.A. Cooper, A. Fasoli, and J.P. Graves. Effects of magnetic ripple on 3D equilibrium and alpha particle confinement in the european DEMO. *Nuclear Fusion*, 56(11):112002, 2016.

- [PCGM14] D. Pfefferl , W.A. Cooper, J.P. Graves, and C. Misev. VENUS-LEVIS and its spline-fourier interpolation of 3D toroidal magnetic field representation for guiding-centre and full-orbit simulations of charged energetic particles. *Computer Physics Communications*, 185(12):3127 – 3140, 2014.
- [Pfe15] D. Pfefferl . *Energetic ion dynamics and confinement in 3D saturated MHD configurations*. PhD thesis, cole Polytechnique F d rale de Lausanne, 2015.
- [PGC⁺14] D. Pfefferl , J.P. Graves, W.A. Cooper, C. Misev, I.T. Chapman, M. Turnyanskiy, and S. Sangaroon. NBI fast ion confinement in the helical core of MAST hybrid-like plasmas. *Nuclear Fusion*, 54(6):064020, 2014.
- [PGC15] D Pfefferl , J P Graves, and W A Cooper. Hybrid guiding-centre/full-orbit simulations in non-axisymmetric magnetic geometry exploiting general criterion for guiding-centre accuracy. *Plasma Physics and Controlled Fusion*, 57(5):054017, 2015.
- [PMA⁺04] Alexei Pankin, Douglas McCune, Robert Andre, Glenn Bateman, and Arnold Kritz. The tokamak monte carlo fast ion module NUBEAM in the national transport code collaboration library. *Computer Physics Communications*, 159(3):157 – 184, 2004.
- [PMCG15] David Pfefferl , Cyril Misev, Wilfred A. Cooper, and Jonathan P. Graves. Impact of RMP magnetic field simulation models on fast ion losses. *Nuclear Fusion*, 55(1):012001, 2015.
- [PMP⁺07] E.J. Doyle (Chair Transport Physics), W.A. Houlberg (Chair Confinement Da Modelling), Y. Kamada (Chair Pedestal, Edge), V. Mukhovatov (co Chair Transport Physics), T.H. Osborne (co Chair Pedestal, Edge), A. Polevoi (co Chair Confinement Da Modelling), G Bateman, J.W Connor, J.G. Cordey (retired), T Fujita, X Garbet, T.S Hahm, L.D Horton, A.E Hubbard, F Imbeaux, F Jenko, J.E Kinsey, Y Kishimoto, J Li, T.C Luce, Y Martin, M Ossipenko, V Parail, A Peeters, T.L Rhodes, J.E Rice, C.M Roach, V Rozhansky, F Ryter, G Saibene, R Sartori, A.C.C Sips, J.A Snipes, M Sugihara, E.J Synakowski, H Takenaga, T Takizuka, K Thomsen, M.R Wade, H.R Wilson, ITPA Transport Physics Topical Group, ITPA Confinement Database, Model Group, ITPA Pedestal, and Edge Topical Group. Chapter 2: Plasma confinement and transport. *Nuclear Fusion*, 47(6):S18–S127, jun 2007.
- [PSBZ92] F. Porcelli, R. Stankiewicz, H. L. Berk, and Y. Z. Zhang. Internal kink stabilization by high-energy ions with nonstandard orbits. *Physics of Fluids B: Plasma Physics*, 4(10):3017–3023, 1992.
- [PSK94] F Porcelli, R Stankiewicz, and W Kerner. Solution of the drift-kinetic equation for global plasma modes and finite particle orbit widths. 470, 1994.

Bibliography

- [QTRL00] H. Qin, W. M. Tang, G. Rewoldt, and W. W. Lee. On the gyrokinetic equilibrium. *Physics of Plasmas*, 7(3):991–1000, 2000.
- [Ram05a] J. J. Ramos. Fluid formalism for collisionless magnetized plasmas. *Physics of Plasmas*, 12(5):1–14, 2005.
- [Ram05b] J. J. Ramos. General expression of the gyroviscous force. *Physics of Plasmas*, 12(11):1–7, 2005.
- [RDR73] Marshall N Rosenbluth, RY Dagazian, and PH Rutherford. Nonlinear properties of the internal $m=1$ kink instability in the cylindrical tokamak. *The Physics of Fluids*, 16(11):1894–1902, 1973.
- [Reb92] PH Rebut. The jet preliminary tritium experiment. *Plasma Physics and Controlled Fusion*, 34(13):1749, 1992.
- [RR59] MN Rosenbluth and Norman Rostoker. Theoretical structure of plasma equations. *The Physics of Fluids*, 2(1):23–30, 1959.
- [RT62] K. V. Roberts and J. B. Taylor. Magnetohydrodynamic equations for finite larmor radius. *Phys. Rev. Lett.*, 8:197–198, Mar 1962.
- [SCM⁺07] M Shimada, D.J Campbell, V Mukhovatov, M Fujiwara, N Kirneva, K Lackner, M Nagami, V.D Pustovitov, N Uckan, J Wesley, N Asakura, A.E Costley, A.J.H Donné, E.J Doyle, A Fasoli, C Gormezano, Y Gribov, O Gruber, T.C Hender, W Houlberg, S Ide, Y Kamada, A Leonard, B Lipschultz, A Loarte, K Miyamoto, V Mukhovatov, T.H Osborne, A Polevoi, and A.C.C Sips. Chapter 1: Overview and summary. *Nuclear Fusion*, 47(6):S1–S17, jun 2007.
- [Sco07] B. D. Scott. Tokamak edge turbulence: background theory and computation. *Plasma Physics and Controlled Fusion*, 49(7):S25–S41, jun 2007.
- [SGM⁺05] E Strumberger, S Günter, P Merkel, S Riondato, E Schwarz, C Tichmann, and H.P Zehrfeld. Numerical MHD stability studies: toroidal rotation, viscosity, resistive walls and current holes. *Nuclear Fusion*, 45(9):1156–1167, sep 2005.
- [Sha58] VD Shafranov. On magnetohydrodynamical equilibrium configurations. *Soviet Physics JETP*, 6(3):1013, 1958.
- [Sha70] VD Shafranov. Hydromagnetic stability of a current-carrying pinch in a strong longitudinal magnetic field. *Soviet Physics Technical Physics*, 15:175, 1970.
- [SKSS10] A. Snicker, T. Kurki-Suonio, and S. K. Sipila. Realistic simulations of fast-ion wall distribution including effects due to finite larmor radius. *IEEE Transactions on Plasma Science*, 38(9):2177–2184, Sept 2010.
- [Smo98] A. I. Smolyakov. Gyroviscous forces in a collisionless plasma with temperature gradients. *Can. J. Phys.*, 76:321–331, 1998.

- [Spi52] L. Spitzer. Equations of motion for an ideal plasma. *Astrophys. J.*, 116:299, 1952.
- [TCH80] W Tang, J Connor, and R Hastie. Kinetic-balloonning-mode theory in general geometry KINETIC-BALLOONING-MODE THEORY IN GENERAL GEOMETRY. 1439, 1980.
- [WAA⁺17] R Wenninger, R Albanese, R Ambrosino, F Arbeiter, J Aubert, C Bachmann, L Barbato, T Barrett, M Beckers, W Biel, et al. The demo wall load challenge. *Nuclear Fusion*, 57(4):046002, 2017.
- [Wae96] F. L. Waelbroeck. Gyroscopic stabilization of the internal kink mode. *Physics of Plasmas*, 3(3):1047–1053, 1996.
- [Wah98] C. Wahlberg. Analytical stability condition for the ideal $m=n=1$ kink mode in a toroidal plasma with elliptic cross section. *Physics of Plasmas*, 5(5):1387–1402, 1998.
- [WB97] C. WAHLBERG and A. BONDESON. Stability analysis of the ideal $m=n=1$ kink mode in toroidal geometry by direct expansion of the hydromagnetic equations. *Journal of Plasma Physics*, 57(2):327–341, 1997.
- [WB00] C. Wahlberg and A. Bondeson. Stabilization of the internal kink mode in a tokamak by toroidal plasma rotation. *Physics of Plasmas*, 7(3):923–930, 2000.
- [WB01] C Wahlberg and A Bondeson. Stabilization of the mercier modes in a tokamak by toroidal plasma rotation. *Physics of Plasmas*, 8(8):3595–3604, 2001.
- [WCG09] C. Wahlberg, I. T. Chapman, and J. P. Graves. Importance of centrifugal effects for the internal kink mode stability in toroidally rotating tokamak plasmas. *Physics of Plasmas*, 16(11):112512, 2009.
- [WG16] Christer Wahlberg and JP Graves. Magnetohydrodynamic theory of the global structure and magnetic components of the geodesic acoustic continuum modes in tokamaks. *Plasma Physics and Controlled Fusion*, 58(7):075014, 2016.
- [WGC13] Christer Wahlberg, JP Graves, and IT Chapman. Analysis of global hydromagnetic instabilities driven by strongly sheared toroidal flows in tokamak plasmas. *Plasma Physics and Controlled Fusion*, 55(10):105004, 2013.
- [ZC14] Fulvio Zonca and Liu Chen. Theory on excitations of drift alfvén waves by energetic particles. ii. the general fishbone-like dispersion relation. *Physics of Plasmas*, 21(7):072121, 2014.

Acknowledgements

This thesis is the product of an extremely enriching experience over the last few years, both intellectually and personally. This work has not been written in isolation, and I am indebted to my mentors, colleagues and friends for their many contributions during this time.

First and foremost, I would like to express my deepest gratitude to my supervisor Jon, for his guidance, for sharing his deep knowledge of the subject, for the many enlightening discussions and all the good advice given along the way. Without his constant encouragement and enthusiasm, this thesis could not have been written.

I would also like to thank the administration at SPC for providing the care-free environment that I have been allowed to experience, for the opportunity to attend several conferences and for keeping the bureaucratic chores to a minimum. In particular, I would like to acknowledge the competent support of the secretaries, Edith, Roxanne, Thushi and Christine, who have always welcomed me with a friendly smile. Similarly, the kind assistance of Pierre Etienne with IT-related concerns was very much appreciated.

It was a great pleasure to meet the many scientists at SPC, always happy to share their insights and to provide the opportunity to learn more about the various facets of plasma physics. I would in particular like to acknowledge insightful discussions and comments from Stephan Brunner, Laurent Villard and Paolo Ricci over the years, which have helped shape this thesis.

I would also like to thank the MHD and fast particle group, for contributing to the stimulating and cheerful environment I have experienced in Lausanne. Special thanks to David for providing the initial impetus for this work and his observant remarks, to Jonathan for his help and incredible patience with many coding and compiling aspects, and to Timothée for his insightful advice. Furthermore, I would like to thank Madhu for many discussions as well as the occasional hike, Eduardo for introducing me to Portuguese traditions and culture, and Hamish and Andreas for their company during various parts of my journey at SPC. I also wish the new generation, Guillermo and Mike, all the best.

My sincere appreciation is also extended to my office neighbours Julien and Paola. It has always been a delight to visit your office for a chat, ask your advice and discuss elements of gyrokinetic theory or the latest numerical problems.

Acknowledgements

I would like to thank all doctoral students at SPC, old and new, for making this time also personally rewarding and for providing the needed balance, usually in the form of the occasional excursion (to SAT or otherwise). In particular, I would like to thank Jérémie, Mengdi, Matteo and Marta, Himank and Isha for their friendship and hospitality.

I am indebted to my family for their support throughout my studies, and for their constant encouragement. I would also like to extend my thanks to my new family overseas for their always warm reception, their thoughts and kind advice.

Words cannot express the gratitude I feel to my wife, Dahye, who has been my rock throughout these years. Without her loving support, this thesis could not have been written.

Zürich, 22 March 2020

S. L.

CURRICULUM VITAE

PERSONAL DATA

NAME: SAMUEL LANTHALER
PLACE AND DATE OF BIRTH: SWITZERLAND | 18 AUGUST 1989
ADDRESS: BUCHHOLZSTRASSE 9, 8053 ZÜRICH, SWITZERLAND

EDUCATION

SINCE 11/2018	PhD student in Mathematics, ETH Zurich, Switzerland ADVISOR: PROF. SIDDHARTHA MISHRA
08/2015 – 03/2020	PhD student in Physics, EPFL, Switzerland ADVISOR: DR. JONATHAN P. GRAVES
03/2015	Master of Science ETH in Mathematics, ETH Zurich THESIS: “COMPUTATION OF MEASURE-VALUED SOLUTIONS OF THE INCOMPRESSIBLE EULER EQUATIONS” ADVISOR: PROF. SIDDHARTHA MISHRA
09/2013	Bachelor of Science ETH in Mathematics, ETH Zurich THESIS: “THE FIRST AND SECOND VARIATION OF AREA AND ITS APPLICATIONS” ADVISOR: PROF. MICHAEL EICHMAIR

PUBLICATIONS

1. “GUIDING-CENTRE THEORY FOR KINETIC-MAGNETOHYDRODYNAMIC MODES IN STRONGLY FLOWING PLASMAS”, S. LANTHALER ET AL., *Plasma Phys. Control. Fusion*, **61**, (2019), 074006
2. “HIGHER ORDER LARMOR RADIUS CORRECTIONS TO GUIDING-CENTRE EQUATIONS AND APPLICATION TO FAST ION EQUILIBRIUM DISTRIBUTIONS”, S. LANTHALER ET AL., *Plasma Phys. Control. Fusion*, **59**, (2017), 044014
3. “ON THE CONVERGENCE OF THE SPECTRAL VISCOSITY METHOD FOR THE TWO-DIMENSIONAL INCOMPRESSIBLE EULER EQUATIONS WITH ROUGH INITIAL DATA”, S. LANTHALER, S. MISHRA, *Found Comput Math*, (2019)
4. “STATISTICAL SOLUTIONS OF HYPERBOLIC CONSERVATION LAWS I: FOUNDATIONS”, U. S. FJORD-HOLM AND S. LANTHALER AND S. MISHRA, *Arch. Ration. Mech. An.*, **226**(2), (2017), 809–849
5. “COMPUTATION OF MEASURE-VALUED SOLUTIONS FOR THE INCOMPRESSIBLE EULER EQUATIONS”, S. LANTHALER, S. MISHRA, *Math. Models and Methods Appl. Sci.*, **25**, (2015), 2043–2088