THEORETICAL AND NUMERICAL STUDIES OF GYROKINETIC MODELS FOR SHAPED TOKAMAK PLASMAS

Edoardo Zoni

Technische Universität München

Fakultät für Mathematik
Lehrstuhl für numerische Methoden der Plasmaphysik

Max-Planck-Institut für Plasmaphysik

Theoretical and numerical studies of gyrokinetic models for shaped Tokamak plasmas

Edoardo Zoni

Vollständiger Abdruck der von der Fakultät für Mathematik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender : Prof. Dr. Oliver Junge

Prüfer der Dissertation : 1. Prof. Dr. Eric Sonnendrücker

2. Prof. Dr. Claudia Negulescu

3. Assoc. Prof. Dr. Lukas Einkemmer

Die Dissertation wurde am 30.07.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Mathematik am 19.11.2019 angenommen.

To Mom and Dad

Contents

Acknowledgments

First, I would like to sincerely thank my supervisor Eric Sonnendrücker for constantly supporting my research, for always involving me in interesting scientific discussions and meetings with several colleagues and external collaborators, and for always giving me the opportunity to participate in international workshops and conferences that contributed significantly to my professional development as a researcher in the areas of applied mathematics and computational physics.

This thesis would not have been possible without the scientific guidance of my mentor Yaman Güçlü. I would like to thank Yaman for supervising my work on a daily basis with extraordinary patience and dedication, for always listening to my questions, and for teaching me all I know about object-oriented programming and numerical mathematics. Yaman is a highly competent and extremely meticulous scientist with great physical intuition and it has been an honor to work under his scientific guidance.

I would like to thank my second mentor Stefan Possanner, who supervised my work on gyrokinetic theory, for constantly helping me with difficult analytical calculations, for being an excellent teacher with the rare ability to make complicated concepts simple and easy to understand, and for always approaching work with a positive and healthy mindset.

I would like to thank my colleagues and external collaborators Michel Mehrenberger, Guillaume Latu, Virginie Grandgirard, Omar Maj, Florian Hindenlang and Bruce Scott for our interesting scientific discussions during informal meetings as well as workshops and conferences. I would like to thank also my initial mentor Katharina Kormann, who was involved in the very early stages of this work.

I express my gratitude to Claudia Negulescu and Lukas Einkemmer for their interest in this thesis and for participating in the examining committee, and I thank Prof. Oliver Junge for presiding at my oral examination.

I would like to thank also Peter Kurz, Peter Manz and Birger Buttenschön for maintaining the high scientific quality of the HEPP Graduate School and for always helping me with regard to administrative matters in a kind and effective way. Moreover, I would like to thank Lydia Weber from the Dean's office of the TUM Department of Mathematics and Ute Stinzel from the TUM Student Service Center for helping me with all the administrative issues related to the submission of this thesis.

The numerical simulations presented in this thesis have been performed on several High Performance Computing (HPC) resources, namely the extension cluster DRACO of the HPC system of the Max Planck Computing and Data Facility and the EUROfusion HPC system MARCONI.

This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

List of Figures

List of Tables

Abstract

Gyrokinetics is a fundamental framework for the study of turbulence in magnetized fusion plasmas. In this thesis, we first propose a new methodology for the derivation of gyrokinetic models, based on polynomial transforms instead of Lie transforms. Moreover, we present novel numerical methods for gyrokinetic simulations, including a strategy for the solution of hyperbolic-elliptic PDEs on 2D singular domains and its implementation in a 4D semi-Lagrangian field-aligned drift-kinetic code.

Zusammenfassung

Gyrokinetische Modelle sind ein unverzichtbares Werkzeug für die Untersuchung von Turbulenz in magnetisierten Fusionsplasmen. In dieser Arbeit untersuchen wir eine neue Methode zur Herleitung gyrokinetischer Modelle, die auf Polynomtransformationen statt Lie-Transformationen basiert. Außerdem untersuchen wir neue numerische Methoden für gyrokinetische Simulationen, nämlich eine Strategie für die Lösung hyperbolischer-elliptischer PDGL auf 2D singulären Gebieten und deren Implementierung in einem 4D Semi-Lagrange Feldlinien-angepassten drift-kinetischen Code.

Chapter 1

Introduction

In this first introductory chapter we state the main motivations of this thesis, outline its contents, and give an overview of the overall context of the work, namely the role of gyrokinetic theory and simulations in the research areas of plasma physics and nuclear fusion.

1.1 Motivation and outline of the thesis

This Ph.D. thesis is the result of about three years of research work conducted by the author at the Max Planck Institute for Plasma Physics under the academic supervision of Prof. Dr. Eric Sonnendrücker and the scientific guidance of Dr. Yaman Güçlü and Dr. Stefan Possanner. This work has been carried out within the International Helmholtz Graduate School for Plasma Physics, a thematic graduate center of the Technical University of Munich.

The ultimate goal of this research work is to contribute to the advancement of the mathematical and computational tools at disposal of physicists, mathematicians, computer scientists and engineers working in the research areas of plasma physics and nuclear fusion, and eventually improve the current understanding of the complex variety of physical phenomena occurring in experimental nuclear fusion reactors. Mathematics and computational science provide invaluable tools to build mathematical models of such physical phenomena and study their behavior and evolution in space and time with the help of computer experiments. In particular, this thesis focuses on the mathematical and computational methods employed for the theoretical derivation of gyrokinetic models and for their numerical solution by means of computer simulations.

In the first part of this thesis (chapter 2) we propose a new methodology for the theoretical derivation of gyrokinetic models. Our technique is based on polynomial phase-space coordinate transformations, which we call polynomial transforms. Our derivation is conceptually simpler than the standard one based on Lie transforms and makes gyrokinetic theory easier to access for non-specialists. In the spirit of asymptotic analysis, we define a rigorous normalization procedure in order to carry out our derivation in a fully non-dimensional framework. Moreover, our ordering assumptions (maximal ordering) are inspired by realistic physical scenarios relevant for existing and future fusion experiments, such as, for example, the Tokamaks ASDEX Upgrade (Max Planck Institute for Plasma Physics) and ITER (ITER Organization). As a result, we derive a set of gyrokinetic Vlasov-Maxwell equations for ions and electrons within a unique ordering.

In the second part of this thesis (chapters 3-5) we develop and discuss novel numerical methods for gyrokinetic simulations. In particular, we propose a comprehensive numerical strategy for the solution of coupled hyperbolic-elliptic partial differential equations on two-dimensional disk-like domains, obtained from a rectangular uniform logical domain through singular coordinate mappings. In particular, we propose a novel set of coordinates, which we call pseudo-Cartesian coordinates, for the integration of the characteristic equations for the semi-Lagrangian solution of hyperbolic equations on such domains, and a finite element method based on globally $\mathcal{C}^1$ smooth splines for the solution of elliptic equations. We also report about the development of a semi-Lagrangian simulation code based on the field-aligned approach. This code addresses the numerical solution of a simplified four-dimensional drift-kinetic electrostatic model. We discuss in detail the field-aligned approach and show how to integrate the numerical machinery developed for two-dimensional problems within the four-dimensional code, presenting preliminary verification tests.

It is our hope that the numerical methods and techniques presented in this thesis, or their extension and improvement, will be eventually implemented and employed in state-of-the-art gyrokinetic codes, as well as other simulation codes used in computational plasma physics.

1.2 Plasma physics and nuclear fusion

Plasma physics and nuclear fusion constitute the overall context of this research work with respect to the physical sciences. The physical system being ultimately investigated consists of a fusion plasma, that is, an ionized gas of electrons and ions at extremely high temperatures and densities inside a nuclear fusion reactor. When nuclei of light atoms, such as hydrogen, collide and fuse together, they produce nuclei of heavier atoms, such as helium, and release enormous amounts of energy. This reaction occurs, for example, in the core of the Sun and all other stars in the universe. The goal of nuclear fusion science is to replicate such reactions in laboratories and experimental facilities on earth, with the ultimate aim of converting the energy produced by fusion reactions into electrical energy. Thanks to the fact that fusion fuels, namely deuterium and tritium, are extensively available and almost inexhaustible on earth, nuclear fusion represents a possible future source of large-scale carbon-free energy, that may help address some of the most urgent challenges facing humanity, namely the energy needs of our species within a safe preservation of our ecosystem. Several aspects make nuclear fusion an attractive safe alternative to nuclear fission as a nuclear energy source. To name a few: fusion does not emit carbon dioxide or other greenhouse gases into the atmosphere, nuclear fusion reactors do not produce long-lived nuclear waste, and there is no risk of a chain reaction or meltdown. The downside is that what makes nuclear fusion inherently safe makes it also extraordinarily challenging from a scientific and technological point of view, as the reaction stops or does not occur at all if any of the physical conditions necessary for it are disturbed in any way. In order for the reaction to occur at a reasonable rate, the plasma inside a fusion reactor needs to be heated at temperatures of about 150 million degrees Celsius, much higher than the Sun's core. It is therefore necessary to confine the hot plasma (which has a tendency to expand) inside the core of a fusion reactor for a sufficiently long time and prevent it from reaching the outer walls of the machine. Several techniques have been invented and investigated in the last decades in order to achieve high plasma confinement in fusion reactors. The most common one is magnetic confinement, which makes use of strong external magnetic fields to contain and control the plasma. The idea of magnetic confinement comes from the simple physical observation that electrically charged particles (such as electrons and ions inside a plasma) exhibit a motion of gyration around the field lines of an external magnetic field. As a consequence, a combination of poloidal and toroidal magnetic coils in a donut-like shaped device can act as a magnetic cage and trap the plasma particles in a well-confined region inside the reactor's vessels. The two main existing types of fusion reactors based on magnetic confinement are the so-called Tokamaks (Artsimovich, 1972) and Stellarators (Spitzer, 1958). They differ mainly because of the shape and the spatial symmetry properties of the magnetic coils used to create the external magnetic fields. This thesis deals with magnetically confined fusion plasmas in Tokamak devices, such as, for example, ASDEX Upgrade, DIII-D, JET, TCV (to name just a few among the more than 200 Tokamak reactors that have been built and operated since the 1950s) and the future machine ITER (Figure 1.1).

1.3 Gyrokinetic theory and simulations

Fusion plasmas consist typically of an extremely large number of charged particles interacting between themselves and with external electromagnetic fields. When a macroscopic physical system consists of a large number of microscopic components, it is common to describe it from a statistical point of view, instead of looking at its individual microscopic degrees of freedom. In the case of a fusion plasma, we may want to describe its physical dynamics with the tools provided by the kinetic theory of gases, rather than looking at the individual equations of motion of each electron and ion in the plasma. The fundamental bit of information in such a kinetic description is typically the probability of finding a particle at a given position, with a given velocity and at a given time. The mathematical object that contains such information is the so-called distribution function of the plasma particles. Such objects

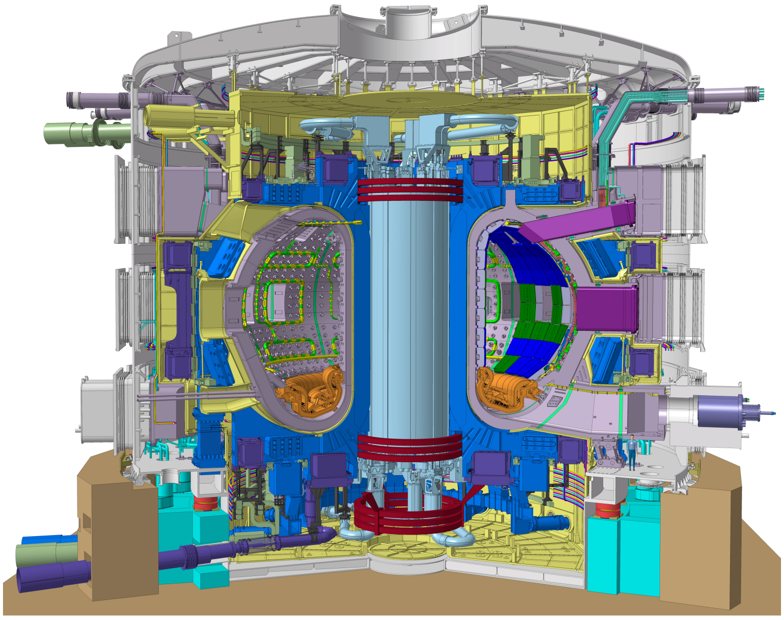


Figure 1.1: A cutaway of the ITER Tokamak (produced by the ITER Design Office in January 2013).

are defined, at a given time, on a six-dimensional phase space (three dimensions are associated to the position of the plasma particles and three to their velocity). The evolution in time of a particle distribution function is therefore described by a six-dimensional phase-space dynamics. When the equations describing such dynamics need to be solved on a computer, as it is usually the case, the high dimensionality of the phase space becomes a computational limitation, as it results in long run times and heavy memory footprints of the corresponding numerical simulations. It becomes therefore useful to develop reduced models capable of retaining a physically meaningful description of the system while at the same time decreasing the computational cost of the model. Gyrokinetic theory is an example of such reduced models: it is derived from a fully kinetic model by applying standard techniques of perturbation theory and results in a set of dynamical equations defined on a five-dimensional phase space, thus enabling more efficient numerical simulations of fusion plasmas.

Several efforts have been made in the last decades in order to develop computer codes for the numerical simulation of fusion plasmas by means of gyrokinetic models. Such codes are typically classified into two different categories, according to the overall discretization approach on which their numerical methods are based. In particle-in-cell (PIC) codes the system is observed from macro-particles and only a computational grid in the three-dimensional configuration space is necessary in order to solve the dynamical equations (more precisely, the equations describing the dynamics of the electromagnetic fields). Examples of gyrokinetic PIC codes include GTC (Ethier et al., 2005), GTS (Wang et al., 2006), XGC1 (Ku et al., 2009), ORB5 (Bottino et al., 2010), and ELMFIRE (Heikkinen et al., 2008). In continuum codes, instead, the dynamical equations are discretized on a five-dimensional phase-space grid. Such codes are further classified into purely Eulerian codes, based on various discretization approaches, such as, for example, finite differences, finite volumes, or discontinuous Galerkin methods, and semi-Lagrangian codes, where the particle distribution function is discretized on a grid but the grid points are advected along the characteristic trajectories of the transport equation describing its dynamics. Examples of gyrokinetic Eulerian codes include GENE (Jenko and Dorland, 2001), COGENT (Dorf et al., 2012), GKEYLL (Shi et al., 2017), GT5D (Idomura et al., 2008); examples of gyrokinetic semi-Lagrangian codes include GySeLa (Grandgirard et al., 2006b) and the code described in (Kwon et al., 2015).

Chapter 2

A new derivation of gyrokinetic theory

Gyrokinetics is one of the major frameworks used in theoretical and numerical studies of low-frequency turbulence in magnetized fusion plasmas (Garbet et al., 2010; Krommes, 2012). Gyrokinetic models are based on a change of phase-space coordinates that separates the fast motion of gyration of the charged plasma particles around the magnetic field lines from the slower motion of their guiding-centers along the direction of the magnetic field lines. The idea is to derive a reduced set of dynamical equations that contain all the information necessary to provide a good enough description of the physical dynamics of the plasma, while at the same time neglecting the dynamics occurring at the fastest time scales, thus reducing the overall dimensionality of the equations to be solved and enabling more efficient numerical simulations (by reducing both their run time and their memory footprint).

From the point of view of mathematics, gyrokinetic theory is the asymptotic study of the Vlasov-Maxwell model in a strong background magnetic field. This requires a suitable reformulation of the equations such that the formal limit of infinite magnetic field strength can be carried out in a meaningful way. The procedure can be understood in the context of averaging the characteristics (or Lagrangian paths) of the Vlasov equation, as described first in (Northrop, 1963). Rigorous mathematical accounts can be found, for example, in (Frénod and Sonnendrücker, 2001; Bostan, 2010; Chartier et al., 2016). A first attempt to average the Vlasov-Maxwell equations taking into account the self-consistent interaction between plasma particles and electromagnetic fields can be found in (Frieman and Chen, 1982). On the other hand, the works of (Littlejohn, 1979; Dubin, 1983; Hahm, 1988; Brizard, 1989) laid the foundation of "structure-preserving" gyrokinetic theory. In these works it was first realized that averaging can be carried out on the level of the variational principle, by transforming the particle Lagrangian that leads, together with the free-field electromagnetic Lagrangian, to the Vlasov-Maxwell equations (or, equivalently, by transforming its corresponding Poisson bracket structure). This strategy has the advantage of preserving symmetries of the plasma equations of motion during the process of averaging. More precisely, the averaged equations will exactly conserve averaged versions of the true constants of the motion, such as, for example, energy and momentum. Most, if not all, of the numerical codes attempting to solve the gyrokinetic equations are currently based on structure-preserving gyrokinetic models and often show improved stability and accuracy.

The derivation of gyrokinetic models has been discussed and reviewed extensively in (Brizard and Hahm, 2007; Krommes, 2012), and references therein. The prevalent methodology is based on Lie transform perturbation theory, as presented for example in (Cary, 1981; Cary and Littlejohn, 1983), and most of the recent derivations have been carried out in this framework, as in (Tronko et al., 2016; Brizard, 2017; Tronko et al., 2017a,b; Tronko and Chandre, 2018). The work described in this thesis suggests a different derivation of the gyrokinetic equations, which we hope will make gyrokinetic theory easier to access for scientists who are not familiar with it. The idea is to derive the gyrokinetic equations without relying on methods based on Lie transform perturbation theory. Such methods, despite being mathematically elegant, are formulated in the language of differential geometry and may thus prevent readers from focusing on the core of the gyrokinetic reduction. The derivation discussed in this thesis is inspired by the guiding-center theory of Littlejohn (Littlejohn, 1983) and a similar approach has been recently suggested in (Parra and Calvo, 2011; Scott, 2017) and worked out in the long-wavelength regime of gyrokinetics in (Possanner, 2018). More precisely, we extend the methodology presented in (Possanner, 2018) to address the description of turbulence on the microscopic scales of magnetized fusion plasmas, such as the ion Larmor radius. As in Lie transform perturbation theory, our method is based on near-identity phase-space coordinate transformations. However, we propose to construct such transformations in a different way. More precisely, our phase-space coordinate transformations are defined as polynomials of finite degree in powers of a given perturbation parameter. The coefficients of such polynomials are the so-called generating functions (or generators) and represent the degrees of freedom that allow us to separate fast and slow scales. We call such phase-space coordinate transformations polynomial transforms, in analogy with Lie transforms. Our polynomial transforms are conceptually simpler than Lie transforms, which are asymptotic series (thus not necessarily convergent) constructed as products of operator exponentials which feature Lie derivatives along the generating vector fields.

In the spirit of asymptotic analysis, our derivation of the gyrokinetic equations is carried out in a fully non-dimensional framework. The perturbation parameter used in our near-identity phase-space coordinate transformations, which we denote by ε, is identified by a rigorous normalization of the Vlasov-Maxwell model. Our ordering in powers of ε is then based on assumptions derived from realistic physical scenarios relevant for existing and future fusion experiments, such as, for example, the Tokamaks ASDEX Upgrade and ITER . In this way we clearly separate the physical assumptions (generally referred to as "ordering" in the gyrokinetic literature) from the mathematical model reduction (averaging with polynomial transforms). The reduction procedure via polynomial transforms can then be viewed as a sort of automatic procedure, which ideally could be carried out by a symbolic computer program, in order to prevent errors in long analytical calculations. Moreover, our methodology based on a priori normalization of the physical equations allows us to formulate a gyrokinetic theory for both ions and electrons within the same physical scenario and assumptions.

This chapter is organized as follows. Section 2.1 introduces the basic equations of the Vlasov-Maxwell model for a non-collisional magnetized plasma, including a field-theoretic Lagrangian formulation. Section 2.2 defines the normalization scheme used for the purpose of non-dimensionalization and derives an ordering pattern based on physical considerations. Section 2.3 outlines the main results of this work, namely a set of gyrokinetic equations for both ions and electrons in maximal ordering with the corresponding sources in Maxwell's equations. Finally, we end the chapter with a summary of our contributions.

2.1 The Vlasov-Maxwell model

We consider a non-collisional plasma composed of ions and electrons described in terms of the distribution functions $f_\mathrm{s}(t, \boldsymbol{x}, \boldsymbol{v})$, where the subscript s denotes the particle species, $t \in \mathbb{R}^+$ denotes the time coordinate, and $(\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$ are position and velocity coordinates in phase space. In this chapter we write all equations in SI units (Taylor, 2008). The distribution functions f_s obey the non-collisional Vlasov equation

$$\frac{\partial f_\mathrm{s}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f_\mathrm{s} + \frac{q_\mathrm{s}}{m_\mathrm{s}} (\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}) \cdot \frac{\partial f_\mathrm{s}}{\partial \boldsymbol{v}} = 0\,, \tag{2.1}$$

where q_s and m_s denote the particle charge and mass, respectively. The electromagnetic fields $\boldsymbol{E}(t, \boldsymbol{x})$ and $\boldsymbol{B}(t, \boldsymbol{x})$ satisfy Maxwell's equations

$$\boldsymbol{\nabla} \cdot \boldsymbol{E} = \frac{\rho}{\varepsilon_0} \qquad \text{(Coulomb's law)}\,, \tag{2.2a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{B} = 0 \qquad \text{(absence of free magnetic poles)}\,, \tag{2.2b}$$

$$\boldsymbol{\nabla} \times \boldsymbol{E} = -\frac{\partial \boldsymbol{B}}{\partial t} \qquad \text{(Faraday's law)}\,, \tag{2.2c}$$

$$\boldsymbol{\nabla} \times \boldsymbol{B} = \mu_0 \boldsymbol{J} + \varepsilon_0 \mu_0 \frac{\partial \boldsymbol{E}}{\partial t} \qquad \text{(Ampère-Maxwell's law)}\,, \tag{2.2d}$$

where ε_0 and μ_0 denote the vacuum electric permittivity and the vacuum magnetic permeability, respectively. The sources $\rho(t, \boldsymbol{x})$ and $\boldsymbol{J}(t, \boldsymbol{x})$ are defined in terms of the distribution functions as

$$\rho := \sum_\mathrm{s} q_\mathrm{s} \int \mathrm{d}^3\boldsymbol{v}\, f_\mathrm{s}\,, \qquad \boldsymbol{J} := \sum_\mathrm{s} q_\mathrm{s} \int \mathrm{d}^3\boldsymbol{v}\, \boldsymbol{v}\, f_\mathrm{s}\,. \tag{2.3}$$

The derivation of the Vlasov-Maxwell system (2.1)-(2.3) from an action principle was recognized first by (Low, 1958). Denoting by $\Phi(t, \boldsymbol{x})$ and $\boldsymbol{A}(t, \boldsymbol{x})$ the electric scalar potential and the magnetic vector potential associated to the electric and magnetic fields $\boldsymbol{E}$ and $\boldsymbol{B}$ via

$$\boldsymbol{E} = -\boldsymbol{\nabla}\Phi - \frac{\partial \boldsymbol{A}}{\partial t} \,, \qquad \boldsymbol{B} = \boldsymbol{\nabla} \times \boldsymbol{A} \,, \tag{2.4}$$

Low's action principle reads

$$\delta \int_{t_0}^{t_1} \mathrm{d}t \; (L_\mathrm{p} + L_\mathrm{EM}) = 0 \,, \tag{2.5}$$

where δ denotes the Fréchet derivative (as usual in the calculus of variations), $L_\mathrm{p}\,(t, \Psi_t, \mathrm{d}\Psi_t/\mathrm{d}t, \Phi, \boldsymbol{A})$ denotes the particle Lagrangian

$$L_\mathrm{p} := \sum_\mathrm{s} \int \mathrm{d}^3\boldsymbol{x}_0 \, \mathrm{d}^3\boldsymbol{v}_0 \, f_\mathrm{s}(t_0, \boldsymbol{x}_0, \boldsymbol{v}_0) \, L_\mathrm{s} \,, \tag{2.6}$$

and $L_\mathrm{EM}\,(\Phi, \boldsymbol{A}, \boldsymbol{\nabla}\Phi, \partial\boldsymbol{A}/\partial t, \boldsymbol{\nabla}\boldsymbol{A})$ denotes the electromagnetic free-field Lagrangian

$$L_\mathrm{EM} := \frac{\varepsilon_0}{2} \int \mathrm{d}^3\boldsymbol{x} \left| \boldsymbol{\nabla}\Phi + \frac{\partial \boldsymbol{A}}{\partial t} \right|^2 - \frac{1}{2\mu_0} \int \mathrm{d}^3\boldsymbol{x} \, \left| \boldsymbol{\nabla} \times \boldsymbol{A} \right|^2 \,. \tag{2.7}$$

In the particle Lagrangian (2.6), $L_\mathrm{s}\,(t, \Psi_t, \mathrm{d}\Psi_t/\mathrm{d}t, \Phi, \boldsymbol{A})$ (sometimes denoted simply by $L_\mathrm{s}\,(\Psi_t, \mathrm{d}\Psi_t/\mathrm{d}t)$, neglecting the dependence on the other variables) denotes the single-particle Lagrangian for species s, defined on the tangent bundle of the single-particle phase space. In the phase-space coordinates $(\boldsymbol{x}, \boldsymbol{v})$, the single-particle Lagrangian $L_\mathrm{s}(t, \boldsymbol{x}, \boldsymbol{v}, \dot{\boldsymbol{x}}, \dot{\boldsymbol{v}})$ reads

$$L_\mathrm{s} := (m_\mathrm{s}\boldsymbol{v} + q_\mathrm{s}\boldsymbol{A}) \cdot \dot{\boldsymbol{x}} - H_\mathrm{s} \quad \text{with} \quad H_\mathrm{s} := \frac{m_\mathrm{s}}{2}|\boldsymbol{v}|^2 + q_\mathrm{s}\Phi \,. \tag{2.8}$$

We remark that L_s depends implicitly on the potentials Φ and $\boldsymbol{A}$ and describes therefore the self-consistent interaction between plasma particles and electromagnetic fields. In general, the single-particle Lagrangian $L_\mathrm{s}\,(\Psi_t, \mathrm{d}\Psi_t/\mathrm{d}t)$ is evaluated at the t-family of maps $\Psi_t : \mathbb{R}^6 \to \mathbb{R}^6$, which depend parametrically on time and denote the flow maps of the characteristics of the Vlasov equation (2.1):

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{v} \,, \\ \boldsymbol{x}(t_0) = \boldsymbol{x}_0 \,, \end{cases} \qquad \begin{cases} \dfrac{\mathrm{d}\boldsymbol{v}}{\mathrm{d}t} = \dfrac{q_\mathrm{s}}{m_\mathrm{s}} \left(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B} \right) \,, \\ \boldsymbol{v}(t_0) = \boldsymbol{v}_0 \,. \end{cases} \tag{2.9}$$

The flow map Ψ_t transports a particle from the phase-space point $\boldsymbol{z}_0 := (\boldsymbol{x}_0, \boldsymbol{v}_0)$ at the initial time t_0 to the phase-space point $\Psi_t(\boldsymbol{z}_0)$ at time t and it is a volume-preserving diffeomorphism. The description of an ensemble of particles via the particle Lagrangian L_p arises from the single-particle picture as follows. Newton's equations of motion for a single particle can be deduced from the variational principle

$$\delta \int_{t_0}^{t_1} \mathrm{d}t \, L_\mathrm{s} \left(\boldsymbol{z}(t), \frac{\mathrm{d}\boldsymbol{z}(t)}{\mathrm{d}t} \right) = 0 \,. \tag{2.10}$$

The extremum defined by (2.10) is denoted by $\boldsymbol{z}(t) := (\boldsymbol{x}(t), \boldsymbol{v}(t))$ and it is the solution of the Euler-Lagrange equations (2.9). Hence, $\boldsymbol{z}(t) = \Psi_t(\boldsymbol{z}_0)$ and (2.10) can be written as

$$\delta \int_{t_0}^{t_1} \mathrm{d}t \int \mathrm{d}^6\boldsymbol{z}_1 \, \delta^6(\boldsymbol{z}_1 - \boldsymbol{z}_0) \, L_\mathrm{s} \left(\Psi_t(\boldsymbol{z}_1), \frac{\mathrm{d}\Psi_t(\boldsymbol{z}_1)}{\mathrm{d}t} \right) = 0 \,. \tag{2.11}$$

Formally, the ensemble description is obtained by replacing the Dirac delta function with the initial particle distribution function $f_\mathrm{s}(0, \boldsymbol{z}_1)$, which yields

$$\delta \int_{t_0}^{t_1} \mathrm{d}t \int \mathrm{d}^6\boldsymbol{z}_1 \, f_\mathrm{s}(t_0, \boldsymbol{z}_1) \, L_\mathrm{s} \left(\Psi_t(\boldsymbol{z}_1), \frac{\mathrm{d}\Psi_t(\boldsymbol{z}_1)}{\mathrm{d}t} \right) = 0 \,. \tag{2.12}$$

The particle Lagrangian (2.6) is then obtained by taking the sum over the species and by relabeling the variables of integration $\boldsymbol{z}_1$ as $\boldsymbol{z}_0$. By construction, the variation (2.12) with respect to Ψ_t yields the characteristics of the Vlasov equation. From a more physical point of view, the particle Lagrangian (2.6) is obtained by multiplying the single-particle Lagrangian L_s by the probability of finding a particle in the phase-space point $\boldsymbol{z}_0$ at the initial time t_0 and by summing over all possible initial phase-space points. The Vlasov equation enters the picture via the following definition of the particle distribution function:

$$f_\mathrm{s}(t, \boldsymbol{z}) := \int \mathrm{d}^6\boldsymbol{z}_0 \, \delta^6(\boldsymbol{z} - \Psi_t(\boldsymbol{z}_0)) \, f_\mathrm{s}(t_0, \boldsymbol{z}_0) \,. \tag{2.13}$$

We remark that $\boldsymbol{z} \in \mathbb{R}^6$ denotes coordinates here and not a path in phase space. Hence, (2.13) makes the link between the Lagrangian paths Ψ_t and the distribution function f_s via $f_\mathrm{s}(t, \boldsymbol{z}) = f_\mathrm{s}(t_0, (\Psi_t)^{-1}(\boldsymbol{z}))$. This implies, in particular, that f_s is constant along the Lagrangian paths,

$$f_\mathrm{s}(t, \Psi_t(\boldsymbol{z}_0)) = f_\mathrm{s}(t_0, \boldsymbol{z}_0)\,, \tag{2.14}$$

which corresponds to the information contained in the Vlasov equation. Let us now write the characteristic equations in the form

$$\begin{cases} \dfrac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = \boldsymbol{F}(\boldsymbol{z})\,, \\ \boldsymbol{z}(t_0) = \boldsymbol{z}_0\,, \end{cases} \tag{2.15}$$

with $\boldsymbol{F}(\boldsymbol{z}) = \boldsymbol{F}(\boldsymbol{x}, \boldsymbol{v}) := (\boldsymbol{v}, \frac{q_\mathrm{s}}{m_\mathrm{s}}(\boldsymbol{E} + \boldsymbol{v} \times \boldsymbol{B}))^T$. From the definition of the distribution function it follows that

$$\int \mathrm{d}^6\boldsymbol{z}_0\, f_\mathrm{s}(t_0, \boldsymbol{z}_0)\, L_\mathrm{s}\left(\Psi_t, \frac{\mathrm{d}\Psi_t}{\mathrm{d}t}\right) = \int \mathrm{d}^6\boldsymbol{z}\, f_\mathrm{s}(t, \boldsymbol{z})\, L_\mathrm{s}(\boldsymbol{z}, \boldsymbol{F}(\boldsymbol{z}))\,. \tag{2.16}$$

This is easily verified by substituting the definition (2.13) into the right-hand side and integrating over $\boldsymbol{z}$. The variational principle (2.5) leads to the characteristics of the Vlasov equation by computing variations with respect to Ψ_t, to Coulomb's law by computing variations with respect to Φ and to Ampère-Maxwell's law by computing variations with respect to $\boldsymbol{A}$. Variations with respect to the electromagnetic potentials are computed in the Eulerian representation of the particle Lagrangian, namely from the right-hand side of (2.16). Only the non-homogeneous Maxwell's equations (2.2a) and (2.2d) can be derived from the variational principle. With appropriate initial/boundary conditions, this results in a well-posed system for $(f_\mathrm{s}, \Phi, \boldsymbol{A})$, which describes the self-consistent interaction between plasma particles and electromagnetic fields.

An important aspect in gyrokinetic theory is the separation of the electromagnetic fields into background and fluctuating parts. More precisely, we assume that the magnetic field consists of dynamic fluctuations added to a static background, while the electric field consists only of dynamic fluctuations (without a static background):

$$\boldsymbol{B}(t, \boldsymbol{x}) = \boldsymbol{B}_0(\boldsymbol{x}) + \boldsymbol{B}_1(t, \boldsymbol{x})\,, \qquad \boldsymbol{E}(t, \boldsymbol{x}) = \boldsymbol{E}_1(t, \boldsymbol{x})\,. \tag{2.17}$$

Similarly, the electromagnetic potentials are written as

$$\boldsymbol{A}(t,\boldsymbol{x}) = \boldsymbol{A}_0(\boldsymbol{x}) + \boldsymbol{A}_1(t,\boldsymbol{x})\,, \qquad \Phi(t,\boldsymbol{x}) = \Phi_1(t,\boldsymbol{x})\,, \tag{2.18}$$

so that $\boldsymbol{B}_0 = \boldsymbol{\nabla} \times \boldsymbol{A}_0$. Therefore, in the variational principle (2.5) variations have to be computed with respect to Ψ_t, Φ_1 and $\boldsymbol{A}_1$, respectively.

2.2 Normalization and ordering

The formulation of the Vlasov-Maxwell system as a perturbation problem requires the non-dimensionalization (or scaling) of the physical equations and the subsequent application of an ordering scheme that allows for a comparison of terms in relation to a small perturbation parameter $\varepsilon \ll 1$. Since gyrokinetics is ultimately the theory of low-frequency dynamics in strongly-magnetized plasmas, the perturbation parameter is typically defined as the ratio between a characteristic ion turbulence frequency $\widehat{\omega}_\mathrm{i}$ and the ion cyclotron frequency ω_ci:

$$\varepsilon := \frac{\widehat{\omega}_\mathrm{i}}{\omega_\mathrm{ci}}\,. \tag{2.19}$$

Gyrokinetic theory can then be considered as the asymptotic analysis of the Vlasov-Maxwell model in the limit $\varepsilon \to 0$. The procedure of non-dimensionalization and the introduction of the scaling parameter ε in the equations are fundamental. Therefore, we give here a detailed description of both steps, which typically are not treated extensively in the gyrokinetic literature. We first present the methodology by means of a one-dimensional example and then introduce the generic normalization of the Low action principle (2.5). Finally, we suggest an ordering scheme that corresponds to a realistic physical scenario relevant for existing and future fusion experimental reactors, such as, for example, the Tokamaks ASDEX Upgrade and ITER.

In magnetized fusion plasmas two length scales play a major role, namely the length scale ℓ_0 of the background magnetic field, that can be approximately measured by looking at the quantity $||\boldsymbol{\nabla} B_0/B_0||^{-1}$, and the length scale ℓ_1 of the turbulent fluctuations, that can be approximately measured by looking at the quantity $||\boldsymbol{\nabla}\Phi_1/\Phi_1||^{-1}$.

The length scale ℓ_1 is typically of the order of the ion gyro-radius $\rho_{\mathrm{i}} := \widehat{v}_{\mathrm{i}}/\omega_{\mathrm{ci}}$ (Brizard and Hahm, 2007), where $\widehat{v}_{\mathrm{i}}$ denotes the ion thermal velocity. The ratio of these length scales determines the importance of the effects related to the curvature of the background magnetic field in the perturbation analysis. The following three scenarios are usually of interest: if the background magnetic field is uniform, we formally have $\ell_1/\ell_0 = 0$; if the background magnetic field is slowly-varying, we typically have $\ell_1/\ell_0 = O(\varepsilon^2)$ in the limit $\varepsilon \to 0$; in the last scenario, which is referred to as maximal ordering, we have instead $\ell_1/\ell_0 = O(\varepsilon)$ in the limit $\varepsilon \to 0$. The majority of gyrokinetic models have been derived in a homogeneous background magnetic field for the sake of conceptual clarity. However, curvature effects play an important role in magnetically confined fusion plasmas and many state-of-the-art gyrokinetic codes address the numerical solution of gyrokinetic models derived in the case of a slowly-varying background magnetic field. In this thesis, we develop a consistent gyrokinetic theory in the maximal ordering, by taking into account curvature terms that appear at the first order of the perturbation theory. As pointed out in (Tronko and Chandre, 2018), according to recent experimental results for the Tokamak Tore Supra (Casati et al., 2009) and according to numerical results obtained with the PIC gyrokinetic code ORB5 for the Tokamaks ITER and DIII-D (Wersal et al., 2012), the maximal ordering is the most relevant scenario for existing and future fusion experimental reactors (including Spheromaks, besides conventional Tokamaks and Stellarators). Uniform or slowly-varying scenarios can be then easily deduced from our more general results in the maximal ordering, by simply moving curvature terms to higher orders in the perturbation theory.

When dealing with different length scales of variations, such as ℓ_0 and ℓ_1, particular care is required during the process of non-dimensionalization (or scaling) of the physical equations, because the perturbation parameter ε will appear in the argument of functions. We explain this by means of a one-dimensional example. Consider the one-dimensional real-valued function

$$g(x) = \widehat{g}\,\sin\left(\frac{2\pi}{\ell}x\right). \tag{2.20}$$

The amplitude $\widehat{g}$ represents the characteristic size of g and the length scale ℓ represents its scale of variation. In a physical problem, ℓ is typically expressed in some physical units (for example, $\ell = 1\,\mathrm{mm}$), and so is x, in order to have a non-dimensional argument in the sine function. We then introduce a second function

$$g'(x') := \sin(2\pi\, x')\,, \tag{2.21}$$

with non-dimensional argument x'. The function g' is non-dimensional, of order $O(1)$ and with variations of order $O(1)$ in the limit $\ell \to 0$. We call such functions "elementary functions" and denote them with a prime. The goal of the scaling procedure is to express all dependent variables of a physical model in terms of elementary functions. From (2.20) and (2.21) it follows that

$$g(x) = \widehat{g}\, g'\left(\frac{x}{\ell}\right)\,, \qquad \frac{\partial g}{\partial x}(x) = \frac{\widehat{g}}{\ell}\frac{\partial g'}{\partial x'}\left(\frac{x}{\ell}\right)\,. \tag{2.22}$$

We also introduce the characteristic length scale $\widehat{x}$ in order to scale the independent variable x via $x = \widehat{x}\, x'$. The length scale $\widehat{x}$ represents a characteristic scale of observation. From (2.22) we obtain

$$g(x) = \widehat{g}\, g'\left(\frac{\widehat{x}}{\ell}\, x'\right)\,, \qquad \frac{\partial g}{\partial x}(x) = \frac{\widehat{g}}{\ell}\frac{\partial g'}{\partial x'}\left(\frac{\widehat{x}}{\ell}\, x'\right)\,, \tag{2.23}$$

where $g'(x')$ is differentiated with respect to x' and then evaluated at what stands in the parentheses. If we fix a scale of observation $\widehat{x}$ and we introduce a perturbation parameter $\varepsilon \ll 1$, we then distinguish between three types of functions, depending on the ratio $\widehat{x}/\ell$:

- *slowly-varying* functions with $\widehat{x}/\ell = O(\varepsilon)$ in the limit $\varepsilon \to 0$: $g = \widehat{g}\, g'\left(\varepsilon\, x'\right)$ (scale of observation much smaller than the scale of variation);
- *standard* functions with $\widehat{x}/\ell = O(1)$ in the limit $\varepsilon \to 0$: $g = \widehat{g}\, g'\left(x'\right)$ (scale of observation comparable to the scale of variation);
- *strongly-varying* functions with $\widehat{x}/\ell = O(1/\varepsilon)$ in the limit $\varepsilon \to 0$: $g = \widehat{g}\, g'(x'/\varepsilon)$ (scale of observation much larger than the scale of variation).

We will use this notation throughout this chapter, adapted to the multi-dimensional case as well as for vector-valued functions.

2.2.1 Normalization of the Vlasov-Maxwell model

In order to write the Vlasov-Maxwell model in non-dimensional form, we introduce reference scales (denoted, as before, by a hat) for time, length, and ion and electron velocities:

$$t = \widehat{t}\,t'\,, \qquad \boldsymbol{x} = \widehat{x}\,\boldsymbol{x}'\,, \qquad \boldsymbol{v} = \begin{cases} \widehat{v}_{\mathrm{i}}\,\boldsymbol{v}' & \text{ions}\,, \\ \widehat{v}_{\mathrm{e}}\,\boldsymbol{v}'' & \text{electrons}\,. \end{cases} \tag{2.24}$$

The characteristic velocities for ions and electrons, $\widehat{v}_{\mathrm{i}}$ and $\widehat{v}_{\mathrm{e}}$, can be chosen differently, which means that $\boldsymbol{v}'$ and $\boldsymbol{v}''$ are different velocity coordinates. Following the procedure described in the previous section, we write the background magnetic field and its corresponding magnetic vector potential as

$$\boldsymbol{B}_0(\boldsymbol{x}) = \widehat{B}_0\,\boldsymbol{B}_0'\left(\frac{\widehat{x}}{\ell_0}\boldsymbol{x}'\right)\,, \qquad \boldsymbol{A}_0(\boldsymbol{x}) = \widehat{A}_0\,\boldsymbol{A}_0'\left(\frac{\widehat{x}}{\ell_0}\boldsymbol{x}'\right)\,, \tag{2.25}$$

where ℓ_0 denotes the length scale of the background magnetic field. Choosing $\widehat{A}_0 := \widehat{B}_0\ell_0$, we obtain $\boldsymbol{B}_0' = \boldsymbol{\nabla}' \times \boldsymbol{A}_0'$ in the scaled variables. We remark that for a uniform background magnetic field the considerations for the vector potential $\boldsymbol{A}_0$ are still valid ($\boldsymbol{A}_0 = (\boldsymbol{B}_0 \times \boldsymbol{x})/2$ in this case). With regard to the dynamic fields $\boldsymbol{B}_1$ and $\boldsymbol{E}_1$, we denote their length and time scales by ℓ_1 and τ_1, respectively. For the magnetic fluctuations and its corresponding magnetic vector potential we write

$$\boldsymbol{B}_1(t,\boldsymbol{x}) = \widehat{B}_1\,\boldsymbol{B}_1'\left(\frac{\widehat{t}}{\tau_1}t', \frac{\widehat{x}}{\ell_1}\boldsymbol{x}'\right)\,, \qquad \boldsymbol{A}_1(t,\boldsymbol{x}) = \widehat{A}_1\,\boldsymbol{A}_1'\left(\frac{\widehat{t}}{\tau_1}t', \frac{\widehat{x}}{\ell_1}\boldsymbol{x}'\right)\,. \tag{2.26}$$

Similarly, for the electric field and its corresponding electric scalar potential we write

$$\boldsymbol{E}_1(t,\boldsymbol{x}) = \widehat{E}_1\,\boldsymbol{E}_1'\left(\frac{\widehat{t}}{\tau_1}t', \frac{\widehat{x}}{\ell_1}\boldsymbol{x}'\right)\,, \qquad \Phi_1(t,\boldsymbol{x}) = \widehat{\Phi}_1\,\Phi_1'\left(\frac{\widehat{t}}{\tau_1}t', \frac{\widehat{x}}{\ell_1}\boldsymbol{x}'\right)\,. \tag{2.27}$$

Choosing $\widehat{A}_1 := \widehat{B}_1\ell_1$ and $\widehat{\Phi}_1 := \widehat{E}_1\ell_1$, we obtain

$$\boldsymbol{B}_1' = \boldsymbol{\nabla}' \times \boldsymbol{A}_1'\,, \qquad \boldsymbol{E}_1' = -\boldsymbol{\nabla}'\Phi_1' - \frac{\widehat{B}_1\ell_1}{\widehat{E}_1\tau_1}\frac{\partial \boldsymbol{A}_1'}{\partial t'}\,, \tag{2.28}$$

in the scaled variables. We remark that the amplitudes $\widehat{A}_1$ and $\widehat{\Phi}_1$ depend on the length scale ℓ_1 of the fluctuations. This means that, if the sizes of the field fluctuations $\widehat{B}_1$ and $\widehat{E}_1$ are fixed, small-scale fluctuations are associated to small potentials, while large-scale fluctuations are associated to large potentials. The size of the potentials, in turn, plays a role in the ordering of terms in the particle Lagrangian, and thus in the overall asymptotic expansion.

Regarding the Low action, we normalize the electromagnetic free-field Lagrangian (2.7) as

$$L_{\mathrm{EM}} = \widehat{n}_{\mathrm{i}}\, k_{\mathrm{B}}\, \widehat{T}_{\mathrm{i}}\, \widehat{x}^3\, L'_{\mathrm{EM}}\,, \tag{2.29}$$

where $\widehat{n}_{\mathrm{i}}$ denotes a reference ion density, $\widehat{T}_{\mathrm{i}}$ a reference ion temperature, and k_{B} is the Boltzmann constant. Therefore, we obtain the non-dimensional electromagnetic free-field Lagrangian

$$\begin{aligned} L'_{\mathrm{EM}} = {} & \frac{\varepsilon_0\, \widehat{E}_1^2}{\widehat{n}_{\mathrm{i}}\, k_{\mathrm{B}}\, \widehat{T}_{\mathrm{i}}} \frac{1}{2} \int \mathrm{d}^3\boldsymbol{x}' \left| \boldsymbol{\nabla}'\phi'_1 + \frac{\widehat{B}_1 \ell_1}{\widehat{E}_1 \tau_1}\, \frac{\partial \boldsymbol{A}'_1}{\partial t'} \right|^2 \\ & - \frac{\widehat{B}_0^2}{\mu_0\, \widehat{n}_{\mathrm{i}}\, k_{\mathrm{B}}\, \widehat{T}_{\mathrm{i}}} \frac{1}{2} \int \mathrm{d}^3\boldsymbol{x}' \left| \boldsymbol{\nabla}' \times \boldsymbol{A}'_0 + \frac{\widehat{B}_1}{\widehat{B}_0}\, \boldsymbol{\nabla}' \times \boldsymbol{A}'_1 \right|^2 . \end{aligned} \tag{2.30}$$

In order to normalize the single-particle Lagrangian (2.8), we note that $\dot{\boldsymbol{x}}$ represents three components of an element $(\dot{\boldsymbol{x}}, \dot{\boldsymbol{v}}) \in \mathbb{R}^3 \times \mathbb{R}^3$ of the tangent space at $(\boldsymbol{x}, \boldsymbol{v}) \in \mathbb{R}^3 \times \mathbb{R}^3$. Tangents $\dot{\boldsymbol{x}}$ have units of velocity and are therefore normalized as $\dot{\boldsymbol{x}} = (\widehat{x}/\widehat{t}\,)\, \dot{\boldsymbol{x}}'$. We then normalize the single-particle Lagrangian (2.8) as

$$L_{\mathrm{s}} = m_{\mathrm{s}}\, \widehat{v}_{\mathrm{s}}^2\, L'_{\mathrm{s}}\,, \tag{2.31}$$

obtaining the non-dimensional single-particle Lagrangian

$$L'_{\mathrm{s}} = \frac{\widehat{x}}{\widehat{v}_{\mathrm{s}}\, \widehat{t}} \left[\boldsymbol{v}' + \frac{q_{\mathrm{s}} \widehat{B}_0\, \ell_0}{m_{\mathrm{s}} \widehat{v}_{\mathrm{s}}} \left(\boldsymbol{A}'_0 + \frac{\widehat{B}_1 \ell_1}{\widehat{B}_0 \ell_0}\, \boldsymbol{A}'_1 \right) \right] \cdot \dot{\boldsymbol{x}}' - \frac{|\boldsymbol{v}'|^2}{2} - \frac{q_{\mathrm{s}} \widehat{E}_1 \ell_1}{m_{\mathrm{s}} \widehat{v}_{\mathrm{s}}^2} \Phi'_1\,. \tag{2.32}$$

Finally, we normalize the particle Lagrangian (2.6) of the Low action in the same way as the electromagnetic free-field Lagrangian, namely as

$$L_{\mathrm{p}} = \widehat{n}_{\mathrm{i}}\, k_{\mathrm{B}}\, \widehat{T}_{\mathrm{i}}\, \widehat{x}^3\, L'_{\mathrm{p}}\,, \tag{2.33}$$

thus obtaining the non-dimensional particle Lagrangian

$$L'_{\mathrm{p}} = \sum_{\mathrm{s}} \frac{m_{\mathrm{s}}\, \widehat{v}_{\mathrm{s}}^5\, \widehat{f}_s}{\widehat{n}_{\mathrm{i}}\, k_{\mathrm{B}} \widehat{T}_{\mathrm{i}}} \int \mathrm{d}^3\boldsymbol{x}'_0\, \mathrm{d}^3\boldsymbol{v}'_0\, f'_{\mathrm{s}}(t'_0, \boldsymbol{x}'_0, \boldsymbol{v}'_0)\, L'_{\mathrm{s}}\,. \tag{2.34}$$

All dependent variables in the Low action are now expressed in terms of elementary functions. Therefore, the size of each term is determined only by the size of the non-dimensional coefficients in front of it. Such coefficients are, in turn, determined by the physical scenario under consideration, as we discuss in the next section.

2.2.2 Physical scenario and ordering

The process of quantifying the size of the non-dimensional coefficients appearing in the physical quantities of interest, such as the Low action, is usually referred to as ordering. Different orderings lead to different perturbation theories and to reduced models with different physical content. An ordering is thus the mathematical expression of a specific physical scenario. Two such scenarios for magnetic confinement fusion experiments are listed in Tables 2.1 and 2.2. We choose as the characteristic length and time scales of observation the minor radius a and the inverse of the ion thermal frequency $\widehat{\omega}_{\mathrm{i}}$, respectively:

$$\widehat{x} := a\,, \qquad \widehat{t} := \frac{1}{\widehat{\omega}_{\mathrm{i}}} = \frac{a}{\widehat{v}_{\mathrm{i}}}\,. \tag{2.35}$$

The ion thermal frequency $\widehat{\omega}_{\mathrm{i}}$, that is, the inverse of the time required for an ion to travel the distance a, is close to the frequency of micro-turbulence observed in Tokamaks (Liewer, 1985; Wootton et al., 1990; Krommes, 2012). Moreover, measurements in Tokamaks have shown that fluctuation levels in turbulent plasmas satisfy (Liewer, 1985; Wootton et al., 1990; Brizard and Hahm, 2007)

$$\frac{\widehat{B}_1}{\widehat{B}_0} \sim \frac{\widehat{E}_1}{\widehat{B}_0 \widehat{v}_{\mathrm{i}}} \approx 10^{-3}\,, \tag{2.36}$$

which tells that fluctuations are small compared to the corresponding background quantities and that the $\boldsymbol{E} \times \boldsymbol{B}$ velocity is small compared to the ion thermal velocity.

		ions	electrons
major radius	R_0	$1.6\ \mathrm{m}$	
minor radius	a	$0.8\ \mathrm{m}$	
toroidal magnetic field	B_T	$3.9\ \mathrm{T}$	
average particle density	$\langle n_\mathrm{s} \rangle$	$2.0\times\ \dfrac{10^{20}}{\mathrm{m}^3}$	$2.0\times\ \dfrac{10^{20}}{\mathrm{m}^3}$
average thermal energy	$k_\mathrm{B}\langle T_\mathrm{s} \rangle$	$8.7\ \mathrm{keV}$	$8.7\ \mathrm{keV}$
cyclotron frequency	$\omega_\mathrm{cs} := \dfrac{q_\mathrm{s}\widehat{B}_0}{m_\mathrm{s}}$	$1.9 \times 10^{8}\ \mathrm{Hz}$	$6.9 \times 10^{11}\ \mathrm{Hz}$
thermal velocity	$\widehat{v}_\mathrm{s} := \sqrt{\dfrac{k_\mathrm{B}\widehat{T}_\mathrm{s}}{m_\mathrm{s}}}$	$6.4 \times 10^{5}\ \dfrac{\mathrm{m}}{\mathrm{s}}$	$3.9 \times 10^{7}\ \dfrac{\mathrm{m}}{\mathrm{s}}$
thermal frequency	$\widehat{\omega}_\mathrm{s} := \dfrac{\widehat{v}_\mathrm{s}}{a}$	$8.0 \times 10^{5}\ \mathrm{Hz}$	$4.9 \times 10^{7}\ \mathrm{Hz}$
Larmor radius	$\rho_\mathrm{s} := \dfrac{\widehat{v}_\mathrm{s}}{\omega_\mathrm{cs}}$	$3.4 \times 10^{-3}\ \mathrm{m}$	$5.7 \times 10^{-5}\ \mathrm{m}$
Debye length	$\lambda_\mathrm{s} := \sqrt{\dfrac{\varepsilon_0\, k_\mathrm{B}\, \widehat{T}_\mathrm{s}}{q_\mathrm{s}^2\, \widehat{n}_\mathrm{s}}}$	$4.9 \times 10^{-5}\ \mathrm{m}$	$4.9 \times 10^{-5}\ \mathrm{m}$
$\widehat{\omega}_\mathrm{s}/\omega_\mathrm{cs}$		3.4×10^{-3}	5.7×10^{-5}
$\widehat{v}_\mathrm{s}^2/c^2$		4.6×10^{-6}	1.7×10^{-2}
λ_s^2/a^2		3.7×10^{-9}	3.7×10^{-9}
$\beta_\mathrm{s} := \mu_0\, \widehat{n}_\mathrm{s}\, k_\mathrm{B}\, \widehat{T}_\mathrm{s}/\widehat{B}_0^2$		2.3×10^{-2}	2.3×10^{-2}

Table 2.1: Physical scenarios for magnetic confinement fusion experiments: parameters for the Tokamak ASDEX Upgrade. Note that we choose $\widehat{B}_0 = B_\mathrm{T}$, $\widehat{n}_\mathrm{s} = \langle n_\mathrm{s} \rangle$ and $\widehat{T}_\mathrm{s} = \langle T_\mathrm{s} \rangle$.

		ions	electrons
major radius	R_0	$6.2\ \mathrm{m}$	
minor radius	a	$2.0\ \mathrm{m}$	
toroidal magnetic field	B_T	$5.3\ \mathrm{T}$	
average particle density	$\langle n_\mathrm{s} \rangle$	$1.0 \times \dfrac{10^{20}}{\mathrm{m}^3}$	$1.0 \times \dfrac{10^{20}}{\mathrm{m}^3}$
average thermal energy	$k_\mathrm{B} \langle T_\mathrm{s} \rangle$	$8.0\ \mathrm{keV}$	$8.8\ \mathrm{keV}$
cyclotron frequency	$\omega_\mathrm{cs} := \dfrac{q_\mathrm{s} \widehat{B}_0}{m_\mathrm{s}}$	$2.5 \times 10^8\ \mathrm{Hz}$	$9.3 \times 10^{11}\ \mathrm{Hz}$
thermal velocity	$\widehat{v}_\mathrm{s} := \sqrt{\dfrac{k_\mathrm{B} \widehat{T}_\mathrm{s}}{m_\mathrm{s}}}$	$6.2 \times 10^5\ \dfrac{\mathrm{m}}{\mathrm{s}}$	$3.9 \times 10^7\ \dfrac{\mathrm{m}}{\mathrm{s}}$
thermal frequency	$\widehat{\omega}_\mathrm{s} := \dfrac{\widehat{v}_\mathrm{s}}{a}$	$3.1 \times 10^5\ \mathrm{Hz}$	$2.0 \times 10^7\ \mathrm{Hz}$
Larmor radius	$\rho_\mathrm{s} := \dfrac{\widehat{v}_\mathrm{s}}{\omega_\mathrm{cs}}$	$2.4 \times 10^{-3}\ \mathrm{m}$	$4.2 \times 10^{-5}\ \mathrm{m}$
Debye length	$\lambda_\mathrm{s} := \sqrt{\dfrac{\varepsilon_0\, k_\mathrm{B}\, \widehat{T}_\mathrm{s}}{q_\mathrm{s}^2\, \widehat{n}_\mathrm{s}}}$	$6.7 \times 10^{-5}\ \mathrm{m}$	$7.0 \times 10^{-5}\ \mathrm{m}$
$\widehat{\omega}_\mathrm{s} / \omega_\mathrm{cs}$		1.2×10^{-3}	2.1×10^{-5}
$\widehat{v}_\mathrm{s}^2 / c^2$		4.3×10^{-6}	1.7×10^{-2}
$\lambda_\mathrm{s}^2 / a^2$		1.1×10^{-9}	1.2×10^{-9}
$\beta_\mathrm{s} := \mu_0\, \widehat{n}_\mathrm{s}\, k_\mathrm{B}\, \widehat{T}_\mathrm{s} / \widehat{B}_0^2$		5.7×10^{-3}	6.3×10^{-3}

Table 2.2: Physical scenarios for magnetic confinement fusion experiments: parameters for the Tokamak ITER (Sips et al., 2005). Note that we choose $\widehat{B}_0 = B_\mathrm{T}$, $\widehat{n}_\mathrm{s} = \langle n_\mathrm{s} \rangle$ and $\widehat{T}_\mathrm{s} = \langle T_\mathrm{s} \rangle$.

If we consider electrons and deuterium ions, we also have $q_\mathrm{e}/q_\mathrm{i} = -1$ and

$$\frac{m_\mathrm{e}}{m_\mathrm{i}} \approx 2.7 \times 10^{-3} \,. \tag{2.37}$$

Defining the perturbation parameter as in (2.19), we notice an ordering pattern in powers of $\varepsilon \approx 10^{-3}$ in the normalized plasma parameters of Tables 2.1 and 2.2. We then apply this ordering to the normalized Low action. By choosing

$$\ell_0 := a \,, \qquad \ell_1 := \rho_\mathrm{i} \,, \qquad \tau_1 := \frac{1}{\widehat{\omega}_\mathrm{i}} \,, \tag{2.38}$$

we satisfy the maximal ordering $\ell_1/\ell_0 = O(\varepsilon)$ in the limit $\varepsilon \to 0$. The normalized background magnetic field (2.25) and its corresponding magnetic vector potential become standard functions of $\boldsymbol{x}'$, namely $\boldsymbol{B}_0'(\boldsymbol{x}')$ and $\boldsymbol{A}_0'(\boldsymbol{x}')$. On the other hand, since $\ell_1/\widehat{x} = O(\varepsilon)$, the fluctuating electric and magnetic fields and their corresponding potentials become strongly-varying functions of $\boldsymbol{x}'$, namely depending on $\boldsymbol{x}'/\varepsilon$. These assumptions reflect the conventional gyrokinetic ordering $|\boldsymbol{k}_\perp|\rho_\mathrm{i} = O(1)$ and $k_\parallel \rho_\mathrm{i} = O(\varepsilon)$ in the limit $\varepsilon \to 0$, where $\boldsymbol{k} := (k_\parallel, \boldsymbol{k}_\perp)$ denotes a characteristic wave vector. Moreover, for the coefficient appearing in (2.28) we have

$$\frac{\widehat{B}_1 \ell_1}{\widehat{E}_1 \tau_1} = \frac{\widehat{B}_1 \rho_\mathrm{i}\, \widehat{\omega}_\mathrm{i}}{\widehat{E}_1} = \frac{\widehat{\omega}_\mathrm{i}}{\omega_\mathrm{ci}} \frac{\widehat{B}_1 \widehat{v}_\mathrm{i}}{\widehat{E}_1} = \varepsilon \frac{\widehat{B}_1}{\widehat{B}_0} \frac{\widehat{B}_0 \widehat{v}_\mathrm{i}}{\widehat{E}_1} = O(\varepsilon) \,, \tag{2.39}$$

in the limit $\varepsilon \to 0$, which yields

$$\boldsymbol{E}_1' = -\boldsymbol{\nabla}' \Phi_1' - \varepsilon \frac{\partial \boldsymbol{A}_1'}{\partial t'} \,. \tag{2.40}$$

The non-dimensional coefficients in the normalized electromagnetic free-field Lagrangian (2.30) have the following size in the limit $\varepsilon \to 0$:

$$\frac{\varepsilon_0 \widehat{E}_1^2}{\widehat{n}_\mathrm{i} k_\mathrm{B} \widehat{T}_\mathrm{i}} = \frac{\widehat{v}_\mathrm{i}^2}{c^2} \frac{\widehat{E}_1^2}{\widehat{B}_0^2 \widehat{v}_\mathrm{i}^2} \frac{\widehat{B}_0^2}{\mu_0 \widehat{n}_\mathrm{i} k_\mathrm{B} \widehat{T}_\mathrm{i}} = O(\varepsilon^3) \,, \qquad \frac{\widehat{B}_0^2}{\mu_0 \widehat{n}_\mathrm{i} k_\mathrm{B} \widehat{T}_\mathrm{i}} =: \frac{1}{\beta_\mathrm{i}} = O\left(\frac{1}{\varepsilon}\right) \,. \tag{2.41}$$

In the normalized particle Lagrangian (2.34) we set $k_\mathrm{B}\widehat{T}_\mathrm{e} = k_\mathrm{B}\widehat{T}_\mathrm{i}$ and $\widehat{v}_\mathrm{s}^3 \widehat{f}_\mathrm{s} = \widehat{n}_\mathrm{i}$ to fix the characteristic size of the distribution function. Therefore, in our ordering the normalized ion and electron single-particle Lagrangians (2.32) read

$$L'_{\mathrm{i}} = \left(\boldsymbol{v}' + \frac{\boldsymbol{A}'_0}{\varepsilon} + \varepsilon\,\boldsymbol{A}'_1\right)\cdot\dot{\boldsymbol{x}}' - \frac{|\boldsymbol{v}'|^2}{2} - \varepsilon\,\Phi'_1\,, \tag{2.42a}$$

$$L'_{\mathrm{e}} = \left(\sqrt{\frac{m_{\mathrm{e}}}{m_{\mathrm{i}}}}\,\boldsymbol{v}' - \frac{\boldsymbol{A}'_0}{\varepsilon} - \varepsilon\,\boldsymbol{A}'_1\right)\cdot\dot{\boldsymbol{x}}' - \frac{|\boldsymbol{v}'|^2}{2} + \varepsilon\,\Phi'_1\,. \tag{2.42b}$$

The only difference between ions and electrons is given by the square root of the mass ratio multiplying the term $\boldsymbol{v}'\cdot\dot{\boldsymbol{x}}'$ in the symplectic part of the Lagrangian. This defines an intermediate scale that is not an integer power of ε:

$$\sqrt{\frac{m_{\mathrm{e}}}{m_{\mathrm{i}}}} = O(\sqrt{\varepsilon})\,, \tag{2.43}$$

in the limit $\varepsilon \to 0$. As a result of the ordering procedure, we obtain the normalized Low action principle

$$\delta\int_{t'_0}^{t'_1}\mathrm{d}t'\,(L'_{\mathrm{p}} + L'_{\mathrm{EM}}) = 0\,, \tag{2.44}$$

with the Lagrangians L'_{p} and L'_{EM} given by (omitting the primes for a simpler notation)

$$\begin{aligned} L'_{\mathrm{p}} = &\int\mathrm{d}^6\boldsymbol{z}_0\,f_{\mathrm{i}}(t_0,\boldsymbol{z}_0)\,\left[\left(\boldsymbol{v} + \frac{\boldsymbol{A}_0}{\varepsilon} + \varepsilon\,\boldsymbol{A}_1\right)\cdot\frac{\mathrm{d}\Psi^{\mathrm{i}}_t}{\mathrm{d}t} - \frac{|\boldsymbol{v}|^2}{2} - \varepsilon\,\Phi_1\right]_{(\boldsymbol{x},\boldsymbol{v})=\Psi^{\mathrm{i}}_t} \\ &+\int\mathrm{d}^6\boldsymbol{z}_0\,f_{\mathrm{e}}(t_0,\boldsymbol{z}_0)\,\left[\left(\sqrt{\varepsilon}\,\boldsymbol{v} - \frac{\boldsymbol{A}_0}{\varepsilon} - \varepsilon\,\boldsymbol{A}_1\right)\cdot\frac{\mathrm{d}\Psi^{\mathrm{e}}_t}{\mathrm{d}t} - \frac{|\boldsymbol{v}|^2}{2} + \varepsilon\,\Phi_1\right]_{(\boldsymbol{x},\boldsymbol{v})=\Psi^{\mathrm{e}}_t}\,, \\ L'_{\mathrm{EM}} = &\,\frac{\varepsilon^3}{2}\int\mathrm{d}^3\boldsymbol{x}\,\left|\boldsymbol{\nabla}\Phi_1 + \varepsilon\,\frac{\partial\boldsymbol{A}_1}{\partial t}\right|^2 - \frac{1}{2\,\varepsilon}\int\mathrm{d}^3\boldsymbol{x}\,|\boldsymbol{\nabla}\times\boldsymbol{A}_0 + \varepsilon\,\boldsymbol{\nabla}\times\boldsymbol{A}_1|^2\,. \end{aligned}$$

Taking variations with respect to Ψ_t^{i}, Ψ_t^{e}, Φ_1 and $\boldsymbol{A}_1$ in (2.44) we obtain the normalized Vlasov-Maxwell equations (again omitting the primes)

$$\frac{\partial f_{\mathrm{i}}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f_{\mathrm{i}} + \left[\boldsymbol{E}_1 + \boldsymbol{v} \times \left(\frac{\boldsymbol{B}_0}{\varepsilon} + \boldsymbol{B}_1\right)\right] \cdot \frac{\partial f_{\mathrm{i}}}{\partial \boldsymbol{v}} = 0\,, \tag{2.46a}$$

$$\sqrt{\varepsilon}\,\frac{\partial f_{\mathrm{e}}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} f_{\mathrm{e}} - \left[\boldsymbol{E}_1 + \boldsymbol{v} \times \frac{1}{\sqrt{\varepsilon}}\left(\frac{\boldsymbol{B}_0}{\varepsilon} + \boldsymbol{B}_1\right)\right] \cdot \frac{\partial f_{\mathrm{e}}}{\partial \boldsymbol{v}} = 0\,, \tag{2.46b}$$

$$\varepsilon^2\, \boldsymbol{\nabla} \cdot \boldsymbol{E}_1 = \rho\,, \tag{2.46c}$$

$$\boldsymbol{\nabla} \times \left(\frac{\boldsymbol{B}_0}{\varepsilon} + \boldsymbol{B}_1\right) = \boldsymbol{J} + \varepsilon^3\,\frac{\partial \boldsymbol{E}_1}{\partial t}\,, \tag{2.46d}$$

where the normalized charge and current densities are given by

$$\rho = \int \mathrm{d}^3\boldsymbol{v}\, f_{\mathrm{i}} - \int \mathrm{d}^3\boldsymbol{v}\, f_{\mathrm{e}}\,, \qquad \boldsymbol{J} = \int \mathrm{d}^3\boldsymbol{v}\,\boldsymbol{v}\, f_{\mathrm{i}} - \frac{1}{\sqrt{\varepsilon}}\int \mathrm{d}^3\boldsymbol{v}\,\boldsymbol{v}\, f_{\mathrm{e}}\,. \tag{2.47}$$

The factor $1/\sqrt{\varepsilon}$ in front of the electron current density comes from the different choice of scales for the ion and electron thermal velocities. From the two Vlasov equations (2.46a) and (2.46b) we deduce the charge conservation law

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{J} = 0\,, \tag{2.48}$$

also known as the continuity equation in electrodynamics. This is also a solvability condition for Maxwell's equations. Indeed, (2.48) can be derived by taking the divergence of Ampère-Maxwell's law (2.46d), recalling that $\boldsymbol{E}_1(t, \boldsymbol{x}/\varepsilon)$ is strongly-varying in space, and then inserting Coulomb's law (2.46c).

The normalized variational principle (2.44) is a suitable starting point for our perturbation analysis of the Vlasov-Maxwell model, which we discuss in detail in the next section.

2.3 Gyrokinetic reduction

The basic idea of gyrokinetic theory is to replace the exact trajectories of the plasma particles by the trajectories of their gyrocenters, which move on the time scale of the thermal frequency $\widehat{\omega}_\mathrm{i}$ or slower. The dynamics occurring at scales equal to or faster than the cyclotron frequency ω_ci are “averaged out” in the gyrocenter picture. However, some effects of the fast motion of gyration are still present in form of drifts of the gyrocenters. In this section we make these concepts more precise by analyzing the asymptotic limit $\varepsilon \to 0$ in the normalized single-particle Lagrangians (2.42a)-(2.42b). From the reduced Lagrangians we then derive the gyrokinetic Vlasov equation for ions and electrons and define gyrocenter charge and current densities with polarization corrections, thus coupling plasma particles and electromagnetic fields from the normalized gyrocenter action principle. Primes are omitted from now on, in order to keep the notation as simple as possible.

Following (Sugama, 2000), we intend to replace the particle Lagrangian L_p in the Low action principle (2.44) by its gyrocenter representation $\mathcal{L}_\mathrm{p}$:

$$\int \mathrm{d}^6\boldsymbol{z}_0\ \left[f_\mathrm{i}(0,\boldsymbol{z}_0)\,L_\mathrm{i} + f_\mathrm{e}(0,\boldsymbol{z}_0)\,L_\mathrm{e}\right] \approx \int \mathrm{d}^6\boldsymbol{\mathcal{Z}}_0\ \left[B^*_{\|\mathrm{i}}F_\mathrm{i}(0,\boldsymbol{\mathcal{Z}}_0)\,\mathcal{L}_\mathrm{i} + B^*_{\|\mathrm{e}}F_\mathrm{e}(0,\boldsymbol{\mathcal{Z}}_0)\,\mathcal{L}_\mathrm{e}\right]\ ,$$

where $\mathrm{d}^6\boldsymbol{\mathcal{Z}}_0 := \mathrm{d}^3\boldsymbol{X}_0\,\mathrm{d}P_{\|0}\,\mathrm{d}\mu_0\,\mathrm{d}\Theta_0$ denotes the measure in gyrocenter phase space, F_s denotes the gyrocenter distribution function and $\mathcal{L}_\mathrm{s}$ is the corresponding gyrocenter single-particle Lagrangian, to be derived below. The gyrocenter coordinates are the gyrocenter position $\boldsymbol{X} \in \mathbb{R}^3$, the gyrocenter parallel momentum $P_\| \in \mathbb{R}$, the gyrocenter magnetic moment $\mu \in \mathbb{R}^+$ and the gyro-angle $\Theta \in \mathbb{R}$. The single-particle dynamics in the new coordinates is such that the time evolution of the gyro-angle Θ is decoupled from the rest of the coordinates, leading to a closed system of equations for the “slow” variables $(\boldsymbol{X}, P_\|)$, where μ is a constant of the motion. The slow system represents the averaged dynamics.

The phase-space coordinate transformation $\boldsymbol{\mathcal{Z}} := (\boldsymbol{X}, P_\|, \mu, \Theta) \mapsto (\boldsymbol{x}, \boldsymbol{v})$, with Jacobian determinant denoted by $B^*_\|$, is the central object of gyrokinetic theory. It is usually derived in terms of (canonical) Lie transforms of the fundamental one-form associated to the single-particle Lagrangian L_s (Hahm, 1988; Brizard, 1989;

Tronko and Chandre, 2018). Despite being an elegant mathematical framework, Lie transform perturbation theory introduces many formal complications, which seem not to be strictly necessary for averaging. In this thesis we replace Lie transforms with polynomials of finite degree in ε, algebraic in the generating functions. We show that also with this different ansatz for the phase-space near-identity coordinate transformation it is possible to remove the gyro-angle dependence from the single-particle Lagrangian up to the desired order in ε, without changing its symplectic part (and thus the Jacobian $B_\parallel^*$ of the coordinate transformation). Our polynomial transforms are well-defined coordinate transformations (locally invertible), in contrast to the asymptotic series in Lie transform perturbation theory, where it is difficult to prove convergence and existence of the transforms. It is our hope that the simpler derivation based on polynomial transforms will enable more rigorous mathematical studies of gyrokinetic theory in the future.

2.3.1 Preliminary transformations

The phase-space coordinate transformation $\boldsymbol{\mathcal{Z}} \mapsto (\boldsymbol{x}, \boldsymbol{v})$ is a composition of several coordinate changes, which we summarize in Tables 2.3 and 2.4 for ions and electrons, respectively. The first transformation moves the magnetic vector potential $\boldsymbol{A}_1$ from the symplectic part of the single-particle Lagrangian to the Hamiltonian, by defining the "momentum"

$$\boldsymbol{p} := \begin{cases} \boldsymbol{v} + \varepsilon\, \boldsymbol{A}_1 & \text{ions}\,, \\ \boldsymbol{v} - \sqrt{\varepsilon}\boldsymbol{A}_1 & \text{electrons}\,. \end{cases} \tag{2.49}$$

This is a near-identity transformation in $\boldsymbol{v}$ in the limit $\varepsilon \to 0$, with unit Jacobian determinant. It resembles the usual transformation to canonical coordinates, but it does not contain the background magnetic vector potential $\boldsymbol{A}_0$. The new exact

physical coordinates $(\boldsymbol{x}, \boldsymbol{v})$	
Hamiltonian picture $(\boldsymbol{x}, \boldsymbol{p})$	$\boldsymbol{p} := \boldsymbol{v} + \varepsilon\, \boldsymbol{A}_1$
angle coordinates $(\boldsymbol{x}, p_\parallel, \upmu, \theta)$	$p_\parallel := \boldsymbol{p} \cdot \boldsymbol{b}_0$ $\upmu := \dfrac{1}{2B_0} \lvert \boldsymbol{b}_0 \times \boldsymbol{p} \times \boldsymbol{b}_0 \rvert^2$ $\theta := \arctan\left(\dfrac{\boldsymbol{p} \cdot \boldsymbol{e}_1}{\boldsymbol{p} \cdot \boldsymbol{e}_2}\right)$
guiding-center coordinates $\bar{\boldsymbol{Z}} := (\overline{\boldsymbol{X}}, \bar{p}_\parallel, \bar{\upmu}, \bar{\theta})$	$\boldsymbol{x} := \overline{\boldsymbol{X}} + \varepsilon\, \bar{\boldsymbol{\rho}}_{1\mathrm{i}} + \varepsilon^2 \bar{\boldsymbol{\rho}}_{2\mathrm{i}} + \varepsilon^3 \bar{\boldsymbol{\rho}}_{3\mathrm{i}} + \varepsilon^4 \bar{\boldsymbol{\rho}}_{4\mathrm{i}}$ $p_\parallel := \bar{p}_\parallel + \varepsilon\, \bar{G}^{\parallel}_{1\mathrm{i}} + \varepsilon^2 \bar{G}^{\parallel}_{2\mathrm{i}} + \varepsilon^3 \bar{G}^{\parallel}_{3\mathrm{i}}$ $\upmu := \bar{\upmu} + \varepsilon\, \bar{G}^{\upmu}_{1\mathrm{i}} + \varepsilon^2 \bar{G}^{\upmu}_{2\mathrm{i}} + \varepsilon^3 \bar{G}^{\upmu}_{3\mathrm{i}}$ $\theta := \bar{\theta} + \varepsilon\, \bar{G}^{\Theta}_{1\mathrm{i}} + \varepsilon^2 \bar{G}^{\Theta}_{2\mathrm{i}} + \varepsilon^3 \bar{G}^{\Theta}_{3\mathrm{i}}$
preliminary gyrocenter coordinates $\boldsymbol{Z} := (\boldsymbol{X}, P_\parallel, \hat{\upmu}, \Theta)$	$\overline{\boldsymbol{X}} := \boldsymbol{X} + \varepsilon^2 \boldsymbol{\rho}_{2\mathrm{i}} + \varepsilon^3 \boldsymbol{\rho}_{3\mathrm{i}} + \varepsilon^4 \boldsymbol{\rho}_{4\mathrm{i}}$ $\bar{p}_\parallel := P_\parallel + \varepsilon\, G^{\parallel}_{1\mathrm{i}} + \varepsilon^2 G^{\parallel}_{2\mathrm{i}} + \varepsilon^3 G^{\parallel}_{3\mathrm{i}}$ $\bar{\upmu} := \hat{\upmu} + \varepsilon\, G^{\upmu}_{1\mathrm{i}} + \varepsilon^2 G^{\upmu}_{2\mathrm{i}} + \varepsilon^3 G^{\upmu}_{3\mathrm{i}}$ $\bar{\theta} := \Theta + \varepsilon\, G^{\Theta}_{1\mathrm{i}} + \varepsilon^2 G^{\Theta}_{2\mathrm{i}} + \varepsilon^3 G^{\Theta}_{3\mathrm{i}}$
gyrocenter coordinates $\boldsymbol{\mathcal{Z}} := (\boldsymbol{X}, P_\parallel, \mu, \Theta)$	$\mu := \hat{\upmu} + \varepsilon \left\langle \gamma_2^{\Theta} \right\rangle + \varepsilon^2 \left\langle \gamma_3^{\Theta} \right\rangle$

Table 2.3: Coordinate changes involved in the phase-space coordinate transformation $(\boldsymbol{X}, P_\parallel, \mu, \Theta) \mapsto (\boldsymbol{x}, \boldsymbol{v})$ relating physical coordinates and gyrocenter coordinates for ions.

physical coordinates $(\boldsymbol{x}, \boldsymbol{v})$	
Hamiltonian picture $(\boldsymbol{x}, \boldsymbol{p})$	$\boldsymbol{p} := \boldsymbol{v} - \sqrt{\varepsilon}\, \boldsymbol{A}_1$
angle coordinates $(\boldsymbol{x}, p_\parallel, \upmu, \theta)$	$p_\parallel := \boldsymbol{p} \cdot \boldsymbol{b}_0$ $\upmu := \frac{1}{2B_0} \left\lvert \boldsymbol{b}_0 \times \boldsymbol{p} \times \boldsymbol{b}_0 \right\rvert^2$ $\theta := \arctan\left(\frac{\boldsymbol{p} \cdot \boldsymbol{e}_1}{\boldsymbol{p} \cdot \boldsymbol{e}_2}\right)$
guiding-center coordinates $\bar{\boldsymbol{Z}} := (\overline{\boldsymbol{X}}, \bar{p}_\parallel, \bar{\upmu}, \bar{\theta})$	$\boldsymbol{x} := \overline{\boldsymbol{X}} + \varepsilon\, \bar{\boldsymbol{\rho}}_{1\mathrm{e}} + \varepsilon^2 \bar{\boldsymbol{\rho}}_{2\mathrm{e}} + \varepsilon^3 \bar{\boldsymbol{\rho}}_{3\mathrm{e}} + \varepsilon^4 \bar{\boldsymbol{\rho}}_{4\mathrm{e}}$ $p_\parallel := \bar{p}_\parallel + \varepsilon\, \bar{G}^{\parallel}_{1\mathrm{e}} + \varepsilon^2 \bar{G}^{\parallel}_{2\mathrm{e}} + \varepsilon^3 \bar{G}^{\parallel}_{3\mathrm{e}}$ $\upmu := \bar{\upmu} + \varepsilon\, \bar{G}^{\upmu}_{1\mathrm{e}} + \varepsilon^2 \bar{G}^{\upmu}_{2\mathrm{e}} + \varepsilon^3 \bar{G}^{\upmu}_{3\mathrm{e}}$ $\theta := \bar{\theta} + \varepsilon\, \bar{G}^{\Theta}_{1\mathrm{e}} + \varepsilon^2 \bar{G}^{\Theta}_{2\mathrm{e}} + \varepsilon^3 \bar{G}^{\Theta}_{3\mathrm{e}}$
preliminary gyrocenter coordinates $\boldsymbol{Z} := (\boldsymbol{X}, P_\parallel, \hat{\upmu}, \Theta)$	$\overline{\boldsymbol{X}} := \boldsymbol{X} + \varepsilon^2 \boldsymbol{\rho}_{2\mathrm{e}} + \varepsilon^{\frac{5}{2}} \boldsymbol{\rho}_{\frac{5}{2}\mathrm{e}} + \varepsilon^3 \boldsymbol{\rho}_{3\mathrm{e}}\,,$ $\bar{p}_\parallel := P_\parallel + \varepsilon\, G^{\parallel}_{1\mathrm{e}} + \varepsilon^2 G^{\parallel}_{2\mathrm{e}}\,,$ $\bar{\upmu} := \hat{\upmu} + \sqrt{\varepsilon}\, G^{\upmu}_{\frac{1}{2}\mathrm{e}} + \varepsilon\, G^{\upmu}_{1\mathrm{e}}\,,$ $\bar{\theta} := \Theta + \sqrt{\varepsilon}\, G^{\Theta}_{\frac{1}{2}\mathrm{e}} + \varepsilon\, G^{\Theta}_{1\mathrm{e}}\,.$
gyrocenter coordinates $\boldsymbol{\mathcal{Z}} := (\boldsymbol{X}, P_\parallel, \mu, \Theta)$	$\mu := \hat{\upmu} + \sqrt{\varepsilon} \left\langle G^{\upmu}_{\frac{1}{2}\mathrm{e}} \right\rangle - \varepsilon \left\langle \gamma^{\Theta}_3 \right\rangle$

Table 2.4: Coordinate changes involved in the phase-space coordinate transformation $(\boldsymbol{X}, P_\parallel, \mu, \Theta) \mapsto (\boldsymbol{x}, \boldsymbol{v})$ relating physical coordinates and gyrocenter coordinates for electrons.

single-particle Lagrangians read

$$L_{\mathrm{i}} = \left(\boldsymbol{p} + \frac{\boldsymbol{A}_0}{\varepsilon}\right) \cdot \dot{\boldsymbol{x}} - \left(\frac{|\boldsymbol{p}|^2}{2} + \varepsilon\,\Psi_1 + \varepsilon^2 \frac{|\boldsymbol{A}_1|^2}{2}\right), \tag{2.50a}$$

$$L_{\mathrm{e}} = \left(\sqrt{\varepsilon}\,\boldsymbol{p} - \frac{\boldsymbol{A}_0}{\varepsilon}\right) \cdot \dot{\boldsymbol{x}} - \left(\frac{|\boldsymbol{p}|^2}{2} - \sqrt{\varepsilon}\,\Psi_1 + \varepsilon\,\frac{|\boldsymbol{A}_1|^2}{2}\right), \tag{2.50b}$$

where the generalized potential $\Psi_1(t, \boldsymbol{x}, \boldsymbol{p})$ is defined as

$$\Psi_1 := \begin{cases} \Phi_1 - \boldsymbol{p} \cdot \boldsymbol{A}_1 & \text{ions}\,, \\ \sqrt{\varepsilon}\,\Phi_1 - \boldsymbol{p} \cdot \boldsymbol{A}_1 & \text{electrons}\,. \end{cases} \tag{2.51}$$

We remark that the invariance of the single-particle Lagrangians with respect to gauge transformations of the electric and magnetic potentials is broken by this preliminary phase-space coordinate transformation. We then introduce local cylindrical coordinates in $\boldsymbol{p}$-space, namely

$$p_\parallel := \boldsymbol{p} \cdot \boldsymbol{b}_0\,, \tag{2.52a}$$

$$\mu := \frac{1}{2B_0} |\boldsymbol{b}_0 \times \boldsymbol{p} \times \boldsymbol{b}_0|^2\,, \tag{2.52b}$$

$$\theta := \arctan\left(\frac{\boldsymbol{p} \cdot \boldsymbol{e}_1}{\boldsymbol{p} \cdot \boldsymbol{e}_2}\right), \tag{2.52c}$$

where $(\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{b}_0)$ represents a local static orthonormal basis of $\mathbb{R}^3$, given an arbitrary unit vector $\boldsymbol{e}_1$ perpendicular to $\boldsymbol{b}_0$. Denoting by $\boldsymbol{p}_\perp := \boldsymbol{b}_0 \times \boldsymbol{p} \times \boldsymbol{b}_0$ the component of $\boldsymbol{p}$ perpendicular to the local background magnetic field, we have $\boldsymbol{p} = p_\parallel \boldsymbol{b}_0 + \boldsymbol{p}_\perp$, with $\boldsymbol{p}_\perp = (\boldsymbol{p} \cdot \boldsymbol{e}_1)\boldsymbol{e}_1 + (\boldsymbol{p} \cdot \boldsymbol{e}_2)\boldsymbol{e}_2$. From the definition of θ we can write $(\boldsymbol{p} \cdot \boldsymbol{e}_1) = -\sqrt{2\mu B_0}\sin\theta$ and $(\boldsymbol{p} \cdot \boldsymbol{e}_2) = -\sqrt{2\mu B_0}\cos\theta$ and thus define a second θ-dependent orthonormal basis $(\boldsymbol{a}_0, \boldsymbol{b}_0, \boldsymbol{c}_0)$, with

$$\boldsymbol{a}_0 := +\boldsymbol{e}_1 \cos\theta - \boldsymbol{e}_2 \sin\theta\,, \tag{2.53a}$$

$$\boldsymbol{c}_0 := -\boldsymbol{e}_1 \sin\theta - \boldsymbol{e}_2 \cos\theta\,. \tag{2.53b}$$

We note that the basis vectors $\boldsymbol{a}_0$ and $\boldsymbol{c}_0$ satisfy the properties

$$\frac{\partial \boldsymbol{a}_0}{\partial \theta} = \boldsymbol{c}_0\,, \qquad \frac{\partial \boldsymbol{c}_0}{\partial \theta} = -\boldsymbol{a}_0\,, \tag{2.54}$$

which we will use oftentimes in later calculations. The transformation to angle coordinates thus reads

$$\boldsymbol{p} = p_\parallel \boldsymbol{b}_0 + \sqrt{2\mu B_0}\,\boldsymbol{c}_0\,, \tag{2.55}$$

with Jacobian determinant B_0. This leads to the exact Lagrangians

$$L_\mathrm{i} = \left(p_\parallel \boldsymbol{b}_0 + \sqrt{2\mu B_0}\,\boldsymbol{c}_0 + \frac{\boldsymbol{A}_0}{\varepsilon}\right)\cdot\dot{\boldsymbol{x}} - \left(\frac{p_\parallel^2}{2} + \mu B_0 + \varepsilon\,\Psi_1 + \varepsilon^2\,\frac{|\boldsymbol{A}_1|^2}{2}\right)\,, \tag{2.56a}$$

$$L_\mathrm{e} = \left(\sqrt{\varepsilon}\,p_\parallel \boldsymbol{b}_0 + \sqrt{\varepsilon}\sqrt{2\mu B_0}\,\boldsymbol{c}_0 - \frac{\boldsymbol{A}_0}{\varepsilon}\right)\cdot\dot{\boldsymbol{x}} - \left(\frac{p_\parallel^2}{2} + \mu B_0 - \sqrt{\varepsilon}\,\Psi_1 + \varepsilon\,\frac{|\boldsymbol{A}_1|^2}{2}\right)\,. \tag{2.56b}$$

2.3.2 Guiding-center coordinates

The guiding-center phase-space coordinate transformation dates back to the pioneering work of (Littlejohn, 1983) and has the purpose of removing the gyro-angle dependence from those parts of the Lagrangians (2.56a)-(2.56b) that do not depend on the fluctuating potentials Φ_1 and $\boldsymbol{A}_1$. For this reason, the gyrokinetic literature often describes the guiding-center phase-space coordinate transformation as a transformation acting on a single-particle Lagrangian that involves only quantities related to the background magnetic field $\boldsymbol{B}_0$ and does not feature any fluctuating fields. These are said to be added at a later stage, after the guiding-center coordinate transformation has been performed. We believe that this description is slightly misleading, as it seems to suggest the idea that the single-particle Lagrangian is modified by adding terms related to the fluctuating fields during the process of transforming the phase-space coordinates. In fact, the fluctuating fields are present in the single-particle Lagrangian since the beginning of the derivation (as it should be, once we identify the physical system that we want to describe), but their gyro-angle dependence is simply treated at a later stage, after the guiding-center coordinate transformation has been performed. Following (Possanner, 2018), we define the guiding-center coordinate transformation as a polynomial transform of the form

$$\begin{aligned} \boldsymbol{x} &:= \overline{\boldsymbol{X}} + \varepsilon\, \bar{\boldsymbol{\rho}}_{1\mathrm{s}} + \varepsilon^2 \bar{\boldsymbol{\rho}}_{2\mathrm{s}} + \varepsilon^3 \bar{\boldsymbol{\rho}}_{3\mathrm{s}} + \varepsilon^4 \bar{\boldsymbol{\rho}}_{4\mathrm{s}}\,, && \text{(2.57a)} \\ p_\parallel &:= \bar{p}_\parallel + \varepsilon\, \bar{G}^\parallel_{1\mathrm{s}} + \varepsilon^2 \bar{G}^\parallel_{2\mathrm{s}} + \varepsilon^3 \bar{G}^\parallel_{3\mathrm{s}}\,, && \text{(2.57b)} \\ \mu &:= \bar{\mu} + \varepsilon\, \bar{G}^\mu_{1\mathrm{s}} + \varepsilon^2 \bar{G}^\mu_{2\mathrm{s}} + \varepsilon^3 \bar{G}^\mu_{3\mathrm{s}}\,, && \text{(2.57c)} \\ \theta &:= \bar{\theta} + \varepsilon\, \bar{G}^\Theta_{1\mathrm{s}} + \varepsilon^2 \bar{G}^\Theta_{2\mathrm{s}} + \varepsilon^3 \bar{G}^\Theta_{3\mathrm{s}}\,, && \text{(2.57d)} \end{aligned}$$

where $\bar{\boldsymbol{\rho}}_{n\mathrm{s}}$, $\bar{G}^\parallel_{n\mathrm{s}}$, $\bar{G}^\mu_{n\mathrm{s}}$ and $\bar{G}^\Theta_{n\mathrm{s}}$ denote the generators of the coordinate transformation for the respective particle species. The generators are functions of the guiding-center coordinates $\bar{\boldsymbol{Z}} := (\overline{\boldsymbol{X}}, \bar{p}_\parallel, \bar{\mu}, \bar{\theta})$ and may additionally depend on time. The guiding-center coordinates are the guiding-center position $\overline{\boldsymbol{X}}$, the guiding-center parallel momentum $\bar{p}_\parallel$, the guiding-center magnetic moment $\bar{\mu}$ and the guiding-center angle variable $\bar{\theta}$. The idea is then to substitute the coordinate transformation (2.57) in the single-particle Lagrangians (2.56a)-(2.56b), by using the transformation law of vector fields

$$\dot{\boldsymbol{x}} = \dot{\overline{\boldsymbol{X}}} + \sum_{n=1}^{4} \varepsilon^n\, \dot{\bar{\boldsymbol{\rho}}}_{n\mathrm{s}} = \dot{\overline{\boldsymbol{X}}} + \sum_{n=1}^{4} \varepsilon^n \left(\frac{\partial \bar{\boldsymbol{\rho}}_{n\mathrm{s}}}{\partial \bar{\boldsymbol{Z}}} \cdot \dot{\bar{\boldsymbol{Z}}} + \frac{\partial \bar{\boldsymbol{\rho}}_{n\mathrm{s}}}{\partial t} \right)\,, \qquad \text{(2.58)}$$

to obtain the corresponding guiding-center Lagrangians. The work of (Possanner, 2018) showed that the gyro-angle dependence due to the term $\sqrt{2\mu B_0}\, \boldsymbol{c}_0$ can be indeed removed via polynomial transforms in maximal ordering. The resulting guiding-center single-particle Lagrangians read

$$\begin{aligned} L_\mathrm{i} \sim{} & \left(\bar{p}_\parallel \boldsymbol{b}_0 + \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\overline{\boldsymbol{X}}} + \varepsilon\, \bar{\mu}\, \dot{\bar{\theta}} \\ & - \left[\bar{H}_{0\mathrm{i}} + \varepsilon\, \bar{H}_{1\mathrm{i}} + \varepsilon^2 \bar{H}_{2\mathrm{i}} + \bar{O}(\varepsilon^3) \right] + O(\varepsilon^4)\,, \end{aligned} \qquad \text{(2.59a)}$$

$$\begin{aligned} L_\mathrm{e} \sim{} & \left(\sqrt{\varepsilon}\, \bar{p}_\parallel \boldsymbol{b}_0 - \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\overline{\boldsymbol{X}}} - \varepsilon^2\, \bar{\mu}\, \dot{\bar{\theta}} \\ & - \left[\bar{H}_{0\mathrm{e}} + \sqrt{\varepsilon}\, \bar{H}_{\frac{1}{2}\mathrm{e}} + \varepsilon\, \bar{H}_{1\mathrm{e}} + \varepsilon^{\frac{3}{2}} \bar{H}_{\frac{3}{2}\mathrm{e}} + \bar{O}(\varepsilon^2) \right] + O(\varepsilon^4)\,. \end{aligned} \qquad \text{(2.59b)}$$

The symbol "$\sim$" denotes the equivalence between Lagrangians, namely the fact that two Lagrangians differ only by the total differential of some scalar function. Moreover, the symbol $\bar{O}(\varepsilon^p)$ denotes corrections to the Hamiltonians of order $O(\varepsilon^p)$ that are independent of the guiding-center angle $\bar{\theta}$. The ion and electron guiding-center Hamiltonians in (2.59a)-(2.59b) read

$$\begin{aligned}
&\bar{H}_{0\mathrm{i}} := \frac{\bar{p}_\parallel^2}{2} + \bar{\mu} B_0 \,, && \bar{H}_{0\mathrm{e}} := \frac{\bar{p}_\parallel^2}{2} + \bar{\mu} B_0 \,, && (2.60\mathrm{a})\\
&\bar{H}_{1\mathrm{i}} := \Psi_1 + \delta H_1 \,, && \bar{H}_{\frac{1}{2}\mathrm{e}} := \boldsymbol{p}\cdot\boldsymbol{A}_1 \,, && (2.60\mathrm{b})\\
&\bar{H}_{2\mathrm{i}} := \frac{1}{2}\left|\boldsymbol{A}_1\right|^2 + \delta H_2 \,, && \bar{H}_{1\mathrm{e}} := -\Phi_1 + \frac{1}{2}\left|\boldsymbol{A}_1\right|^2 \,, && (2.60\mathrm{c})\\
& && \bar{H}_{\frac{3}{2}\mathrm{e}} := -\delta H_1 \,. && (2.60\mathrm{d})
\end{aligned}$$

We remark the following comments about the guiding-center single-particle Lagrangians (2.59a)-(2.59b):

- the dynamic potentials in the Hamiltonians (2.60) are evaluated at the physical particle position $\boldsymbol{x}/\varepsilon$:

  $$\Phi_1\left(t, \frac{\boldsymbol{x}}{\varepsilon}\right) = \Phi_1\left(t, \frac{\overline{\boldsymbol{X}}}{\varepsilon} + \bar{\boldsymbol{\rho}}_{1\mathrm{s}} + \varepsilon\, \bar{\boldsymbol{\rho}}_{2\mathrm{s}} + O(\varepsilon^2)\right) \,, \qquad (2.61\mathrm{a})$$

  $$\boldsymbol{A}_1\left(t, \frac{\boldsymbol{x}}{\varepsilon}\right) = \boldsymbol{A}_1\left(t, \frac{\overline{\boldsymbol{X}}}{\varepsilon} + \bar{\boldsymbol{\rho}}_{1\mathrm{s}} + \varepsilon\, \bar{\boldsymbol{\rho}}_{2\mathrm{s}} + O(\varepsilon^2)\right) \,. \qquad (2.61\mathrm{b})$$

  The gyro-angle dependence in the generators $\bar{\boldsymbol{\rho}}_{n\mathrm{s}}$ occurring in the arguments of the fluctuating potentials will be removed eventually by the phase-space coordinate transformation from guiding-center to gyrocenter coordinates, as discussed in detail in the next section;

- due to our assumption of maximal ordering, the guiding-center Hamiltonians feature geometric terms related to the curvature of the background magnetic field, in particular

  $$\delta H_1 := \bar{\mu}\left[\frac{\bar{p}_\parallel}{2}\left(\boldsymbol{\nabla}\times\boldsymbol{b}_0\right)\cdot\boldsymbol{b}_0 - \bar{p}_\parallel\left(\boldsymbol{\nabla}\boldsymbol{a}_0\cdot\boldsymbol{c}_0\right)\cdot\boldsymbol{b}_0\right] \,, \qquad (2.62)$$

  and a cumbersome term δH_2, which we do not write explicitly. The two terms in (2.62) are usually referred to as Baños drift (Baños, 1967) and gyro-gauge term, respectively. The curvature terms are less important for the electrons, where δH_1 appears at order $O(\varepsilon^{\frac{3}{2}})$, because of the mass ratio between ions and electrons of order $O(\sqrt{\varepsilon})$;

- for electrons, the magnetic perturbations $\boldsymbol{A}_1$ are $O(\sqrt{\varepsilon})$ larger than the electric perturbations Φ_1. This can be already foreseen in the normalized Vlasov-Maxwell equations (2.46a)-(2.46c) and is due to the mass ratio between ions and electrons. Moreover, it shows the importance of electron dynamics in electromagnetic (rather than electrostatic) gyrokinetic simulations of fusion plasmas;
- due to the error term $\bar{O}(\varepsilon^2)$ in the electron Hamiltonian, the electron guiding-center single-particle Lagrangian (2.59b) is less accurate than the ion guiding-center single-particle Lagrangian (2.59a). This is due to the fact that the guiding-center magnetic moment $\bar{\mu}$ has been computed with less precision for electrons than for ions. We could easily improve the accuracy of $\bar{\mu}$ for electrons, but, as we can see from (2.59b), the dynamic potentials Φ_1 and $\boldsymbol{A}_1$ play a more prominent role than any curvature terms. In the Hamiltonian, the term $\boldsymbol{p} \cdot \boldsymbol{A}_1$ of Ψ_1 appears at order $O(\sqrt{\varepsilon})$ and the term $|\boldsymbol{A}_1|^2/2$ appears at order $O(\varepsilon)$, whereas the first curvature terms appear at $O(\varepsilon^{\frac{3}{2}})$. This is in contrast to the ions, where the first curvature term δH_1 appears already at order $O(\varepsilon)$, which is the same order as Ψ_1 and one order lower than the quadratic term $|\boldsymbol{A}_1|^2/2$. In order to achieve an equally accurate description for the electrons, we should truncate the electron single-particle Lagrangian at order $O(\varepsilon^5)$: this is beyond the scope of the work presented in this thesis, but does not represent a limitation of the method in general;
- the Jacobian determinants J_s of the guiding-center coordinate transformation $\bar{\boldsymbol{Z}} \mapsto (\boldsymbol{x}, \boldsymbol{v})$ can be computed directly from the symplectic part of the guiding-center single-particle Lagrangians (2.59a)-(2.59b):

$$J_\mathrm{i} = B_0 + \varepsilon\, \bar{p}_\parallel \, (\boldsymbol{\nabla} \times \boldsymbol{b}_0) \cdot \boldsymbol{b}_0 \,, \tag{2.63a}$$

$$J_\mathrm{e} = B_0 - \varepsilon^{\frac{3}{2}} \bar{p}_\parallel \, (\boldsymbol{\nabla} \times \boldsymbol{b}_0) \cdot \boldsymbol{b}_0 \,. \tag{2.63b}$$

Such Jacobian determinants are exact because the symplectic forms in (2.59a)-(2.59b) remain the same at any order of the guiding-center expansion, as only the guiding-center Hamiltonian changes with increased order of accuracy (see, for example, (Possanner, 2018) for a proof of this statement). The Jacobian determinants (2.63a) confirm that geometric terms related to the curvature of the background magnetic field appear at order $O(\varepsilon)$ for the ions and at order $O(\varepsilon^{\frac{3}{2}})$ for the electrons, in accordance with the guiding-center Hamiltonians (2.60).

2.3.3 Gyrocenter coordinates

The guiding-center single-particle Lagrangians (2.59a)-(2.59b) obtained from the guiding-center coordinate transformation still carry a dependence on the guiding-center angle $\bar{\theta}$ (the fast variable) in the arguments of the dynamic potentials (2.61). Consequently, the guiding-center magnetic moment $\bar{\mu}$ is not a constant of the motion and the dynamics of slow and fast variables are still coupled in the guiding-center phase space. The purpose of the gyrocenter phase-space coordinate transformation is to remove this residual dependence on the angle variable $\bar{\theta}$ from the Lagrangians, in particular from the Hamiltonians given in (2.60). As for the guiding-center coordinate transformation (2.57), we define the gyrocenter coordinate transformation for the ions as a polynomial transform of the form

$$\overline{\boldsymbol{X}} := \boldsymbol{X} + \varepsilon^2 \boldsymbol{\rho}_{2\mathrm{i}} + \varepsilon^3 \boldsymbol{\rho}_{3\mathrm{i}} + \varepsilon^4 \boldsymbol{\rho}_{4\mathrm{i}}\,, \tag{2.64a}$$

$$\bar{p}_\parallel := P_\parallel + \varepsilon\, G^\parallel_{1\mathrm{i}} + \varepsilon^2 G^\parallel_{2\mathrm{i}} + \varepsilon^3 G^\parallel_{3\mathrm{i}}\,, \tag{2.64b}$$

$$\bar{\mu} := \widehat{\mu} + \varepsilon\, G^\mu_{1\mathrm{i}} + \varepsilon^2 G^\mu_{2\mathrm{i}} + \varepsilon^3 G^\mu_{3\mathrm{i}}\,, \tag{2.64c}$$

$$\bar{\theta} := \Theta + \varepsilon\, G^\Theta_{1\mathrm{i}} + \varepsilon^2 G^\Theta_{2\mathrm{i}} + \varepsilon^3 G^\Theta_{3\mathrm{i}}\,, \tag{2.64d}$$

and the gyrocenter coordinate transformation for the electrons as a polynomial transform of the form

$$\overline{\boldsymbol{X}} := \boldsymbol{X} + \varepsilon^2 \boldsymbol{\rho}_{2\mathrm{e}} \quad + \varepsilon^{\frac{5}{2}} \boldsymbol{\rho}_{\frac{5}{2}\mathrm{e}} + \varepsilon^3 \boldsymbol{\rho}_{3\mathrm{e}} \,, \tag{2.65a}$$

$$\bar{p}_\parallel := P_\parallel + \varepsilon\, G^\parallel_{1\mathrm{e}} \quad + \varepsilon^2 G^\parallel_{2\mathrm{e}} \,, \tag{2.65b}$$

$$\bar{\mu} := \hat{\mu} + \sqrt{\varepsilon}\, G^\mu_{\frac{1}{2}\mathrm{e}} + \varepsilon\, G^\mu_{1\mathrm{e}} \,, \tag{2.65c}$$

$$\bar{\theta} := \Theta + \sqrt{\varepsilon}\, G^\Theta_{\frac{1}{2}\mathrm{e}} + \varepsilon\, G^\Theta_{1\mathrm{e}} \,. \tag{2.65d}$$

Here, $\boldsymbol{Z} := (\boldsymbol{X}, P_\parallel, \hat{\mu}, \Theta)$ denote preliminary gyrocenter coordinates, and $\boldsymbol{\rho}_{n\mathrm{s}}, G^\parallel_{n\mathrm{s}}, G^\mu_{n\mathrm{s}}$ and $G^\Theta_{n\mathrm{s}}$ (with n integer or half-integer) denote the generators of the coordinate transformation for the respective particle species. Our preliminary gyrocenter coordinates are the gyrocenter position $\boldsymbol{X}$, the gyrocenter parallel momentum $P_\parallel$, the preliminary gyrocenter magnetic moment $\hat{\mu}$ and the gyrocenter angle variable (also called simply gyro-angle) Θ. The polynomial transform for the electrons is defined by polynomials in powers of $\sqrt{\varepsilon}$ because of the mass ratio between ions and electrons. Moreover, it consists of fewer terms than the ion coordinate transformation because of the lower accuracy of the electron guiding-center single-particle Lagrangian. We also set $\rho_{1\mathrm{s}} = 0$ a priori: it is, in principle, possible to keep these first-order generators in the calculations and then find out that they can be set to zero without loss of generality. We remark the conceptual simplicity of the polynomial transform $\boldsymbol{Z} \mapsto \bar{\boldsymbol{Z}}$ defined in (2.64)-(2.65) compared to Lie transforms (Brizard and Hahm, 2007): for each coordinate, the transformation is a polynomial of finite degree in ε (the degree being adapted to the desired accuracy of the transformation) and it is moreover linear and algebraic in the generators. By substituting (2.64)-(2.65) in the guiding-center single-particle Lagrangians (2.59a)-(2.59b), the gyrocenter generators can be chosen in order to eliminate the residual dependence on the gyro-angle Θ. The method is analogous to the guiding-center transformation and it is discussed in detail in (Possanner, 2018) for the long-wavelength regime, that is, the case of dynamic potentials with spatial variations on the macroscopic length scale $\hat{x}$. In this thesis we apply the same methodology to the short-wavelength (strongly-varying) regime expressed in (2.61).

The exact same ideas and computations of the guiding-center transformation can be applied also for the gyrocenter transformation. In particular, we make use of the equivalence of Lagrangians under the addition of the total differential $\dot{S}$ of arbitrary scalar functions $S(t, \boldsymbol{Z})$ and write

$$L_\mathrm{i} \sim L_\mathrm{i} \;+ \varepsilon^2 \dot{S}_{2\mathrm{i}} \;+ \varepsilon^3 \dot{S}_{3\mathrm{i}}\,, \tag{2.66a}$$

$$L_\mathrm{e} \sim L_\mathrm{e} + \varepsilon^{\frac{5}{2}} \dot{S}_{\frac{5}{2}\mathrm{e}} + \varepsilon^3 \dot{S}_{3\mathrm{e}}\,, \tag{2.66b}$$

where the total differential $\dot{S}_{ns}$ reads

$$\dot{S}_{ns} = \frac{1}{\varepsilon} \boldsymbol{\nabla}_\perp S_{ns} \cdot \dot{\boldsymbol{X}} + \nabla_\parallel S_{ns}\, \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} + \frac{\partial S_{ns}}{\partial P_\parallel} \dot{P}_\parallel + \frac{\partial S_{ns}}{\partial \widehat{\mu}} \dot{\widehat{\mu}} + \frac{\partial S_{ns}}{\partial \Theta} \dot{\Theta} + \frac{\partial S_{ns}}{\partial t}\,. \tag{2.67}$$

Here, $\nabla_\parallel := \boldsymbol{b}_0 \cdot \boldsymbol{\nabla}$ and $\boldsymbol{\nabla}_\perp := \boldsymbol{b}_0 \times \boldsymbol{\nabla} \times \boldsymbol{b}_0$ denote the gradients with respect to the direction parallel and perpendicular to the background magnetic field, respectively. We remark that in the derivation of the gyrokinetic Lagrangians, the scalar functions S_{ns} turn out to be functions of the fluctuating potentials Φ_1 and $\boldsymbol{A}_1$ and thus have strong variations in the perpendicular directions, which has to be expressed in (2.67) by means of the factor $1/\varepsilon$ in front of $\boldsymbol{\nabla}_\perp$.

Gyrokinetic ion Lagrangian

We summarize our results for ions in the following three propositions. The species index is omitted for more readability.

Proposition 1. The ion guiding-center single-particle Lagrangian (2.59a), expressed in the preliminary gyrocenter coordinates $(\boldsymbol{X}, P_\parallel, \widehat{\mu}, \Theta)$ via the polynomial transform (2.64), is equivalent to

$$L_\mathrm{i} \sim \left(P_\parallel \boldsymbol{b}_0 + \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\boldsymbol{X}} - H_0 + \sum_{n=1}^{3} \varepsilon^n L_n + O(\varepsilon^4)\,, \tag{2.68}$$

where $H_0 := P_\parallel^2/2 + \widehat{\mu}\, B_0$ is the lowest-order Hamiltonian and the Lagrangians L_n, for $n = 1, 2, 3$, read

$$L_n := \boldsymbol{\gamma}_n^{\boldsymbol{X}} \cdot \dot{\boldsymbol{X}} + \gamma_n^{\parallel}\, \dot{P}_{\parallel} + \gamma_n^{\mu}\, \dot{\mu} + \gamma_n^{\Theta}\, \dot{\Theta} - H_n\,. \tag{2.69}$$

The components $\boldsymbol{\gamma}_n^{\boldsymbol{X}}$, for $n = 1, 2, 3$, are given by

$$\boldsymbol{\gamma}_1^{\boldsymbol{X}} := G_1^{\parallel} \boldsymbol{b}_0 - \boldsymbol{\rho}_2 \times \boldsymbol{B}_0 + \boldsymbol{\nabla}_{\perp} S_2\,, \tag{2.70a}$$

$$\boldsymbol{\gamma}_2^{\boldsymbol{X}} := G_2^{\parallel} \boldsymbol{b}_0 - \boldsymbol{\rho}_3 \times \boldsymbol{B}_0 + \boldsymbol{\nabla}_{\perp} S_3 + \nabla_{\parallel} S_2\, \boldsymbol{b}_0 + \boldsymbol{F}_{\perp} + \delta \boldsymbol{\gamma}_2^{\boldsymbol{X}}\,, \tag{2.70b}$$

$$\boldsymbol{\gamma}_3^{\boldsymbol{X}} := G_3^{\parallel} \boldsymbol{b}_0 - \boldsymbol{\rho}_4 \times \boldsymbol{B}_0 + \nabla_{\parallel} S_3\, \boldsymbol{b}_0 + F_{\parallel} \boldsymbol{b}_0 + \delta \boldsymbol{\gamma}_3^{\boldsymbol{X}}\,, \tag{2.70c}$$

where the terms $\boldsymbol{F}_{\perp}$ and $F_{\parallel}$ are defined as

$$\boldsymbol{F}_{\perp} := G_1^{\parallel} \boldsymbol{\nabla}_{\perp} \boldsymbol{\rho}_2 \cdot \boldsymbol{b}_0 - \frac{1}{2} (\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \boldsymbol{\nabla}_{\perp} \boldsymbol{\rho}_2 + G_1^{\mu} \boldsymbol{\nabla}_{\perp} G_1^{\Theta}\,, \tag{2.71a}$$

$$F_{\parallel} := G_1^{\parallel} \nabla_{\parallel} \boldsymbol{\rho}_2 \cdot \boldsymbol{b}_0 - \frac{1}{2} (\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \nabla_{\parallel} \boldsymbol{\rho}_2 + G_1^{\mu} \nabla_{\parallel} G_1^{\Theta}\,, \tag{2.71b}$$

and the terms $\delta \boldsymbol{\gamma}_n^{\boldsymbol{X}}$, for $n = 2, 3$, contain terms related to the curvature of the background magnetic field and are given by

$$\delta \boldsymbol{\gamma}_2^{\boldsymbol{X}} := -\, P_{\parallel}\, \boldsymbol{\rho}_2 \times (\boldsymbol{\nabla} \times \boldsymbol{b}_0)\,, \tag{2.72a}$$

$$\begin{aligned} \delta \boldsymbol{\gamma}_3^{\boldsymbol{X}} := &-\, P_{\parallel}\, \boldsymbol{\rho}_3 \times (\boldsymbol{\nabla} \times \boldsymbol{b}_0) + P_{\parallel}\, (\boldsymbol{\rho}_3 \cdot \boldsymbol{\nabla}) \boldsymbol{b}_0 \\ &+ G_1^{\parallel} (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla}) \boldsymbol{b}_0 + \frac{1}{2} (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla} \boldsymbol{B}_0) \times \boldsymbol{\rho}_2\,. \end{aligned} \tag{2.72b}$$

The components $\gamma_n^{\parallel}$, for $n = 1, 2, 3$, are given by

$$\gamma_1^{\parallel} := 0\,, \tag{2.73a}$$

$$\gamma_2^{\parallel} := -\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_2 + \frac{\partial S_2}{\partial P_{\parallel}}\,, \tag{2.73b}$$

$$\gamma_3^{\parallel} := -\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_3 + \frac{\partial S_3}{\partial P_{\parallel}} + \left(G_1^{\parallel} \boldsymbol{b}_0 - \frac{1}{2}\, \boldsymbol{\rho}_2 \times \boldsymbol{B}_0 \right) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial P_{\parallel}} + G_1^{\mu} \frac{\partial G_1^{\Theta}}{\partial P_{\parallel}}\,. \tag{2.73c}$$

The components γ_n^{μ}, for $n = 1, 2, 3$, are given by

$$\gamma_1^{\mu} := 0\,, \tag{2.74a}$$

$$\gamma_2^{\mu} := -G_1^{\Theta} + \frac{\partial S_2}{\partial \widehat{\mu}}\,, \tag{2.74b}$$

$$\gamma_3^{\mu} := -G_2^{\Theta} + \frac{\partial S_3}{\partial \widehat{\mu}} + \left(G_1^{\parallel} \boldsymbol{b}_0 - \frac{1}{2}\, \boldsymbol{\rho}_2 \times \boldsymbol{B}_0 \right) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \widehat{\mu}} + G_1^{\mu} \frac{\partial G_1^{\Theta}}{\partial \widehat{\mu}}\,. \tag{2.74c}$$

The components γ_n^{Θ}, for $n = 1, 2, 3$, are given by

$$\gamma_1^{\Theta} := \widehat{\mu}\,, \tag{2.75a}$$

$$\gamma_2^{\Theta} := G_1^{\mu} + \frac{\partial S_2}{\partial \Theta}\,, \tag{2.75b}$$

$$\gamma_3^{\Theta} := G_2^{\mu} + \frac{\partial S_3}{\partial \Theta} + \left(G_1^{\parallel} \boldsymbol{b}_0 - \frac{1}{2}\, \boldsymbol{\rho}_2 \times \boldsymbol{B}_0 \right) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \Theta} + G_1^{\mu} \frac{\partial G_1^{\Theta}}{\partial \Theta}\,. \tag{2.75c}$$

Finally, the Hamiltonians H_n, for $n = 1, 2, 3$, are given by

$$H_1 := G_1^{\mu} B_0 + P_{\parallel} G_1^{\parallel} + \Psi_1 + \delta H_1\,, \tag{2.76a}$$

$$\begin{aligned} H_2 := {} & G_2^{\mu} B_0 + \widehat{\mu}\, \boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla} B_0 + P_{\parallel} G_2^{\parallel} + \frac{1}{2} (G_1^{\parallel})^2 + \Bigg[(\bar{\boldsymbol{\rho}}_2 + \boldsymbol{\rho}_2) \cdot \boldsymbol{\nabla} \\ & + (\bar{G}_1^{\parallel} + G_1^{\parallel}) \frac{\mathrm{d}}{\mathrm{d}P_{\parallel}} + (\bar{G}_1^{\mu} + G_1^{\mu}) \frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}} + (\bar{G}_1^{\Theta} + G_1^{\Theta}) \frac{\mathrm{d}}{\mathrm{d}\Theta} \Bigg] \Psi_1 \\ & + \left(G_1^{\parallel} \frac{\mathrm{d}}{\mathrm{d}P_{\parallel}} + G_1^{\mu} \frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}} \right) \delta H_1 + \delta H_2 + \frac{1}{2} |\boldsymbol{A}_1|^2 + \frac{\partial S_2}{\partial t}\,, \end{aligned} \tag{2.76b}$$

$$H_3 := G_3^{\mu} B_0 + \delta H_3\,, \tag{2.76c}$$

where the generalized potential reads

$$\begin{aligned} \Psi_1 \left(t, \frac{\boldsymbol{X}}{\varepsilon} + \bar{\boldsymbol{\rho}}_1(\boldsymbol{Z}), P_{\parallel}, \widehat{\mu}, \Theta \right) = {} & \Phi_1 \left(t, \frac{\boldsymbol{X}}{\varepsilon} + \bar{\boldsymbol{\rho}}_1(\boldsymbol{Z}) \right) - P_{\parallel} A_{1\parallel} \left(t, \frac{\boldsymbol{X}}{\varepsilon} + \bar{\boldsymbol{\rho}}_1(\boldsymbol{Z}) \right) \\ & - \sqrt{2 \widehat{\mu} B_0(\boldsymbol{X})}\, \boldsymbol{c}_0(\boldsymbol{X}, \Theta) \cdot \boldsymbol{A}_{1\perp} \left(t, \frac{\boldsymbol{X}}{\varepsilon} + \bar{\boldsymbol{\rho}}_1(\boldsymbol{Z}) \right)\,, \end{aligned} \tag{2.77}$$

with $A_{1\|} := \boldsymbol{A}_1 \cdot \boldsymbol{b}_0$ and $\boldsymbol{A}_{1\perp} := \boldsymbol{b}_0 \times \boldsymbol{A}_1 \times \boldsymbol{b}_0$. Moreover, δH_1 is the curvature term introduced in (2.62), we do not write the explicit expression of δH_2, and the explicit expression of δH_3 is not relevant for our order of accuracy.

Proof 1. The result is obtained by substituting the gyrocenter coordinate transformation (2.64) into the guiding-center single-particle Lagrangian (2.59a) and computing its Taylor expansion in powers of ε up to order ε^3 (starting from $1/\varepsilon$). We first denote by Γ the symplectic part of the guiding-center Lagrangian (2.59a):

$$\Gamma := \left(\bar{p}_\| \boldsymbol{b}_0 + \frac{\boldsymbol{A}_0}{\varepsilon}\right) \cdot \dot{\bar{\boldsymbol{X}}} + \varepsilon\, \bar{\mu}\, \dot{\bar{\theta}}. \tag{2.78}$$

The coefficients Γ_n, for $n = -1, 0, 1, 2, 3$, of the Taylor expansion of Γ read

$$\Gamma_{-1} := A_0 \cdot \dot{\boldsymbol{X}}\,, \tag{2.79a}$$

$$\Gamma_0 := P_\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}}\,, \tag{2.79b}$$

$$\Gamma_1 := G_1^\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} + (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} + \boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_2 + \widehat{\mu}\, \dot{\Theta}\,, \tag{2.79c}$$

$$\begin{aligned}\Gamma_2 := {} & \Big[G_2^\| + P_\|(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\Big] \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} + (\boldsymbol{\rho}_3 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} + \boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_3 \\ & + P_\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{\rho}}_2 + G_1^\mu \dot{\Theta} + \widehat{\mu}\, \dot{G}_1^\Theta\,,\end{aligned} \tag{2.79d}$$

$$\begin{aligned}\Gamma_3 := {} & \Big[G_3^\| + P_\|(\boldsymbol{\rho}_3 \cdot \boldsymbol{\nabla}) + G_1^\|(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\Big] \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} + (\boldsymbol{\rho}_4 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} \\ & + \boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_4 + \frac{1}{2}(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})^2 \boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} + \Big[G_1^\| \boldsymbol{b}_0 + (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0\Big] \cdot \dot{\boldsymbol{\rho}}_2 \\ & + P_\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{\rho}}_3 + G_2^\mu \dot{\Theta} + \widehat{\mu}\, \dot{G}_2^\Theta + G_1^\mu \dot{G}_1^\Theta\,.\end{aligned} \tag{2.79e}$$

The results for the symplectic part follow by using the equivalence relations (for generic generators $\boldsymbol{\rho}$ and G^{Θ})

$$(\boldsymbol{\rho}\cdot\boldsymbol{\nabla})\boldsymbol{A}_0\cdot\dot{\boldsymbol{X}}+\boldsymbol{A}_0\cdot\dot{\boldsymbol{\rho}}\sim-(\boldsymbol{\rho}\times\boldsymbol{B}_0)\cdot\dot{\boldsymbol{X}}\,, \tag{2.80a}$$

$$P_{\parallel}\boldsymbol{b}_0\cdot\dot{\boldsymbol{\rho}}\sim-P_{\parallel}(\boldsymbol{\nabla}\boldsymbol{b}_0\cdot\boldsymbol{\rho})\cdot\dot{\boldsymbol{X}}-(\boldsymbol{b}_0\cdot\boldsymbol{\rho})\dot{P}_{\parallel}\,, \tag{2.80b}$$

$$\begin{aligned}(\boldsymbol{\rho}\cdot\boldsymbol{\nabla})\boldsymbol{A}_0\cdot\dot{\boldsymbol{\rho}}+\frac{1}{2}(\boldsymbol{\rho}\cdot\boldsymbol{\nabla})^2\boldsymbol{A}_0\cdot\dot{\boldsymbol{X}}\sim{}&-\frac{1}{2}(\boldsymbol{\rho}\times\boldsymbol{B}_0)\cdot\dot{\boldsymbol{\rho}}\\&+\frac{1}{2}\left[(\boldsymbol{\rho}\cdot\boldsymbol{\nabla}\boldsymbol{B}_0)\times\boldsymbol{\rho}\right]\cdot\dot{\boldsymbol{X}}\,,\end{aligned} \tag{2.80c}$$

$$\hat{\mu}\,\dot{G}^{\Theta}\sim-G^{\Theta}\,\dot{\hat{\mu}}\,, \tag{2.80d}$$

together with the vector identity $(\boldsymbol{\rho}\cdot\boldsymbol{\nabla})\boldsymbol{b}_0-\boldsymbol{\nabla}\boldsymbol{b}_0\cdot\boldsymbol{\rho}=-\boldsymbol{\rho}\times(\boldsymbol{\nabla}\times\boldsymbol{b}_0)$, and by adding the terms corresponding to the total differentials $\dot{S}_2$ and $\dot{S}_3$. For the Hamiltonian part, we note that the fluctuating potential Ψ_1 must be first transformed to the guiding-center coordinates $\bar{\boldsymbol{Z}}$ and then to the preliminary gyrocenter coordinates $\boldsymbol{Z}$. We first recall that in physical coordinates we have

$$\begin{aligned}\Psi_1\left(t,\frac{\boldsymbol{x}}{\varepsilon},p_{\parallel},\mu,\theta\right)={}&\Phi_1\left(t,\frac{\boldsymbol{x}}{\varepsilon}\right)-p_{\parallel}A_{1\parallel}\left(t,\frac{\boldsymbol{x}}{\varepsilon}\right)\\&-\sqrt{2\mu\,B_0(\boldsymbol{x})}\,\boldsymbol{c}_0(\boldsymbol{x},\theta)\cdot\boldsymbol{A}_{1\perp}\left(t,\frac{\boldsymbol{x}}{\varepsilon}\right)\,.\end{aligned} \tag{2.81}$$

We shall first substitute the guiding-center coordinate transformation (2.57). Using that Φ_1 and $\boldsymbol{A}_1$ in (2.81) are normalized functions with size and variations of order $O(1)$ in the limit $\varepsilon\to 0$, we can safely expand in a Taylor series around $(\overline{\boldsymbol{X}}/\varepsilon+\bar{\boldsymbol{\rho}}_1,\bar{p}_{\parallel},\bar{\mu},\bar{\theta})$ and obtain

$$\begin{aligned}\Psi_1\left(t,\frac{\boldsymbol{x}}{\varepsilon},p_{\parallel},\mu,\theta\right)={}&\Psi_1\left(t,\frac{\overline{\boldsymbol{X}}}{\varepsilon}+\bar{\boldsymbol{\rho}}_1(\bar{\mu},\bar{\theta}),\bar{p}_{\parallel},\bar{\mu},\bar{\theta}\right)\\&+\varepsilon\left(\bar{\boldsymbol{\rho}}_2\cdot\boldsymbol{\nabla}+\bar{G}_1^{\parallel}\frac{\mathrm{d}}{\mathrm{d}\bar{p}_{\parallel}}+\bar{G}_1^{\mu}\frac{\mathrm{d}}{\mathrm{d}\bar{\mu}}+\bar{G}_1^{\Theta}\frac{\mathrm{d}}{\mathrm{d}\bar{\theta}}\right)\Psi_1+O(\varepsilon^2)\,.\end{aligned} \tag{2.82}$$

The same reasoning applies when we substitute the gyrocenter coordinate transformation (2.64) into (2.82), yielding

$$\begin{aligned}\Psi_1\left(t,\frac{\boldsymbol{x}}{\varepsilon},p_\parallel,\mathrm{\mu},\theta\right) = \Psi_1\left(t,\frac{\boldsymbol{X}}{\varepsilon}+\bar{\boldsymbol{\rho}}_1(\widehat{\mathrm{\mu}},\Theta),P_\parallel,\widehat{\mathrm{\mu}},\Theta\right)& \\ +\varepsilon\bigg[(\bar{\boldsymbol{\rho}}_2+\boldsymbol{\rho}_2)\cdot\boldsymbol{\nabla}+(\bar{G}_1^\parallel+G_1^\parallel)\frac{\mathrm{d}}{\mathrm{d}P_\parallel}+(\bar{G}_1^{\mathrm{\mu}}+G_1^{\mathrm{\mu}})\frac{\mathrm{d}}{\mathrm{d}\widehat{\mathrm{\mu}}}& \\ +(\bar{G}_1^\Theta+G_1^\Theta)\frac{\mathrm{d}}{\mathrm{d}\Theta}\bigg]\Psi_1+O(\varepsilon^2)\,.&\end{aligned} \tag{2.83}$$

This explains the second line of the second-order Hamiltonian H_2 in (2.76) as well as the expression for Ψ_1 given in (2.77).

Proposition 2. In the ion single-particle Lagrangian (2.68) the generators of the polynomial transform can be chosen such that

$$L_{\mathrm{i}} \sim \left(P_\parallel\boldsymbol{b}_0+\frac{\boldsymbol{A}_0}{\varepsilon}\right)\cdot\dot{\boldsymbol{X}}+\varepsilon\left(\widehat{\mathrm{\mu}}+\varepsilon\left\langle\gamma_2^\Theta\right\rangle+\varepsilon^2\left\langle\gamma_3^\Theta\right\rangle\right)\dot{\Theta}-H_0+O(\varepsilon^4)\,, \tag{2.84}$$

where, for a given function $g(\Theta)$, $\langle g\rangle$ denotes its gyro-average and is defined as

$$\langle g\rangle := \frac{1}{2\pi}\int_0^{2\pi}\mathrm{d}\Theta\,g(\Theta)\,. \tag{2.85}$$

Proof 2. The components $\boldsymbol{\gamma}_n^{\boldsymbol{X}}$, for $n=1,2,3$, in (2.70) vanish if and only if we set

$$G_1^\parallel = 0\,, \tag{2.86a}$$

$$G_2^\parallel = -\nabla_\parallel S_2-\delta\gamma_{2\parallel}^{\boldsymbol{X}}\,, \tag{2.86b}$$

$$G_3^\parallel = -\nabla_\parallel S_3-F_\parallel-\delta\gamma_{3\parallel}^{\boldsymbol{X}}\,, \tag{2.86c}$$

as well as

$$\boldsymbol{\rho}_{2\perp} = \frac{\boldsymbol{b}_0}{B_0} \times \boldsymbol{\nabla}_\perp S_2 \,, \tag{2.87a}$$

$$\boldsymbol{\rho}_{3\perp} = \frac{\boldsymbol{b}_0}{B_0} \times \left(\boldsymbol{\nabla}_\perp S_3 + \boldsymbol{F}_\perp + \delta\boldsymbol{\gamma}^{\boldsymbol{X}}_{2\perp}\right) \,, \tag{2.87b}$$

$$\boldsymbol{\rho}_{4\perp} = \frac{\boldsymbol{b}_0}{B_0} \times \delta\boldsymbol{\gamma}^{\boldsymbol{X}}_{3\perp} \,. \tag{2.87c}$$

The components $\gamma^{\parallel}_n$, for $n = 2, 3$, in (2.73) vanish if and only if we set

$$\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_2 = \frac{\partial S_2}{\partial P_\parallel} \,, \tag{2.88a}$$

$$\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_3 = \frac{\partial S_3}{\partial P_\parallel} + \left(G^{\parallel}_1 \boldsymbol{b}_0 - \frac{1}{2}\,\boldsymbol{\rho}_2 \times \boldsymbol{B}_0\right) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial P_\parallel} + G^{\mu}_1 \frac{\partial G^{\Theta}_1}{\partial P_\parallel} \,. \tag{2.88b}$$

The components γ^{μ}_n, for $n = 2, 3$, in (2.74) vanish if and only if we set

$$G^{\Theta}_1 = \frac{\partial S_2}{\partial \hat{\mu}} \,, \tag{2.89a}$$

$$G^{\Theta}_2 = \frac{\partial S_3}{\partial \hat{\mu}} + \left(G^{\parallel}_1 \boldsymbol{b}_0 - \frac{1}{2}\,\boldsymbol{\rho}_2 \times \boldsymbol{B}_0\right) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \hat{\mu}} + G^{\mu}_1 \frac{\partial G^{\Theta}_1}{\partial \hat{\mu}} \,. \tag{2.89b}$$

The Hamiltonians H_n, for $n = 1, 2, 3$, in (2.76) vanish if and only if we set

$$G^{\mu}_1 = -\,\frac{1}{B_0}\left(P_\parallel G^{\parallel}_1 + \Psi_1 + \delta H_1\right) \,, \tag{2.90a}$$

$$\begin{aligned} G^{\mu}_2 = -\,\frac{1}{B_0}\Bigg\{&\hat{\mu}\,\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla} B_0 + P_\parallel G^{\parallel}_2 + \frac{1}{2}(G^{\parallel}_1)^2 + \Bigg[(\bar{\boldsymbol{\rho}}_2 + \boldsymbol{\rho}_2) \cdot \boldsymbol{\nabla} \\ &+ (\bar{G}^{\parallel}_1 + G^{\parallel}_1)\frac{\mathrm{d}}{\mathrm{d}P_\parallel} + (\bar{G}^{\mu}_1 + G^{\mu}_1)\frac{\mathrm{d}}{\mathrm{d}\hat{\mu}} + (\bar{G}^{\Theta}_1 + G^{\Theta}_1)\frac{\mathrm{d}}{\mathrm{d}\Theta}\Bigg]\Psi_1 \\ &+ \left(G^{\parallel}_1 \frac{\mathrm{d}}{\mathrm{d}P_\parallel} + G^{\mu}_1 \frac{\mathrm{d}}{\mathrm{d}\hat{\mu}}\right)\delta H_1 + \delta H_2 + \frac{1}{2}|\boldsymbol{A}_1|^2 + \frac{\partial S_2}{\partial t}\Bigg\} \,, \end{aligned} \tag{2.90b}$$

$$G^{\mu}_3 = -\,\frac{\delta H_3}{B_0} \,. \tag{2.90c}$$

The only degrees of freedom left are the arbitrary scalar functions S_2 and S_3. Since these functions must be 2π-periodic in the gyro-angle Θ, we cannot eliminate γ_2^Θ and γ_3^Θ, given by (2.75), entirely from the Lagrangian. The reason is that the equation $\partial S_n / \partial\Theta = g$, for a given function g, has 2π-periodic solutions S_n if and only if $\langle g \rangle = 0$, where $\langle g \rangle$ denotes the gyro-average of g defined in (2.85). Denoting by $\widetilde{g}$ the fluctuating part of g (with zero gyro-average),

$$\widetilde{g} := g - \langle g \rangle \,, \tag{2.91}$$

the dependence on the gyro-angle Θ can be removed from (2.75) by setting, for $n = 2, 3$,

$$\gamma_n^\Theta = \left\langle \gamma_n^\Theta \right\rangle \,, \tag{2.92}$$

or, equivalently, by requiring that S_n, for $n = 2, 3$, satisfy the differential equations

$$\frac{\partial S_2}{\partial\Theta} = -\widetilde{G_1^{\mu}} \,, \tag{2.93a}$$

$$\frac{\partial S_3}{\partial\Theta} = -\widetilde{G_2^{\mu}} - \widetilde{G_1^{\parallel} \boldsymbol{b}_0 \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial\Theta}} + \frac{1}{2} \widetilde{\boldsymbol{\rho}_2 \times \boldsymbol{B}_0} \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial\Theta} - \widetilde{G_1^{\mu} \frac{\partial G_1^\Theta}{\partial\Theta}} \,. \tag{2.93b}$$

The solutions of (2.93) read (with arbitrary lower bound Θ_0 of integration)

$$S_2(\Theta) = S_2(\Theta_0) - \int_{\Theta_0}^{\Theta} \mathrm{d}\Theta' \, \widetilde{G_1^{\mu}} \,, \tag{2.94a}$$

$$\begin{aligned} S_3(\Theta) &= S_3(\Theta_0) \\ &\quad - \int_{\Theta_0}^{\Theta} \mathrm{d}\Theta' \left(\widetilde{G_2^{\mu}} + \widetilde{G_1^{\parallel} \boldsymbol{b}_0 \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial\Theta}} - \frac{1}{2} \widetilde{\boldsymbol{\rho}_2 \times \boldsymbol{B}_0} \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial\Theta} + \widetilde{G_1^{\mu} \frac{\partial G_1^\Theta}{\partial\Theta}} \right) \,. \end{aligned} \tag{2.94b}$$

Proposition 3. The generalized gyrocenter magnetic moment

$$\mu := \widehat{\mu} + \varepsilon \left\langle \gamma_2^\Theta \right\rangle + \varepsilon^2 \left\langle \gamma_3^\Theta \right\rangle \tag{2.95}$$

is a constant of the motion, accurate up to order $O(\varepsilon^2)$, with respect to the dynamics induced by the preliminary gyrocenter single-particle Lagrangian (2.84). Moreover, there is a one-to-one correspondence $\mu \mapsto \widehat{\mu}$, which implies that the gyrocenter

Hamiltonian $\mathcal{H}$ is given by

$$\mathcal{H} := \frac{P_\parallel^2}{2} + \hat{\upmu}(\mu) B_0 \,. \tag{2.96}$$

In other words, the gyrocenter Hamiltonian is obtained from H_0 by inverting the transformation $\hat{\upmu} \mapsto \mu$ defined in (2.95). By expressing (2.84) in terms of the new gyrocenter coordinates $\boldsymbol{\mathcal{Z}} := (\boldsymbol{X}, P_\parallel, \mu, \Theta)$, we obtain the ion gyrocenter single-particle Lagrangian

$$\mathcal{L}_\mathrm{i} \sim \left(P_\parallel \boldsymbol{b}_0 + \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\boldsymbol{X}} + \varepsilon \mu \, \dot{\Theta} - \mathcal{H} + \bar{O}(\varepsilon^3) + O(\varepsilon^4) \,, \tag{2.97}$$

where the gyrocenter Hamiltonian is defined as $\mathcal{H} := \mathcal{H}_0 + \varepsilon \, \mathcal{H}_1 + \varepsilon^2 \mathcal{H}_2$, with $\mathcal{H}_n$, for $n = 0, 1, 2$, given by

$$\mathcal{H}_0 := \frac{P_\parallel^2}{2} + \mu B_0 \,, \tag{2.98a}$$

$$\mathcal{H}_1 := \langle \Psi_1 \rangle + \delta H_1 \,, \tag{2.98b}$$

$$\begin{aligned} \mathcal{H}_2 := {} & \frac{1}{2} \left\langle |\boldsymbol{A}_1|^2 \right\rangle - \frac{1}{2B_0} \frac{\mathrm{d}}{\mathrm{d}\mu} \left\langle \widetilde{\Psi_1}^2 \right\rangle \\ & - \frac{1}{2B_0^2} \left\langle \left(\boldsymbol{b}_0 \times \nabla_\perp \widetilde{\Psi_1} \right) \cdot \boldsymbol{\nabla}_\perp \int^\Theta \mathrm{d}\Theta' \, \widetilde{\Psi_1} \right\rangle \\ & - \frac{1}{B_0} \left\langle \nabla_\parallel \widetilde{\Psi_1} \int^\Theta \mathrm{d}\Theta' \, \widetilde{A_{1\parallel}} \right\rangle + \delta \mathcal{G}_2 + \delta H_2 \,. \end{aligned} \tag{2.98c}$$

Here, we introduced a new term $\delta\mathcal{G}_2$ related to the curvature of the background magnetic field, besides δH_1 and δH_2:

$$\begin{aligned} \delta \mathcal{G}_2 := \delta G_2 & - \frac{1}{2} \left\langle \frac{\widetilde{\Psi_1}}{B_0} \left(\frac{\boldsymbol{b}_0}{B_0} \times \boldsymbol{\nabla}_\perp B_0 \right) \cdot \boldsymbol{\nabla}_\perp \int^\Theta \mathrm{d}\Theta' \, \frac{\widetilde{\Psi_1}}{B_0} \right\rangle \\ & + \frac{1}{2} \left\langle \left(\frac{\boldsymbol{b}_0}{B_0^3} \times \boldsymbol{\nabla}_\perp \widetilde{\Psi_1} \right) \cdot \boldsymbol{\nabla}_\perp B_0 \int^\Theta \mathrm{d}\Theta' \, \widetilde{\Psi_1} \right\rangle \,, \end{aligned} \tag{2.99}$$

with δG_2 defined as

$$\delta G_2 := \left\langle \left(\bar{\boldsymbol{\rho}}_2 \cdot \boldsymbol{\nabla} + \bar{G}_1^\parallel \frac{\mathrm{d}}{\mathrm{d}P_\parallel} + \bar{G}_1^\upmu \frac{\mathrm{d}}{\mathrm{d}\hat{\upmu}} + \bar{G}_1^\Theta \frac{\mathrm{d}}{\mathrm{d}\Theta} \right) \Psi_1 \right\rangle \,. \tag{2.100}$$

The term δG_2 is linear in the fluctuating potential Ψ_1 and couples to higher-order generators of the guiding-center transformation (Parra and Calvo, 2011). Moreover, the symbol $\bar{O}(\varepsilon^3)$ in (2.97) denotes terms of order $O(\varepsilon^3)$ that are independent of the gyro-angle Θ.

Proof 3. Computing the Euler-Lagrange equation

$$\frac{\partial L_{\mathrm{i}}}{\partial \Theta} = \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L_{\mathrm{i}}}{\partial \dot{\Theta}}\,, \tag{2.101}$$

for the preliminary gyrocenter single-particle Lagrangian (2.84), and noting that $\bar{O}(\varepsilon^3)$ is independent of Θ, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\hat{\upmu} + \varepsilon\left\langle\gamma_2^\Theta\right\rangle + \varepsilon^2\left\langle\gamma_3^\Theta\right\rangle\right) = O(\varepsilon^3)\,. \tag{2.102}$$

Hence, the gyrocenter magnetic moment μ defined in (2.95) is conserved with second-order accuracy in ε. Let us now compute the terms $\left\langle\gamma_2^\Theta\right\rangle$ and $\left\langle\gamma_3^\Theta\right\rangle$ that define the transformation $\hat{\upmu} \mapsto \mu$ in (2.95). From (2.75) and (2.90) we have

$$\left\langle\gamma_2^\Theta\right\rangle = \langle G_1^{\upmu}\rangle = -\frac{\langle\Psi_1\rangle}{B_0} - \frac{\delta H_1}{B_0}\,, \tag{2.103}$$

where we used the result $G_1^{\parallel} = 0$ from (2.86) and the fact that the geometric term δH_1 does not depend on the gyro-angle. Moreover, from (2.75) we have

$$\left\langle\gamma_3^\Theta\right\rangle = \langle G_2^{\upmu}\rangle - \frac{1}{2}\left\langle\frac{\partial \boldsymbol{\rho}_2}{\partial\Theta}\cdot(\boldsymbol{\rho}_2\times\boldsymbol{B}_0)\right\rangle + \left\langle\frac{\partial G_1^\Theta}{\partial\Theta}G_1^{\upmu}\right\rangle\,. \tag{2.104}$$

The gyro-average of $G_2^{\upmu}$ can be computed from (2.90), obtaining

$$\begin{aligned}\langle G_2^{\upmu}\rangle = &-\frac{1}{B_0}\left\langle P_\parallel G_2^\parallel + \left(\boldsymbol{\rho}_2\cdot\boldsymbol{\nabla} + G_1^{\upmu}\frac{\mathrm{d}}{\mathrm{d}\hat{\upmu}} + G_1^\Theta\frac{\mathrm{d}}{\mathrm{d}\Theta}\right)\Psi_1 + \frac{1}{2}|\boldsymbol{A}_1|^2\right\rangle\\ &-\frac{\langle G_1^{\upmu}\rangle}{B_0}\frac{\mathrm{d}}{\mathrm{d}\hat{\upmu}}\delta H_1 - \frac{\delta H_2}{B_0} - \frac{\delta G_2}{B_0}\,,\end{aligned} \tag{2.105}$$

where we used $\langle\boldsymbol{\rho}_2\rangle = 0$. In order to compute the second term on the right-hand side of (2.104), the generator $\boldsymbol{\rho}_2$ is determined by the function S_2 via (2.87) and

(2.88). Omitting the arbitrary lower bound of integration Θ_0 in (2.94), we have

$$S_2 = -\int^{\Theta} \mathrm{d}\Theta' \, \widetilde{G_1^{\mu}} = \int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{\Psi_1}}{B_0} \, , \tag{2.106}$$

and, recalling the functional form of Ψ_1 in (2.81), we obtain

$$\frac{\partial S_2}{\partial P_\parallel} = -\int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{A_{1\parallel}}}{B_0} \, . \tag{2.107}$$

Therefore, the generator $\boldsymbol{\rho}_2$ reads

$$\boldsymbol{\rho}_2 = \frac{\boldsymbol{b}_0}{B_0} \times \boldsymbol{\nabla}_\perp \int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{\Psi_1}}{B_0} - \boldsymbol{b}_0 \int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{A_{1\parallel}}}{B_0} \, , \tag{2.108}$$

which leads to

$$-\frac{1}{2} \left\langle \frac{\partial \boldsymbol{\rho}_2}{\partial \Theta} \cdot (\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \right\rangle = -\frac{1}{2} \left\langle \left(\frac{\boldsymbol{b}_0}{B_0} \times \boldsymbol{\nabla}_\perp \frac{\widetilde{\Psi_1}}{B_0} \right) \cdot \boldsymbol{\nabla}_\perp \int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{\Psi_1}}{B_0} \right\rangle \, . \tag{2.109}$$

In order to compute the last term in (2.104), we get from (2.89)

$$G_1^{\Theta} = \frac{\partial S_2}{\partial \hat{\mu}} = \int^{\Theta} \mathrm{d}\Theta' \, \frac{\mathrm{d}}{\mathrm{d}\hat{\mu}} \frac{\widetilde{\Psi_1}}{B_0} \, , \tag{2.110}$$

yielding

$$\left\langle \frac{\partial G_1^{\Theta}}{\partial \Theta} G_1^{\mu} \right\rangle = -\frac{1}{2 B_0^2} \frac{\mathrm{d}}{\mathrm{d}\hat{\mu}} \left\langle \widetilde{\Psi_1}^2 \right\rangle \, . \tag{2.111}$$

In order to get an explicit expression for $\left\langle \gamma_3^{\Theta} \right\rangle$, we need to compute the right-hand side of (2.105) term by term. Using (2.86) and the fact that $\left\langle S_2 \right\rangle = 0$ and $\left\langle \delta\gamma_2^{\boldsymbol{X}} \right\rangle = 0$

from (2.72), we find $\left\langle G_2^{\parallel} \right\rangle = 0$. The second to fourth terms in (2.105) read

$$\begin{aligned} -\frac{1}{B_0}\left\langle \boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla}\Psi_1 \right\rangle = & \left\langle \left(\frac{\boldsymbol{b}_0}{B_0^2} \times \boldsymbol{\nabla}_{\perp}\widetilde{\Psi_1} \right) \cdot \boldsymbol{\nabla}_{\perp} \int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{\Psi_1}}{B_0} \right\rangle \\ & + \frac{1}{B_0} \left\langle \nabla_{\parallel}\widetilde{\Psi_1} \int^{\Theta} \mathrm{d}\Theta' \, \frac{\widetilde{A_{1\parallel}}}{B_0} \right\rangle , \end{aligned} \tag{2.112a}$$

$$-\frac{1}{B_0}\left\langle G_1^{\mu} \frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}} \Psi_1 \right\rangle = \frac{1}{2B_0^2}\frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}}\left\langle \Psi_1^2 \right\rangle + \frac{\delta H_1}{B_0^2}\frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}}\left\langle \Psi_1 \right\rangle , \tag{2.112b}$$

$$-\frac{1}{B_0}\left\langle G_1^{\Theta} \frac{\mathrm{d}}{\mathrm{d}\Theta} \Psi_1 \right\rangle = \frac{1}{2B_0^2}\frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}}\left\langle \widetilde{\Psi_1}^2 \right\rangle , \tag{2.112c}$$

where we integrated by parts in order to obtain the last equality. Substitution of (2.109), (2.111)-(2.112) into (2.104) yields

$$\begin{aligned} \left\langle \gamma_3^{\Theta} \right\rangle = & -\frac{1}{2B_0}\left\langle |\boldsymbol{A}_1|^2 \right\rangle + \frac{1}{2B_0^2}\frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}}\left\langle \Psi_1^2 \right\rangle \\ & + \frac{1}{2}\left\langle \left(\frac{\boldsymbol{b}_0}{B_0^3} \times \boldsymbol{\nabla}_{\perp}\widetilde{\Psi_1} \right) \cdot \boldsymbol{\nabla}_{\perp} \int^{\Theta} \mathrm{d}\Theta' \, \widetilde{\Psi_1} \right\rangle \\ & + \frac{1}{B_0^2}\left\langle \nabla_{\parallel}\widetilde{\Psi_1} \int^{\Theta} \mathrm{d}\Theta' \, \widetilde{A_{1\parallel}} \right\rangle - \frac{\delta \mathcal{G}_2}{B_0} \\ & + \frac{1}{B_0^2}\frac{\mathrm{d}}{\mathrm{d}\widehat{\mu}}\left(\left\langle \Psi_1 \right\rangle \delta H_1 + \frac{1}{2}\delta H_1^2 \right) - \frac{\delta H_2}{B_0} . \end{aligned} \tag{2.113}$$

The generalized magnetic moment μ can now be computed explicitly as a function of the fluctuating potentials from (2.103) and (2.113). It remains to identify the gyrocenter Hamiltonian. For this purpose, we need to invert (2.95) and substitute the result into H_0. From

$$\mu = \widehat{\mu} + \varepsilon \left\langle \gamma_2^{\Theta} \right\rangle (\widehat{\mu}) + \varepsilon^2 \left\langle \gamma_3^{\Theta} \right\rangle (\widehat{\mu}) , \tag{2.114}$$

we obtain

$$\begin{aligned}\hat{\upmu} &= \mu - \varepsilon \left\langle \gamma_2^\Theta \right\rangle \big(\mu - \varepsilon \left\langle \gamma_2^\Theta \right\rangle (\mu)\big) - \varepsilon^2 \left\langle \gamma_3^\Theta \right\rangle (\mu) + \bar{O}(\varepsilon^3) \\ &= \mu - \varepsilon \left\langle \gamma_2^\Theta \right\rangle (\mu) + \varepsilon^2 \left\langle \gamma_2^\Theta \right\rangle (\mu) \frac{\mathrm{d}}{\mathrm{d}\mu} \left\langle \gamma_2^\Theta \right\rangle (\mu) - \varepsilon^2 \left\langle \gamma_3^\Theta \right\rangle (\mu) + \bar{O}(\varepsilon^3)\,.\end{aligned} \tag{2.115}$$

At order $O(\varepsilon^2)$ we compute the difference

$$\left\langle \gamma_3^\Theta \right\rangle - \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\mu} \left\langle \gamma_2^\Theta \right\rangle^2 = \left\langle \gamma_3^\Theta \right\rangle - \frac{1}{2B_0^2}\frac{\mathrm{d}}{\mathrm{d}\mu} \left(\langle \Psi_1 \rangle + \delta H_1 \right)^2 \,, \tag{2.116}$$

and obtain, using the explicit formulas (2.103) and (2.113) for $\left\langle \gamma_2^\Theta \right\rangle$ and $\left\langle \gamma_3^\Theta \right\rangle$,

$$\begin{aligned}\hat{\upmu} = \mu &+ \frac{\varepsilon}{B_0}(\langle \Psi_1 \rangle + \delta H_1) + \frac{\varepsilon^2}{B_0}\Bigg(\frac{1}{2}\left\langle |\boldsymbol{A}_1|^2 \right\rangle - \frac{1}{2B_0}\frac{\mathrm{d}}{\mathrm{d}\hat{\upmu}} \left\langle \widetilde{\Psi_1}^2 \right\rangle \\ &- \frac{1}{2B_0^2} \left\langle \left(\boldsymbol{b}_0 \times \boldsymbol{\nabla}_\perp \widetilde{\Psi_1} \right) \cdot \boldsymbol{\nabla}_\perp \int^\Theta \mathrm{d}\Theta' \, \widetilde{\Psi_1} \right\rangle \\ &- \frac{1}{B_0} \left\langle \nabla_\parallel \widetilde{\Psi_1} \int^\Theta \mathrm{d}\Theta' \, \widetilde{A_{1\parallel}} \right\rangle + \delta \mathcal{G}_2 + \delta H_2 \Bigg) + \bar{O}(\varepsilon^3)\,,\end{aligned} \tag{2.117}$$

where we used the fact that $\langle g \rangle^2 - \langle g^2 \rangle = -\langle \widetilde{g}^2 \rangle$, for a given function $g(\Theta)$. Substituting this into the Hamiltonian H_0 completes the proof.

Gyrokinetic electron Lagrangian

We summarize our results for electrons in the following three propositions. The species index is again omitted for more readability.

Proposition 1. The electron guiding-center single-particle Lagrangian (2.59b), expressed in the preliminary gyrocenter coordinates $(\boldsymbol{X}, P_\parallel, \hat{\upmu}, \Theta)$ via the polynomial transform (2.65), is equivalent to

$$L_\mathrm{e} \sim \left(\sqrt{\varepsilon} P_\parallel \boldsymbol{b}_0 - \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\boldsymbol{X}} - H_0 - \sqrt{\varepsilon} H_{\frac{1}{2}} + \sum_{n=2}^{6} \varepsilon^{\frac{n}{2}} L_{\frac{n}{2}} + O(\varepsilon^{\frac{7}{2}})\,, \tag{2.118}$$

where $H_0 := P_\|^2/2 + \widehat{\mu}\, B_0$ is the lowest-order Hamiltonian, the Hamiltonian $H_{\frac{1}{2}}$ reads

$$H_{\frac{1}{2}} := G^{\mu}_{\frac{1}{2}} B_0 + P_\| A_{1\|} + \sqrt{2\widehat{\mu} B_0}\, \boldsymbol{c}_0 \cdot \boldsymbol{A}_{1\perp} \,, \tag{2.119}$$

and the Lagrangians $L_{\frac{n}{2}}$, for $n = 2, \ldots, 6$, read

$$L_{\frac{n}{2}} := \boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{n}{2}} \cdot \dot{\boldsymbol{X}} + \gamma^{\|}_{\frac{n}{2}}\, \dot{P}_\| + \gamma^{\mu}_{\frac{n}{2}}\, \dot{\widehat{\mu}} + \gamma^{\Theta}_{\frac{n}{2}}\, \dot{\Theta} - H_{\frac{n}{2}} \,. \tag{2.120}$$

The components $\boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{n}{2}}$, for $n = 2, \ldots, 6$, are given by

$$\boldsymbol{\gamma}^{\boldsymbol{X}}_1 := \boldsymbol{\rho}_2 \times \boldsymbol{B}_0 \,, \tag{2.121a}$$

$$\boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{3}{2}} := G^{\|}_1 \boldsymbol{b}_0 + \boldsymbol{\rho}_{\frac{5}{2}} \times \boldsymbol{B}_0 + \boldsymbol{\nabla}_\perp S_{\frac{5}{2}} \,, \tag{2.121b}$$

$$\boldsymbol{\gamma}^{\boldsymbol{X}}_2 := \boldsymbol{\rho}_3 \times \boldsymbol{B}_0 + \boldsymbol{\nabla}_\perp S_3 + \boldsymbol{F}_{2\perp} \,, \tag{2.121c}$$

$$\boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{5}{2}} := G_2 \| \boldsymbol{b}_0 + \nabla_\| S_{\frac{5}{2}}\, \boldsymbol{b}_0 + \boldsymbol{F}_{\frac{5}{2}\perp} + \delta \boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{5}{2}} \,, \tag{2.121d}$$

$$\boldsymbol{\gamma}^{\boldsymbol{X}}_3 := \nabla_\| S_3\, \boldsymbol{b}_0 + \boldsymbol{F}_{3\perp} + F_{3\|}\, \boldsymbol{b}_0 + \delta \boldsymbol{\gamma}^{\boldsymbol{X}}_3 \,, \tag{2.121e}$$

where the terms $\boldsymbol{F}_{\frac{n}{2}\perp}$, for $n = 4, 5, 6$, and $F_{3\|}$ are defined as

$$\boldsymbol{F}_{2\perp} := \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_2 - G^{\mu}_{\frac{1}{2}} \boldsymbol{\nabla}_\perp G^{\Theta}_{\frac{1}{2}} \,, \tag{2.122a}$$

$$\begin{aligned} \boldsymbol{F}_{\frac{5}{2}\perp} := {} & G^{\|}_1 \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_2 \cdot \boldsymbol{b}_0 - G^{\mu}_{\frac{1}{2}} \boldsymbol{\nabla}_\perp G^{\Theta}_1 - G^{\mu}_1 \boldsymbol{\nabla}_\perp G^{\Theta}_{\frac{1}{2}} \\ & - (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla}) \boldsymbol{A}_0 \cdot \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_{\frac{5}{2}} - (\boldsymbol{\rho}_{\frac{5}{2}} \cdot \boldsymbol{\nabla}) \boldsymbol{A}_0 \cdot \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_2 \,, \end{aligned} \tag{2.122b}$$

$$\begin{aligned} \boldsymbol{F}_{3\perp} := {} & G^{\|}_1 \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_{\frac{5}{2}} \cdot \boldsymbol{b}_0 + \frac{1}{2}(\boldsymbol{\rho}_{\frac{5}{2}} \times \boldsymbol{B}_0) \cdot \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_{\frac{5}{2}} - G^{\mu}_1 \boldsymbol{\nabla}_\perp G^{\Theta}_1 \\ & - (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla}) \boldsymbol{A}_0 \cdot \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_3 - (\boldsymbol{\rho}_3 \cdot \boldsymbol{\nabla}) \boldsymbol{A}_0 \cdot \boldsymbol{\nabla}_\perp \boldsymbol{\rho}_2 \,, \end{aligned} \tag{2.122c}$$

$$F_{3\|} := \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \nabla_\| \boldsymbol{\rho}_2 - G^{\mu}_{\frac{1}{2}} \nabla_\| G^{\Theta}_{\frac{1}{2}} \,, \tag{2.122d}$$

and the terms $\delta \boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{n}{2}}$, for $n = 5, 6$, contain terms related to the curvature of the background magnetic field and are given by

$$\delta\boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{5}{2}} := -P_\parallel\,\boldsymbol{\rho}_2 \times (\boldsymbol{\nabla} \times \boldsymbol{b}_0)\,, \tag{2.123a}$$

$$\delta\boldsymbol{\gamma}^{\boldsymbol{X}}_{3} := -P_\parallel\,\boldsymbol{\rho}_{\frac{5}{2}} \times (\boldsymbol{\nabla} \times \boldsymbol{b}_0) - \frac{1}{2}(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla}\boldsymbol{B}_0) \times \boldsymbol{\rho}_2\,. \tag{2.123b}$$

The components $\gamma^{\parallel}_{\frac{n}{2}}$, for $n = 2, \ldots, 6$, are given by

$$\gamma^{\parallel}_{1} := 0\,, \tag{2.124a}$$

$$\gamma^{\parallel}_{\frac{3}{2}} := 0\,, \tag{2.124b}$$

$$\gamma^{\parallel}_{2} := 0\,, \tag{2.124c}$$

$$\gamma^{\parallel}_{\frac{5}{2}} := -\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_2 + \frac{\partial S_{\frac{5}{2}}}{\partial P_\parallel}\,, \tag{2.124d}$$

$$\gamma^{\parallel}_{3} := -\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_{\frac{5}{2}} + \frac{\partial S_3}{\partial P_\parallel} + \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial P_\parallel} - G^{\mu}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial P_\parallel}\,. \tag{2.124e}$$

The components $\gamma^{\mu}_{\frac{n}{2}}$, for $n = 2, \ldots, 6$, are given by

$$\gamma^{\mu}_{1} := 0\,, \tag{2.125a}$$

$$\gamma^{\mu}_{\frac{3}{2}} := 0\,, \tag{2.125b}$$

$$\gamma^{\mu}_{2} := 0\,, \tag{2.125c}$$

$$\gamma^{\mu}_{\frac{5}{2}} := G^{\Theta}_{\frac{1}{2}} + \frac{\partial S_{\frac{5}{2}}}{\partial \widehat{\mu}}\,, \tag{2.125d}$$

$$\gamma^{\mu}_{3} := G^{\Theta}_{1} + \frac{\partial S_3}{\partial \widehat{\mu}} + \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \widehat{\mu}} - G^{\mu}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial \widehat{\mu}}\,. \tag{2.125e}$$

The components $\gamma^{\Theta}_{\frac{n}{2}}$, for $n = 2, \ldots, 6$, are given by

$$\gamma^{\Theta}_{1} := 0\,, \tag{2.126a}$$

$$\gamma^{\Theta}_{\frac{3}{2}} := 0\,, \tag{2.126b}$$

$$\gamma^{\Theta}_{2} := -\widehat{\mu}\,, \tag{2.126c}$$

$$\gamma^{\Theta}_{\frac{5}{2}} := -G^{\mu}_{\frac{1}{2}} + \frac{\partial S_{\frac{5}{2}}}{\partial \Theta}\,, \tag{2.126d}$$

$$\gamma^{\Theta}_{3} := -G^{\mu}_{1} + \frac{\partial S_{3}}{\partial \Theta} + \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \Theta} - G^{\mu}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial \Theta}\,. \tag{2.126e}$$

Finally, the Hamiltonian H_1 is given by

$$\begin{aligned} H_1 := {} & G^{\mu}_{1} B_0 + P_{\parallel} G^{\parallel}_{1} - \Phi_1 + \frac{1}{2}|\boldsymbol{A}_1|^2 \\ & + \sqrt{\frac{B_0}{2\widehat{\mu}}} G^{\mu}_{\frac{1}{2}} \boldsymbol{c}_0 \cdot \boldsymbol{A}_{1\perp} - \sqrt{2\widehat{\mu} B_0}\, G^{\Theta}_{\frac{1}{2}}\, \boldsymbol{a}_0 \cdot \boldsymbol{A}_{1\perp} \\ & - \sqrt{\frac{2\widehat{\mu}}{B_0}} \boldsymbol{a}_0 \cdot \boldsymbol{\nabla} \boldsymbol{A}_1 \cdot \Big(P_{\parallel} \boldsymbol{b}_0 + \sqrt{2\widehat{\mu} B_0} \boldsymbol{c}_0 \Big) + O(\sqrt{\varepsilon})\,, \end{aligned} \tag{2.127}$$

where $\boldsymbol{A}_{1\perp} := \boldsymbol{b}_0 \times \boldsymbol{A}_1 \times \boldsymbol{b}_0$, as before, and the fluctuating potentials Φ_1 and $\boldsymbol{A}_1$ are evaluated at the position $\boldsymbol{X}/\varepsilon$. We remark that the higher-order Hamiltonians H_2 and H_3 are not relevant for our order of accuracy.

Proof 1. The result is obtained by substituting the gyrocenter coordinate transformation (2.65) into the guiding-center single-particle Lagrangian (2.59b) and computing its Taylor expansion in powers of $\sqrt{\varepsilon}$ up to order ε^3 (starting from $1/\varepsilon$). We denote again by Γ the symplectic part of the guiding-center Lagrangian (2.59b),

$$\Gamma := \left(\sqrt{\varepsilon}\, \bar{p}_{\parallel} \boldsymbol{b}_0 - \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\overline{\boldsymbol{X}}} - \varepsilon^2\, \bar{\mu}\, \dot{\bar{\theta}}\,. \tag{2.128}$$

The coefficients $\Gamma_{\frac{n}{2}}$, for $n = -2, \dots, 6$, of the Taylor expansion of Γ read

$$\Gamma_{-1} := -A_0 \cdot \dot{\boldsymbol{X}} \,, \tag{2.129a}$$

$$\Gamma_{-\frac{1}{2}} := 0 \,, \tag{2.129b}$$

$$\Gamma_0 := 0 \,, \tag{2.129c}$$

$$\Gamma_{\frac{1}{2}} := P_\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} \,, \tag{2.129d}$$

$$\Gamma_1 := -(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} - \boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_2 \,, \tag{2.129e}$$

$$\Gamma_{\frac{3}{2}} := G_1^\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} - (\boldsymbol{\rho}_{\frac{5}{2}} \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} - \boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_{\frac{5}{2}} \,, \tag{2.129f}$$

$$\Gamma_2 := -(\boldsymbol{\rho}_3 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} - \boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_3 - \hat{\mu}\,\dot{\Theta} \,, \tag{2.129g}$$

$$\Gamma_{\frac{5}{2}} := \Big(G_2^\| + P_\|(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\Big)\, \boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} + P_\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{\rho}}_2 - G_{\frac{1}{2}}^{\mu}\dot{\Theta} - \hat{\mu}\,\dot{G}_{\frac{1}{2}}^{\Theta} \,, \tag{2.129h}$$

$$\begin{aligned}\Gamma_3 := {} & P_\|(\boldsymbol{\rho}_{\frac{5}{2}} \cdot \boldsymbol{\nabla})\boldsymbol{b}_0 \cdot \dot{\boldsymbol{X}} - \frac{1}{2}(\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})^2 \boldsymbol{A}_0 \cdot \dot{\boldsymbol{X}} - (\boldsymbol{\rho}_2 \cdot \boldsymbol{\nabla})\boldsymbol{A}_0 \cdot \dot{\boldsymbol{\rho}}_2 \\ & + P_\| \boldsymbol{b}_0 \cdot \dot{\boldsymbol{\rho}}_{\frac{5}{2}} - G_1^{\mu}\dot{\Theta} - \hat{\mu}\,\dot{G}_1^{\Theta} - G_{\frac{1}{2}}^{\mu}\dot{G}_{\frac{1}{2}}^{\Theta} \,.\end{aligned} \tag{2.129i}$$

The results for the symplectic part follow by using the same equivalence relations used for ions and by adding the terms corresponding to the total differentials $\dot{S}_{\frac{5}{2}}$ and $\dot{S}_3$. For the Hamiltonian part, we note that, because of the guiding-center generator

$$\bar{\boldsymbol{\rho}}_{1\mathrm{e}} = -\sqrt{\varepsilon}\sqrt{\frac{2\bar{\mu}}{B_0}}\,\boldsymbol{a}_0 \,, \tag{2.130}$$

the fluctuating potentials Φ_1 and $\boldsymbol{A}_1$ can be expanded in a Taylor series around $\boldsymbol{X}/\varepsilon$, yielding

$$\begin{aligned}\Phi_1\left(t, \frac{\boldsymbol{x}}{\varepsilon}\right) &= \Phi_1\left(t, \frac{\overline{\boldsymbol{X}}}{\varepsilon} - \sqrt{\varepsilon}\sqrt{\frac{2\bar{\mu}}{B_0}}\,\boldsymbol{a}_0(\overline{\boldsymbol{X}}, \bar{\theta}) + O(\varepsilon)\right) \\ &= \Phi_1\left(t, \frac{\boldsymbol{X}}{\varepsilon}\right) - \sqrt{\varepsilon}\sqrt{\frac{2\hat{\mu}}{B_0}}\,\boldsymbol{a}_0(\boldsymbol{X}, \Theta) \cdot \boldsymbol{\nabla}\Phi_1\left(t, \frac{\boldsymbol{X}}{\varepsilon}\right) + O(\varepsilon) \,,\end{aligned} \tag{2.131}$$

and the same for $\boldsymbol{A}_1$. This explains the particular form of the Hamiltonian (2.127).

Proposition 2. In the electron single-particle Lagrangian (2.118) the generators of the polynomial transform can be chosen such that

$$L_{\mathrm{e}} \sim \left(\sqrt{\varepsilon} P_{\|} \boldsymbol{b}_0 - \frac{\boldsymbol{A}_0}{\varepsilon}\right) \cdot \dot{\boldsymbol{X}} + \varepsilon^2 \left(\left\langle \gamma_2^{\Theta} \right\rangle + \varepsilon \left\langle \gamma_3^{\Theta} \right\rangle\right) \dot{\Theta} - H_0 + O(\varepsilon^{\frac{7}{2}}) \,, \tag{2.132}$$

where, for a given function $g(\Theta)$, $\langle g \rangle$ denotes its gyro-average as defined in (2.85).

Proof 2. The components $\boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{n}{2}}$, for $n = 2, \ldots, 6$, in (2.121) vanish if and only if we set

$$G_1^{\|} = 0 \,, \tag{2.133a}$$

$$G_2^{\|} = -\nabla_{\|} S_{\frac{5}{2}} - \delta\gamma^{\boldsymbol{X}}_{\frac{5}{2}\|} - \sqrt{\varepsilon}\left(\nabla_{\|} S_3 + F_{3\|} + \delta\gamma^{\boldsymbol{X}}_{3\|}\right) \,, \tag{2.133b}$$

as well as

$$\boldsymbol{\rho}_{2\perp} = 0 \,, \tag{2.134a}$$

$$\boldsymbol{\rho}_{\frac{5}{2}\perp} = -\frac{\boldsymbol{b}_0}{B_0} \times \boldsymbol{\nabla}_{\perp} S_{\frac{5}{2}} \,, \tag{2.134b}$$

$$\begin{aligned} \boldsymbol{\rho}_{3\perp} = -\frac{\boldsymbol{b}_0}{B_0} \times \Bigg[& \boldsymbol{\nabla}_{\perp} S_3 + \boldsymbol{F}_{2\perp} \\ & + \sqrt{\varepsilon}\left(\boldsymbol{F}_{\frac{5}{2}\perp} + \delta\boldsymbol{\gamma}^{\boldsymbol{X}}_{\frac{5}{2}\perp}\right) + \varepsilon\left(\boldsymbol{F}_{3\perp} + \delta\boldsymbol{\gamma}^{\boldsymbol{X}}_{3\perp}\right)\Bigg] \,. \end{aligned} \tag{2.134c}$$

The components $\gamma^{\|}_{\frac{n}{2}}$, for $n = 5, 6$, in (2.124) vanish if and only if we set

$$\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_2 = \frac{\partial S_{\frac{5}{2}}}{\partial P_{\|}} \,, \tag{2.135a}$$

$$\boldsymbol{b}_0 \cdot \boldsymbol{\rho}_{\frac{5}{2}} = \frac{\partial S_3}{\partial P_{\|}} + \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial P_{\|}} - G^{\mathrm{\mu}}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial P_{\|}} \,. \tag{2.135b}$$

The components $\gamma^{\mu}_{\frac{n}{2}}$, for $n = 5, 6$, in (2.125) vanish if and only if we set

$$G^{\Theta}_{\frac{1}{2}} = -\frac{\partial S_{\frac{5}{2}}}{\partial \widehat{\mu}} \,, \tag{2.136a}$$

$$G^{\Theta}_{1} = -\frac{\partial S_3}{\partial \widehat{\mu}} - \frac{1}{2}(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \widehat{\mu}} + G^{\mu}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial \widehat{\mu}} \,. \tag{2.136b}$$

The Hamiltonians $H_{\frac{n}{2}}$, for $n = 1, 2$, in (2.119) and (2.127) vanish if and only if we set

$$G^{\mu}_{\frac{1}{2}} = -\frac{1}{B_0} \left(P_{\|} A_{1\|} + \sqrt{2\widehat{\mu} B_0}\, \boldsymbol{c}_0 \cdot \boldsymbol{A}_{1\perp} \right) \,, \tag{2.137a}$$

$$\begin{aligned} G^{\mu}_{1} = -\frac{1}{B_0} \Bigg\{ & P_{\|} G^{\|}_{1} - \Phi_1 + \frac{1}{2} |\boldsymbol{A}_1|^2 \\ & + \sqrt{\frac{B_0}{2\widehat{\mu}}} G^{\mu}_{\frac{1}{2}} \boldsymbol{c}_0 \cdot \boldsymbol{A}_{1\perp} - \sqrt{2\widehat{\mu} B_0} G^{\Theta}_{\frac{1}{2}}\, \boldsymbol{a}_0 \cdot \boldsymbol{A}_{1\perp} \\ & - \sqrt{\frac{2\widehat{\mu}}{B_0}} \boldsymbol{a}_0 \cdot \boldsymbol{\nabla} \boldsymbol{A}_1 \cdot \left(P_{\|} \boldsymbol{b}_0 + \sqrt{2\widehat{\mu} B_0} \boldsymbol{c}_0 \right) + O(\sqrt{\varepsilon}) \Bigg\} \,. \end{aligned} \tag{2.137b}$$

The only degrees of freedom left are the arbitrary scalar functions $S_{\frac{5}{2}}$ and S_3. As for ions, the dependence on the gyro-angle Θ can be removed from (2.126) by setting, for $n = 5, 6$,

$$\gamma^{\Theta}_{\frac{n}{2}} = \left\langle \gamma^{\Theta}_{\frac{n}{2}} \right\rangle \,, \tag{2.138}$$

or, equivalently, by requiring that $S_{\frac{n}{2}}$, for $n = 5, 6$, satisfy the differential equations

$$\frac{\partial S_{\frac{5}{2}}}{\partial \Theta} = \widetilde{G^{\mu}_{\frac{1}{2}}} \,, \tag{2.139a}$$

$$\frac{\partial S_3}{\partial \Theta} = \widetilde{G^{\mu}_{1}} - \frac{1}{2} \widetilde{(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \Theta}} + \widetilde{G^{\mu}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial \Theta}} \,. \tag{2.139b}$$

The solutions of (2.139) read (with arbitrary lower bound Θ_0 of integration)

$$S_{\frac{5}{2}}(\Theta) = S_{\frac{5}{2}}(\Theta_0) + \int_{\Theta_0}^{\Theta} \mathrm{d}\Theta' \, \widetilde{G^{\upmu}_{\frac{1}{2}}} \,, \tag{2.140a}$$

$$S_3(\Theta) = S_3(\Theta_0) + \int_{\Theta_0}^{\Theta} \mathrm{d}\Theta' \left(\widetilde{G^{\upmu}_1} - \frac{1}{2} \widetilde{(\boldsymbol{\rho}_2 \times \boldsymbol{B}_0) \cdot \frac{\partial \boldsymbol{\rho}_2}{\partial \Theta}} + \widetilde{G^{\upmu}_{\frac{1}{2}} \frac{\partial G^{\Theta}_{\frac{1}{2}}}{\partial \Theta}} \right) . \tag{2.140b}$$

Proposition 3. The generalized gyrocenter magnetic moment

$$\mu := \widehat{\upmu} + \sqrt{\varepsilon} \left\langle G^{\upmu}_{\frac{1}{2}} \right\rangle - \varepsilon \left\langle \gamma^{\Theta}_3 \right\rangle \tag{2.141}$$

is a constant of the motion, accurate up to order $O(\varepsilon)$, with respect to the dynamics induced by the preliminary gyrocenter single-particle Lagrangian (2.132). Moreover, there is a one-to-one correspondence $\mu \mapsto \widehat{\upmu}$, which implies that the gyrocenter Hamiltonian $\mathcal{H}$ is given by

$$\mathcal{H} := \frac{P_{\|}^2}{2} + \widehat{\upmu}(\mu) B_0 \,. \tag{2.142}$$

In other words, the gyrocenter Hamiltonian is obtained from H_0 by inverting the transformation $\widehat{\upmu} \mapsto \mu$ defined in (2.141). By expressing (2.132) in terms of the new gyrocenter coordinates $\boldsymbol{\mathcal{Z}} := (\boldsymbol{X}, P_{\|}, \mu, \Theta)$, we obtain the electron gyrocenter single-particle Lagrangian

$$\mathcal{L}_{\mathrm{e}} \sim \left(\sqrt{\varepsilon} P_{\|} \boldsymbol{b}_0 - \frac{\boldsymbol{A}_0}{\varepsilon} \right) \cdot \dot{\boldsymbol{X}} - \varepsilon^2 \mu \, \dot{\Theta} - \mathcal{H} + \bar{O}(\varepsilon^{\frac{3}{2}}) + O(\varepsilon^{\frac{7}{2}}) \,, \tag{2.143}$$

where the gyrocenter Hamiltonian is defined as $\mathcal{H} := \mathcal{H}_0 + \sqrt{\varepsilon}\, \mathcal{H}_{\frac{1}{2}} + \varepsilon\, \mathcal{H}_1$, with $\mathcal{H}_{\frac{n}{2}}$, for $n = 0, 1, 2$, given by

$$\mathcal{H}_0 := \frac{P_{\|}^2}{2} + \mu B_0 \,, \tag{2.144a}$$

$$\mathcal{H}_{\frac{1}{2}} := P_{\|} A_{1\|} \,, \tag{2.144b}$$

$$\mathcal{H}_1 := -\Phi_1 + \frac{1}{2} A_{1\|}^2 + \mu \, (\boldsymbol{\nabla} \times \boldsymbol{A}_1) \cdot \boldsymbol{b}_0 \,. \tag{2.144c}$$

As for ions, the symbol $\bar{O}(\varepsilon^{\frac{3}{2}})$ in (2.143) denotes terms of order $O(\varepsilon^{\frac{3}{2}})$ that are independent of the gyro-angle Θ. Moreover, the fluctuating potentials Φ_1 and $\boldsymbol{A}_1$ are evaluated at $\boldsymbol{X}/\varepsilon$.

Proof 3. Using that

$$\left\langle \gamma_2^{\Theta} \right\rangle = -\widehat{\upmu} - \sqrt{\varepsilon} \left\langle G_{\frac{1}{2}}^{\upmu} \right\rangle , \tag{2.145}$$

the Euler-Lagrange equation

$$\frac{\partial L_{\mathrm{e}}}{\partial \Theta} = \frac{\mathrm{d}}{\mathrm{d}t} \frac{\partial L_{\mathrm{e}}}{\partial \dot{\Theta}} , \tag{2.146}$$

for the preliminary gyrocenter single-particle Lagrangian (2.132) yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \left(\widehat{\upmu} + \sqrt{\varepsilon} \left\langle G_{\frac{1}{2}}^{\upmu} \right\rangle - \varepsilon \left\langle \gamma_3^{\Theta} \right\rangle \right) = O(\varepsilon^{\frac{3}{2}}) . \tag{2.147}$$

Hence, the gyrocenter magnetic moment μ defined in (2.141) is conserved with first-order accuracy in ε. Let us now compute the terms $\left\langle G_{\frac{1}{2}}^{\upmu} \right\rangle$ and $\left\langle \gamma_3^{\Theta} \right\rangle$ that define the transformation $\widehat{\upmu} \mapsto \mu$ in (2.141). From (2.137) we have

$$\left\langle G_{\frac{1}{2}}^{\upmu} \right\rangle = -\frac{1}{B_0} P_{\|} A_{1\|} , \tag{2.148a}$$

$$\begin{aligned} \left\langle G_1^{\upmu} \right\rangle = &-\frac{1}{B_0} \bigg(-\Phi_1 + \frac{1}{2} |\boldsymbol{A}_1|^2 - \left\langle (\boldsymbol{c}_0 \cdot \boldsymbol{A}_{1\perp})^2 \right\rangle \\ &- \left\langle (\boldsymbol{a}_0 \cdot \boldsymbol{A}_{1\perp})^2 \right\rangle - 2\widehat{\upmu} \left\langle \boldsymbol{a}_0 \cdot \boldsymbol{\nabla} \boldsymbol{A}_1 \cdot \boldsymbol{c}_0 \right\rangle \bigg) + \bar{O}(\sqrt{\varepsilon}) . \end{aligned} \tag{2.148b}$$

Moreover, from (2.136) we have

$$\left\langle G_{\frac{1}{2}}^{\upmu} \frac{\partial G_{\frac{1}{2}}^{\Theta}}{\partial \Theta} \right\rangle = -\frac{1}{B_0} \left\langle (\boldsymbol{c}_0 \cdot \boldsymbol{A}_{1\perp})^2 \right\rangle . \tag{2.149}$$

By computing the additional terms

$$\left\langle (\boldsymbol{a}_0 \cdot \boldsymbol{A}_{1\perp})^2 \right\rangle = \frac{1}{2} \left[(\boldsymbol{e}_1 \cdot \boldsymbol{A}_{1\perp})^2 + (\boldsymbol{e}_2 \cdot \boldsymbol{A}_{1\perp})^2 \right] = \frac{1}{2} |\boldsymbol{A}_{1\perp}|^2 , \tag{2.150}$$

and

$$\begin{aligned}\left\langle \boldsymbol{a}_0 \cdot \boldsymbol{\nabla}\boldsymbol{A}_1 \cdot \boldsymbol{c}_0 \right\rangle &= \frac{1}{2}\left(\boldsymbol{e}_2 \cdot \boldsymbol{\nabla}\boldsymbol{A}_1 \cdot \boldsymbol{e}_1 - \boldsymbol{e}_1 \cdot \boldsymbol{\nabla}\boldsymbol{A}_1 \cdot \boldsymbol{e}_2\right) \\ &= \frac{1}{2}\boldsymbol{e}_2 \cdot [\boldsymbol{e}_1 \times (\boldsymbol{\nabla} \times \boldsymbol{A}_1)] = -\frac{1}{2}(\boldsymbol{\nabla} \times \boldsymbol{A}_1) \cdot \boldsymbol{b}_0\,,\end{aligned} \tag{2.151}$$

we finally obtain

$$\left\langle \gamma_3^\Theta \right\rangle = \frac{1}{B_0}\left[-\Phi_1 + \frac{1}{2}A_{1\|}^2 + \widehat{\mu}\,(\boldsymbol{\nabla} \times \boldsymbol{A}_1) \cdot \boldsymbol{b}_0\right] + \bar{O}(\sqrt{\varepsilon})\,, \tag{2.152}$$

where we used the results $G_1^\| = 0$ and $\boldsymbol{\rho}_{2\perp} = 0$ from (2.133) and (2.134), respectively. It remains to identify the gyrocenter Hamiltonian. For this purpose, we need to invert (2.141) and substitute the result into H_0. Thanks to (2.148), the inversion is trivial and yields

$$\widehat{\mu} = \mu + \sqrt{\varepsilon}\frac{1}{B_0}P_\| A_{1\|} + \varepsilon\left\langle \gamma_3^\Theta \right\rangle(\mu) + \bar{O}(\varepsilon^{\frac{3}{2}})\,. \tag{2.153}$$

Substituting this into the Hamiltonian H_0 completes the proof.

2.3.4 Gyrokinetic Vlasov-Maxwell model

We first remark that the Jacobian determinants $\mathcal{J}_\mathrm{s}$ of the gyrocenter coordinate transformation $\boldsymbol{\mathcal{Z}} \mapsto (\boldsymbol{x}, \boldsymbol{v})$ can be computed directly from the symplectic part of the gyrocenter single-particle Lagrangians (2.97) and (2.143) for ions and electrons, respectively:

$$\mathcal{J}_\mathrm{i} = B_0 + \varepsilon\, P_\|\,(\boldsymbol{\nabla} \times \boldsymbol{b}_0) \cdot \boldsymbol{b}_0\,, \tag{2.154a}$$

$$\mathcal{J}_\mathrm{e} = B_0 - \varepsilon^{\frac{3}{2}} P_\|\,(\boldsymbol{\nabla} \times \boldsymbol{b}_0) \cdot \boldsymbol{b}_0\,. \tag{2.154b}$$

Such Jacobian determinants are exact because the symplectic forms in (2.97) and (2.143) remain the same at any order of the gyrocenter expansion, as only the gyrocenter Hamiltonians change with increased order of accuracy (as for the guiding-center coordinate transformation). The Jacobian determinants (2.154) confirm that geometric terms related to the curvature of the background magnetic field appear again at order $O(\varepsilon)$ for the ions and at order $O(\varepsilon^{\frac{3}{2}})$ for the electrons. The gyrokinetic equations of motion for ions derived from the ion gyrocenter single-particle Lagrangian (2.97) read

$$\dot{\boldsymbol{X}} = \frac{1}{B_\parallel^*}\left(\varepsilon\, \boldsymbol{b}_0 \times \boldsymbol{\nabla}\mathcal{H} + \frac{\partial \mathcal{H}}{\partial P_\parallel}\boldsymbol{B}^*\right) + O(\varepsilon^3)\,, \tag{2.155a}$$

$$\dot{P}_\parallel = -\frac{\boldsymbol{B}^*}{B_\parallel^*}\cdot \boldsymbol{\nabla}\mathcal{H} + O(\varepsilon^3)\,, \tag{2.155b}$$

$$\dot{\mu} = O(\varepsilon^3)\,, \tag{2.155c}$$

$$\dot{\Theta} = \frac{1}{\varepsilon}\frac{\partial \mathcal{H}}{\partial \mu} + O(\varepsilon^2)\,, \tag{2.155d}$$

where the modified magnetic field $\boldsymbol{B}^*$ is defined as

$$\boldsymbol{B}^* := \boldsymbol{B}_0 + \varepsilon\, P_\parallel\, \boldsymbol{\nabla} \times \boldsymbol{b}_0\,, \tag{2.156}$$

and its parallel component is defined as $B_\parallel^* := \boldsymbol{B}^* \cdot \boldsymbol{b}_0$, which is the ion Jacobian (2.154a). We note that, as expected, the gyro-angle Θ, whose dynamics is decoupled from the other phase-space variables, is the phase-space coordinate associated with the fastest scale of the physical particle dynamics. The gyrokinetic equations of motion for electrons derived from the electron gyrocenter single-particle Lagrangian (2.143) read

$$\dot{\boldsymbol{X}} = -\frac{1}{\sqrt{\varepsilon}}\frac{1}{B_\parallel^*}\left(\varepsilon^{\frac{3}{2}}\, \boldsymbol{b}_0 \times \boldsymbol{\nabla}\mathcal{H} - \frac{\partial \mathcal{H}}{\partial P_\parallel}\boldsymbol{B}^*\right) + O(\varepsilon)\,, \tag{2.157a}$$

$$\dot{P}_\parallel = -\frac{1}{\sqrt{\varepsilon}}\frac{\boldsymbol{B}^*}{B_\parallel^*}\cdot \boldsymbol{\nabla}\mathcal{H} + O(\varepsilon)\,, \tag{2.157b}$$

$$\dot{\mu} = O(\varepsilon^{\frac{3}{2}})\,, \tag{2.157c}$$

$$\dot{\Theta} = -\frac{1}{\varepsilon^2}\frac{\partial \mathcal{H}}{\partial \mu} + O(\varepsilon^{-\frac{1}{2}})\,, \tag{2.157d}$$

where the modified magnetic field $\boldsymbol{B}^*$ is defined as

$$\boldsymbol{B}^* := \boldsymbol{B}_0 - \varepsilon^{\frac{3}{2}}\, P_\| \,\boldsymbol{\nabla} \times \boldsymbol{b}_0\,, \tag{2.158}$$

and its parallel component is defined again as $B_\|^* := \boldsymbol{B}^* \cdot \boldsymbol{b}_0$, which is the electron Jacobian (2.154b). We note again that, as expected, the gyro-angle Θ, whose dynamics is decoupled from the other phase-space variables, is the phase-space coordinate associated with the fastest scale of the physical particle dynamics. The non-homogeneous gyrokinetic Maxwell's equations can be derived from the variational principle by taking variations of the Low Lagrangian with respect to the electromagnetic fluctuating potentials Φ_1 and $\boldsymbol{A}_1$. The weak form of gyrokinetic Coulomb's law is obtained by taking variations with respect to Φ_1 and reads

$$\begin{aligned}
0 = \varepsilon \int \mathrm{d}^6 \boldsymbol{\mathcal{Z}}_0 \Big(& B_{\|\mathrm{e}}^* F_\mathrm{e}(0, \boldsymbol{\mathcal{Z}}_0)\, \langle \delta\Phi_1 \rangle - B_{\|\mathrm{i}}^* F_\mathrm{i}(0, \boldsymbol{\mathcal{Z}}_0)\, \langle \delta\Phi_1 \rangle \Big) \\
& + \varepsilon^2 \int \mathrm{d}^6 \boldsymbol{\mathcal{Z}}_0\, B_{\|\mathrm{i}}^* F_\mathrm{i}(0, \boldsymbol{\mathcal{Z}}_0) \Bigg(\frac{1}{B_0} \frac{\mathrm{d}}{\mathrm{d}\mu} \left\langle \widetilde{\Psi_1}\, \widetilde{\delta\Phi_1} \right\rangle \\
& + \frac{1}{2B_0^2} \left\langle \Big(\boldsymbol{b}_0 \times \nabla_\perp \widetilde{\delta\Phi_1} \Big) \cdot \int^\Theta \mathrm{d}\Theta'\, \nabla_\perp \widetilde{\Psi_1} \right\rangle \\
& + \frac{1}{2B_0^2} \left\langle \Big(\boldsymbol{b}_0 \times \nabla_\perp \widetilde{\Psi_1} \Big) \cdot \int^\Theta \mathrm{d}\Theta'\, \nabla_\perp \widetilde{\delta\Phi_1} \right\rangle \\
& + \frac{1}{B_0} \left\langle \nabla_\| \delta\Phi_1 \int^\Theta \mathrm{d}\Theta'\, \widetilde{A_{1\|}} \right\rangle + \frac{\delta(\delta\mathcal{G}_2)}{\delta\Phi_1}(\delta\Phi_1) \Bigg)\,,
\end{aligned}$$

where $\delta\Phi_1$ denotes an arbitrary test function. Here we neglected all terms of order higher than $O(\varepsilon^2)$ and $O(\varepsilon)$ from the ion and electron Hamiltonians, respectively. The terms of order $O(\varepsilon)$ and $O(\varepsilon^2)$ represent the gyrocenter charge density and the gyrocenter polarization density, respectively. The weak form of Ampère-Maxwell's law is obtained by taking variations with respect to $\boldsymbol{A}_1$ and reads

$$\begin{aligned}
0 = &- \int \mathrm{d}^3\boldsymbol{x} \left(\nabla \times \boldsymbol{B}_0 + \varepsilon \, \nabla \times (\nabla \times \boldsymbol{A}_1) \right) \cdot \delta\boldsymbol{A}_1 \\
&- \sqrt{\varepsilon} \int \mathrm{d}^6\boldsymbol{\mathcal{Z}}_0 \, B^*_{\|\mathrm{e}} F_\mathrm{e}(0, \boldsymbol{\mathcal{Z}}_0) \, P_\| \, \delta A_{1\|} + \varepsilon \int \mathrm{d}^6\boldsymbol{\mathcal{Z}}_0 \Bigg[B^*_{\|\mathrm{i}} F_\mathrm{i}(0, \boldsymbol{\mathcal{Z}}_0) \, \langle \boldsymbol{P} \cdot \delta\boldsymbol{A}_1 \rangle \\
&- B^*_{\|\mathrm{e}} F_\mathrm{e}(0, \boldsymbol{\mathcal{Z}}_0) \Big(A_{1\|} \delta A_{1\|} + \mu \, (\boldsymbol{\nabla} \times \delta\boldsymbol{A}_1) \cdot \boldsymbol{b}_0 \Big) \Bigg] \\
&- \varepsilon^2 \int \mathrm{d}^6\boldsymbol{\mathcal{Z}}_0 \, B^*_{\|\mathrm{i}} F_\mathrm{i}(0, \boldsymbol{\mathcal{Z}}_0) \Bigg(\frac{1}{B_0} \frac{\mathrm{d}}{\mathrm{d}\mu} \left\langle \widetilde{\Psi_1} \, \boldsymbol{P} \cdot \widetilde{\delta\boldsymbol{A}_1} \right\rangle + \langle \boldsymbol{A}_1 \cdot \delta\boldsymbol{A}_1 \rangle \\
&+ \frac{1}{2B_0^2} \left\langle \Big(\boldsymbol{b}_0 \times \nabla_\perp (\boldsymbol{P} \cdot \widetilde{\delta\boldsymbol{A}_1}) \Big) \cdot \int^\Theta \mathrm{d}\Theta' \, \nabla_\perp \widetilde{\Psi_1} \right\rangle \\
&+ \frac{1}{2B_0^2} \left\langle \Big(\boldsymbol{b}_0 \times \nabla_\perp \widetilde{\Psi_1} \Big) \cdot \int^\Theta \mathrm{d}\Theta' \, \nabla_\perp (\boldsymbol{P} \cdot \widetilde{\delta\boldsymbol{A}_1}) \right\rangle \\
&+ \frac{1}{B_0} \left\langle \nabla_\| (\boldsymbol{P} \cdot \delta\boldsymbol{A}_1) \int^\Theta \mathrm{d}\Theta' \, \widetilde{A_{1\|}} \right\rangle + \frac{1}{B_0} \left\langle \nabla_\| \Psi_1 \int^\Theta \mathrm{d}\Theta' \, \widetilde{\delta A_{1\|}} \right\rangle \\
&+ \frac{\delta(\delta\mathcal{G}_2)}{\delta\boldsymbol{A}_1} (\delta\boldsymbol{A}_1) \Bigg) \,,
\end{aligned}$$

where $\delta\boldsymbol{A}_1$ denotes an arbitrary test function and $\boldsymbol{P} := P_\| \boldsymbol{b}_0 + \sqrt{2\mu B_0} \, \boldsymbol{c}_0$. As before, we neglected all terms of order higher than $O(\varepsilon^2)$ and $O(\varepsilon)$ from the ion and electron Hamiltonians, respectively.

2.3.5 Discussion of the results

The main results of this work are summarized in Proposition 3 for both ions and electrons, stating the final ion and electron single-particle gyrocenter Lagrangians (2.97) and (2.143), obtained by applying our gyrokinetic reduction to the normalized ion and electron single-particle Lagrangians (2.42a) and (2.42b), in the maximal ordering defined in section 2.2. The results obtained in (Tronko and Chandre, 2018) and in the previous works of (Hahm, 1988) and (Brizard, 1989) for ions are recovered and augmented by terms related to the assumption of maximal ordering. In particular, novel terms are the geometric first-order and second-order corrections δH_1, δH_2 and $\delta\mathcal{G}_2$, appearing in the first-order and second-order ion gyrocenter Hamiltonians $\mathcal{H}_1$ and $\mathcal{H}_2$ in (2.98), respectively. Moreover, the term

$$-\frac{1}{B_0}\left\langle \nabla_\parallel \widetilde{\Psi_1} \int^{\Theta} \mathrm{d}\Theta' \, \widetilde{A_{1\parallel}} \right\rangle , \tag{2.159}$$

appearing in the second-order ion gyrocenter Hamiltonian $\mathcal{H}_2$, is also new. On the other hand, electrons turn out to be insensitive to the effects related to the curvature of the background magnetic field up to second order in ε, even in maximal ordering. However, terms related to the parallel component of the magnetic vector potential appear in the electron gyrocenter Hamiltonian already at order $O(\sqrt{\varepsilon})$.

The gyrokinetic equations of motion (2.155) derived for ions are accurate up to order $O(\varepsilon^3)$, including the conservation of the gyrocenter magnetic moment μ (thanks to the fact that the error terms of order $O(\varepsilon^3)$ in the ion gyrocenter Lagrangian (2.97) do not depend on the gyro-angle Θ). Similarly, the gyrokinetic equations of motion (2.157) derived for electrons are accurate up to order $O(\varepsilon)$ and $O(\varepsilon^{\frac{3}{2}})$ for the conservation of the gyrocenter magnetic moment μ (thanks to the fact that the error terms of order $O(\varepsilon^{\frac{3}{2}})$ in the electron gyrocenter Lagrangian (2.143) do not depend on the gyro-angle Θ).

The gyrokinetic electron Lagrangian, and the corresponding gyrokinetic equations of motion, have been derived within an ordering consistent with the ions, despite the order of accuracy of the results being different for the two species (due to the fact that the gyrocenter magnetic moment μ has been computed with less precision for electrons than for ions). We conclude that it is possible to derive a set of gyrokinetic Vlasov-Maxwell equations for ions and electrons within unique ordering assumptions relevant for realistic fusion scenarios (maximal ordering) by considering phase-space coordinate transformations based on polynomial transforms. Our technique is alternative to the use of Lie transforms and, combined with our rigorous normalization procedure, can provide useful insights into the derivation of gyrokinetic models.

2.4 Summary of the contributions

In this chapter we described a new method for the derivation of gyrokinetic models. Our technique is based on a rigorous normalization and ordering of the physical equations of interest and makes use of polynomial transforms, instead of Lie transforms, in order to define the phase-space coordinate transformations involved in the gyrokinetic reduction. Our main contributions can be summarized as follows:

- Normalization and ordering of the Vlasov-Maxwell model according to physical scenarios relevant for existing and future fusion experiments, such as ASDEX Upgrade and ITER;
- Derivation of gyrokinetic equations for both ions and electrons within one common ordering;
- Use of polynomial transforms as an alternative to Lie transforms for the gyrokinetic reduction.

The material presented in this chapter is also described and discussed in a manuscript written in collaboration with Dr. Stefan Possanner, which is currently in preparation.

Chapter 3

Towards field-aligned semi-Lagrangian gyrokinetic simulations

In this chapter and the following two we present and discuss the results of the research that we conducted on the investigation and development of numerical methods for gyrokinetic simulations. These simulations are a well-established tool for the study of plasma turbulence in nuclear fusion devices and it is our hope that the numerical methods presented in this thesis, or their extension and improvement, will be eventually employed in state-of-the-art gyrokinetic codes. All the numerical investigations and computer experiments discussed in this thesis have been conducted within the framework of the software library SeLaLib. The primary aim of this software library is to provide mathematical and computational tools for kinetic, drift-kinetic and gyrokinetic simulations based on either semi-Lagrangian or particle-in-cell methods. In this thesis we focus on the development of numerical methods and algorithms for gyrokinetic simulations based on the semi-Lagrangian method. Therefore, our target state-of-the-art gyrokinetic code is the semi-Lagrangian gyrokinetic code GySeLa (Grandgirard et al., 2006a,b, 2016), developed at the Cadarache Center of the French Commission for Atomic Energy and Alternative Energies (CEA) in collaboration with several research partners.

In this chapter we introduce the basic features of a field-aligned semi-Lagrangian drift-kinetic electrostatic code that we developed within the SeLaLib library. Our code addresses the numerical solution of a 4D drift-kinetic electrostatic model that can be derived from the full gyrokinetic Vlasov-Maxwell equations under specific assumptions. A more detailed description of the model is given in chapter 5. However, the numerical features presented here are fairly general and do not depend on the simplifications behind the model addressed by our code. In particular, the features discussed in this chapter apply to any advection equation with characteristic equations similar to the ones derived for ions and electrons in chapter 2 (equations (2.155) and (2.157), respectively).

This chapter is organized as follows. Section 3.1 describes briefly the basic features of our target physical model. Section 3.2 defines the curvilinear coordinate systems that we will employ in the next chapters. Section 3.3 introduces the field-aligned approach, which constitutes a major feature of our numerical methods, and defines, in particular, the field-aligned splitting of the advection equation in our model. Section 3.4 describes our time advancing strategy and how to combine it with the field-aligned splitting. Section 3.5 reviews briefly the basics of the semi-Lagrangian method for the numerical solution of advection equations. Section 3.6 describes the discretization of the continuous problem into a numerical model with a finite number of degrees of freedom. Section 3.7 reviews briefly the basics of the interpolation schemes that we will employ, namely 1D Lagrange interpolation and 1D and 2D spline interpolation. Sections 3.8 and 3.9 describe two additional methods that characterize our field-aligned approach, namely field-aligned interpolation and field-aligned differentiation. Finally, we end the chapter with a summary of our contributions.

3.1 The physical model

We consider a gyrokinetic Vlasov equation of the form

$$\frac{\partial F}{\partial t} + \boldsymbol{U} \cdot \boldsymbol{\nabla} F + U_{P_\parallel} \frac{\partial F}{\partial P_\parallel} = 0\,. \tag{3.1}$$

The unknown in (3.1) is the particle distribution function $F(t, \boldsymbol{X}, P_\|)$, defined on the 4D gyrocenter phase space described by the gyrocenter position and parallel momentum $\boldsymbol{X}$ and $P_\|$. The time evolution of F is typically coupled to the time evolution of some electromagnetic potentials, whose dynamics is described by the gyrokinetic Maxwell's equations or some approximations of them. The advection fields $\boldsymbol{U}(t, \boldsymbol{X}, P_\|)$ and $U_{P_\|}(t, \boldsymbol{X}, P_\|)$ in (3.1) are defined as

$$\boldsymbol{U} := \frac{1}{B_\|^*}\left(\boldsymbol{b}_0 \times \boldsymbol{\nabla} H + \frac{\partial H}{\partial P_\|}\boldsymbol{B}^*\right) , \tag{3.2a}$$

$$U_{P_\|} := -\frac{\boldsymbol{B}^*}{B_\|^*} \cdot \boldsymbol{\nabla} H \,, \tag{3.2b}$$

where the modified magnetic field $\boldsymbol{B}^*(\boldsymbol{X}, P_\|)$ is defined as

$$\boldsymbol{B}^* := \boldsymbol{B}_0 + P_\| \boldsymbol{\nabla} \times \boldsymbol{b}_0 \,, \tag{3.3}$$

its parallel component $B_\|^*(\boldsymbol{X}, P_\|)$ is defined as $B_\|^* := \boldsymbol{B}^* \cdot \boldsymbol{b}_0$, and $H(t, \boldsymbol{X}, P_\|)$ denotes the Hamiltonian of the system. The particular form of H is not relevant for the methods presented in this chapter. The advection fields in (3.2) are as general as the ones derived for ions in chapter 2 (where H was denoted as $\mathcal{H}$), with the only difference that the small perturbation parameter ε has been absorbed into the physical quantities and does not appear explicitly in the equations.

3.2 Coordinate systems

The positions $\boldsymbol{X}$ in (3.1) are Cartesian coordinates in the 3D gyrocenter configuration space. However, depending on the specific geometry of the physical domain considered, which usually represents an approximation of the geometry of a realistic physical device, such as a Tokamak, it may be convenient to address the numerical solution of (3.1) choosing a different system of coordinates, typically a curvilinear coordinate system. The basics of vector algebra and analysis in curvilinear coordinates are reviewed in Appendix A. In this chapter we denote by (η^1, η^2, η^3) the logical coordinates in the configuration space, where the spatial degrees of freedom of our problem are represented. Moreover, we denote by η^4 the logical coordinate corresponding to the parallel momentum $P_\parallel$. Poloidal planes are spanned by the coordinates (η^1, η^2), where η^1 is a non-periodic radial-like coordinate and η^2 is a 2π-periodic angle-like coordinate. On the other hand, flux surfaces are spanned by the coordinates (η^2, η^3) at a given fixed value of η^1. The coordinate η^3 is a 2π-periodic angle coordinate, which we call toroidal angle. We assume that the magnetic field $\boldsymbol{B}_0$ is mostly aligned to the toroidal direction and always satisfies the condition $B_0^{\eta^3} \neq 0$, where $B_0^{\eta^3}$ is the third contravariant component of $\boldsymbol{B}_0$ with respect to the tangent basis $(\boldsymbol{e}_{\eta^1}, \boldsymbol{e}_{\eta^2}, \boldsymbol{e}_{\eta^3})$ of the logical coordinate system. We then introduce a second system of coordinates, which we denote by $(\zeta^1, \zeta^2, \zeta^3)$, and we assume that the magnetic field $\boldsymbol{B}_0$ is given in terms of its contravariant components $(B_0^{\zeta^1}, B_0^{\zeta^2}, B_0^{\zeta^3})$ with respect to the tangent basis $(\boldsymbol{e}_{\zeta^1}, \boldsymbol{e}_{\zeta^2}, \boldsymbol{e}_{\zeta^3})$ of this coordinate system. As for the logical coordinates, we assume that the magnetic field $\boldsymbol{B}_0$ is mostly aligned to the direction of the coordinate ζ^3 and always satisfies the condition $B_0^{\zeta^3} \neq 0$. In other words, the coordinates ζ^3 and η^3 are related by a linear homogeneous transformation of the form $\zeta^3 = c\,\eta^3$, for a given constant c, and poloidal planes are also entirely spanned by the coordinates (ζ^1, ζ^2).

For example, in a cylindrical geometry, such as the one shown in Figure 3.2 (top), we choose as logical coordinates the cylindrical coordinates (r, θ, φ) defined by the transformation

$$\begin{aligned} x(r,\theta,\varphi) &= r\cos\theta\,, \\ y(r,\theta,\varphi) &= r\sin\theta\,, \\ z(r,\theta,\varphi) &= R_0\,\varphi\,, \end{aligned} \tag{3.4}$$

where R_0 is related to the length L of the cylinder via $L = 2\pi R_0$. Moreover, we assume that the magnetic field $\boldsymbol{B}_0$ is given in terms of its contravariant components (B_0^x, B_0^y, B_0^z) with respect to the tangent basis $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ of the coordinate system defined by the Cartesian coordinates (x, y, z). In a toroidal geometry with circular poloidal cross sections, such as the one shown in Figure 3.2 (bottom), we choose as logical coordinates the elementary toroidal coordinates (r, θ, φ) defined by the

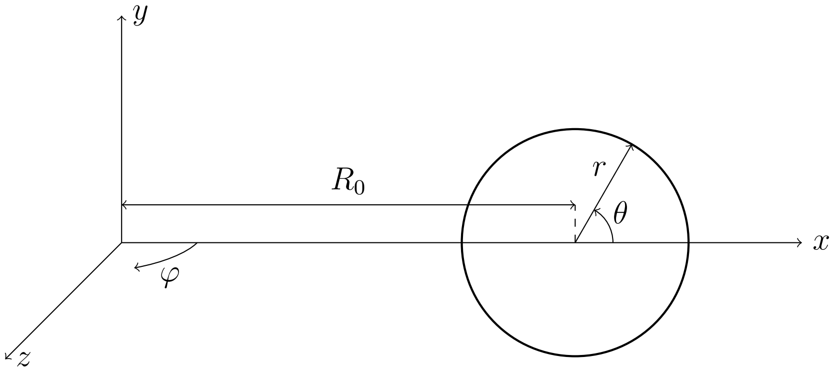


Figure 3.1: Elementary toroidal coordinates (r, θ, φ).

transformation

$$\begin{aligned} x(r,\theta,\varphi) &= (R_0 + r\cos\theta)\cos\varphi\,, \\ y(r,\theta,\varphi) &= r\sin\theta\,, \\ z(r,\theta,\varphi) &= (R_0 + r\cos\theta)\sin\varphi\,, \end{aligned} \tag{3.5}$$

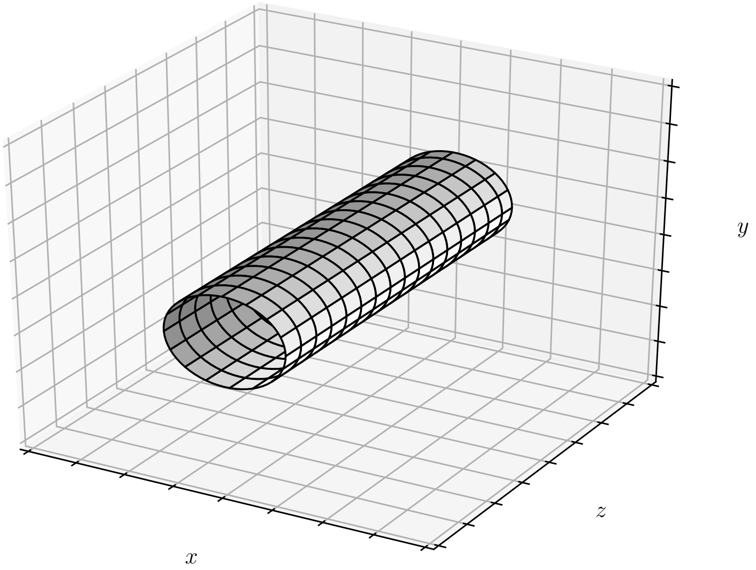
where R_0 denotes the major radius of the torus. Moreover, we assume that the magnetic field $\boldsymbol{B}_0$, irrespective of the shape of the poloidal cross sections, is given in terms of its contravariant components $(B_0^R, B_0^Z, B_0^\varphi)$ with respect to the tangent basis $(\boldsymbol{e}_R, \boldsymbol{e}_Z, \boldsymbol{e}_\varphi)$ of the coordinate system defined by the cylindrical coordinates (R, Z, φ). These are defined, in turn, by the transformation

$$\begin{aligned} x(R,Z,\varphi) &= R\cos\varphi\,, \\ y(R,Z,\varphi) &= Z\,, \\ z(R,Z,\varphi) &= R\sin\varphi\,. \end{aligned} \tag{3.6}$$

The elementary toroidal coordinates defined in (3.5) are shown in Figure 3.1.

When dealing with curvilinear coordinates, a mathematical object that contains useful geometric information is the Jacobian matrix of the coordinate system, defined in (A.9). In cylindrical geometry the Jacobian matrix $J_\eta(r,\theta,\varphi)$ of the logical coordinate system reads

$$J_\eta(r,\theta,\varphi) = \begin{bmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & R_0 \end{bmatrix}, \tag{3.7}$$



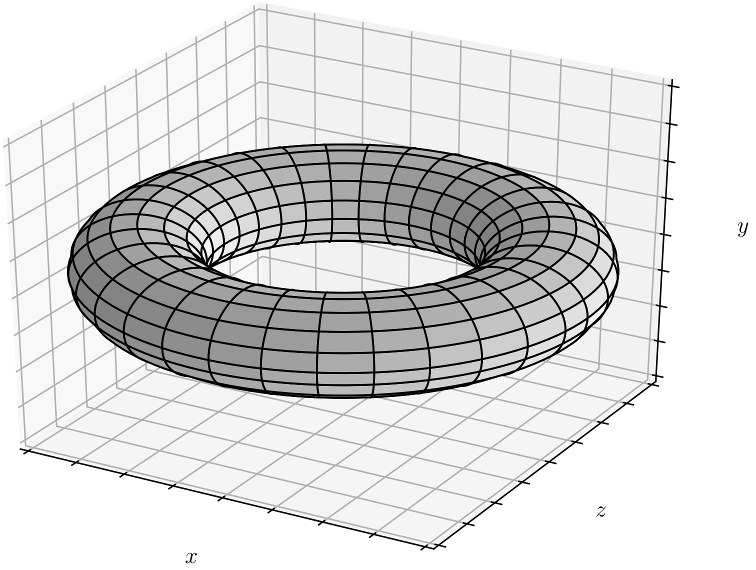


Figure 3.2: Cylindrical and circular toroidal grids.

with determinant $\det J_\eta(r,\theta,\varphi) = R_0\, r$, while the Jacobian matrix $J_\zeta(x,y,z)$ of the coordinate system defined by the coordinates (x,y,z) reads simply

$$J_\zeta(x,y,z) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{3.8}$$

with determinant $\det J_\zeta(x,y,z) = 1$. In toroidal geometry with circular poloidal cross sections the Jacobian matrix $J_\eta(r,\theta,\varphi)$ of the logical coordinate system reads

$$J_\eta(r,\theta,\varphi) = \begin{bmatrix} \cos\theta\cos\varphi & -r\sin\theta\cos\varphi & -(R_0 + r\cos\theta)\sin\varphi \\ \sin\theta & r\cos\theta & 0 \\ \cos\theta\sin\varphi & -r\sin\theta\sin\varphi & (R_0 + r\cos\theta)\cos\varphi \end{bmatrix}, \tag{3.9}$$

with determinant $\det J_\eta(r,\theta,\varphi) = (R_0 + r\cos\theta)\, r$, while the Jacobian matrix $J_\zeta(R,Z,\varphi)$ of the coordinate system defined by the coordinates (R,Z,φ) reads

$$J_\zeta(R,Z,\varphi) = \begin{bmatrix} \cos\varphi & 0 & -R\sin\varphi \\ 0 & 1 & 0 \\ \sin\varphi & 0 & R\cos\varphi \end{bmatrix}, \tag{3.10}$$

with determinant $\det J_\zeta(R,Z,\varphi) = R$.

3.3 Field-aligned splitting

One of the major features of the numerical methods developed in the context of our semi-Lagrangian drift-kinetic code is what we call the field-aligned approach. The idea behind this approach, first investigated in (Ottaviani, 2011; Hariri and Ottaviani, 2013) and recently employed in (Latu et al., 2018) in a context similar to the one discussed in this thesis, is to exploit at the numerical level the anisotropy of the turbulent plasma structures stretched along the direction of the magnetic field. This is achieved by introducing a field-aligned splitting of the 4D Vlasov equation (3.1) and by replacing interpolation on the flux surfaces and differentiation along the toroidal direction, to which the magnetic field $\boldsymbol{B}_0$ is mostly aligned, with field-aligned interpolation and field-aligned differentiation. The idea is to choose any arbitrarily fine discretization of the poloidal planes and interpolate our solution between adjacent poloidal planes following the magnetic field lines. We now present the field-aligned splitting and leave the discussion on field-aligned interpolation and field-aligned differentiation to sections 3.8 and 3.9, respectively.

The idea of the field-aligned splitting is to decompose the advection field $\boldsymbol{U}$ in (3.1) into a 2D component on the poloidal plane and a 1D component aligned with the direction of the magnetic field $\boldsymbol{B}_0$. This is done by expressing the third vector of the tangent basis $(\boldsymbol{e}_{\zeta^1}, \boldsymbol{e}_{\zeta^2}, \boldsymbol{e}_{\zeta^3})$ in terms of the unit vector $\boldsymbol{b}_0 := \boldsymbol{B}_0/B_0$, with contravariant components $(b_0^{\zeta^1}, b_0^{\zeta^2}, b_0^{\zeta^3})$, thanks to the assumption $b_0^{\zeta^3} \neq 0$:

$$\boldsymbol{e}_{\zeta^3} = -\frac{b_0^{\zeta^1}}{b_0^{\zeta^3}}\boldsymbol{e}_{\zeta^1} - \frac{b_0^{\zeta^2}}{b_0^{\zeta^3}}\boldsymbol{e}_{\zeta^2} + \frac{\boldsymbol{b}_0}{b_0^{\zeta^3}} \,. \tag{3.11}$$

Since the advection field $\boldsymbol{U}$ is typically written in terms of the magnetic field $\boldsymbol{B}_0$, its curl, and its cross product with the gradient of the Hamiltonian $\boldsymbol{\nabla} H$, it is natural to assume that $\boldsymbol{U}$, as $\boldsymbol{B}_0$, is also given in terms of its contravariant components $(U^{\zeta^1}, U^{\zeta^2}, U^{\zeta^3})$ with respect to the tangent basis $(\boldsymbol{e}_{\zeta^1}, \boldsymbol{e}_{\zeta^2}, \boldsymbol{e}_{\zeta^3})$, provided that the components of $\boldsymbol{\nabla} H$ with respect to the same basis are also available. We then split $\boldsymbol{U}$ as

$$\boldsymbol{U} = \boldsymbol{U}_{\mathrm{p}} + U_{\mathrm{b}}\boldsymbol{b}_0 \,, \tag{3.12}$$

where $\boldsymbol{U}_{\mathrm{p}}$, which lies on the poloidal planes spanned by (ζ^1, ζ^2), and U_{b} are defined as

$$\boldsymbol{U}_{\mathrm{p}} := \left(U^{\zeta^1} - \frac{b_0^{\zeta^1}}{b_0^{\zeta^3}}U^{\zeta^3}\right)\boldsymbol{e}_{\zeta^1} + \left(U^{\zeta^2} - \frac{b_0^{\zeta^2}}{b_0^{\zeta^3}}U^{\zeta^3}\right)\boldsymbol{e}_{\zeta^2} \,, \tag{3.13a}$$

$$U_{\mathrm{b}} := \frac{U^{\zeta^3}}{b_0^{\zeta^3}} \,. \tag{3.13b}$$

Therefore, we split the 4D Vlasov equation (3.1) into the following three separate lower-dimensional advection problems:

1. a 2D advection on poloidal planes:

$$\frac{\partial F}{\partial t} + \boldsymbol{U}_{\mathrm{p}} \cdot \boldsymbol{\nabla} F = 0\,; \tag{3.14}$$

2. a 2D field-aligned advection on flux surfaces:

$$\frac{\partial F}{\partial t} + U_{\mathrm{b}} \boldsymbol{b}_0 \cdot \boldsymbol{\nabla} F = 0\,; \tag{3.15}$$

3. a 1D advection in momentum space:

$$\frac{\partial F}{\partial t} + U_{P_\parallel} \frac{\partial F}{\partial P_\parallel} = 0\,. \tag{3.16}$$

From (3.11) we can also express derivatives with respect to the toroidal coordinate ζ^3 as derivatives with respect to the coordinates (ζ^1, ζ^2) and parallel derivatives $\nabla_\parallel := \boldsymbol{b}_0 \cdot \boldsymbol{\nabla}$:

$$\frac{\partial}{\partial \zeta^3} = -\frac{b_0^{\zeta^1}}{b_0^{\zeta^3}} \frac{\partial}{\partial \zeta^1} - \frac{b_0^{\zeta^2}}{b_0^{\zeta^3}} \frac{\partial}{\partial \zeta^2} + \frac{1}{b_0^{\zeta^3}} \nabla_\parallel\,. \tag{3.17}$$

How the parallel derivatives $\nabla_\parallel$, which we also call field-aligned derivatives, are computed numerically is discussed in section 3.9. The explicit expressions of the advection fields $\boldsymbol{U}_{\mathrm{p}}$, U_{b} and $U_{P_\parallel}$ in the coordinate system $(\zeta^1, \zeta^2, \zeta^3)$ can be computed from (3.2). The contravariant components U^{ζ^i} of $\boldsymbol{U}$, for $i = 1, 2, 3$, read

$$U^{\zeta^i} = \frac{1}{B_\parallel^*} \left(\frac{\varepsilon^{ijk}}{\det J_\zeta} (b_0)_{\zeta^j} \frac{\partial H}{\partial \zeta^k} + \frac{\partial H}{\partial P_\parallel} (B^*)^{\zeta^i} \right)\,, \tag{3.18}$$

where ε^{ijk} denotes the Levi-Civita symbol, defined in (A.13), the contravariant components $(B^*)^{\zeta^i}$ read

$$(B^*)^{\zeta^i} = B_0^{\zeta^i} + P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^i}\,, \tag{3.19}$$

the parallel component $B_\parallel^*$ reads

$$B_\parallel^* = B_0 + P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^i} (b_0)_{\zeta^i}\,, \tag{3.20}$$

and we employ the Einstein summation convention. Here, $(b_0)_{\zeta^i}$, for $i = 1, 2, 3$, denote the covariant components of $\boldsymbol{b}_0$ with respect to the co-tangent basis $(\boldsymbol{e}^{\zeta^1}, \boldsymbol{e}^{\zeta^2}, \boldsymbol{e}^{\zeta^3})$. From (3.17) and (3.18) it follows that

$$\begin{aligned} U_{\mathrm{p}}^{\zeta^1} = \frac{1}{B_\parallel^*}\frac{1}{b_0^{\zeta^3}}\Bigg[& \frac{1}{\det J_\zeta}\left(-\frac{\partial H}{\partial \zeta^2} + (b_0)_{\zeta^2}\nabla_\parallel H\right) \\ & + \frac{\partial H}{\partial P_\parallel} P_\parallel \left(b_0^{\zeta^3} (\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^1} - b_0^{\zeta^1}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\Bigg], \end{aligned} \tag{3.21a}$$

$$\begin{aligned} U_{\mathrm{p}}^{\zeta^2} = \frac{1}{B_\parallel^*}\frac{1}{b_0^{\zeta^3}}\Bigg[& \frac{-1}{\det J_\zeta}\left(-\frac{\partial H}{\partial \zeta^1} + (b_0)_{\zeta^1}\nabla_\parallel H\right) \\ & + \frac{\partial H}{\partial P_\parallel} P_\parallel \left(b_0^{\zeta^3} (\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^2} - b_0^{\zeta^2}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\Bigg]. \end{aligned} \tag{3.21b}$$

The advection field U_{b} for the field-aligned advection on flux surfaces is obtained directly dividing (3.18), with $i = 3$, by $b_0^{\zeta^3}$ and reads

$$\begin{aligned} U_{\mathrm{b}} = \frac{1}{B_\parallel^*}\frac{1}{b_0^{\zeta^3}}\Bigg[& \frac{1}{\det J_\zeta}\left((b_0)_{\zeta^1}\frac{\partial H}{\partial \zeta^2} - (b_0)_{\zeta^2}\frac{\partial H}{\partial \zeta^1}\right) \\ & + \frac{\partial H}{\partial P_\parallel}\left(B_0^{\zeta^3} + P_\parallel(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\Bigg]. \end{aligned} \tag{3.22}$$

Finally, the advection field $U_{P_\parallel}$ for the advection in momentum space reads

$$U_{P_\parallel} = -\frac{1}{B_\parallel^*}(B^*)^{\zeta^i}\frac{\partial H}{\partial \zeta^i}, \tag{3.23}$$

which, using (3.17), can be written as

$$\begin{aligned} U_{P_\parallel} = -\frac{1}{B_\parallel^*}\Bigg[& \left(B_0 + \frac{P_\parallel}{b_0^{\zeta^3}}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\nabla_\parallel H \\ & + \frac{P_\parallel}{b_0^{\zeta^3}}\left(b_0^{\zeta^3}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^1} - b_0^{\zeta^1}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\frac{\partial H}{\partial \zeta^1} \\ & + \frac{P_\parallel}{b_0^{\zeta^3}}\left(b_0^{\zeta^3}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^2} - b_0^{\zeta^2}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\frac{\partial H}{\partial \zeta^2}\Bigg]. \end{aligned} \tag{3.24}$$

3.4 Time-advancing strategy

The idea for advancing in time the Vlasov equation (3.1) is to choose, within each time step, an approximate time-independent Hamiltonian (and the corresponding approximate time-independent advection fields) and solve the resulting linear advection equation with the numerical method of choice. More precisely, we employ a second-order predictor-corrector scheme similar to the one described in (Latu et al., 2018, section 4.2), that we now review briefly. Let us denote by $F(t)$ and $H(t)$ the distribution function F and the Hamiltonian H at time t, omitting the dependence on the other phase-space variables. In the predictor step we advance $F(t)$ for half the time step with the Hamiltonian $H(t)$ (and the corresponding advection fields) computed from $F(t)$, and obtain $F(t+\Delta t/2)$. In the corrector step we advance $F(t)$ for the entire time step with the Hamiltonian $H(t+\Delta t/2)$ (and the corresponding advection fields) computed from $F(t+\Delta t/2)$, and obtain $F(t+\Delta t)$. This scheme is represented graphically in Figure 3.3. Our time advancing strategy is then combined

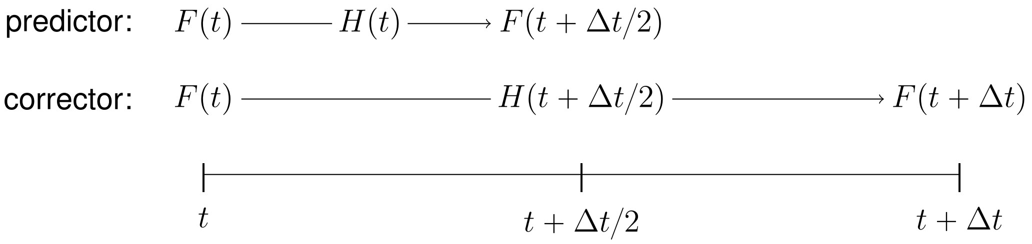


Figure 3.3: Graphical representation of the second-order predictor-corrector scheme employed to solve the Vlasov equation (3.1).

with the field-aligned splitting as follows. In the predictor step we start from $F(t)$, compute $H(t)$ from it, perform the steps

1. solve (3.15) in the time interval $[t, t+\Delta t/2]$,
2. solve (3.16) in the time interval $[t, t+\Delta t/2]$,
3. solve (3.14) in the time interval $[t, t+\Delta t/2]$,

and obtain $F(t+\Delta t/2)$. This corresponds to a first-order Lie splitting of the 4D Vlasov equation. In the corrector step we start again from $F(t)$, compute $H(t+\Delta t/2)$

from $F(t+\Delta t/2)$, perform the steps

1. solve (3.15) in the time interval $[t, t+\Delta t/2]$,
2. solve (3.16) in the time interval $[t, t+\Delta t/2]$,
3. solve (3.14) in the time interval $[t, t+\Delta t]$,
4. solve (3.16) in the time interval $[t+\Delta t/2, t+\Delta t]$,
5. solve (3.15) in the time interval $[t+\Delta t/2, t+\Delta t]$,

and obtain $F(t+\Delta t)$. This corresponds to a second-order Strang splitting of the 4D Vlasov equation.

3.5 Semi-Lagrangian method

In both the predictor and the corrector steps, each of the lower-dimensional advection equations, (3.14), (3.15) and (3.16), is solved with the semi-Lagrangian method, which we briefly review in this section for a generic 2D advection problem (the corresponding scheme for 1D advection equations is obtained by simply replacing vector quantities with scalar quantities). The idea of the semi-Lagrangian method is to combine the Eulerian and the Lagrangian approaches, by keeping the good properties of both schemes (namely, the fact that Eulerian methods work well on regular Cartesian meshes and the fact that Lagrangian methods allow for large time steps) and trying to get rid of their disadvantages (namely, the fact that Eulerian methods often require restrictive time steps in order to ensure computational stability and the fact that Lagrangian methods exhibit noise in the case of particle methods and strong grid deformation in the case of mesh-based methods). The semi-Lagrangian method was originally developed within the context of numerical weather prediction (Fjørtoft, 1952, 1955; Wiin-Nielsen, 1959; Krishnamurti, 1962; Sawyer, 1963; Leith, 1964; Purnell, 1976) (see, for example, (Staniforth and Côté, 1991) for a comprehensive review). The method was applied later on to Vlasov-like transport equations and drift-kinetic and gyrokinetic models in the context of plasma physics by (Cheng and Knorr, 1976; Gagné and Shoucri, 1977; Sonnendrücker et al., 1999; Filbet et al., 2001; Besse and Sonnendrücker, 2003; Crouseilles et al., 2010). We now review how the semi-Lagrangian scheme works for a 2D advection equation of the form

$$\frac{\partial g}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} g = 0\,, \tag{3.25}$$

for a given advection field $\boldsymbol{u}(t, \boldsymbol{x})$. The characteristics of (3.25) are the solutions of the dynamical system

$$\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}t} = \boldsymbol{u}\,, \tag{3.26}$$

with given initial conditions. The information contained in (3.25) is that its solution g is conserved along the characteristics obtained by solving (3.26). Typically, when we solve (3.25) numerically, we have a discrete set of mesh points $\{\boldsymbol{x}_{ij}\}_{ij}$ and we are interested in knowing the value of g at a given time t and a given mesh point $\boldsymbol{x}_{kl}$. The information we have is the set of values of g at the previous time $t - \Delta t$ at each mesh point $\boldsymbol{x}_{ij}$. We then first solve (3.26) backward in time, starting from the initial condition $\boldsymbol{x}_{kl}$ at time t, and find the origin $\boldsymbol{x}^*_{kl}$ of the characteristic at time $t - \Delta t$. Since g is conserved along the characteristics of (3.25), we have

$$g(t, \boldsymbol{x}_{kl}) = g(t - \Delta t, \boldsymbol{x}^*_{kl})\,. \tag{3.27}$$

Since we are integrating the characteristics backward in time, the algorithm described here is usually referred to as backward semi-Lagrangian method, as opposed to forward semi-Lagrangian methods, where the characteristics are integrated forward in time. Typically, the origin $\boldsymbol{x}^*_{kl}$ of the characteristic does not coincide with a mesh point and the value $g(t - \Delta t, \boldsymbol{x}^*_{kl})$ in (3.27) is not immediately available. Instead, it is obtained by interpolating the values of g at the same time $t - \Delta t$ and at mesh points in some neighborhood of $\boldsymbol{x}^*_{kl}$. The number and collocation of such mesh points with respect to $\boldsymbol{x}^*_{kl}$ depend on the choice of the interpolation scheme. The semi-Lagrangian method just described is represented graphically in Figure 3.4. The specific interpolation schemes and numerical methods employed for each of the lower-dimensional advection equations, (3.14), (3.15) and (3.16), will be discussed later on in this thesis.

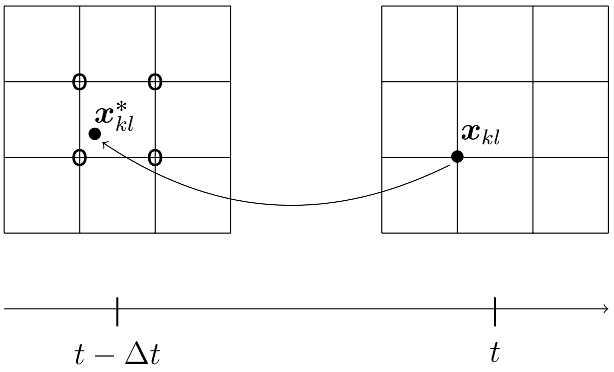


Figure 3.4: Semi-Lagrangian method: the value of g at a given time t and a given mesh point $\boldsymbol{x}_{kl}$ is obtained by first finding the origin $\boldsymbol{x}^*_{kl}$ of the characteristic at time $t-\Delta t$ and then interpolating the values of g at the same time $t-\Delta t$ at mesh points in some neighborhood of $\boldsymbol{x}^*_{kl}$ (empty bullets).

3.6 Discretization

In order to discuss the numerical methods and algorithms that we employ to solve our target model (3.1), we need to first present how we define the discrete degrees of freedom corresponding to the continuous degrees of freedom of the model. When we solve (3.1) numerically with the semi-Lagrangian method, we have a discrete set of mesh points corresponding to each of the logical phase-space coordinates, which we denote by $\{\eta^1_{i_1}\}_{i_1=1}^{n_1}$, $\{\eta^2_{i_2}\}_{i_2=1}^{n_2}$, $\{\eta^3_{i_3}\}_{i_3=1}^{n_3}$ and $\{\eta^4_{i_4}\}_{i_4=1}^{n_4}$, respectively. The logical mesh points corresponding to the poloidal coordinates (η^1,η^2) and the parallel momentum η^4 are defined as the interpolation points that we choose for the reconstruction of relevant quantities in each of these directions. On the other hand, the logical mesh points corresponding to the toroidal angle η^3 are simply defined from a uniform mesh in the domain of η^3.

On the poloidal planes and along the direction of the parallel momentum we interpolate using splines. More precisely, we employ 2D tensor-product splines on the poloidal planes spanned by the coordinates (η^1,η^2) and 1D splines in the direction of the parallel momentum η^4. In general, we work on a Schönberg space of maximum regularity. Given a decomposition of a 1D domain into cells, this is the space of piecewise polynomials of degree p with $p-1$ continuous derivatives across the domain. The B-splines that we are going to define constitute a basis for this space and have some desirable properties, most notably: positivity, compact support, and partition of unity. The two types of B-splines that we consider in order to define our spline bases are clamped (non-periodic) B-splines and periodic B-splines. Given a domain $[a, b]$, we decompose it into cells, whose limit points are named break points. Denoting by n the number of degrees of freedom and by p the degree of the B-splines, the number n^{c} of cells for clamped and periodic B-splines is defined as

$$n^{\mathrm{c}} := \begin{cases} n - p & \text{clamped}\,, \\ n & \text{periodic}\,, \end{cases} \tag{3.28}$$

with $n^{\mathrm{b}} := n^{\mathrm{c}} + 1$ break points, which we denote by $\{\bar{\eta}_i\}_{i=1}^{n^{\mathrm{b}}}$. The domain $[a, b]$ is thus decomposed as

$$[a, b] = \bigcup_{i=1}^{n^{\mathrm{b}}-1} [\bar{\eta}_i, \bar{\eta}_{i+1}]\,. \tag{3.29}$$

From the break points $\{\bar{\eta}_i\}_{i=1}^{n^{\mathrm{b}}}$ we define a knot sequence of $n^{\mathrm{k}} := n^{\mathrm{b}} + 2p$ knots $\{t_i\}_{i=1-p}^{n^{\mathrm{b}}+p}$. More precisely, for clamped B-splines we define the open knots

$$t_i := \begin{cases} \bar{\eta}_1 & i = 1 - p, \dots, 0\,, \\ \bar{\eta}_i & i = 1, \dots, n^{\mathrm{b}}\,, \\ \bar{\eta}_{n^{\mathrm{b}}} & i = n^{\mathrm{b}} + 1, \dots, n^{\mathrm{b}} + p\,, \end{cases} \tag{3.30}$$

whereas for periodic B-splines we define the periodic knots

$$t_i := \begin{cases} \bar{\eta}_{n^{\mathrm{b}}+i} - 2\pi & i = 1 - p, \dots, 0\,, \\ \bar{\eta}_i & i = 1, \dots, n^{\mathrm{b}}\,, \\ \bar{\eta}_{i-n^{\mathrm{b}}} + 2\pi & i = n^{\mathrm{b}} + 1, \dots, n^{\mathrm{b}} + p\,, \end{cases} \tag{3.31}$$

assuming that the B-splines are 2π-periodic. From the knot sequence $\{t_i\}_{i=1-p}^{n^{\mathrm{b}}+p}$ the B-splines B_i^p of degree p are defined by the recurrence relations (de Boor, 1980)

$$B_i^p(\eta) := \frac{\eta - t_i}{t_{i+p} - t_i} B_i^{p-1}(\eta) + \frac{t_{i+p+1} - \eta}{t_{i+p+1} - t_{i+1}} B_{i+1}^{p-1}(\eta)\,, \tag{3.32}$$

with $\eta \in [a,b]$ and the B-splines B_i^0 of degree $p=0$ defined as

$$B_i^0(\eta) := \begin{cases} 1 & \eta \in [t_i, t_{i+1})\,, \\ 0 & \text{elsewhere}\,. \end{cases} \tag{3.33}$$

We then define the interpolation points $\{\eta_i\}_{i=1}^n$ as the Greville averages

$$\eta_i := \frac{1}{p} \sum_{j=i+1-p}^{i} t_j\,. \tag{3.34}$$

Such points are averages of the knots generally lying near the values corresponding to the maximum of the basis functions (Gordon and Riesenfeld, 1974; Farin, 1988). In the case of a periodic domain the Greville averages reduce to either the break points themselves or the mid-points of each cell, depending on whether the degree of the B-splines is odd or even, respectively. In formulas, the interpolation points for periodic B-splines are given by

$$\eta_i = \begin{cases} \bar{\eta}_i & p \text{ odd}\,, \\ \dfrac{\bar{\eta}_i + \bar{\eta}_{i+1}}{2} & p \text{ even}\,. \end{cases} \tag{3.35}$$

On the poloidal planes we then define a 2D tensor-product spline basis of clamped B-splines of degree p_1 in $\eta^1 \in [0, \eta^1_{\max}]$ and 2π-periodic B-splines of degree p_2 in $\eta^2 \in [0, 2\pi)$:

$$\{B^{p_1 p_2}_{i_1 i_2}(\eta^1, \eta^2) := B^{p_1}_{i_1}(\eta^1) B^{p_2}_{i_2}(\eta^2)\}_{i_1, i_2 = 1}^{n_1, n_2}\,. \tag{3.36}$$

Along the direction of the parallel momentum we define a 1D spline basis of clamped B-splines of degree p_4 in $\eta^4 \in [\eta^4_{\min}, \eta^4_{\max}]$ (with typically $\eta^4_{\min} = -\eta^4_{\max}$):

$$\{B^{p_4}_{i_4}(\eta^4)\}_{i_4=1}^{n_4}\,. \tag{3.37}$$

Finally, along the 2π-periodic toroidal direction η^3 we define the logical mesh by decomposing the periodic domain $[0, 2\pi)$ into $n_3^{\mathrm{c}} := n_3$ cells with $n_3^{\mathrm{b}} := n_3^{\mathrm{c}} + 1$ break points $\{\bar{\eta}^3_{i_3}\}_{i_3=1}^{n_3^{\mathrm{b}}}$,

$$[0, 2\pi) = \bigcup_{i_3=1}^{n_3^{\mathrm{b}}-1} [\bar{\eta}^3_{i_3}, \bar{\eta}^3_{i_3+1})\,, \tag{3.38}$$

and choosing the logical mesh points $\{\eta^3_{i_3}\}_{i_3=1}^{n_3}$ as the break points: $\eta^3_{i_3} := \bar{\eta}^3_{i_3}$. Examples of logical meshes in the directions $\eta^1 \in [0,1]$, $\eta^2 \in [0,2\pi)$, $\eta^3 \in [0,2\pi)$ and $\eta^4 \in [-1,1]$, with $n_1 = n_2 = n_3 = n_4 = 8$ and $p_1 = p_2 = p_4 = 3$, are shown in Figure 3.5.

3.7 Interpolation

Interpolation is one of the main building blocks of the semi-Lagrangian method. In this section we review briefly the two types of standard interpolation that we employ in this thesis, namely 1D Lagrange interpolation and 1D and 2D spline interpolation.

For 1D Lagrange interpolation we employ the Neville's algorithm (Neville, 1934). Given n data points $\{(x_i, z_i)\}_{i=1}^{n}$, the value $P(x)$ of the interpolating Lagrange polynomial P of degree $n-1$ at the point x is given by $P(x) = P_{0,n-1}(x)$, where $P_{i,j}$ denotes the Lagrange polynomial of degree $j-i$ interpolating the data points $\{(x_k, z_k)\}_{k=i}^{j}$ and satisfying the recurrence relations

$$P_{i,i}(x) = z_i \qquad 0 \leqslant i \leqslant n-1\,, \tag{3.39a}$$

$$P_{i,j}(x) = \frac{(x-x_j)P_{i,j-1}(x) - (x-x_i)P_{i+1,j}(x)}{x_i - x_j} \qquad 0 \leqslant i < j \leqslant n-1\,. \tag{3.39b}$$

For 1D spline interpolation, let us introduce a generic 1D spline basis of B-splines of degree p:

$$\{B^p_i(x)\}_{i=1}^{n}\,. \tag{3.40}$$

Given n data points $\{(x_i, z_i)\}_{i=1}^{n}$, corresponding to the values of a given quantity of interest on the logical mesh defined by the interpolation points, an interpolatory spline

$$S(x) := \sum_{i=1}^{n} c_i B^p_i(x) \tag{3.41}$$

satisfies the interpolation conditions $S(x_i) = z_i$, for all $i = 1, \ldots, n$. These can be

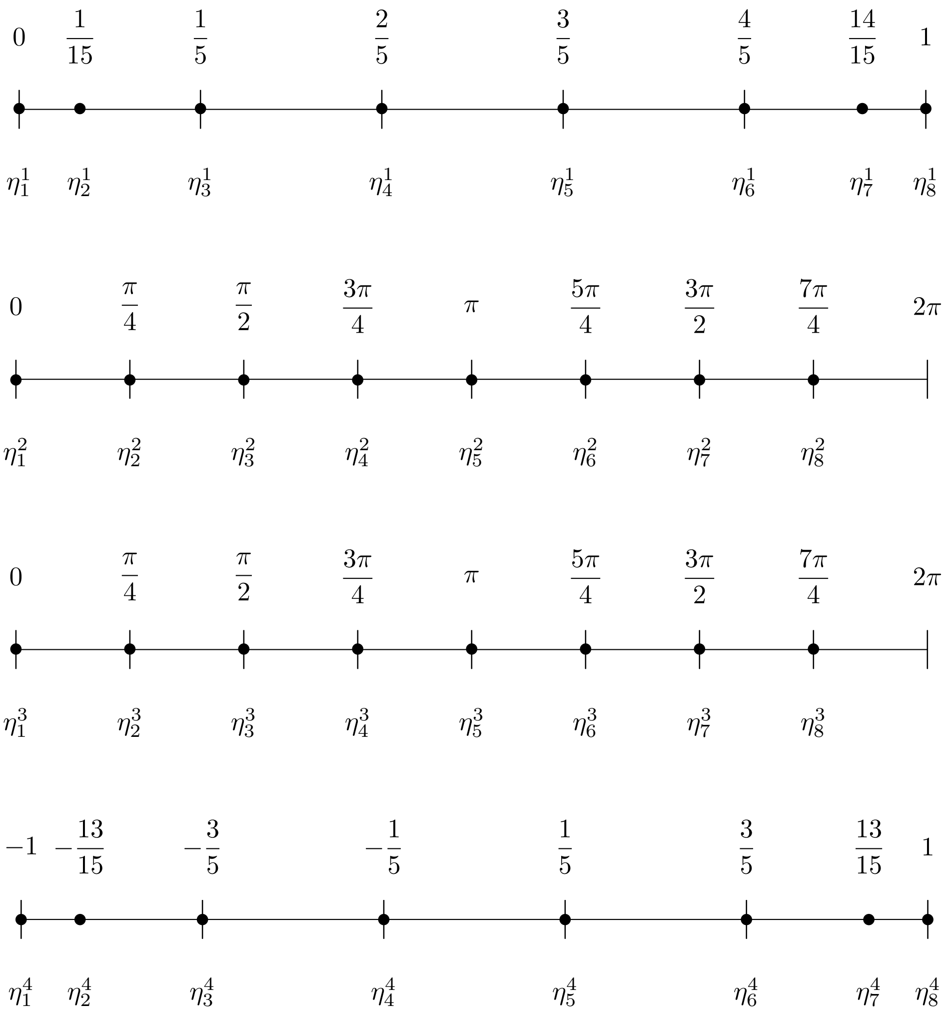


Figure 3.5: Examples of logical meshes in the directions $\eta^1 \in [0,1]$, $\eta^2 \in [0,2\pi)$, $\eta^3 \in [0,2\pi)$ and $\eta^4 \in [-1,1]$, with $n_1 = n_2 = n_3 = n_4 = 8$ and $p_1 = p_2 = p_4 = 3$. The vertical lines represent the break points ($n_1^{\mathrm{b}} = n_4^{\mathrm{b}} = 6$ break points in the directions η^1 and η^4, and $n_2^{\mathrm{b}} = n_3^{\mathrm{b}} = 9$ break points in the directions η^2 and η^3).

written in the matrix form

$$\begin{bmatrix} B_1^p(x_1) & \cdots & B_n^p(x_1) \\ \vdots & \ddots & \vdots \\ B_1^p(x_n) & \cdots & B_n^p(x_n) \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix}, \tag{3.42}$$

and the resulting linear system can be solved with a method of choice in order to obtain the coefficients $\{c_i\}_{i=1}^n$. We remark that, for any given location x in the domain, only $p+1$ B-splines are non-zero at x. Therefore, the linear system (3.42) is in general sparse (given that $n \gg p$ typically) and banded.

For 2D spline interpolation, let us introduce a generic 2D tensor-product spline basis of B-splines of degrees p and q, respectively:

$$\{B_{ij}^{pq}(x,y) := B_i^p(x) B_j^q(y)\}_{i,j=1}^{n,m} \,. \tag{3.43}$$

Given $n \times m$ data points $\{(x_i, y_j, z_{ij})\}_{i,j=1}^{n,m}$, corresponding to the values of a given quantity of interest on the 2D logical mesh defined by the interpolation points, an interpolatory spline

$$S(x,y) := \sum_{i=1}^{n} \sum_{j=1}^{m} c_{ij} B_i^p(x) B_j^q(y) \tag{3.44}$$

satisfies the interpolation conditions $S(x_i, y_j) = z_{ij}$, for all $i = 1, \dots, n$ and $j = 1, \dots, m$. These can be solved by decomposing the problem into two separate 1D interpolation problems. More precisely, for each location y_j, with $j = 1, \dots, m$, we first solve the 1D interpolation problem along x

$$\begin{bmatrix} B_1^p(x_1) & \cdots & B_n^p(x_1) \\ \vdots & \ddots & \vdots \\ B_1^p(x_n) & \cdots & B_n^p(x_n) \end{bmatrix} \begin{bmatrix} \bar{c}_{1j} \\ \vdots \\ \bar{c}_{nj} \end{bmatrix} = \begin{bmatrix} z_{1j} \\ \vdots \\ z_{nj} \end{bmatrix}, \quad j = 1, \dots, m\,, \tag{3.45}$$

thus obtaining the intermediate coefficients $\{\bar{c}_{ij}\}_{i,j=1}^{n,m}$. Then, for each location x_i, with $i = 1, \dots, n$, we solve the 1D interpolation problem along y

$$\begin{bmatrix} B_1^q(y_1) & \cdots & B_m^q(y_1) \\ \vdots & \ddots & \vdots \\ B_1^q(y_m) & \cdots & B_m^q(y_m) \end{bmatrix} \begin{bmatrix} c_{i1} \\ \vdots \\ c_{im} \end{bmatrix} = \begin{bmatrix} \bar{c}_{i1} \\ \vdots \\ \bar{c}_{im} \end{bmatrix} , \quad i = 1, \dots, n , \tag{3.46}$$

thus obtaining the coefficients $\{c_{ij}\}_{i,j=1}^{n,m}$.

3.8 Field-aligned interpolation

Interpolation on the flux surfaces (the 2D surfaces spanned by the coordinates (η^2, η^3) for a fixed value of η^1) is performed as described in (Latu et al., 2018, section 2.1). The idea of field-aligned interpolation comes from the observation that standard interpolation based on a centered rectangular stencil of interpolation points, as shown in Figure 3.6, requires a fine mesh in the toroidal direction η^3, in order to achieve good accuracy. The idea of field-aligned interpolation is instead to define a stencil of interpolation points which is approximately aligned to the magnetic field lines, as shown in Figure 3.7, thus allowing for a coarser mesh in the toroidal direction η^3 and reducing the memory footprint of the overall numerical scheme. Figures 3.6 and 3.7 show centered and field-aligned stencils that we would employ in the case of third-order Lagrange interpolation, in order to clarify the general idea. When using splines, field-aligned interpolation allows us to replace 2D spline interpolation in (η^2, η^3) with 1D spline interpolation in η^2 and Lagrange interpolation along the magnetic field lines. More precisely, if we want to interpolate a function $g(\eta^2, \eta^3)$ at a given point (η^2_*, η^3_*) on the flux surface (not belonging to the logical mesh), we first identify a set of adjacent lines at fixed grid values of η^3 in the neighborhood of η^3_*, which we denote by $\{\eta^3_\ell\}$. It is then possible to compute the values of the coordinate η^2 corresponding to the intersections between the magnetic field line passing through the point (η^2_*, η^3_*) and the lines at $\eta^3 = \eta^3_\ell$. We denote such values of the coordinate η^2 as $\xi(\eta^2_*, \eta^3_*, \ell)$, following a notation similar to (Latu et al., 2018, section 2.1). The values of g at these intersection points, $g(\xi(\eta^2_*, \eta^3_*, \ell), \eta^3_\ell)$, are computed by spline interpolation along the direction η^2. The

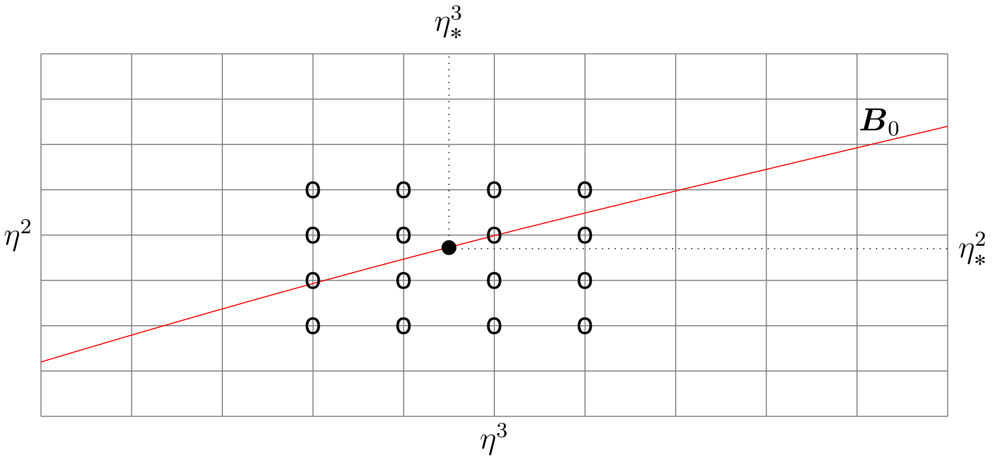


Figure 3.6: Standard interpolation on a flux surface, based on a centered rectangular stencil of interpolation points (empty bullets).

final value at the point (η^2_*, η^3_*) is then computed by Lagrange interpolation from the values $g(\xi(\eta^2_*, \eta^3_*, \ell), \eta^3_\ell)$ obtained in the previous step. In particular, we typically employ a centered stencil and Lagrange polynomials of odd degree.

3.9 Field-aligned differentiation

The parallel derivatives $\nabla_\parallel := \boldsymbol{b}_0 \cdot \boldsymbol{\nabla}$, which we also call field-aligned derivatives, are computed in the same way as described in (Latu et al., 2018, section 2.2). For a given function $g(\eta^2, \eta^3)$ on a flux surface, its parallel derivative at the mesh point $(\eta^2_{i_2}, \eta^3_{i_3})$ is computed with central finite differences of order 6, yielding

$$\nabla_\parallel g(\eta^2_{i_2}, \eta^3_{i_3}) = \frac{b_0^{\eta^3}}{\Delta\eta^3} \sum_{\ell=-3}^{3} w_\ell \, \bar{g}(\xi(\eta^2_{i_2}, \eta^3_{i_3}, i_3 + \ell), \eta^3_{i_3+\ell}) \,, \tag{3.47}$$

where $b_0^{\eta^3}$ denotes the third contravariant component of $\boldsymbol{b}_0$ with respect to the tangent basis $(\boldsymbol{e}_{\eta^1}, \boldsymbol{e}_{\eta^2}, \boldsymbol{e}_{\eta^3})$, $\Delta\eta^3 := 2\pi/n_3$ represents the uniform size of each cell along the toroidal direction η^3, the weights $\{w_\ell\}_{\ell=-3}^3$ are defined as

$$w_0 = 0 \,, \quad w_{-1} = -w_1 = -\frac{3}{4} \,, \quad w_{-2} = -w_2 = \frac{3}{20} \,, \quad w_{-3} = -w_3 = -\frac{1}{60} \,,$$

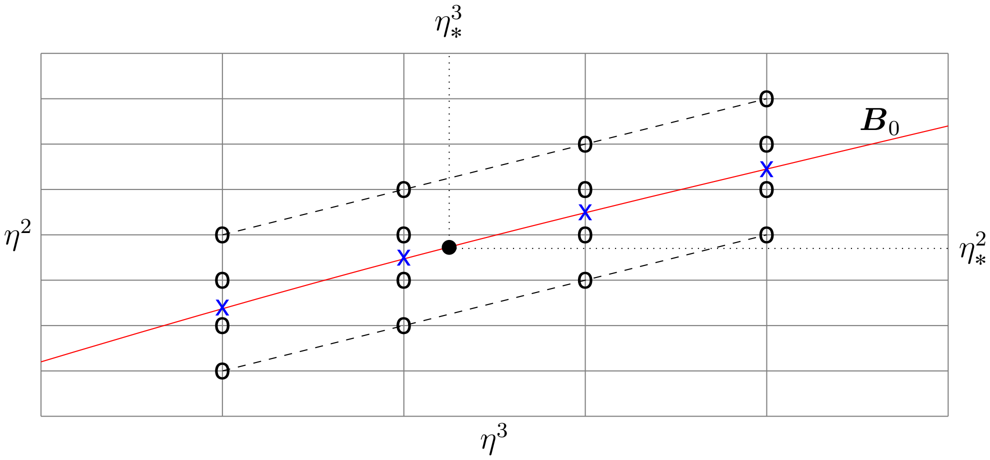


Figure 3.7: Field-aligned interpolation on a flux surface, based on a field-aligned stencil of interpolation points (empty bullets). The blue crosses represent the points $(\xi(\eta^2_*, \eta^3_*, \ell), \eta^3_\ell)$.

and the values $\bar{g}(\xi(\eta^2_{i_2}, \eta^3_{i_3}, i_3 + \ell), \eta^3_{i_3+\ell})$ are computed by spline interpolation along the direction η^2. The algorithm for computing field-aligned derivatives is represented graphically in Figure 3.8.

3.10 Summary of the contributions

In this chapter we introduced the basic features of the semi-Lagrangian drift-kinetic code that we developed within the SeLaLib library. Among all the topics described, our main contributions are:

- the extension of the field-aligned splitting discussed in (Latu et al., 2018) to a general setting based on curvilinear coordinates;
- the improvement of the capabilities offered by the SeLaLib library with respect to our spline discretization, which led to the implementation of a small embedded spline library, offering the possibility to define spline bases on periodic or non-periodic domains, based on uniform or non-uniform meshes, and interpolation tools employing different types of boundary conditions, such as periodic boundary conditions, interpolation on Greville points, and Hermite boundary

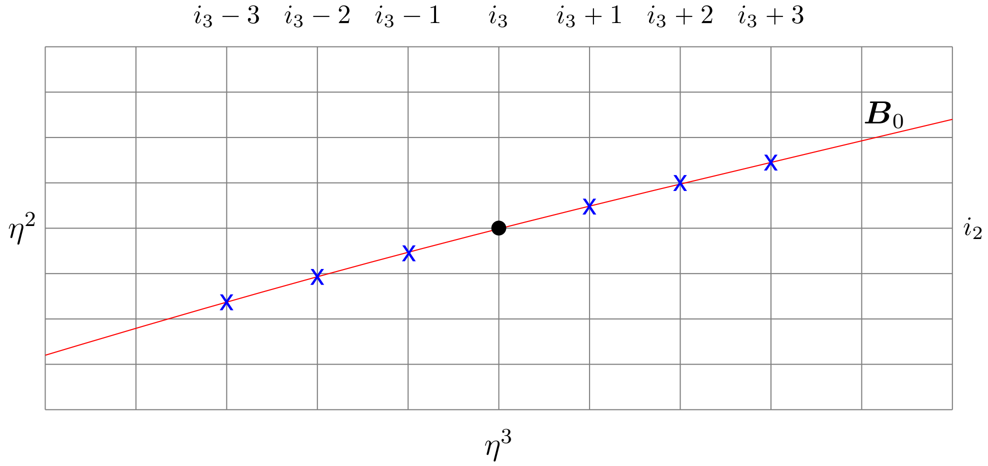


Figure 3.8: Field-aligned derivative of a function g at the mesh point $(\eta^2_{i_2}, \eta^3_{i_3})$ on a flux surface. The blue crosses represent the points $(\xi(\eta^2_{i_2}, \eta^3_{i_3}, i_3+\ell), \eta^3_{i_3+\ell})$, for $l=-3,\ldots,3$.

conditions (which we do not employ, thus neither describe, in this thesis).

The rest of the material presented in this chapter refers to either background mathematical notions or mathematical and computational tools that were already available before this research started and that may have been only slightly improved within the scope of this thesis.

Chapter 4

Coupled hyperbolic-elliptic problems in 2D complex geometries

The field-aligned splitting strategy discussed in the previous chapter allows us to split the 4D Vlasov equation of our drift-kinetic electrostatic model into separate lower-dimensional advection problems. In this chapter we look at the 2D advection on poloidal planes in more detail, including its coupling with the quasi-neutrality equation (described in more detail in chapter 5). If we want to consider realistic geometries, such as, for example, D-shaped Tokamaks, the domains corresponding to the poloidal planes of such geometries can be described, from a mathematical point of view, as singular mapped disk-like domains. The term “mapped” refers to the fact that such domains can be obtained from a rectangular uniform logical domain by applying a given coordinate mapping and the term “singular” refers to the fact that one edge of the rectangular logical domain collapses, through the coordinate mapping, to one point of the physical domain. We refer sometimes to this point as the pole of the domain. In this chapter we suggest a numerical strategy to solve coupled hyperbolic-elliptic problems, similar to the 2D poloidal advection equation coupled to the quasi-neutrality equation of our original 4D model, in such singular domains.

More precisely, we are interested in solving the hyperbolic part of our problem with the semi-Lagrangian method and the elliptic part with a finite element method based on B-splines. One advantage of the semi-Lagrangian method is to avoid any limitation related to the Courant-Friedrichs-Lewy condition (Courant et al., 1928) in the region close to the pole, where the mesh cells become smaller and smaller. Since the method is based on the integration of the characteristics (backward in time), the choice of coordinates to be used while performing this integration turns out to be crucial: such coordinates need indeed to be well-defined in the whole domain, including the pole. The choice of coordinates that we propose fulfills this aim without affecting the robustness, efficiency and accuracy of the numerical scheme. The same coordinates can be used as well for the integration forward in time of the characteristic trajectories of point charges or point-like vortices. On the other hand, the elliptic equation is solved with a finite element method based on B-splines. We require the elliptic solver to return a solution at least $\mathcal{C}^1$ smooth everywhere in the domain. This is difficult to achieve at the pole. Therefore, we follow the approach recently developed by (Toshniwal et al., 2017) to define a set of globally $\mathcal{C}^1$ smooth spline basis functions on singular mapped disk-like domains. A higher degree of smoothness, consistent with the spline degree, may be imposed as well, if needed.

This chapter is organized as follows. Section 4.1 introduces the 2D model that we choose as a test-bed for our numerical methods. Section 4.2 introduces some basic notation that we employ consistently throughout this chapter. Section 4.3 describes singular mapped disk-like domains in more detail. Section 4.4 presents our numerical strategy to solve advection problems on disk-like domains, including numerical tests. Section 4.5 describes our finite element elliptic solver based on globally $\mathcal{C}^1$ smooth splines, including numerical tests. Section 4.6 describes how to couple the two numerical schemes in order to solve a self-consistent hyperbolic-elliptic problem and presents the results of different numerical tests in various disk-like domains. Finally, we end the chapter with a summary of our contributions.

4.1 The physical model

The 2D model that we choose as a test-bed for the numerical methods presented in this chapter is the guiding-center model (O'Neil, 1985; Dubin and O'Neil, 1988)

$$\begin{cases} \dfrac{\partial \rho}{\partial t} - E^y \dfrac{\partial \rho}{\partial x} + E^x \dfrac{\partial \rho}{\partial y} = 0\,, \\ -\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \Phi = \rho\,, \end{cases} \quad \text{with} \quad \begin{cases} \rho(0,x,y) = \rho_{\mathsf{IN}}(x,y)\,, \\ \Phi(t,x,y) = 0 \text{ on } \partial\Omega\,. \end{cases} \tag{4.1}$$

In the context of plasma physics, (4.1) is typically used to describe low-density non-neutral plasmas (Davidson, 2001; Driscoll et al., 2002; Sengupta and Ganesh, 2014, 2015) in a uniform magnetic field $\boldsymbol{B}$ aligned with the direction perpendicular to the (x,y) plane. The unknowns in (4.1) are the density of the plasma particles ρ and the electric scalar potential Φ, related to the electric field via $\boldsymbol{E} = (E^x, E^y)^T = -\boldsymbol{\nabla}\Phi$. The advection field $(-E^y, E^x)$, responsible for the transport of ρ in (4.1), represents the $\boldsymbol{E} \times \boldsymbol{B}$ drift velocity. This model is also equivalent, from a mathematical point of view, to the 2D Euler equations for incompressible inviscid fluids, with $-\rho$ representing the vorticity of the fluid and Φ a stream function. Indeed, (4.1) has been investigated also in the fluid dynamics community for a variety of studies related to vortex dynamics and turbulence (Schecter and Dubin, 1999; Schecter et al., 1999; Schecter and Dubin, 2001; Ganesh and Lee, 2002).

4.2 Notation

In this chapter we denote by $\widehat{\Omega} := [0,1] \times [0,2\pi)$ the logical domain (Figure 4.1) and by $\Omega \subset \mathbb{R}^2$ the physical domain, which is the image of $\widehat{\Omega}$ through a given coordinate transformation $\boldsymbol{F} : \widehat{\Omega} \to \mathbb{R}^2$. In other words, $\Omega := \boldsymbol{F}(\widehat{\Omega})$. We denote by $\boldsymbol{\eta} := (s,\theta) \in \widehat{\Omega}$ and $\boldsymbol{x} := (x,y) \in \Omega$ the logical and Cartesian coordinates, respectively. Moreover, all quantities defined on the logical domain $\widehat{\Omega}$ are denoted by placing a hat over their symbols. On the other hand, the corresponding quantities defined on the physical domain Ω are denoted by the same symbols without the hat. For example, denoting by α a scalar quantity of interest, we have $\widehat{\alpha} : \widehat{\Omega} \to \mathbb{R}$

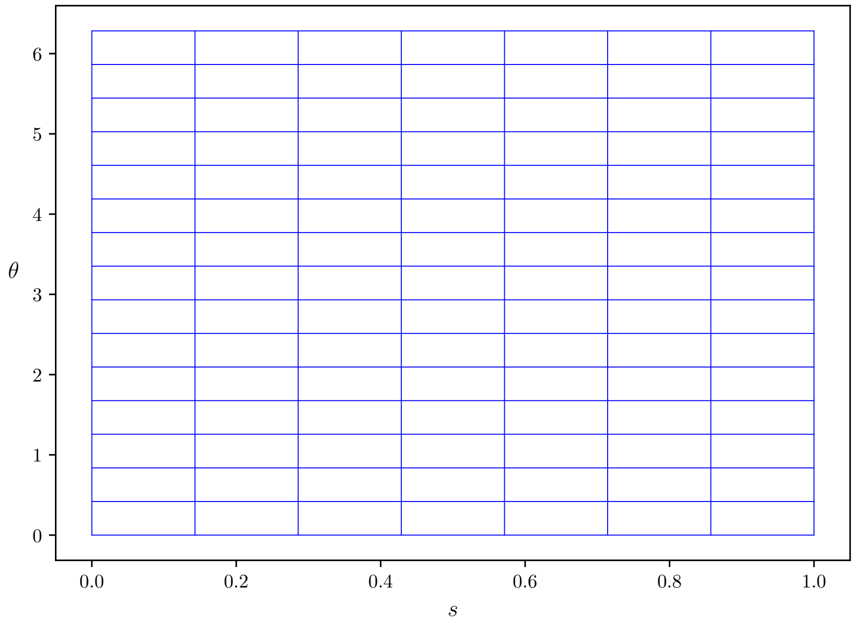


Figure 4.1: Rectangular uniform logical domain $\widehat{\Omega}$ spanned by the logical coordinates (s, θ).

and $\alpha : \Omega \rightarrow \mathbb{R}$, and the two functions are related via $\widehat{\alpha} = \alpha \circ \boldsymbol{F}$. For time-dependent quantities, the domain $\widehat{\Omega}$ (respectively, Ω) is replaced by $\mathbb{R}^+ \times \widehat{\Omega}$ (respectively, $\mathbb{R}^+ \times \Omega$). Moreover, for vector quantities, the codomain $\mathbb{R}$ is replaced by $\mathbb{R}^2$.

4.3 Complex geometries

As mentioned above, we consider here logical domains with a singularity at a unique pole, where the edge $s = 0$ of $\widehat{\Omega}$ collapses to the pole (x_0, y_0) of the physical domain through the mapping $\boldsymbol{F}$: $\boldsymbol{F}(0, \theta) = (x_0, y_0)$ for all θ. We now provide two analytical examples of such mappings. The first mapping is defined by (Bouzat et al., 2018) as

$$\begin{aligned} x(s,\theta) &:= x_0 + (1-\kappa) s \cos\theta - \Delta\, s^2 \,, \\ y(s,\theta) &:= y_0 + (1+\kappa) s \sin\theta \,, \end{aligned} \tag{4.2}$$

where κ and Δ denote the elongation and the Shafranov shift, respectively. For $s = 0$ the mapping collapses to the pole (x_0, y_0). The Jacobian matrix of the mapping, denoted by $J_{\boldsymbol{F}}$, reads

$$J_{\boldsymbol{F}}(s,\theta) = \begin{bmatrix} (1-\kappa)\cos\theta - 2\,\Delta\, s & (\kappa-1)s\sin\theta \\ (1+\kappa)\sin\theta & (1+\kappa)s\cos\theta \end{bmatrix}, \tag{4.3}$$

with determinant

$$\det J_{\boldsymbol{F}}(s,\theta) = s(1+\kappa)[(1-\kappa) - 2\,\Delta\, s\cos\theta]\,, \tag{4.4}$$

which vanishes at the pole. The Jacobian matrix of the inverse transformation reads

$$J_{\boldsymbol{F}}^{-1}(s,\theta) = \frac{1}{\det J_{\boldsymbol{F}}(s,\theta)} \begin{bmatrix} (1+\kappa)s\cos\theta & (1-\kappa)s\sin\theta \\ -(1+\kappa)\sin\theta & (1-\kappa)\cos\theta - 2\,\Delta\, s \end{bmatrix}, \tag{4.5}$$

and it is singular at the pole. The second mapping is defined by (Czarny and Huysmans, 2008) as

$$\begin{aligned} x(s,\theta) &:= \frac{1}{\varepsilon}\left(1 - \sqrt{1 + \varepsilon(\varepsilon + 2\, s\cos\theta)}\right), \\ y(s,\theta) &:= y_0 + \frac{e\,\xi\, s\sin\theta}{2 - \sqrt{1+\varepsilon(\varepsilon + 2\, s\cos\theta)}} = y_0 + \frac{e\,\xi\, s\sin\theta}{1 + \varepsilon\, x(s,\theta)}, \end{aligned} \tag{4.6}$$

where ε and e denote the inverse aspect ratio and the ellipticity, respectively, and $\xi := 1/\sqrt{1-\varepsilon^2/4}$. For $s = 0$ the mapping collapses to the pole $\left(\left(1 - \sqrt{1+\varepsilon^2}\right)/\varepsilon, y_0\right)$. The Jacobian matrix of the mapping reads

$$J_{\boldsymbol{F}}(s,\theta) = \frac{e\,\xi}{1+\varepsilon\, x(s,\theta)} \begin{bmatrix} -\dfrac{1+\varepsilon\, x(s,\theta)}{1-\varepsilon\, x(s,\theta)}\dfrac{\cos\theta}{e\,\xi} & \dfrac{1+\varepsilon\, x(s,\theta)}{1-\varepsilon\, x(s,\theta)}\dfrac{s\sin\theta}{e\,\xi} \\ \sin\theta + \dfrac{\varepsilon\, s\sin\theta\cos\theta}{1-\varepsilon^2\, x^2(s,\theta)} & s\cos\theta - \dfrac{\varepsilon\, s^2\sin^2\theta}{1-\varepsilon^2\, x^2(s,\theta)} \end{bmatrix}, \tag{4.7}$$

with determinant

$$\det J_{\boldsymbol{F}}(s,\theta) = \frac{s}{\varepsilon\, x(s,\theta) - 1}, \tag{4.8}$$

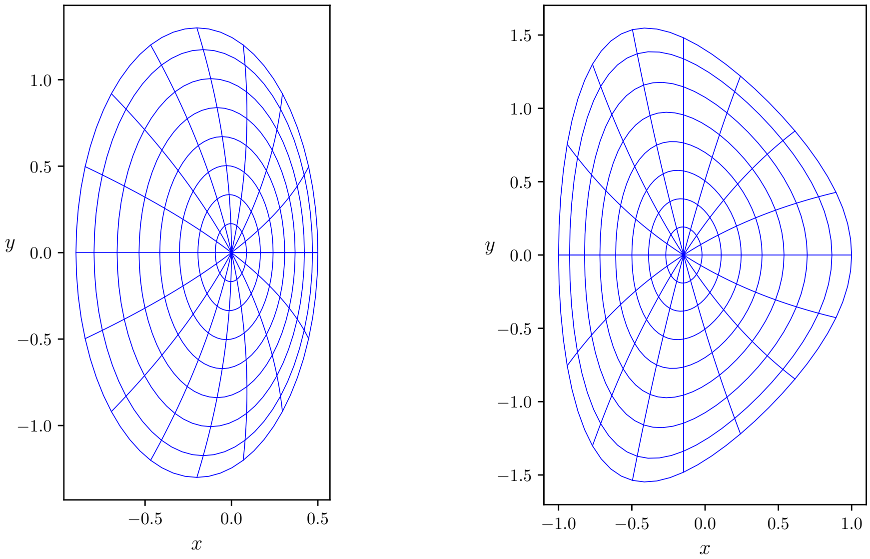


Figure 4.2: Disk-like domains defined by the mappings (4.2) (left) and (4.6) (right). The lines originating from the pole are isolines at constant θ, while the lines concentric around the pole are isolines at constant s.

which vanishes at the pole. The Jacobian matrix of the inverse transformation reads

$$J_{\boldsymbol{F}}^{-1}(s,\theta) = \frac{1}{\det J_{\boldsymbol{F}}(s,\theta)} \begin{bmatrix} s\cos\theta - \dfrac{\varepsilon\, s^2 \sin^2\theta}{1-\varepsilon^2\, x^2(s,\theta)} & -\dfrac{1+\varepsilon\, x(s,\theta)}{1-\varepsilon\, x(s,\theta)} \dfrac{s\sin\theta}{e\,\xi} \\ -\sin\theta - \dfrac{\varepsilon\, s \sin\theta\cos\theta}{1-\varepsilon^2\, x^2(s,\theta)} & -\dfrac{1+\varepsilon\, x(s,\theta)}{1-\varepsilon\, x(s,\theta)} \dfrac{\cos\theta}{e\,\xi} \end{bmatrix}, \tag{4.9}$$

and it is again singular at the pole. In all the numerical tests considered in this thesis, mapping (4.2) is set up with the parameters

$$(x_0, y_0) = (0,0)\,, \quad \kappa = 0.3\,, \quad \Delta = 0.2\,, \tag{4.10}$$

and mapping (4.6) is set up with the parameters

$$y_0 = 0\,, \quad \varepsilon = 0.3\,, \quad e = 1.4\,, \tag{4.11}$$

which yield $x_0 \approx -0.15$ and $\xi \approx 1.0114$. Figure 4.2 shows the physical domains obtained with these mappings.

In practical applications it may be not possible to have an analytical description of the mapping that represents the physical domain of interest, as in the examples discussed above. Moreover, we are going to solve the elliptic equation in (4.1) with a finite element method based on B-splines. This means that our numerical strategy is based on a machinery inherently defined at the discrete level. We therefore need to have a discrete representation of the analytical mappings discussed above. Based on the 2D spline basis defined in section 3.6, we define a discrete representation of our analytical mappings as

$$\begin{aligned}
x(s,\theta) &:= \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2} c^x_{i_1 i_2}\widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta) = x_0\,\widehat{B}^s_1(s) + \sum_{i_1=2}^{n_1}\sum_{i_2=1}^{n_2} c^x_{i_1 i_2}\widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta)\,, \\
y(s,\theta) &:= \sum_{i_1=1}^{n_1}\sum_{i_2=1}^{n_2} c^y_{i_1 i_2}\widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta) = y_0\,\widehat{B}^s_1(s) + \sum_{i_1=2}^{n_1}\sum_{i_2=1}^{n_2} c^y_{i_1 i_2}\widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta)\,.
\end{aligned} \tag{4.12}$$

We remark again a few differences between the notation employed here (and in the following) and the corresponding notation employed in section 3.6. The logical coordinates are now denoted as (s,θ) instead of (η^1,η^2), a hat is placed over the symbols denoting the basis functions, when they are defined on the logical domain, and the superscripts indicating the degrees p_1 and p_2 have been removed in favor of superscripts indicating the functional dependence of the basis functions (either on s or on θ).

In all the numerical tests considered in this thesis, the control points $\{(c^x_{i_1 i_2}, c^y_{i_1 i_2})\}_{i_1,i_2=1}^{n_1,n_2}$ are obtained by interpolating a given analytical mapping on the interpolation points defined in section 3.6. For more general practical applications, they could be given as an input from any code capable of constructing a mesh conformal to the flux surfaces of a given equilibrium magnetic field, such as, for example, the software Tokamesh (Guillard et al., 2018). Finally, we note that all the control points at $i_1 = 1$ are equal to the pole, $(c^x_{1 i_2}, c^y_{1 i_2}) = (x_0, y_0)$, which is another way of saying that the edge $s = 0$ of the logical domain collapses to the pole of the physical domain.

In order to compute integrals on the logical domain $\widehat{\Omega}$ (in the subsequent sections), $p_1 + 1$ Gauss-Legendre quadrature points and weights are introduced in each cell of the domain $[0,1]$ along s and $p_2 + 1$ Gauss-Legendre quadrature points and weights are introduced in each cell of the domain $[0, 2\pi)$ along θ.

4.4 Semi-Lagrangian advection solver

We now focus on the advection equation in (4.1) and write it in terms of a general advection field $\boldsymbol{U} := (-E^y, E^x)^T$ as

$$\frac{\partial \rho}{\partial t} + \boldsymbol{U} \cdot \boldsymbol{\nabla} \rho = 0 \,. \tag{4.13}$$

We now discuss the optimal choice of coordinates for the integration backward in time of the characteristics of (4.13). We recall that the advection field available at the discrete level is the advection field $\widehat{\boldsymbol{U}}$ defined on the logical domain. It is natural to think of integrating the characteristic equations in either Cartesian or logical coordinates. However, both choices present some drawbacks. The characteristic equations in Cartesian coordinates $\boldsymbol{x}$ read

$$\dot{\boldsymbol{x}} = \widehat{\boldsymbol{U}}(t, \boldsymbol{F}^{-1}(\boldsymbol{x})) \,. \tag{4.14}$$

These equations are well defined everywhere in the domain, but they become computationally expensive if the mapping $\boldsymbol{F}$ is not easy to invert. On the other hand, the characteristic equations in logical coordinates $\boldsymbol{\eta}$ read

$$\dot{\boldsymbol{\eta}} = J_{\boldsymbol{F}}^{-1} \widehat{\boldsymbol{U}}(t, \boldsymbol{\eta}) \,. \tag{4.15}$$

These equations are not defined at the pole $s = 0$, because $J_{\boldsymbol{F}}^{-1}$ is singular there. We then suggest to introduce the new coordinates $\boldsymbol{X} := (X, Y)$, defined by the polar transformation

$$\begin{aligned} X(s, \theta) &:= s \cos \theta \,, \\ Y(s, \theta) &:= s \sin \theta \,, \end{aligned} \tag{4.16}$$

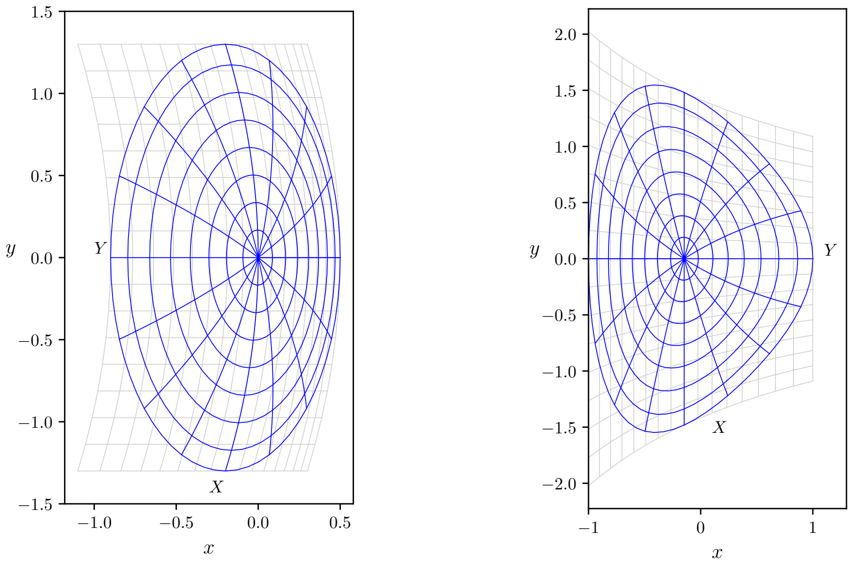


Figure 4.3: Pseudo-Cartesian coordinates: the light-gray grids represent the grids in the pseudo-Cartesian coordinates (X, Y) for disk-like domains defined by the mappings (4.2) (left) and (4.6) (right).

which we name pseudo-Cartesian coordinates. We denote by $\boldsymbol{G} : \widehat{\Omega} \rightarrow \mathbb{R}^2$ the new mapping defined by $\boldsymbol{G}(\boldsymbol{\eta}) := \boldsymbol{X}$, and by $J_{\boldsymbol{G}}$ its Jacobian, given by

$$J_{\boldsymbol{G}}(s, \theta) = \begin{bmatrix} \cos\theta & -s\sin\theta \\ \sin\theta & s\cos\theta \end{bmatrix} . \tag{4.17}$$

The pseudo-Cartesian coordinates for the domains defined by the mappings (4.2) and (4.6) are shown in Figure 4.3. In the simplest case of a circular mapping, they reduce to standard Cartesian coordinates. The characteristic equations in pseudo-Cartesian coordinates $\boldsymbol{X}$ read

$$\dot{\boldsymbol{X}} = (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1} \widehat{\boldsymbol{U}}(t, \boldsymbol{G}^{-1}(\boldsymbol{X})) , \tag{4.18}$$

where $J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1}$ represents the Jacobian of the composite mapping $\boldsymbol{F} \circ \boldsymbol{G}^{-1}$ defined by $\boldsymbol{F} \circ \boldsymbol{G}^{-1}(\boldsymbol{X}) = \boldsymbol{x}$. For a circular mapping, $\boldsymbol{F} \circ \boldsymbol{G}^{-1}$ reduces to the identity and (4.18) reduces to (4.14), which works well because $\boldsymbol{F}^{-1}$ (inverse polar transformation) is easy to compute. For more complex non-circular mappings, (4.18) is more convenient than (4.14) because the mapping $\boldsymbol{G}$ is easier to invert than the original mapping $\boldsymbol{F}$. More precisely, the inverse mapping $\boldsymbol{G}^{-1}$ is analytical and reads

$$\begin{aligned} s(X,Y) &= \sqrt{X^2+Y^2}\,, \\ \theta(X,Y) &= \mathsf{atan}2(Y,X)\,, \end{aligned} \tag{4.19}$$

where $\mathsf{atan}2(Y,X)$ returns the principal value of the argument function applied to the complex number $X+iY$ in the range $(-\pi,\pi]$ (which then must be shifted appropriately to the domain $[0,2\pi)$). Moreover, the inverse Jacobian matrix $(J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}$ in (4.18) turns out to be well-behaved everywhere in the physical domain, including the pole. More precisely, the singularity of the inverse Jacobian matrix

$$J_{\boldsymbol{G}}^{-1}(s,\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\dfrac{1}{s}\sin\theta & \dfrac{1}{s}\cos\theta \end{bmatrix}, \tag{4.20}$$

in the limit $s\to 0^+$, is cancelled by the matrix elements of $J_{\boldsymbol{F}}$. The product $J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1}$ in general reads

$$J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1}(s,\theta) = \begin{bmatrix} \dfrac{\partial x}{\partial s}\cos\theta - \dfrac{1}{s}\dfrac{\partial x}{\partial\theta}\sin\theta & \dfrac{\partial x}{\partial s}\sin\theta + \dfrac{1}{s}\dfrac{\partial x}{\partial\theta}\cos\theta \\ \dfrac{\partial y}{\partial s}\cos\theta - \dfrac{1}{s}\dfrac{\partial y}{\partial\theta}\sin\theta & \dfrac{\partial y}{\partial s}\sin\theta + \dfrac{1}{s}\dfrac{\partial y}{\partial\theta}\cos\theta \end{bmatrix}. \tag{4.21}$$

From an analytical point of view, (4.21) holds for all values of s except at the pole $s=0$. However, the products $\dfrac{1}{s}\dfrac{\partial x}{\partial\theta}$ and $\dfrac{1}{s}\dfrac{\partial y}{\partial\theta}$ are finite and well-defined in the limit $s\to 0^+$. From a numerical point of view, (4.21) holds for all values of s sufficiently far from the pole, as far as the factor $1/s$ does not become too large. Therefore, we assume that (4.21) holds for $s \geqslant \epsilon$, for a given small ϵ. More precisely, the derivatives $\partial x/\partial\theta$ and $\partial y/\partial\theta$ vanish for $s=0$. Hence, expanding in s around $s=0$, we have

$$\begin{aligned} \frac{\partial x}{\partial\theta}(s,\theta) &= s\,\frac{\partial^2 x}{\partial s\,\partial\theta}(0,\theta) + O(s^2)\,, \\ \frac{\partial y}{\partial\theta}(s,\theta) &= s\,\frac{\partial^2 y}{\partial s\,\partial\theta}(0,\theta) + O(s^2)\,, \end{aligned} \tag{4.22}$$

which yields

$$\begin{aligned}\lim_{s\to0^+}\frac{1}{s}\frac{\partial x}{\partial\theta}(s,\theta) &= \frac{\partial^2 x}{\partial s\,\partial\theta}(0,\theta)\,,\\ \lim_{s\to0^+}\frac{1}{s}\frac{\partial y}{\partial\theta}(s,\theta) &= \frac{\partial^2 y}{\partial s\,\partial\theta}(0,\theta)\,.\end{aligned}\tag{4.23}$$

Therefore, the product $J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1}$ at the pole $s=0$ reads

$$J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1}(0,\theta) = \begin{bmatrix} \frac{\partial x}{\partial s}(0,\theta)\cos\theta - \frac{\partial^2 x}{\partial s\,\partial\theta}(0,\theta)\sin\theta & \frac{\partial x}{\partial s}(0,\theta)\sin\theta + \frac{\partial^2 x}{\partial s\,\partial\theta}(0,\theta)\cos\theta \\ \frac{\partial y}{\partial s}(0,\theta)\cos\theta - \frac{\partial^2 y}{\partial s\,\partial\theta}(0,\theta)\sin\theta & \frac{\partial y}{\partial s}(0,\theta)\sin\theta + \frac{\partial^2 y}{\partial s\,\partial\theta}(0,\theta)\cos\theta \end{bmatrix}.\tag{4.24}$$

For example, in the case of mapping (4.2) we get

$$(J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1})^{-1}(0,\theta) = \begin{bmatrix} \dfrac{1}{1-\kappa} & 0 \\ 0 & \dfrac{1}{1+\kappa} \end{bmatrix},\tag{4.25}$$

and, similarly, in the case of mapping (4.6) we get

$$(J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1})^{-1}(0,\theta) = \begin{bmatrix} -\sqrt{1+\varepsilon^2} & 0 \\ 0 & \dfrac{2-\sqrt{1+\varepsilon^2}}{e\,\xi} \end{bmatrix}.\tag{4.26}$$

In order to connect (4.21) and (4.24) in a smooth way, for $0<s<\epsilon$ we interpolate linearly the value at the pole $s=0$ and the value at $s=\epsilon$, obtaining

$$(J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1})^{-1}(s,\theta) = \left(1-\frac{s}{\epsilon}\right)(J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1})^{-1}(0,\theta) + \frac{s}{\epsilon}(J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1})^{-1}(\epsilon,\theta)\,.\tag{4.27}$$

We remark that the result obtained in (4.24) needs to be single-valued, and hence should not depend on the angle-like variable θ. This is true if we consider analytical mappings such as (4.2) and (4.6), as demonstrated by (4.25) and (4.26), respectively. If we consider, instead, a discrete representation of the above-mentioned mappings, defined, for example, in terms of splines, we observe a residual dependence of (4.24) on θ. It is possible to measure the discrepancy between the matrix elements of $(J_{\boldsymbol{F}}J_{\boldsymbol{G}}^{-1})^{-1}(0,\theta)$, computed by inverting (4.24) (with the derivatives

	Circular mapping		Mapping (4.2)		Mapping (4.6)	
$n_1 \times n_2$	Error	Order	Error	Order	Error	Order
16×32	8.30×10^{-6}		1.19×10^{-5}		8.66×10^{-6}	
32×64	5.17×10^{-7}	4.01	7.38×10^{-7}	4.01	5.39×10^{-7}	4.01
64×128	3.23×10^{-8}	4.00	4.61×10^{-8}	4.00	3.37×10^{-8}	4.00
128×256	2.02×10^{-9}	4.00	2.88×10^{-9}	4.00	2.94×10^{-9}	3.52
256×512	1.26×10^{-10}	4.00	1.80×10^{-10}	4.00	3.69×10^{-10}	3.00

Table 4.1: Convergence of the product $(J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}$ to the θ-independent analytical values for a circular mapping and for the mappings (4.2) and (4.6).

evaluated from the discrete spline mapping (4.12)), and the corresponding analytical θ-independent matrix elements. As a measure of the error, we consider the maximum among all matrix elements and all values of θ for a given interpolation grid. The results in Table 4.1 show that such errors become asymptotically small as the computational mesh is refined (that is, as the number of interpolation points is increased). Such errors do not constitute a problem if they turn out to be smaller than the overall numerical accuracy of our scheme. However, we suggest to guarantee that (4.24) is truly single-valued by taking an average of (4.24) over all values of θ in the interpolation grid. This may become particularly useful if implicit integration schemes are used, when the magnitude of the above-mentioned errors may become comparable to the tolerances chosen for the implicit methods of choice. We also remark that the parameter ϵ can be chosen arbitrarily small, as far as it avoids underflows and overflows in floating point arithmetic. For all the numerical tests presented in this thesis we set $\epsilon = 10^{-12}$.

4.4.1 Numerical results

We test our advection solver for the stationary rotating advection field

$$\boldsymbol{U}(x,y) := \omega \begin{pmatrix} y_c - y \\ x - x_c \end{pmatrix} , \tag{4.28}$$

with $\omega = 2\pi$ and $(x_c, y_c) = (0.25, 0)$. The numerical test is performed on mapping (4.6) with the parameters in (4.11). The flow field corresponding to the advection field (4.28) can be computed analytically and reads

$$\begin{aligned} x(t+\Delta t) &= x_c + (x(t) - x_c)\cos(\omega \Delta t) - (y(t) - y_c)\sin(\omega \Delta t) \, , \\ y(t+\Delta t) &= y_c + (x(t) - x_c)\sin(\omega \Delta t) + (y(t) - y_c)\cos(\omega \Delta t) \, , \end{aligned} \tag{4.29}$$

where Δt denotes the time step. Therefore, the numerical solution can be compared to the exact one obtained from the analytical flow field by the method of characteristics, $\rho_{\text{ex}}(t, x(t), y(t)) = \rho(0, x(0), y(0))$, where the initial positions $x(0)$ and $y(0)$ are obtained from (4.29) with $\Delta t = -t$. The initial condition is set to a superposition of cosine bells with elliptical cross sections:

$$\rho(0,x,y) := \frac{1}{2} \left[\mathcal{G}\left(r_1(x,y)\right) + \mathcal{G}\left(r_2(x,y)\right) \right] \, , \tag{4.30}$$

where $\mathcal{G}(r)$ is defined as

$$\mathcal{G}(r) := \begin{cases} \cos^4\left(\dfrac{\pi r}{2a}\right) & r < a \, , \\ 0 & \text{elsewhere} \, , \end{cases} \tag{4.31}$$

with $a = 0.3$, and $r_1(x,y)$ and $r_2(x,y)$ are defined as

$$\begin{aligned} r_1(x,y) &:= \sqrt{(x-\bar{x})^2 + 8(y-\bar{y})^2} \, , \\ r_2(x,y) &:= \sqrt{8(x-\bar{x})^2 + (y-\bar{y})^2} \, . \end{aligned} \tag{4.32}$$

This test case is inspired by one presented in (Güçlü et al., 2014, section 5.2): the non-Gaussian initial condition allows us to possibly detect any deformation of the initial density perturbation while rotating under the action of the advection field (4.28). Denoting by $\Delta\rho := \rho - \rho_{\text{ex}}$ the numerical error, that is, the difference between the numerical solution and the exact one, measures of the error of our numerical scheme are obtained by taking the L^∞-norm in time of the spatial L^2-norm of $\Delta\rho$,

$$\begin{aligned} \max_t ||\Delta\rho||_{L^2} &:= \max_t \left(\sqrt{\int_\Omega \mathrm{d}x\,\mathrm{d}y\ [\Delta\rho(t,x,y)]^2} \right) \\ &= \max_t \left(\sqrt{\int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}(s,\theta)|\,[\Delta\widehat{\rho}(t,s,\theta)]^2} \right) , \end{aligned} \tag{4.33}$$

computed using the Gauss-Legendre quadrature points and weights mentioned in section 4.3, and the L^∞-norm in time of the spatial L^∞-norm of $\Delta\rho$,

$$\max_t ||\Delta\rho||_{L^\infty} := \max_t \max_{(x,y)\in\Omega} |\Delta\rho(t,x,y)| = \max_t \max_{(s,\theta)\in\widehat{\Omega}} |\Delta\widehat{\rho}(t,s,\theta)| , \tag{4.34}$$

computed on the Greville points. We remark that the pole is included when we estimate the spatial L^∞-norm (4.34). Table 4.2 shows the convergence of our numerical scheme while decreasing the time step Δt and correspondingly refining the spatial mesh by increasing the number of points n_1 in the direction s and the number of points n_2 in the direction θ (in order to keep the CFL number constant), using cubic splines and an explicit third-order Runge-Kutta method for the integration of the characteristics. We note that there are no effects of order reduction due to the singularity at the pole. Standard tensor-product spline interpolation turns out to work well in the case of analytical advection fields, provided our choice of coordinates for the integration of the characteristics.

The time integration algorithm is as follows. Starting from a mesh point $\boldsymbol{\eta}_{ij} := (s_i, \theta_j)$ with pseudo-Cartesian coordinates $\boldsymbol{X}_{ij} := \boldsymbol{G}(\boldsymbol{\eta}_{ij})$, we compute the first-stage,

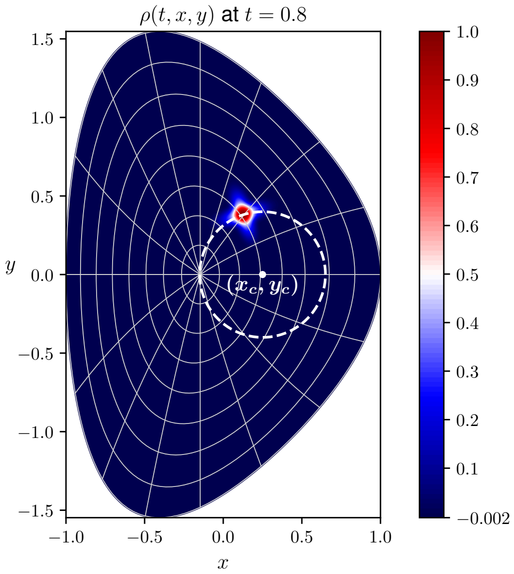


Figure 4.4: Numerical test of the advection solver: contour plot of the density $\rho(t, x, y)$ at time $t = 0.8$. The dashed circle represents the trajectory that the initial density perturbation is expected to follow under the action of the rotating advection field (4.28).

Δt	$n_1 \times n_2$	$\max_t \|\|\Delta\rho\|\|_{L^2}$	Order	$\max_t \|\|\Delta\rho\|\|_{L^\infty}$	Order
0.1	64×128	3.25×10^{-2}		3.53×10^{-1}	
$0.1/2$	128×256	4.10×10^{-3}	2.99	4.34×10^{-2}	3.02
$0.1/4$	256×512	5.11×10^{-4}	3.00	5.09×10^{-3}	3.09
$0.1/8$	512×1024	6.39×10^{-5}	3.00	6.13×10^{-4}	3.05
$0.1/16$	1024×2048	7.98×10^{-6}	3.00	7.52×10^{-5}	3.03

Table 4.2: Third-order convergence of the advection solver using cubic splines and an explicit third-order Runge-Kutta method for the integration of the characteristics.

second-stage and third-stage derivatives and solutions

$$\begin{aligned}
1.\quad & \dot{\boldsymbol{X}}_{ij}^{(1)} := (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij})\, \widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}) \\
& \boldsymbol{X}_{ij}^{(1)} := \boldsymbol{X}_{ij} - \frac{\Delta t}{2} \dot{\boldsymbol{X}}_{ij}^{(1)} \\
& \boldsymbol{\eta}_{ij}^{(1)} := \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(1)})
\end{aligned} \tag{4.35a}$$

$$\begin{aligned}
2.\quad & \dot{\boldsymbol{X}}_{ij}^{(2)} := (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij}^{(1)})\, \widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}^{(1)}) \\
& \boldsymbol{X}_{ij}^{(2)} := \boldsymbol{X}_{ij} - \Delta t \left[2\dot{\boldsymbol{X}}_{ij}^{(2)} - \dot{\boldsymbol{X}}_{ij}^{(1)} \right] \\
& \boldsymbol{\eta}_{ij}^{(2)} := \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(2)})
\end{aligned} \tag{4.35b}$$

$$\begin{aligned}
3.\quad & \dot{\boldsymbol{X}}_{ij}^{(3)} := (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij}^{(2)})\, \widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}^{(2)}) \\
& \boldsymbol{X}_{ij}^{(3)} := \boldsymbol{X}_{ij} - \frac{\Delta t}{6} \left[\dot{\boldsymbol{X}}_{ij}^{(1)} + 4\dot{\boldsymbol{X}}_{ij}^{(2)} + \dot{\boldsymbol{X}}_{ij}^{(3)} \right] \\
& \boldsymbol{\eta}_{ij}^{(3)} := \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(3)})
\end{aligned} \tag{4.35c}$$

The logical coordinates $\boldsymbol{\eta}_{ij}^{(3)}$ obtained represent the origin of the characteristic at time $t - \Delta t$ passing through the point $\boldsymbol{\eta}_{ij}$ at time t.

4.5 Spline finite element elliptic solver

We now focus on the elliptic equation in (4.1) and write it in the more general form

$$-\boldsymbol{\nabla} \cdot (\alpha \boldsymbol{\nabla} \Phi) + \beta\, \Phi = \rho\,, \tag{4.36}$$

which reduces to Poisson's equation $-\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\Phi = \rho$ for $\alpha = 1$ and $\beta = 0$. We want to solve this equation with a finite element method based on B-splines. Following an isogeometric approach, we use the same spline basis used to construct the discrete spline mappings (4.12) as a basis for our finite element method. Moreover, our aim is to obtain a potential Φ which is at least $\mathcal{C}^1$ smooth everywhere in the physical domain, including the pole, so that the corresponding advection fields for the transport of ρ are at least continuous. This is achieved by imposing appropriate $\mathcal{C}^1$ smoothness constraints on the spline basis while solving the linear system obtained from the weak form of (4.36). A systematic approach to define a set of globally $\mathcal{C}^1$ smooth spline basis functions on singular mapped disk-like domains was suggested in (Toshniwal et al., 2017) and we now recall its basic ideas.

4.5.1 $\mathcal{C}^1$ smooth polar splines

The idea is to satisfy the $\mathcal{C}^1$ smoothness requirements by imposing appropriate constraints on the $2n_2$ degrees of freedom corresponding to $i_1 = 1, 2$ for all i_2. More precisely, the $2n_2$ basis functions corresponding to these degrees of freedom are replaced by only three new basis functions, defined as linear combinations of the existing ones. In order to guarantee the properties of partition of unity and positivity, (Toshniwal et al., 2017) suggests to use barycentric coordinates to construct these linear combinations. Taking an equilateral triangle enclosing the pole and the first row of control points $(c^x_{2\,i_2}, c^y_{2\,i_2})$, with vertices

$$V_1 := (x_0 + \tau, y_0)\,, \quad V_2 := \left(x_0 - \frac{\tau}{2}, y_0 + \frac{\sqrt{3}}{2}\tau\right), \quad V_3 := \left(x_0 - \frac{\tau}{2}, y_0 - \frac{\sqrt{3}}{2}\tau\right),$$

with τ defined as

$$\begin{aligned} \tau := \max \Bigg[& \max_{i_2}(-2(c^x_{2\,i_2} - x_0)), \\ & \max_{i_2}((c^x_{2\,i_2} - x_0) - \sqrt{3}(c^y_{2\,i_2} - y_0)), \\ & \max_{i_2}((c^x_{2\,i_2} - x_0) + \sqrt{3}(c^y_{2\,i_2} - y_0))\Bigg], \end{aligned} \tag{4.37}$$

we denote by $(\lambda_1, \lambda_2, \lambda_3)$ the barycentric coordinates of any point with respect to the vertices of this triangle:

$$\lambda_1(x, y) := \frac{1}{3} + \frac{2}{3}\frac{1}{\tau}(x - x_0)\,, \tag{4.38a}$$

$$\lambda_2(x, y) := \frac{1}{3} - \frac{1}{3}\frac{1}{\tau}(x - x_0) + \frac{\sqrt{3}}{3}\frac{1}{\tau}(y - y_0)\,, \tag{4.38b}$$

$$\lambda_3(x, y) := \frac{1}{3} - \frac{1}{3}\frac{1}{\tau}(x - x_0) - \frac{\sqrt{3}}{3}\frac{1}{\tau}(y - y_0)\,. \tag{4.38c}$$

Then, the three new basis functions are denoted by $\widehat{\mathcal{B}}_l$, for $l = 1, 2, 3$, and defined as

$$\widehat{\mathcal{B}}_l(s,\theta) := \sum_{i_1=1}^{2}\sum_{i_2=1}^{n_2} \lambda_l(c^x_{i_1 i_2}, c^y_{i_1 i_2})\widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta)\,. \tag{4.39}$$

It can be shown that these basis functions are positive, $\widehat{\mathcal{B}}_l(s,\theta) \geqslant 0\ \forall(s,\theta)$ and $\forall l$, and that they satisfy the partition of unity property, namely

$$\sum_{l=1}^{3}\widehat{\mathcal{B}}_l(s,\theta) + \sum_{i_1=3}^{n_1}\sum_{i_2=1}^{n_2} \widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta) = 1 \quad \forall(s,\theta)\,. \tag{4.40}$$

Moreover, the new basis functions $\mathcal{B}_l$, related to $\widehat{\mathcal{B}}_l$ via $\widehat{\mathcal{B}}_l = \mathcal{B}_l \circ \boldsymbol{F}$, are $\mathcal{C}^1$ smooth everywhere in the physical domain.

4.5.2 Finite element solver

We now consider a more general version of the elliptic equation (4.36) which includes a finite set of n_c point charges, denoted with the label c, of charges q_c and positions (x_c, y_c). Denoting by ρ_{SL} and ρ_{PIC} the semi-Lagrangian density and the particle density, respectively, we rewrite (4.36) as

$$-\boldsymbol{\nabla}\cdot(\alpha\boldsymbol{\nabla}\Phi) + \beta\,\Phi = \rho_{\mathrm{SL}} + \rho_{\mathrm{PIC}}\,, \tag{4.41}$$

with the particle density ρ_{PIC} defined as

$$\rho_{\mathrm{PIC}}(x,y) := \sum_{c=1}^{n_c} q_c\,\delta(x-x_c)\delta(y-y_c)\,, \tag{4.42}$$

and homogeneous Dirichlet boundary conditions $\Phi(x,y) = 0$ on $\partial\Omega$ (omitting the time dependence of Φ). We impose these boundary conditions by removing the corresponding basis functions from both the test space and the trial space. More precisely, we choose as test and trial spaces the space defined by the tensor-product spline basis $\{\widehat{B}_{i_1 i_2}(s,\theta) := \widehat{B}^s_{i_1}(s)\widehat{B}^\theta_{i_2}(\theta)\}_{i_1,i_2=1}^{n_1-1,n_2}$, where we remove the last n_2 basis functions corresponding to $i_1 = n_1$. Hence, the weak form of (4.41) reads

$$\int_\Omega \mathrm{d}x\,\mathrm{d}y\;\left(\alpha\boldsymbol{\nabla}\Phi\cdot\boldsymbol{\nabla}B_{i_1i_2} + \beta\,\Phi B_{i_1i_2}\right) = \int_\Omega \mathrm{d}x\,\mathrm{d}y\,\rho_{\mathrm{SL}}\,B_{i_1i_2} + \sum_{c=1}^{n_c} q_c B_{i_1i_2}(x_c, y_c)\,,$$

for all $i_1 = 1,\dots,n_1-1$ and $i_2 = 1,\dots,n_2$. We now expand Φ on the trial space and ρ_{SL} on the full tensor-product space (without removing the last n_2 basis functions, as the space where ρ_{SL} is defined is completely independent from the test and trial spaces):

$$\Phi = \sum_{j_1=1}^{n_1-1}\sum_{j_2=1}^{n_2}\Phi_{j_1j_2}\,B_{j_1j_2}\,,\qquad \rho_{\mathrm{SL}} = \sum_{k_1=1}^{n_1}\sum_{k_2=1}^{n_2}\rho_{k_1k_2}\,B_{k_1k_2}\,. \tag{4.43}$$

To sum up, the following integer indices are being used:

$$\begin{aligned} &i_1 = 1,\dots,n_1-1 && i_2 = 1,\dots,n_2 && \text{(test space)}\\ &j_1 = 1,\dots,n_1-1 && j_2 = 1,\dots,n_2 && \text{(trial space)}\\ &k_1 = 1,\dots,n_1 && k_2 = 1,\dots,n_2 && \text{(space of } \rho_{\mathrm{SL}}\text{)} \end{aligned} \tag{4.44}$$

Hence, we obtain

$$\begin{aligned} &\sum_{j_1=1}^{n_1-1}\sum_{j_2=1}^{n_2}\Phi_{j_1j_2}\int_\Omega \mathrm{d}x\,\mathrm{d}y\;\left(\alpha\boldsymbol{\nabla}B_{j_1j_2}\cdot\boldsymbol{\nabla}B_{i_1i_2} + \beta\,B_{j_1j_2}B_{i_1i_2}\right)\\ &\qquad\qquad = \sum_{k_1=1}^{n_1}\sum_{k_2=1}^{n_2}\rho_{k_1k_2}\int_\Omega \mathrm{d}x\,\mathrm{d}y\,B_{k_1k_2}B_{i_1i_2} + \sum_{c=1}^{n_c} q_c B_{i_1i_2}(x_c, y_c)\,, \end{aligned} \tag{4.45}$$

for all $i_1 = 1,\dots,n_1-1$ and $i_2 = 1,\dots,n_2$. We now introduce the tensors

$$\begin{aligned} S_{i_1i_2j_1j_2} &:= \int_\Omega \mathrm{d}x\,\mathrm{d}y\;\left(\alpha\boldsymbol{\nabla}B_{j_1j_2}\cdot\boldsymbol{\nabla}B_{i_1i_2} + \beta\,B_{j_1j_2}B_{i_1i_2}\right)\\ &= \int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}|\;\left(\widehat{\alpha}\widehat{\boldsymbol{\nabla}}\widehat{B}_{j_1j_2}\cdot G^{-1}\cdot\widehat{\boldsymbol{\nabla}}\widehat{B}_{i_1i_2} + \widehat{\beta}\,\widehat{B}_{j_1j_2}\widehat{B}_{i_1i_2}\right)\,, \end{aligned} \tag{4.46a}$$

$$M_{i_1i_2k_1k_2} := \int_\Omega \mathrm{d}x\,\mathrm{d}y\,B_{k_1k_2}\,B_{i_1i_2} = \int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}|\,\widehat{B}_{k_1k_2}\,\widehat{B}_{i_1i_2}\,, \tag{4.46b}$$

where $\widehat{\boldsymbol{\nabla}}$ denotes the gradient in the logical domain, defined as $\widehat{\boldsymbol{\nabla}}\widehat{g} := \left(\dfrac{\partial\widehat{g}}{\partial s}, \dfrac{\partial\widehat{g}}{\partial\theta}\right)^T$ for any function $\widehat{g} \in \mathcal{C}^1(\widehat{\Omega})$, and G^{-1} denotes the inverse of the metric matrix of the logical coordinate system (defined by the metric coefficients (A.7)). Such integrals are computed using the Gauss-Legendre quadrature points and weights mentioned in section 4.3. We then obtain

$$\sum_{j_1=1}^{n_1-1}\sum_{j_2=1}^{n_2} S_{i_1i_2j_1j_2}\,\Phi_{j_1j_2} = \sum_{k_1=1}^{n_1}\sum_{k_2=1}^{n_2} M_{i_1i_2k_1k_2}\,\rho_{k_1k_2} + \sum_{c=1}^{n_c} q_c\widehat{B}_{i_1i_2}(s_c,\theta_c)\,, \tag{4.47}$$

for all $i_1 = 1,\dots,n_1-1$ and $i_2 = 1,\dots,n_2$. Here, the basis functions $\widehat{B}_{i_1i_2}$ in the last term are evaluated at the positions $(s_c,\theta_c) = \boldsymbol{F}^{-1}(x_c,y_c)$ of the point charges in the logical domain. We remark that, when the elliptic equation is coupled to the advection equation for ρ in the guiding-center model, $\boldsymbol{F}^{-1}(x_c,y_c)$ needs only to be computed at the beginning of a simulation: later on, the particle equations of motion are integrated using the pseudo-Cartesian coordinates (X_c,Y_c) and therefore $(s_c,\theta_c) = \boldsymbol{G}^{-1}(X_c,Y_c)$. Equation (4.47) can be written in matrix form as follows. Defining the new integer indices

$$\begin{aligned} i &:= (i_1-1)n_2 + i_2 &&= 1,\dots,(n_1-1)n_2 &&\text{(test space)}\\ j &:= (j_1-1)n_2 + j_2 &&= 1,\dots,(n_1-1)n_2 &&\text{(trial space)}\\ k &:= (k_1-1)n_2 + k_2 &&= 1,\dots,n_1n_2 &&\text{(space of } \rho_{\mathsf{SL}}) \end{aligned} \tag{4.48}$$

we can write (4.47) as

$$S\boldsymbol{\Phi} = M\boldsymbol{\rho}_{\mathsf{SL}} + \boldsymbol{\rho}_{\mathsf{PIC}}\,, \tag{4.49}$$

where we introduced the matrices S and M with elements $(S)_{ij} := S_{i_1i_2j_1j_2}$ and $(M)_{ik} := M_{i_1i_2k_1k_2}$, and the vectors $\boldsymbol{\Phi}$, $\boldsymbol{\rho}_{\mathsf{SL}}$ and $\boldsymbol{\rho}_{\mathsf{PIC}}$ with elements $(\boldsymbol{\Phi})_j := \Phi_{j_1j_2}$, $(\boldsymbol{\rho}_{\mathsf{SL}})_k := \rho_{k_1k_2}$ and $(\boldsymbol{\rho}_{\mathsf{PIC}})_i := \sum_{c=1}^{n_c} q_c\widehat{B}_{i_1i_2}(s_c,\theta_c)$. The $\mathcal{C}^1$ smoothness constraint is imposed by applying to the tensor-product spline basis of the test and trial spaces the restriction operator (using a notation similar to (Toshniwal et al., 2017, section 3.3))

$$E^T := \begin{pmatrix} \bar{E}^T & 0\\ 0 & I \end{pmatrix}\,, \tag{4.50}$$

where $\bar{E}$ contains the barycentric coordinates of the pole and of the first row of control points. More precisely, $\bar{E}$ is a $2n_2 \times 3$ matrix with elements $\bar{E}_{il} := \lambda_l(c^x_{i_1 i_2}, c^y_{i_1 i_2})$ and I is the identity matrix of size $[(n_1 - 3)n_2] \times [(n_1 - 3)n_2]$. Hence, the restriction operator E is a matrix of size $[(n_1 - 1)n_2] \times [3 + (n_1 - 3)n_2]$. Therefore, (4.49) becomes

$$\bar{S}\bar{\mathbf{\Phi}} = E^T(M\boldsymbol{\rho}_{\mathsf{SL}} + \boldsymbol{\rho}_{\mathsf{PIC}})\,, \tag{4.51}$$

where $\bar{S} := E^T S E$ and the solution vector $\bar{\mathbf{\Phi}}$ is of size $[3 + (n_1 - 3)n_2]$. The matrix $\bar{S}$ is symmetric and positive-definite, hence we can solve the linear system (4.51) with the conjugate gradient method (Hestenes and Stiefel, 1952; Quarteroni et al., 2006). The resulting solution is then prolonged back to the trial space via $\mathbf{\Phi} = E\bar{\mathbf{\Phi}}$.

4.5.3 Numerical results

We first test our elliptic solver on Poisson's equation

$$-\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\Phi = \rho\,, \tag{4.52}$$

with the method of manufactured solutions, looking for an exact solution of the form

$$\widehat{\Phi}_{\mathsf{ex}}(s,\theta) := (1 - s^2)\cos\left(2\pi\, x(s,\theta)\right)\sin\left(2\pi\, y(s,\theta)\right)\,, \tag{4.53}$$

on the physical domain defined by mapping (4.2). Denoting by $\Delta\Phi := \Phi - \Phi_{\mathsf{ex}}$ the numerical error, that is the difference between the numerical solution and the exact one, measures of the error are obtained by computing the spatial L^2-norm of $\Delta\Phi$,

$$||\Delta\Phi||_{L^2} := \sqrt{\int_\Omega \mathrm{d}x\,\mathrm{d}y\;[\Delta\Phi(x,y)]^2} = \sqrt{\int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}(s,\theta)|\,[\Delta\widehat{\Phi}(s,\theta)]^2}\,,$$

computed using the Gauss-Legendre quadrature points and weights mentioned in section 4.3, and the spatial L^∞-norm of $\Delta\Phi$,

$$||\Delta\Phi||_{L^\infty} := \max_{(x,y)\in\Omega} |\Delta\Phi(x,y)| = \max_{(s,\theta)\in\widehat{\Omega}} \left|\Delta\widehat{\Phi}(s,\theta)\right|\,, \tag{4.54}$$

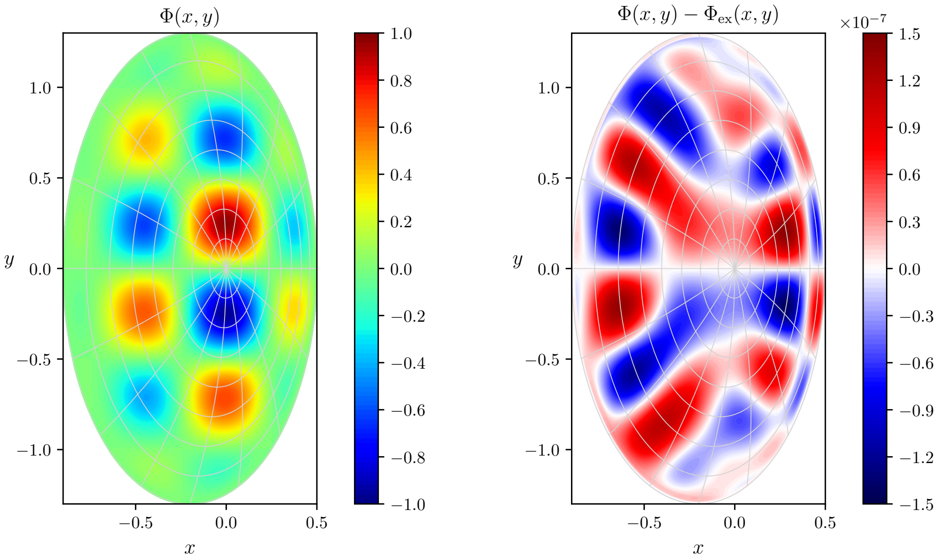


Figure 4.5: Numerical solution of Poisson's equation on a disk-like domain defined by mapping (4.2): contour plots of the numerical solution (left) and error (right), obtained with $n_1 \times n_2 = 128 \times 256$ and cubic splines.

computed on the Greville points. We remark again that the pole is included when we estimate the spatial L^∞-norm (4.54). Numerical results are shown in Figure 4.5. Table 4.3 shows the convergence of the solver while increasing the mesh size using cubic splines.

We then test our solver on the quasi-neutrality equation

$$-\boldsymbol{\nabla} \cdot (\alpha \boldsymbol{\nabla} \Phi) + \beta \Phi = \rho \,, \tag{4.55}$$

with θ-independent profiles $\alpha(s)$ and $\beta(s)$ defined as (Figure 4.6)

$$\alpha(s) := \exp\left[-\tanh\left(\frac{s-0.5}{0.1}\right)\right] , \qquad \beta(s) := \exp\left[\tanh\left(\frac{s-0.5}{0.2}\right)\right] ,$$

and look again for an exact solution of the form (4.53). Measures of the error are obtained, as before, by computing the spatial L^2-norm and L^∞-norm of the numerical error, $||\Delta\Phi||_{L^2}$ and $||\Delta\Phi||_{L^\infty}$. Numerical results are shown in Figure 4.7. Table 4.4 shows the convergence of the solver while increasing the mesh size using

$n_1 \times n_2$	$\|\|\Delta\Phi\|\|_{L^2}$	Order	$\|\|\Delta\Phi\|\|_{L^\infty}$	Order
32×64	7.10×10^{-5}		4.17×10^{-5}	
64×128	3.87×10^{-6}	4.20	2.31×10^{-6}	4.17
128×256	2.33×10^{-7}	4.05	1.41×10^{-7}	4.03
256×512	1.44×10^{-8}	4.02	8.78×10^{-9}	4.01
512×1024	8.99×10^{-10}	4.00	5.48×10^{-10}	4.00

Table 4.3: Numerical solution of Poisson's equation on a disk-like domain defined by mapping (4.2): fourth-order convergence of the elliptic solver using cubic splines.

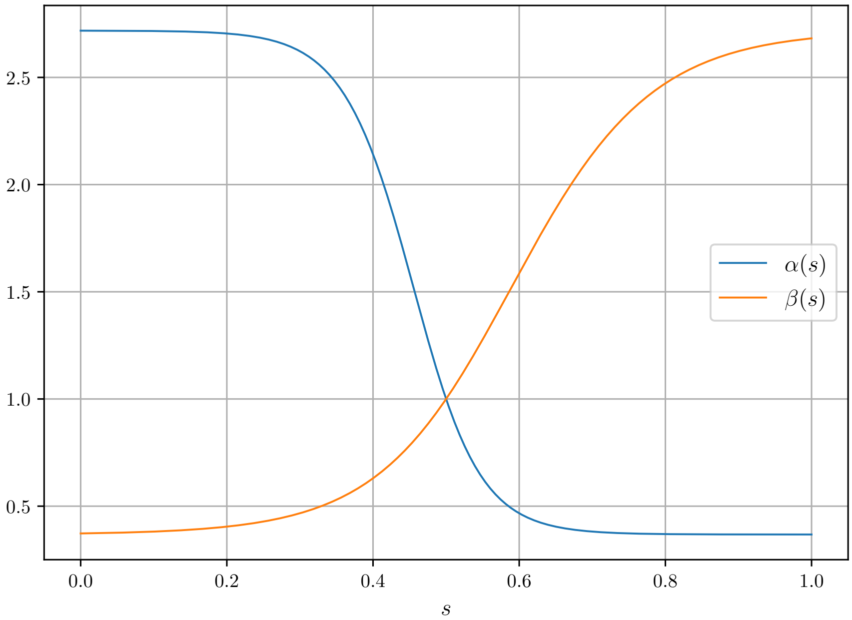


Figure 4.6: θ-independent profiles used in the quasi-neutrality equation (4.55).

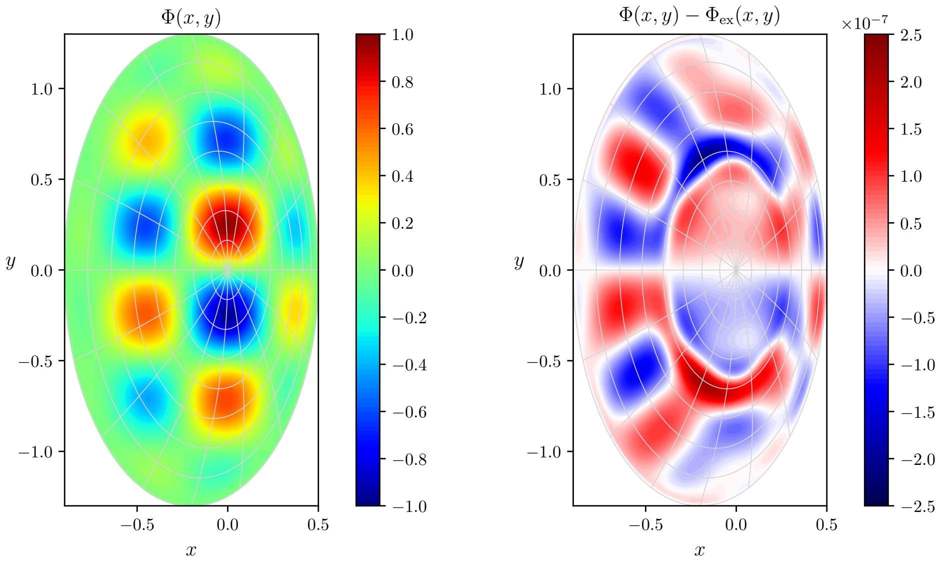


Figure 4.7: Numerical solution of the quasi-neutrality equation on a disk-like domain defined by mapping (4.2): contour plots of the numerical solution (left) and error (right), obtained with $n_1 \times n_2 = 128 \times 256$ and cubic splines.

cubic splines.

4.5.4 Evaluation of the electric field

The advection fields for the transport of ρ in (4.1) are obtained from the potential Φ by means of derivatives. Here we suggest a strategy to evaluate the Cartesian components of the gradient of Φ while taking into account the singularity at the pole. We denote again by $\widehat{\boldsymbol{\nabla}}\widehat{\Phi}$ the gradient of $\widehat{\Phi}$ in the logical domain. The Cartesian components of the gradient are obtained from the logical ones by applying the inverse of the transposed Jacobian matrix:

$$\widehat{\boldsymbol{\nabla}\Phi}(s,\theta) = (J_{\boldsymbol{F}}^{-1})^T(s,\theta)\,\widehat{\boldsymbol{\nabla}}\widehat{\Phi}(s,\theta)\,. \tag{4.56}$$

From an analytical point of view, (4.56) holds for all values of $s > 0$ and its limit as $s \to 0^+$ is finite and unique. From a numerical point of view, (4.56) holds for all values of s sufficiently far from the pole, as far as the inverse Jacobian does not

$n_1 \times n_2$	$\|\|\Delta\Phi\|\|_{L^2}$	Order	$\|\|\Delta\Phi\|\|_{L^\infty}$	Order
32×64	7.11×10^{-5}		8.36×10^{-5}	
64×128	3.82×10^{-6}	4.22	3.78×10^{-6}	4.47
128×256	2.29×10^{-7}	4.06	2.13×10^{-7}	4.15
256×512	1.41×10^{-8}	4.02	1.28×10^{-8}	4.06
512×1024	8.81×10^{-10}	4.00	7.87×10^{-10}	4.02

Table 4.4: Numerical solution of the quasi-neutrality equation on a disk-like domain defined by mapping (4.2): fourth-order convergence of the elliptic solver using cubic splines.

become too large. Therefore, we assume that (4.56) holds for $s \geqslant \epsilon$, for a given small ϵ. For $s = 0$ the partial derivative with respect to θ vanishes and all the information is contained in the partial derivative with respect to s, which takes a different value for each value of θ. Recalling that a partial derivative has the geometrical meaning of a directional derivative along a vector of the tangent basis, the idea is to combine two given values corresponding to two different values of θ and extract from them the Cartesian components of the gradient at the pole. The two chosen values of θ must correspond to linearly independent directions, so that from

$$\frac{\partial\widehat{\Phi}}{\partial s}(0,\theta_1) = \widehat{\boldsymbol{\nabla}\Phi} \cdot \boldsymbol{e}_s = (\widehat{\boldsymbol{\nabla}\Phi})_x \frac{\partial x}{\partial s}(0,\theta_1) + (\widehat{\boldsymbol{\nabla}\Phi})_y \frac{\partial y}{\partial s}(0,\theta_1)\,, \tag{4.57a}$$

$$\frac{\partial\widehat{\Phi}}{\partial s}(0,\theta_2) = \widehat{\boldsymbol{\nabla}\Phi} \cdot \boldsymbol{e}_s = (\widehat{\boldsymbol{\nabla}\Phi})_x \frac{\partial x}{\partial s}(0,\theta_2) + (\widehat{\boldsymbol{\nabla}\Phi})_y \frac{\partial y}{\partial s}(0,\theta_2)\,, \tag{4.57b}$$

the two components $(\widehat{\boldsymbol{\nabla}\Phi})_x$ and $(\widehat{\boldsymbol{\nabla}\Phi})_y$ can be obtained. Each possible couple of linearly independent directions produces the same result. In order to connect the two approaches in a smooth way, for $0 < s < \epsilon$ we interpolate linearly the value at the pole $s = 0$ and the value at $s = \epsilon$:

$$\widehat{\boldsymbol{\nabla}\Phi}(s,\theta) = \left(1 - \frac{s}{\epsilon}\right)\widehat{\boldsymbol{\nabla}\Phi}(0,\theta) + \frac{s}{\epsilon}\widehat{\boldsymbol{\nabla}\Phi}(\epsilon,\theta)\,. \tag{4.58}$$

We remark that the parameter ϵ can be chosen arbitrarily small, as far as it avoids underflows and overflows in floating point arithmetic. For all the numerical tests presented in this thesis we set $\epsilon = 10^{-12}$.

4.6 The 2D guiding-center model

We now address the solution of our 2D guiding-center model

$$\begin{cases} \dfrac{\partial \rho}{\partial t} - E^y \dfrac{\partial \rho}{\partial x} + E^x \dfrac{\partial \rho}{\partial y} = 0\,, \\ -\boldsymbol{\nabla} \cdot \boldsymbol{\nabla} \Phi = \rho\,, \end{cases} \quad \text{with} \quad \begin{cases} \rho(0,x,y) = \rho_{\mathsf{IN}}(x,y)\,, \\ \Phi(t,x,y) = 0 \text{ on } \partial\Omega\,. \end{cases} \tag{4.59}$$

Physical quantities conserved by the model are the total mass and energy

$$M(t) := \int_{\Omega} \mathrm{d}x\,\mathrm{d}y\,\rho(t,x,y) \quad = \int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}(s,\theta)|\,\widehat{\rho}(t,s,\theta)\,, \tag{4.60a}$$

$$W(t) := \int_{\Omega} \mathrm{d}x\,\mathrm{d}y\,|\boldsymbol{E}(t,x,y)|^2 = \int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}(s,\theta)|\,|\widehat{\boldsymbol{E}}(t,s,\theta)|^2\,. \tag{4.60b}$$

These integrals are computed using the Gauss-Legendre quadrature points and weights mentioned in section 4.3. The relative errors for the conservation of the invariants (4.60) are defined as

$$\delta M(t) := \frac{|M(0) - M(t)|}{|M(0)|}\,, \qquad \delta W(t) := \frac{|W(0) - W(t)|}{|W(0)|}\,. \tag{4.61}$$

Before describing the numerical tests considered for this model, we present our time-advancing strategy and how we deal with the problem of defining an equilibrium density on complex mappings while initializing our simulations.

4.6.1 Time integration

We present here two different time integration schemes, one explicit and one implicit, that may be chosen according to the particular physical dynamics described by model (4.59). Both integration schemes are based on a predictor-corrector procedure. In the numerical tests discussed in this section, the explicit scheme is our default choice, because of its low computational cost. However, there are situations (as, for example, the test case simulating the merger of two macroscopic vortices presented in section 4.6.3) where the dynamics described by model (4.59) is such that the explicit scheme would require very small time steps in order to produce correct results. Instead, the implicit trapezoidal scheme that we describe here has proven capable of capturing the correct dynamics with much larger time steps, thanks to its symmetry and adjoint-symplecticity (Hairer et al., 2012).

Second-order explicit scheme

The explicit time integration scheme is the second-order integrator described in (Xiong et al., 2018, section 2.2). Since it will be used also for test cases involving point charges, we denote here again by ρ_{SL} and ρ_{PIC} the semi-Lagrangian density and the particle density, respectively. Moreover, following the notation of section 4.4, we denote by $\boldsymbol{X}_{ij} := \boldsymbol{G}(\boldsymbol{\eta}_{ij})$ the pseudo-Cartesian coordinates of a given mesh point and by $\boldsymbol{X}_c := \boldsymbol{G}(\boldsymbol{\eta}_c)$ the pseudo-Cartesian coordinates of a given point charge, respectively. The first-order prediction (superscript "(P)") is given by

$$
\begin{aligned}
1.\quad \dot{\boldsymbol{X}}_{ij}^{(P)} &:= (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij})\, \widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}) \\
\boldsymbol{X}_{ij}^{(P)} &:= \boldsymbol{X}_{ij} - \Delta t\, \dot{\boldsymbol{X}}_{ij}^{(P)} \\
\boldsymbol{\eta}_{ij}^{(P)} &:= \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(P)})
\end{aligned}
\tag{4.62a}
$$

$$
\begin{aligned}
2.\quad \dot{\boldsymbol{X}}_{c}^{(P)} &:= (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{c})\, \widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{c}) \\
\boldsymbol{X}_{c}^{(P)} &:= \boldsymbol{X}_{c} + \Delta t\, \dot{\boldsymbol{X}}_{c}^{(P)} \\
\boldsymbol{\eta}_{c}^{(P)} &:= \boldsymbol{G}^{-1}(\boldsymbol{X}_{c}^{(P)})
\end{aligned}
\tag{4.62b}
$$

We then compute the intermediate semi-Lagrangian and particle densities $\rho_{\mathsf{SL}}^{(P)}$ and $\rho_{\mathsf{PIC}}^{(P)}$ and obtain the intermediate electric potential $\Phi^{(P)}$ by solving Poisson's equation. Denoting by $\widehat{\boldsymbol{U}}^{(P)}$ the corresponding intermediate advection field, the second-order correction (superscript "(C)") is given by

$$
\begin{aligned}
1.\quad \dot{\boldsymbol{X}}_{ij}^{(C)} &:= (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij}^{(P)})\,\widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}^{(P)}) + (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij})\,\widehat{\boldsymbol{U}}^{(P)}(\boldsymbol{\eta}_{ij}) \\
\boldsymbol{X}_{ij}^{(C)} &:= \boldsymbol{X}_{ij} - \frac{\Delta t}{2}\dot{\boldsymbol{X}}_{ij}^{(C)} \\
\boldsymbol{\eta}_{ij}^{(C)} &:= \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(C)})
\end{aligned}
\tag{4.63a}
$$

$$
\begin{aligned}
2.\quad \dot{\boldsymbol{X}}_{c}^{(C)} &:= (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{c})\,\widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{c}) + (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{c}^{(P)})\,\widehat{\boldsymbol{U}}^{(P)}(\boldsymbol{\eta}_{c}^{(P)}) \\
\boldsymbol{X}_{c}^{(C)} &:= \boldsymbol{X}_{c} + \frac{\Delta t}{2}\dot{\boldsymbol{X}}_{c}^{(C)} \\
\boldsymbol{\eta}_{c}^{(C)} &:= \boldsymbol{G}^{-1}(\boldsymbol{X}_{c}^{(C)})
\end{aligned}
\tag{4.63b}
$$

For point charges, this second-order scheme is equivalent to Heun's method (improved Euler's method (Süli and Mayers, 2003)).

Second-order implicit scheme

The implicit time integration scheme is based on the implicit trapezoidal rule and it will not be used for test cases involving point charges. We denote again by ρ_{SL} the semi-Lagrangian density and by $\boldsymbol{X}_{ij} := \boldsymbol{G}(\boldsymbol{\eta}_{ij})$ the pseudo-Cartesian coordinates of a given mesh point. The second-order prediction (superscript "(P)") is given by $\boldsymbol{X}_{ij}^{(P)} := \boldsymbol{X}_{ij}^{(k)}$ and $\boldsymbol{\eta}_{ij}^{(P)} := \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(P)})$, where the k-th iteration is computed as

$$
\begin{aligned}
\dot{\boldsymbol{X}}_{ij}^{(k)} &:= (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij})\,\widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}) + (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij}^{(k-1)})\,\widehat{\boldsymbol{U}}(\boldsymbol{\eta}_{ij}^{(k-1)}) \\
\boldsymbol{X}_{ij}^{(k)} &:= \boldsymbol{X}_{ij} - \frac{\Delta t}{4}\dot{\boldsymbol{X}}_{ij}^{(k)} \\
\boldsymbol{\eta}_{ij}^{(k)} &:= \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(k)})
\end{aligned}
$$

with $\boldsymbol{X}_{ij}^{(0)} := \boldsymbol{X}_{ij}$ and $\boldsymbol{\eta}_{ij}^{(0)} := \boldsymbol{\eta}_{ij}$, provided that $|\boldsymbol{X}_{ij}^{(k)} - \boldsymbol{X}_{ij}^{(k-1)}|^2 \leqslant \tau^2$, where the tolerance τ is defined as $\tau := \tau_A + \tau_R\,|\boldsymbol{X}_{ij}|$, for given absolute and relative tolerances τ_A and τ_R. We then compute the intermediate semi-Lagrangian density $\rho_{\mathsf{SL}}^{(P)}$ and obtain the intermediate electric potential $\Phi^{(P)}$ by solving Poisson's equation. Denoting by $\widehat{\boldsymbol{U}}^{(P)}$ the corresponding intermediate advection field, the second-order correction (superscript "(C)") is given by $\boldsymbol{X}_{ij}^{(C)} := \boldsymbol{X}_{ij}^{(k)}$ and $\boldsymbol{\eta}_{ij}^{(C)} := \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(C)})$, where the k-th iteration is computed as

$$\begin{aligned}
\dot{\boldsymbol{X}}_{ij}^{(k)} &:= (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij})\, \widehat{\boldsymbol{U}}^{(P)}(\boldsymbol{\eta}_{ij}) + (J_{\boldsymbol{F}} J_{\boldsymbol{G}}^{-1})^{-1}(\boldsymbol{\eta}_{ij}^{(k-1)})\, \widehat{\boldsymbol{U}}^{(P)}(\boldsymbol{\eta}_{ij}^{(k-1)}) \\
\boldsymbol{X}_{ij}^{(k)} &:= \boldsymbol{X}_{ij} - \frac{\Delta t}{2} \dot{\boldsymbol{X}}_{ij}^{(k)} \\
\boldsymbol{\eta}_{ij}^{(k)} &:= \boldsymbol{G}^{-1}(\boldsymbol{X}_{ij}^{(k)})
\end{aligned}$$

with $\boldsymbol{X}_{ij}^{(0)} := \boldsymbol{X}_{ij}$ and $\boldsymbol{\eta}_{ij}^{(0)} := \boldsymbol{\eta}_{ij}$, provided that $|\boldsymbol{X}_{ij}^{(k)} - \boldsymbol{X}_{ij}^{(k-1)}|^2 \leqslant \tau^2$.

4.6.2 Numerical equilibria

Defining an equilibrium density ρ and a corresponding equilibrium potential Φ for the system (4.59) becomes non-trivial on domains defined by complex non-circular mappings, such as (4.2) and (4.6). In the case of circular mappings, any axisymmetric density independent of the angle variable θ turns out to be an equilibrium for the transport equation in (4.59). For more complex mappings we follow the numerical procedure suggested by (Takeda and Tokuda, 1991), and references therein, to compute an equilibrium couple (ρ, Φ). The equilibrium is determined by the eigenvalue problem of finding (σ, Φ) such that $-\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\Phi = \sigma\, g(\Phi)$, with given g such that $g'(\Phi) \neq 0$ in some limited domain. Given initial data $(\sigma^{(0)}, \Phi^{(0)})$, the i-th iteration, with $i \geqslant 1$, is computed with the following steps:

1. compute $\rho^{(i)} := \sigma^{(i-1)} g(\Phi^{(i-1)})$;
2. compute $\Phi_*^{(i)}$ by solving $-\boldsymbol{\nabla} \cdot \boldsymbol{\nabla}\Phi_*^{(i)} = \rho^{(i)}$;
3. if a maximum value $\Phi_{\max}$ is given, compute $c^{(i)}$ by setting $c^{(i)} := \Phi_{\max}/||\Phi_*^{(i)}||_{L^\infty}$;
   if a maximum value $\rho_{\max}$ is given, compute $c^{(i)}$ by solving $c^{(i)} g(c^{(i)}\, ||\Phi_*^{(i)}||_{L^\infty}) = \dfrac{\rho_{\max}}{\sigma^{(i-1)}}$;
4. compute $(\sigma^{(i)}, \Phi^{(i)}) := c^{(i)}\, (\sigma^{(i-1)}, \Phi_*^{(i)})$.

The iterative procedure stops when $|\sigma^{(i)} - \sigma^{(i-1)}| \leqslant \tau$, for a given tolerance τ. The eigenvalue problem does not have a unique solution, but the algorithm is expected to converge to the ground state, that is the eigenstate with minimum eigenvalue. Figure 4.8 illustrates, for example, the equilibrium obtained in this way with $g(\Phi) = \Phi^2$

and $\rho_{\max} = 1$ on domains defined by a circular mapping and by the mappings (4.2) and (4.6).

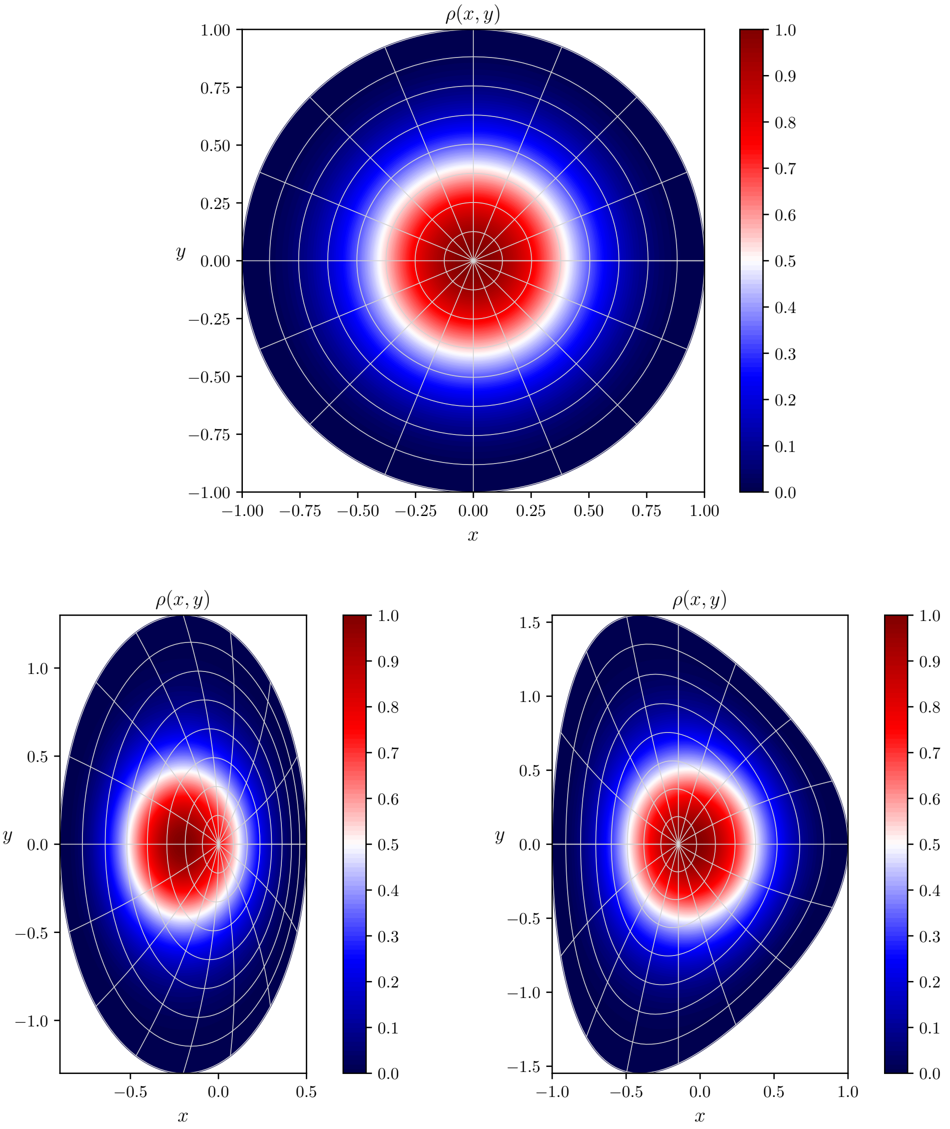


Figure 4.8: Numerical equilibrium density $\rho(x, y)$ obtained with $g(\Phi) = \Phi^2$ and $\rho_{\max} = 1$ on disk-like domains defined by a circular mapping (top) and by the mappings (4.2) (bottom left) and (4.6) (bottom right).

4.6.3 Numerical results

We present here various numerical tests performed on model (4.59), in order to verify the validity and correctness of the numerical schemes discussed in the previous sections.

Diocotron instability

As a first test we investigate the evolution of the diocotron instability on a domain defined by a circular mapping. From a physical point of view, this corresponds to studying a non-neutral plasma in cylindrical geometry, where the plasma particles are confined radially by a uniform axial magnetic field with a cylindrical conducting wall located at the outer boundary (Levy, 1965). Following (Davidson, 2001), we consider the initial density profile

$$\widehat{\rho}(0,s,\theta) := \widehat{\rho}_0(s) + \widehat{\rho}_1(0,s,\theta) := \begin{cases} 1+\epsilon\cos(m\theta) & s^- \leqslant s \leqslant s^+\,, \\ 0 & \text{elsewhere}\,. \end{cases} \tag{4.64}$$

This corresponds to a θ-independent equilibrium $\widehat{\rho}_0$ (an annular charged layer) with a density perturbation $\widehat{\rho}_1$ of azimuthal mode number m and small amplitude ϵ. The linear dispersion relation for a complex eigenfrequency ω reads (Davidson, 2001)

$$\left(\frac{\omega}{\omega_D}\right)^2 - b_m \frac{\omega}{\omega_D} + c_m = 0\,, \tag{4.65}$$

where ω_D is the diocotron frequency ($\omega_D = 1/2$ in our units), and b_m and c_m are defined as

$$b_m := m\left[1-\left(\frac{s^-}{s^+}\right)^2\right] + (s^+)^{2m} - (s^-)^{2m}\,, \tag{4.66a}$$

$$c_m := m\left[1-\left(\frac{s^-}{s^+}\right)^2\right]\left[1-(s^-)^{2m}\right] - \left[1-\left(\frac{s^-}{s^+}\right)^{2m}\right]\left[1-(s^+)^{2m}\right]\,. \tag{4.66b}$$

If $4c_m > b_m^2$, the oscillation frequencies resulting from (4.65) form complex conjugate pairs. The solution with $\operatorname{Im}\omega > 0$ corresponds to the diocotron instability and describes how rapidly the electric potential grows. The quantity of interest, in this regard, is the L^2-norm of the perturbed electric potential

$$\begin{aligned} ||\Phi - \Phi_0||_{L^2} &= \sqrt{\int_\Omega \mathrm{d}x\,\mathrm{d}y\ [\Phi(t,x,y) - \Phi_0(x,y)]^2} \\ &= \sqrt{\int_{\widehat{\Omega}} \mathrm{d}s\,\mathrm{d}\theta\,|\det J_{\boldsymbol{F}}(s,\theta)|\, \left[\widehat{\Phi}(t,s,\theta) - \widehat{\Phi}_0(s,\theta)\right]^2}\,, \end{aligned} \tag{4.67}$$

where Φ_0 denotes the equilibrium electric potential and the integration is performed again using the Gauss-Legendre quadrature points and weights. In order to represent the initial density in the finite-dimensional space of tensor-product splines, we modify (4.64) by a radial smoothing to avoid discontinuities:

$$\widehat{\rho}(0,s,\theta) := \begin{cases} [1+\epsilon\cos(m\theta)]\exp\left[-\left(\dfrac{s-\bar{s}}{d}\right)^p\right] & s^- \leqslant s \leqslant s^+\,, \\ 0 & \text{elsewhere}\,, \end{cases} \tag{4.68}$$

with $\bar{s} := (s^+ + s^-)/2$ and $d := (s^+ - s^-)/2$. If the smoothing layer is small enough, we can still rely on the analytical result obtained for the dispersion relation in the case of the sharp annular layer (4.64). The numerical results have been verified against the analytical dispersion relation for a perturbation with azimuthal mode number $m = 9$ and amplitude $\epsilon = 10^{-4}$. The numerical growth rate is in good agreement with the analytical one, $\operatorname{Im}\omega \approx 0.18$, for the time interval $20 \lesssim t \lesssim 50$, which corresponds to the linear phase. At time $t \approx 50$, the system enters its non-linear phase. The simulation is run with $n_1 \times n_2 = 128 \times 256$ and $\Delta t = 0.1$, with the explicit time integrator described in section 4.6.1. Additional parameters defining the initial condition (4.68) are set to $s^- = 0.45$, $s^+ = 0.50$ and $p = 50$. Numerical results are illustrated in Figures 4.9-4.12. For the conservation of mass and energy we get

$$\max_{t\in[0,70]} \delta M(t) \approx 5.8\times 10^{-4}\,, \qquad \max_{t\in[0,70]} \delta W(t) \approx 1.8\times 10^{-3}\,. \tag{4.69}$$

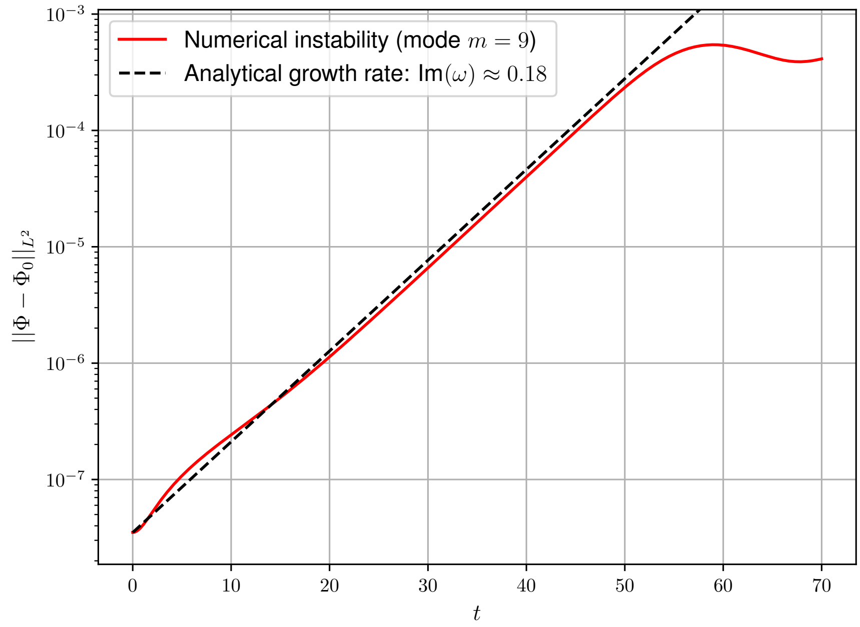


Figure 4.9: Numerical simulation of the diocotron instability: L^2-norm of the perturbed electric potential.

Figures 4.10-4.12 show that in this test case nothing significant happens in the region close to the pole. The effect of using $\mathcal{C}^1$ smooth polar splines in such situations is not particularly evident, but they do ensure continuity of the advection field responsible for the transport of ρ (the electric field) everywhere in the domain. Moreover, pseudo-Cartesian coordinates reduce to standard Cartesian coordinates, as the physical domain is defined by a simple circular mapping. The interest of this test case lies primarily in the fact that it provides the valuable possibility of easily verifying the implementation of our numerical scheme by comparing the numerical results with an analytical dispersion relation.

Vortex merger

In the context of incompressible inviscid 2D Euler fluids, we simulate the merger of two macroscopic vortices by setting up initial conditions qualitatively similar to those described in (Driscoll et al., 2002, section 3). Unlike the diocotron instability, the

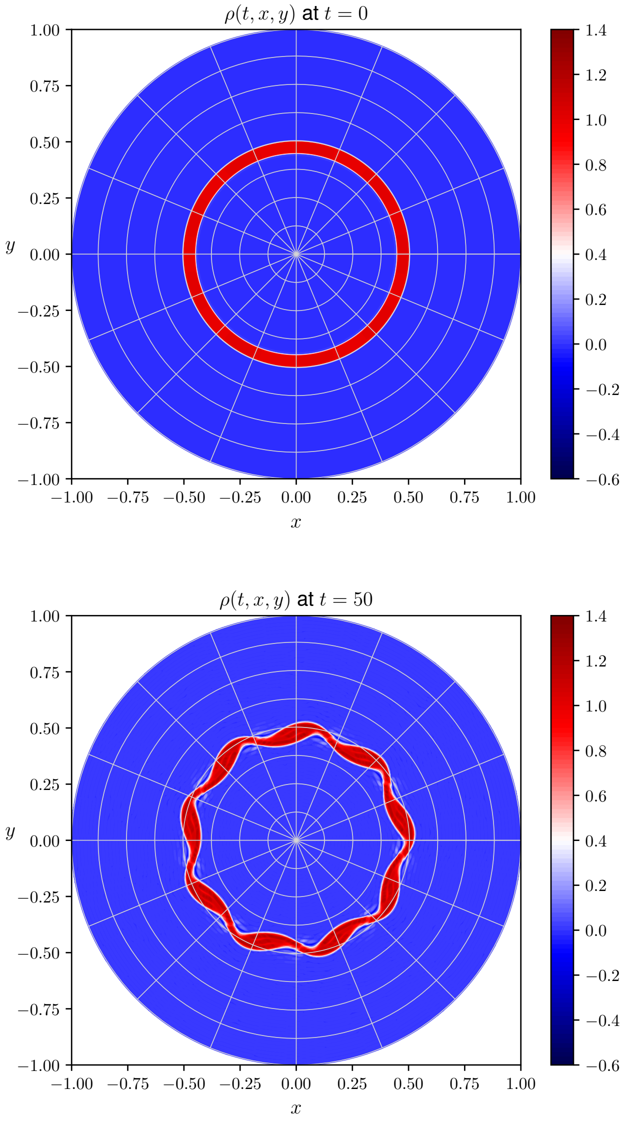


Figure 4.10: Numerical simulation of the diocotron instability: contour plots of the density $\rho(t, x, y)$ at times $t = 0$ (beginning of the simulation) and $t = 50$ (end of the linear phase).

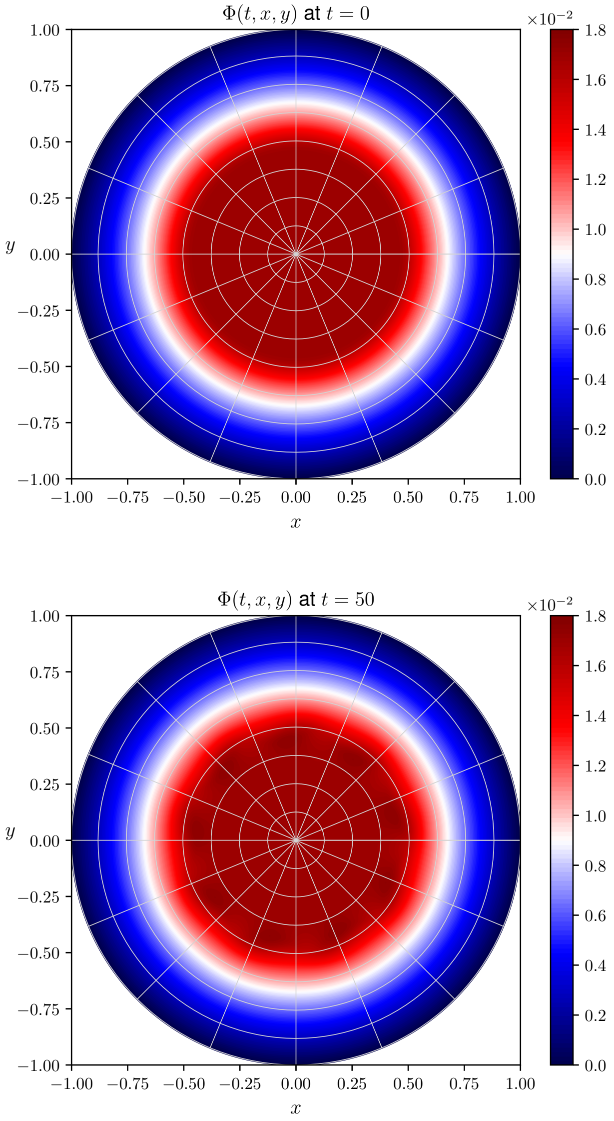


Figure 4.11: Numerical simulation of the diocotron instability: contour plots of the electric potential $\Phi(t, x, y)$ at times $t = 0$ (beginning of the simulation) and $t = 50$ (end of the linear phase).

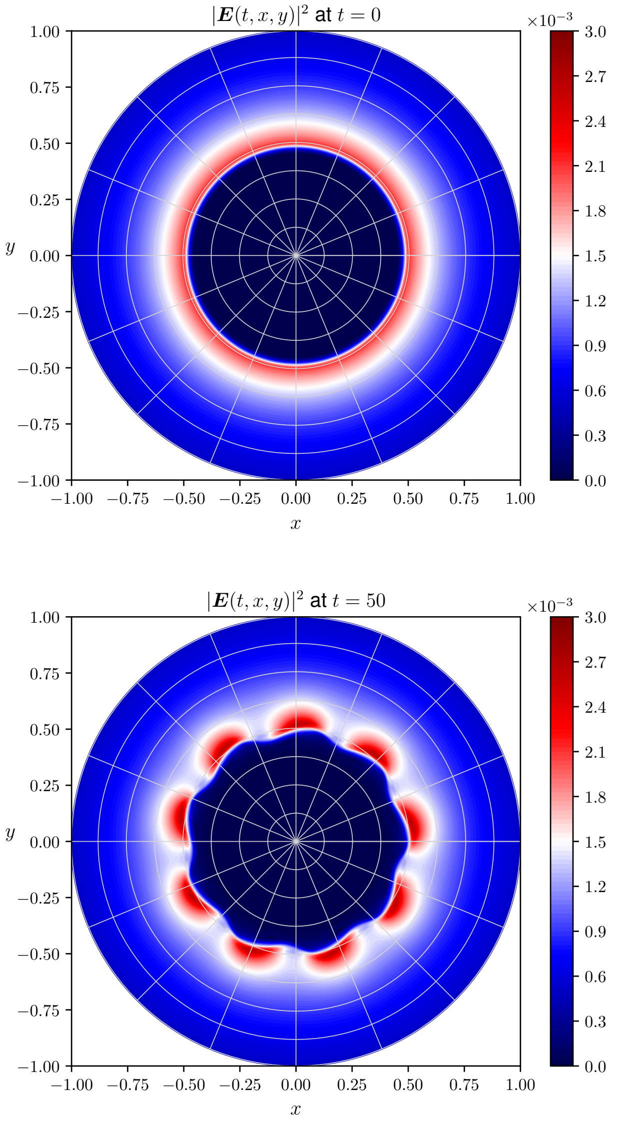


Figure 4.12: Numerical simulation of the diocotron instability: contour plots of the electric energy density $|\boldsymbol{E}(t,x,y)|^2$ at times $t=0$ (beginning of the simulation) and $t=50$ (end of the linear phase).

interest of this test case lies primarily in the fact that the relevant dynamics occurs in a region close the pole of the physical domain. We consider an equilibrium ρ_0 obtained with the numerical procedure described in section 4.6.2 with $g(\Phi) = \Phi^2$ and $\Phi_{\max} = 1$, and perturb it with two Gaussian perturbations,

$$\begin{aligned} \rho(0,x,y) &:= \rho_0(x,y) + \rho_1(0,x,y) \\ &:= \rho_0(x,y) + \epsilon\Bigg(\exp\left[-\frac{(x-x_1^*)^2 + (y-y_1^*)^2}{2\sigma^2}\right] \\ &\qquad + \exp\left[-\frac{(x-x_2^*)^2 + (y-y_2^*)^2}{2\sigma^2}\right]\Bigg)\,, \end{aligned} \tag{4.70}$$

with amplitude $\epsilon = 10^{-4}$, width $\sigma = 0.08$ and centered in $(x_1^*, y_1^*) = (+0.08, -0.14)$ and $(x_2^*, y_2^*) = (-0.08, +0.14)$, respectively. The time evolution of the initial perturbation ρ_1 is shown in Figures 4.13-4.14. The simulation is run with $n_1 \times n_2 = 128 \times 256$ and time step $\Delta t = 0.1$, with the second-order implicit time integrator described in section 4.6.1. The explicit time integrator would require in this case very small time steps in order to capture the correct dynamics. Two different aspects play a role in the choice of the time integrator for this particular test case. On the one hand, the error in the integration of the characteristics, which scales with Δt^2 for the second-order explicit scheme described in section 4.6.1, must not be larger than the amplitude of the perturbation on the advection field caused by the density perturbation. In other words, for the explicit scheme, the choice of the time step would be dependent on the amplitude ϵ of the density perturbation. On the other hand, committing an error in the integration of closed trajectories (as it would be when using the explicit scheme even for stationary advection fields) seems to disrupt the dynamics, preventing the simulation from correctly predicting the merger of the two macroscopic vortices. For the conservation of mass and energy we get

$$\max_{t\in[0,10]} \delta M(t) \approx 2.8 \times 10^{-9}\,, \qquad \max_{t\in[0,10]} \delta W(t) \approx 4.9 \times 10^{-9}\,. \tag{4.71}$$

The results of a convergence analysis of the numerical results while decreasing the time step are shown in Table 4.5, where $\Delta\rho$ denotes the difference between the vorticity ρ and a reference vorticity obtained by running a simulation with time step $\Delta t = 0.1/16$.

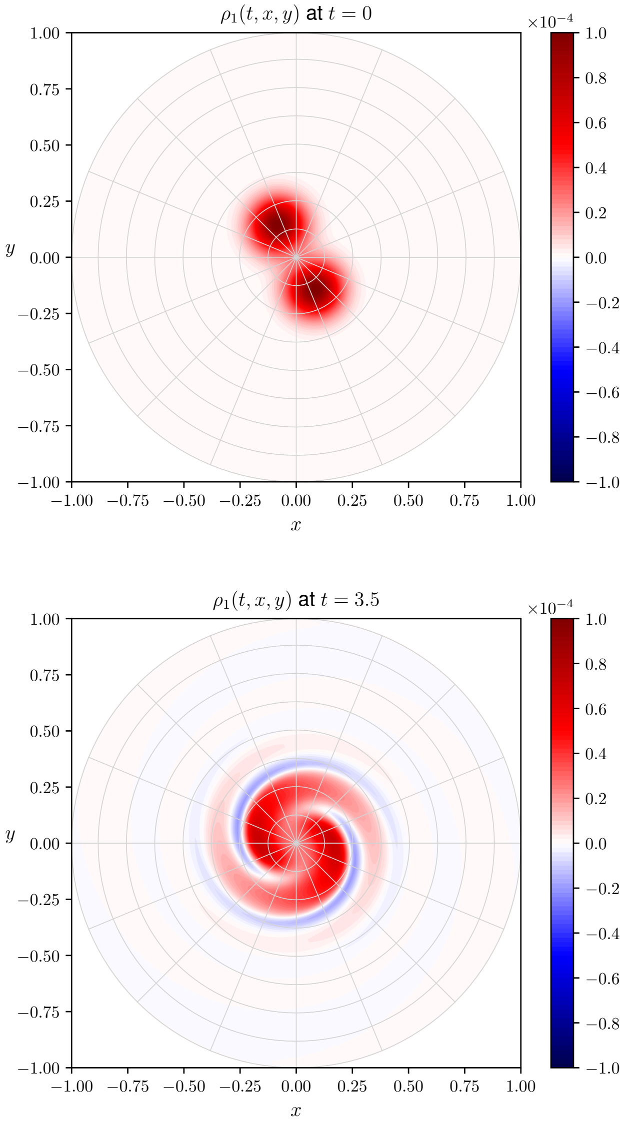


Figure 4.13: Numerical simulation of the merger of two vortices: contour plots of the density perturbation $\rho_1(t,x,y)$ at times $t=0$ and $t=3.5$.

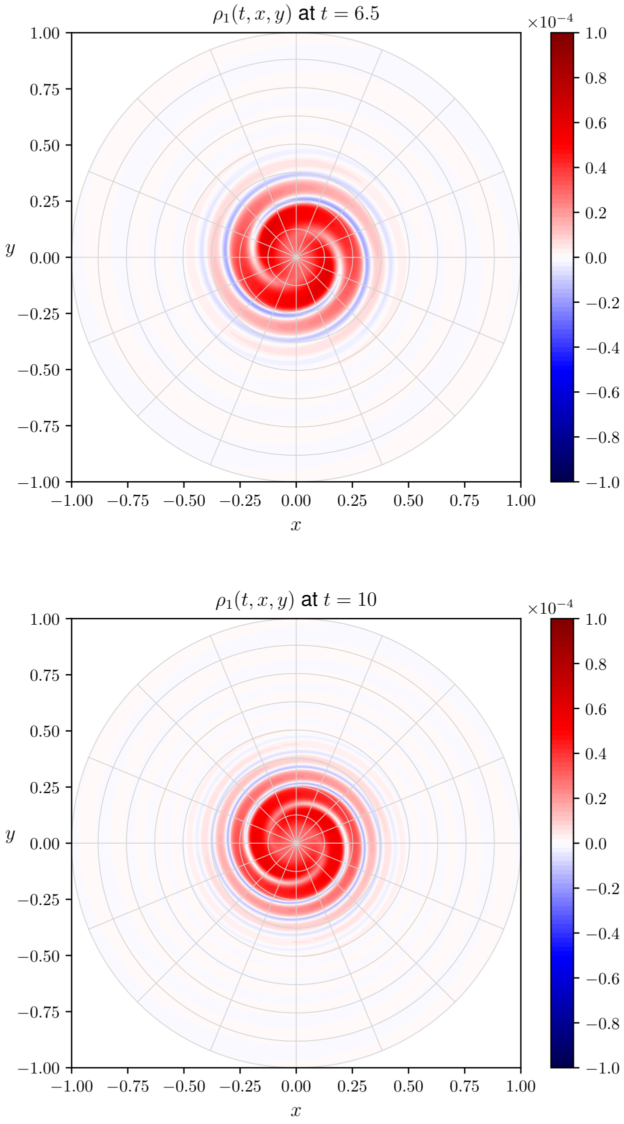


Figure 4.14: Numerical simulation of the merger of two vortices: contour plots of the density perturbation $\rho_1(t, x, y)$ at times $t = 6.5$ and $t = 10$.

Δt	$\|\|\Delta\rho\|\|_{L^\infty}$	Order	$\max_t \delta M(t)$	Order	$\max_t \delta W(t)$	Order
0.1	3.04×10^{-5}		2.84×10^{-9}		4.92×10^{-9}	
$0.1/2$	8.44×10^{-6}	1.85	1.42×10^{-9}	1.00	2.47×10^{-9}	0.99
$0.1/4$	2.05×10^{-6}	2.04	7.14×10^{-10}	0.99	1.22×10^{-9}	1.02
$0.1/8$	4.12×10^{-7}	2.32	3.46×10^{-10}	1.05	6.08×10^{-10}	1.01

Table 4.5: Convergence in time of the numerical results for the vortex merger with respect to reference results obtained with time step $\Delta t = 0.1/16$. The mesh size $n_1 \times n_2 = 128 \times 256$ is kept fixed in this convergence analysis.

Point-like vortex dynamics

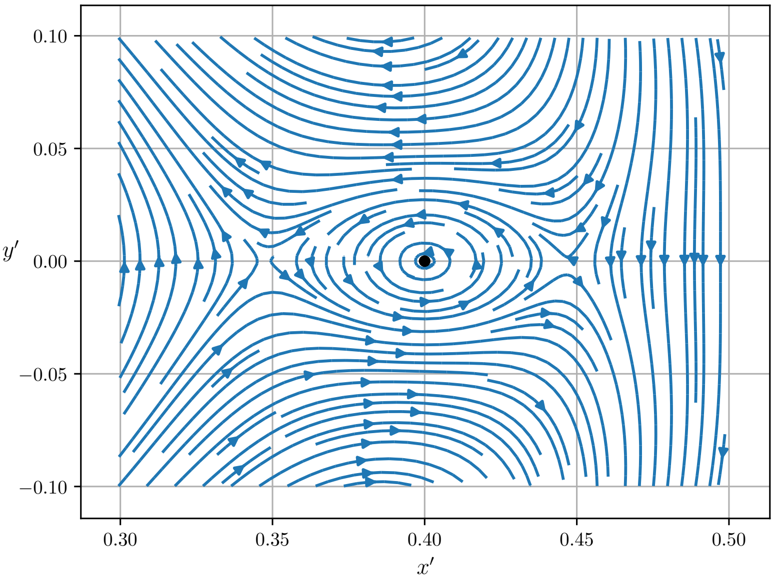
Again in the context of incompressible inviscid 2D Euler fluids, we also investigate the dynamics of point-like vortices (or point charges) on a non-uniform equilibrium, following the discussion in (Schecter and Dubin, 2001). The numerical tests presented in this section show that the numerical approaches suggested in this thesis can be applied straightforwardly in the context of particle-in-cell methods. This makes our numerical methods interesting also for numerical codes based on such methods, as for example the gyrokinetic PIC codes GTC (Ethier et al., 2005), GTS (Wang et al., 2006), XGC1 (Ku et al., 2009), ORB5 (Bottino et al., 2010), and ELMFIRE (Heikkinen et al., 2008). The examples discussed here can be considered as limit cases of usual particle-in-cell simulations, as we will include only one single point-like vortex (or point charge) in the system. Since our strategy turns out to work well for this extreme scenario, we do not expect issues to appear when dealing with the usual case of large numbers of particles. The point-like vortex contributes to the total charge density as described in equation (4.42). Moreover, the position of the point-like vortex is evolved following the same advection fields $(-E^y, E^x)^T$ responsible for the transport of ρ. Integration in time is performed with the second-order explicit scheme described in section 4.6.1. For a domain defined by a circular mapping, we consider an equilibrium vorticity of the form

$$\widehat{\rho}_0(s) := \begin{cases} 1 - 1.25\,s & s \leqslant 0.8\,, \\ 0 & s > 0.8\,, \end{cases} \tag{4.72}$$

identical to the one considered in (Schecter and Dubin, 2001, section IV). Figure 4.15 shows the local stream lines of the advection field near positive and negative point-like vortices at the initial time $t = 0$ in a rotating frame where the point-like vortices are initially at rest. This is obtained by transforming given coordinates (x, y) at time t to the rotated coordinates

$$\begin{aligned} x' &:= x\cos(-\omega t) - y\sin(-\omega t)\,, \\ y' &:= x\sin(-\omega t) + y\cos(-\omega t)\,, \end{aligned} \tag{4.73}$$

and by transforming the advection field $(-E^y, E^x)^T$ to the rotated advection field $(E^y + \omega\, y, E^x - \omega\, x)^T$, where $\omega = 0.3332$ represents the angular velocity of the background at $t = 0$ and $s = 0.4$. Figures 4.16-4.17 show results for a point-like vortex of intensity $q = \pm 0.0025$ at the initial position $s = 0.4$ and $\theta = 0$, viewed in a rotating frame. Time is here normalized as $t' = 0.1668\,t$ (as in (Schecter and Dubin, 2001), where t' is denoted as T). The results shown in Figures 4.16-4.17 are in agreement with the ones shown in Figures 7a and 10a of (Schecter and Dubin, 2001). As explained in (Schecter and Dubin, 2001), positive point-like vortices drift transverse to the shear flow, up the background vorticity gradient, while negative point-like vortices drift down the gradient. Figure 4.18 shows, in this regard, the time evolution of the radial position of the vortices, in agreement with the results shown in Figures 7b and 10b of (Schecter and Dubin, 2001). The simulation is run with $n_1 \times n_2 = 256 \times 512$ and time step $\Delta t = 0.005$. The time step is chosen small enough to resolve the oscillations due to the self-force experienced by the point-like vortices. Moreover, the computational mesh needs to be finer than the previous test cases in order to capture correctly the complex nonlinear dynamics of the interaction between the point-like vortices and the background vorticity. For a rough comparison, the vortex-in-cell simulations discussed in (Schecter and Dubin, 2001) require as well a large computational rectangular grid of size 1025×1025. Special techniques may be used to reduce self-force effects on non-uniform meshes (or even unstructured



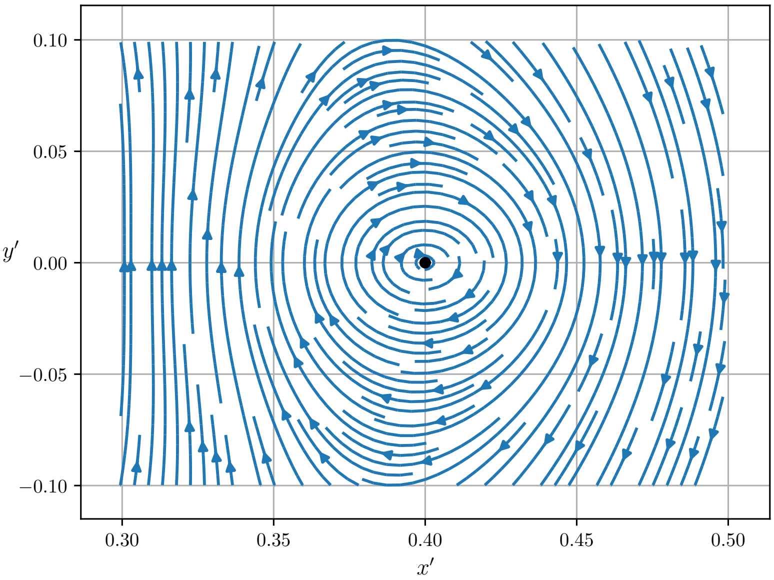


Figure 4.15: Local stream lines of the advection field near a positive (top) and negative (bottom) point-like vortex in a rotating frame (x', y').

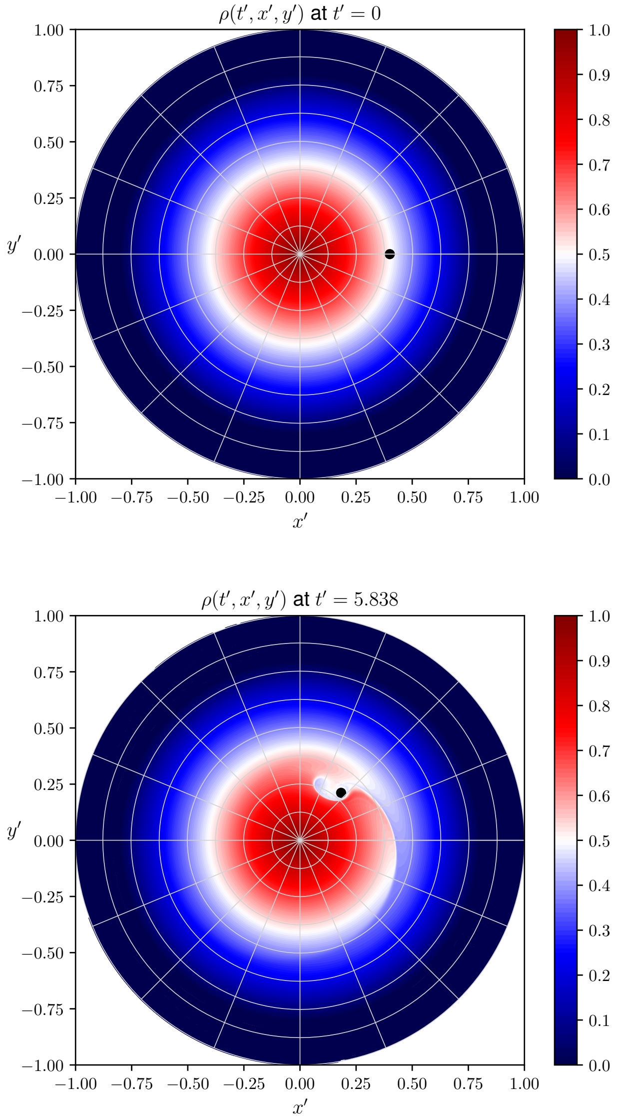


Figure 4.16: Dynamics of a positive point-like vortex on a domain defined by a circular mapping: contour plots of the density $\rho(t', x', y')$ at times $t' = 0$ and $t' = 5.838$ in a rotating frame (x', y').

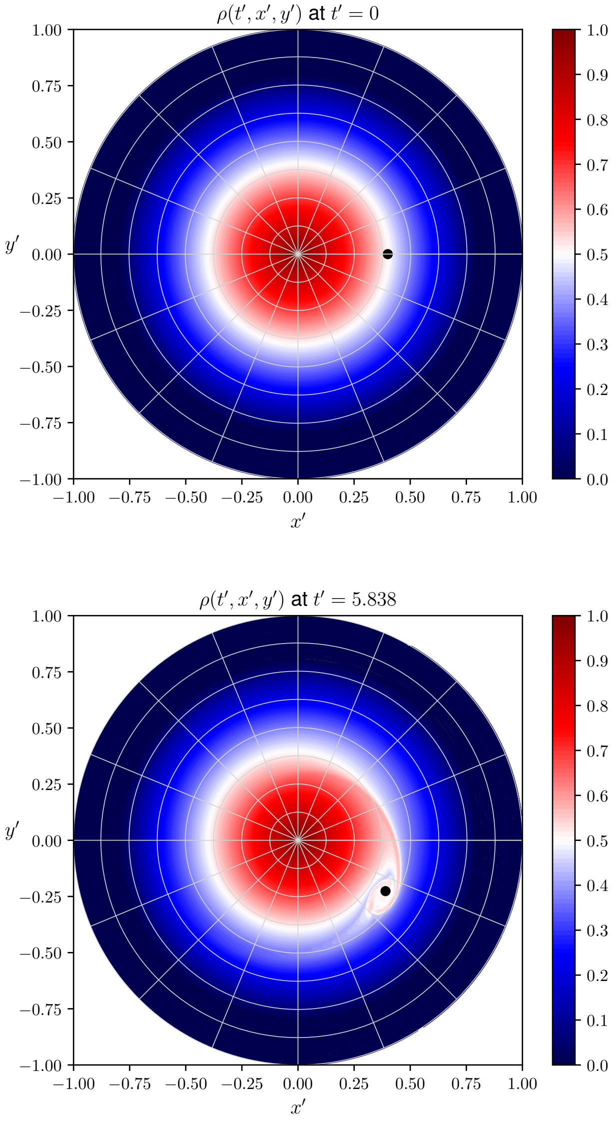


Figure 4.17: Dynamics of a negative point-like vortex on a domain defined by a circular mapping: contour plots of the density $\rho(t', x', y')$ at times $t' = 0$ and $t' = 5.838$ in a rotating frame (x', y').

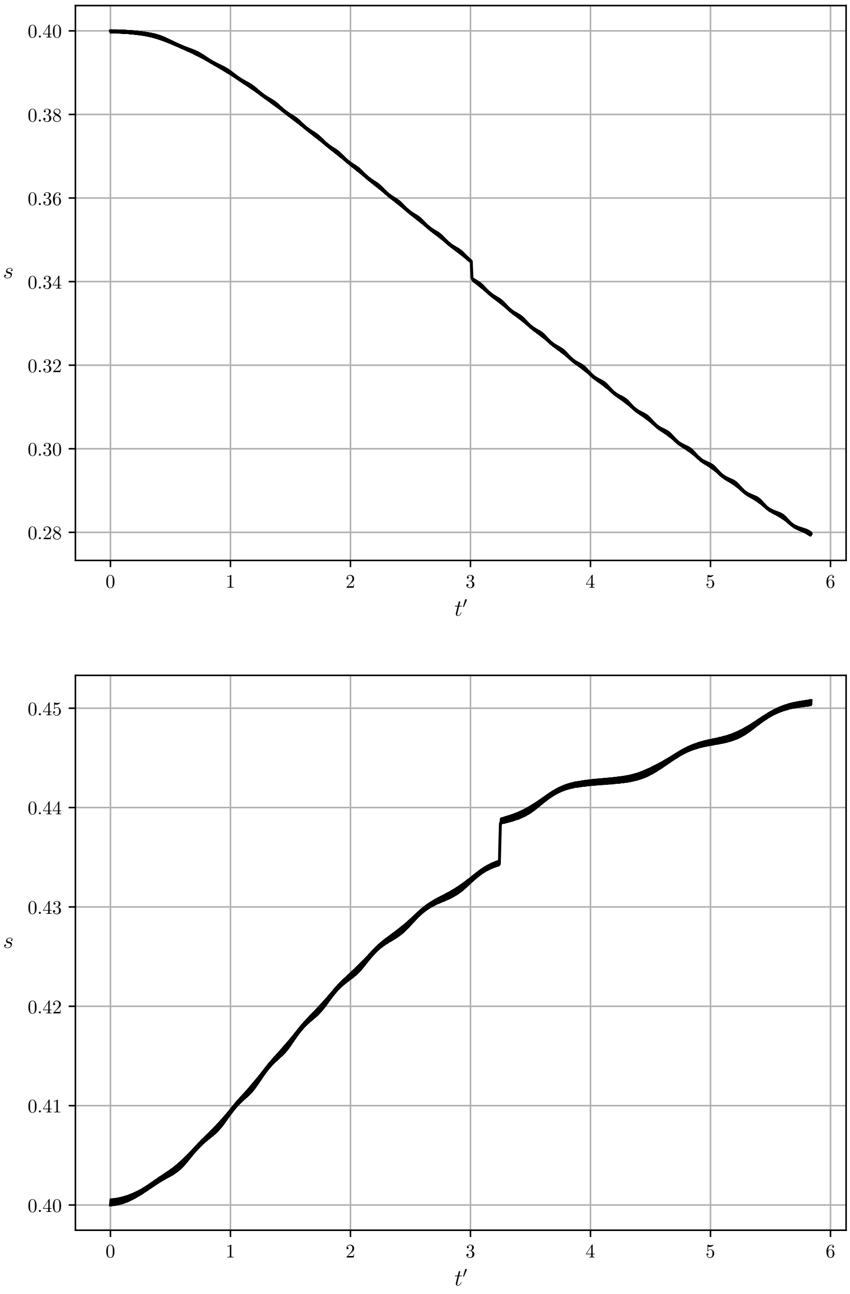


Figure 4.18: Time evolution of the radial position of a positive (top) and negative (bottom) point-like vortex.

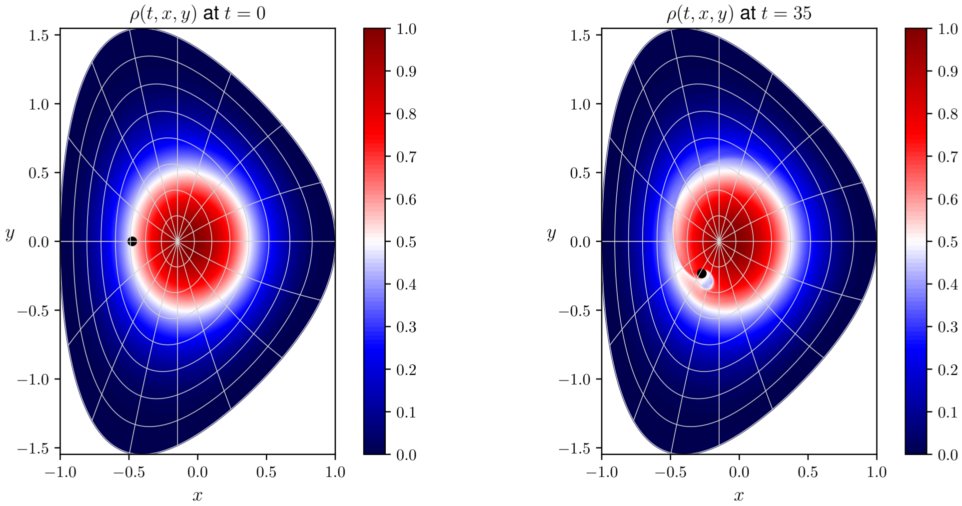


Figure 4.19: Dynamics of a positive point-like vortex on a disk-like domain defined by mapping (4.6): contour plots of the density $\rho(t,x,y)$ at times $t=0$ and $t=35$.

meshes), such as, for example, the ones suggested in (Bettencourt, 2014), but they have not been considered in this thesis. For the conservation of mass and energy we get

$$\max_{t'\in[0,5.838]} \delta M(t') \approx 6.9\times 10^{-6}\,, \qquad \max_{t'\in[0,5.838]} \delta W(t') \approx 8.4\times 10^{-3}\,, \tag{4.74}$$

for the positive point-like vortex and

$$\max_{t'\in[0,5.838]} \delta M(t') \approx 6.9\times 10^{-6}\,, \qquad \max_{t'\in[0,5.838]} \delta W(t') \approx 7.3\times 10^{-3}\,, \tag{4.75}$$

for the negative point-like vortex.

Similar simulations on a domain defined by mapping (4.6), initialized with an equilibrium vorticity obtained with the numerical procedure described in section 4.6.2 with $g(\Phi)=\Phi^2$ and $\rho_{\max}=1$, show the same qualitative behavior: a positive point-like vortex drifts towards the center of the domain, while a negative point-like vortex drifts towards the boundary (Figures 4.19-4.20). The final time $t=35$ corresponds to the normalized time $t'=5.838$ considered before. For the conservation of mass and

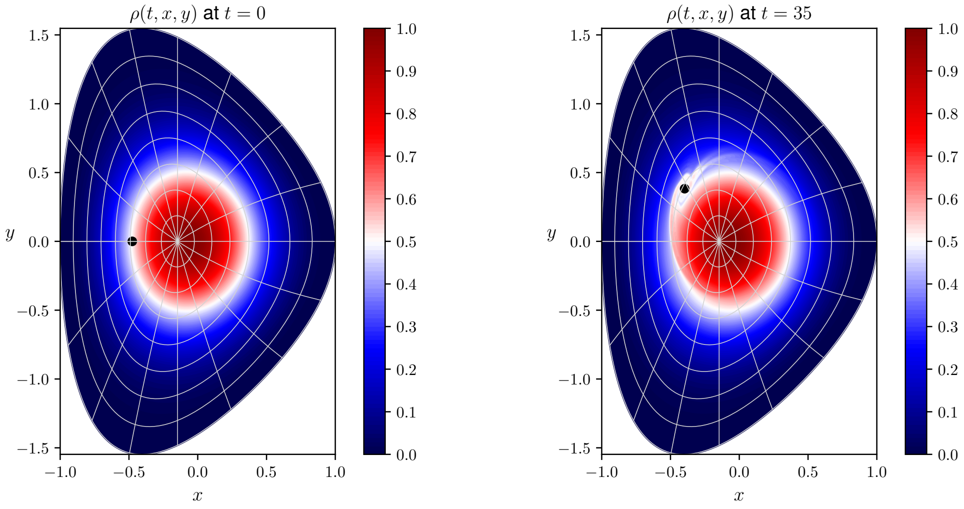


Figure 4.20: Dynamics of a negative point-like vortex on a disk-like domain defined by mapping (4.6): contour plots of the density $\rho(t,x,y)$ at times $t=0$ and $t=35$.

energy we get

$$\max_{t\in[0,35]} \delta M(t) \approx 1.6 \times 10^{-5}\,, \qquad \max_{t\in[0,35]} \delta W(t) \approx 6.9 \times 10^{-3}\,, \tag{4.76}$$

for the positive point-like vortex and

$$\max_{t\in[0,35]} \delta M(t) \approx 3.6 \times 10^{-6}\,, \qquad \max_{t\in[0,35]} \delta W(t) \approx 5.7 \times 10^{-3}\,, \tag{4.77}$$

for the negative point-like vortex.

4.7 Summary of the contributions

In this chapter we presented a comprehensive numerical strategy for the solution of systems of coupled hyperbolic-elliptic partial differential equations on singular mapped disk-like domains. We introduced a novel set of coordinates, named pseudo-Cartesian coordinates, for the integration of the characteristics of the hyperbolic equation of the system. Such coordinates are well-defined everywhere in the computational domain, including the pole, and provide a straightforward and relatively simple solution for dealing with singularities when solving advection problems in complex geometries. They reduce to standard Cartesian coordinates in the case of a circular mapping. Moreover, we developed a finite element elliptic solver based on globally $\mathcal{C}^1$ smooth splines (Toshniwal et al., 2017). In this thesis we considered only $\mathcal{C}^1$ smoothness, but higher-order smoothness, consistent with the spline degree, may be considered as well. We tested our solvers on several test cases in the simplest case of a circular domain and in more complex geometries. The numerical methods presented here show high-order convergence in the space discretization parameters, uniformly across the computational domain, including the pole. Moreover, the techniques discussed in this chapter can be easily applied in the context of particle-in-cell methods and are not necessarily restricted to semi-Lagrangian schemes, which were here discussed in more detail. The range of physical problems that can be approached following the ideas presented in this chapter includes the study of turbulence in magnetized fusion plasmas by means of Vlasov-Poisson fully kinetic models as well as drift-kinetic and gyrokinetic models, and turbulence models for incompressible inviscid Euler fluids in the context of fluid dynamics.

The methods discussed here were developed with the primary aim of being applied later on within our 4D field-aligned semi-Lagrangian drift-kinetic code. How this is done constitutes part of the subjects covered by the next chapter.

All the mathematical and computational tools described in this chapter, except for background material, were derived and tested within the scope of this thesis. The code development was done entirely in the framework of the SeLaLib library. The material presented in this chapter is also described and discussed in a manuscript written in collaboration with Dr. Yaman Güçlü, which is currently under consideration for publication in an international peer-reviewed scientific journal.

Chapter 5

Field-aligned semi-Lagrangian 4D drift-kinetic simulations

In this chapter we consider a 4D drift-kinetic electrostatic model and discuss its numerical solution in two different geometries, namely a cylindrical geometry with a screw pinch magnetic field and a toroidal geometry with circular concentric poloidal cross sections, conformal to the circular concentric flux surfaces of the magnetic field. Our physical model describes the electromagnetic interaction of positive ions of a single atomic species with adiabatic electrons. The model is among the simplest that can be derived under specific assumptions from the full gyrokinetic Vlasov-Maxwell equations. The word “electrostatic” refers to the fact that we consider only electric perturbations (time-dependent dynamic fluctuations of the electric scalar potential), without treating magnetic perturbations (time-dependent dynamic fluctuations of the magnetic vector potential). We note that this meaning is in contrast to the usual meaning of the word “electrostatic” in the context of classical electrodynamics, where it describes physical processes where the electric field and potential do not vary in time. Despite these simplifications, the model is still complex enough to ensure that the numerical methods and algorithms developed within its scope can be easily employed also within simulation codes that address the solution of more complex gyrokinetic models.

Our primary aim is to show that it is quite straightforward to integrate the 2D numerical machinery developed in the previous chapter within our 4D field-aligned semi-Lagrangian drift-kinetic code. It is our hope that this will serve as a basis for the future implementation of the same machinery in state-of-the-art gyrokinetic codes, such as, for example, the semi-Lagrangian gyrokinetic code GySeLa (Grandgirard et al., 2006b). The numerical tests that we performed in order to verify our implementation should be considered as first simple tests and can be certainly improved and extended to cover more complex and realistic scenarios.

This chapter is organized as follows. Section 5.1 describes briefly our 4D drift-kinetic electrostatic model. Section 5.2 presents in more detail the numerical methods employed for the solution of each of the lower-dimensional advection problems obtained by applying the field-aligned splitting discussed in chapter 3. Section 5.3 defines the scalar quantities that we look at as diagnostic tools in order to verify our implementation and analyze the results of our numerical simulations. Section 5.4 describes our physical model in cylindrical geometry with a screw pinch magnetic field and presents the results obtained for the numerical simulation of an ITG instability, verified through a linear dispersion analysis. Section 5.5 describes our physical model in a toroidal geometry with circular concentric poloidal cross sections and present a first basic numerical test with the primary aim of verifying the conservation properties of the model. We also review how to define an appropriate kinetic equilibrium in such geometry and suggest a numerical strategy to deal with the problem of particles crossing the boundary of the computational domain, when solving the advection equation on poloidal planes. Finally, we end the chapter with a summary of our contributions.

5.1 The physical model

The target model that we want to solve is the 4D drift-kinetic electrostatic model

$$\begin{cases} \dfrac{\partial F}{\partial t} + \boldsymbol{U} \cdot \boldsymbol{\nabla} F + U_{P_\parallel} \dfrac{\partial F}{\partial P_\parallel} = 0\,, \\ -\boldsymbol{\nabla}_\perp \cdot \left(\dfrac{\rho_0}{B_0^2} \boldsymbol{\nabla}_\perp \Phi_1 \right) + \dfrac{n_\mathrm{e}}{T_\mathrm{e}} \Phi_1 = \rho_1\,. \end{cases} \tag{5.1}$$

The unknowns in (5.1) are the particle distribution function $F(t, \boldsymbol{X}, P_\parallel)$, defined on the 4D gyrocenter phase space described by the gyrocenter position and parallel momentum $\boldsymbol{X}$ and $P_\parallel$, and the perturbed electric scalar potential $\Phi_1(t, \boldsymbol{X})$, defined on the 3D gyrocenter configuration space. Boundary conditions for (5.1) are an initial distribution function $F(0, \boldsymbol{X}, P_\parallel)$ and homogeneous Dirichlet boundary conditions for Φ_1. Both F and Φ_1 depend on time and model (5.1) describes their self-consistent evolution in time. We sometimes refer to the first equation in (5.1) as the 4D drift-kinetic Vlasov equation (or simply the Vlasov equation, if the overall context is clear enough) and to the second equation in (5.1) as the quasi-neutrality equation, as it can be derived under the assumption of quasi-neutrality of the plasma from the Vlasov-Maxwell system (see, for example, (Tronko et al., 2016) for a comprehensive explanation of the quasi-neutrality approximation in the derivation of gyrokinetic models). The word "self-consistent" refers to the fact that the evolution in time of F is coupled to the evolution in time of Φ_1. More precisely, the evolution of Φ_1, computed by solving the quasi-neutrality equation in (5.1), is coupled to the evolution of F through the perturbed charge density $\rho_1(t, \boldsymbol{X})$, defined as

$$\rho_1 := \int_{-\infty}^{+\infty} \mathrm{d}P_\parallel \, B_\parallel^* \, (F - F_0) \,, \tag{5.2}$$

where $F_0(\boldsymbol{X}, P_\parallel)$ denotes the equilibrium distribution function, while the evolution of F, computed by solving the Vlasov equation in (5.1), is coupled to the evolution of Φ_1 through the advection fields $\boldsymbol{U}(t, \boldsymbol{X}, P_\parallel)$ and $U_{P_\parallel}(t, \boldsymbol{X}, P_\parallel)$, defined as

$$\boldsymbol{U} := \frac{1}{B_\parallel^*} \left(\boldsymbol{b}_0 \times \boldsymbol{\nabla} H + \frac{\partial H}{\partial P_\parallel} \boldsymbol{B}^* \right) \,, \tag{5.3a}$$

$$U_{P_\parallel} := -\frac{\boldsymbol{B}^*}{B_\parallel^*} \cdot \boldsymbol{\nabla} H \,, \tag{5.3b}$$

where the modified magnetic field $\boldsymbol{B}^*(\boldsymbol{X}, P_\parallel)$ is defined as

$$\boldsymbol{B}^* := \boldsymbol{B}_0 + P_\parallel \boldsymbol{\nabla} \times \boldsymbol{b}_0 \,, \tag{5.4}$$

its parallel component $B_\parallel^*(\boldsymbol{X}, P_\parallel)$ is defined as $B_\parallel^* := \boldsymbol{B}^* \cdot \boldsymbol{b}_0$, and the drift-kinetic electrostatic Hamiltonian $H(t, \boldsymbol{X}, P_\parallel)$ is defined as

$$H := \frac{P_\parallel^2}{2} + \Phi_1 \,. \tag{5.5}$$

We note that the perpendicular gradient $\boldsymbol{\nabla}_\perp$ in (5.1) denotes the gradient computed with respect to the coordinates spanning the poloidal planes of the 3D geometry considered. The notation is therefore slightly imprecise, as such gradient does not entirely lie on a plane perpendicular to the direction of the magnetic field. Model (5.1) has been considered, for example, in (Latu et al., 2018), in a context similar to the one of this thesis.

5.2 Numerical methods

In this chapter we denote the logical coordinates and domains by

$$(r, \theta, \varphi, P_\parallel) \in [0, a] \times [0, 2\pi) \times [0, 2\pi) \times [-\bar{P}_\parallel, \bar{P}_\parallel] \,, \tag{5.6}$$

where (r, θ, φ) refer to either cylindrical coordinates in cylindrical geometry or elementary toroidal coordinates in toroidal geometry, $a > 0$ denotes the maximum radius (called the minor radius in toroidal geometry, in order to distinguish it from the major radius R_0) and $\bar{P}_\parallel > 0$ denotes the maximum parallel momentum. We also denote the corresponding logical meshes by $\{r_i\}_{i=1}^{n_1}$, $\{\theta_j\}_{j=1}^{n_2}$, $\{\varphi_k\}_{k=1}^{n_3}$, and $\{P_{\parallel\ell}\}_{\ell=1}^{n_4}$, respectively.

According to the field-aligned splitting described in chapter 3, we split the 4D Vlasov equation in (5.1) into the following three separate lower-dimensional advection problems:

1. a 2D advection on poloidal planes, for constant values $\varphi = \varphi_k$ for all $k = 1, \dots, n_3$ and $P_\parallel = P_{\parallel\ell}$ for all $\ell = 1, \dots, n_4$:

$$\frac{\partial F}{\partial t} + \boldsymbol{U}_\mathrm{p} \cdot \boldsymbol{\nabla} F = 0 \,; \tag{5.7}$$

2. a 2D field-aligned advection on flux surfaces, for constant values $r = r_i$ for all $i = 1, \ldots, n_1$ and $P_\parallel = P_{\parallel \ell}$ for all $\ell = 1, \ldots, n_4$:

$$\frac{\partial F}{\partial t} + U_\mathrm{b} \boldsymbol{b}_0 \cdot \boldsymbol{\nabla} F = 0\,; \tag{5.8}$$

3. a 1D advection in momentum space, for constant values $(r, \theta, \varphi) = (r_i, \theta_j, \varphi_k)$ for all $i = 1, \ldots, n_1$, $j = 1, \ldots, n_2$ and $k = 1, \ldots, n_3$:

$$\frac{\partial F}{\partial t} + U_{P_\parallel} \frac{\partial F}{\partial P_\parallel} = 0\,. \tag{5.9}$$

We recall that the contravariant components $(U_\mathrm{p}^{\zeta^1}, U_\mathrm{p}^{\zeta^2})$ of the poloidal advection field $\boldsymbol{U}_\mathrm{p}$ with respect to the tangent basis $(\boldsymbol{e}_{\zeta^1}, \boldsymbol{e}_{\zeta^2}, \boldsymbol{e}_{\zeta^3})$ of the coordinate system defined by the coordinates $(\zeta^1, \zeta^2, \zeta^3)$ (Cartesian coordinates in cylindrical geometry or cylindrical coordinates in toroidal geometry) read

$$\begin{aligned} U_\mathrm{p}^{\zeta^1} = \frac{1}{B_\parallel^*} \frac{1}{b_0^{\zeta^3}} \Bigg[& \frac{1}{\det J_\zeta} \left(-\frac{\partial \Phi_1}{\partial \zeta^2} + (b_0)_{\zeta^2} \nabla_\parallel \Phi_1 \right) \\ & + P_\parallel^2 \left(b_0^{\zeta^3} (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^1} - b_0^{\zeta^1} (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^3} \right) \Bigg]\,, \end{aligned} \tag{5.10a}$$

$$\begin{aligned} U_\mathrm{p}^{\zeta^2} = \frac{1}{B_\parallel^*} \frac{1}{b_0^{\zeta^3}} \Bigg[& \frac{-1}{\det J_\zeta} \left(-\frac{\partial \Phi_1}{\partial \zeta^1} + (b_0)_{\zeta^1} \nabla_\parallel \Phi_1 \right) \\ & + P_\parallel^2 \left(b_0^{\zeta^3} (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^2} - b_0^{\zeta^2} (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^3} \right) \Bigg]\,, \end{aligned} \tag{5.10b}$$

the advection field U_b for the field-aligned advection on flux surfaces reads

$$\begin{aligned} U_\mathrm{b} = \frac{1}{B_\parallel^*} \frac{1}{b_0^{\zeta^3}} \Bigg[& \frac{1}{\det J_\zeta} \left((b_0)_{\zeta^1} \frac{\partial \Phi_1}{\partial \zeta^2} - (b_0)_{\zeta^2} \frac{\partial \Phi_1}{\partial \zeta^1} \right) \\ & + P_\parallel \left(B_0^{\zeta^3} + P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^{\zeta^3} \right) \Bigg]\,, \end{aligned} \tag{5.11}$$

and the advection field $U_{P_\parallel}$ for the advection in momentum space reads

$$\begin{aligned} U_{P_\parallel} = \frac{-1}{B_\parallel^*}\Bigg[& \left(B_0 + \frac{P_\parallel}{b_0^{\zeta^3}}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\nabla_\parallel\Phi_1 \\ & + \frac{P_\parallel}{b_0^{\zeta^3}}\left(b_0^{\zeta^3}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^1} - b_0^{\zeta^1}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\frac{\partial\Phi_1}{\partial\zeta^1} \\ & + \frac{P_\parallel}{b_0^{\zeta^3}}\left(b_0^{\zeta^3}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^2} - b_0^{\zeta^2}(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^{\zeta^3}\right)\frac{\partial\Phi_1}{\partial\zeta^2}\Bigg] , \end{aligned} \tag{5.12}$$

where we also inserted the explicit expression of the drift-kinetic electrostatic Hamiltonian (5.5). We now look at each separate advection problem and discuss in more detail the numerical methods employed for its solution.

We remark that in the following three paragraphs we do not write the explicit dependence of the various advection fields on all variables and we leave instead only the dependence on the variables involved in the particular advection described. In other words, the poloidal advection field $\boldsymbol{U}_{\rm p}$ is considered to be a function of the poloidal coordinates (r,θ) only, the advection field $U_{\rm b}$ for the field-aligned advection on flux surfaces is considered to be a function of the flux-surface coordinates (θ,φ) only, and the advection field $U_{P_\parallel}$ for the advection in momentum space is considered to be a function of the momentum coordinate $P_\parallel$ only.

The poloidal advection (5.7) is solved by following the strategy described in section 4.4 for the 2D guiding-center model. In this context we denote the pseudo-Cartesian coordinates by $\boldsymbol{\mathcal{X}} := (\mathcal{X},\mathcal{Y})$, in order to avoid confusion with the Cartesian coordinates $\boldsymbol{X}$ in the 3D gyrocenter configuration space. The pseudo-Cartesian coordinates $\boldsymbol{\mathcal{X}}$ are defined as

$$\begin{aligned} \mathcal{X}(r,\theta) &:= r\cos\theta\,, \\ \mathcal{Y}(r,\theta) &:= r\sin\theta\,. \end{aligned} \tag{5.13}$$

We denote again by $\boldsymbol{G}$ the mapping defined by $\boldsymbol{G}(r,\theta) := (\mathcal{X},\mathcal{Y})$, and by $J_{\boldsymbol{G}}$ its Jacobian, using the same notation of section 4.4. The characteristic equations for the poloidal advection (5.7) in pseudo-Cartesian coordinates read

$$\dot{\boldsymbol{\mathcal{X}}} = (JJ_{\boldsymbol{G}}^{-1})^{-1}\boldsymbol{U}_{\rm p}\,, \tag{5.14}$$

where J is the 2×2 upper-left block extracted from the Jacobian matrix of the coordinate transformation $(r, \theta, \varphi) \mapsto (\zeta^1, \zeta^2, \zeta^3)$,

$$J := \begin{bmatrix} \dfrac{\partial \zeta^1}{\partial r} & \dfrac{\partial \zeta^1}{\partial \theta} \\ \dfrac{\partial \zeta^2}{\partial r} & \dfrac{\partial \zeta^2}{\partial \theta} \end{bmatrix} . \tag{5.15}$$

We recall again that the coordinates $(\zeta^1, \zeta^2, \zeta^3)$ are chosen to be the Cartesian coordinates (x, y, z) in cylindrical geometry or the cylindrical coordinates (R, Z, φ) in toroidal geometry. The poloidal components of the electric field $\boldsymbol{E} = -\boldsymbol{\nabla}\Phi_1$ with respect to the tangent basis $(\boldsymbol{e}_{\zeta^1}, \boldsymbol{e}_{\zeta^2}, \boldsymbol{e}_{\zeta^3})$, appearing in $U_\mathrm{p}^{\zeta^1}$ and $U_\mathrm{p}^{\zeta^2}$, are computed by following the procedure described in section 4.5. We then integrate the characteristic equations (5.14) with an implicit 2D trapezoidal scheme. Suppose that we want to find the origin (r^*, θ^*) at time $t - \Delta t$ of the characteristic passing through the mesh point (r_i, θ_j) at time t. We denote by $\boldsymbol{\mathcal{X}}_{ij} := (\mathcal{X}_{ij}, \mathcal{Y}_{ij}) := \boldsymbol{G}(r_i, \theta_j)$ the pseudo-Cartesian coordinates corresponding to (r_i, θ_j). The origin (r^*, θ^*) is given by $(r^*, \theta^*) = (r_i^{(q)}, \theta_j^{(q)}) = \boldsymbol{G}^{-1}(\mathcal{X}_{ij}^{(q)}, \mathcal{Y}_{ij}^{(q)})$, where the q-th iteration is computed as

$$\begin{aligned} \dot{\boldsymbol{\mathcal{X}}}_{ij}^{(q)} &:= (J J_{\boldsymbol{G}}^{-1})^{-1}(r_i, \theta_j)\, \boldsymbol{U}_\mathrm{p}(r_i, \theta_j) + (J J_{\boldsymbol{G}}^{-1})^{-1}(r_i^{(q-1)}, \theta_j^{(q-1)})\, \boldsymbol{U}_\mathrm{p}(r_i^{(q-1)}, \theta_j^{(q-1)}) \\ \boldsymbol{\mathcal{X}}_{ij}^{(q)} &:= \boldsymbol{\mathcal{X}}_{ij} - \frac{\Delta t}{2} \dot{\boldsymbol{\mathcal{X}}}_{ij}^{(q)} \\ (r_i^{(q)}, \theta_j^{(q)}) &:= \boldsymbol{G}^{-1}(\boldsymbol{\mathcal{X}}_{ij}^{(q)}) \end{aligned}$$

with $\boldsymbol{\mathcal{X}}_{ij}^{(0)} := (\mathcal{X}_{ij}^{(0)}, \mathcal{Y}_{ij}^{(0)}) := (\mathcal{X}_{ij}, \mathcal{Y}_{ij})$ and $(r_i^{(0)}, \theta_j^{(0)}) := (r_i, \theta_j)$, provided that $|\boldsymbol{\mathcal{X}}_{ij}^{(q)} - \boldsymbol{\mathcal{X}}_{ij}^{(q-1)}|^2 \leqslant \tau^2$, where the tolerance τ is defined as $\tau := \tau_\mathrm{A} + \tau_\mathrm{R} |\boldsymbol{\mathcal{X}}_{ij}|$, for given absolute and relative tolerances τ_A and τ_R.

The field-aligned advection on flux surfaces (5.8) is solved as follows. Suppose that we want to find the origin (θ^*, φ^*) at time $t - \Delta t$ of the characteristic passing through the mesh point (θ_j, φ_k) at time t. The toroidal angle coordinate φ satisfies the ordinary differential equation

$$\frac{\mathrm{d}\varphi}{\mathrm{d}t} = U_\mathrm{b}\, b_0^\varphi = U^\varphi , \tag{5.16}$$

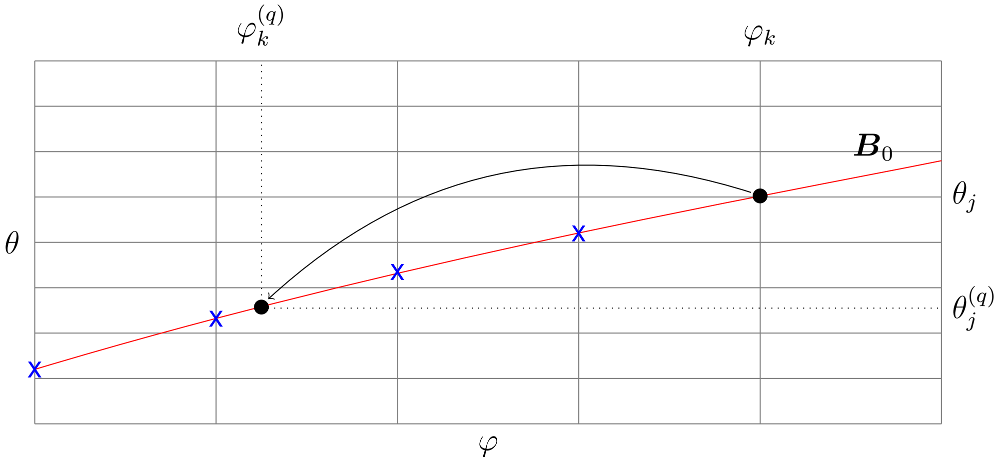


Figure 5.1: Field-aligned advection on a flux surface. The blue crosses represent the points $(\xi^{(q)}(\theta_j, \varphi_k, \ell), \varphi_\ell^{(q)})$ at the q-th iteration of the implicit trapezoidal scheme.

where b_0^φ represents the third contravariant component of $\boldsymbol{b}_0$ with respect to the tangent basis $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$ of the logical coordinate system. This scalar equation is solved with an implicit trapezoidal scheme. For $q \geqslant 1$, we denote by $\varphi_k^{(q)}$ the value of φ obtained at the q-th iteration and by $\{\varphi_\ell^{(q)}\}$ a set of grid values of φ in some neighborhood of $\varphi_k^{(q)}$. It is then possible to compute the values of the coordinate θ corresponding to the intersections between the magnetic field line passing through the starting point (θ_j, φ_k) and the lines at $\varphi = \varphi_\ell^{(q)}$. We denote such values of the coordinate θ as $\xi^{(q)}(\theta_j, \varphi_k, \ell)$, as in section 3.8. The values of F and U_b at these intersection points (blue crosses in Figure 5.1) are computed by spline interpolation along the direction θ. We also denote by $\mathcal{U}_\mathrm{b}^{(q)}$ the value obtained by interpolating with Lagrange polynomials the values of $U_\mathrm{b}\, b_0^\varphi$ at the intersection points. We then compute the value of φ at the q-th iteration as

$$\varphi_k^{(q)} := \varphi_k - \frac{\Delta t}{2} \left[U_\mathrm{b}(\theta_j, \varphi_k) b_0^\varphi(\theta_j, \varphi_k) + \mathcal{U}_\mathrm{b}^{(q-1)} \right] , \tag{5.17}$$

with $\varphi_k^{(0)} := \varphi_k$. We then set the origin of the characteristic to the point obtained at the q-th iteration, $\varphi^* = \varphi_k^{(q)}$, if $|\varphi_k^{(q)} - \varphi_k^{(q-1)}| \leqslant \tau$, for a given tolerance τ.

Finally, the advection in momentum space (5.9) is solved with an implicit 1D trapezoidal scheme. The origin $P_\parallel^*$ at time $t - \Delta t$ of the characteristic passing through the mesh point $P_{\|\ell}$ at time t is given by $P_{\|}^{*} = P_{\|\ell}^{(q)}$, where the q-th iteration is computed as

$$P_{\|\ell}^{(q)} := P_{\|\ell} - \frac{\Delta t}{2}\left[U_{P_{\|}}(P_{\|\ell}) + U_{P_{\|}}(P_{\|\ell}^{(q-1)})\right]\,, \tag{5.18}$$

with $P_{\|\ell}^{(0)} := P_{\|\ell}$, provided that $|P_{\|\ell}^{(q)} - P_{\|\ell}^{(q-1)}| \leqslant \tau$, for a given tolerance τ. If for any $q \geqslant 1$ we have either $P_{\|\ell}^{(q)} < -\bar{P}_{\|}$ or $P_{\|\ell}^{(q)} > \bar{P}_{\|}$, we set $P_{\|\ell}^{(q)} = -\bar{P}_{\|}$ or $P_{\|\ell}^{(q)} = \bar{P}_{\|}$, respectively.

5.3 Scalar diagnostics

The scalar diagnostics currently implemented in our 4D field-aligned semi-Lagrangian drift-kinetic code are the total mass (or total number of particles) and the L^2-norms of F and Φ_1:

$$M(t) := \int_0^a \mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi \int_{-\bar{P}_{\|}}^{\bar{P}_{\|}} \mathrm{d}P_{\|}\, |\det J_\eta|\, B_{\|}^{*}\, F\,, \tag{5.19a}$$

$$||F||_{L^2}(t) := \left[\int_0^a \mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi \int_{-\bar{P}_{\|}}^{\bar{P}_{\|}} \mathrm{d}P_{\|}\, |\det J_\eta|\, B_{\|}^{*}\, F^2\right]^{1/2}\,, \tag{5.19b}$$

$$||\Phi_1||_{L^2}(t) := \left[\int_0^a \mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \int_0^{2\pi} \mathrm{d}\varphi\, |\det J_\eta|\, \Phi_1^2\right]^{1/2}\,. \tag{5.19c}$$

These scalar quantities are computed numerically by performing the above integrations on the logical meshes. More precisely, the integration in r is performed with the weights

$$w_i^r := \frac{1}{2\Delta r}\begin{cases} r_{i+1} - r_i & i = 1\,, \\ r_{i+1} - r_{i-1} & i = 2, \dots, n_1 - 1\,, \\ r_i - r_{i-1} & i = n_1\,, \end{cases} \tag{5.20}$$

where $\Delta r := a/(n_1 - 1)$ (as for a uniform mesh); the integrations in θ and φ are performed with the trapezoidal rule on the respective periodic meshes (defining the cell spacings $\Delta\theta := 2\pi/n_2$ and $\Delta\varphi := 2\pi/n_3$); finally, the integration in $P_{\|}$ is performed with the weights

$$w_\ell^{P_\parallel} := \frac{1}{2\Delta P_\parallel} \begin{cases} P_{\parallel \ell+1} - P_{\parallel \ell} & \ell = 1\,, \\ P_{\parallel \ell+1} - P_{\parallel \ell-1} & \ell = 2, \dots, n_4 - 1\,, \\ P_{\parallel \ell} - P_{\parallel \ell-1} & \ell = n_4\,, \end{cases} \tag{5.21}$$

where $\Delta P_\parallel := 2\bar{P}_\parallel/(n_4 - 1)$ (as for a uniform mesh). Therefore, the total mass and the L^2-norms of F and Φ_1 are computed as

$$\begin{aligned} M(t) &= \sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{\ell=1}^{n_4} |\det J_\eta(r_i,\theta_j,\varphi_k)|\, B_\parallel^*(r_i,\theta_j,\varphi_k,P_{\parallel\ell})\, F(t,r_i,\theta_j,\varphi_k,P_{\parallel\ell}) \times \\ &\qquad \times w_i^r\, w_\ell^{P_\parallel}\, \Delta r\, \Delta\theta\, \Delta\varphi\, \Delta P_\parallel\,, \\ ||F||_{L^2}(t) &= \Bigg[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3}\sum_{\ell=1}^{n_4} |\det J_\eta(r_i,\theta_j,\varphi_k)|\, B_\parallel^*(r_i,\theta_j,\varphi_k,P_{\parallel\ell})\, F^2(t,r_i,\theta_j,\varphi_k,P_{\parallel\ell}) \times \\ &\qquad \times w_i^r\, w_\ell^{P_\parallel}\, \Delta r\, \Delta\theta\, \Delta\varphi\, \Delta P_\parallel \Bigg]^{1/2}\,, \\ ||\Phi_1||_{L^2}(t) &= \Bigg[\sum_{i=1}^{n_1}\sum_{j=1}^{n_2}\sum_{k=1}^{n_3} |\det J_\eta(r_i,\theta_j,\varphi_k)|\, \Phi_1^2(t,r_i,\theta_j,\varphi_k)\, w_i^r\, \Delta r\, \Delta\theta\, \Delta\varphi \Bigg]^{1/2}\,. \end{aligned}$$

The relative errors for the conservation of the total mass and the L^2-norm of F are defined as

$$\delta M(t) := \frac{|M(0) - M(t)|}{|M(0)|}\,, \qquad \delta||F||_{L^2}(t) := \frac{|\,||F||_{L^2}(0) - ||F||_{L^2}(t)|}{||F||_{L^2}(0)}\,. \tag{5.23}$$

The quadrature formulas presented in this section, especially regarding the integrations in the non-periodic domains along r and $P_\parallel$, could be improved by using appropriate Gauss-Legendre quadrature points and weights defined in each cell of the respective domains, consistently with the degrees of the spline basis functions along r and $P_\parallel$.

5.4 The 4D drift-kinetic model in cylindrical geometry

We now describe model (5.1) in cylindrical geometry with a screw pinch magnetic field. In this case we choose the coordinates $(\zeta^1, \zeta^2, \zeta^3)$ as the Cartesian coordinates (x, y, z), related to the logical coordinates (r, θ, φ) via the transformation

$$\begin{aligned} x(r,\theta,\varphi) &= r\cos\theta\,, \\ y(r,\theta,\varphi) &= r\sin\theta\,, \\ z(r,\theta,\varphi) &= R_0\,\varphi\,, \end{aligned} \tag{5.24}$$

where R_0 is related to the length L of the cylinder via $L = 2\pi R_0$. The contravariant components (B_0^x, B_0^y, B_0^z) of the magnetic field $\boldsymbol{B}_0$ with respect to the tangent basis $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ of the Cartesian coordinate system, expressed as functions of the logical coordinates (r, θ, φ), read

$$\begin{aligned} B_0^x(r,\theta) &:= -\bar{B}_0 \frac{r\sin\theta}{R_0\,q(r)}\,, \\ B_0^y(r,\theta) &:= \bar{B}_0 \frac{r\cos\theta}{R_0\,q(r)}\,, \\ B_0^z(r,\theta) &:= \bar{B}_0\,, \end{aligned} \tag{5.25}$$

where $\bar{B}_0$ is given and $q(r)$ denotes the safety factor profile. We note that $\boldsymbol{B}_0$ is axisymmetric and therefore does not depend on the toroidal angle coordinate φ. The magnitude B_0 of the magnetic field reads

$$B_0(r) = \sqrt{\frac{\bar{B}_0^2\,(R_0^2\,q^2(r) + r^2)}{R_0^2\,q^2(r)}}\,, \tag{5.26}$$

and depends only on r. The contravariant components (b_0^x, b_0^y, b_0^z) of the unit vector $\boldsymbol{b}_0 := \boldsymbol{B}_0/B_0$ with respect to the tangent basis $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ of the Cartesian coordinate system, expressed as functions of the logical coordinates (r, θ, φ), read

$$\begin{aligned} b_0^x(r,\theta) &:= \frac{B_0^x(r,\theta)}{B_0(r)} = -\frac{\bar{B}_0}{B_0(r)} \frac{r \sin\theta}{R_0\, q(r)}\,, \\ b_0^y(r,\theta) &:= \frac{B_0^y(r,\theta)}{B_0(r)} = \frac{\bar{B}_0}{B_0(r)} \frac{r \cos\theta}{R_0\, q(r)}\,, \\ b_0^z(r,\theta) &:= \frac{B_0^z(r,\theta)}{B_0(r)} = \frac{\bar{B}_0}{B_0(r)}\,. \end{aligned} \tag{5.27}$$

Moreover, the contravariant components of the curl of $\boldsymbol{b}_0$ with respect to the tangent basis $(\boldsymbol{e}_x, \boldsymbol{e}_y, \boldsymbol{e}_z)$ of the Cartesian coordinate system, expressed as functions of the logical coordinates (r, θ, φ), read

$$\begin{aligned} (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^x &:= \frac{\partial}{\partial y}(b_0)_z - \frac{\partial}{\partial z}(b_0)_y = \frac{q(r) - r\, q'(r)}{R_0^2\, q^2(r) + r^2} R_0\, b_0^x(r,\theta)\,, \\ (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^y &:= \frac{\partial}{\partial z}(b_0)_x - \frac{\partial}{\partial x}(b_0)_z = \frac{q(r) - r\, q'(r)}{R_0^2\, q^2(r) + r^2} R_0\, b_0^y(r,\theta)\,, \\ (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^z &:= \frac{\partial}{\partial x}(b_0)_y - \frac{\partial}{\partial y}(b_0)_x = \frac{r^2 + 2R_0^2\, q^2(r) - R_0^2\, r\, q(r)\, q'(r)}{(R_0^2\, q^2(r) + r^2) R_0\, q(r)}\, b_0^z(r,\theta)\,, \end{aligned}$$

where $q'(r) = \mathrm{d}q(r)/\mathrm{d}r$. Here we used the fact that $(b_0)_x = b_0^x$, $(b_0)_y = b_0^y$ and $(b_0)_z = b_0^z$, and expressed derivatives with respect to the Cartesian coordinates in terms of derivatives with respect to the logical coordinates as

$$\begin{aligned} \frac{\partial}{\partial x} &= \frac{\partial r}{\partial x}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial x}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial x}\frac{\partial}{\partial \varphi} = \cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\,, \\ \frac{\partial}{\partial y} &= \frac{\partial r}{\partial y}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial y}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial y}\frac{\partial}{\partial \varphi} = \sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\,, \\ \frac{\partial}{\partial z} &= \frac{\partial r}{\partial z}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial z}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial z}\frac{\partial}{\partial \varphi} = \frac{1}{R_0}\frac{\partial}{\partial \varphi}\,. \end{aligned} \tag{5.28}$$

The equation of a magnetic field line in logical coordinates reads

$$\frac{B_0^\theta}{\mathrm{d}\theta} = \frac{B_0^\varphi}{\mathrm{d}\varphi}\,, \tag{5.29}$$

which gives

$$\mathrm{d}\theta = \frac{\mathrm{d}\varphi}{q(r)}\,. \tag{5.30}$$

Here we used the fact that the contravariant components $(B_0^r, B_0^\theta, B_0^\varphi)$ of the magnetic field $\boldsymbol{B}_0$ with respect to the tangent basis $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$ of the logical coordinate system can be computed from the corresponding contravariant components (B_0^x, B_0^y, B_0^z) using (A.18):

$$\begin{aligned}
B_0^r &= \cos\theta\, B_0^x + \sin\theta\, B_0^y\,,\\
B_0^\theta &= -\frac{\sin\theta}{r}\, B_0^x + \frac{\cos\theta}{r}\, B_0^y\,,\\
B_0^\varphi &= \frac{B_0^z}{R_0}\,.
\end{aligned} \tag{5.31}$$

The integration of the right-hand side of (5.30) in the interval $[\varphi_0, \varphi]$ yields

$$\int_{\varphi_0}^{\varphi} \frac{\mathrm{d}\varphi'}{q(r)} = \frac{\varphi - \varphi_0}{q(r)}\,. \tag{5.32}$$

The integration of the left-hand side of (5.30) in the interval $[\theta_0, \theta]$ yields

$$\int_{\theta_0}^{\theta} \mathrm{d}\theta' = \theta - \theta_0\,. \tag{5.33}$$

Therefore, a magnetic field line passing through the point (r, θ_0, φ_0) can be described by the relation between θ and φ, which reads

$$\theta(\varphi) = \theta_0 + \frac{\varphi - \varphi_0}{q(r)}\,. \tag{5.34}$$

The resulting field lines, shown in Figure 5.2 on the flux surface at $r = a/2$, are simply straight lines, with a slope depending only on r.

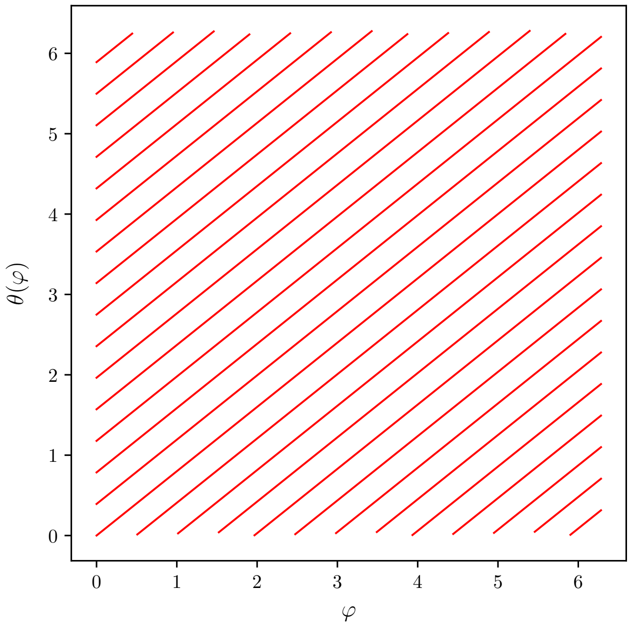


Figure 5.2: Magnetic field lines for the magnetic field (5.25) on the flux surface at $r = a/2$.

5.4.1 Linear dispersion analysis

We perform now a linear dispersion analysis of (5.1) and compute the corresponding dispersion relation. We split the distribution function F into an equilibrium part $F_0(r, P_\parallel)$ and a linear perturbation $F_1(t, r, \theta, \varphi, P_\parallel)$, of the same order as the perturbed electric potential Φ_1:

$$F = F_0 + F_1 \,. \tag{5.35}$$

The equilibrium distribution function F_0 is defined as

$$F_0 := \frac{1}{B_0} \frac{\rho_0}{\sqrt{2\pi T_\mathrm{i}}} \exp\left(-\frac{P_\parallel^2}{2T_\mathrm{i}}\right) \,, \tag{5.36}$$

where B_0, ρ_0 and T_i depend only on r, and it is such that

$$\int_{-\infty}^{+\infty} \mathrm{d}P_\parallel \, B_\parallel^* \, F_0 = \rho_0 \,. \tag{5.37}$$

We note that in this section we assume an infinite momentum domain $[-\infty, \infty]$, so that we can easily perform some analytical calculations. The idea is that the finite domain $[-\bar{P}_\parallel, \bar{P}_\parallel]$ is chosen in such a way that integrals over $[-\bar{P}_\parallel, \bar{P}_\parallel]$ approximate well enough the corresponding integrals over $[-\infty, \infty]$. Similarly to (5.35), we split the advection fields $\boldsymbol{U}$ and $U_{P_\parallel}$ as

$$\boldsymbol{U} = \boldsymbol{U}_0 + \boldsymbol{U}_1 \,, \qquad U_{P_\parallel} = U_{P_\parallel 0} + U_{P_\parallel 1} \,, \tag{5.38}$$

with $\boldsymbol{U}_0$, $\boldsymbol{U}_1$, $U_{P_\parallel 0}$ and $U_{P_\parallel 1}$ defined as

$$\boldsymbol{U}_0 := \frac{P_\parallel}{B_\parallel^*}\boldsymbol{B}^* \,, \qquad \boldsymbol{U}_1 := \frac{\boldsymbol{b}_0}{B_\parallel^*} \times \boldsymbol{\nabla}\Phi_1 \,, \tag{5.39a}$$

$$U_{P_\parallel 0} := 0 \,, \qquad U_{P_\parallel 1} := -\frac{\boldsymbol{B}^*}{B_\parallel^*} \cdot \boldsymbol{\nabla}\Phi_1 \,. \tag{5.39b}$$

Neglecting terms of order higher than linear, the linearized Vlasov equation reads

$$\frac{\partial F_1}{\partial t} + \boldsymbol{U}_0 \cdot \boldsymbol{\nabla} F_1 = -\boldsymbol{U}_0 \cdot \boldsymbol{\nabla} F_0 - \boldsymbol{U}_1 \cdot \boldsymbol{\nabla} F_0 - U_{P_\parallel 1}\frac{\partial F_0}{\partial P_\parallel} \,. \tag{5.40}$$

The corresponding (already) linearized quasi-neutrality equation reads

$$-\boldsymbol{\nabla}_\perp \cdot \left(\frac{\rho_0}{B_0^2}\boldsymbol{\nabla}_\perp \Phi_1\right) + \frac{n_{\mathrm{e}}}{T_{\mathrm{e}}}\,\Phi_1 = \int_{-\infty}^{+\infty} \mathrm{d}P_\parallel \, B_\parallel^* \, F_1 \,. \tag{5.41}$$

Introducing the Fourier coefficients

$$\widehat{F_1} := \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\varphi}{2\pi} F_1 e^{-im\theta - in\varphi} \,, \quad \widehat{\Phi_1} := \int_0^{2\pi} \frac{\mathrm{d}\theta}{2\pi} \int_0^{2\pi} \frac{\mathrm{d}\varphi}{2\pi} \Phi_1 e^{-im\theta - in\varphi} \,, \tag{5.42}$$

with $m \in \mathbb{Z}$ and $n \in \mathbb{Z}$, multiplying the linearized equations by $e^{-im\theta - in\varphi}$ and integrating with respect to θ and φ, we obtain

$$\frac{\partial \widehat{F_1}}{\partial t} + i\boldsymbol{k} \cdot \boldsymbol{U}_0 \, \widehat{F_1} = i\left(\frac{1}{B_\parallel^*}\frac{\partial F_0}{\partial r}(\boldsymbol{k} \times \boldsymbol{b}_0)^r + \frac{1}{B_\parallel^*}\frac{\partial F_0}{\partial P_\parallel}\boldsymbol{k} \cdot \boldsymbol{B}^*\right)\widehat{\Phi_1} \,, \tag{5.43}$$

where the vector $\boldsymbol{k}$ has the covariant components $(k_r, k_\theta, k_\varphi) := (0, m, n)$, and

$$-\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\rho_0}{B_0^2}\frac{\partial\widehat{\Phi_1}}{\partial r}\right) + \frac{m^2}{r^2}\frac{\rho_0}{B_0^2}\widehat{\Phi_1} + \frac{n_{\mathrm{e}}}{T_{\mathrm{e}}}\,\widehat{\Phi_1} = \int_{-\infty}^{+\infty} \mathrm{d}P_\parallel\, B_\parallel^*\,\widehat{F_1}\,. \tag{5.44}$$

Here we used $U_0^r = 0$, the fact that U_0^θ and U_0^φ depend only on r and not on θ and φ, and expressed the differential operator on the left-hand side of the quasi-neutrality equation in the curvilinear coordinates (r, θ) (recalling that $\boldsymbol{\nabla}_\perp$ denotes the gradient computed with respect to the coordinates spanning the poloidal planes). Introducing the Laplace transforms

$$\widetilde{F_1} := \int_0^\infty \mathrm{d}t\,\widehat{F_1}e^{i\omega t}\,, \qquad \widetilde{\Phi_1} := \int_0^\infty \mathrm{d}t\,\widehat{\Phi_1}e^{i\omega t}\,, \tag{5.45}$$

with $\omega \in \mathbb{C}$, multiplying the linearized equations by $e^{i\omega t}$ and integrating with respect to t, we obtain

$$i\left(-\omega + \boldsymbol{k}\cdot\boldsymbol{U}_0\right)\widetilde{F_1} = \widehat{F_1}_{\mathsf{IN}} + i\left(\frac{1}{B_\parallel^*}\frac{\partial F_0}{\partial r}(\boldsymbol{k}\times\boldsymbol{b}_0)^r + \frac{1}{B_\parallel^*}\frac{\partial F_0}{\partial P_\parallel}\boldsymbol{k}\cdot\boldsymbol{B}^*\right)\widetilde{\Phi_1}\,, \tag{5.46}$$

where $\widehat{F_1}_{\mathsf{IN}}$ denotes the value of $\widehat{F_1}$ at the initial time $t = 0$, and

$$-\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\rho_0}{B_0^2}\frac{\partial\widetilde{\Phi_1}}{\partial r}\right) + \frac{m^2}{r^2}\frac{\rho_0}{B_0^2}\widetilde{\Phi_1} + \frac{n_{\mathrm{e}}}{T_{\mathrm{e}}}\,\widetilde{\Phi_1} = \int_{-\infty}^{+\infty} \mathrm{d}P_\parallel\, B_\parallel^*\,\widetilde{F_1}\,. \tag{5.47}$$

Inserting $\widetilde{F_1}$ obtained from (5.46) into (5.47) we obtain the linear dispersion relation

$$D(\boldsymbol{k}, \omega)\widetilde{\Phi_1} = N(\boldsymbol{k}, \omega)\,, \tag{5.48}$$

where $N(\boldsymbol{k},\omega)$ and the differential operator $D(\boldsymbol{k},\omega)$ are defined as

$$N(\boldsymbol{k},\omega) := \int_{-\infty}^{+\infty} \mathrm{d}P_\parallel \, \frac{-i\, B_\parallel^* \, \widehat{F_{1\mathsf{IN}}}}{-\omega + \boldsymbol{k}\cdot\boldsymbol{U}_0} \,, \tag{5.49a}$$

$$\begin{aligned} D(\boldsymbol{k},\omega) := &-\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\rho_0}{B_0^2}\frac{\partial}{\partial r}\right) + \frac{m^2}{r^2}\frac{\rho_0}{B_0^2} + \frac{n_\mathrm{e}}{T_\mathrm{e}} \\ &- \int_{-\infty}^{+\infty} \mathrm{d}P_\parallel \, \frac{\dfrac{\partial F_0}{\partial r}(\boldsymbol{k}\times\boldsymbol{b}_0)^r + \dfrac{\partial F_0}{\partial P_\parallel}\boldsymbol{k}\cdot\boldsymbol{B}^*}{-\omega + \boldsymbol{k}\cdot\boldsymbol{U}_0} \,. \end{aligned} \tag{5.49b}$$

The idea is that the Fourier coefficient $\widehat{\Phi_1}$ can be obtained by inverting the Laplace transform (5.45) as

$$\widehat{\Phi_1} = \frac{1}{2i\pi}\int_{-\infty+iu}^{\infty+iu} \mathrm{d}\omega \, \widetilde{\Phi_1} e^{-i\omega t} \,, \tag{5.50}$$

where the integration is done along the line $\mathrm{Re}\,\omega = u$ in the complex plane and u is greater than the real part of all singularities of $\widetilde{\Phi_1}$. The integral (5.50) can be computed with the residue theorem by closing the contour by a half-circle in the lower complex half-plane and letting the radius of the half-circle go to infinity. If $\widetilde{\Phi_1}$ is analytical everywhere in the domain of integration except a finite number of poles $\{\omega_i\}_{i=1}^p$ and if the integral on the half-circle vanishes in the limit of infinite radius, we have

$$\widehat{\Phi_1} = \sum_{i=1}^{p} \mathrm{Res}(\widetilde{\Phi_1},\omega_i) e^{-i\omega_i t} \,. \tag{5.51}$$

We refer to (Ahlfors, 1979) for more details about path integrals in complex variables. From the dispersion relation (5.48) we conclude that in order to find the poles of $\widetilde{\Phi_1}$ it is necessary to find the values of ω such that $D(\boldsymbol{k},\omega) = 0$, for a given $\boldsymbol{k}$. Such zeros are computed with the mathematical software package ZEAL (Kravanja et al., 2000), integrated within the SeLaLib library. We also note that the integral

$$\int_{-\infty}^{+\infty} \mathrm{d}P_\parallel \, \frac{\dfrac{\partial F_0}{\partial r}(\boldsymbol{k}\times\boldsymbol{b}_0)^r + \dfrac{\partial F_0}{\partial P_\parallel}\boldsymbol{k}\cdot\boldsymbol{B}^*}{-\omega + \boldsymbol{k}\cdot\boldsymbol{U}_0} \,, \tag{5.52}$$

appearing in the differential operator $D(\boldsymbol{k},\omega)$, can be computed analytically and equals the integral $K(\alpha_0,\dots,\alpha_n,\beta_0,\beta_1,\beta_2,\gamma)$, defined in Appendix B, where $n=3$, $\gamma := 1/(2\,T_\mathrm{i})$, the coefficients β_0, β_1 and β_2 are defined as

$$\beta_0 := -\omega B_0\,, \tag{5.53a}$$

$$\beta_1 := \boldsymbol{k}\cdot\boldsymbol{B}_0 - \omega\left[(\boldsymbol{\nabla}\times\boldsymbol{b}_0)\cdot\boldsymbol{b}_0\right]\,, \tag{5.53b}$$

$$\beta_2 := \boldsymbol{k}\cdot(\boldsymbol{\nabla}\times\boldsymbol{b}_0)\,, \tag{5.53c}$$

and the coefficients α_0, α_1, α_2 and α_3 are defined as

$$\alpha_0 := B_0\frac{\partial}{\partial r}\left(\frac{1}{B_0}\frac{\rho_0}{\sqrt{2\pi T_\mathrm{i}}}\right)(\boldsymbol{k}\times\boldsymbol{b}_0)^r\,, \tag{5.54a}$$

$$\alpha_1 := \frac{1}{B_0}\left[(\boldsymbol{\nabla}\times\boldsymbol{b}_0)\cdot\boldsymbol{b}_0\right]\alpha_0 - \frac{1}{T_\mathrm{i}}\frac{\rho_0}{\sqrt{2\pi T_\mathrm{i}}}\boldsymbol{k}\cdot\boldsymbol{B}_0\,, \tag{5.54b}$$

$$\alpha_2 := \frac{1}{T_\mathrm{i}}\frac{\rho_0}{\sqrt{2\pi T_\mathrm{i}}}\left(\frac{1}{2}\frac{1}{T_\mathrm{i}}\frac{\partial T_\mathrm{i}}{\partial r}(\boldsymbol{k}\times\boldsymbol{b}_0)^r - \beta_2 - (\boldsymbol{k}\cdot\boldsymbol{b}_0)\left[(\boldsymbol{\nabla}\times\boldsymbol{b}_0)\cdot\boldsymbol{b}_0\right]\right)\,, \tag{5.54c}$$

$$\alpha_3 := \frac{1}{B_0}\left[(\boldsymbol{\nabla}\times\boldsymbol{b}_0)\cdot\boldsymbol{b}_0\right]\left(\alpha_2 + \frac{1}{T_\mathrm{i}}\frac{\rho_0}{\sqrt{2\pi T_\mathrm{i}}}(\boldsymbol{k}\cdot\boldsymbol{b}_0)\left[(\boldsymbol{\nabla}\times\boldsymbol{b}_0)\cdot\boldsymbol{b}_0\right]\right)\,. \tag{5.54d}$$

5.4.2 Numerical results

We present here the numerical results obtained from the simulation of an ITG instability. From a physical point of view, we are looking at a plasma instability driven by the gradient of the temperature profile of the ions. The instability exhibits a linear phase of exponential growth, that eventually reaches a saturation level, followed by a non-linear evolution. The parameters defining the cylindrical domain considered here are non-physical, as the size of the system is only about 10 times bigger than the ion Larmor radius. However, such geometry has the effect of amplifying the physical phenomena under study, thus making numerical issues related to our numerical scheme more visible and easier to fix. We therefore think that the geometrical setting considered here, despite being non-physical, is still highly valuable from the point of view of testing our numerical methods and algorithms. More precisely, the minor radius of the cylinder is $a = 14.5$ and its length is $L = 2\pi R_0$, with $R_0 = 72.5$. This corresponds to an aspect ratio $A = R_0/a = 5$, which is only slightly larger than the corresponding realistic physical values for toroidal devices ($A \approx 2.0$ for ASDEX Upgrade, $A \approx 2.4$ for JET, $A \approx 3.1$ for ITER). The density and temperature profiles of ions and electrons are defined as

$$\rho_0(r) = n_\mathrm{e}(r) := \bar{\rho}_0 \exp\left[- k_{\rho_0} w_{\rho_0} \frac{a}{R_0} \tanh\left(\frac{r - \bar{r}}{w_{\rho_0} a} \right) \right] , \tag{5.55a}$$

$$T_\mathrm{i}(r) := \overline{T}_\mathrm{i} \exp\left[- k_{T_\mathrm{i}} w_{T_\mathrm{i}} \frac{a}{R_0} \tanh\left(\frac{r - \bar{r}}{w_{T_\mathrm{i}} a} \right) \right] , \tag{5.55b}$$

$$T_\mathrm{e}(r) := \overline{T}_\mathrm{e} \exp\left[- k_{T_\mathrm{e}} w_{T_\mathrm{e}} \frac{a}{R_0} \tanh\left(\frac{r - \bar{r}}{w_{T_\mathrm{e}} a} \right) \right] , \tag{5.55c}$$

where $\bar{r} := a/2$ and the parameters are set as

$$\bar{\rho}_0 = 0.15 \,, \quad \bar{T}_\mathrm{i} = \bar{T}_\mathrm{e} = 1.0 \,, \tag{5.56a}$$

$$k_{\rho_0} = 4.0 \,, \quad k_{T_\mathrm{i}} = k_{T_\mathrm{e}} = 20.0 \,, \tag{5.56b}$$

$$w_{\rho_0} = 0.2 \,, \quad w_{T_\mathrm{i}} = w_{T_\mathrm{e}} = 0.1 \,. \tag{5.56c}$$

The profiles in (5.55) are shown in Figure 5.3. The safety factor profile $q(r)$ is chosen to be uniform in the whole radial domain, $q(r) = 1.25$, and $\bar{B}_0 = 1$. The values of the parameters defining the initial setup of our numerical test are chosen in order to match the initial setup of the numerical simulations presented in (Latu et al., 2018, section 4.3). The initial distribution function is defined as a small perturbation added on an equilibrium local Maxwellian:

$$F(0, r, \theta, \varphi, P_\parallel) := F_0(r, P_\parallel) \left[1 + \epsilon \exp\left(- \frac{(r - \bar{r})^2}{4 w_{\rho_0} / w_{T_\mathrm{i}}} \right) \cos(m\theta + n\varphi) \right] , \tag{5.57}$$

where the equilibrium local Maxwellian F_0 is defined as in (5.36). The numerical results have been verified against the analytical dispersion relation for a perturbation with poloidal mode number $m = 9$, toroidal mode number $n = -8$ and amplitude $\epsilon = 10^{-8}$. The exponential growth rate observed in the numerical experiment is in good agreement with the analytical one, $\mathrm{Im}\, \omega \approx 0.005$, for the time interval $1000 \lesssim t \lesssim 3000$, which corresponds to the linear phase, as shown in Figure 5.4. At time $t \approx 3000$ the system enters its non-linear phase. The simulation is run

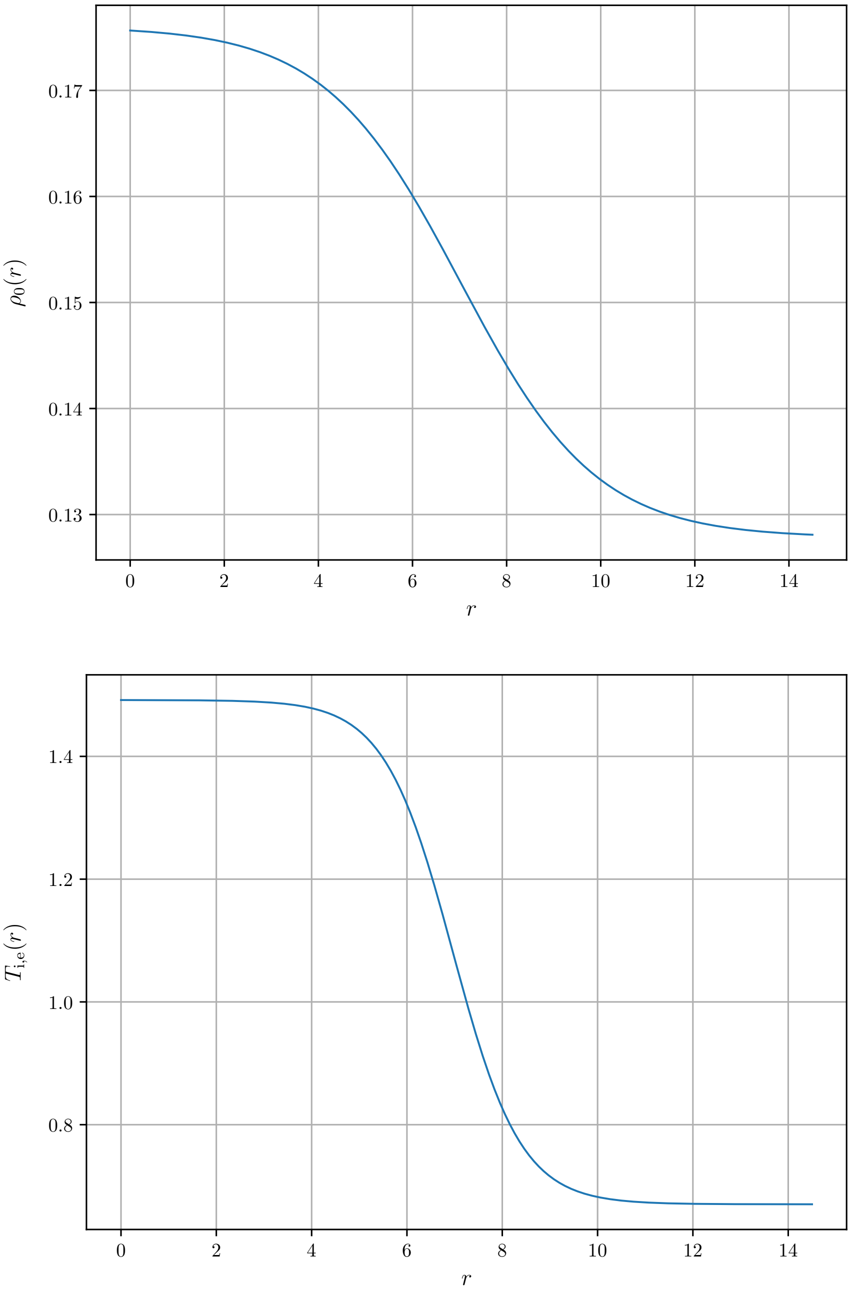


Figure 5.3: Numerical simulation of an ITG instability in cylindrical geometry with a screw pinch magnetic field: density and temperature profiles used to initialize the simulation.

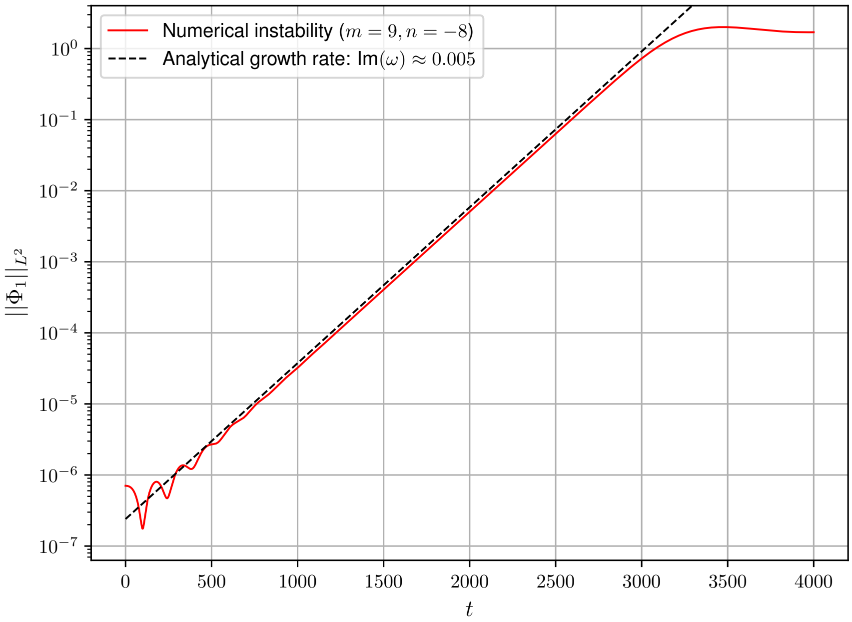


Figure 5.4: Numerical simulation of an ITG instability in cylindrical geometry with a screw pinch magnetic field: L^2-norm of the perturbed electric potential.

with $n_1 \times n_2 \times n_3 \times n_4 = 256 \times 512 \times 32 \times 128$, $\Delta t = 2.0$ and a momentum domain $[-8.0, 8.0]$. Numerical results are shown in Figures 5.4-5.8. For the conservation of the total mass $M(t)$ and the L^2-norm of F we get

$$\max_{t \in [0,4000]} \delta M(t) \approx 1.6 \times 10^{-7}\,, \qquad \max_{t \in [0,4000]} \delta ||F||_{L^2}(t) \approx 2.7 \times 10^{-5}\,. \tag{5.58}$$

5.5 The 4D drift-kinetic model in toroidal geometries

We now describe model (5.1) in a toroidal geometry with circular concentric poloidal cross sections, conformal to the circular concentric flux surfaces of the magnetic field. In this case we choose the coordinates $(\zeta^1, \zeta^2, \zeta^3)$ as the cylindrical coordi-

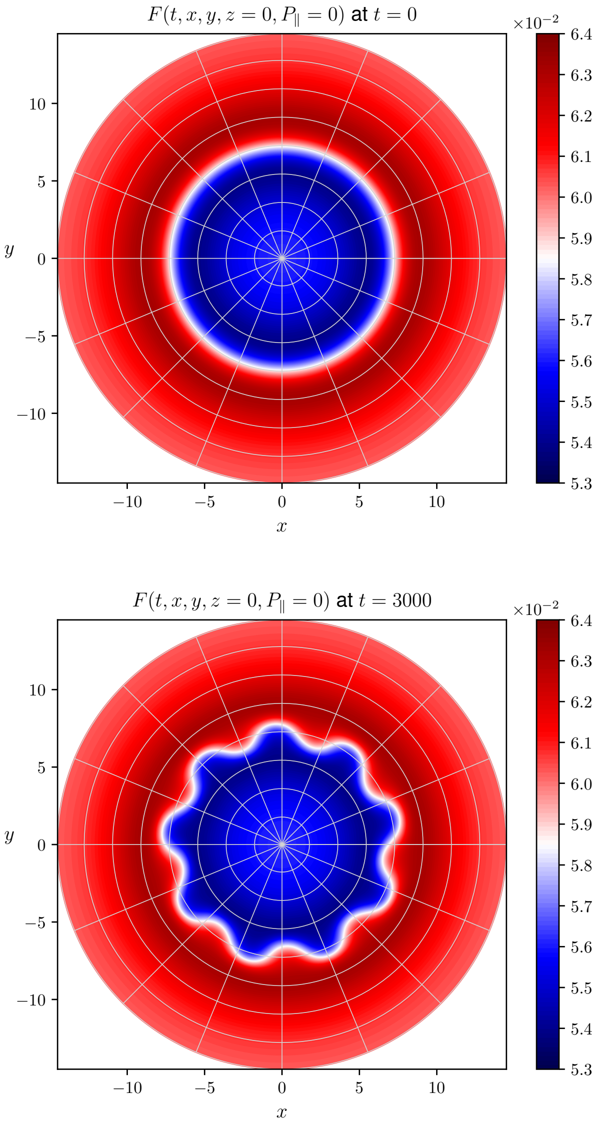


Figure 5.5: Numerical simulation of an ITG instability in cylindrical geometry with a screw pinch magnetic field: contour plots of the distribution function $F(t, x, y, z = 0, P_\parallel = 0)$ at times $t = 0$ (beginning of the simulation) and $t = 3000$ (end of the linear phase).

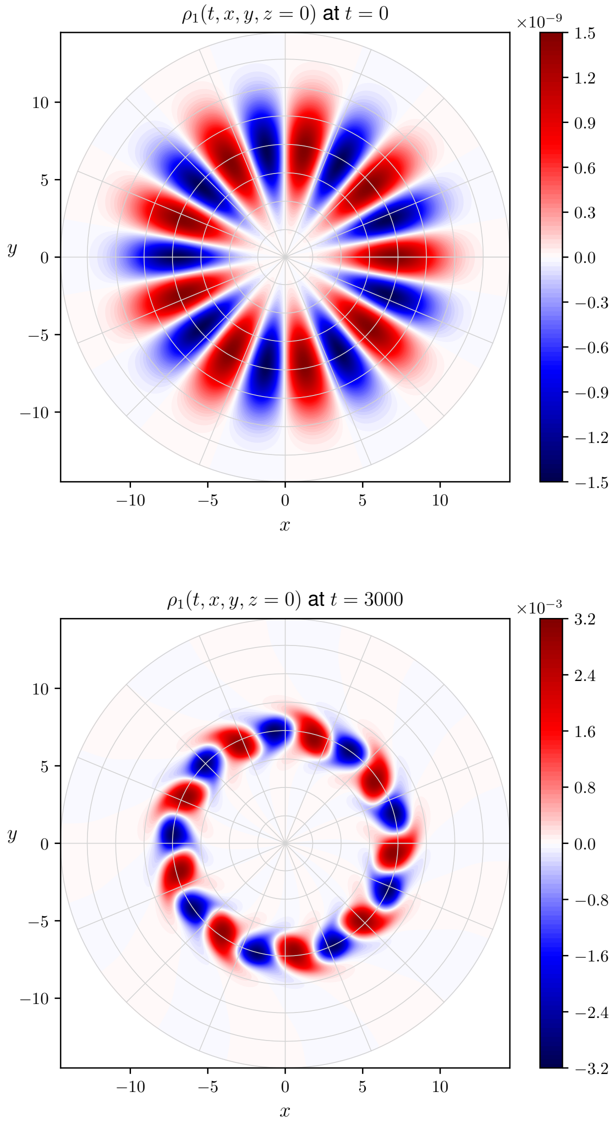
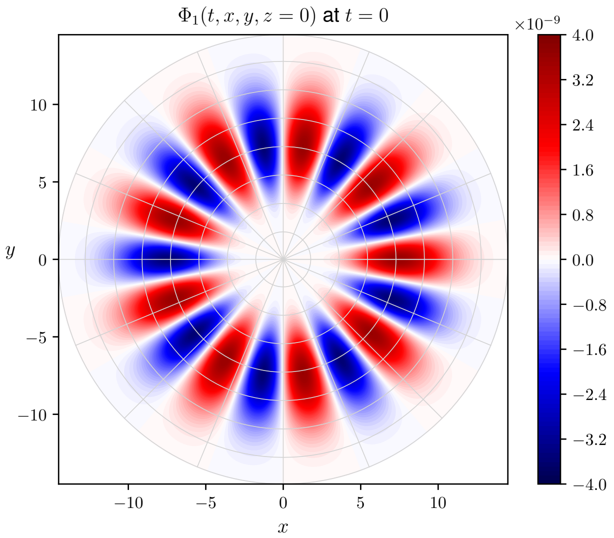


Figure 5.6: Numerical simulation of an ITG instability in cylindrical geometry with a screw pinch magnetic field: contour plots of the density perturbation $\rho_1(t, x, y, z = 0)$ at times $t = 0$ (beginning of the simulation) and $t = 3000$ (end of the linear phase).



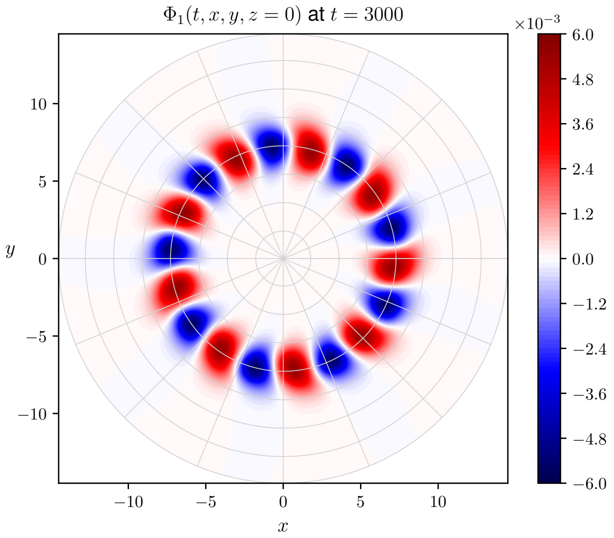


Figure 5.7: Numerical simulation of an ITG instability in cylindrical geometry with a screw pinch magnetic field: contour plots of the perturbed electric potential $\Phi_1(t, x, y, z = 0)$ at times $t = 0$ (beginning of the simulation) and $t = 3000$ (end of the linear phase).

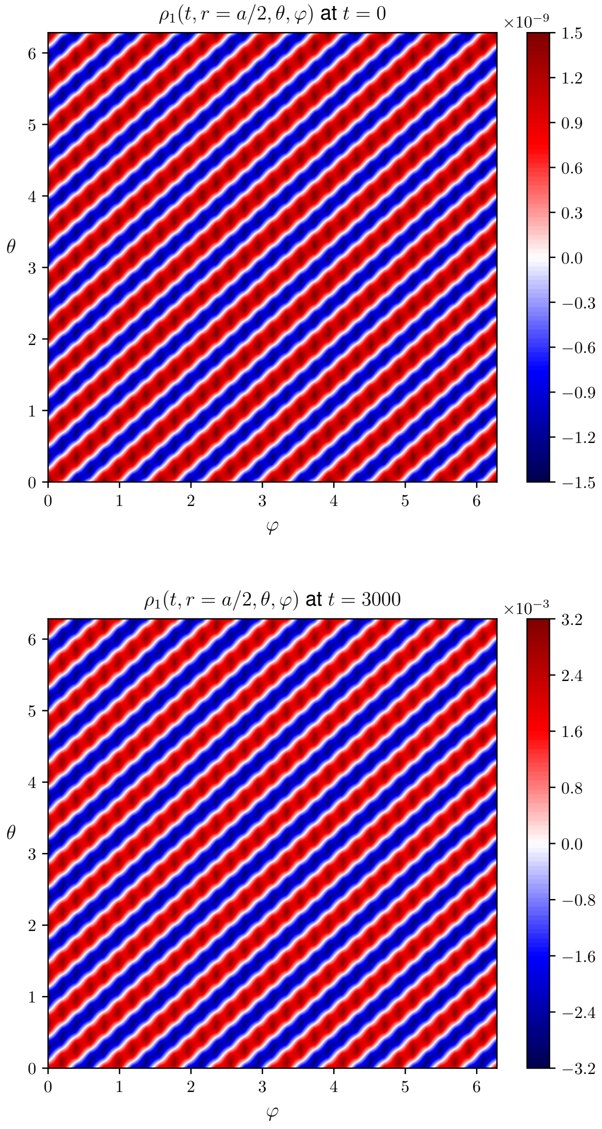


Figure 5.8: Numerical simulation of an ITG instability in cylindrical geometry with a screw pinch magnetic field: contour plots of the density perturbation $\rho_1(t, r = a/2, \theta, \varphi)$ on the flux surface at $r = a/2$ at times $t = 0$ (beginning of the simulation) and $t = 3000$ (end of the linear phase). The oblique lines in the top figure are exactly straight: any oscillation is an artifact of the interpolation (not field-aligned) used for visualization.

nates (R, Z, φ), related to the logical coordinates (r, θ, φ) via the transformation

$$\begin{aligned} R(r,\theta,\varphi) &= R_0 + r\cos\theta\,, \\ Z(r,\theta,\varphi) &= r\sin\theta\,, \\ \varphi(r,\theta,\varphi) &= \varphi\,, \end{aligned} \tag{5.59}$$

where R_0 denotes the major radius of the torus. The contravariant components $(B_0^R, B_0^Z, B_0^\varphi)$ of the magnetic field $\boldsymbol{B}_0$ with respect to the tangent basis $(\boldsymbol{e}_R, \boldsymbol{e}_Z, \boldsymbol{e}_\varphi)$ of the cylindrical coordinate system, expressed as functions of the logical coordinates (r, θ, φ), read

$$\begin{aligned} B_0^R(r,\theta) &:= -\bar{B}_0 \frac{r\sin\theta}{(R_0 + r\cos\theta)q(r)}\,, \\ B_0^Z(r,\theta) &:= \bar{B}_0 \frac{r\cos\theta}{(R_0 + r\cos\theta)q(r)}\,, \\ B_0^\varphi(r,\theta) &:= \bar{B}_0 \frac{R_0}{(R_0 + r\cos\theta)^2}\,, \end{aligned} \tag{5.60}$$

where $\bar{B}_0$ is given and $q(r)$ denotes the safety factor profile. We note that $\boldsymbol{B}_0$ is axisymmetric and therefore does not depend on the toroidal angle coordinate φ. The magnitude B_0 of the magnetic field reads

$$B_0(r,\theta) = \sqrt{\frac{\bar{B}_0^2\,(R_0^2\,q^2(r) + r^2)}{(R_0 + r\cos\theta)^2 q^2(r)}}\,. \tag{5.61}$$

The contravariant components $(b_0^R, b_0^Z, b_0^\varphi)$ of the unit vector $\boldsymbol{b}_0 := \boldsymbol{B}_0/B_0$ with respect to the tangent basis $(\boldsymbol{e}_R, \boldsymbol{e}_Z, \boldsymbol{e}_\varphi)$ of the cylindrical coordinate system, expressed as functions of the logical coordinates (r, θ, φ), read

$$\begin{aligned} b_0^R(r,\theta) &:= \frac{B_0^R(r,\theta)}{B_0(r,\theta)} = -\frac{\bar{B}_0}{B_0(r,\theta)} \frac{r\sin\theta}{(R_0 + r\cos\theta)q(r)}\,, \\ b_0^Z(r,\theta) &:= \frac{B_0^Z(r,\theta)}{B_0(r,\theta)} = \frac{\bar{B}_0}{B_0(r,\theta)} \frac{r\cos\theta}{(R_0 + r\cos\theta)q(r)}\,, \\ b_0^\varphi(r,\theta) &:= \frac{B_0^\varphi(r,\theta)}{B_0(r,\theta)} = \frac{\bar{B}_0}{B_0(r,\theta)} \frac{R_0}{(R_0 + r\cos\theta)^2}\,. \end{aligned} \tag{5.62}$$

Moreover, the contravariant components of the curl of $\boldsymbol{b}_0$ with respect to the tangent basis $(\boldsymbol{e}_R, \boldsymbol{e}_Z, \boldsymbol{e}_\varphi)$ of the cylindrical coordinate system, expressed as functions of the logical coordinates (r, θ, φ), read

$$\begin{aligned}
(\boldsymbol{\nabla} \times \boldsymbol{b}_0)^R &:= \frac{1}{R}\left(\frac{\partial}{\partial Z}(b_0)_\varphi - \frac{\partial}{\partial \varphi}(b_0)_Z\right) = \frac{q(r) - r\,q'(r)}{R_0^2\,q^2(r) + r^2} R_0\, b_0^R(r,\theta)\,, \\
(\boldsymbol{\nabla} \times \boldsymbol{b}_0)^Z &:= \frac{1}{R}\left(\frac{\partial}{\partial \varphi}(b_0)_R - \frac{\partial}{\partial R}(b_0)_\varphi\right) = \frac{q(r) - r\,q'(r)}{R_0^2\,q^2(r) + r^2} R_0\, b_0^Z(r,\theta) - b_0^\varphi(r,\theta)\,, \\
(\boldsymbol{\nabla} \times \boldsymbol{b}_0)^\varphi &:= \frac{1}{R}\left(\frac{\partial}{\partial R}(b_0)_Z - \frac{\partial}{\partial Z}(b_0)_R\right) = \frac{r^2 + 2R_0^2\,q^2(r) - R_0^2\,r\,q(r)\,q'(r)}{(R_0^2\,q^2(r) + r^2)R_0\,q(r)}\, b_0^\varphi(r,\theta)\,,
\end{aligned}$$

where $q'(r) = \mathrm{d}q(r)/\mathrm{d}r$. Here we used the fact that $(b_0)_R = b_0^R$, $(b_0)_Z = b_0^Z$ and $(b_0)_\varphi = (R_0 + r\cos\theta)^2 b_0^\varphi$, and expressed derivatives with respect to the cylindrical coordinates in terms of derivatives with respect to the logical coordinates as

$$\begin{aligned}
\frac{\partial}{\partial R} &= \frac{\partial r}{\partial R}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial R}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial R}\frac{\partial}{\partial \varphi} = \cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\,, \\
\frac{\partial}{\partial Z} &= \frac{\partial r}{\partial Z}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial Z}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial Z}\frac{\partial}{\partial \varphi} = \sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\,, \\
\frac{\partial}{\partial \varphi} &= \frac{\partial r}{\partial \varphi}\frac{\partial}{\partial r} + \frac{\partial \theta}{\partial \varphi}\frac{\partial}{\partial \theta} + \frac{\partial \varphi}{\partial \varphi}\frac{\partial}{\partial \varphi} = \frac{\partial}{\partial \varphi}\,.
\end{aligned} \tag{5.63}$$

The equation of a magnetic field line in logical coordinates reads

$$\frac{B_0^\theta}{\mathrm{d}\theta} = \frac{B_0^\varphi}{\mathrm{d}\varphi}\,, \tag{5.64}$$

which gives

$$\frac{\mathrm{d}\theta}{R_0 + r\cos\theta} = \frac{\mathrm{d}\varphi}{R_0\,q(r)}\,. \tag{5.65}$$

Here we used the fact that the contravariant components $(B_0^r, B_0^\theta, B_0^\varphi)$ of the magnetic field $\boldsymbol{B}_0$ with respect to the tangent basis $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$ of the logical coordinate system can be computed from the corresponding contravariant components

$(B_0^R, B_0^Z, B_0^\varphi)$ using (A.18):

$$\begin{aligned} B_0^r &= \cos\theta \, B_0^R + \sin\theta \, B_0^Z \, , \\ B_0^\theta &= -\frac{\sin\theta}{r} \, B_0^R + \frac{\cos\theta}{r} \, B_0^Z \, , \\ B_0^\varphi &= B_0^\varphi \, . \end{aligned} \tag{5.66}$$

The integration of the right-hand side of (5.65) in the interval $[\varphi_0, \varphi]$ yields

$$\int_{\varphi_0}^{\varphi} \frac{\mathrm{d}\varphi'}{R_0 \, q(r)} = \frac{\varphi - \varphi_0}{R_0 \, q(r)} \, . \tag{5.67}$$

The integration of the left-hand side of (5.65) in the interval $[\theta_0, \theta]$,

$$\int_{\theta_0}^{\theta} \frac{\mathrm{d}\theta'}{R_0 + r\cos\theta'} \, , \tag{5.68}$$

can be performed using the trigonometric identity

$$\cos\theta = \frac{1 - \tan^2(\theta/2)}{1 + \tan^2(\theta/2)} \, , \tag{5.69}$$

and introducing the change of variable

$$\xi' := \sqrt{\frac{R_0 - r}{R_0 + r}} \tan\frac{\theta'}{2} \, , \tag{5.70}$$

which eventually yields

$$\begin{aligned} \int_{\theta_0}^{\theta} \frac{\mathrm{d}\theta'}{R_0 + r\cos\theta'} = \frac{2}{\sqrt{R_0^2 - r^2}} \Bigg[& \arctan\left(\sqrt{\frac{R_0 - r}{R_0 + r}} \tan\frac{\theta}{2} \right) \\ & - \arctan\left(\sqrt{\frac{R_0 - r}{R_0 + r}} \tan\frac{\theta_0}{2} \right) \Bigg] . \end{aligned} \tag{5.71}$$

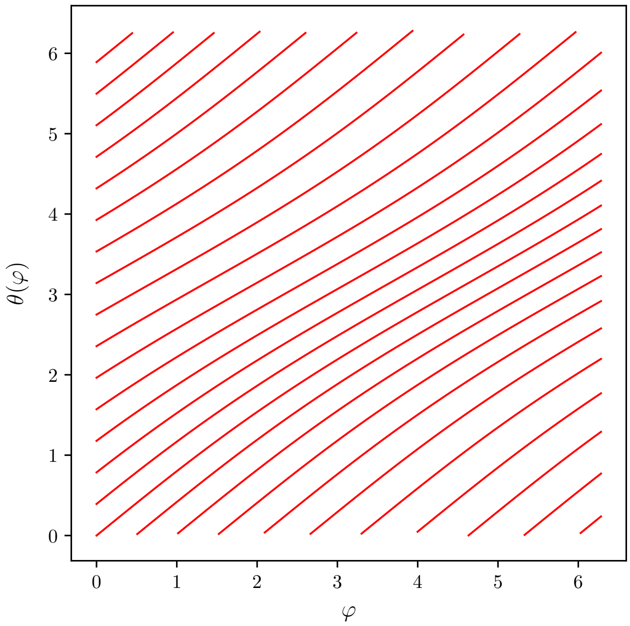


Figure 5.9: Magnetic field lines for the magnetic field (5.60) on the flux surface at $r = a/2$.

Therefore, a magnetic field line passing through the point (r, θ_0, φ_0) can be described by the relation between θ and φ, which reads

$$\theta(\varphi) = 2\arctan\left(\sqrt{\frac{R_0 + r}{R_0 - r}} \tan\left\{\frac{\sqrt{R_0^2 - r^2}}{2} \times \right.\right.$$
$$\left.\left. \times \left[\frac{2}{\sqrt{R_0^2 - r^2}} \arctan\left(\sqrt{\frac{R_0 - r}{R_0 + r}} \tan\frac{\theta_0}{2}\right) + \frac{\varphi - \varphi_0}{R_0\, q(r)}\right]\right\}\right).$$

Figure 5.9 shows these magnetic field lines on the flux surface at $r = a/2$.

5.5.1 Kinetic equilibrium

We now briefly review how to define a kinetic equilibrium distribution function, denoted again by F_0, for model (5.1) in a toroidal geometry with circular poloidal cross sections with the magnetic field (5.60). This problem has been addressed and discussed, for example, in (Idomura et al., 2003; Angelino et al., 2006; Dif-Pradalier et al., 2008a,b; Di Troia, 2012). By definition, the equilibrium F_0 does not depend on time. Moreover, we assume that F_0 is axisymmetric and thus depends only on $(r, \theta, P_\parallel)$. Such equilibrium distribution function is a stationary solution of the Vlasov equation in (5.1), where only the equilibrium advection fields $\boldsymbol{U}_0$ and $U_{P_\parallel 0}$, given by

$$\boldsymbol{U}_0 = \frac{P_\parallel}{B^*_\parallel} \boldsymbol{B}^* \,, \qquad U_{P_\parallel 0} = 0 \,, \tag{5.72}$$

are considered. Under these assumptions, the equation for F_0 reads simply

$$\boldsymbol{U}_0 \cdot \boldsymbol{\nabla} F_0 = 0 \,, \tag{5.73}$$

which in logical coordinates yields

$$P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^r \frac{\partial F_0}{\partial r} + \left[B_0^\theta + P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^\theta \right] \frac{\partial F_0}{\partial \theta} = 0 \,. \tag{5.74}$$

Following (Angelino et al., 2006), we assume that F_0 depends on the variables $(r, \theta, P_\parallel)$ only implicitly through two new variables $\widehat{\Psi}(r, \theta, P_\parallel)$ and $K(P_\parallel)$. More precisely, we can write $F_0(r, \theta, P_\parallel) = \widetilde{F}_0(\widehat{\Psi}(r, \theta, P_\parallel), K(P_\parallel))$. The idea behind this assumption is to choose a distribution function which depends only on the canonical invariants of (5.73), namely the poloidal magnetic flux and the kinetic energy, denoted here by $\widehat{\Psi}$ and K, respectively. For this reason, the kinetic equilibrium distribution function that will result from our analysis is usually referred to as canonical Maxwellian, as it exhibits a functional form similar to a standard local Maxwellian, but it is modified in order to depend only on the above-mentioned canonical invariants. We then write (5.74) as

$$\frac{\partial \widetilde{F}_0}{\partial \widehat{\Psi}} \left(P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^r \frac{\partial \widehat{\Psi}}{\partial r} + \left[B_0^\theta + P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^\theta \right] \frac{\partial \widehat{\Psi}}{\partial \theta} \right) = 0 \,. \tag{5.75}$$

If we assume that $\widehat{\Psi}$ has the form

$$\widehat{\Psi}(r, \theta, P_\parallel) := \Psi(r) + P_\parallel (R_0 + r \cos\theta) \chi(r) \,, \tag{5.76}$$

it follows that (5.75) holds if and only if the following two ordinary differential equations are satisfied by Ψ and χ, respectively:

$$\frac{\mathrm{d}\Psi}{\mathrm{d}r} = \frac{r\sin\theta B_0^\theta}{(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^r}\chi\,, \tag{5.77a}$$

$$\frac{\mathrm{d}\chi}{\mathrm{d}r} + \left(\frac{\cos\theta}{R_0 + r\cos\theta} - \frac{(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^\theta}{(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^r}\frac{r\sin\theta}{R_0 + r\cos\theta}\right)\chi = 0\,. \tag{5.77b}$$

Here we can use again the fact that the contravariant components of the curl of $\boldsymbol{b}_0$ with respect to the tangent basis $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$ of the logical coordinate system can be computed from the corresponding contravariant components with respect to the tangent basis $(\boldsymbol{e}_R, \boldsymbol{e}_Z, \boldsymbol{e}_\varphi)$ of the cylindrical coordinate system using (A.18):

$$\begin{aligned}
(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^r &= \cos\theta\,(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^R + \sin\theta\,(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^Z\,,\\
(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^\theta &= -\frac{\sin\theta}{r}\,(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^R + \frac{\cos\theta}{r}\,(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^Z\,,\\
(\boldsymbol{\nabla}\times\boldsymbol{b}_0)^\varphi &= (\boldsymbol{\nabla}\times\boldsymbol{b}_0)^\varphi\,.
\end{aligned} \tag{5.78}$$

Therefore, the equation for χ reads

$$\frac{\mathrm{d}\chi}{\mathrm{d}r} + \frac{\xi\,\xi'}{1+\xi^2}\chi = 0\,, \tag{5.79}$$

where the auxiliary function ξ is defined as

$$\xi(r) := \frac{r}{R_0\,q(r)}\,. \tag{5.80}$$

The solution of (5.79) is

$$\chi(r) = \frac{\chi_0}{\sqrt{1+\xi^2(r)}}\,, \tag{5.81}$$

where χ_0 is a constant. The function Ψ is then given by

$$\Psi(r) = -|\bar{B}_0|\chi_0\int_0^r \mathrm{d}r'\,\frac{r'}{q(r')}\,. \tag{5.82}$$

Any function of the two variables $\widehat{\Psi}$ and K, with

$$\widehat{\Psi}(r,\theta,P_\parallel) = -|\bar{B}_0|\chi_0 \int_0^r \mathrm{d}r' \, \frac{r'}{q(r')} \; + \; P_\parallel (R_0 + r\cos\theta) \frac{\chi_0}{\sqrt{1+\xi^2(r)}} \, , \tag{5.83}$$

is an equilibrium distribution function for (5.1). A shift $\delta\widehat{\Psi}(P_\parallel) := -R_0 P_\parallel$ on the poloidal magnetic flux $\widehat{\Psi}$ can be introduced in order to minimize parallel flows (Angelino et al., 2006). To sum up, a canonical Maxwellian distribution function is given by

$$F_0(\widehat{\Psi},K) := \frac{1}{\widehat{B}_0} \frac{\widehat{\rho}_0}{\sqrt{2\pi \, \widehat{T}_\mathrm{i}}} \exp\left(-\frac{K}{\widehat{T}_\mathrm{i}}\right) \, , \tag{5.84}$$

where $\widehat{\rho}_0 = \rho_0(\widehat{r})$, $\widehat{T}_\mathrm{i} = T_\mathrm{i}(\widehat{r})$, $\widehat{B}_0 = B_0(\widehat{r},\theta=\pi/2)$, with $\widehat{r}$ defined as the inverse function $r(\Psi)$ evaluated at $\Psi = \widehat{\Psi}$. If q is uniform, i.e. $q(r) = q_0$, we have (setting $\chi_0 = 1$)

$$\Psi(r) = -|\bar{B}_0| \frac{r^2}{2q_0} \, , \tag{5.85}$$

with the corresponding inverse function

$$r(\Psi) = \sqrt{-\frac{2q_0}{|\bar{B}_0|}\Psi} \, . \tag{5.86}$$

If q has the form $q(r) = q_0 + q_2 r^2$, we have (setting again $\chi_0 = 1$)

$$\Psi(r) = -\frac{|\bar{B}_0|}{2q_2} \log\frac{q(r)}{q_0} \, , \tag{5.87}$$

with the corresponding inverse function

$$r(\Psi) = \sqrt{\frac{q_0}{q_2}\left[\exp\left(-\frac{2q_2}{|\bar{B}_0|}\Psi\right) - 1\right]} \, . \tag{5.88}$$

In general, for a given profile $q(r)$, the flux function Ψ can be obtained by computing numerically the integral (5.82). The corresponding inverse function $r(\Psi)$ can be then computed by a numerical inversion: for example, for any given value $\Psi = \Psi^*$, we could use a Newton method to find $r = r^*$ that solves $\Psi(r) - \Psi^* = 0$.

5.5.2 Boundary conditions

We now address another issue that comes into play when studying our drift-kinetic model in toroidal geometry, namely the problem of particles crossing the boundary of the computational domain, when solving the 2D advection problem (5.7) on a given poloidal plane. Such problem occurs even when we consider only the equilibrium advection field $\boldsymbol{U}_0$. By looking at the kinetic equilibrium distribution function F_0 given by the canonical Maxwellian (5.84), we indeed notice that the contour lines of F_0 on a given poloidal plane are not conformal to the circular concentric flux surfaces of the magnetic field. This is shown in Figure 5.10 with the parameters of the highly non-physical scenario considered for the numerical simulations described in section 5.4.2 and with the more realistic parameters considered for the numerical simulations described in the next section. As a consequence, the characteristics obtained by solving (5.7), even when we consider only the equilibrium advection field $\boldsymbol{U}_0$, can go across the boundary of the poloidal plane. This corresponds to having particles that enter the computational domain from the region outside the poloidal plane, as we integrate the characteristics backward in time. We therefore need a strategy for the integration of (5.7) in the region outside the poloidal plane. For this purpose, we introduce an outer computational domain $(a, r_\text{o}] \times [0, 2\pi)$, for a given $r_\text{o} > a$. The strategy we suggest is based on the observation that the electric field becomes very small when we approach the boundary of the poloidal plane and consists in integrating (5.7) by considering only the equilibrium advection field $\boldsymbol{U}_0$. In other words, we assume that the electric field is zero in the outer region. Since the radial electric field is small but not exactly zero when we approach the boundary from the inner region, in order for our strategy to work at the numerical level, we impose by hand zero radial electric field at $r = a$, thus connecting the inner and outer regions in a smooth way without discontinuities. Moreover, when we need to interpolate the distribution function at a given point in the outer region, starting from its values at neighboring mesh points, we choose to interpolate the equilibrium distribution function F_0. One computational advantage of our strategy is that the values of the equilibrium advection field $\boldsymbol{U}_0$ and of the equilibrium distribution function F_0 at the mesh points in the outer region can be pre-computed only once. However, when pre-computing the values of F_0 we have to take into account the fact that, according

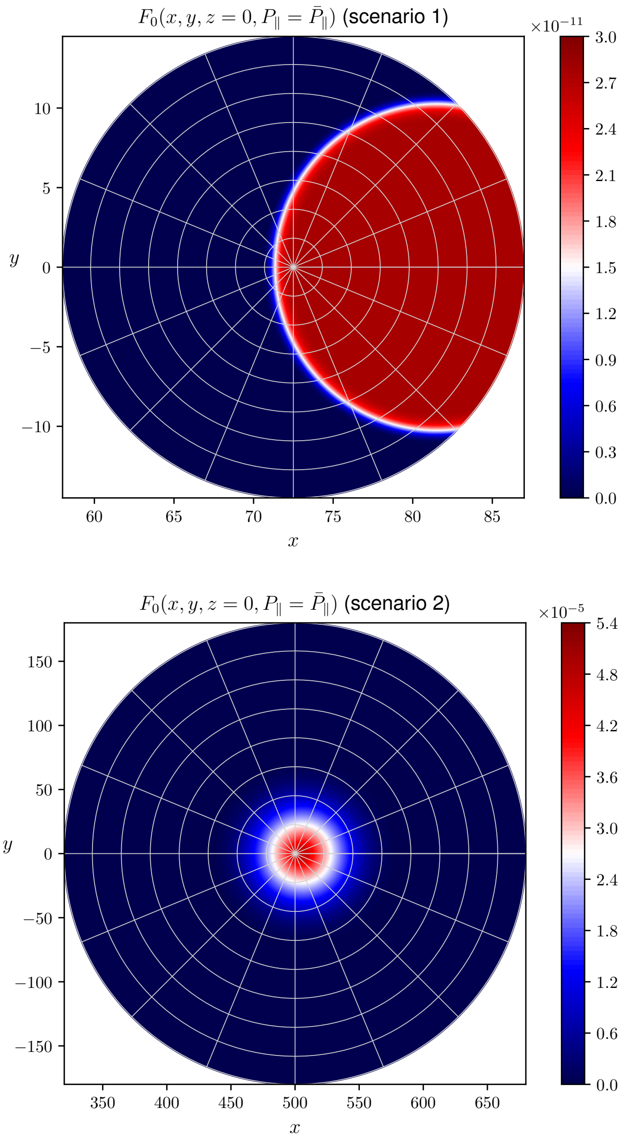


Figure 5.10: Contour plots of the equilibrium canonical Maxwellian distribution function F_0 in toroidal geometry with circular poloidal cross sections, with the parameters of the highly non-physical scenario considered for the numerical simulations described in section 5.4.2 (top) and with the more realistic parameters considered for the numerical simulations described in section 5.5.3 (bottom). In both cases, the contour lines of F_0 are not conformal to the circular concentric flux surfaces of the magnetic field.

to our field-aligned splitting combined with the time-advancing strategy described in section 3.4, the poloidal advection (5.7) is solved after solving the field-aligned advection on flux surfaces within half a time step. Therefore, in order to be consistent with our splitting, the equilibrium distribution F_0 should also undergo $\Delta t/2$ of field-aligned advection: when pre-computing the value of F_0 at a given phase-space point $(r_i, \theta_j, \varphi_k, P_{\parallel\ell})$ in the outer region $a < r_i \leqslant r_o$, we set

$$F_0(r_i, \theta_j, \varphi_k, P_{\parallel\ell}) = F_0(r_i, \theta'_j, \varphi'_k, P_{\parallel\ell}) \,, \tag{5.89}$$

where the modified coordinates θ'_j and φ'_k are obtained from θ_j and φ_k by integrating the characteristics of (5.8) on the flux surface at $r = r_i$ within half a time step:

$$\theta'_j := \theta_j - \frac{\Delta t}{2} \frac{b_0^\theta}{b_0^\varphi} U_0^\varphi \,, \qquad \varphi'_k := \varphi_k - \frac{\Delta t}{2} U_0^\varphi \,, \tag{5.90}$$

where U_0^φ denotes the third contravariant component of $\boldsymbol{U}_0$ with respect to the tangent basis $(\boldsymbol{e}_r, \boldsymbol{e}_\theta, \boldsymbol{e}_\varphi)$ of the logical coordinate system and is given by

$$U_0^\varphi = \frac{P_\parallel}{B_\parallel^*} \left(B_0^\varphi + P_\parallel (\boldsymbol{\nabla} \times \boldsymbol{b}_0)^\varphi \right) \,. \tag{5.91}$$

5.5.3 Numerical results

We present here the numerical results obtained from the first tests performed on model (5.1) in a toroidal geometry with circular poloidal cross sections. We note that this is also the geometry currently available in the gyrokinetic code GySeLa (Grandgirard et al., 2006a,b, 2016), with the difference that the center of the poloidal plane is included in our case, thanks to the numerical machinery that we developed for 2D singular geometries (described in the previous chapter). The parameters defining the toroidal geometry considered and the initial density and temperature profiles of ions and electrons are identical to the ones of the test case presented in (Görler et al., 2016, section III), based in turn on the cyclone base case (Greenfield et al., 1997; Dimits et al., 2000). More precisely, the minor and major radius of the torus are set to $a = 180$ and $R_0 = 500$, thus yielding an aspect ratio $A = 2.78$, which is quite close to the realistic physical values ($A \approx 2.0$ for ASDEX Upgrade, $A \approx 2.4$ for JET, $A \approx 3.1$ for ITER). The density and temperature profiles of ions and electrons, as for the tests performed in cylindrical geometry, are defined as

$$\rho_0(r) = n_\mathrm{e}(r) := \bar{\rho}_0 \exp\left[- k_{\rho_0} w_{\rho_0} \frac{a}{R_0} \tanh\left(\frac{r - \bar{r}}{w_{\rho_0} a} \right) \right] , \tag{5.92a}$$

$$T_\mathrm{i}(r) := \overline{T}_\mathrm{i} \exp\left[- k_{T_\mathrm{i}} w_{T_\mathrm{i}} \frac{a}{R_0} \tanh\left(\frac{r - \bar{r}}{w_{T_\mathrm{i}} a} \right) \right] , \tag{5.92b}$$

$$T_\mathrm{e}(r) := \overline{T}_\mathrm{e} \exp\left[- k_{T_\mathrm{e}} w_{T_\mathrm{e}} \frac{a}{R_0} \tanh\left(\frac{r - \bar{r}}{w_{T_\mathrm{e}} a} \right) \right] , \tag{5.92c}$$

where $\bar{r} := a/2$ and the parameters are set as

$$\bar{\rho}_0 = 1.0 \ , \quad \overline{T}_\mathrm{i} = \overline{T}_\mathrm{e} = 1.0 \ , \tag{5.93a}$$

$$k_{\rho_0} = 2.23 \, , \quad k_{T_\mathrm{i}} = k_{T_\mathrm{e}} = 6.96 \, , \tag{5.93b}$$

$$w_{\rho_0} = 0.3 \ , \quad w_{T_\mathrm{i}} = w_{T_\mathrm{e}} = 0.3 \ . \tag{5.93c}$$

The profiles (5.92) are shown in Figure 5.11. The safety factor profile $q(r)$ is chosen as $q(r) := q_0 + q_2 r^2$, with $q_0 = 0.84$ and $q_2 = 2.26/a^2$, again very close to the one considered in the test case presented in (Görler et al., 2016, section III), and $\bar{B}_0 = 1$. The numerical test consists in loading a kinetic equilibrium distribution function F_0 for model (5.1) in toroidal geometry and verifying that we are capable of maintaining the equilibrium while we let the system evolve in time. The initial distribution function is defined as the equilibrium canonical Maxwellian (5.84). The simulation is run with $n_1 \times n_2 \times n_3 \times n_4 = 256 \times 512 \times 32 \times 128$, $\Delta t = 2.0$ and a momentum domain $[-6.0, 6.0]$. Numerical results are shown in Figures 5.12-5.14. For the conservation of the total mass $M(t)$ and the L^2-norm of F we get

$$\max_{t \in [0,4000]} \delta M(t) \approx 6.8 \times 10^{-10} \, , \qquad \max_{t \in [0,4000]} \delta ||F||_{L^2}(t) \approx 4.7 \times 10^{-10} \, . \tag{5.94}$$

It is also interesting to measure within this simple numerical test the residual electric potential Φ_1 after one time step. This gives information on how small a perturbation of the kinetic equilibrium distribution function could be in order to produce an initial condition that makes sense from a numerical point of view. More precisely, the

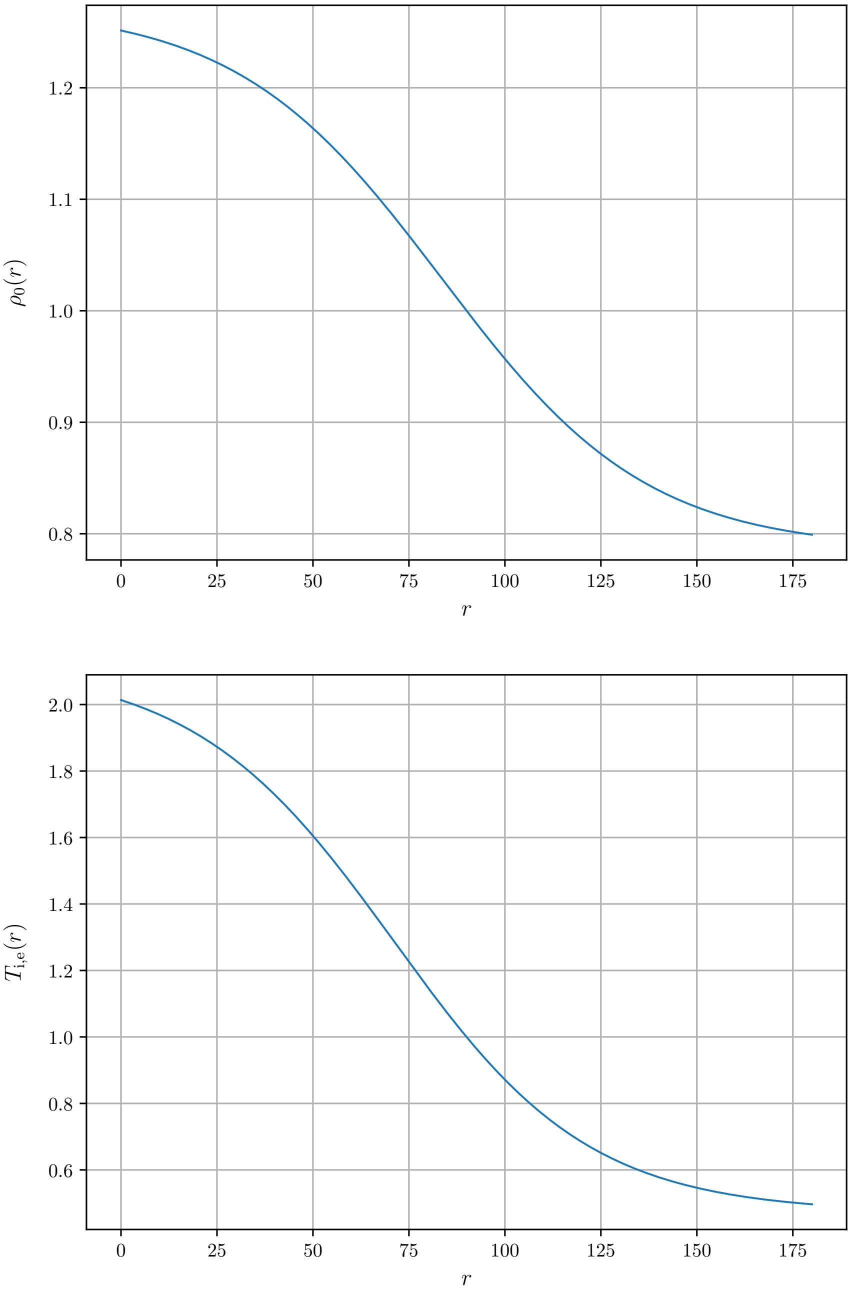


Figure 5.11: Numerical test of the 4D drift-kinetic electrostatic model in toroidal geometry with circular poloidal cross sections: density and temperature profiles used to initialize the simulation.

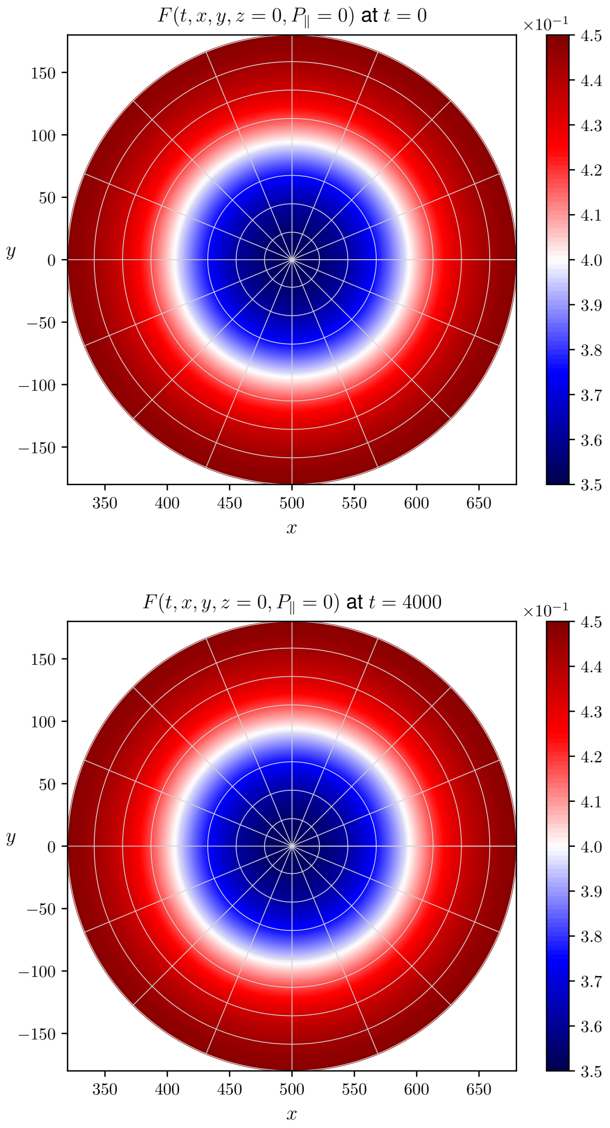


Figure 5.12: Numerical test of the 4D drift-kinetic electrostatic model in toroidal geometry with circular poloidal cross sections: contour plots of the distribution function $F(t, x, y, z = 0, P_\parallel = 0)$ at times $t = 0$ (beginning of the simulation) and $t = 4000$ (end of the simulation). A kinetic equilibrium distribution function is loaded and well maintained during the time evolution.

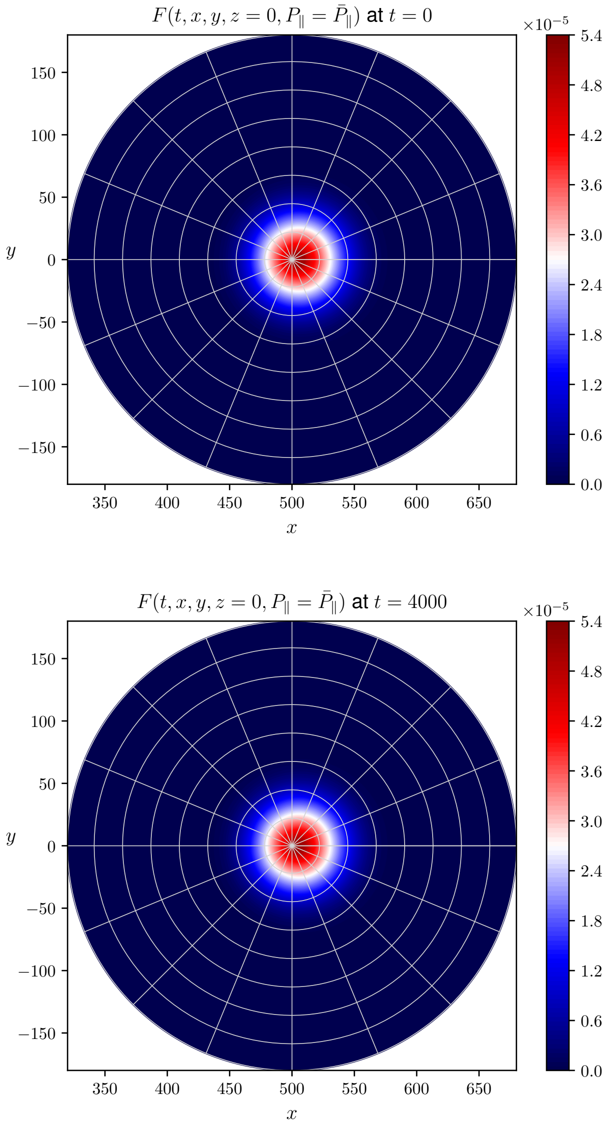


Figure 5.13: Numerical test of the 4D drift-kinetic electrostatic model in toroidal geometry with circular poloidal cross sections: contour plots of the distribution function $F(t, x, y, z = 0, P_\parallel = \bar{P}_\parallel)$ at times $t = 0$ (beginning of the simulation) and $t = 4000$ (end of the simulation). A kinetic equilibrium distribution function is loaded and well maintained during the time evolution.

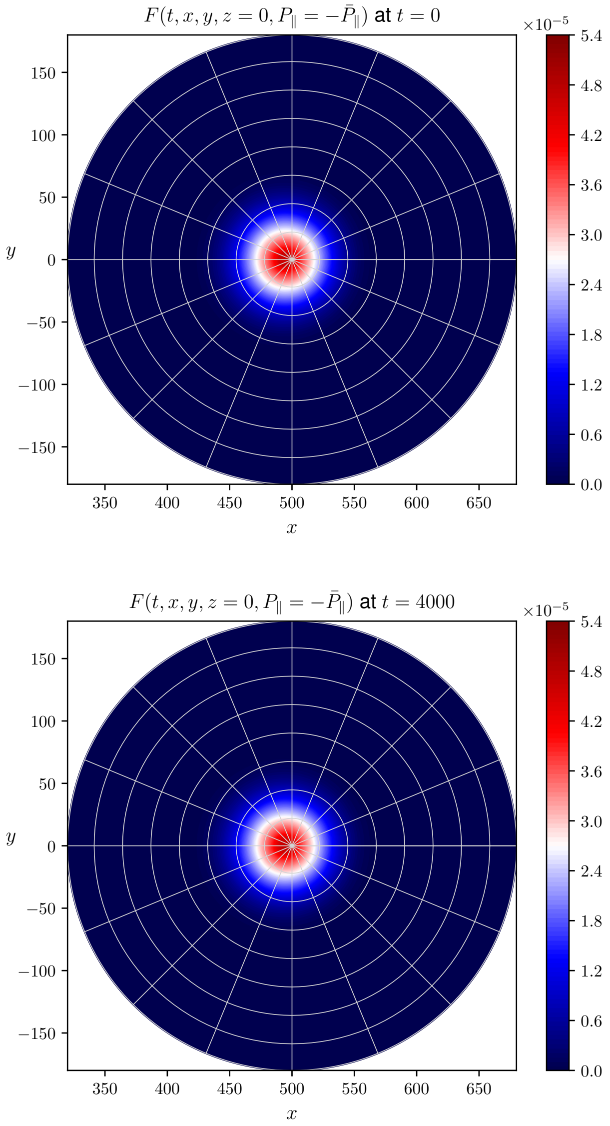


Figure 5.14: Numerical test of the 4D drift-kinetic electrostatic model in toroidal geometry with circular poloidal cross sections: contour plots of the distribution function $F(t,x,y,z=0,P_\parallel=-\bar{P}_\parallel)$ at times $t=0$ (beginning of the simulation) and $t=4000$ (end of the simulation). A kinetic equilibrium distribution function is loaded and well maintained during the time evolution.

amplitude of a perturbation should be large enough to ensure that the electric potential resulting from it is larger than the residual electric potential obtained without the perturbation. As a measure of such residual electric potential, we can compute, for example, the spatial L^∞-norm of $\Phi_1(t, x, y, z = 0)$ at the initial time $t = 0$ (where Φ_1 is expected to be zero, because there is no initial perturbation) and at time $t = 2$, after one time step:

$$\max_{(x,y)} |\Phi_1(t = 0, x, y, z = 0)| \approx 1.3 \times 10^{-16}\,, \tag{5.95a}$$

$$\max_{(x,y)} |\Phi_1(t = 2, x, y, z = 0)| \approx 3.4 \times 10^{-7}\,. \tag{5.95b}$$

As a result, we observe that the minimum amplitude ϵ for a perturbation of the kinetic equilibrium distribution function would be approximately $\epsilon \approx 10^{-4}$. Smaller perturbations would produce an electric potential which is too small for the accuracy resulting from our time splitting.

5.6 Summary of the contributions

In this chapter we discussed the latest developments of our field-aligned semi-Lagrangian drift-kinetic code. The numerical methods implemented in the code were described in detail, as well as the numerical results obtained from simulations in two different 3D geometries, namely a cylindrical geometry and a toroidal geometry with circular concentric poloidal cross sections. We showed that the numerical machinery developed for the solution of coupled hyperbolic-elliptic problems in 2D singular geometries can be applied to higher-dimensional kinetic, drift-kinetic and gyrokinetic models in a straightforward way, therefore making our numerical methods ready to be implemented in state-of-the-art gyrokinetic codes, such as, for example, the semi-Lagrangian gyrokinetic code GySeLa (Grandgirard et al., 2006a,b, 2016). The simulations performed in cylindrical geometries enabled us to provide a robust verification of our implementation, thanks to the semi-analytical results that can be derived from a linear dispersion analysis of our drift-kinetic electrostatic model. The numerical results obtained in toroidal geometry should be considered as first simple tests and can be certainly improved and extended to cover more complex and realistic scenarios.

The starting point at the beginning of our developments was a field-aligned semi-Lagrangian code addressing the solution of a simplified version of our 4D drift-kinetic electrostatic model in cylindrical geometry. Such code (and the corresponding model) is described, for example, in (Latu et al., 2018, section 4). It was the primary aim of this thesis to improve that code and extend its capabilities, with respect to both the physical model addressed and the available 3D geometries. Our main contributions can be summarized as follows:

- the extension of the physical model (and its linear dispersion analysis) from the simplified one presented in (Latu et al., 2018, section 4) to a more comprehensive drift-kinetic electrostatic model, closer to the physical model implemented, for example, in the semi-Lagrangian gyrokinetic code GySeLa;
- the extension of the geometric capabilities of the existing code to toroidal geometries;
- the integration in the existing code of the numerical machinery developed for the solution of coupled hyperbolic-elliptic problems in 2D singular geometries;
- the treatment of boundary conditions in toroidal geometries, handling the problem of particles crossing the boundaries of the poloidal planes.

The rest of the material presented in this chapter refers to either background mathematical notions or mathematical and computational tools that were already available before this research started and that may have been only slightly improved within the scope of this thesis.

Chapter 6

Conclusions and outlook

In this thesis we presented and discussed novel mathematical and computational tools for theoretical and computational gyrokinetics, which we briefly summarize in the following.

With regard to theoretical gyrokinetics, we presented and discussed a new method for the derivation of gyrokinetic models, based on polynomial transforms instead of Lie transforms. Our technique is based on a rigorous normalization of the physical equations within ordering assumptions relevant for existing and future fusion experiments, such as ASDEX Upgrade and ITER. In particular, we derived a set of gyrokinetic equations for both ions and electrons in maximal ordering. Our method is alternative to the use of Lie transforms and can provide useful insights into the derivation of gyrokinetic models. It is our hope that the derivation presented in this thesis will make gyrokinetic theory more accessible for scientists who are not familiar with it and for students at the early undergraduate level, without requiring any solid knowledge of differential geometry. Moreover, our rigorous non-dimensional approach may be of particular interest for mathematicians who want to apply techniques of error analysis and asymptotic methods to gyrokinetic models of magnetized fusion plasmas.

With regard to computational gyrokinetics, we presented and discussed novel numerical methods with the ultimate aim of providing useful computational techniques to be employed in state-of-the-art gyrokinetic codes. In chapter 3 we introduced the basic features of a field-aligned semi-Lagrangian 4D drift-kinetic code developed within the software library SeLaLib. In particular, we extended the field-aligned splitting discussed in (Latu et al., 2018) to a general setting based on curvilinear coordinates and also improved the capabilities offered by the SeLaLib library with respect to our spline discretization. In chapter 4 we presented a comprehensive numerical strategy for the solution of coupled hyperbolic-elliptic PDEs on singular mapped disk-like domains. In particular, we introduced a novel set of coordinates, named pseudo-Cartesian coordinates, for the integration in time of the characteristics of hyperbolic PDEs on such domains. Moreover, we described a finite element elliptic solver based on the $\mathcal{C}^1$ smooth polar splines presented in (Toshniwal et al., 2017). The resulting numerical techniques are fairly general and can be also applied to problems relevant for areas of physics and engineering other than plasma physics and nuclear fusion. In chapter 5 we described the latest developments on our field-aligned semi-Lagrangian 4D drift-kinetic code, which may be of interest for similar codes in the computational plasma physics community. In particular, we showed that the numerical machinery described in chapter 4 for 2D problems can be applied to higher-dimensional models in a straightforward way, therefore making our numerical methods ready to be implemented in state-of-the-art gyrokinetic codes, such as, for example, the semi-Lagrangian gyrokinetic code GySeLa (Grandgirard et al., 2006a,b, 2016).

All the numerical tests presented in this thesis were described in detail, as they represent a fundamental building block of code development. The field-aligned semi-Lagrangian 4D drift-kinetic code can be further developed and improved in several ways. In order to consider singular disk-like poloidal domains resembling realistic Tokamak geometries, as for the 2D setting described in chapter 4, the capability of constructing a discrete mesh from a given magnetic field should be added. For example, an interface with the software Tokamesh (Guillard et al., 2018) could be built, as the two codes are based on a very similar discretization. Moreover, a fifth dimension (corresponding to the magnetic moment) should be added, in order to be able to perform all the standard test cases used for verification and validation purposes between different gyrokinetic codes, such as the ones discussed in (Görler et al., 2016). In order for our model to correspond precisely to the ones considered for such test cases, the quasi-neutrality equation for the electric potential should be also modified by including the so-called “zonal-flow” term, namely the magnetic flux surface average of the electric potential. The method described in (Crouseilles et al., 2012, section 6) could be implemented in order to avoid non-locality in the finite element solution of the quasi-neutrality equation, thus preserving the overall efficiency of the code. With regard to the code efficiency, it would be certainly helpful to parallelize the serial finite element elliptic solver described in chapter 4 and implemented at the moment in our 4D code. An interesting approach to tackle this problem could be to build an interface for the similar parallel finite element elliptic solver currently implemented in the novel HPC PSYDAC framework based on Python (Ratnani et al., 2019).

Appendix A

Curvilinear coordinates

In this appendix we review briefly the basics of vector algebra and analysis in curvilinear coordinates and collect several useful relations that are used consistently throughout this thesis. It is not within the scope of this appendix to derive all relations within the context of differential geometry. For this purpose and for a practical and more comprehensive treatment of the topic, we refer to (D'haeseleer et al., 1991).

A set of curvilinear coordinates (η^1, η^2, η^3) is defined by a transformation $\boldsymbol{T}$ of the form

$$\boldsymbol{T}: \begin{aligned} x &= x(\eta^1, \eta^2, \eta^3)\,, \\ y &= y(\eta^1, \eta^2, \eta^3)\,, \\ z &= z(\eta^1, \eta^2, \eta^3)\,, \end{aligned} \tag{A.1}$$

where (x, y, z) denote Cartesian coordinates in 3D space. The tangent basis of the curvilinear coordinate system defined by the transformation $\boldsymbol{T}$ is denoted by $(\boldsymbol{e}_{\eta^1}, \boldsymbol{e}_{\eta^2}, \boldsymbol{e}_{\eta^3})$ and defined as

$$\boldsymbol{e}_{\eta^i} := \frac{\partial \boldsymbol{T}}{\partial \eta^i}\,, \tag{A.2}$$

for $i = 1, 2, 3$. The corresponding cotangent basis is denoted by $(\boldsymbol{e}^{\eta^1}, \boldsymbol{e}^{\eta^2}, \boldsymbol{e}^{\eta^3})$ and defined as

$$\boldsymbol{e}^{\eta^i} := \boldsymbol{\nabla} \eta^i\,, \tag{A.3}$$

for $i = 1, 2, 3$. Here, the gradient $\boldsymbol{\nabla}\eta^i$ is such that $\boldsymbol{\nabla}\eta^i \cdot \boldsymbol{e}_{\eta^j} = \delta^i_j$, where δ^i_j is the

Kronecker delta, defined as

$$\delta^i_j := \begin{cases} 1 & \text{if } i = j\,, \\ 0 & \text{if } i \neq j\,, \end{cases} \tag{A.4}$$

for $i, j = 1, 2, 3$. The contravariant components of a vector $\boldsymbol{V}$ with respect to the tangent basis $(\boldsymbol{e}_{\eta^1}, \boldsymbol{e}_{\eta^2}, \boldsymbol{e}_{\eta^3})$ are denoted by $(V^{\eta^1}, V^{\eta^2}, V^{\eta^3})$ and defined as

$$V^{\eta^i} := \boldsymbol{V} \cdot \boldsymbol{e}^{\eta^i}\,, \tag{A.5}$$

for $i = 1, 2, 3$. Similarly, the covariant components of $\boldsymbol{V}$ with respect to the cotangent basis $(\boldsymbol{e}^{\eta^1}, \boldsymbol{e}^{\eta^2}, \boldsymbol{e}^{\eta^3})$ are denoted by $(V_{\eta^1}, V_{\eta^2}, V_{\eta^3})$ and defined as

$$V_{\eta^i} := \boldsymbol{V} \cdot \boldsymbol{e}_{\eta^i}\,, \tag{A.6}$$

for $i = 1, 2, 3$. We note that contravariant and covariant components do not differ in a Cartesian coordinate system. The metric coefficients of the curvilinear coordinate system defined by the coordinates (η^1, η^2, η^3) are denoted by $g_{ij}(\eta^1, \eta^2, \eta^3)$, for $i, j = 1, 2, 3$, and defined as

$$g_{ij} := \boldsymbol{e}_{\eta^i} \cdot \boldsymbol{e}_{\eta^j}\,. \tag{A.7}$$

The covariant component of a vector $\boldsymbol{V}$ can be expressed in terms of its contravariant components by means of the metric coefficients as

$$V_{\eta^i} = g_{ij} V^{\eta^j}\,, \tag{A.8}$$

for $i = 1, 2, 3$. Here and in the following we employ the Einstein summation convention (sum over repeated indices). The Jacobian matrix of the curvilinear coordinate system defined by the coordinates (η^1, η^2, η^3) is denoted by $J_\eta(\eta^1, \eta^2, \eta^3)$ and defined as

$$J_\eta := \begin{bmatrix} \dfrac{\partial x}{\partial \eta^1} & \dfrac{\partial x}{\partial \eta^2} & \dfrac{\partial x}{\partial \eta^3} \\ \dfrac{\partial y}{\partial \eta^1} & \dfrac{\partial y}{\partial \eta^2} & \dfrac{\partial y}{\partial \eta^3} \\ \dfrac{\partial z}{\partial \eta^1} & \dfrac{\partial z}{\partial \eta^2} & \dfrac{\partial z}{\partial \eta^3} \end{bmatrix}\,. \tag{A.9}$$

The determinant $\det J_\eta$ of the Jacobian matrix J_η can be expressed in terms of the determinant $\det G$ of the metric matrix G (defined by the metric coefficients g_{ij}) as

$$\det J_\eta = \sqrt{\det G}\,. \tag{A.10}$$

The dot product between two vectors $\boldsymbol{V}$ and $\boldsymbol{W}$ can be expressed in terms of their contravariant and covariant components as

$$\boldsymbol{V} \cdot \boldsymbol{W} = V^{\eta^i} W_{\eta^i} = V_{\eta^i} W^{\eta^i}\,. \tag{A.11}$$

The contravariant components of the cross product between two vectors $\boldsymbol{V}$ and $\boldsymbol{W}$ can be expressed in terms of their individual covariant components as

$$(\boldsymbol{V} \times \boldsymbol{W})^{\eta^k} = \frac{\varepsilon^{ijk}}{\det J_\eta} V_{\eta^i} W_{\eta^j}\,, \tag{A.12}$$

for $k = 1, 2, 3$. Here ε^{ijk} denotes the Levi-Civita symbol, defined as

$$\varepsilon^{ijk} := \begin{cases} +1 & \text{if } (i,j,k) \text{ even permutation of } (1,2,3)\,, \\ -1 & \text{if } (i,j,k) \text{ odd permutation of } (1,2,3)\,, \\ 0 & \text{otherwise}\,, \end{cases} \tag{A.13}$$

for $i, j, k = 1, 2, 3$. Similarly, the covariant components of the cross product between $\boldsymbol{V}$ and $\boldsymbol{W}$ can be expressed in terms of their individual contravariant components as

$$(\boldsymbol{V} \times \boldsymbol{W})_{\eta^k} = \varepsilon_{ijk} (\det J_\eta) V^{\eta^i} W^{\eta^j}\,, \tag{A.14}$$

for $k = 1, 2, 3$. Here ε_{ijk} is defined as in (A.13). When we consider the gradient of a scalar function $g(\eta^1, \eta^2, \eta^3)$, the derivatives of g with respect to the coordinates (η^1, η^2, η^3) define the covariant components of its gradient. In formulas:

$$(\boldsymbol{\nabla} g)_{\eta^i} = \frac{\partial g}{\partial \eta^i}\,, \tag{A.15}$$

for $i = 1, 2, 3$. The divergence of a vector $\boldsymbol{V}$ can be expressed in terms of its contravariant components as

$$\boldsymbol{\nabla} \cdot \boldsymbol{V} = \frac{1}{\det J_\eta} \frac{\partial}{\partial \eta^i} \left[(\det J_\eta) V^{\eta^i} \right] . \tag{A.16}$$

The contravariant components of the curl of a vector $\boldsymbol{V}$ can be expressed in terms of its covariant components as

$$(\boldsymbol{\nabla} \times \boldsymbol{V})^{\eta^k} = \frac{\varepsilon^{ijk}}{\det J_\eta} \frac{\partial V_{\eta^j}}{\partial \eta^i} , \tag{A.17}$$

for $k = 1, 2, 3$. The corresponding covariant components can be obtained by lowering the indices with the metric coefficients g_{ij}, using (A.8). Finally, we end the appendix by recalling how to obtain the contravariant (respectively, covariant) components of a vector $\boldsymbol{V}$ with respect to a curvilinear coordinate system defined by the coordinates (η^1, η^2, η^3) starting from the corresponding contravariant (respectively, covariant) components of the same vector with respect to a different curvilinear coordinate system defined by the coordinates $(\bar{\eta}^1, \bar{\eta}^2, \bar{\eta}^3)$. The contravariant components of $\boldsymbol{V}$ in the two curvilinear coordinate systems are related by

$$V^{\bar{\eta}^i} = \frac{\partial \bar{\eta}^i}{\partial \eta^j} V^{\eta^j} , \tag{A.18}$$

for $i = 1, 2, 3$. Similarly, the covariant components of $\boldsymbol{V}$ in the two curvilinear coordinate systems are related by

$$V_{\bar{\eta}^i} = \frac{\partial \eta^j}{\partial \bar{\eta}^i} V_{\eta^j} , \tag{A.19}$$

for $i = 1, 2, 3$.

Appendix B

Useful analytical integrals

We show here how to compute some recurrent integrals frequently occurring in linear dispersion analyses. Let us first consider the integrals

$$I_i := \int_{-\infty}^{+\infty} \mathrm{d}x \, x^i e^{-x^2} \,, \tag{B.1}$$

with $i \in \mathbb{N}$. If $i = 0$, we have

$$I_0 = \int_{-\infty}^{+\infty} \mathrm{d}x \, e^{-x^2} = \sqrt{\pi} \,. \tag{B.2}$$

If i is odd, an odd integrand function is integrated over a domain symmetric with respect to 0, yielding $I_i = 0$. If i is even, there exists a unique integer $j > 0$ such that $i = 2j$. Therefore, we have

$$I_i = \int_{-\infty}^{+\infty} \mathrm{d}x \, x^{2j} e^{-x^2} \,. \tag{B.3}$$

Due to the parity of the integrand function, we have

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, x^{2j} e^{-x^2} = 2 \int_{0}^{+\infty} \mathrm{d}x \, x^{2j} e^{-x^2} \,. \tag{B.4}$$

Via the change of variable $t := x^2$, we obtain

$$2\int_0^{+\infty} \mathrm{d}x\, x^{2j} e^{-x^2} = \int_0^{+\infty} \frac{\mathrm{d}t}{\sqrt{t}} t^j e^{-t} \,. \tag{B.5}$$

We now define $k := j + 1/2$ and obtain

$$\int_0^{+\infty} \frac{\mathrm{d}t}{\sqrt{t}} t^j e^{-t} = \int_0^{+\infty} \mathrm{d}t\, t^{k-1} e^{-t} = \Gamma(k) \,, \tag{B.6}$$

where Γ is the Euler Gamma function. Summing up, we have

$$I_i = \begin{cases} \sqrt{\pi} & i = 0 \,, \\ 0 & i \text{ odd} \,, \\ \Gamma\left(\dfrac{i+1}{2}\right) & i \text{ even} \,. \end{cases} \tag{B.7}$$

Let us now consider the integrals

$$J_i(z) := \int_{-\infty}^{+\infty} \mathrm{d}x \frac{x^i e^{-x^2}}{x - z} \,, \tag{B.8}$$

for $i \in \mathbb{N}$ and $z \in \mathbb{C}$. If $i > 0$, we have

$$\begin{aligned} J_i(z) &= \int_{-\infty}^{+\infty} \mathrm{d}x\, x^{i-1} e^{-x^2} + z \int_{-\infty}^{+\infty} \mathrm{d}x \frac{x^{i-1} e^{-x^2}}{x - z} \\ &= I_{i-1} + z J_{i-1}(z) \,. \end{aligned} \tag{B.9}$$

If i is even, then $J_i(z) = z J_{i-1}(z)$. If i is odd, we have

$$J_i(z) = \Gamma\left(\frac{i}{2}\right) + z J_{i-1}(z) \,. \tag{B.10}$$

Summing up, for $i > 0$ we have the recursive formula

$$J_i(z) = \begin{cases} \Gamma\left(\dfrac{i}{2}\right) + z J_{i-1}(z) & i \text{ odd} \,, \\ z J_{i-1}(z) & i \text{ even} \,. \end{cases} \tag{B.11}$$

In order to compute the first integral $J_0(z)$,

$$J_0(z) = \int_{-\infty}^{+\infty} \mathrm{d}x \frac{e^{-x^2}}{x-z}\,, \tag{B.12}$$

we introduce the complex function (Barton et al., 1965) (sometimes called Faddeeva function)

$$\begin{aligned} w(z) &:= e^{-z^2}\left(1 + \frac{2i}{\sqrt{\pi}}\int_0^z \mathrm{d}t\, e^{t^2}\right) \\ &=: e^{-z^2}\mathrm{erfc}(-iz) =: e^{-z^2}\left[1 - \mathrm{erf}(-iz)\right]\,, \end{aligned} \tag{B.13}$$

for any $z \in \mathbb{C}$. If $\mathrm{Im}(z) > 0$, the function $w(z)$ has also the integral representation

$$w(z) = \frac{i}{\pi}\int_{-\infty}^{+\infty} \mathrm{d}t \frac{e^{-t^2}}{z-t}\,. \tag{B.14}$$

Therefore, if $\mathrm{Im}(z) > 0$, the integral $J_0(z)$ is given by

$$J_0(z) = i\pi\, w(z)\,. \tag{B.15}$$

If $\mathrm{Im}(z) < 0$, we introduce the new variable $t := -x$ and obtain

$$J_0(z) = \int_{-\infty}^{+\infty} \mathrm{d}x \frac{e^{-x^2}}{x-z} = -\int_{-\infty}^{+\infty} \mathrm{d}t \frac{e^{-t^2}}{t+z} = -i\pi\, w(-z)\,, \tag{B.16}$$

since now $\mathrm{Im}(-z) > 0$. Summing up, the integral $J_0(z)$ is given by

$$\begin{aligned} J_0(z) &= \begin{cases} i\pi\, w(z) & \mathrm{Im}(z) > 0\,, \\ -i\pi\, w(-z) & \mathrm{Im}(z) < 0\,, \end{cases} \\ &= \mathrm{sgn}\left[\mathrm{Im}(z)\right] i\pi\, w\left(\mathrm{sgn}\left[\mathrm{Im}(z)\right] z\right)\,. \end{aligned} \tag{B.17}$$

The derivative of $J_i(z)$ for $i > 0$ can be computed from the recursive formula (B.11), yielding

$$\frac{\mathrm{d}}{\mathrm{d}z}J_i(z) = \sum_{j=1}^{i} z^{j-1} J_{i-j}(z) + z^i \frac{\mathrm{d}}{\mathrm{d}z}J_0(z)\,. \tag{B.18}$$

The derivative of $J_0(z)$ comes then from (B.17), independently of the sign of $\mathrm{Im}(z)$:

$$\frac{\mathrm{d}}{\mathrm{d}z} J_0(z) = i\pi \, \frac{\mathrm{d}}{\mathrm{d}z} w(z) \, . \tag{B.19}$$

Let us now consider the integral

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, \beta_2, \gamma) := \int_{-\infty}^{+\infty} \mathrm{d}x \frac{\left(\sum_{i=0}^{n} \alpha_i x^i\right) e^{-\gamma x^2}}{\beta_0 + \beta_1 x + \beta_2 x^2} \, , \tag{B.20}$$

with $n \in \mathbb{N}$, $\alpha_i \in \mathbb{R}$ for $i = 0, \ldots, n$, $\gamma \in \mathbb{R}$ with $\gamma \neq 0$, $\beta_0 \in \mathbb{C}$ with $\beta_0 \neq 0$, $\beta_1 \in \mathbb{C}$ and $\beta_2 \in \mathbb{R}$. If $\beta_1 = 0$ and $\beta_2 = 0$, we have

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, 0, 0, \gamma) = \frac{1}{\beta_0} \sum_{i=0}^{n} \alpha_i \int_{-\infty}^{+\infty} \mathrm{d}x \, x^i e^{-\gamma x^2} \, . \tag{B.21}$$

We introduce the change of variable $y := \sqrt{\gamma}\, x$ and obtain

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, 0, 0, \gamma) = \frac{1}{\beta_0} \sum_{i=0}^{n} \alpha_i \gamma^{-(i+1)/2} I_i \, . \tag{B.22}$$

If only $\beta_2 = 0$, we have

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, 0, \gamma) = \sum_{i=0}^{n} \alpha_i \int_{-\infty}^{+\infty} \mathrm{d}x \frac{x^i e^{-\gamma x^2}}{\beta_0 + \beta_1 x} \, . \tag{B.23}$$

We first rewrite the integral as

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, 0, \gamma) = \frac{1}{\beta_1} \sum_{i=0}^{n} \alpha_i \int_{-\infty}^{+\infty} \mathrm{d}x \frac{x^i e^{-\gamma x^2}}{x + \beta_0/\beta_1} \, . \tag{B.24}$$

Introducing again the change of variable $y := \sqrt{\gamma}\, x$, we define $z := -\sqrt{\gamma}\, \beta_0/\beta_1$ and obtain

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, 0, \gamma) = \frac{1}{\beta_1} \sum_{i=0}^{n} \alpha_i \gamma^{-i/2} J_i(z) \, . \tag{B.25}$$

If $\beta_2 \neq 0$, we define $\Delta \in \mathbb{C}$ as

$$\Delta := \beta_1^2 - 4\beta_2\beta_0 \, , \tag{B.26}$$

and $u_\pm \in \mathbb{C}$ as

$$u_\pm := \frac{1}{2\beta_2}(-\beta_1 \pm \sqrt{\Delta}) \,. \tag{B.27}$$

We note that

$$\begin{aligned} u_+u_- &= \frac{1}{4\beta_2^2}(\beta_1^2 - \Delta) = \frac{\beta_0}{\beta_2} \,, \\ u_+ + u_- &= -\frac{\beta_1}{\beta_2} \,. \end{aligned} \tag{B.28}$$

Therefore, we have

$$\beta_2 x^2 + \beta_1 x + \beta_0 = \beta_2(x - u_+)(x - u_-) \,. \tag{B.29}$$

The integral can be then rewritten as

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, \beta_2, \gamma) = \frac{1}{\beta_2}\sum_{i=0}^{n} \alpha_i \int_{-\infty}^{+\infty} \mathrm{d}x \frac{x^i e^{-\gamma x^2}}{(x - u_+)(x - u_-)} \,. \tag{B.30}$$

Via the change of variable $y := \sqrt{\gamma} x$, we can write the integral as

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, \beta_2, \gamma) = \frac{1}{\beta_2}\sum_{i=0}^{n} \alpha_i \gamma^{-(i-1)/2} \int_{-\infty}^{+\infty} \mathrm{d}y \frac{y^i e^{-y^2}}{(y - z_+)(y - z_-)} \,, \tag{B.31}$$

where we have defined $z_\pm := \sqrt{\gamma}\, u_\pm$. We now note that

$$\frac{1}{(y - z_+)(y - z_-)} = \frac{1}{z_+ - z_-}\left(\frac{1}{y - z_+} - \frac{1}{y - z_-}\right) \,, \tag{B.32}$$

with $z_+ - z_-$ given by

$$z_+ - z_- = \sqrt{\gamma}(u_+ - u_-) = \sqrt{\gamma}\frac{\sqrt{\Delta}}{\beta_2} \,. \tag{B.33}$$

Therefore, we obtain

$$K(\alpha_0, \ldots, \alpha_n, \beta_0, \beta_1, \beta_2, \gamma) = \frac{1}{\sqrt{\Delta}}\sum_{i=0}^{n} \alpha_i \gamma^{-i/2} \left[J_i(z_+) - J_i(z_-)\right] \,. \tag{B.34}$$

List of Publications

1. Edoardo Zoni and Yaman Güçlü. Solving hyperbolic-elliptic problems on singular mapped disk-like domains with the method of characteristics and spline finite elements. Submitted to *Journal of Computational Physics*, 2018 (under second round of review).
2. Edoardo Zoni and Stefan Possanner. Gyrokinetic theory with polynomial transforms: a model for ions and electrons in maximal ordering. Manuscript in preparation, 2019.

Bibliography

L. Ahlfors. *Complex Analysis : An Introduction to The Theory of Analytic Functions of One Complex Variable*. McGraw-Hill, 1979. ISBN 9780070006577.

P. Angelino, A. Bottino, R. Hatzky, S. Jolliet, O. Sauter, T.M. Tran, and L. Villard. On the definition of a kinetic equilibrium in global gyrokinetic simulations. *Physics of Plasmas*, 13(5):052304, 2006. ISSN 1070-664X. doi:10.1063/1.2193947.

L.A. Artsimovich. Tokamak devices. *Nuclear Fusion*, 12(2):215–252, 1972. ISSN 0029-5515. doi:10.1088/0029-5515/12/2/012.

A. Baños. The guiding centre approximation in lowest order. *Journal of Plasma Physics*, 1(3):305–316, 1967. ISSN 0022-3778. doi:10.1017/S0022377800003317.

D.E. Barton, M. Abramovitz, and I.A. Stegun. Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables. *Journal of the Royal Statistical Society. Series A (General)*, 128(4):593, 1965. ISSN 00359238. doi:10.2307/2343473.

N. Besse and E. Sonnendrücker. Semi-Lagrangian schemes for the Vlasov equation on an unstructured mesh of phase space. *Journal of Computational Physics*, 191 (2):341–376, 2003. ISSN 00219991. doi:10.1016/S0021-9991(03)00318-8.

M.T. Bettencourt. Controlling Self-Force for Unstructured Particle-in-Cell (PIC) Codes. *IEEE Transactions on Plasma Science*, 42(5):1189–1194, 2014. ISSN 0093-3813. doi:10.1109/TPS.2014.2313515.

M. Bostan. Gyrokinetic Vlasov Equation in Three Dimensional Setting. Second Order Approximation. *Multiscale Modeling & Simulation*, 8(5):1923–1957, 2010. ISSN 1540-3459. doi:10.1137/090777621.

A. Bottino, B. Scott, S. Brunner, B.F. McMillan, T.M. Tran, T. Vernay, L. Villard, S. Jolliet, R. Hatzky, and A.G. Peeters. Global Nonlinear Electromagnetic Simulations of Tokamak Turbulence. *IEEE Transactions on Plasma Science*, 38(9):2129–2135, 2010. ISSN 0093-3813. doi:10.1109/TPS.2010.2055583.

N. Bouzat, C. Bressan, V. Grandgirard, G. Latu, and M. Mehrenberger. Targeting Realistic Geometry in Tokamak Code Gysela. *ESAIM: Proceedings and Surveys*, 63:179–207, 2018. ISSN 2267-3059. doi:10.1051/proc/201863179.

A.J. Brizard. Nonlinear gyrokinetic Maxwell-Vlasov equations using magnetic coordinates. *Journal of Plasma Physics*, 41(3):541–559, 1989. ISSN 0022-3778. doi:10.1017/S0022377800014070.

A.J. Brizard. Variational principle for the parallel-symplectic representation of electromagnetic gyrokinetic theory. *Physics of Plasmas*, 24(8):081201, 2017. ISSN 1070-664X. doi:10.1063/1.4997484.

A.J. Brizard and T.S. Hahm. Foundations of nonlinear gyrokinetic theory. *Reviews of Modern Physics*, 79(2):421–468, 2007. ISSN 0034-6861. doi:10.1103/RevModPhys.79.421.

J. Cary. Lie transform perturbation theory for Hamiltonian systems. *Physics Reports*, 79(2):129–159, 1981. ISSN 03701573. doi:10.1016/0370-1573(81)90175-7.

J.R. Cary and R.G. Littlejohn. Noncanonical Hamiltonian mechanics and its application to magnetic field line flow. *Annals of Physics*, 151(1):1–34, 1983. ISSN 00034916. doi:10.1016/0003-4916(83)90313-5.

A. Casati, T. Gerbaud, P. Hennequin, C. Bourdelle, J. Candy, F. Clairet, X. Garbet, V. Grandgirard, Ö.D. Gürcan, S. Heuraux, G.T. Hoang, C. Honoré, F. Imbeaux, R. Sabot, Y. Sarazin, L. Vermare, and R.E. Waltz. Turbulence in the TORE SUPRA Tokamak: Measurements and Validation of Nonlinear Simulations. *Physical Review Letters*, 102(16):165005, 2009. ISSN 0031-9007. doi:10.1103/PhysRevLett.102.165005.

P. Chartier, N. Crouseilles, and M. Lemou. Averaging of highly-oscillatory transport equations. 2016. URL `http://arxiv.org/abs/1609.09819`.

C.Z. Cheng and G. Knorr. The integration of the vlasov equation in configuration space. *Journal of Computational Physics*, 22(3):330–351, 1976. ISSN 00219991. doi:10.1016/0021-9991(76)90053-X.

R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Mathematische Annalen*, 100(1):32–74, 1928. ISSN 0025-5831. doi:10.1007/BF01448839.

N. Crouseilles, M. Mehrenberger, and E. Sonnendrücker. Conservative semi-Lagrangian schemes for Vlasov equations. *Journal of Computational Physics*, 229(6):1927–1953, 2010. ISSN 00219991. doi:10.1016/j.jcp.2009.11.007.

N. Crouseilles, A. Ratnani, and E. Sonnendrücker. An Isogeometric Analysis approach for the study of the gyrokinetic quasi-neutrality equation. *Journal of Computational Physics*, 231(2):373–393, 2012. ISSN 00219991. doi:10.1016/j.jcp.2011.09.004.

O. Czarny and G. Huysmans. Bézier surfaces and finite elements for MHD simulations. *Journal of Computational Physics*, 227(16):7423–7445, 2008. ISSN 00219991. doi:10.1016/j.jcp.2008.04.001.

R.C. Davidson. *Physics of Nonneutral Plasmas*. Co-published with WORLD SCIENTIFIC PUBLISHING CO, 2001. ISBN 978-1-86094-302-7. doi:10.1142/p251.

C. de Boor. A Practical Guide to Splines. *Mathematics of Computation*, 34(149): 325, 1980. ISSN 00255718. doi:10.2307/2006241.

W.D. D'haeseleer, W.N.G. Hitchon, J.D. Callen, and J.L. Shohet. *Flux Coordinates and Magnetic Field Structure: A Guide to a Fundamental Tool of Plasma Theory*. Springer-Verlag, 1991. ISBN 3540524193.

C. Di Troia. From the orbit theory to a guiding center parametric equilibrium distribution function. *Plasma Physics and Controlled Fusion*, 54(10):105017, 2012. ISSN 0741-3335. doi:10.1088/0741-3335/54/10/105017.

G. Dif-Pradalier, V. Grandgirard, Y. Sarazin, X. Garbet, and P. Ghendrih. Defining an equilibrium state in global full-f gyrokinetic models. *Communications in Nonlinear Science and Numerical Simulation*, 13(1):65–71, 2008a. ISSN 10075704. doi:10.1016/j.cnsns.2007.05.004.

G. Dif-Pradalier, V. Grandgirard, Y. Sarazin, X. Garbet, P. Ghendrih, and P. Angelino. On the influence of initial state on gyrokinetic simulations. *Physics of Plasmas*, 15(4):042315, 2008b. ISSN 1070664X. doi:10.1063/1.2904901.

A.M. Dimits, G. Bateman, M.A. Beer, B.I. Cohen, W. Dorland, G.W. Hammett, C. Kim, J.E. Kinsey, M. Kotschenreuther, A.H. Kritz, L.L. Lao, J. Mandrekas, W.M. Nevins, S.E. Parker, A.J. Redd, D.E. Shumaker, R. Sydora, and J. Weiland. Comparisons and physics basis of tokamak transport models and turbulence simulations. *Physics of Plasmas*, 7(3):969–983, 2000. ISSN 1070-664X. doi:10.1063/1.873896.

M.A. Dorf, R.H. Cohen, J.C. Compton, M. Dorr, T.D. Rognlien, J. Angus, S. Krasheninnikov, P. Colella, D. Martin, and P. McCorquodale. Progress with the COGENT Edge Kinetic Code: Collision Operator Options. *Contributions to Plasma Physics*, 52(5-6):518–522, 2012. ISSN 08631042. doi:10.1002/ctpp.201210042.

C.F. Driscoll, D.Z. Jin, D.A. Schecter, and D.H.E. Dubin. Vortex dynamics of 2D electron plasmas. *Physica C: Superconductivity*, 369(1-4):21–27, 2002. ISSN 09214534. doi:10.1016/S0921-4534(01)01216-3.

D.H.E. Dubin. Nonlinear gyrokinetic equations. *Physics of Fluids*, 26(12):3524, 1983. ISSN 00319171. doi:10.1063/1.864113.

D.H.E. Dubin and T.M. O'Neil. Two-dimensional guiding-center transport of a pure electron plasma. *Physical Review Letters*, 60(13):1286–1289, 1988. ISSN 0031-9007. doi:10.1103/PhysRevLett.60.1286.

S. Ethier, W.M. Tang, and Z. Lin. Gyrokinetic particle-in-cell simulations of plasma microturbulence on advanced computing platforms. *Journal of Physics: Conference Series*, 16(1):1–15, 2005. ISSN 1742-6588. doi:10.1088/1742-6596/16/1/001.

G.E. Farin. *Curves and surfaces for computer aided geometric design : a practical guide*. Academic Press, 1988. ISBN 9781483296999.

F. Filbet, E. Sonnendrücker, and P. Bertrand. Conservative Numerical Schemes for the Vlasov Equation. *Journal of Computational Physics*, 172(1):166–187, 2001. ISSN 00219991. doi:10.1006/jcph.2001.6818.

R. Fjørtoft. On a Numerical Method of Integrating the Barotropic Vorticity Equation. *Tellus*, 4(3):179–194, aug 1952. ISSN 00402826. doi:10.1111/j.2153-3490.1952.tb01003.x.

R. Fjørtoft. On the Use of Space-Smoothing in Physical Weather Forecasting. *Tellus*, 7(4):462–480, 1955. ISSN 00402826. doi:10.1111/j.2153-3490.1955.tb01185.x.

E. Frénod and E. Sonnendrücker. The Finite Larmor Radius Approximation. *SIAM Journal on Mathematical Analysis*, 32(6):1227–1247, 2001. ISSN 0036-1410. doi:10.1137/S0036141099364243.

E.A. Frieman and L. Chen. Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *Physics of Fluids*, 25(3):502, 1982. ISSN 00319171. doi:10.1063/1.863762.

R.R.J. Gagné and M.M. Shoucri. A splitting scheme for the numerical solution of a one-dimensional Vlasov equation. *Journal of Computational Physics*, 24(4): 445–449, 1977. ISSN 00219991. doi:10.1016/0021-9991(77)90032-8.

R. Ganesh and J.K. Lee. Formation of quasistationary vortex and transient hole patterns through vortex merger. *Physics of Plasmas*, 9(11):4551–4559, 2002. ISSN 1070-664X. doi:10.1063/1.1513154.

X. Garbet, Y. Idomura, L. Villard, and T.H. Watanabe. Gyrokinetic simulations of turbulent transport. *Nuclear Fusion*, 50(4):043002, 2010. ISSN 0029-5515. doi:10.1088/0029-5515/50/4/043002.

W.J. Gordon and R.F. Riesenfeld. B-SPLINE CURVES AND SURFACES. In *Computer Aided Geometric Design*, pages 95–126. Elsevier, 1974. ISBN 9780120790500. doi:10.1016/B978-0-12-079050-0.50011-4.

T. Görler, N. Tronko, W.A. Hornsby, A. Bottino, R. Kleiber, C. Norscini, V. Grandgirard, F. Jenko, and E. Sonnendrücker. Intercode comparison of gyrokinetic global electromagnetic modes. *Physics of Plasmas*, 23(7):072503, 2016. ISSN 1070-664X. doi:10.1063/1.4954915.

V. Grandgirard, M. Brunetti, P. Bertrand, N. Besse, X. Garbet, P. Ghendrih, G. Manfredi, Y. Sarazin, O. Sauter, E. Sonnendrücker, J. Vaclavik, and L. Villard. A drift-kinetic Semi-Lagrangian 4D code for ion turbulence simulation. *Journal of Computational Physics*, 217(2):395–423, 2006a. ISSN 00219991. doi:10.1016/j.jcp.2006.01.023.

V. Grandgirard, Y. Sarazin, X. Garbet, G. Dif-Pradalier, P. Ghendrih, N. Crouseilles, G. Latu, E. Sonnendrücker, N. Besse, and P. Bertrand. GYSELA, a full-f global gyrokinetic Semi-Lagrangian code for ITG turbulence simulations. In *AIP Conference Proceedings*, volume 871, pages 100–111. AIP, 2006b. doi:10.1063/1.2404543.

V. Grandgirard, J. Abiteboul, J. Bigot, T. Cartier-Michaud, N. Crouseilles, G. Dif-Pradalier, C. Ehrlacher, D. Esteve, X. Garbet, P. Ghendrih, G. Latu, M. Mehrenberger, C. Norscini, C. Passeron, F. Rozar, Y. Sarazin, E. Sonnendrücker, A. Strugarek, and D. Zarzoso. A 5D gyrokinetic full-f global semi-Lagrangian code for flux-driven ion turbulence simulations. *Computer Physics Communications*, 207: 35–68, 2016. ISSN 00104655. doi:10.1016/j.cpc.2016.05.007.

C.M. Greenfield, J.C. DeBoo, T.H. Osborne, F.W. Perkins, M.N. Rosenbluth, and D. Boucher. Enhanced fusion performance due to plasma shape modification of simulated ITER discharges in DIII-D. *Nuclear Fusion*, 37(9):1215–1228, 1997. ISSN 0029-5515. doi:10.1088/0029-5515/37/9/I03.

Y. Güçlü, A.J. Christlieb, and W.N.G. Hitchon. Arbitrarily high order Convected Scheme solution of the Vlasov-Poisson system. *Journal of Computational Physics*, 270:711–752, 2014. ISSN 10902716. doi:10.1016/j.jcp.2014.04.003.

H. Guillard, J. Lakhlili, A. Loseille, A. Loyer, B. Nkonga, A. Ratnani, and A. Elarif. Tokamesh : A software for mesh generation in Tokamaks. 2018. URL `https://hal.inria.fr/hal-01948060/`.

T.S. Hahm. Nonlinear gyrokinetic equations for tokamak microturbulence. *Physics of Fluids*, 31(9):2670–2673, 1988. ISSN 0031-9171. doi:10.1063/1.866544.

E. Hairer, C. Lubich, and G. Wanner. Geometric numerical integration. Structure-preserving Algorithms for Ordinary Differential Equations. Springer Series in Computational Mathematics. Springer-Verlag, Berlin/Heidelberg, 2012. ISBN 3-540-30663-3. doi:10.1017/cbo9780511995569.008.

F. Hariri and M. Ottaviani. A flux-coordinate independent field-aligned approach to plasma turbulence simulations. *Computer Physics Communications*, 184(11): 2419–2429, 2013. ISSN 00104655. doi:10.1016/j.cpc.2013.06.005.

J.A. Heikkinen, S.J. Janhunen, T.P. Kiviniemi, and F. Ogando. Full f gyrokinetic method for particle simulation of tokamak transport. *Journal of Computational Physics*, 227(11):5582–5609, 2008. ISSN 00219991. doi:10.1016/j.jcp.2008.02.013.

M.R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49(6):409, 1952. ISSN 0091-0635. doi:10.6028/jres.049.044.

Y. Idomura, S. Tokuda, and Y. Kishimoto. Global gyrokinetic simulation of ion temperature gradient driven turbulence in plasmas using a canonical Maxwellian distribution. *Nuclear Fusion*, 43(4):234–243, 2003. ISSN 0029-5515. doi:10.1088/0029-5515/43/4/303.

Y. Idomura, M. Ida, T. Kano, N. Aiba, and S. Tokuda. Conservative global gyrokinetic toroidal full-f five-dimensional Vlasov simulation. *Computer Physics Communications*, 179(6):391–403, 2008. ISSN 00104655. doi:10.1016/j.cpc.2008.04.005.

ITER Organization. ITER. URL `https://www.iter.org/`.

F. Jenko and W. Dorland. Nonlinear electromagnetic gyrokinetic simulations of tokamak plasmas. *Plasma Physics and Controlled Fusion*, 43(12A):A141–A150, 2001. ISSN 0741-3335. doi:10.1088/0741-3335/43/12A/310.

P. Kravanja, M. Van Barel, O. Ragos, M.N. Vrahatis, and F.A. Zafiropoulos. ZEAL: A mathematical software package for computing zeros of analytic functions.

Computer Physics Communications, 124(2-3):212–232, 2000. ISSN 00104655. doi:10.1016/S0010-4655(99)00429-4.

T.N. Krishnamurti. Numerical Integration of Primitive Equations by a Quasi-Lagrangian Advective Scheme. *Journal of Applied Meteorology*, 1(4):508–521, 1962. ISSN 0021-8952. doi:10.1175/1520-0450(1962)001<0508:NIOPEB>2.0.CO;2.

J.A. Krommes. The Gyrokinetic Description of Microturbulence in Magnetized Plasmas. *Annual Review of Fluid Mechanics*, 44(1):175–201, 2012. ISSN 0066-4189. doi:10.1146/annurev-fluid-120710-101223.

S. Ku, C.S. Chang, and P.H. Diamond. Full-f gyrokinetic particle simulation of centrally heated global ITG turbulence from magnetic axis to edge pedestal top in a realistic tokamak geometry. *Nuclear Fusion*, 49(11):115021, 2009. ISSN 0029-5515. doi:10.1088/0029-5515/49/11/115021.

J. Kwon, D. Yi, X. Piao, and P. Kim. Development of semi-Lagrangian gyrokinetic code for full-f turbulence simulation in general tokamak geometry. *Journal of Computational Physics*, 283:518–540, 2015. ISSN 00219991. doi:10.1016/j.jcp.2014.12.017.

G. Latu, M. Mehrenberger, Y. Güçlü, M. Ottaviani, and E. Sonnendrücker. Field-Aligned Interpolation for Semi-Lagrangian Gyrokinetic Simulations. *Journal of Scientific Computing*, 74(3):1601–1650, 2018. ISSN 0885-7474. doi:10.1007/s10915-017-0509-5.

C.E. Leith. Lagrangian Advection in an Atmospheric Model. Technical report, 1964.

R. H. Levy. Diocotron Instability in a Cylindrical Geometry. *Physics of Fluids*, 8(7): 1288, 1965. ISSN 00319171. doi:10.1063/1.1761400.

P.C. Liewer. Measurements of microturbulence in tokamaks and comparisons with theories of turbulence and anomalous transport. *Nuclear Fusion*, 25(5):543–621, 1985. ISSN 0029-5515. doi:10.1088/0029-5515/25/5/004.

R.G. Littlejohn. A guiding center Hamiltonian: A new approach. *Jour-*

nal of Mathematical Physics, 20(12):2445–2458, 1979. ISSN 0022-2488. doi:10.1063/1.524053.

R.G. Littlejohn. Variational principles of guiding centre motion. *Journal of Plasma Physics*, 29(1):111–125, 1983. ISSN 0022-3778. doi:10.1017/S002237780000060X.

F.E. Low. A Lagrangian formulation of the Boltzmann-Vlasov equation for plasmas. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 248(1253):282–287, 1958. ISSN 2053-9169. doi:10.1098/rspa.1958.0244.

Max Planck Institute for Plasma Physics. ASDEX Upgrade. URL `https://www.ipp.mpg.de/16195/asdex`.

E.H. Neville. Iterative Interpolation. *J. Indian Math. Soc.*, 20:87–120, 1934.

T.G. Northrop. Adiabatic Charged-Particle Motion. Technical report, 1963.

T.M. O'Neil. New Theory of Transport Due to Like-Particle Collisions. *Physical Review Letters*, 55(9):943–946, 1985. ISSN 0031-9007. doi:10.1103/PhysRevLett.55.943.

M. Ottaviani. An alternative approach to field-aligned coordinates for plasma turbulence simulations. *Physics Letters A*, 375(15):1677–1685, 2011. ISSN 03759601. doi:10.1016/j.physleta.2011.02.069.

F. I. Parra and I. Calvo. Phase-space Lagrangian derivation of electrostatic gyrokinetics in general geometry. *Plasma Physics and Controlled Fusion*, 53(4), 2011. ISSN 07413335. doi:10.1088/0741-3335/53/4/045001.

S. Possanner. Gyrokinetics from variational averaging: Existence and error bounds. *Journal of Mathematical Physics*, 59(8):082702, 2018. ISSN 0022-2488. doi:10.1063/1.5018354.

D.K. Purnell. Solution of the Advective Equation by Upstream Interpolation with a Cubic Spline. *Monthly Weather Review*, 104(1):42–48, 1976. ISSN 0027-0644. doi:10.1175/1520-0493(1976)104<0042:SOTAEB>2.0.CO;2.

A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathmatics*. Springer, 2006. ISBN 9783540346586.

A. Ratnani, Y. Güçlü, J. Lakhlili, and S. Hadjout. Automated symbolic and HPC isogeometric analysis using Python. *Manuscript in preparation*, 2019.

J.S. Sawyer. A semi-Lagrangian method of solving the vorticity advection equation. *Tellus*, 15(4):336–342, 1963. ISSN 0040-2826. doi:10.3402/tellusa.v15i4.8862.

D.A. Schecter and D.H.E. Dubin. Vortex Motion Driven by a Background Vorticity Gradient. *Physical Review Letters*, 83(11):2191–2194, 1999. ISSN 0031-9007. doi:10.1103/PhysRevLett.83.2191.

D.A. Schecter and D.H.E. Dubin. Theory and simulations of two-dimensional vortex motion driven by a background vorticity gradient. *Physics of Fluids*, 13(6):1704–1723, 2001. ISSN 1070-6631. doi:10.1063/1.1359763.

D.A. Schecter, D.H.E. Dubin, K.S. Fine, and C.F. Driscoll. Vortex crystals from 2D Euler flow: Experiment and simulation. *Physics of Fluids*, 11(4):905–914, 1999. ISSN 1070-6631. doi:10.1063/1.869961.

B.D. Scott. Gyrokinetic Field Theory as a Gauge Transform or: gyrokinetic theory without Lie transforms. 2017. URL `http://arxiv.org/abs/1708.06265`.

M. Sengupta and R. Ganesh. Inertia driven radial breathing and nonlinear relaxation in cylindrically confined pure electron plasma. *Physics of Plasmas*, 21(2):022116, 2014. ISSN 1070-664X. doi:10.1063/1.4866022.

M. Sengupta and R. Ganesh. Linear and nonlinear evolution of the ion resonance instability in cylindrical traps: A numerical study. *Physics of Plasmas*, 22(7):072112, 2015. ISSN 1070-664X. doi:10.1063/1.4927126.

E.L. Shi, G.W. Hammett, T. Stoltzfus-Dueck, and A. Hakim. Gyrokinetic continuum simulation of turbulence in a straight open-field-line plasma. *Journal of Plasma Physics*, 83(3):905830304, 2017. ISSN 0022-3778. doi:10.1017/S002237781700037X.

A.C.C. Sips, for the Steady State Operation, and the Transport Physics topical group Activity. Advanced scenarios for ITER operation. *Plasma Physics and Con-*

trolled Fusion, 47(5A):A19–A40, 2005. ISSN 0741-3335. doi:10.1088/0741-3335/47/5A/003.

E. Sonnendrücker, J. Roche, P. Bertrand, and A. Ghizzo. The Semi-Lagrangian Method for the Numerical Resolution of the Vlasov Equation. *Journal of Computational Physics*, 149(2):201–220, 1999. ISSN 00219991. doi:10.1006/jcph.1998.6148.

L. Spitzer. The Stellarator Concept. *Physics of Fluids*, 1(4):253, 1958. ISSN 00319171. doi:10.1063/1.1705883.

A. Staniforth and J. Côté. Semi-Lagrangian Integration Schemes for Atmospheric Models—A Review. *Monthly Weather Review*, 119(9):2206–2223, 1991. ISSN 0027-0644. doi:10.1175/1520-0493(1991)119<2206:SLISFA>2.0.CO;2.

H. Sugama. Gyrokinetic field theory. *Physics of Plasmas*, 7(2):466–480, 2000. ISSN 1070-664X. doi:10.1063/1.873832.

E. Süli and D.F. Mayers. *An Introduction to Numerical Analysis*. Cambridge University Press, Cambridge, 2003. ISBN 9780511801181. doi:10.1017/CBO9780511801181.

T. Takeda and S. Tokuda. Computation of MHD equilibrium of tokamak plasma. *Journal of Computational Physics*, 93(1):1–107, 1991. ISSN 00219991. doi:10.1016/0021-9991(91)90074-U.

B.N. Taylor. The international system of units (SI). Technical report, National Institute of Standards and Technology, 2008.

D. Toshniwal, H. Speleers, R.R. Hiemstra, and T.J.R. Hughes. Multi-degree smooth polar splines: A framework for geometric modeling and isogeometric analysis. *Computer Methods in Applied Mechanics and Engineering*, 316:1005–1061, 2017. ISSN 00457825. doi:10.1016/j.cma.2016.11.009.

N. Tronko and C. Chandre. Second-order nonlinear gyrokinetic theory: from the particle to the gyrocentre. *Journal of Plasma Physics*, 84(3):925840301, 2018. ISSN 0022-3778. doi:10.1017/S0022377818000430.

N. Tronko, A. Bottino, and E. Sonnendrücker. Second order gyrokinetic theory for particle-in-cell codes. *Physics of Plasmas*, 23(8):082505, 2016. ISSN 1070-664X. doi:10.1063/1.4960039.

N. Tronko, A. Bottino, C. Chandre, and E. Sonnendrücker. Hierarchy of second order gyrokinetic Hamiltonian models for particle-in-cell codes. *Plasma Physics and Controlled Fusion*, 59(6):064008, 2017a. ISSN 0741-3335. doi:10.1088/1361-6587/aa68af.

N. Tronko, A. Bottino, T. Görler, E. Sonnendrücker, D. Told, and L. Villard. Verification of Gyrokinetic codes: Theoretical background and applications. *Physics of Plasmas*, 24(5):056115, 2017b. ISSN 1070-664X. doi:10.1063/1.4982689.

W.X. Wang, Z. Lin, W.M. Tang, W.W. Lee, S. Ethier, J.L.V. Lewandowski, G. Rewoldt, T.S. Hahm, and J. Manickam. Gyro-kinetic simulation of global turbulent transport properties in tokamak experiments. *Physics of Plasmas*, 13(9):092505, 2006. ISSN 1070-664X. doi:10.1063/1.2338775.

C. Wersal, A. Bottino, P. Angelino, and B.D. Scott. Fluid moments and spectral diagnostics in global particle-in-cell simulations. *Journal of Physics: Conference Series*, 401(1):012025, 2012. ISSN 1742-6588. doi:10.1088/1742-6596/401/1/012025.

A. Wiin-Nielsen. On the Application of Trajectory Methods in Numerical Forecasting. *Tellus*, 11(2):180–196, 1959. ISSN 0040-2826. doi:10.3402/tellusa.v11i2.9300.

A.J. Wootton, B.A. Carreras, H. Matsumoto, K. McGuire, W.A. Peebles, C.P. Ritz, P.W. Terry, and S.J. Zweben. Fluctuations and anomalous transport in tokamaks. *Physics of Fluids B: Plasma Physics*, 2(12):2879–2903, 1990. ISSN 0899-8221. doi:10.1063/1.859358.

T. Xiong, G. Russo, and J. Qiu. High Order Multi-dimensional Characteristics Tracing for the Incompressible Euler Equation and the Guiding-Center Vlasov Equation. *Journal of Scientific Computing*, 77(1):263–282, 2018. ISSN 0885-7474. doi:10.1007/s10915-018-0705-y.