

Modelling and Optimising Turbulence in 3D Magnetic Geometries for Enhanced Microstability



Georgia O. Acton
Merton College
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2025

Abstract

This thesis investigates microturbulence in three-dimensional magnetised plasma geometries, with a specific focus on the impact of turbulence to plasma stability. Microturbulence is the small-scale turbulent fluctuations in a plasma caused by instabilities at or near the ion and electron gyroradius scales. Microturbulence remains a critical challenge within the field of plasma physics, and in order to control the effects of turbulence, we first need to accurately model it. This task is particularly complex in non-axisymmetric geometries, where mode-coupling occurs across non-identical field lines, leading to potential inconsistencies in the common representation of zonal flows.

The first part of the thesis focuses on the development and validation of a novel δf -gyrokinetic code, designed to model microturbulence across a full flux-annulus in non-axisymmetric magnetic configurations. This is especially important when considering turbulent dynamics in stellarators and tokamaks with 3D magnetic perturbations, as capturing the coupling of modes across different field lines is necessary to accurately resolve ‘zonal’ modes, that are constant across a flux surface. The code is benchmarked against the current flux-tube implementation of the **stella** code, along with other existing gyrokinetic codes. Novel results incorporating kinetic electrons are also presented.

The second part of the thesis addresses microstability, which poses unique optimisation challenges due to the large number of tunable parameters in magnetic confinement devices. To tackle this problem, an efficient method for calculating the derivative of the linear growth rate with respect to multiple externally-controllable parameters, has been developed using an adjoint method. The theoretical framework is derived in the limit of no field-line coupling, including electromagnetic effects and collisions. Numerical simulations then demonstrate the efficacy of the approach in the electrostatic, collisionless regime.

Acknowledgements

While countless individuals deserve my gratitude, space does not allow for me to mention them all. Over the last decade at Oxford, I have encountered many remarkable individuals who have left a lasting impression on my life. I am deeply thankful to everyone who has shaped my personal and academic journey during this time. I would especially like to thank and acknowledge the physics tutors at Merton College, in particular Alex Scheckochihin, who unknowingly inspired me to continue in my pursuit of theoretical physics through particularly difficult times, and sparked my initial interest in fusion. Thank you to all who have taught me, and graciously entertained my endless questions.

I extend my heartfelt thanks to my supervisors, Michael Barnes and Sarah Newton, for tolerating more discussions about fluffy cows and cycling than any supervisor in theoretical physics should have to endure. Your wisdom, encouragement, and support has motivated me to persevere, even when faced with the unfortunate mishaps of supercomputers flooding and catching fire. I am also incredibly grateful for the input and insightful discussions with so many people along the way, including John Omotani, Hanne Theinpondt, Michael Hardman, Peter Catto, William Dorland, and Felix Parra. I would also like to extend thanks to colleagues and friends who have lifted me up during challenging moments—Mantas Abazorius, William Clarke, Megan Evans, Plamen Ivanov, Lucas McConnell, Michael Nastac, Juan Ruiz Ruiz, Leonard Turica, and Yujia Zhang.

This journey would have been impossible without the kindness and support of countless friends. I owe eternal gratitude to Toby Adkins, Becky Strain, Robert Ewart, Catherine Felce, and Esther Hung for their unwavering friendship and encouragement. I'd also like to thank Oxford University Cycling Club (OUCC) for offering a much-needed escape from the world of physics, into the mountains. I have made lifelong friends and created countless cherished memories.

My biggest thanks, and deepest appreciation goes to my family. To my parents—thank you for your patience, understanding, and unconditional love. Your unwavering support has carried me through difficult times, and I am endlessly grateful. A special thank you to my sister for your empathy and friendship. You are truly the best sister I could ever ask for, and I love you deeply.

Lastly, to Andrea—thank you for bringing more joy into my life than I had ever imagined possible. You are my best friend, and I cannot wait to share all of life's adventures with you. Thank you for making me a better person, and for making my life infinitely better; Andromeda and back isn't far enough!

Contents

1	Introduction	1
1.1	Motivation	3
1.1.1	Linear Instabilities	4
1.1.2	Zonal Flows	6
1.1.3	Modelling Plasma Turbulence	7
1.2	Plasma Parameters and Assumptions	8
1.3	Overview	9
2	Plasma Description	10
2.1	Introduction	10
2.2	Magnetic Geometry	11
2.2.1	Axisymmetric Geometries	13
2.2.1.1	Grad–Shafranov Equation	15
2.2.1.2	Miller Geometry	16
2.2.2	Non-Axisymmetric Geometries	17
2.3	Plasma Evolution Equations	19
2.3.1	System Equations	24
2.3.2	Local Approximation	27
2.3.3	Electrostatic and Boltzmann Electron Response Limit	29
2.4	Concluding Remarks	30
3	Gyrokinetic Modelling using a δf code	32
3.1	Introduction	32
3.2	Coordinates	35
3.3	Normalisations	36
3.4	Flux Tube Equations in stella	39
3.5	Boundary Conditions	41
3.5.1	Parallel Boundary Condition; twist-and-shift	41
3.5.2	Extended z -domain	43
3.6	Time Advance Algorithms	45

3.6.1	Operator Splitting	46
3.6.2	Explicit Time Advance	47
3.6.3	Mirror Advance	48
3.6.4	Parallel Streaming	49
4	Full Flux-Annulus δf-Gyrokinetic Code	52
4.1	Introduction	52
4.2	Numerical Implementation	53
4.2.1	Coordinates	54
4.2.2	Normalisation	55
4.2.3	Fourier Transform	56
4.2.4	Gyroaverage	56
4.3	Time Advance Algorithms: full flux-annulus	57
4.3.1	Explicit Time Advance	57
4.3.2	Mirror Advance	58
4.3.3	Intermediate Field Solve	58
4.3.4	Parallel Streaming	59
4.3.5	Boltzmann Electrons	64
4.4	Numerical Results	65
4.4.1	Expectation	66
4.4.2	Axisymmetric Benchmark Cases	68
4.4.2.1	Miller Geometry: kinetic electron results	69
4.4.2.2	VMEC Geometry: ρ_* convergence	73
4.4.3	Non-Axisymmetric Benchmark Case	74
4.4.3.1	Linear Results with Boltzmann Electrons	74
4.4.3.2	Nonlinear Results with Boltzmann Electrons	77
4.4.3.3	Numerical Cost	83
4.4.4	Nonlinear Results with Kinetic Electron Results	84
5	Optimisation of Linear Microstability Using an Adjoint Method	89
5.1	Introduction	89
5.2	Adjoint Method Introduction and Procedure	92
5.3	Example: the Heat Conduction Equation	93
5.3.1	Gradient Computation via the Adjoint Method	94
5.4	Adjoint Method Introduction and Procedure	96
5.5	Adjoint Method for Gyrokinetics	99
5.5.1	General Formalism	99

5.5.2	Electrostatic, Collisionless Limit	104
5.6	Normalisations and Magnetic Geometry	104
5.6.1	Gyrokinetic Normalisations	105
5.6.2	Adjoint Normalisation	105
5.6.3	Electrostatic Collisionless Limit	107
5.6.4	Miller Geometry	107
5.7	Numerical Implementation	107
5.7.1	Initial Simulation	108
5.7.2	Adjoint Simulation	109
5.7.3	Optimisation Loop	110
5.8	Numerical Results	111
5.8.1	Initial Benchmark	111
5.8.2	Increasing the Critical Temperature Gradient	113
5.8.3	Negative Triangularity	115
5.9	Numerical Efficiency	118
5.10	Conclusion and Discussion	120
6	Summary and Discussion	122
6.1	Future work	123
	Appendix A: Full Flux-Annulus	125
A.1	Flux-Surface Average	125
A.1.1	Particle Flux	125
A.1.2	Momentum Flux	126
A.1.3	Heat Flux	128
	Appendix B: Adjoint	128
B.1	Decomposition of Operators	129
B.2	Adjoint of Operators	130
B.3	Simplifying Adjoint Equations	131
B.4	Geometry Implementation	132
	Bibliography	133

List of Figures

1.1	Schematic of toroidal ion temperature gradient driven instability . . .	5
1.2	Schematic of zonal flow in a tokamak	6
2.1	Schematic representation of nested flux surfaces	12
2.2	Diagram depicting the Clebsch representation	12
2.3	Axisymmetric flux surface	13
2.4	Magnetic field of an axisymmetric flux surface in Boozer coordinates .	15
2.5	Miller geometry	17
2.6	Miller geometry examples	18
2.7	Non-axisymmetric flux surface	19
2.8	Magnetic field of an axisymmetric flux surface in Boozer coordinates .	19
2.9	Illustration of the approximate motion of a charged particle in a mag- netic field	22
2.10	Illustration of particle completing multiple gyro-orbits and sampling non-uniform magnetic field lines	26
3.1	Schematic of a flux tube for an axisymmetric device	39
3.2	An illustration of the connections in the extended z -domain	45
4.1	Flux annulus domain	55
4.2	Flow chart showing the iteration scheme employed for implicitly ad- vancing the parallel streaming term in the full flux annulus algorithm	63
4.3	Iterative implicit scheme normalised convergence with number of iter- ations	64
4.4	Sketch of the expectation for the growth rate spectrum for an axisym- metric device	67
4.5	Sketch of the expectation for the growth rate spectrum for a non- axisymmetric device	68
4.6	The magnetic geometry for the CBC as considered by the Miller and VMEC implementations	69

4.7	Comparison of the linear growth rate and real frequency spectra against binormal mode number for CBC using Miller geometry with Boltzmann and kinetic electrons	71
4.8	Time trace of the nonlinear heat flux for the flux-tube and flux-annulus simulations for CBC using Miller geometry with kinetic electrons . .	72
4.9	Spectra of the squared electrostatic potential and heat flux for CBC using Miller geometry from nonlinear simulations with kinetic electrons	72
4.10	Scan of linear growth rate and real frequency against inverse ρ_* for CBC with Boltzmann and kinetic electrons	74
4.11	Comparison of linear growth rate and real frequency spectra for the high-mirror W7-X configuration with Boltzmann electrons between the flux-tube code GX and the flux-annulus version of stella	76
4.12	Benchmark of flux-annulus code against other codes	77
4.13	Flux tube geometry for multiple poloidal turns	78
4.14	Time trace of the heat flux from flux-tube simulations with varying parallel domain lengths, compared with the corresponding flux-annulus result	80
4.15	Spectra of the squared electrostatic potential and heat flux from the nonlinear simulations	81
4.16	Zonal potential and zonal flow from the flux-tube and flux-annulus nonlinear Boltzmann electron simulations as a function of the radial position and time	82
4.17	Time trace of ion and electron heat fluxes in W7-X nonlinear simulations with kinetic electrons	85
4.18	Spectra of the squared electrostatic potential and heat flux from the nonlinear simulations with kinetic electrons	85
4.19	Time trace of ion and electron heat fluxes in W7-X nonlinear simulations with kinetic electrons	86
4.20	Spectra of the squared electrostatic potential and heat flux from the second set of nonlinear simulations with kinetic electrons	86
4.21	Zonal potential and zonal flow from the flux-tube and flux-annulus nonlinear kinetic electron simulations as a function of the radial position and time	88
5.1	Linear instability time evolution	91
5.2	Flow chart showing adjoint optimisation loop	108

5.3	Result of adjoint method in two-dimensions; elongation and triangularity	114
5.4	Result of adjoint method with increased temperature gradient	114
5.5	Result of adjoint method for a case with negative triangularity	116
5.6	Result of adjoint method demonstrating the non-uniqueness of a gradient based optimisation	117
5.7	Result of adjoint method for a case with negative triangularity with increased temperature gradient	117
5.8	Poloidal cross section of a flux surface that has been optimised using the adjoint method	119

Chapter 1

Introduction

You can't spell confusion without fusion.

Lucas McConnell

Among various world energy alternatives, thermonuclear fusion—the process that powers the Sun—presents a promising solution to provide clean, safe, and efficient energy. For this process, hydrogen is heated until it ionises, turning into a plasma composed of ions and electrons. These charged particles can be influenced by electric and magnetic fields, motivating the concept of magnetic confinement fusion (MCF), which uses a magnetic cage to confine the plasma, allowing it to be heated to sufficiently high temperatures for fusion to occur. However, to achieve fusion, we require temperatures in excess of one hundred million degrees Celsius due to the Coulomb repulsion between ions. To design a viable device, we must reach these high temperatures in the core while ensuring that the temperature at the edges of our plasma is low enough for the materials in the walls of the device to withstand. This necessitates a temperature gradient between the centre and edges of the device and requires one to set up a profile of the parameters across the plasma. Since the inception of MCF research in the 1940s, substantial progress has been made, primarily focused on the tokamak, which is the most advanced MCF design to date. These toroidally axisymmetric devices rely on a substantial plasma current to produce the poloidal magnetic field required for confinement, making them vulnerable to current-driven instabilities, that can impede plasma performance, or even terminate the discharge.

As research and computational tools have advanced, geometric considerations are being used to constrain the plasma in ways that promote stability. This has application in resonant magnetic perturbations (RMPs) to tokamaks [100], and even the deliberate breaking of axisymmetry when constructing the initial plasma equilibrium. Stellarators are a class of devices designed for the magnetic confinement of

plasmas without a toroidally symmetric equilibrium, and have become a competitive alternative to tokamaks. Rather than relying on a large plasma current, the poloidal magnetic field in a stellarator is generated by external current-carrying coils, which are often complicated and highly shaped [16]. Unlike tokamaks, which rely on inductive currents that require pulsed operation and can lead to disruptions, stellarators do not rely on any internal current for confinement, making them inherently less susceptible to current-driven instabilities and disruptions. Experimental evidence suggests that stellarator configurations may operate above the linear MHD stability pressure threshold expected for tokamaks, and achieve higher densities than those set by the Greenwald limit for equivalent tokamaks [49, 124, 42]. This capability could allow stellarators to function within parameter regimes that are most favourable to fusion.

In contrast to tokamaks, the lack of axisymmetry in stellarators means that they typically exhibit high levels of neoclassical transport at low collisionality [63]. This enhanced transport arises because particle orbits no longer average out to produce zero net radial motion. One can insist on a magnetic geometry such that drift of particles when averaged over time vanishes, a condition known as omnigeneity. This equates to eliminating the net radial drift of collisionless particle trajectories [23, 52], to ensure single-particle confinement. Near-omnigeneous stellarators have been one successful example of plasma shaping in stellarators, enforced through geometric considerations alone. Targeting omnigeneity in non-axisymmetric configurations has been shown to significantly enhance plasma confinement [54]. Recognising the importance of omnigeneity in such configurations has led to substantial progress in both theoretical understanding, and in the optimisation of magnetic fields for improved omnigeneity in experimental designs [55]. Particular subsets of omnigeneity, such as quasi-symmetric configurations, have been demonstrated to offer guiding-centre confinement and neoclassical transport properties comparable to those of equivalent tokamaks [15]. Consequently, there have been considerable advancements in the design and construction of these optimised stellarator devices [64, 59, 6, 77, 101, 46]. Additional advances in coil design has made it easier to build these designs with higher error tolerances [82, 94, 79].

However, despite recent advances and their appealing prospects, the parameter space of devices without exact axisymmetry is vast, and hence numerical optimisation becomes necessary as it is often impractical to explore all alternative possibilities. Additionally, the modelling of core turbulence—which is required to optimise confinement—for non-axisymmetric geometries is significantly more challenging. This is true for both stellarators and tokamaks with magnetic perturbations. Therefore

the ability to accurately model turbulence in generic magnetic field configurations is of broad relevance to the MCF community. The motivation for this thesis centres on accurately modelling the plasma turbulence, and numerically optimising for turbulent energy confinement in configurations with a large parameter space. The aim with both of these endeavors is to improve confinement and overall fusion performance.

1.1 Motivation

To properly predict plasma profiles, one should solve the transport equations in which nonlinear, turbulent fluxes determine profile evolution in the presence of given sources and the boundary conditions. The turbulent fluxes are driven by gradients of temperature and density, which act as sources of free energy that are capable of driving microscale instabilities, should they exceed critical values (see e.g. [105, 98, 37, 114, 103, 50, 28] for ion temperature gradient (ITG) driven instabilities, [36, 37] for microtearing modes (MTMs), [5, 114] for kinetic ballooning mode (KBM) descriptions, and [34, 69] for electron temperature gradient (ETG) driven instabilities respectively). The threshold values of the plasma profile parameters for the onset of linear instabilities are known as the ‘linear critical gradients’; a comprehensive analysis of ITG critical gradients and their onset can be found in Zocco (2018) [125]. Beyond these critical gradients, the linear modes can lead to strong, nonlinear turbulence, with the fluxes rapidly increasing with increasing temperature gradient. This would in turn require a large additional power input to maintain the temperature gradient, and is described as a regime of ‘stiff’ transport [32]. These large fluxes result in the increased transport of particles, momentum, and energy out of the device. Due to this stiff transport, the critical gradients predicted by linear instability analysis often serve as a reasonable first approximation to experimental results [35], since increased heating power tends to have little effect on the plasma profile parameters, which remain pinned near these values. It is important to note that, unlike in neutral fluids, linear instabilities in plasmas often act as the primary drivers of nonlinear turbulence, making them essential to understanding the full nonlinear plasma dynamics.

However, although linear calculations are often good predictors, this picture can be modified in various ways by the nonlinear physics. First, stiffness is not absolute and is not always present, so an understanding of nonlinear saturation may be required to determine achievable gradients for a given input power, particle source and torque. Second, when the fluctuation amplitudes associated with the linear instabilities grow large enough, they trigger ‘secondary’ instabilities that cause ‘zonal’ modes to grow.

The energy injection into the zonal modes, and their associated zonal shear flows can suppress turbulence completely in a region near the linear critical gradients. This shift to the onset of stiff transport, known as the ‘Dimitis shift’ (see e.g., [32, 102]), leads to a ‘nonlinear critical gradient’ that exceeds the linear value. This defines the onset of stiff transport in a fully developed nonlinear system, and is often up-shifted from the linear gradient. Finally, sub-critical turbulence, in which linearly stable systems lead to a finite amplitude turbulent steady-state, also exist¹ [87, 56]. Hence, to reliably predict even the critical gradients for turbulent onset, nonlinear simulations are required. Thus, while linear analysis has its merits, the ability to accurately model nonlinear turbulence and its interactions with zonal flow is crucial for making better predictions about plasma evolution within MCF devices.

1.1.1 Linear Instabilities

The archetypal linear microinstability in the core of MCF plasmas is the toroidal ion temperature gradient (ITG) instability [105, 98, 37, 114, 103, 50, 28]. It has been identified as one of the most prominent and ubiquitous instabilities, being one of the main sources of turbulent heat transport in the core of tokamaks [27, 120, 84]. Its basic features are outlined here, to give a flavour of the physics captured and to use as a case study in stability optimisation. This type of instability occurs on the ‘bad curvature’ side of MCF devices, where the temperature and magnetic field strength gradients align. For an axisymmetric device, this is on the outboard side of the device. A schematic of the instability in the electrostatic limit is shown in Figure 1.1. It is primarily driven by the free energy stored in the ion temperature gradient. Fluctuations in temperature that naturally arise in the plasma lead to density fluctuations, due to the difference in the magnetic drift velocities between colder and hotter ions. This sets up an electric field perturbation, which applies a force to the charged ions. When a force is applied to a charged particle in the presence of a magnetic field, it drifts in a direction that is perpendicular to both the force and the magnetic field. Hence, the electric field perturbation leads to the so-called ‘ $\mathbf{E} \times \mathbf{B}$ drift’ which acts to enhance the initial perturbation, leading to an instability. On the ‘good curvature side’ of the device—where the temperature and magnetic field strength gradients are oppositely aligned—this $\mathbf{E} \times \mathbf{B}$ drift acts to reduce the amplitude of the initial perturbation, stabilising the ITG instability. For an axisymmetric device, this occurs on the inboard side. A similar analysis applies to the electrons, whose magnetic drift is in the opposite direction to ions due to their opposite charge.

¹For the optimisations performed in this work, sub-critical turbulence will not be considered.

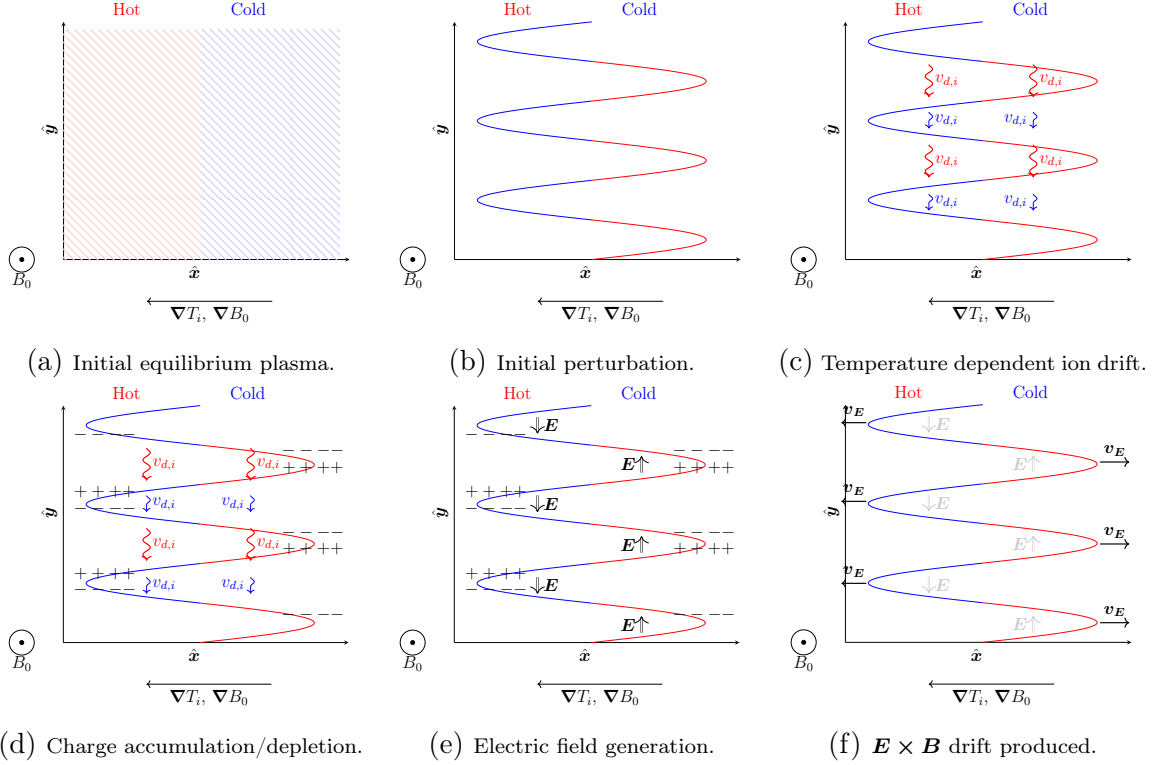


Figure 1.1: Schematic of the toroidal ion temperature gradient (ITG) driven instability. Figure (a) shows the equilibrium plasma, where the gradient of the ion temperature, T_i , and the gradient of equilibrium magnetic field strength, B_0 , act in the same direction. Here the equilibrium magnetic field is directed out of the plane of the page. Figure (b) shows the initial perturbation introduced into the plasma, causing hotter regions of plasma to extend into regions of colder background plasma, and vice versa. In Figure (c) the ions drift with a speed $v_{d,i}$ that depends on their temperature—hotter ions drift faster compared with colder ones. It can be seen in (d) that this leads to accumulation (depletion) of positive charge as ions are drifting into the region faster (slower) than they are drifting out of that region, due to the difference in drift velocities between hot and cold ions. This separation of charge generates a perturbed electric field, $\mathbf{E}$, shown in (e). This electric field generates an $\mathbf{E} \times \mathbf{B}$ drift that acts to reinforce the original perturbation, as shown in (f). This feedback mechanism leads to instability.

The linear growth rates of such instabilities can depend on numerous parameters in the system, such as the magnetic geometry of the device. Given the high-dimensionality of the parameter space involved, optimising a plasma configuration for microstability can prove computationally prohibitive. As such, it is of importance to develop efficient methods allowing for the linear growth rate to be minimised with respect to experimentally-controllable parameters, in order to practically aim for plasma microstability.

1.1.2 Zonal Flows

The gyrokinetic system captures nonlinear processes that determine the turbulent saturation levels and transport within MCF devices. One important example of this is the so-called ‘zonal flow’, which has been shown to significantly impact the evolution of the turbulence [14, 104, 81, 20, 68, 67, 88]. Zonal flow is defined as flow that is constant within a magnetic flux surface, but is able to vary across flux surfaces. It is distinguished from the equilibrium flow in that it is self-organised by the turbulence over the turbulence time-scale and can have a radial variation over scales comparable to turbulent eddy sizes, i.e., the Larmor radius [31]. Depending on the plasma parameters, the zonal flow may be coherent, providing an effective modification to the equilibrium that can affect instability, or can be fluctuating in time. In most gyrokinetic simulations the zonal flow plays a crucial role in determining the amplitudes at which turbulence saturates [111]. A schematic illustration of a zonal flow in a tokamak is shown in Figure 1.2, where the red and blue shaded regions represent bulk plasma flows in the clockwise and counterclockwise directions, respectively.

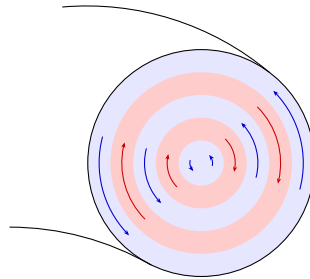


Figure 1.2: A schematic of the poloidal cross-section of a tokamak illustrating the zonal flow structure. The red (blue) bands represent bulk plasma flow in the clockwise (anticlockwise) direction. This flow pattern is consistent across the entire device.

The fact that the zonal flow is constant across all field lines in a flux surface is a challenge for local simulations of non-axisymmetric plasmas [53, 4, 107]. Many of the current simulation approaches model the turbulence locally by considering a domain

that follows the magnetic field line. This domain is elongated in the direction parallel to the magnetic field, but is thin in the directions perpendicular to the magnetic field, and is known as a flux-tube. However, modelling the turbulence locally in a flux-tube may limit the ability to capture the generation, evolution, and effects of zonal flows in non-axisymmetric devices: in general, the zonal flow determined locally, using a flux-tube code, for two different field lines on the same flux surface will not agree. If not properly represented, the dynamics of zonal flows may result in an incomplete or inaccurate depiction of turbulence and transport processes, meaning simulations may be unable to capture the true turbulence saturation levels. Consequently, an efficient means of coupling zonal flow and the corresponding turbulence across multiple field lines is thus crucial for studying turbulent saturation in non-axisymmetric devices. Efficiently capturing all of this numerically can put a large demand on computing resources, potentially requiring very large processing times on parallelised systems.

1.1.3 Modelling Plasma Turbulence

The gyrokinetic system of equations serves as a foundational theoretical framework for analysing magnetised plasmas. Gyrokinetic simulations have become valuable tools within the field for modelling the plasma on turbulent length and time scales in order to predict the associated transport. However, achieving solutions to the full nonlinear, 5-dimensional, integro-differential gyrokinetic equation, incorporating comprehensive physics—such as kinetic electron dynamics, collisionality, and electromagnetic fluctuations—in realistic, non-axisymmetric geometries places substantial demands on computational resources. Consequently, reduced descriptions of the system, which suppress certain physics, have become routine.

The main goal of this thesis will be to discuss novel numerical schemes aimed at addressing these complexities, with the aim to improve the modelling capabilities for plasma dynamics, and help inform the design of future fusion reactors. The research presented in this thesis has two main objectives: the first is to accurately model gyrokinetic turbulence on a flux-annulus in order to capture geometric effects in non-axisymmetric plasmas, particularly in relation to correctly modelling the zonal flow in such devices. The second is to develop tools to explore how controllable parameters of the system can be used to enhance microstability, thereby improving conditions for fusion. This involves the development of an efficient framework for optimising the linear growth rates—which can serve as a proxy for the entire nonlinear system—to achieve microstability by considering a wide range of parameters that can be feasibly adjusted in the design of an MCF device.

1.2 Plasma Parameters and Assumptions

In the presence of magnetic fields, charged particles execute a helical motion around the magnetic field lines. In the absence of an electric field, and in a uniform magnetic field, the perpendicular component of this motion reduces to a simple circular path, with a frequency termed the gyrofrequency, Ω_ν , where ν denotes the species index. Here $\nu = i$ and $\nu = e$ indicate the bulk ion and electron species respectively. The gyrofrequency is given by $\Omega_\nu = Z_\nu e B / m_\nu c$, where B is the magnitude of the magnetic field, m_ν is the species' mass, c is the speed of light in a vacuum, and $Z_\nu e$ is the species charge, with e the proton charge.. The radius of this circular motion is known as the Larmor radius, ρ_ν .

When deriving a model of the plasma behaviour accounting for turbulence, we focus on plasma conditions characterised by weak coupling, strong magnetisation, low frequencies, and small amplitude turbulent fluctuations. We shall detail each of these assumptions:

- **Weak Coupling:** This condition is expressed as $n_e \lambda_{De}^3 \gg 1$, where n_e is the equilibrium electron density. The Debye length, λ_{De} , is defined in cgs units by:

$$\lambda_{De} = \left(\sum_\nu \frac{4\pi n_\nu (Z_\nu e)^2}{T_\nu} \right)^{-\frac{1}{2}}. \quad (1.1)$$

This is a statement that there are numerous particles within a ‘Debye sphere’—a sphere whose radius is the Debye length—meaning that the plasma shields out the effect of charge perturbations over length scales on the order of λ_{De} . This leads to an overall quasineutral plasma, for which the perturbed density of each species, δn_ν , obeys the following relation

$$\sum_\nu Z_\nu \delta n_\nu = 0. \quad (1.2)$$

- **Strong Magnetisation:** The analysis in this thesis is limited to the regime in which the Larmor radius, ρ_ν , is considerably smaller than the characteristic system size, L : $\rho_\nu / L \ll 1$. This implies that the magnetic field influence is dominant over particle motion on macroscopic scales, which is required for plasma confinement.
- **Low Frequencies:** The characteristic frequencies of fluctuations within the plasma are taken to be much smaller than the gyrofrequency, which is a consequence of strong magnetisation. This allows the rapid gyromotion to be temporally separated from the slower dynamics of the system.

- **Near Equilibrium:** It is assumed that the turbulent fluctuations within the plasma are small in amplitude compared to the mean field values, meaning perturbative methods may be used to analyse the plasma behaviour.

1.3 Overview

This thesis aims to address the complex challenge of modelling turbulence in a magnetically confined fusion plasma and optimising for gyrokinetic microstability. The structure of the remainder of this thesis is as follows. In Chapter 2, the foundational framework upon which this research is based is introduced, presenting the gyrokinetic equations along with the associated field equations that govern the system behaviour. Chapter 3 explores the various considerations that must be taken into account when numerically modelling plasma behaviour, specifying the particular choices made for the δf -gyrokinetic code **stella**. Chapter 4 introduces a novel numerical extension of the existing flux-tube version of the **stella** code, to include the capability of modelling turbulence across a full flux-annulus, allowing for a more accurate modelling of turbulence in non-axisymmetric plasmas. The algorithms that have been implemented are given, and numerical benchmarks are provided to demonstrate the efficacy of this work. This code has been developed in order to effectively capture the effects of varying magnetic geometry and accurately model zonal flow, and the resulting differences between flux-tube and the full flux-annulus versions are shown. In Chapter 5, the adjoint method is detailed, which provides an efficient approach for computing gradients in a high-dimensional parameter space. Its application to the linearised gyrokinetic system is demonstrated, highlighting the capability for optimising for microstability within a plasma, and improving plasma profile parameters favorable for fusion. It is demonstrated that such an optimisation can be achieved by varying externally controllable experimental parameters, and that this method is capable of efficiently doing so for a large parameter space. This chapter also includes numerical validation of the proposed theory using the flux-tube version of the **stella** code. The material used in this chapter has been published by Acton et al. (2024) [3]. By correctly modelling the turbulence and optimising for microstability, this work aims to contribute to the global effort to move forward fusion as a viable clean energy source. Finally, Chapter 6 presents a summary of the key work in this thesis and outlines potential directions for future research.

Chapter 2

Plasma Description

[The equations] are generally accepted by most theoreticians, they are developed with the most sophisticated mathematical methods and it is only the plasma itself which does not understand how beautiful the theories are and absolutely refuses to obey them. It is now obvious that we have to start a second approach from widely different starting points.

Hannes Alfvén

2.1 Introduction

In theory, it is possible to use classical physics to solve the equations of motion for every particle in a plasma, and so predict the fluxes of particles, momentum and energy out of an MCF device for given sources. However, storing such information for approximately $\mathcal{O}(10^{23})$ particles is not feasible. Instead, a statistical description is more practical. This involves considering the distribution functions of each particle species, $f_\nu(\mathbf{x}, \mathbf{v}, t)$, which characterises the distribution for a given species ν in terms of spatial position $\mathbf{x}$, velocity $\mathbf{v}$, and time t . To fully determine f_ν , one would need to solve the complete nonlinear Vlasov–Maxwell system of equations, a task that is itself complex. To make further progress, a fundamental small parameter for each species is identified in a strongly magnetised plasma, $\rho_{*,\nu}$, and the separation of spatial and temporal scales inherent in plasma systems, that were outlined in Section 1.2, can be exploited. The first nonlinear gyrokinetic equation was derived using this method by Frieman and Chen in 1982 [41]. There have since been many other works, which have extended the derivation to include generalised magnetic geometries, electromagnetic effects, and the use of gyrocenter coordinates (e.g., see [65, 1, ?, 25]). The derivation covered here closely follows that of Abel (2013) [1] which uses a perturbative expansion in gyrocenter coordinates to derive the full nonlinear gyrokinetic equations,

including electromagnetic effects. These derivations take advantage of the scale separation present in most MCF plasmas, where the plasma evolves across varying spatial and temporal scales. The resistive evolution of the magnetic geometry, for example, evolves over slow temporal and large spatial scales, while turbulent fluctuations appear on shorter temporal and spatial scales.

2.2 Magnetic Geometry

In order to consider how turbulent fluctuations of the plasma evolve, one must first consider the magnetic geometry of the plasma. In MCF devices, achieving a stable plasma equilibrium is essential for operation. Global plasma equilibrium is governed by a combination of Maxwell's equations and the ideal MHD equations. The non-relativistic version of Maxwell's equations are

$$\nabla \cdot \mathbf{E} = 4\pi\varrho, \quad (2.1a)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (2.1b)$$

$$\nabla \times \mathbf{E} = -\frac{1}{c} \frac{\partial \mathbf{B}}{\partial t}, \quad (2.1c)$$

$$\nabla \times \mathbf{B} = \frac{4\pi}{c} \mathbf{j}, \quad (2.1d)$$

where ϱ is the electric charge density, $\mathbf{j}$ is the current density, and c is the speed of light. The fields $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, respectively.

In a steady state, with no sonic flows, the equilibrium condition for the plasma, treated as a single fluid, reduces to simple force balance. The forces on the plasma, at the equilibrium scale, involve the thermal pressure gradient force, the magnetic tension, and the magnetic pressure

$$\frac{1}{c} \mathbf{j} \times \mathbf{B} = \nabla p, \quad (2.2)$$

where p is the plasma pressure. Equation (2.2) illustrates that $\mathbf{B}$ and $\mathbf{j}$ must lie on surfaces of constant pressure, giving rise to the equilibrium conditions $\mathbf{B} \cdot \nabla p = 0$ and $\mathbf{j} \cdot \nabla p = 0$. To ensure that $\mathbf{B}$ remains tangential to a non-vanishing and singularity-free pressure gradient, the magnetic field lines must lie on a torus [118, 52]. Given that $\mathbf{B}$ must lie on surfaces of constant pressure with toroidal topology, we can make the additional assumption for the existence of nested magnetic surfaces of the configuration. In a poloidal cross-section, these nested surfaces resemble a series of concentric magnetic surfaces, illustrated by Figure 2.1. The degenerate surface at the

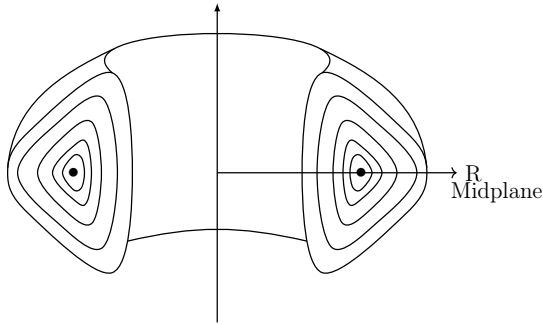


Figure 2.1: Schematic representation of a poloidal cross section through a torus with nested flux surfaces.

center of the nested magnetic surfaces is called the magnetic axis, and is marked by the black dot. The magnetic field $\mathbf{B}$ can be expressed using the Clebsch representation [75]:

$$\mathbf{B} = \nabla\alpha \times \nabla\psi, \quad (2.3)$$

where ψ is the flux surface label, and α labels field lines. A diagram depicting these coordinates on a flux surface can be seen in Figure 2.2.

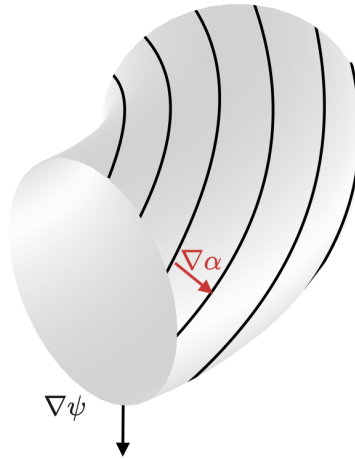


Figure 2.2: Diagram illustrating the Clebsch representation on a single flux surface. The coordinate ψ labels the flux surfaces, and the gradient vector $\nabla\psi$ points in the direction of greatest change in ψ —i.e., normal to the flux surface. The coordinate α labels magnetic field lines, with the gradient $\nabla\alpha$ lying within the flux surface and pointing in the direction of maximum variation in field-line labeling.

These labels help in defining a magnetic coordinate system that aligns well with the geometry of the magnetic surfaces and facilitates the mathematical treatment of

equilibrium and stability analyses [18]. Together with a third coordinate, that characterises the position along the magnetic field, this forms a non-orthogonal curvilinear coordinate system. The choice of these coordinates is not unique, and as such they are currently left unspecified.

2.2.1 Axisymmetric Geometries

Within toroidally axisymmetric geometries, the coordinate labels (θ, ζ, ψ) are commonly used. Here, θ represents the poloidal angle, which measures angular distance along the shorter path around a torus, while ζ represents the toroidal angle, which measures the angular distance along the longer path around the torus. The ψ coordinate denotes the flux surface label. Figure 2.3 illustrates a flux surface of an axisymmetric device, with this choice of coordinates. Here the colours indicate the magnetic field strength—with red signifying the strongest field strength, and blue representing the weakest.

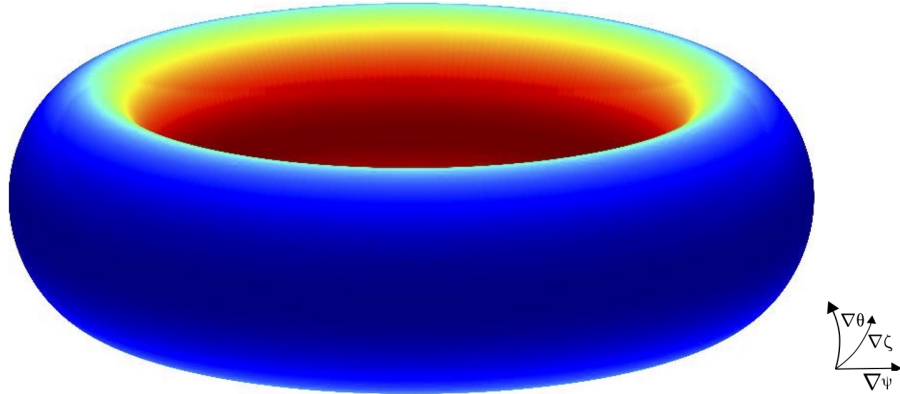


Figure 2.3: An example of an axisymmetric flux surface with the associated coordinate system orientations. Here the colour indicates the strength of the magnetic field, with red denoting a stronger field and blue indicating a weaker field. Image created using stellapy [115].

In general, θ and ζ can be taken as any 2π -periodic coordinates that advance in the poloidal and toroidal directions, respectively. Thus, there are two position-dependent degrees of freedom when choosing (θ, ζ) , as one is always at liberty to redefine them:

$$\begin{aligned}\tilde{\theta} &= \theta + \lambda(\psi, \theta, \zeta), \\ \tilde{\zeta} &= \zeta + \nu(\psi, \theta, \zeta),\end{aligned}\tag{2.4}$$

where λ and ν are any single-valued functions that are 2π -periodic in θ , and ζ [52, 66]. The specification of these functions is a gauge choice, and each gauge results in

analogous ‘poloidal’ and ‘toroidal’ angles that are 2π periodic and refer to the ‘short’ and ‘long’ ways around the torus respectively. This provides freedom to define the coordinate system such that the magnetic field has certain desirable properties. To relate this back to the Clebsch definition of the magnetic field, let $\alpha = \alpha(\theta, \zeta)$:

$$\mathbf{B} = \frac{\partial \alpha}{\partial \theta} \nabla \theta \times \nabla \psi + \frac{\partial \alpha}{\partial \zeta} \nabla \zeta \times \nabla \psi. \quad (2.5)$$

This helical magnetic field in tokamaks is primarily created by the interaction of two components: a toroidal magnetic field generated by external magnets surrounding the plasma vessel, and a poloidal magnetic field produced by a toroidal current within the plasma. In tokamak devices, this toroidal current is typically induced by a central solenoid, which generates a secondary current in the plasma. Additional methods for generating the toroidal current include neutral beam injection, microwave heating, and naturally occurring currents within the plasma itself, driven by the system gradients. The poloidal magnetic field plays a crucial role in mitigating net particle drifts, by linking different vertical positions within the plasma and thereby preventing the formation of a vertical electric field. The toroidal magnetic field is essential to prevent large-scale plasma instabilities. One characteristic of the magnetic field that describes the relative size of the toroidal to poloidal magnetic field is the safety factor, q , which represents the number of toroidal turns a magnetic field line must complete to achieve a full poloidal rotation,

$$q(\psi) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\mathbf{B} \cdot \nabla \zeta}{\mathbf{B} \cdot \nabla \theta} d\theta. \quad (2.6)$$

This safety factor indicates the relative increase of the toroidal angle to the poloidal angle of the field line (i.e., the ‘pitch’ of the field line), with $q = 0$ corresponding to zero toroidal field and $q = \infty$ indicating zero poloidal field. Substituting (2.5) into this definition gives

$$q(\psi) = -\frac{1}{2\pi} \int_0^{2\pi} \frac{\partial \alpha / \partial \theta}{\partial \alpha / \partial \zeta} d\theta. \quad (2.7)$$

Given the degrees of freedom in the choice of θ and ζ , one can enforce that the integrand is θ -independent, giving $\partial \alpha / \partial \theta + q \partial \alpha / \partial \zeta = 0$. This is an advection equation with solutions of the form $\alpha = \alpha(\psi, \zeta - q\theta)$. The choice $\alpha = \zeta - q\theta$ is made, noting that any α of the form $\hat{\alpha}(\psi)(\zeta - q\theta)$ can be manipulated such that $\hat{\alpha}(\psi)$ is absorbed into the definition of ψ . Using $\alpha = \zeta - q\theta$ and defining $-q(\psi)\nabla \theta \times \nabla \psi = I(\psi)\nabla \zeta$, the magnetic field can be rewritten as

$$\mathbf{B} = \nabla \zeta \times \nabla \psi + I(\psi)\nabla \zeta, \quad (2.8)$$

where $I(\psi)$ is a flux function defined by $I(\psi) = \mathbf{B} \cdot \nabla \zeta = RB_\zeta$, which can be interpreted as the current inside the flux surface ψ . This coordinate choice results in a field-aligned coordinate system where field lines are straight in the (θ, ζ) plane. The remaining degree of freedom allows for tailoring desirable qualities of the magnetic field, which gives rise to different coordinate choices. One such choice is Boozer coordinates [17], $(\theta_B, \zeta_B, \psi)$,

$$\mathbf{B} = \nabla \zeta_B \times \nabla \psi + q \nabla \psi \times \nabla \theta_B. \quad (2.9)$$

Boozer coordinates are a useful coordinate choice because they preserve symmetry. In such coordinates the poloidal and toroidal components of the magnetic field are constant on flux surfaces. Figure 2.4 gives an example of magnetic field lines in an axisymmetric system represented in these Boozer coordinates. For an axisymmetric system, it can be seen that every field line that extends 2π in θ_B will sample the same geometry.

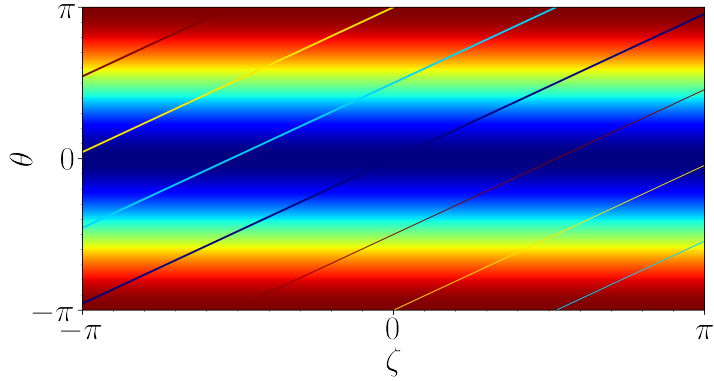


Figure 2.4: The magnetic field strength of the configuration in Figure 2.3 represented in Boozer coordinates. Here the colour indicates the strength of the magnetic field, with red denoting a stronger field and blue indicating a weaker field. The coloured straight lines represent magnetic field lines, all of which sample identical geometric configurations. Image created using stellapy [115].

2.2.1.1 Grad–Shafranov Equation

For toroidal geometries with axisymmetry, if magnetic surfaces exist, the coordinate ψ in equation (2.8) can be taken to be the poloidal magnetic flux, which is defined as

$$\psi = \int_{S_P(\psi)} \mathbf{B} \cdot \hat{\mathbf{n}} \, d^2r, \quad (2.10)$$

where $\hat{n}$ is a unit vector in the poloidal direction, and $S_P(\psi)$ is a ribbon-like surface extending from the magnetic axis to the flux coordinate ψ at fixed poloidal angle. $I(\psi)$ can be interpreted as the poloidal current density as a function of the poloidal magnetic flux, ψ , in the plasma. The ψ -component of the plasma equilibrium force balance, (2.2), is used with Maxwell's equations, (2.1), to derive the Grad–Shafranov equation [108, 40]

$$\nabla^2\psi - \frac{2}{R}\frac{\partial\psi}{\partial R} = -R^2\frac{dp}{d\psi} - I\frac{dI}{d\psi}, \quad (2.11)$$

where R is the major radius, as labelled in Figure 2.1. The quantities $p = p(\psi)$, and $I = I(\psi)$ are both flux functions. Typically the shape of the outer boundary, or the coil currents are given along with the pressure and current profiles, $p(\psi)$ and $I(\psi)$. With this, the function ψ for all inner flux surfaces is solved for, which governs the shape of the magnetic surfaces in axisymmetric devices, like tokamaks, through equation (2.8). Once the equilibrium magnetic field geometry has been determined, it can then be tested for MHD stability [?].

2.2.1.2 Miller Geometry

Miller geometry [86] is a formalism capable of describing the local magnetic geometry of a flux surface within axisymmetric systems. The Miller approach ensures that the Grad–Shafranov equation is locally satisfied in ψ by using a set of independent parameters to define a single flux surface in an axisymmetric device. It is worth noting that satisfying the Grad–Shafranov equation locally does not guarantee consistency with a global equilibrium, but rather parameterises a local flux surface. The model equations describing the shape of the flux surface—with flux label r defined as the half-diameter of the surface at the height of the magnetic axis—are:

$$R(r, \theta) = R_0(r) + r_0(r) \cos(\theta + \sin(\theta) \arcsin[\bar{\delta}(r)]), \quad (2.12)$$

$$Z(r, \theta) = r_0(r) \kappa(r) \sin(\theta), \quad (2.13)$$

where $R(r, \theta)$ and $Z(r, \theta)$ define the major radial and vertical locations for a given poloidal location, θ , illustrated in Figure 2.5. Here, $\bar{\delta}$ and κ indicate the triangularity and elongation of the surface, respectively. An illustration of how variations in elongation and triangularity affect the shape of the flux surface is shown in Figure 2.6.

By varying θ over a 2π interval, equations (2.12) and (2.13) define a closed line in the poloidal plane. This cross section is revolved 2π in the toroidal angle to map out the magnetic flux surface. The Miller geometry formalism allows for

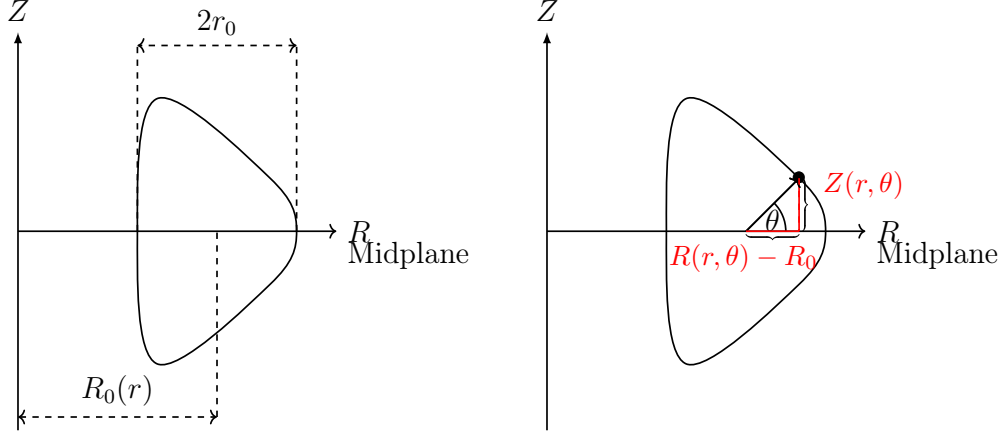


Figure 2.5: A poloidal cross section of a flux surface, illustrating the definition of the quantities used in the Miller representation of a flux surface.

the study of local stability and confinement properties without the need to repeatedly construct a self-consistent global equilibrium. The set of Miller parameters are $\{r_0, R, \Delta, q, \hat{s}, \kappa, \kappa', I, \bar{\delta}, \bar{\delta}', \beta'\}$, where primes indicate a derivative with respect to r . These correspond to markers for the minor and major radii, horizontal Shafranov shift ($\Delta = R'$), safety factor, magnetic shear ($\hat{s} \doteq (r_0/q)q'$), elongation and its radial derivative, the poloidal current density, triangularity and its radial derivative, and the radial pressure derivative respectively, with each being specified at the flux surface of interest.

2.2.2 Non-Axisymmetric Geometries

Non-axisymmetric magnetic confinement devices, such as stellarators, present unique challenges and opportunities for plasma confinement. Unlike tokamaks, which maintain symmetry about a central axis, non-axisymmetric devices feature three-dimensional magnetic field structures, such as the one shown in Figure 2.7. Unlike tokamaks, stellarators do not rely on an internal plasma current to generate a helical magnetic field, and instead achieve this by breaking axisymmetry using external coils that generate the magnetic field; this can be done by rotating elongated structures into a flux surface, or by making the magnetic axis non-planar. For non-axisymmetric devices, the rotational transform is defined as

$$\iota(\psi) = \frac{1}{2\pi} \int_0^{2\pi} \frac{\mathbf{B} \cdot \nabla \theta}{\mathbf{B} \cdot \nabla \zeta} d\theta, \quad (2.14)$$

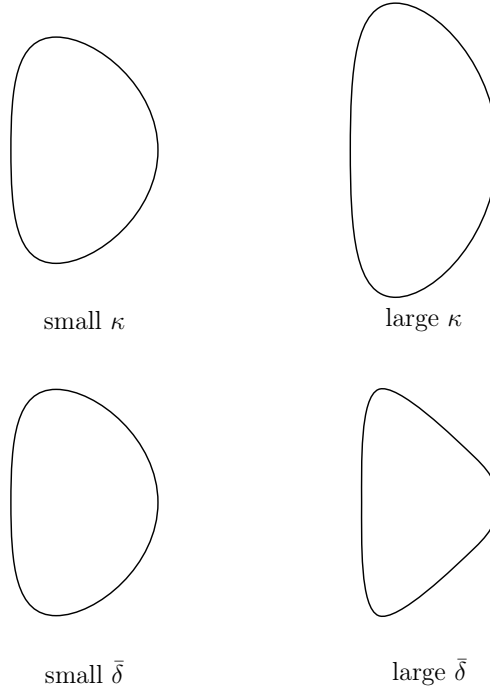


Figure 2.6: Poloidal cross-sections of a flux surface illustrating how variations in elongation, κ , and triangularity, $\bar{\delta}$, affect the shape of the flux surface.

which is analogous to the inverse of the safety factor for axisymmetric devices. The same analysis as in Section 2.2.1 gives $\alpha = \theta - \iota\zeta$. It should be noted that, as with axisymmetric devices, this specific choice for α is not unique. Different choices may result in varying representations of the magnetic field that can yield useful insights, depending on the context of the analysis or simulation. For non-axisymmetric devices a common coordinate choice is (α, ζ, ψ) . The magnetic field takes the following form using this coordinate choice

$$\mathbf{B} = (\nabla\theta \times \nabla\psi - \iota\nabla\zeta \times \nabla\psi). \quad (2.15)$$

With this choice of α we adopt the field-aligned representation for the magnetic field, an example of which is shown in Figure 2.8. Note that different field lines will sample different parts of the magnetic geometry for a 2π extension in either θ or ζ . This is in contrast with axisymmetric geometries, where all field lines are geometrically identical.

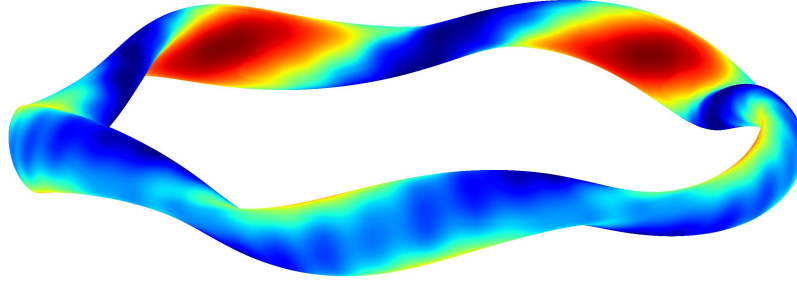


Figure 2.7: Example of a flux surface for the stellarator W7-X. Here the colours represent the magnetic field strength, with red indicating a stronger field, and blue indicating a weaker field. Image created using stellapy [115].

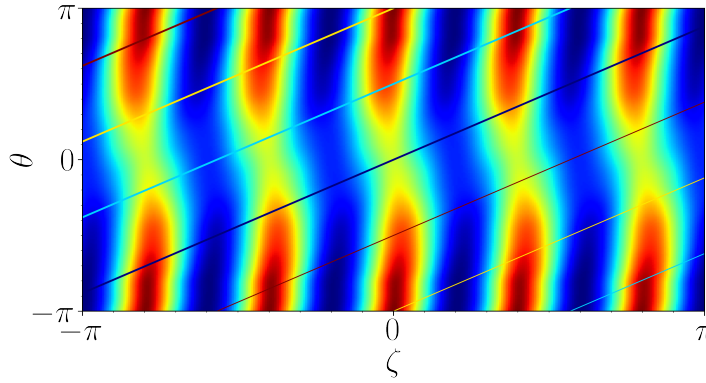


Figure 2.8: The magnetic field strength of the configuration in Figure 2.7 represented in Boozer coordinates. Here the colour indicates the strength of the magnetic field, with red denoting a stronger field and blue indicating a weaker field. The coloured straight lines represent different magnetic field lines, which are sampling different parts of the geometric configuration. Image created using stellapy [115].

2.3 Plasma Evolution Equations

We wish to consider how plasma turbulence evolves in the core of an MCF device. To model this evolution of plasma fluctuations we use the δf -gyrokinetic equation [24, 5, 1, 41, 26, 65], which one derives by assuming a spatial and temporal scale separation. This model is appropriate for the core of many MCF devices because of the strong magnetisation and the modest plasma profile gradients. However, scale separation may fail in certain situations, such as in the pedestal region, where plasma profile gradients are significant, or in devices with sufficiently weak magnetisation where the gyroradius of a thermal species may be comparable to the device size [39]. Therefore, we restrict our considerations to scenarios where the scale separation, both

in space and time, is present.

The starting point is the Vlasov–Maxwell system of equations including collisions; for this, combine Maxwell’s equations, given by (2.1), with the Vlasov–Fokker–Planck equation for the evolution of the species distribution functions

$$\frac{df_\nu}{dt} = \sum_{\nu'} C_{\nu\nu'}[f_\nu, f_{\nu'}], \quad (2.16)$$

where $C_{\nu\nu'}[f_\nu, f_{\nu'}]$ accounts for the effects on species ν from Coulomb collisions with species ν' . The charge and current densities are related to the particle distribution function, f_ν , via the velocity space integrals

$$\varrho = \sum_{\nu} Z_{\nu} e \int d^3v f_{\nu}, \quad (2.17a)$$

$$\mathbf{j} = \sum_{\nu} Z_{\nu} e \int d^3v \mathbf{v} f_{\nu}. \quad (2.17b)$$

Gyrokinetics is formulated to study the turbulent evolution of plasmas, but plasma dynamics is fundamentally a multiscale problem. As a result, one must consider the physics driving the system and tease out the different scales that may be present to exploit a separation of scales and simplify the system.

On the length scales of the equilibrium, the plasma profile parameters, such as temperature and pressure, vary smoothly. The fundamental assumption is that these equilibrium quantities have slow spatial and temporal evolution; the dynamics are governed by slow transport time scales, τ_E , and variations are on large spatial scales that are on the order of the macroscopic device size, L .

Conversely, turbulence exists on short spatial and temporal scales. The natural length scale of turbulence is the Larmor radius, $\rho_\nu = v_\perp/\Omega_\nu$, where $v_\perp$ is the magnitude of the particle velocity in the plane perpendicular to the magnetic field. The Larmor radius of particles is taken to be much smaller than the system size; $\rho_\nu \ll L$. We define, for use henceforth, the thermalised gyroradius of a species, defined as $\rho_{\text{th},\nu} \doteq v_{\text{th},\nu}/\Omega_{r,\nu}$, where $v_{\text{th},\nu} \doteq \sqrt{2T_\nu/m_\nu}$ is the species’ thermal velocity, with T_ν the mean species’ temperature. Here $\Omega_{r,\nu}$ is the species reference gyrofrequency which is evaluated at some characteristic magnetic field strength. It is important to note that the electron Debye length is much smaller than the ion turbulent length scale, $\lambda_{De} \ll \rho_i$, allowing the plasma to be treated as quasineutral over turbulent scales. Also, while the turbulence evolves much more rapidly than the equilibrium, the frequency of fluctuations, ω , is taken to be much less than the gyrofrequency of particles; $\omega \ll \Omega_\nu$.

This separation of scales can be exploited to decompose quantities into their mean and fluctuating parts. The mean components determine the evolution of the background plasma, and are found by averaging over all fluctuations. The average of a given quantity, h , over all fluctuations is represented by $\langle h \rangle_{\text{turb}}$, defined as

$$\langle h(t) \rangle_{\text{turb}} = \frac{1}{T} \int_{t-T/2}^{t+T/2} dt' \langle h(t') \rangle_{\perp}, \quad (2.18)$$

where T is some intermediate time shorter than the (transport) time scale associated with mean profile evolution, and longer than time scales associated with typical fluctuations, such that $\omega^{-1} \ll T \ll L/v_{\text{th},i}$. Here $\langle \cdot \rangle_{\perp}$ is an appropriately defined spatial average over a surface perpendicular to the magnetic field with a characteristic length l , satisfying $\rho_{\nu} \ll l \ll L$ [1]. The distribution function is decomposed as $f_{\nu} = F_{\nu} + \delta f_{\nu}$ with $F_{\nu} = \langle f_{\nu} \rangle_{\text{turb}}$ determining the distribution function of the plasma equilibrium, and δf_{ν} the contribution from plasma fluctuations.

The Vlasov–Maxwell equations including collisions, given by equations (2.1) and (2.16), are expanded in the small parameter $\epsilon \sim \rho_* \ll 1$, with $\rho_* = \rho_{\text{th},i}/L$ the ratio of the thermal ion gyroradius to the system size, and terms of equivalent order equated. The terms are ordered as follows

$$\epsilon \sim \rho_* \doteq \frac{\rho_{\text{th},i}}{L} \sim \frac{\omega}{\Omega_i} \sim \frac{k_{\parallel}}{k_{\perp}} \sim \frac{e\phi}{T_{\nu}} \sim \frac{\delta B}{B_0} \sim \frac{\delta f_{\nu}}{F_{0,\nu}} \ll 1 \quad \text{and} \quad k_{\perp} \rho_i \sim 1. \quad (2.19)$$

In the above orderings the magnetic field has been decomposed into the equilibrium contribution, $\mathbf{B}_0$, and the magnetic fluctuations, $\delta \mathbf{B}$, such that the full magnetic field is $\mathbf{B} = \mathbf{B}_0 + \delta \mathbf{B}$. The parallel and perpendicular wavenumbers have also been introduced, $k_{\parallel} = \mathbf{k} \cdot \hat{\mathbf{b}}$ and $\mathbf{k}_{\perp} = (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \mathbf{k}$, with $\hat{\mathbf{b}}$ the unit vector in the direction of the equilibrium magnetic field, and $\mathbf{I}$ the identity matrix. The perturbed electric potential has been introduced as ϕ . The mean and fluctuating components of the distribution function are also expanded

$$\begin{aligned} F_{\nu} &= F_{0,\nu} + F_{1,\nu} + F_{2,\nu} \cdots, \\ \delta f_{\nu} &= \delta f_{1,\nu} + \delta f_{2,\nu} + \cdots, \end{aligned} \quad (2.20)$$

with $F_{0,\nu} \sim f_{\nu}$, $F_{1,\nu} \sim \delta f_{1,\nu} \sim \epsilon f_{\nu}$, $F_{2,\nu} \sim \delta f_{2,\nu} \sim \epsilon^2 f_{\nu}$, and so on. Equilibrium quantities are taken to have characteristic length scales of order L , and evolve slowly on the long transport time scale $\tau_E^{-1} \sim \epsilon^3 \Omega_i$; they are thus understood to be static during these considerations. Small-scale fluctuations, captured in δf_{ν} , have characteristic length scales of the order $\rho_{\text{th},i} \sim \epsilon L$ and frequencies $\omega \sim \epsilon \Omega_i$.

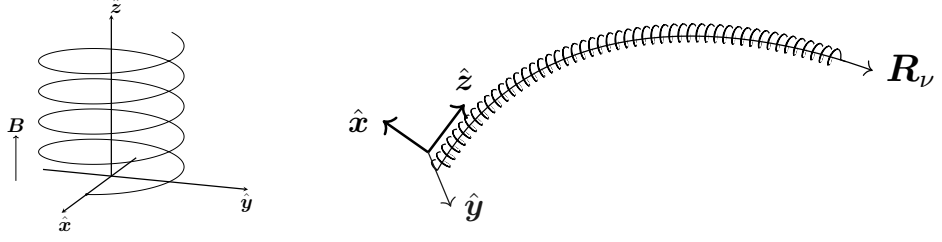


Figure 2.9: Illustration of the approximate motion of a charged particle in a magnetic field. Here the helical path indicates the approximate particle trajectory, whilst $\mathbf{R}_\nu$ gives the gyrocenter of the particle motion.

The charged particles follow magnetic field lines and perform rapid gyration in the plane perpendicular to the field, a schematic of which is shown in Figure 2.9. The particle velocities can be decomposed into the parallel velocity, $v_\parallel$, the perpendicular velocity, $v_\perp$, and the gyrophase, ϑ ,

$$\mathbf{v} = v_\parallel \hat{\mathbf{b}} + v_\perp [\cos(\vartheta) \hat{\mathbf{x}} + \sin(\vartheta) \hat{\mathbf{y}}], \quad (2.21)$$

where $\hat{\mathbf{x}}$, and $\hat{\mathbf{y}}$ are unit vectors that locally form an orthonormal basis with $\hat{\mathbf{b}}$. The notation $\mathbf{R}_\nu$ in Figure 2.9 is introduced as the gyrocentre for particles of species ν , and $\mathbf{r}$ is used to indicate the spatial position vector for any given particle. These descriptions of particle location are related through $\mathbf{R}_\nu = \mathbf{r} - \boldsymbol{\rho}_\nu(\vartheta)$, with $\boldsymbol{\rho}_\nu(\vartheta) = \hat{\mathbf{b}} \times \mathbf{v} / \Omega_\nu$ the velocity-dependent vector gyroradius. The gyrophase, ϑ , characterises the gyromotion, and has a large associated frequency $|\dot{\vartheta}| \approx \Omega_\nu$, where the dot on the gyrophase indicates a temporal derivative. A natural approach is to average over these fast oscillations by introducing a gyroaverage, defined through

$$\langle h(\mathbf{r}) \rangle_{\mathbf{R}_\nu} = \langle h(\mathbf{R}_\nu + \boldsymbol{\rho}_\nu(\vartheta)) \rangle_{\mathbf{R}_\nu} = \frac{1}{2\pi} \int_0^{2\pi} h(\mathbf{R}_\nu + \boldsymbol{\rho}_\nu(\vartheta)) \, d\vartheta, \quad (2.22)$$

$$\langle h(\mathbf{R}_\nu) \rangle_{\mathbf{r}} = \langle h(\mathbf{r} - \boldsymbol{\rho}_\nu(\vartheta)) \rangle_{\mathbf{r}} = \frac{1}{2\pi} \int_0^{2\pi} h(\mathbf{r} - \boldsymbol{\rho}_\nu(\vartheta)) \, d\vartheta, \quad (2.23)$$

where $\mathbf{R}_\nu$ and $\mathbf{r}$ are held constant when performing the ϑ integrations in equations (2.22) and (2.23) respectively.

Expanding the Vlasov–Maxwell equations, then equating terms which are ordered as $\mathcal{O}(\epsilon^{-1} v_{\text{th},i} F_{0,i} / L)$ one obtains the equation $\partial F_{0,\nu} / \partial \vartheta = 0$ at fixed $\mathbf{R}_\nu$, which demands $F_{0,\nu}$ be independent of gyrophase. In the presence of modest collisionality¹, zeroth-order terms, $\mathcal{O}(\epsilon^0 v_{\text{th},i} F_{0,i} / L)$, provide the further constraint that the equilibrium component is a Maxwellian in velocities:

$$F_{0,\nu} := \frac{n_\nu}{(\pi v_{\text{th},\nu}^2)^{\frac{3}{2}}} e^{-v^2 / v_{\text{th},\nu}^2}, \quad (2.24)$$

¹This requires the collisionality $\nu_* \gtrsim \rho_* \omega$ – a regime within which this work is considered.

where n_ν represents the species' mean density.

Equating terms ordered $\mathcal{O}(\epsilon^1 v_{\text{th},i} F_{0,i}/L)$ gives the evolution equation for first-order perturbations. This exists within a six-dimensional phase space with coordinate choice $\{\mathbf{R}_\nu, v_\parallel, \mu_\nu, \vartheta\}$. Here, μ_ν is the magnetic moment defined as $\mu_\nu = m_\nu v_\perp^2/2B$ and is a conserved quantity to the order of consideration. Gyrophase dependence is removed by gyroaveraging the full equation, reducing the phase-space dimensionality by one from 6D to 5D. For convenience, we define the following guiding-centre-dependent distribution function, in terms of δf_ν , to be $g_\nu(\mathbf{R}_\nu, v_\parallel, \mu_\nu) \doteq \delta f_\nu + (Z_\nu e/T_\nu) F_{0,\nu}[\phi - \langle \chi \rangle_{\mathbf{R}_\nu}]$, such that g_ν is gyrophase independent. The gyrokinetic potential, $\chi = \phi - \mathbf{v} \cdot \mathbf{A}/c$, is also introduced, with $\mathbf{A} = A_\parallel \hat{\mathbf{b}} + \mathbf{A}_\perp$ the fluctuating magnetic vector potential, $\delta \mathbf{B} = \nabla \times \mathbf{A}$, which has been decomposed into components parallel and perpendicular to the equilibrium magnetic field. The Coulomb gauge is imposed on the magnetic vector potential by requiring $\nabla \cdot \mathbf{A} = 0$.

The total time derivative in (2.16) is expanded in terms of partial derivatives in $\{t, v_\parallel, \mu_\nu, \vartheta, \mathbf{R}_\nu\}$ – with each partial derivative taken assuming all other variables are held fixed, unless explicitly stated otherwise. A ‘low-flow’, or ‘drift’, ordering is defined to be when the flow speed is ordered as ρ_* small compared with the thermal speed. In this ordering the resulting nonlinear, electromagnetic gyrokinetic equation including collisions is symbolically represented as

$$\frac{\partial g_\nu}{\partial t} + \mathcal{S}_\nu[g_\nu, \chi] + \mathcal{M}_\nu[g_\nu] + \mathcal{D}_\nu[g_\nu, \chi] + \mathcal{G}_\nu[\chi] + \mathcal{A}_\nu[A_\parallel] + \mathcal{N}_\nu[g_\nu, \chi] = \mathcal{C}_\nu[\{g_\nu, \langle \chi \rangle_{\mathbf{R}_\nu}\}], \quad (2.25)$$

with operations defined by

$$\mathcal{S}_\nu[g_\nu, \chi] \doteq v_\parallel \left(\hat{\mathbf{b}} \cdot \nabla g_\nu + \frac{Z_\nu e}{T_\nu} F_{0,\nu} \hat{\mathbf{b}} \cdot \nabla \langle \chi \rangle_{\mathbf{R}_\nu} \right), \quad (2.26a)$$

$$\mathcal{M}_\nu[g_\nu] \doteq - \frac{\mu_\nu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla B_0 \frac{\partial g_\nu}{\partial v_\parallel}, \quad (2.26b)$$

$$\mathcal{D}_\nu[g_\nu, \chi] \doteq \mathbf{v}_{M,\nu} \cdot \left(\nabla g_\nu + \frac{Z_\nu e}{T_\nu} F_{0,\nu} \nabla \langle \chi \rangle_{\mathbf{R}_\nu} \right), \quad (2.26c)$$

$$\begin{aligned} \mathcal{G}_\nu[\chi] &\doteq \langle \mathbf{v}_\chi \rangle_{\mathbf{R}_\nu} \cdot \nabla|_{\mathcal{E}} F_{0,\nu} \\ &= c \frac{\partial \langle \chi \rangle_{\mathbf{R}_\nu}}{\partial \alpha} F_{0,\nu} \left[\frac{\partial \ln n_\nu}{\partial \psi} + \frac{\partial \ln T_\nu}{\partial \psi} \left(\frac{m_\nu v_\parallel^2}{2T_\nu} + \frac{\mu_\nu B_0}{T_\nu} - \frac{3}{2} \right) \right], \end{aligned} \quad (2.26d)$$

$$\mathcal{A}_\nu[A_\parallel] \doteq \frac{Z_\nu e}{T_\nu} \frac{\mu_\nu}{m_\nu c} \hat{\mathbf{b}} \cdot \nabla B_0 F_{0,\nu} \langle A_\parallel \rangle_{\mathbf{R}_\nu}, \quad (2.26e)$$

$$\mathcal{N}_\nu[g_\nu, \chi] \doteq \langle \mathbf{v}_\chi \rangle_{\mathbf{R}_\nu} \cdot \nabla_\perp g_\nu = c \left(\frac{\partial \langle \chi \rangle_{\mathbf{R}_\nu}}{\partial \alpha} \frac{\partial g_\nu}{\partial \psi} - \frac{\partial \langle \chi \rangle_{\mathbf{R}_\nu}}{\partial \psi} \frac{\partial g_\nu}{\partial \alpha} \right), \quad (2.26f)$$

and $\mathcal{C}_\nu[\{g_{\nu'}, \langle \chi \rangle_{\mathbf{R}_{\nu'}}\}]$ is the corresponding collision operator which depends on $g_{\nu'}$ and $\langle \chi \rangle_{\mathbf{R}_{\nu'}}$ of other species.

Here $C_\nu = \sum_{\nu'} \langle C_{\nu, \nu'}[F_{0, \nu}, \delta f_{\nu'}] + C_{\nu, \nu'}[\delta f_\nu, F_{0, \nu'}] \rangle_{\mathbf{R}_\nu}$ is the linearised collision operator. The collision operator is assumed to be self-adjoint, such that $C_\nu^\dagger = C_\nu$ holds, but otherwise its form is left unspecified². The gradient acting on the Maxwellian appears as $\nabla|_{\mathcal{E}}$. This indicates that the derivative has been taken at constant kinetic energy, $\mathcal{E} = m_\nu v^2/2$, rather than at fixed $\{v_\parallel, \mu_\nu\}$ variables, in contrast to the other spatial gradients appearing in equation (2.25). Finally, $\mathbf{v}_{M, \nu}$ and $\mathbf{v}_\chi$ are the magnetic and generalised $\mathbf{E} \times \mathbf{B}$ drifts defined through

$$\mathbf{v}_{M, \nu} = \frac{1}{\Omega_\nu} \hat{\mathbf{b}} \times \left(\frac{\mu_\nu}{m_\nu} \nabla B_0 + v_\parallel^2 \boldsymbol{\kappa} \right), \quad (2.27)$$

$$\mathbf{v}_\chi = \frac{c}{B} \hat{\mathbf{b}} \times \nabla_\perp \chi, \quad (2.28)$$

with $\boldsymbol{\kappa} = \hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}$ the equilibrium magnetic field curvature.

The system is closed by the field equations consisting of quasineutrality, given by (1.2), and the low-frequency Ampère's law, $\nabla \times \delta \mathbf{B} = (4\pi/c) \delta \mathbf{J}$, with $\delta \mathbf{B}$ and $\delta \mathbf{J}$ the fluctuating magnetic field and current density respectively. When written in terms of the distribution function these relations become:

$$\sum_\nu Z_\nu \int d^3v \left[\langle g_\nu \rangle_{\mathbf{r}} + \frac{Z_\nu e}{T_\nu} F_{0, \nu} (\langle \langle \chi \rangle_{\mathbf{R}_\nu} \rangle_{\mathbf{r}} - \phi) \right] = 0, \quad (2.29)$$

$$\nabla_\perp^2 A_\parallel - \frac{4\pi}{c} \sum_\nu Z_\nu e \int d^3v v_\parallel \left[\langle g_\nu \rangle_{\mathbf{r}} + \frac{Z_\nu e}{T_\nu c} F_{0, \nu} v_\parallel \langle \langle A_\parallel \rangle_{\mathbf{R}_\nu} \rangle_{\mathbf{r}} \right] = 0, \quad (2.30)$$

$$\nabla_\perp^2 \delta B_\parallel - \frac{4\pi}{c} \sum_\nu Z_\nu e \int d^3v \nabla \cdot \left\langle \left(g_\nu + \frac{Z_\nu e}{T_\nu} F_{0, \nu} \langle \chi \rangle_{\mathbf{R}_\nu} \right) (\hat{\mathbf{b}} \times \mathbf{v}_\perp) \right\rangle_{\mathbf{r}} = 0, \quad (2.31)$$

where $\delta B_\parallel = \delta \mathbf{B} \cdot \hat{\mathbf{b}}$ is the parallel component of the perturbed magnetic field, and equation (2.31) has been obtained by considering $\nabla \cdot (\hat{\mathbf{b}} \times (\nabla \times \delta \mathbf{B}))$. With this choice of coordinates, the velocity space integrals take the following form:

$$\int d^3v f \doteq \int_0^{2\pi} d\vartheta \int_{-\infty}^{\infty} dv_\parallel \int_0^{\infty} d\mu_\nu \frac{B_0}{m_\nu} f.$$

2.3.1 System Equations

The gyrokinetic orderings, (2.19), imply that the turbulence is characterised by long parallel wavelengths and short perpendicular wavelengths, and that the parallel dynamics along the field line is fast compared with the drifts of particles across the

²It should be noted that this encapsulates a broad range of collision operators including the linearised Landau, and the linearised Fokker-Planck collision operators.

field line. Over a sufficiently large domain—much larger than the scale of the fluctuations but still smaller than the equilibrium length scale—the turbulent fluctuations become decorrelated in ψ , and thus statistically identical at the boundaries. If the full poloidal domain is also considered, then the turbulence is also periodic in α . With this, one can take the discrete Fourier transform in the perpendicular domain. The discrete Fourier transform in $\{\psi, \alpha\}$ is

$$g_\nu(\psi, \alpha, z, v_\parallel, \mu_\nu, t) = \sum_{k_\psi, k_\alpha} \hat{g}_{\mathbf{k}, \nu}(z, v_\parallel, \mu_\nu, t) e^{(ik_\psi \psi + ik_\alpha \alpha)}. \quad (2.32)$$

This is justified provided the condition $k_\psi |\nabla \psi| \sim k_\alpha |\nabla \alpha| \gg 1/L$ is satisfied, allowing the use of ideal statistical periodicity to enforce periodic boundary conditions in $\{\psi, \alpha\}$ [11].

The Fourier transform of the evolution equation (2.25), and field equations (2.29)–(2.31), is taken, allowing for geometric dependence in the binormal direction, whilst taking the local limit in the radial direction. As a result of this, when the Fourier transform is taken the resulting equation contains convolutions over the binormal mode number, k_y ,

$$\begin{aligned} \frac{\partial \hat{g}_{\mathbf{k}, \nu}}{\partial t} + \sum_{k'_\alpha} \Big\{ & \hat{\mathcal{S}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{g}_{(k_\psi, k'_\alpha), \nu}, \hat{\chi}_{(k_\psi, k'_\alpha), \nu}] + \hat{\mathcal{M}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{g}_{(k_\psi, k'_\alpha), \nu}, \hat{A}_{\parallel, k_\psi, k'_\alpha, \nu}] \\ & + \hat{\mathcal{D}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{g}_{(k_\psi, k'_\alpha), \nu}, \hat{\chi}_{(k_\psi, k'_\alpha), \nu}] + \hat{\mathcal{G}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{\chi}_{(k_\psi, k'_\alpha), \nu}] \\ & + \hat{\mathcal{A}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{A}_{\parallel, \mathbf{k}}] + \hat{\mathcal{N}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{g}_{(k_\psi, k'_\alpha), \nu}, \hat{\chi}_{(k_\psi, k'_\alpha), \nu}] \Big\} \\ & = \sum_{k'_\alpha} \hat{\mathcal{C}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\{\hat{g}_{(k_\psi, k'_\alpha), \nu'}, \hat{\chi}_{(k_\psi, k'_\alpha), \nu'}\}]. \end{aligned} \quad (2.33)$$

Here $\mathcal{F} [\langle \chi \rangle_{\mathbf{R}_\nu}]_{\mathbf{k}} \doteq \hat{\chi}_{\mathbf{k}, \nu}$ defines the Fourier components of $\langle \chi \rangle_{\mathbf{R}_\nu}$. For clarity, consider just the electrostatic potential here. The gyroaverage of ϕ can be written in terms of Fourier components as follows:

$$\varphi_\nu = \langle \phi \rangle_{\mathbf{R}_\nu} = \sum_{\mathbf{k}''} e^{i\mathbf{k}'' \cdot \mathbf{R}_\nu} J_0(a_{\mathbf{k}'', \nu}) \hat{\phi}_{\mathbf{k}''}. \quad (2.34)$$

The variables J_n are the n^{th} -order Bessel functions of the first kind. The Bessel functions arise naturally as a result of the gyroaverages, and have argument

$$a_{\mathbf{k}'', \nu} = \frac{ck_\perp''(\alpha, z)}{Z_\nu e} \sqrt{\frac{2m_\nu \mu}{B_0(\alpha, z)}}, \quad (2.35)$$

which itself may have α -dependence. For axisymmetric systems, the α dependence is absent as all field lines are identical, so gyroaveraging becomes a local operation in k_α -space; the α dependence arising from non-axisymmetry introduces coupling between

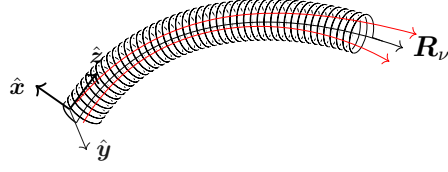


Figure 2.10: An illustration of the path of a particle completing multiple gyro-orbits and sampling multiple magnetic field lines, each with slightly varying structure.

modes with different k_α . The zeroth-order Bessel function, J_0 , can be expanded in Fourier harmonics as

$$J_0(a_{\mathbf{k}'',\nu}) = \sum_{k_\alpha'''} \hat{J}_{0,\mathbf{k}'',k_\alpha''',\nu}(z,\mu) e^{ik_\alpha''' \alpha}. \quad (2.36)$$

With this, the Fourier transform of the gyroaveraged electrostatic potential becomes

$$\hat{\varphi}(k_\psi, k'_\alpha, \nu) = \int d^2 \mathbf{R}_\nu \sum_{\mathbf{k}'', k_\alpha'''} e^{i(k_\psi'' - k_\psi) \psi} e^{i(k_\alpha'' + k_\alpha''' - k'_\alpha) \alpha} \hat{J}_{0,\mathbf{k}'',k_\alpha''',\nu} \hat{\phi}_{\mathbf{k}''}. \quad (2.37)$$

Making use of the orthogonality of the Fourier harmonics, we find $k_\psi'' = k_\psi$ and $k_\alpha'' = (k'_\alpha - k_\alpha''')$ such that equation (2.37) simplifies to

$$\hat{\varphi}(k_\psi, k'_\alpha, \nu) = \sum_{k_\alpha'''} \hat{J}_{0,(k_\psi, k'_\alpha - k_\alpha'''), k_\alpha''', \nu} \hat{\phi}(k_\psi, k'_\alpha - k_\alpha'''). \quad (2.38)$$

Using the above procedure, one obtains an expression for the Fourier-transformed, gyroaveraged gyrokinetic potential

$$\mathcal{F}_{(k_\psi, k'_\alpha)} [\langle \chi \rangle_{\mathbf{R}_\nu}] \doteq \hat{\chi}_{(k_\psi, k'_\alpha), \nu} = \sum_{k_\alpha'''} \left[\hat{J}_{0,(k_\psi, k'_\alpha - k_\alpha'''), k_\alpha''', \nu} \hat{\phi}(k_\psi, k'_\alpha - k_\alpha''') - \frac{v_\parallel}{c} \hat{J}_{0,(k_\psi, k'_\alpha - k_\alpha'''), k_\alpha''', \nu} \hat{A}_{\parallel, (k_\psi, k'_\alpha - k_\alpha''')} + 2 \frac{\mu_\nu}{Z_\nu e} \frac{\hat{J}_{1,(k_\psi, k'_\alpha - k_\alpha'''), k_\alpha''', \nu}}{a_{(k_\psi, k'_\alpha - k_\alpha'''), k_\alpha''', \nu}} \delta \hat{B}_{\parallel, (k_\psi, k'_\alpha - k_\alpha''')} \right]. \quad (2.39)$$

It is clear that gyroaveraging couples different Fourier modes in α . This process can be interpreted physically as the particle sampling multiple field lines with different geometries. This is illustrated in Figure 2.10, where the black central line represents the field line that the particle is following, while the red lines depict nearby field lines that are not parallel to the original. The tight helical pattern illustrates the particle's gyromotion. The gyroaverage induces non-local interactions in k_α due to the geometric variation across field lines. Similar convolutions arise from the geometric

dependence of coefficients in (2.33). For example, the parallel streaming operator becomes

$$\begin{aligned} \hat{\mathcal{S}}_{(k_\psi, k_\alpha - k'_\alpha), \nu} [\hat{g}_{(k_\psi, k_{\alpha'}), \nu}, \hat{\chi}_{(k_\psi, k_{\alpha'}), \nu}] = & \mathcal{F}_{(k_\psi, k_\alpha - k_{\alpha'})} \left[v_\parallel \hat{\mathbf{b}} \cdot \nabla z \right] \frac{\partial \hat{g}_{(k_\psi, k_{\alpha'}), \nu}}{\partial z} \\ & + \mathcal{F}_{(k_x, k_\alpha - k_{\alpha'})} \left[v_\parallel \frac{Z_\nu e}{T_\nu} \hat{\mathbf{b}} \cdot \nabla z F_{0, \nu} \right] \frac{\partial \hat{\chi}_{(k_\psi, k_{\alpha'}), \nu}}{\partial z}. \end{aligned} \quad (2.40)$$

2.3.2 Local Approximation

In the limit where the α -variation of the geometry is slow, it is possible to consider the geometric quantities to be constant in α within the domain. The α -coupling arising from geometric coefficients and gyroaveraging then vanishes. Under this assumption, the convolutions in the system equations reduce to local operations that do not couple different k_α modes. This is a good approximation in axisymmetric devices—where all field lines share the same geometric variation—or in the limit $\rho_* \rightarrow 0$, where turbulent fluctuations become decorrelated over scales much shorter than the equilibrium scale such that the turbulent evolution on each field line can be considered independently. Consequently, the computational domain no longer needs to span the full α -domain, a reduced domain that considers a single field line is sufficient.

In these regimes, the local-in- α approximation is valid, though it may introduce inconsistencies in the treatment of zonal flows across different field lines. This is because, in a flux tube simulation, turbulence drives the generation of zonal flows, which are assumed to be constant across the simulated domain. However, in non-axisymmetric devices, performing flux tube simulations on different field lines can lead to different zonal flow responses, due to the local variation in magnetic geometry. This leads to an inconsistency, since zonal flows are, by definition, flux-surface averaged quantities and must be constant across a given flux surface. However, despite its potential limitations in capturing zonal flow dynamics, it is still a powerful tool used to predict turbulent fluxes in devices with geometry that has slow α -variation. The equations outlined in Section 2.3.1 simplify greatly, which eases the associated computational cost of solving them, and consequently this is the limit taken by many existing codes to model the turbulence. Hence it is of interest to consider this limit.

One such place that this simplification can be seen is in the Bessel functions. In the local approximation, the geometric quantities are considered constant-in- α across the domain. This means that the α -dependence in the arguments of the Bessel functions, given in equation (2.35), is absent, and gyroaveraging thus becomes a local operation

in k_α space. Hence, the gyroaveraged gyrokinetic potential reduces to

$$\mathcal{F}_k [\langle \chi \rangle_{\mathbf{R}_\nu}] \doteq \hat{\chi}_{\mathbf{k},\nu} = \left[J_{0,\mathbf{k},\nu} \hat{\phi}_{\mathbf{k}} - \frac{v_\parallel}{c} J_{0,\mathbf{k},\nu} \hat{A}_{\parallel,\mathbf{k}} + 2 \frac{\mu_\nu}{Z_\nu e} \frac{J_{1,\mathbf{k},\nu}}{a_{\mathbf{k},\nu}} \delta \hat{B}_{\parallel,\mathbf{k}} \right]. \quad (2.41)$$

An additional simplification can be seen in the system equations. Here, the geometric coefficients become constants in k_α space, meaning the convolutions seen in (2.33) reduce to local operations in k_α . The resulting gyrokinetic equation is

$$\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} + \hat{\mathcal{S}}_{\mathbf{k},\nu} + \hat{\mathcal{M}}_{\mathbf{k},\nu} + \hat{\mathcal{D}}_{\mathbf{k},\nu} + \hat{\mathcal{G}}_{\mathbf{k},\nu} + \hat{\mathcal{A}}_{\mathbf{k},\nu} + \hat{\mathcal{N}}_{(k_\psi, k_\alpha - k'_\alpha),\nu} = \hat{\mathcal{C}}_{\mathbf{k},\nu} [\{\hat{g}_{\mathbf{k},\nu'}, \hat{\chi}_{\mathbf{k},\nu'}\}], \quad (2.42)$$

with the local-in- k_α operators defined as

$$\begin{aligned} \hat{\mathcal{S}}_{\mathbf{k},\nu} [\hat{g}_{\mathbf{k},\nu}, \hat{\chi}_{\mathbf{k},\nu}] &= v_\parallel \hat{\mathbf{b}} \cdot \nabla z \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial z} + \frac{Z_\nu e}{T_\nu} \frac{\partial \hat{\chi}_{\mathbf{k},\nu}}{\partial z} F_{0,\nu} \right), \\ \hat{\mathcal{M}}_{\mathbf{k},\nu} [\hat{g}_{\mathbf{k},\nu}] &= - \frac{\mu_\nu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla B_0 \frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial v_\parallel}, \\ \hat{\mathcal{D}}_{\mathbf{k},\nu} [\hat{g}_{\mathbf{k},\nu}, \hat{\chi}_{\mathbf{k},\nu}] &= i \omega_{d,\mathbf{k},\nu} \left(\hat{g}_{\mathbf{k},\nu} + \frac{Z_\nu e}{T_\nu} \hat{\chi}_{\mathbf{k},\nu} F_{0,\nu} \right), \\ \hat{\mathcal{G}}_{\mathbf{k},\nu} [\hat{\chi}_{\mathbf{k},\nu}] &= i \omega_{*,\mathbf{k},\nu} F_{0,\nu} \hat{\chi}_{\mathbf{k},\nu}, \\ \hat{\mathcal{A}}_{\mathbf{k},\nu} [\hat{A}_{\parallel,\mathbf{k}}] &= \frac{Z_\nu e}{T_\nu} \frac{\mu_\nu}{m_\nu c} (\hat{\mathbf{b}} \cdot \nabla B_0) F_{0,\nu} J_{0,\mathbf{k},\nu} \hat{A}_{\parallel,\mathbf{k}}, \\ \hat{\mathcal{N}}_{\mathbf{k},\nu} [\hat{g}_{\mathbf{k},\nu}, \hat{\chi}_{\mathbf{k},\nu}] &= c \mathcal{F}_{\mathbf{k}} [\mathcal{F}_{\mathbf{k}}^{-1} [i k_\alpha \hat{\chi}_{\mathbf{k},\nu}] [i k_\psi \hat{g}_{\mathbf{k},\nu}] - \mathcal{F}_{\mathbf{k}}^{-1} [i k_\psi \hat{\chi}_{\mathbf{k},\nu}] [i k_\alpha \hat{g}_{\mathbf{k},\nu}]], \end{aligned} \quad (2.43)$$

where $\mathcal{F}_{\mathbf{k}}^{-1}[\cdot]$ is the inverse Fourier transform, and $\omega_{d,\mathbf{k},\nu}$ and $\omega_{*,\mathbf{k},\nu}$ are the magnetic drift frequencies resulting from the gradient and curvature of the magnetic field and the diamagnetic drift respectively. The expressions for these are given by

$$\omega_{d,\mathbf{k},\nu} = \frac{1}{\Omega_\nu} (v_\parallel^2 \hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) + \mu_\nu \hat{\mathbf{b}} \times \nabla B_0) \cdot (k_\psi \nabla \psi + k_\alpha \nabla \alpha), \quad (2.44a)$$

$$\omega_{*,\mathbf{k},\nu} = \frac{c k_\alpha}{B_0} \frac{d \ln F_{0,\nu}}{d \psi}. \quad (2.44b)$$

The corresponding transformed field equations are

$$\sum_{\nu} Z_{\nu} e \left\{ \frac{2\pi B_0}{m_{\nu}} \int d^2 v J_{0,\mathbf{k},\nu} \hat{g}_{\mathbf{k},\nu} + \frac{Z_{\nu} e n_{\nu}}{T_{\nu}} (\Gamma_{0,\mathbf{k},\nu} - 1) \hat{\phi}_{\mathbf{k}} + \frac{n_{\nu}}{B_0} \Gamma_{1,\mathbf{k},\nu} \delta \hat{B}_{\parallel,\mathbf{k}} \right\} = 0, \quad (2.45)$$

$$\frac{4\pi}{k_{\perp}^2 c} \sum_{\nu} Z_{\nu} e \frac{2\pi B_0}{m_{\nu}} \int d^2 v v_{\parallel} J_{0,\mathbf{k},\nu} \hat{g}_{\mathbf{k},\nu} - \left[1 + \frac{4\pi}{k_{\perp}^2 c^2} \sum_{\nu} \frac{(Z_{\nu} e)^2 n_{\nu}}{m_{\nu}} \Gamma_{0,\mathbf{k},\nu} \right] \hat{A}_{\parallel,\mathbf{k}} = 0, \quad (2.46)$$

$$\begin{aligned} 8\pi \sum_{\nu} \frac{2\pi B_0}{m_{\nu}} \int d^2 v \frac{J_{1,\mathbf{k},\nu}}{a_{\mathbf{k},\nu}} \mu_{\nu} \hat{g}_{\mathbf{k},\nu} + \left[4\pi \sum_{\nu} \frac{Z_{\nu} e n_{\nu}}{B_0} \Gamma_{1,\mathbf{k},\nu} \right] \hat{\phi}_{\mathbf{k}} \\ + \left[1 + 16\pi \sum_{\nu} \frac{n_{\nu} T_{\nu}}{B_0^2} \Gamma_{2,\mathbf{k},\nu} \right] \delta \hat{B}_{\parallel,\mathbf{k}} = 0, \end{aligned} \quad (2.47)$$

with $\int d^2 v \doteq \int d\mu \int dv_{\parallel}$. In the above, the following functions have been introduced:

$$\begin{aligned} \Gamma_{0,\mathbf{k},\nu} &= I_0(b_{\mathbf{k},\nu}) e^{-b_{\mathbf{k},\nu}}, \\ \Gamma_{1,\mathbf{k},\nu} &= [I_0(b_{\mathbf{k},\nu}) - I_1(b_{\mathbf{k},\nu})] e^{-b_{\mathbf{k},\nu}}, \\ \Gamma_{2,\mathbf{k},\nu} &= I_1(b_{\mathbf{k},\nu}) e^{-b_{\mathbf{k},\nu}}, \end{aligned}$$

where I_0 and I_1 are modified Bessel functions of the first kind, and $b_{\mathbf{k},\nu} = k_{\perp}^2 \rho_{\text{th},\nu}^2 / 2$.

2.3.3 Electrostatic and Boltzmann Electron Response Limit

When the ratio of plasma to magnetic pressure, known as plasma β , tends to zero, the magnetic field perturbations also tend to zero. This ‘electrostatic’ limit is a reasonable approximation in many MCF plasmas due to the strong confining magnetic field. In order to extract the electrostatic limit from the equations, one can simply set $A_{\parallel} = \delta B_{\parallel} = 0$. With this, the distribution function reduces to the electrostatic limit $g_{\nu} = \delta f_{\nu} + (Z_{\nu} e / T_{\nu}) F_{0,\nu} [\phi - \langle \phi \rangle_{\mathbf{R}_{\nu}}] = \langle \delta f_{\nu} \rangle_{\mathbf{R}_{\nu}}$.

A convenient feature of the electrostatic limit is that a further simplification can be made for the electron dynamics. Taking the limit of $\sqrt{m_e / m_i} \ll 1$, the electron inertia becomes negligible, and one finds that the parallel force balance for the electrons reduces to

$$e \nabla_{\parallel} \phi = T_e \nabla_{\parallel} \ln(n_{\text{total},e}), \quad (2.48)$$

which gives a Boltzmann distribution for the density: $n_{\text{total},e} = \bar{n}_e \exp(e\phi / T_e)$, where $\bar{n}_e$ is a constant of integration. Recalling the assumption $e\phi / T_e \ll 1$, then $n_{\text{total},e} \approx \bar{n}_e (1 + e\phi / T_e)$, such that the equilibrium and perturbed densities can be defined as

$n_e = \bar{n}_e - c_0$ and $\delta n_e \approx n_e(e\phi/T_e) + c_0$ respectively, where c_0 is a constant that one is free to define. With this, the non-Boltzmann contribution to the electron distribution function evaluates to zero and $\delta f_e = (e\phi/T_e + c_0/n_e)F_{0,e}$. This corresponds to the electrons being of sufficiently light mass that they can instantaneously respond to the perturbed electrostatic potential set up by the ions along the field line. This defines the perturbed electron density up to a constant. By further insisting that the flux surface average of the perturbed density is zero (i.e., it is not contributing to the equilibrium density), we uniquely define the density through the ‘modified Boltzmann response’, $\delta n_e = en_e(\phi - \Phi)/T_e$ ³, where Φ indicates a flux-surface average of ϕ , defined as

$$\Phi \doteq \frac{\int d\alpha dz \mathcal{J}(\alpha, z) \phi}{\int d\alpha dz \mathcal{J}(\alpha, z)}, \quad (2.49)$$

with $\mathcal{J}$ the α -dependent Jacobian. Thus, the electron distribution function is not evaluated explicitly; rather, it is determined from the electric potential. The electrostatic, Boltzmann-electron limit of quasineutrality is:

$$\sum_{\nu \neq e} Z_\nu e \left\{ \frac{2\pi B_0}{m_\nu} \int d^2v J_{0,\mathbf{k},\nu} \hat{g}_{\mathbf{k},\nu} + \frac{Z_\nu en_\nu}{T_\nu} (\Gamma_{0,\mathbf{k},\nu} - 1) \hat{\phi}_{\mathbf{k}} \right\} - \frac{e^2 n_e}{T_e} (\hat{\phi}_{\mathbf{k}} - \hat{\Phi}_{\mathbf{k}}) = 0. \quad (2.50)$$

2.4 Concluding Remarks

This chapter has outlined the magnetic geometry in MCF devices, and provided the fundamental assumptions and orderings required to derive the gyrokinetic system through an asymptotic expansion in ρ_* . The gyrokinetic approach significantly enhances the tractability of predicting plasma evolution, by averaging over the fast timescales associated with the particle’s gyromotion, and reducing the dimensionality by one. This has established gyrokinetics as a powerful tool for predicting the evolution of plasma turbulence. The method assumes a separation of spatial and temporal scales within the plasma, allowing the equilibrium state to be treated independently from the turbulent fluctuations, with turbulence evolving on top of an established equilibrium.

However, the full nonlinear gyrokinetic system is inherently complex and cannot be solved analytically in its entirety. Consequently, the use of gyrokinetic codes is essential for obtaining numerical solutions to these equations. This topic will be the primary focus of Chapter 3.

³Note that $c_0 = 0$ is another common choice in defining the equilibrium and perturbed density contributions, and is known as the ‘adiabatic’ electron response

A further limit of these equations, which is often applicable and is in many gyrokinetic codes, has also been considered; the local-in- α limit. This is the limit that flux-tube codes, which will be discussed in Section 3.4, assume. Whilst this is an important limit to consider, it has its limitations in predicting turbulent evolution within non-axisymmetric devices, and may lead to inconsistencies in the zonal flow evolution. The numerical extension to include these effects will be the focus of Chapter 4.

Chapter 3

Gyrokinetic Modelling using a δf code

We must think outside the axisymmetric box.

Georgia Acton

3.1 Introduction

In recent years, gyrokinetic modeling has emerged as a powerful tool for understanding plasma behavior in fusion devices. Within the current landscape of gyrokinetic modelling, there exists a need for codes that can efficiently capture turbulent phenomena in complex magnetic geometries. As the fusion community progresses toward more sophisticated devices, such as stellarators and advanced tokamaks, the demand for reliable simulation tools grows. This chapter explores the construction of a δf -code for gyrokinetic simulations. In an ideal scenario, the turbulent fluxes computed from these codes would be input into a transport solver, enabling the equilibrium plasma to be self-consistently evolved for comprehensive modelling of the full MCF device (e.g., TGYRO [21], Trinity [7], Tango [91], GENE+Tango [30], Trinity3D (T3D) [99]). Consequently, it is important to develop these codes efficiently to allow for such analysis on current computing resources.

When constructing a δf -code for gyrokinetic modeling, several critical algorithmic decisions must be addressed. Selecting the appropriate framework for particle representation is an important consideration, with two primary options: particle-in-cell (PIC) [80, 93, 33], or continuum methods.

For a PIC code one adopts a Lagrangian approach, following particle trajectories. As it is not feasible to do this for every particle in the plasma, ‘macro-particles’ are used to represent many particles, with time-evolving weights. These codes work in real space, and can be used locally or to capture global effects across the full

device. They are comparatively easy to implement and parallelise, and can produce accurate results with enough particles per cell. PIC codes can often take larger time steps than continuum codes without associated numerical instabilities. However they can be computationally intensive due to the need for numerous particles per cell to achieve statistical accuracy. Without sufficient numbers of particles per cell, PIC codes are affected by particle noise, arising from numerical errors encountered when evaluating the moments of the distribution function using particles in phase space. PIC codes may also lead to particle bunching, resulting in sparse numbers of particles in some regions of the domain, under-resolving the full dynamics, and it can introduce statistical fluctuations in low-density regions. Finally, PIC codes work in real space; while this can ease computational implementation, it also means they can incur challenges when evaluating particle gyromotion. In order to accurately capture the Finite Larmor Radius (FLR) effects arising from particle gyro-orbits, a dense grid perpendicular to the magnetic field is necessary. Issues associated with gyro-orbits can also appear at the boundaries of the domain where a particle is unable to complete a full gyro-orbit.

Continuum codes instead solve the kinetic equation to model the distribution function using Eulerian or semi-Lagrangian methods. These have the advantage that they can use high-order velocity-space integration algorithms and spectral methods for improved accuracy. However they can be more complicated to implement and pose a greater challenge with parallelisation. The choice between spectral and real space methods is an important consideration when developing such a continuum code. Spectral methods, which can be applied in both velocity space and in the spatial domain, can significantly improve accuracy when evaluating derivatives, and improve computational efficiency and memory management within the code. In particular, a spatial spectral method in the perpendicular domain allows for a more precise treatment of the gyromotion of particles at a given numerical resolution, enabling FLR effects to be captured accurately. Careful consideration is required to balance the benefits of spectral methods against the potential computational complexities they may introduce.

Beyond the choice of spectral or real-space particle representation, the choice between using an Eulerian or a semi-Lagrangian method also needs to be considered, as each offers different advantages and disadvantages. Semi-Lagrangian methods provide a hybrid approach, combining the strengths of both Eulerian and Lagrangian techniques. These allow one to use a spectral approach—associated with Eulerian schemes—with the enhanced numerical stability associated with Lagrangian schemes.

The algorithms are relatively simple to implement, but they require interpolations which can be slow and inaccurate, and can also introduce diffusion [112]. Their dependence on careful grid point selection in velocity space can also pose challenges. They also have limited applicability, being useful only in the context of advancing advection equations.

Eulerian methods can be implemented using explicit or implicit methods. Explicit time-stepping methods are typically simpler to implement and computationally lighter, enabling faster simulations; yet, they often require smaller time steps to maintain numerical stability, which can limit application and efficiency. Explicit methods can lead to numerical instabilities if the Courant–Friedrichs–Lewy (CFL) condition is violated, and may also introduce artificial numerical diffusion. In contrast implicit time-stepping methods can allow larger time steps without compromising numerical stability, making them particularly useful in scenarios with stringent CFL restrictions. For gyrokinetic simulations, explicit time-integration methods are constrained to timesteps typically ranging from 10^{-6} to 10^{-5} seconds. Conversely, implicit approaches enable substantially longer timesteps, often of the order 10^{-3} to 10^{-2} seconds. However, despite their ability to handle larger timesteps, implicit methods often require solving complex algebraic equations at each time step, which can significantly increase the difficulty of implementation. Thus, while Eulerian frameworks offer a robust platform, the choice between implicit and explicit time-stepping is another important consideration that influences the overall performance and accuracy of the code.

Another critical decision involves choosing the coordinate and grid system to work in, particularly for the velocity space coordinates [8]. This is especially crucial when working in the weakly collisional regime. In the absence of collisions or other forms of dissipation, the distribution of particles in velocity space can form arbitrarily small-scale structures, necessitating a large number of grid points to accurately resolve these features [74]. One common choice is to use quantities that are constants of motion, such as energy, $\mathcal{E}$, and a quantity related to the first adiabatic invariant, $\lambda = \mu_\nu/\mathcal{E}$ [73, 22]. This choice removes all velocity space derivatives from the collisionless gyrokinetic equation, thereby simplifying the equations. However, this can lead to difficulties when trying to capture the bounce points of trapped electrons in non-axisymmetric devices. Either the spacing of the velocity space grid points needs to be chosen such that the particle bounce points are included in the grid, or interpolation is required in the velocity grid to accurately capture the effects of trapped

electrons. As there can be many magnetic wells on a single field line, and the locations of these wells will vary across field lines, this choice of velocity coordinates makes the treatment of multiple field lines for a non-axisymmetric device potentially complicated. Conversely, using $(v_{\parallel}, \mu_{\nu})$ coordinates for the velocity-space simplifies this process, but consequently introduces velocity derivatives into the system equations, and can complicate the form of the collision operator. Each of these decisions carries implications for the accuracy, efficiency, and applicability of the gyrokinetic code.

In this section the formalism for a pseudo-spectral, mixed implicit-explicit δf -gyrokinetic continuum code in $(k_{\alpha}, k_{\psi}, z, v_{\parallel}, \mu_{\nu})$ coordinates is outlined. These choices yield the advantage of maintaining spectral accuracy in the perpendicular derivatives and gyroaverages, while also facilitating fast parallel dynamics through an implicit algorithm for advancing motion along the magnetic field, and allowing for treatment of the complex parallel boundary condition. The specific coordinate choices, normalisations, and algorithmic decisions of the code **stella** are introduced. This will serve as the foundation for developing the full flux-annulus gyrokinetic code, designed to capture the α -coupling effects described in Section 2.3.1. The development of this capability is detailed in Chapter 4. The flux-tube version of the code is also used to perform the numerical benchmarks in Chapter 5.

It is important to note that many other gyrokinetic codes are available. Notable examples for PIC methods are GTC [81], ORB5 [78], GTS [121], XGC [76], EUTERPE [72], and see also [92]. For continuum methods, notable examples include GS2 [73], GENE [70], CGYRO [22], GKV [122], GYSELA [48], GKW [95], GENE-3D [85], and GX [83].

3.2 Coordinates

We work in field-aligned coordinates [11], (x, y, z) , where z measures the position along a given magnetic field line, and (x, y) represents position in the plane perpendicular to $\hat{\mathbf{z}} \equiv \hat{\mathbf{b}}$. One advantage of using field-aligned coordinates is that turbulent structures are often highly elongated along the field lines, i.e., $k_{\perp}/k_{\parallel} \gg 1$. Consequently, a much coarser grid can be employed in the coordinate parallel to the magnetic field. The coordinates (x, y) are related to the flux coordinates (ψ, α) via

$$x = \frac{dx}{d\psi} (\psi - \psi_0), \quad (3.1a)$$

$$y = \frac{dy}{d\alpha} (\alpha - \alpha_0), \quad (3.1b)$$

with (ψ_0, α_0) denoting the values of (ψ, α) at the centre of the perpendicular domain. The only other generic constraint that is placed on the coordinates is that x increases monotonically from the magnetic axis out to the plasma edge. As described in [9], the exact choice of (x, y) depends on the choice of model geometry. The parallel dynamics are treated using the real space coordinate z to correctly capture the parallel boundary conditions, which will be described in Section 3.5.1. Common choices for z are the toroidal angle, ζ , the poloidal angle, θ , and the arc length along the field line. Together the unit vectors $\{\hat{\mathbf{x}}, \hat{\mathbf{y}}, \hat{\mathbf{z}}\}$ form a left-handed, orthonormal basis, and a uniform grid in each of the coordinates in (k_x, k_y, z) -space is taken.

The choice of velocity-space coordinates is $(v_{\parallel}, \mu_{\nu})$, with $v_{\parallel}$ the velocity parallel to the magnetic field, $v_{\parallel} = \mathbf{v} \cdot \hat{\mathbf{b}}$, and μ_{ν} the magnetic moment. A uniformly spaced grid in $v_{\parallel}$ is taken, whilst the option of either a uniform grid, or a Gauss–Laguerre grid [2] is available in the magnetic moment coordinate. This latter choice provides spectral accuracy in the μ_{ν} integral, and allows one to more densely sample the region near $\mu_{\nu} = 0$ for better resolution.

3.3 Normalisations

When implementing the gyrokinetic equation (2.25) into a code, the variables are first normalised based on physical considerations, to make all terms of order unity. A list of the normalised quantities are provided in Table 3.1. It is important to note the two different normalisations for the perpendicular coordinates; quantities relating to the equilibrium, such as background gradients, are normalised by L_r , whilst turbulent quantities are normalised using $\rho_r = v_{\text{th},r}/\Omega_r$, which is defined in terms of a user-specified reference mass m_r , and temperature T_r , with the reference length L_r and magnetic field strength B_r being determined by the magnetic geometry specified. The reference density, n_r , is also a user-specified input.

The gyrokinetic equation is normalised by taking its product with the factor $(L_r^2/\rho_r v_{\text{th},r})\exp(-v^2/v_{\text{th},\nu}^2)/F_{0,\nu}$. In real space, this is schematically written as:

$$\frac{\partial \tilde{g}_{\nu}}{\partial \tilde{t}} + \tilde{\mathcal{S}}_{\nu}[\tilde{g}_{\nu}, \tilde{\chi}_{\nu}] + \tilde{\mathcal{M}}_{\nu}[\tilde{g}_{\nu}] + \tilde{\mathcal{D}}_{\nu}[\tilde{g}_{\nu}, \tilde{\chi}_{\nu}] + \tilde{\mathcal{G}}_{\nu}[\tilde{\chi}_{\nu}] + \tilde{\mathcal{A}}[\tilde{A}_{\parallel,\nu}] + \tilde{\mathcal{N}}_{\nu}[\tilde{g}_{\nu}, \tilde{\chi}_{\nu}] = \tilde{\mathcal{C}}_{\nu}[\{\tilde{g}_{\nu'}, \tilde{\chi}_{\nu'}\}], \quad (3.2)$$

Table 3.1: List of normalised parameters and variables used in the δf -gyrokinetic code **stella**.

<i>Normalised Parameters</i>		
<i>Parameter</i>	<i>Normalisation</i>	<i>Normalised Quantity</i>
$\tilde{t}$	$t L_r / v_{th,r}$	Time
$\tilde{x}$	x / L_r	Radial perpendicular coordinate
x^{tb}	x / ρ_r	Turbulent radial coordinate
$\tilde{y}$	x / L_r	Binormal perpendicular coordinate
y^{tb}	y / ρ_r	Turbulent binormal coordinate
$\tilde{z}$	z / L_r	Parallel coordinate
$\tilde{k}_x$	$k_x \rho_r$	Radial wavenumber
$\tilde{k}_y$	$k_y \rho_r$	Binormal wavenumber
$\tilde{k}_\perp$	$k_\perp \rho_r$	Perpendicular wavenumber
$\tilde{\nabla}$	$L_r \nabla$	Gradient operator
$\tilde{\psi}$	$\psi / (L_r^2 B_r)$	Magnetic flux
$\tilde{v}_\parallel$	$v_\parallel / v_{th,\nu}$	Parallel velocity
$\tilde{\mu}_\nu$	$\mu_\nu B_r / 2T_\nu$	Magnetic moment
$\tilde{v}_{th,\nu}$	$v_{th,\nu} / v_{th,r}$	Species thermal velocity
$\tilde{T}_\nu$	T_ν / T_r	Species temperature
$\tilde{m}_\nu$	m_ν / m_r	Species mass
$\tilde{n}_\nu$	n_ν / n_r	Species density
$\tilde{B}_0$	B_0 / B_r	Magnetic field
$\delta \tilde{n}_\nu$	$(L_r / \rho_r n_r) \delta n_\nu$	Perturbed species density
$\tilde{F}_{0,\nu}$	$(v_{th,\nu}^3 / n_\nu) F_{0,\nu}$	Equilibrium distribution function
$\tilde{g}_\nu$	$(L_r / \rho_r) (\exp(-\tilde{v}^2) / F_{0,\nu}) g_\nu$	Guiding centre distribution function
$\tilde{\phi}$	$(e / T_r) (L_r / \rho_r) \phi$	Perturbed electrostatic potential
$\tilde{\varphi}_\nu$	$\langle \tilde{\phi} \rangle_{\mathbf{R}_\nu}$	Gyroaveraged electrostatic potential
$\tilde{A}_\parallel$	$(L_r / B_r \rho_r^2) A_\parallel$	Perturbed parallel magnetic potential
$\delta \tilde{B}_\parallel$	$(L_r / B_r \rho_r) \delta B_\parallel$	Perturbed parallel magnetic field
$\tilde{\chi}_\nu$	$\tilde{\phi} - 2\tilde{\mathbf{v}} \cdot \tilde{\mathbf{A}}$	Perturbed gyrokinetic potential

with normalised operator definitions

$$\tilde{\mathcal{S}}_\nu[\tilde{g}_\nu, \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}] \doteq \tilde{v}_{\text{th},\nu} \tilde{v}_\parallel \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z} \left(\frac{\partial \tilde{g}_\nu}{\partial \tilde{z}} + \frac{Z_\nu}{\tilde{T}_\nu} e^{-\tilde{v}_\nu^2} \frac{\partial \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}}{\partial \tilde{z}} \right), \quad (3.3a)$$

$$\tilde{\mathcal{M}}_\nu[\tilde{g}_\nu] \doteq -\tilde{v}_{\text{th},\nu} \tilde{\mu}_\nu \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{B}_0 \frac{\partial \tilde{g}_\nu}{\partial \tilde{v}_\parallel}, \quad (3.3b)$$

$$\begin{aligned} \tilde{\mathcal{D}}_\nu[\tilde{g}_\nu, \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}] &\doteq \frac{\mathbf{v}_{M,\nu}}{v_{\text{th},r}} \cdot \left(\tilde{\nabla} \tilde{g}_\nu + \frac{Z_\nu}{\tilde{T}_\nu} e^{-\tilde{v}_\nu^2} \tilde{\nabla} \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu} \right), \\ &= \frac{\hat{\mathbf{b}}}{\tilde{B}_0} \times \left(\tilde{\mu}_\nu \tilde{\nabla} \tilde{B}_0 + \tilde{v}_\parallel^2 \hat{\mathbf{b}} \cdot \tilde{\nabla} \hat{\mathbf{b}} \right) \cdot \tilde{\nabla} \tilde{x} \left(\frac{\tilde{T}_\nu}{Z_\nu} \frac{\partial \tilde{g}_\nu}{\partial x^{\text{tb}}} + e^{-\tilde{v}_\nu^2} \frac{\partial \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}}{\partial x^{\text{tb}}} \right) \\ &\quad + \frac{\hat{\mathbf{b}}}{\tilde{B}_0} \times \left(\tilde{\mu}_\nu \tilde{\nabla} \tilde{B}_0 + \tilde{v}_\parallel^2 \hat{\mathbf{b}} \cdot \tilde{\nabla} \hat{\mathbf{b}} \right) \cdot \tilde{\nabla} \tilde{y} \left(\frac{\tilde{T}_\nu}{Z_\nu} \frac{\partial \tilde{g}_\nu}{\partial y^{\text{tb}}} + e^{-\tilde{v}_\nu^2} \frac{\partial \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}}{\partial y^{\text{tb}}} \right), \end{aligned} \quad (3.3c)$$

$$\tilde{\mathcal{G}}_\nu[\langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}] \doteq \frac{1}{2} \frac{\partial \tilde{x}}{\partial \tilde{\psi}} \frac{\partial \tilde{y}}{\partial \alpha} \frac{\partial \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}}{\partial y^{\text{tb}}} e^{-\tilde{v}_\nu^2} \left(\frac{\partial \ln \tilde{n}_\nu}{\partial \tilde{x}} + \frac{\partial \ln \tilde{T}_\nu}{\partial \tilde{x}} \left(\tilde{v}_\parallel^2 + 2\tilde{\mu}_\nu \tilde{B}_0 - \frac{3}{2} \right) \right), \quad (3.3d)$$

$$\tilde{\mathcal{A}}_\nu[\tilde{A}_{\parallel,\nu}] \doteq 2 \frac{Z_\nu}{\tilde{m}_\nu} \tilde{\mu}_\nu \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{B}_0 e^{-\tilde{v}_\nu^2} J_{0,\mathbf{k},\nu} \tilde{A}_{\parallel,\mathbf{k}}, \quad (3.3e)$$

$$\tilde{\mathcal{N}}_\nu[\tilde{g}_\nu, \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}] \doteq \frac{\partial y^{\text{tb}}}{\partial \alpha} \frac{\partial x^{\text{tb}}}{\partial \tilde{\psi}} \frac{B_r \rho_r^2}{2} \left(\frac{\partial \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}}{\partial y^{\text{tb}}} \frac{\partial \tilde{g}_\nu}{\partial x^{\text{tb}}} - \frac{\partial \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu}}{\partial x^{\text{tb}}} \frac{\partial \tilde{g}_\nu}{\partial y^{\text{tb}}} \right), \quad (3.3f)$$

and $\tilde{\mathcal{C}}_\nu[\{\tilde{g}_{\nu'}, \tilde{\varphi}_{\nu'}\}]$ the appropriately normalised collision operator. The quasineutrality constraint, equation (2.29), is normalised by taking its product with the factor $(L_r/\rho_r e n_r)$,

$$\frac{2\tilde{B}_0}{\pi^{1/2}} \sum_\nu Z_\nu \tilde{n}_\nu \int_{-\infty}^{\infty} d\tilde{v}_\parallel \int_0^{\infty} d\tilde{\mu}_\nu \left(\langle \tilde{g}_\nu \rangle_{\mathbf{r}} + \frac{Z_\nu}{\tilde{T}_\nu} \left(\langle \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu} \rangle_{\mathbf{r}} - \tilde{\phi} \right) e^{-\tilde{v}^2} \right) = 0. \quad (3.4)$$

To normalise parallel Ampère's law, equation (2.30), we take its product with the factor (L_r/B_r)

$$\tilde{\nabla}_\perp^2 \tilde{A}_{\parallel,\nu} = \frac{2\tilde{B}_0 \beta_r}{\pi^{1/2}} \sum_\nu Z_\nu \tilde{n}_\nu \tilde{v}_{\text{th},\nu} \int_{-\infty}^{\infty} d\tilde{v}_\parallel \int_0^{\infty} d\tilde{\mu}_\nu \tilde{v}_\parallel \left(\langle \tilde{g}_\nu \rangle_{\mathbf{r}} + \frac{Z_\nu \tilde{v}_{\text{th},\nu}}{\tilde{T}_\nu} \tilde{v}_\parallel \left\langle \langle \tilde{A}_{\parallel,\nu} \rangle_{\mathbf{R}_\nu} \right\rangle_{\mathbf{r}} e^{-\tilde{v}^2} \right), \quad (3.5)$$

with $\beta_r = 8\pi n_r T_r/B_r^2$. To normalise perpendicular Ampère's law, equation (2.31), we take its product with the factor $(L_r \rho_r/B_r)$

$$\tilde{\nabla}_\perp^2 \delta \tilde{B}_\parallel = \frac{2\tilde{B}_0 \beta_r}{\pi^{1/2}} \sum_\nu Z_\nu \tilde{n}_\nu \int_{-\infty}^{\infty} d\tilde{v}_\parallel \int_0^{\infty} d\tilde{\mu}_\nu \tilde{\nabla} \cdot \left\langle \left(\tilde{g}_\nu + \frac{Z_\nu}{\tilde{T}_\nu} \langle \tilde{\chi}_\nu \rangle_{\mathbf{R}_\nu} e^{-\tilde{v}^2} \right) (\hat{\mathbf{b}} \times \tilde{\mathbf{v}}_\perp) \right\rangle_{\mathbf{r}}. \quad (3.6)$$

For simplicity, the tilde notation denoting the normalisations will be dropped in the remainder of this chapter, and until Chapter 5¹.

3.4 Flux Tube Equations in stella

Modelling turbulence in a plasma can be computationally intensive, so simplifying the domain is advantageous for reducing costs. One such simplification is the use of a flux tube [51, 12]; this is a ‘local’ domain that follows a magnetic field line, as depicted in Figure 3.1. This approximation results in the equations becoming local-in- α , reducing to those given in Section 2.3.2.

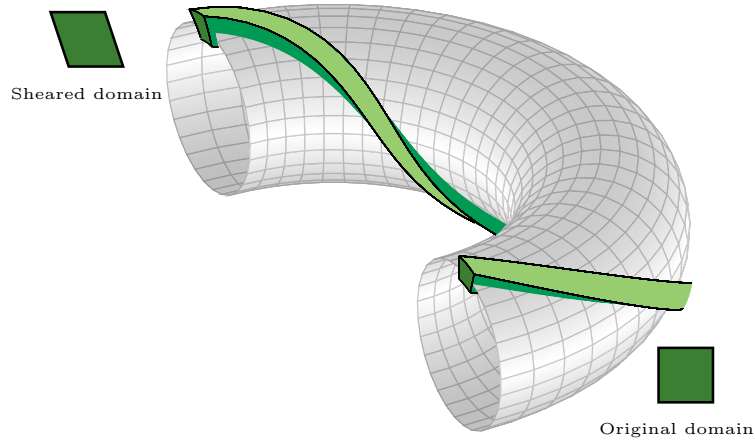


Figure 3.1: Schematic of a flux tube for an axisymmetric device. The meshed surface depicts a flux surface, while the green cuboid represents the flux-tube simulation domain. As one follows the field line, the perpendicular domain experiences shearing. The original and sheared domains are shown.

In directions perpendicular to the magnetic field, $\mathbf{B}$, the flux tube is considered to be sufficiently wide (on the order of several thermal ion Larmor radii, $\rho_{\text{th},i}$) so that turbulent fluctuations on one side of the domain become decorrelated with those on the opposite side. However, the flux tube must remain much narrower than the equilibrium length scale, L_r , allowing equilibrium properties and associated drive terms to be treated as constant across the perpendicular domain. This condition corresponds to the limit in which $\rho_* \rightarrow 0$. This local approximation guarantees that fluctuations at either end of the perpendicular domain are not only decorrelated but also statistically identical, as the same driving factors produce equivalent turbulence

¹The electromagnetic version of **stella** uses a mixed formalism, that includes $\tilde{g}_\nu$ and $\tilde{h}_\nu = \tilde{g}_\nu + Z_\nu/T_\nu e^{-\tilde{v}_\nu^2} \langle \tilde{\chi} \rangle$, to solve these sets of equations. This will not be detailed here.

on average over time, meaning the perpendicular directions can be treated spectrally via a discrete Fourier transform. Computationally, this has the benefit that spectral accuracy may be retained in the perpendicular derivatives, and can capture FLR effects correctly without the need for a dense grid in the perpendicular domain. To implement the Fourier treatment, the evolution equations are decomposed into Fourier harmonics in (x, y) space, and the discrete Fourier transform is taken in x and y :

$$g_\nu(x, y, z, v_\parallel, \mu_\nu, t) = \sum_{k_x, k_y} \hat{g}_{\mathbf{k}, \nu}(z, v_\parallel, \mu_\nu, t) e^{i(k_x x + k_y y)}. \quad (3.7)$$

Next, the local-in- α evolution equations (2.42)-(2.47) are normalised, and all variables are taken to the left hand side to define the objective functions $\hat{G}_{\mathbf{k}, \nu}$, $\hat{Q}_{\mathbf{k}}$, $\hat{M}_{\mathbf{k}}$, and $\hat{N}_{\mathbf{k}}$ that are identically zero. The transformed kinetic evolution equation for species ν , written in terms of normalised variables, is given by

$$\hat{G}_{\mathbf{k}, \nu} \doteq \hat{\mathcal{S}}_{\mathbf{k}, \nu} + \hat{\mathcal{M}}_{\mathbf{k}, \nu} + \hat{\mathcal{D}}_{\mathbf{k}, \nu} + \hat{\mathcal{G}}_{\mathbf{k}, \nu} + \hat{\mathcal{A}}_{\mathbf{k}, \nu} + \hat{\mathcal{N}}_\nu - \hat{\mathcal{C}}_\nu = 0, \quad (3.8)$$

with associated normalised operator definitions

$$\hat{\mathcal{S}}_{\mathbf{k}, \nu} = v_{\text{th}, \nu} v_\parallel \hat{\mathbf{b}} \cdot \nabla z \left[\frac{\partial \hat{g}_{\mathbf{k}, \nu}}{\partial z} + \frac{Z_\nu}{T_\nu} \frac{\partial \langle \hat{\chi} \rangle_{\mathbf{k}, \nu}}{\partial z} e^{-v_\nu^2} \right], \quad (3.9a)$$

$$\hat{\mathcal{M}}_{\mathbf{k}, \nu} = -v_{\text{th}, \nu} \mu_\nu \hat{\mathbf{b}} \cdot \nabla \hat{B}_0 \frac{\partial \hat{g}_{\mathbf{k}, \nu}}{\partial v_\parallel}, \quad (3.9b)$$

$$\hat{\mathcal{D}}_{\mathbf{k}, \nu} = i\omega_{d, \mathbf{k}, \nu} \left[\hat{g}_{\mathbf{k}, \nu} + \frac{Z_\nu}{T_\nu} \langle \hat{\chi} \rangle_{\mathbf{k}, \nu} e^{-v_\nu^2} \right], \quad (3.9c)$$

$$\hat{\mathcal{G}}_{\mathbf{k}, \nu} = i\omega_{*, \mathbf{k}, \nu} e^{-v_\nu^2} \langle \hat{\chi} \rangle_{\mathbf{k}, \nu}, \quad (3.9d)$$

$$\hat{\mathcal{A}}_{\mathbf{k}, \nu} = 2 \frac{Z_\nu}{m_\nu} \mu \hat{\mathbf{b}} \cdot \nabla B_0 e^{-v_\nu^2} J_{0, \mathbf{k}, \nu} \hat{A}_{\parallel, \mathbf{k}}, \quad (3.9e)$$

$$\hat{\mathcal{N}}_{\mathbf{k}, \nu} = \frac{B_r}{2} \frac{dy}{d\alpha} \frac{dx}{d\psi} \mathcal{F}_{\mathbf{k}} \left[\mathcal{F}_{\mathbf{k}}^{-1} [ik_y \rho_r \hat{\chi}_{\mathbf{k}, \nu}] [ik_x \rho_r \hat{g}_{\mathbf{k}, \nu}] - \mathcal{F}_{\mathbf{k}}^{-1} [ik_x \rho_r \hat{\chi}_{\mathbf{k}, \nu}] [ik_y \rho_r \hat{g}_{\mathbf{k}, \nu}] \right], \quad (3.9f)$$

and $\hat{\mathcal{C}}_\nu$ is the normalised collision operator. The corresponding transformed, nor-

malised field equations are given by

$$\hat{Q}_{\mathbf{k}} = \sum_{\nu} Z_{\nu} n_{\nu} \left\{ \frac{2B_0}{\sqrt{\pi}} \int d^2v J_{0,\mathbf{k},\nu} \hat{g}_{\mathbf{k},\nu} + \frac{Z_{\nu}}{T_{\nu}} (\Gamma_{0,\mathbf{k},\nu} - 1) \hat{\phi}_{\mathbf{k}} + \frac{1}{B_0} \Gamma_{1,\mathbf{k},\nu} \delta \hat{B}_{\parallel,\mathbf{k}} \right\}, \quad (3.10)$$

$$\hat{M}_{\mathbf{k}} = - \frac{\beta_r}{(k_{\perp} \rho_r)^2} \sum_{\nu} Z_{\nu} n_{\nu} v_{\text{th}} \frac{2B_0}{\sqrt{\pi}} \int d^2v v_{\parallel} J_{0,\mathbf{k},\nu} \hat{g}_{\mathbf{k},\nu} + \left[1 + \frac{\beta_r}{(k_{\perp} \rho_r)^2} \sum_{\nu} \frac{Z_{\nu} n_{\nu}}{m_{\nu}} \Gamma_{0,\mathbf{k},\nu} \right] \hat{A}_{\parallel,\mathbf{k}}, \quad (3.11)$$

$$\begin{aligned} \hat{N}_{\mathbf{k}} = & 2\beta_r \sum_{\nu} n_{\nu} T_{\nu} \frac{2B_0}{\sqrt{\pi}} \int d^2v \mu_{\nu} \frac{J_{0,\mathbf{k},\nu}}{a_{\mathbf{k},\nu}} \hat{g}_{\mathbf{k},\nu} + \left[\frac{\beta_r}{2B_0} \sum_{\nu} Z_{\nu} n_{\nu} \Gamma_{1,\mathbf{k},\nu} \right] \hat{\phi}_{\mathbf{k}} \\ & + \left[1 + \frac{\beta_r}{2B_0} \sum_{\nu} Z_{\nu} n_{\nu} T_{\nu} \Gamma_{2,\mathbf{k},\nu} \right] \delta \hat{B}_{\parallel,\mathbf{k}}, \end{aligned} \quad (3.12)$$

with the normalised gyrokinetic potential, diamagnetic frequency, and drift frequency given in Table 3.2.

Table 3.2: List of normalised parameters and variables

<i>Normalised Variables</i>	
$\langle \hat{\chi} \rangle_{\mathbf{k},\nu}$	$= J_{0,\mathbf{k},\nu} \hat{\phi}_{\mathbf{k}} - 2v_{\text{th}} v_{\parallel} J_{0,\mathbf{k},\nu} \hat{A}_{\parallel,\mathbf{k}} + 4\mu_{\nu} (T_{\nu}/Z_{\nu}) (J_{1,\mathbf{k},\nu}/a_{\mathbf{k},\nu}) \delta \hat{B}_{\parallel,\mathbf{k}}$
$\omega_{d,\mathbf{k},\nu}$	$= (T_{\nu} \rho_r / Z_{\nu} B_0) (v_{\parallel}^2 \hat{\mathbf{b}} \times (\hat{\mathbf{b}} \cdot \nabla \hat{\mathbf{b}}) + \mu_{\nu} \hat{\mathbf{b}} \times \nabla B_0) \cdot (k_x \nabla x + k_y \nabla y)$
$\omega_{*,\mathbf{k},\nu}$	$= (k_y \rho_r / 2) L_r B_r (dy/d\alpha) (d \ln F_{0,\nu} / d\psi)$

3.5 Boundary Conditions

The spectral flux-tube approach applies a periodic boundary condition in the x and y dimensions. The incoming boundary conditions along the magnetic field on g_{ν} are taken to be $\hat{g}_{\nu}(z \rightarrow -\infty, v_{\parallel} > 0, \mu) \rightarrow 0$, and $\hat{g}_{\nu}(z \rightarrow \infty, v_{\parallel} < 0, \mu) \rightarrow 0$. The boundary condition in $v_{\parallel}$ is $\hat{g}_{\nu}(z, v_{\parallel} \rightarrow \pm\infty, \mu) = 0$.

3.5.1 Parallel Boundary Condition; twist-and-shift

For the parallel boundary conditions one must enforce the constraint that any physical quantity remains single-valued and continuous throughout the domain. Consider a device with magnetic geometry that is periodic in ζ with period $2\pi/N_p$, where N_p defines the period of the device. A segment of that device, with length $2\pi p$ in the z direction, is considered, where $p = M/N_p$ and M is any integer. Any physical quantity, $A(t, x, y(\theta, z), z)$, considered over this segment must itself be physically periodic

across the domain; $A(t, x, y^*, z + 2p\pi) = A(t, x, y(\theta, z), z)$, where $y^* = y(\theta, z + 2p\pi)$. In an axisymmetric device, the ends of the flux tube domain can be placed at any two points with the same poloidal location, since such points share identical geometry and thus the turbulence is statistically equivalent. In contrast, for non-axisymmetric devices, it is more challenging to identify endpoints with matching geometry. To address this, turbulence simulations in non-axisymmetric magnetic configurations often use an up-down symmetric field line, making it easier to end the flux tube in geometrically equivalent locations. The quantity A can be expanded in terms of its Fourier harmonics

$$A(t, x, y, z) = \sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z) e^{ik_y y + ik_x x}, \quad (3.13)$$

and physical periodicity is then enforced. The periodicity condition can then be written as

$$\sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z) e^{ik_y y + ik_x x} = \sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z + 2\pi p) e^{ik_y y^* + ik_x x}. \quad (3.14)$$

Now expand

$$y^* = y + \frac{\partial y}{\partial z} 2\pi p = y + 2\pi p \frac{\partial y}{\partial \alpha} \frac{\partial \alpha}{\partial z} = y - 2\pi p \iota(\psi) \frac{\partial y}{\partial \alpha},$$

where ι is the rotational transform introduced in equation (2.14), and $\iota(\psi) = \iota(\psi_0) + \iota'(\partial\psi/\partial x)x$, with $\iota' = (\partial\iota/\partial\psi)$. From this, a relationship between the coordinates $y^* = y - 2\pi p \iota(\psi_0)(\partial y/\partial \alpha) - 2x\pi p \iota'(\partial y/\partial \alpha)(\partial\psi/\partial x)$ is determined. Substituting this into (3.14) gives

$$\sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z) e^{ik_y y + ik_x x} = \sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z + 2\pi p) e^{ik_y y + i(k_x - \delta k_x)x} e^{i\Delta}. \quad (3.15)$$

Using the orthogonality of Fourier components, it is determined that $\hat{A}_{k_x, k_y}(t, z) = \hat{A}_{k_x + \delta k_x, k_y}(t, z + 2\pi p)$, with

$$\delta k_x = 2\pi p \iota' \frac{\partial y}{\partial \alpha} \frac{\partial \psi}{\partial x} k_y \quad \text{and} \quad \Delta = -2\pi p k_y \iota \frac{\partial y}{\partial \alpha}.$$

Physically, this captures the idea that as an eddy travels around the fusion device, it becomes stretched and tilted by the magnetic shear, which pushes the radial wavenumber to larger values as the eddy develops finer structure. It can also be seen that the amplitude, $\hat{A}_{k_x, k_y}(t, z)$ picks up a phase factor at the end of the z -domain. This encapsulates the fact that the field line does not travel back to the exact same α location

after traversing $2\pi p$ in z , but instead returns to a different location. This phase factor accounts for the geometric shift in the α coordinate at the ends of the domain.

The perpendicular domain consists of a uniform grid in k_x and k_y . The spacing in each grid is set by $\Delta k_x = 2\pi/L_x$ and $\Delta k_y = 2\pi/L_y$ where L_x and L_y are the respective sizes of the simulation domain in (x, y) . For each k_y , different k_x modes are linked together to form an extended domain in z , which is discussed in the next section, such that physical quantities are periodic at $z = 0$ and $z = 2\pi p$, before and after experiencing net magnetic shear respectively. A zero-incoming boundary condition is applied at the extrema of the extended z -domain, at $z = \pm z_{\text{max, ext}}$.

It is worth noting that in a flux-tube simulation the factor $pk_y\iota(\partial y/\partial\alpha)$ can be set to an integer value by considering a different flux surface that is a small distance of order $\mathcal{O}(\rho_*)$ away, such that $\iota(\psi)$ is sufficiently altered to make $pk_y\iota(\partial y/\partial\alpha) \in \mathbb{Z}$ [109]. This adjustment has the advantage of setting the phase factor to unity, which helps reduce the complexity of the flux tube code. In the case of small magnetic shear this approach may be problematic, as it can be seen that δk_x becomes very small, meaning one must either use a large number of poloidal turns, or must use a large box in the x -domain. This can be problematic from the perspective of memory and efficiency constraints as more k_x -modes are necessary. In this case, we can use the ‘phase-shift-periodic parallel boundary condition’ described by St-Onge et al. in [110]. In this case, the phase shift is retained, but δk_x is taken to zero

$$\sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z) e^{ik_y y + ik_x x} = \sum_{k_x, k_y} \hat{A}_{k_x, k_y}(t, z + 2\pi p) e^{ik_y y + ik_x x} e^{i\Delta}, \quad (3.16)$$

such that $\hat{A}_{k_x, k_y}(t, z) = \hat{A}_{k_x, k_y}(t, z + 2\pi p)$. Variations of periodic boundary conditions have been used in gyrokinetic simulations for stellarators with low magnetic shear [38], and these boundary conditions will be used in the simulations presented in the results section of Chapter 4².

3.5.2 Extended z -domain

It was shown in Section 3.5.1 that, when magnetic shear is present, modes with a given k_y are sheared to higher radial wavenumbers as the turbulent eddy is advanced along the field line. Although, in the flux-tube domain, the local value of k_x does not change, the flux tube domain is being sheared, so the ‘lab’ wavenumber is changed. Hence, a mode that has an incoming lab wavenumber, k_x , at the simulation boundary

²Other parallel boundary conditions exist within **stella**, such as the ‘flux-tube train’ [123], and a zero boundary condition on the non-extended z -domain, which will not be detailed here.

$z_{\min}$ will have a different value for the outgoing lab wavenumber, $z_{\max}$. In order to capture this shearing effect, one needs to couple different k_x modes across the ends of the flux tube.

Computationally, the physical periodicity condition—described by the ‘twist-and-shift’ boundary condition—imposes a numerical constraint on the k_x grid spacing. This arises because the modes at the end of the z -domain must be ‘linked’ to the incoming modes at the start of the domain. To ensure these outgoing modes are numerically captured, they must align with grid points in k_x . As a consequence, the quantity $\delta k_x / \Delta k_x = 2\pi \hat{s} L_x / L_y \doteq \text{jtwist}$ must be an integer, where $\hat{s} = -2x_0 / \iota (d\iota/dx)$ is the global magnetic shear, and x_0 is the value at the center of the x -domain. This condition ensures that the smallest change in radial wavenumber matches the grid spacing in the k_x -domain.

The twist-and-shift condition is applied at the end of each 2π segment in z , linking together segments with different k_x values. This results in an extended z -domain with $N_{z+} = N_z \times N_{\text{seg}}$ grid points, where N_z is the number of grid points in each of the 2π segments in z , and N_{seg} denotes the number of connected segments—that is, the number of 2π segments that are linked together to form the extended domain. This construction defines the ballooning-space domain upon which the zero incoming boundary condition is imposed.

Numerically, the parameter `jtwist` quantifies the number of connections for a given k_y . If `jtwist` = 1, then all k_x modes are connected for the smallest value of k_y ; if `jtwist` = 2, then every other k_x mode is connected. Figure 3.2 illustrates the connectivity of modes for two different `jtwist` values, for $k_y \in [0, 5]$ and $k_x \in [-5, 5]$. In this illustration, each arch represents which k_x modes are connected across one 2π simulation domain due to the magnetic shear. In order to capture the full dynamics of the system in the range $k_x \in [-5, 5]$ it is necessary to capture each of these connections up to $\pm k_{x,\max}$.

The number of distinct mode chains for a given k_y is given by $N_{\text{eigen}} = \text{jtwist}(k_y / \Delta k_y)$ for $k_y \neq 0$, and $N_{\text{eigen}} = N_{k_x}$ for $k_y = 0$, where N_{k_x} is the number of k_x modes. In Figure 3.2, this corresponds to the number of coloured chains for a given k_y . The maximum number of connected segments required to represent a mode in a given chain is $N_{\text{seg}} = (N_{k_x} - 1) / N_{\text{eigen}}$; in Figure 3.2, this is visualised as the number of arches within each coloured chain. For example, when `jtwist` = 1 and $k_y = 1$, all 11 k_x modes are connected, necessitating 11 segments to fully resolve the mode dynamics.

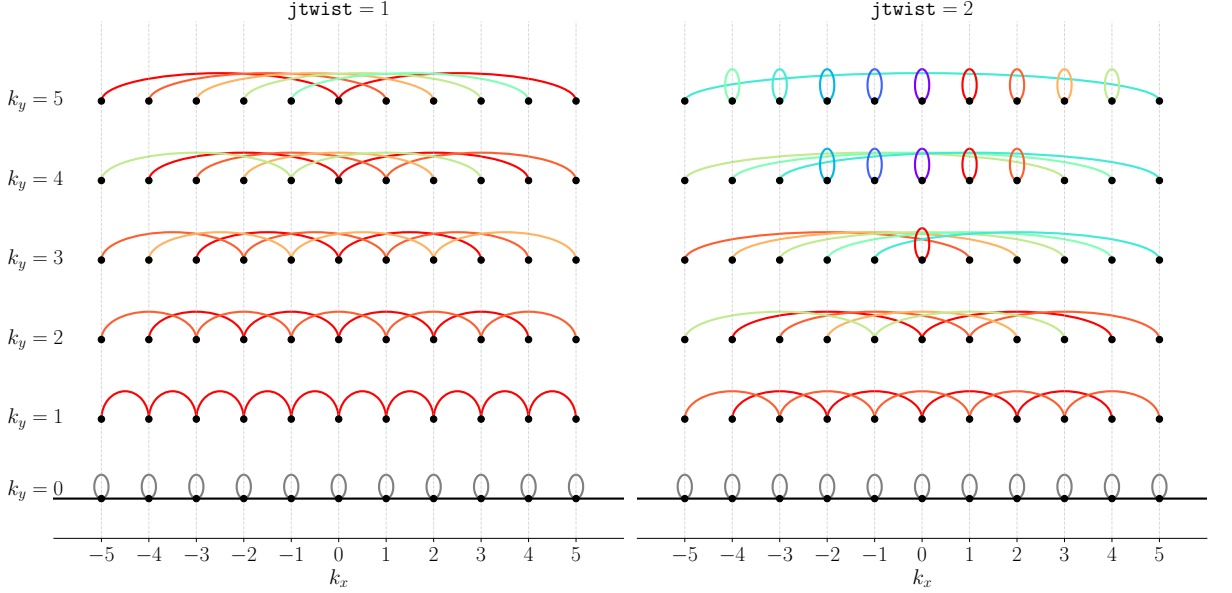


Figure 3.2: An illustration of mode connections on the extended z -domain for two different values of `jtwist`. For a given k_y , the k_x modes linked by each coloured arch form an ‘eigen-chain’, representing modes connected through the twist-and-shift condition. This connectivity arises due to the magnetic shear, which modifies the radial mode number of an eddy at fixed k_y . The number of arches indicates the number of connected 2π segments required to fully resolve the mode dynamics. Loops denote modes that are not connected to any others and are therefore periodic.

3.6 Time Advance Algorithms

Here, the algorithms used in the flux-tube version of `stella` are outlined. A comprehensive summary of all algorithmic choices available, as well as the benchmark results, can be found in Barnes et al. (2019) [9]. Here, we detail only the algorithms used to evolve the electrostatic gyrokinetic system, and which are used in the flux-annulus extension described in Chapter 4. For clarity, the normalised operators are assigned names reflecting the physical processes they represent within the equations. The evolution equation for $\hat{g}_{\mathbf{k},\nu}$, (3.2), in the electrostatic collisionless limit, is broken down as:

$$\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} + \underbrace{\hat{\mathcal{S}}_{\mathbf{k},\nu}[\hat{g}_{\mathbf{k},\nu}, \hat{\phi}]}_{\text{Parallel streaming}} + \underbrace{\hat{\mathcal{M}}_{\mathbf{k},\nu}[\hat{g}_{\mathbf{k},\nu}]}_{\text{Mirror}} + \underbrace{\hat{\mathcal{E}}_{\mathbf{k},\nu}[\hat{g}_{\mathbf{k},\nu}, \hat{\phi}]}_{\text{Explicit}} = 0, \quad (3.17)$$

with

$$\hat{\mathcal{E}}_{\mathbf{k},\nu}[\hat{g}_{\mathbf{k},\nu}, \hat{\phi}] = \underbrace{\hat{\mathcal{D}}_{\mathbf{k},\nu}[\hat{g}_{\mathbf{k},\nu}, \hat{\phi}]}_{\text{Drift}} + \underbrace{\hat{\mathcal{G}}_{\mathbf{k},\nu}[\hat{\phi}]}_{\text{Drive}} + \underbrace{\hat{\mathcal{N}}_{\mathbf{k},\nu}[\hat{g}_{\mathbf{k},\nu}, \hat{\phi}]}_{\text{Nonlinear}}. \quad (3.18)$$

The aim is to treat the fast parallel dynamics implicitly. This is an important feature when treating kinetic electrons—meaning electrons evolve under their own dynamical equation, rather than being assumed to have a Boltzmann response—because the factor $v_{\text{th},e} \sim \sqrt{m_i/m_e}$ present in the streaming and mirror terms increases their relative amplitudes when considering electron dynamics. This has the potential to impose a stringent CFL condition, which constrains the time step if treated with an explicit time advance algorithm. For this, the gyrokinetic equation is split into three pieces using operator splitting. This separates the mirror and streaming terms into individual equations, enabling an efficient implicit treatment to be used.

First, the operator splitting procedure is outlined, and then the algorithms used are detailed for each term split out: explicit advance, mirror advance, field solve, and parallel streaming.

3.6.1 Operator Splitting

As mentioned previously, the mirror and parallel streaming terms are treated implicitly in order to avoid the restrictive CFL condition that arises when treating kinetic electrons. This can be facilitated by using operator splitting to separate the gyrokinetic equation into three parts that are updated successively:

$$\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} = \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} \right)_1 + \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} \right)_2 + \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} \right)_3, \quad (3.19)$$

where equation (3.17) has been split using the Lie-Trotter splitting technique. The split equations are given by

$$\begin{aligned} \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} \right)_1 + \hat{\mathcal{D}}_{\mathbf{k},\nu} + \hat{\mathcal{G}}_{\mathbf{k},\nu} + \hat{\mathcal{N}}_{\mathbf{k},\nu} &= \hat{\mathcal{C}}_{\mathbf{k},\nu}, \\ \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} \right)_2 + \hat{\mathcal{M}}_{\mathbf{k},\nu} &= 0, \\ \left(\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} \right)_3 + \hat{\mathcal{S}}_{\mathbf{k},\nu} &= 0, \end{aligned} \quad (3.20)$$

which is accurate to first order in Δt [19], where Δt is the timestep. The temporal accuracy of this splitting can be readily increased to second order through the application of Strang splitting [113]. The electric potential, $\hat{\phi}_{\mathbf{k}}$, can be thought of as an operator acting on $\hat{g}_{\mathbf{k},\nu}$, by symbolically writing the electrostatic limit of quasineutrality (which closes the system) as

$$\hat{\mathcal{Q}}_{g,\mathbf{k}} [\hat{g}_{\mathbf{k},\nu}] = \hat{\mathcal{Q}}_{\phi,\mathbf{k}} [\hat{\phi}_{\mathbf{k}}]. \quad (3.21)$$

Noting that the electrostatic potential can be pulled out of the operator, as it has no velocity dependence, and multiplying (3.21) by $\hat{\mathcal{Q}}_{\phi,\mathbf{k}}^{-1} = 1/\hat{\mathcal{Q}}_{\phi,\mathbf{k}}$, one obtains the desired expression for $\hat{\phi}_{\mathbf{k}}$:

$$\hat{\phi}_{\mathbf{k}} = \hat{\mathcal{Q}}_{\phi,\mathbf{k}}^{-1} \cdot \hat{\mathcal{Q}}_{g,\mathbf{k}} [\hat{g}_{\mathbf{k},\nu}]. \quad (3.22)$$

Hence, equation (3.20) can be re-cast purely in terms of operators acting on $\hat{g}_{\mathbf{k},\nu}$

$$\frac{\partial \hat{g}_{\mathbf{k},\nu}}{\partial t} = \hat{\mathcal{A}}[\hat{g}_{\mathbf{k},\nu}] + (\hat{\mathcal{B}} + \hat{\mathcal{C}})[\hat{g}_{\mathbf{k},\nu}]. \quad (3.23)$$

Discretising in time gives

$$\hat{g}_{\mathbf{k},\nu}^{n_{\mathcal{E}}} = \hat{g}_{\mathbf{k},\nu}^n + \Delta t \hat{\mathcal{A}}[\hat{g}_{\mathbf{k},\nu}], \quad \hat{g}_{\mathbf{k},\nu}^{n_{\mathcal{M}}} = \hat{g}_{\mathbf{k},\nu}^{n_{\mathcal{E}}} + \Delta t \hat{\mathcal{B}}[\hat{g}_{\mathbf{k},\nu}], \quad \hat{g}_{\mathbf{k},\nu}^{n+1} = \hat{g}_{\mathbf{k},\nu}^{n_{\mathcal{M}}} + \Delta t \hat{\mathcal{C}}[\hat{g}_{\mathbf{k},\nu}]. \quad (3.24)$$

The operation $\hat{\mathcal{A}}$ captures the terms marked explicit in (3.17), $\hat{\mathcal{B}}$ the mirror terms, and $\hat{\mathcal{C}}$ the parallel streaming terms. Here the superscript n gives the time index, with $n_{\mathcal{E}}$ and $n_{\mathcal{M}}$ being intermediate time stages within the time step, and $n+1$ the updated time index. The time discretisation for $\hat{g}_{\mathbf{k},\nu}$ acted on by the operators on the right-hand sides is left unspecified here, as this will depend on the choice of algorithm used. We shall detail these in the coming sections.

The algorithm for each operation is now briefly detailed. For the remainder of this section the subscripts denoting the Fourier mode will be suppressed as all operators, except the nonlinear operator, do not couple together modes of different $\mathbf{k}$ -values. It is assumed that everywhere the equations are advanced in Fourier space, unless explicitly stated otherwise.

3.6.2 Explicit Time Advance

The magnetic drifts, the equilibrium drive term and the nonlinearity are advanced explicitly in time. These are grouped together into the combined operator

$$\hat{\mathcal{E}}_{\nu}[\hat{g}_{\nu}, \hat{\phi}] \doteq \hat{\mathcal{D}}_{\nu}[\hat{g}_{\nu}, \hat{\phi}] + \hat{\mathcal{G}}_{\nu}[\hat{\phi}] + \hat{\mathcal{N}}_{\nu}[\hat{g}_{\nu}, \hat{\phi}]. \quad (3.25)$$

The option to choose from a family of strong-stability-preserving (SSP) Runge-Kutta (RK) schemes is provided for the explicit time advance [47] due to their favorable nonlinear stability properties, as described in the original *stella* paper [9]. These schemes enable more accurate and less diffusive time-stepping while preserving stability properties under the same time step Δt , compared to lower-order methods. The default scheme, which is SSP-RK3, is described here. The updated distribution function following the explicit advance is

$$\hat{g}_{\nu}^{n_{\mathcal{E}}} = \frac{\hat{g}_{\nu}^n}{3} + \frac{\Delta t}{2} G_1^n + \frac{\Delta t}{6} (G_2^n + G_3^n), \quad (3.26)$$

with $G_1^n \doteq \hat{g}_\nu^n + \hat{\mathcal{E}}_\nu[\hat{g}_\nu^n, \hat{\phi}^n]$, $G_2^n \doteq G_1^n + \hat{\mathcal{E}}_\nu[G_1^n]$, and $G_3^n \doteq \hat{\mathcal{E}}_\nu[G_2^n]$. Here ‘ $n_\mathcal{E}$ ’ denotes an intermediate time step that results from advancing all of the explicit terms. A Fourier pseudo-spectral approach is used to evaluate the derivatives in x and y appearing in $\hat{\mathcal{N}}_\nu$ in order to retain spectral accuracy, with de-aliasing applied following the ‘2/3 rule’ [90]. The operators $\hat{\mathcal{D}}_\nu$ and $\hat{\mathcal{G}}_\nu$ are evolved in Fourier space.

3.6.3 Mirror Advance

The mirror terms can be treated explicitly or implicitly, with the latter being crucial for simulations involving kinetic electrons:

$$\hat{g}_\nu^{n\mathcal{M}} = \hat{g}_\nu^{n\mathcal{E}} - \Delta t \hat{\mathcal{M}}_\nu[\hat{g}_\nu^{n*}], \quad (3.27)$$

where n^* indicates a temporal index which depends on the algorithm chosen. In the case of explicit advance in time $\hat{g}_{\nu,m}^{n\mathcal{E}}$ is used in place of $\hat{g}_{\nu,m}^{n*}$. For the implicit case, the subscript n^* indicates a variable time location between t_n and t_{n+1} , defined as

$$t_{n*} = \frac{1 - u_t}{2} t_n + \frac{1 + u_t}{2} t_{n+1}, \quad (3.28)$$

where u_t is a user-specified parameter controlling the time-implicitness. Here $u_t = 0$ indicates a centred-in-time evaluation of the mirror operator, whereas $u_t = 1$ indicates a fully implicit evaluation. The distribution function is evaluated at t_{n*} such that

$$\hat{g}_\nu^{n*} = \frac{1 - u_t}{2} \hat{g}_\nu^{n\mathcal{M}} + \frac{1 + u_t}{2} \hat{g}_\nu^{n\mathcal{E}}. \quad (3.29)$$

The updated distribution function, for both implicit and explicit time advance, is

$$\hat{g}_{\nu,m}^{n\mathcal{M}} = \hat{g}_{\nu,m}^{n\mathcal{E}} + \Delta t v_{\text{th},\nu} \mu_\nu \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \left(\frac{\partial \hat{g}_\nu^{n*}}{\partial v_\parallel} \right)_m, \quad (3.30)$$

where the subscript m denotes that terms are evaluated at grid point $v_{\parallel,m}$. The derivative in $v_\parallel$ is carried out using finite differences with a 3-point stencil. For the explicit case this is done using a third-order upwind scheme, and for the implicit case this is done using

$$\left(\frac{\partial \hat{g}_\nu^{n*}}{\partial v_\parallel} \right)_m = \frac{(1 \mp u_{v_\parallel}) \hat{g}_{\nu,m+1}^{n*} \pm 2u_{v_\parallel} \hat{g}_{\nu,m}^{n*} - (1 \pm u_{v_\parallel}) \hat{g}_{\nu,m-1}^{n*}}{2\Delta v_\parallel}, \quad (3.31)$$

where $\Delta v_\parallel = v_{\parallel,m+1} - v_{\parallel,m}$ is the spacing in $v_\parallel$. The upper (lower) signs are used when the advection speed in $v_\parallel$ ³ is positive (negative). Here, $u_{v_\parallel}$ is a user-specified

³Note that the sign of the advection speed here is given by the sign of $-\partial B_0/\partial z$, as seen in equation (3.9b).

parameter that controls the centring in $v_{\parallel}$. If $u_{v_{\parallel}} = 0$ then the derivative in $v_{\parallel}$ is centred, which is second order accurate in $\Delta v_{\parallel}$. If $u_{v_{\parallel}} = 1$ then the derivative is fully upwinded, resulting in a scheme that is first order accurate in $\Delta v_{\parallel}$.

In the case of explicit advance in time, $\hat{g}_{\nu,m}^{n*} \rightarrow \hat{g}_{\nu,m}^{n\varepsilon}$, and $\hat{g}_{\nu,m}^{n\mathcal{M}}$ is trivially obtained from (3.30). For the implicit scheme, algebraic manipulation and matrix inversion is required. Given that the discretisation in $v_{\parallel}$ uses a 3-point stencil the resulting differentiation matrix is tridiagonal, and the corresponding linear system for implicit advance can be solved using the Thomas algorithm [117].

3.6.4 Parallel Streaming

Physically, parallel streaming describes the ballistic motion of particles along the magnetic field, and is captured by equation (2.26a). Both of the terms, involving $\hat{g}_{\nu}$ and $\hat{\phi}$, account for the parallel dynamics, and are necessary to describe Landau damping, so are treated together. The streaming operator time advance routine is structured as:

$$\hat{g}_{\nu}^{n+1} = \hat{g}_{\nu}^{n\mathcal{M}} - \Delta t \hat{\mathcal{S}}_{\nu}[\hat{g}_{\nu}^{n*}, \hat{\phi}^{n*}], \quad (3.32)$$

where the subscripts denoting the Fourier mode have been suppressed. The z -derivative is evaluated as follows:

$$\left(\frac{\partial \hat{g}}{\partial z} \right)_{i*} = \frac{\hat{g}_{i+1} - \hat{g}_i}{\Delta z}, \quad (3.33)$$

where phase space and species coordinates have been suppressed for simplification of notation. The subscripts i and $i+1$ denote that the distribution function is being evaluated at the z grid locations z_i and z_{i+1} , respectively. Here the subscript $i*$ indicates evaluation at some intermediate grid point defined by

$$z_{i*} = \frac{1 \mp u_z}{2} z_i + \frac{1 \pm u_z}{2} z_{i+1}. \quad (3.34)$$

The $\mp$ captures the sign of the advection velocity, and is determined by the sign of $v_{\parallel}$. The top (bottom) signs are used when the parallel advection speed is positive (negative), and this sign convention will be maintained throughout the remainder of this subsection, and in Section 4.3.4 when the extension of the parallel streaming algorithm to the flux-annulus version is discussed. As with the mirror advance, the user-specified parameter u_z controls spatial centring. Setting $u_z = 0$ results in a centred derivative accurate to second order in Δz , while $u_z = 1$ yields a fully upwinded

derivative accurate to first order in Δz . All other z -dependent quantities are evaluated at z_{i*} using the approximation

$$\hat{g}_{i*} = \frac{1 \mp u_z}{2} \hat{g}_i + \frac{1 \pm u_z}{2} \hat{g}_{i+1}. \quad (3.35)$$

The distribution function vector in z is denoted as $\hat{\mathbf{g}}$, with vector components $\hat{g}_i$, and similarly the quantity $\hat{\phi}$ represents a vector with z -components $\hat{\phi}_i$. The time derivative is discretised as

$$\left(\frac{\partial \hat{\mathbf{g}}}{\partial t} \right)^{n*} = \frac{\hat{\mathbf{g}}^{n+1} - \hat{\mathbf{g}}^{n_{\mathcal{M}}}}{\Delta t}, \quad (3.36)$$

where n is the temporal index, and the index ' $n_{\mathcal{M}}$ ' indicates the distribution function as evaluated after the previous operation (i.e. mirror advance). The subscript $n*$ indicates the same temporal grid point as described in Section 3.6.3. Here we detail only the implicit time advance algorithm. The other time-dependent quantities are evaluated at t_{n*} such that

$$\hat{g}^{n*} = \frac{1 - u_t}{2} \hat{g}^{n_{\mathcal{M}}} + \frac{1 + u_t}{2} \hat{g}^{n+1}. \quad (3.37)$$

To determine $\hat{g}^{n+1}$ using an implicit form of the above equation, $\hat{\phi}^{n+1}$ is needed, which is obtained by inverting the field solve equation, which itself requires $\hat{g}^{n+1}$. To overcome this apparent circular dependency, a Green's function method is used to facilitate the implicit solve [73]. Equation (3.32) can be further split into the following:

$$\hat{\mathbf{g}}_*^{n+1} - \hat{\mathbf{g}}_*^{n_{\mathcal{M}}} = \hat{\mathcal{S}}_g [\hat{\mathbf{g}}^{n*}] + \hat{\mathcal{S}}_\phi [\hat{\phi}^{n*}], \quad (3.38)$$

where $\hat{\mathcal{S}}_g$ and $\hat{\mathcal{S}}_\phi$ are the parts of $\hat{S}$ that operate on $\hat{g}$ and $\hat{\phi}$ respectively. This is further expanded to separate the contribution arising from the time levels $n+1$ and $n_{\mathcal{M}}$:

$$\hat{\mathbf{g}}_*^{n+1} - \hat{\mathbf{g}}_*^{n_{\mathcal{M}}} = \hat{\mathcal{S}}_{g+} [\hat{\mathbf{g}}^{n+1}] + \hat{\mathcal{S}}_{g-} [\hat{\mathbf{g}}^{n_{\mathcal{M}}}] + \hat{\mathcal{S}}_{\phi+} [\hat{\phi}^{n+1}] + \hat{\mathcal{S}}_{\phi-} [\hat{\phi}^{n_{\mathcal{M}}}], \quad (3.39)$$

where we have defined

$$\hat{\mathcal{S}}_{g\pm} = \frac{1 \pm u_t}{\Delta t} \hat{\mathcal{S}}_g, \quad \hat{\mathcal{S}}_{\phi\pm} = \frac{1 \pm u_t}{\Delta t} \hat{\mathcal{S}}_\phi.$$

Quasineutrality may also be symbolically represented at time level $n+1$:

$$\hat{\mathcal{Q}}_g [\hat{\mathbf{g}}^{n+1}] = \hat{\mathcal{Q}}_\phi \hat{\phi}^{n+1}. \quad (3.40)$$

First, the distribution function is split further into a ‘homogeneous’ and an ‘inhomogeneous’ contribution: $\hat{\mathbf{g}}^{n+1} = \hat{\mathbf{g}}_{\text{hom}}^{n+1} + \hat{\mathbf{g}}_{\text{inhom}}^{n+1}$. These contributions are determined by the following equations with normalised operator definitions

$$\hat{\mathbf{g}}_{*,\text{inhom}}^{n+1} - \hat{\mathbf{g}}_*^{n\mathcal{M}} = \hat{\mathcal{S}}_{g+} [\hat{\mathbf{g}}_{\text{inhom}}^{n+1}] + \hat{\mathcal{S}}_{g-} [\hat{\mathbf{g}}^{n\mathcal{M}}] + \hat{\mathcal{S}}_{\phi-} [\hat{\phi}^{n\mathcal{M}}], \quad (3.41a)$$

$$\hat{\mathbf{g}}_{*,\text{hom}}^{n+1} = \hat{\mathcal{S}}_{g+} [\hat{\mathbf{g}}_{\text{hom}}^{n+1}] + \hat{\mathcal{S}}_{\phi+} [\phi^{n+1}]. \quad (3.41b)$$

Equation (3.41a) can be solved for the z -components of $\hat{\mathbf{g}}_{\text{inhom}}^{n+1}$ as $\hat{\mathbf{g}}^{n\mathcal{M}}$ and $\hat{\phi}^n$ are known. To solve (3.41b) for $\hat{\mathbf{g}}_{\text{hom}}^{n+1}$, a Green’s function approach is used. For this a unit impulse is provided as the input for $\hat{\phi}$ at each z -location on the extended z -domain, described in Section 3.5.2, and the corresponding $\hat{\mathbf{g}}_{\text{hom}}^{n+1}$ is calculated. From this we get

$$\hat{g}_i^{n+1} = \underbrace{\sum_{p=1}^{N_{z+}} \frac{\delta \hat{g}_{i,\text{hom}}^{n+1}}{\delta \hat{\phi}_p} \hat{\phi}_p^{n+1}}_{\hat{\mathbf{g}}_{\text{hom}}^{n+1}} + \hat{g}_{i,\text{inhom}}^{n+1}, \quad (3.42)$$

where $\delta \hat{g}_{i,\text{hom}}^{n+1} / \delta \hat{\phi}_p \doteq \mathcal{R}_i$ is the response of $\hat{g}_{i,\text{hom}}^{n+1}$ at grid location z_i to a unit perturbation in $\hat{\phi}_p$ at grid location z_p , and N_{z+} is the size of the extended z -domain, which was described in Section 3.5.2. Substituting this expression for $\hat{\mathbf{g}}^{n+1}$ into quasineutrality, (3.40), gives

$$\left\{ \hat{\mathcal{Q}}_{\phi} - \hat{\mathcal{Q}}_g [\mathcal{R}] \right\} \hat{\phi}^{n+1} = \hat{\mathcal{Q}}_g [\hat{\mathbf{g}}_{\text{inhom}}^{n+1}]. \quad (3.43)$$

The pre-factor on the left hand side of this equation is a time-independent $N_{z+} \times N_{z+}$ matrix that is computed once at the beginning of a simulation, LU decomposed, and stored for computational efficiency. The pre-factor on the left hand side of this equation is a time-independent $N_{z+} \times N_{z+}$ matrix that is computed once at the beginning of a simulation, LU decomposed, and stored for computational efficiency. The computational cost for the LU decomposition of this matrix, and the subsequent solution of linear systems is approximately $\mathcal{O}(N_{z+}^2)$. As N_{z+} typically attains large values, frequently on the order of $\mathcal{O}(10^3)$, the computational cost associated with these operations becomes significant. Similarly, the memory requirements for storing the LU-decomposed factors scale as approximately $\mathcal{O}(N_{z+}^2)$. Consequently, the response matrix can become memory-limiting at large resolutions.

The updated field, $\hat{\phi}^{n+1}$, can then be obtained using back-substitution. With this, the components of the updated distribution function, $\hat{\mathbf{g}}^{n+1}$, can then be solved for using equation (3.39), and both the distribution function and electric potential have then been advanced a full time step.

Chapter 4

Full Flux-Annulus δf -Gyrokinetic Code

The stellarator world is wonderful - just a bit twisted.

William Morris

4.1 Introduction

Gyrokinetic simulations are essential tools in modern plasma physics, enabling the study of turbulence and transport in magnetically confined fusion devices. Over the past two decades, a number of gyrokinetic codes have been developed, which can broadly be categorised into flux-tube (FT) and ‘global’ codes. FT codes—such as GS2 [73], GENE (FT version) [70], CGYRO [22], GKV [122], GKW [95], `stella` [9], and GX [83]—are typically used to model local turbulence in the plasma core. Their efficiency and comparatively lower computational cost make them ideal for high-resolution parameter scans and linear stability analysis. In contrast, some codes—including GTC [81], ORB5 [78], GTS [121], CGYRO [22], GYSELA [48], GT5D [71], XGC [76], `stella` [9], EUTERPE [72], and global versions of GENE-3D [85]—are capable of simulating turbulence across larger radial and poloidal domains, possibly accounting for profile variations, boundary effects, and global mode structures. These tools have been extensively benchmarked and validated against both theoretical predictions and experimental results.

Several comprehensive benchmarking studies have been conducted in recent years comparing leading gyrokinetic codes, including `stella` [45]. A notable example is the work by Sánchez et al. (2021) [106], which employed a variety of standard linear test cases to evaluate the consistency and performance of multiple gyrokinetic codes. The

benchmark demonstrates broad agreement on linear growth rates and zonal saturation levels, while also identifying areas where numerical or physical modeling choices can lead to differences in results.

Despite their capabilities, many of these global codes have historically relied on the adiabatic, or modified-adiabatic electron approximation outlined in Section 2.3.3 in order to reduce computational expense. Although many codes can now simulate kinetic electron effects, doing so remains computationally challenging. While this simplification is sometimes justified in core plasma regimes dominated by ion-scale turbulence, it can lead to inaccurate modelling in scenarios where modes that require kinetic electron dynamics—such as trapped electron modes (TEMs) and electron-scale instabilities—play a significant role, and can lead to large quantitative discrepancies even for ITG, as seen in the CBC (see e.g., [116]). Conversely, FT codes, though more capable of handling the computational demand of kinetic electrons, may be limited in their ability to capture the effects arising from variations across flux surfaces.

In this chapter, we present a new gyrokinetic algorithm designed to address these challenges by implementing a mixed implicit-explicit time advance scheme that includes kinetic electron effects across a flux-annulus (FA). This approach enables a more efficient treatment of electron dynamics in complex magnetic geometries, without incurring too heavy a computational cost. At present, this work is restricted to the electrostatic limit.

4.2 Numerical Implementation

We have developed a novel δf -gyrokinetic algorithm specifically designed to simulate plasma turbulence in an FA for non-axisymmetric magnetic geometries. This algorithm is particularly focused on capturing the α -coupling effects described in Section 2.3.1, which are essential for accurately modelling turbulence and capturing the correct behaviour of the zonal flow when α -variation is present across the equilibrium.

We provide a detailed explanation of the FA algorithms, including an iterative-implicit treatment of the parallel streaming terms. This approach allows for the accurate modelling of kinetic electron behaviour, enabling larger time steps to be taken compared to explicit methods. As a result, we aim to achieve a reduction in the computational cost associated with these turbulent simulations in an FA for non-axisymmetric devices.

This code development is implemented in the existing gyrokinetic code `stella`, the algorithmic choices for which have been outlined in Chapter 3. Many algorithms

remain largely unchanged; however, local k -by- k operations become vector operations that couple different binormal mode numbers k_y . This coupling poses a significant challenge in the parallel streaming step, where the derivative along the field line (z -direction) couples modes of different wavenumbers, as demonstrated in the twist-and-shift derivations in Section 3.5.1. Incorporating the k_y couplings into the response matrix transforms what is originally a matrix of size $N_{z+} \times N_{z+}$, with $N_{z+} = N_z N_{k_x}$, into a matrix of size $N_{z+} \times N_{z+}$ where now $N_{z+} = N_z N_{k_x} N_{k_y}$. Given the $\mathcal{O}(N_{z+}^2)$ scaling in both computational cost and memory requirements, larger response matrices could rapidly become computationally prohibitive. Consequently, the algorithm requires substantial modifications to efficiently handle these couplings without incurring prohibitive increases in computational and memory costs associated with the enlarged response matrix dimension. The following chapter builds upon the existing choices, and outlines the algorithms implemented. Benchmark results demonstrating the efficacy of the code are given in the final section of this chapter.

4.2.1 Coordinates

The coordinate choices for **stella** are outlined in Section 3.2. For FT simulations, **stella** offers a variety of options to model the equilibrium. For FA simulations in stellarators, however, the variational moments equilibrium code, **VMEC** [60, 61, 62], is required to provide the information about the magnetic geometry and is used to calculate all the geometrical quantities appearing in the gyrokinetic equation on a specified flux surface. The flux-surface label used by **VMEC** is defined as $s \doteq \psi_t / \psi_{t,\text{LCFS}}$ with $s = 0$ indicating the magnetic axis, and $s = 1$ the outermost closed flux surface. Here $\psi_{t,\text{LCFS}}$ is the value of the toroidal flux, ψ_t , at the last-closed-flux-surface. In this case $dx/d\psi = -(\psi_t/|\psi_t|)/(\sqrt{s}L_r B_r)$ and $dy/d\alpha = L_r \sqrt{s}$, with L_r the effective minor radius computed by **VMEC**, and $B_r = 2|\psi_{t,\text{LCFS}}|/L_r^2$. From this the identity $(dx/d\psi)(dy/d\alpha) = -(\psi_t/|\psi_t|)/B_r$ is obtained, and

$$x = \frac{L_r}{2} (\sqrt{s} - \sqrt{s_0}), \quad (4.1)$$

$$y = L_r \sqrt{s} (\alpha - \alpha_0), \quad (4.2)$$

with s_0 the value of s at the centre of the simulation domain. A schematic of a flux annulus domain for a tokamak with circular flux surfaces is shown in Figure 4.1. The physical domain size in y spans $\alpha \in [0, 2\pi]$, so that $L_y = 2\pi r$, where $r = L_r \sqrt{s}$ is the radial location of interest. The smallest normalised wave number (see Table 3.1) that exists in the system is set by ρ_* , a user-specified input, as $\Delta k_y = \rho_*/\sqrt{s}$.

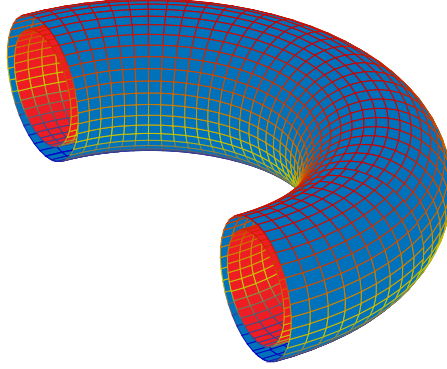


Figure 4.1: Illustration of the flux annulus domain for a tokamak with circular flux surfaces. The simulation region in the full flux annulus model lies between the inner red torus and the outer blue torus.

The z -coordinate is chosen to be the normalised arc-length, such that $\hat{\mathbf{b}} \cdot \nabla z$ is independent of z , by enforcing $z_{\min}(\alpha, \zeta = -\pi) = -\pi$ and $z_{\max}(\alpha, \zeta = \pi) = \pi$. Note that to enforce this condition, the normalisation must itself depend on α . This will be important for enforcing the parallel boundary condition, which was derived in Section 3.5.1, as $z_{\min}$ and $z_{\max}$ will thus be α -independent. For this one must relate z to the physical toroidal angle, ζ ,

$$z(\alpha, \zeta) = z(\alpha, \zeta_0) + \left(\int_{\zeta_0}^{\zeta} \frac{d\zeta'}{\hat{\mathbf{b}} \cdot \nabla \zeta'} \right) \left(\frac{1}{2\pi} \int_{\zeta_{\min}}^{\zeta_{\max}} \frac{d\zeta'}{\hat{\mathbf{b}} \cdot \nabla \zeta'} \right)^{-1}. \quad (4.3)$$

By setting $\zeta_{\min} = -\pi$ and $\zeta_{\max} = \pi$, the total arc length of each field line is guaranteed to be independent of α . Consequently the normalisation, which does depend on α , is

$$\hat{\mathbf{b}} \cdot \nabla z \doteq \left(\frac{1}{2\pi} \int_{\zeta_{\min}}^{\zeta_{\max}} \frac{d\zeta'}{\hat{\mathbf{b}} \cdot \nabla \zeta'} \right)^{-1}. \quad (4.4)$$

4.2.2 Normalisation

All quantities in this chapter are normalised consistently with Table 3.1, unless stated otherwise. For simplicity, the notation for denoting normalised quantities will not be carried, and it is remembered that all variables are normalised. The real-space normalised gyrokinetic equation is given by the electrostatic collisionless limit of equation (3.2), and is repeated here without the notation denoting the normalisations for clarity:

$$\frac{\partial g_\nu}{\partial t} + \mathcal{S}_\nu[g_\nu, \varphi_\nu] + \mathcal{M}_\nu[g_\nu] + \mathcal{D}_\nu[g_\nu, \varphi_\nu] + \mathcal{G}_\nu[\varphi_\nu] + \mathcal{N}_\nu[g_\nu, \varphi_\nu] = 0. \quad (4.5)$$

Similarly, the real-space, normalised version of quasineutrality, (3.4), in the electrostatic limit is

$$\frac{2B_0}{\pi^{1/2}} \sum_{\nu} Z_{\nu} n_{\nu} \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} d\mu_{\nu} \left(\langle g_{\nu} \rangle_{\mathbf{r}} + \frac{Z_{\nu}}{T_{\nu}} (\langle \varphi_{\nu} \rangle_{\mathbf{r}} - \phi) e^{-v^2} \right) = 0, \quad (4.6)$$

where the reader is reminded that $\varphi_{\nu} = \langle \phi \rangle_{\mathbf{R}_{\nu}}$ is the gyroaveraged electrostatic potential.

4.2.3 Fourier Transform

In keeping with the FT code implementation, the FA development of **stella** is spectral in the perpendicular plane, and treats parallel derivatives in real space. The (x, y) -Fourier treatment is important to retain spectral accuracy in the perpendicular derivatives, and to correctly capture the FLR effects without the need for a dense grid in the perpendicular domain. For this we decompose the evolution equations (4.5) and (4.6) in Fourier harmonics in (x, y) -space to obtain the equations used in the δf -gyrokinetic code **stella**. To do this the discrete Fourier transform is taken in x and y :

$$g_{\nu}(x, y, z, v_{\parallel}, \mu_{\nu}, t) = \sum_{k_x, k_y} \hat{g}_{\mathbf{k}, \nu}(z, v_{\parallel}, \mu_{\nu}, t) e^{i(k_x x + k_y y)}, \quad (4.7)$$

which is justified in the α -domain due to physical periodicity, and in the ψ -domain provided the condition $k_x |\nabla x| \gg 1/L_r$ is satisfied, where we assume L_r is comparable to the various equilibrium scale lengths present. This translates to a requirement that the turbulent fluctuations at the edges of the domain in the radial direction are decorrelated, such that they may be considered statistically identical, and periodic boundary conditions can be enforced in (x, y) [11]. As mentioned in Section 2.3.1, when this Fourier transform is taken it results in convolutions of functions in the binormal mode number, k_y , which complicates the resulting gyrokinetic equations.

4.2.4 Gyroaverage

One of the most challenging issues arising from the y -dependent geometry occurs upon gyroaveraging the electrostatic potential. This process results in a Bessel function multiplying the potential, which couples multiple modes in k_y , as described in Section 2.3.1. The gyroaveraged electrostatic potential is expressed in equation (2.34). The argument of the Bessel function, given in equation (2.35), itself contains y -dependence through the y -dependent geometric factors, which leads to the expression given by equation (2.38). This expression demonstrates that the gyroaveraging

procedure couples different Fourier modes, complicating computation. In the FT code this is a local-in- k_y operation. However due to this non-locality in k_y this operation turns into a matrix multiplication of non-diagonal matrices, making it more challenging to manage. To address this, the Fourier harmonics of the Bessel functions are calculated once at the beginning of the simulation and stored. The number of harmonics retained for the gyroaveraging procedure is determined to ensure that the total spectral energy remains within a user-specified tolerance, with the default being one part in a million.

4.3 Time Advance Algorithms: full flux-annulus

In this section, we present the algorithms employed to extend the δf -gyrokinetic code `stella` to incorporate FA effects. Many of the algorithmic techniques discussed in Section 3.6 are used here. The FA code takes advantage of the operator splitting technique described in Section 3.6.1, and we outline the algorithms used in each ‘split’ term here. The structure follows the order of the previous chapter, detailing the explicit advance, mirror advance, field solve, and parallel streaming.

4.3.1 Explicit Time Advance

The magnetic drifts, the equilibrium drive term and the nonlinearity are advanced explicitly in time. These are grouped together into the combined operator, identical to that given in equation (3.25), and the family of SSP-RK schemes is also provided for the FA explicit advance. However, the presence of convolutions in these operators poses a problem. In a similar manner to the nonlinear terms in the FT code, a Fourier pseudo-spectral approach is used to update these operators. The derivatives with respect to x and y are evaluated in (k_x, k_y) -space to retain spectral accuracy, and the multiplication with the geometric coefficients is carried out in (x, y) -space. Care must be taken not only with the nonlinearity but also for terms with y -dependent coefficients to avoid aliasing, which may be done by following the ‘2/3 rule’ [90]. In both cases the arrays are padded with zero’s in (k_x, k_y) -space before transforming to real space to carry out the multiplication of the y -dependent quantities. For the geometric coefficients, this process is performed during the initialisation by computing the coefficients on a finer y -mesh and filtering the highest third of the wavenumbers. Any additional Fourier coefficients introduced are removed when transforming back to (k_x, k_y) -space. The time advance is achieved in the same manner as in Section 3.6.2,

but now with the explicit operator taken to be in (x, y) -space,

$$\mathcal{E}_\nu[g_\nu, \varphi_\nu] \doteq \mathcal{D}_\nu[g_\nu, \varphi_\nu] + \mathcal{G}_\nu[\varphi_\nu] + \mathcal{N}_\nu[g_\nu, \varphi_\nu]. \quad (4.8)$$

For SSP-RK3 the real-space explicit time advance is done analogously to the Fourier explicit time advance:

$$g_\nu^{n\varepsilon} = \frac{g_\nu^n}{3} + \frac{\Delta t}{2} G_1^n + \frac{\Delta t}{6} (G_2^n + G_3^n), \quad (4.9)$$

with $G_1^n \doteq g_\nu^n + \mathcal{E}_\nu[g_\nu^n, \hat{\phi}^n]$, $G_2^n \doteq G_1^n + \mathcal{E}_\nu[G_1^n]$, and $G_3^n = \mathcal{E}_\nu[G_2^n]$. Note that the benefit here is that the derivatives and gyroaverages are computed in Fourier space, meaning the associated spectral accuracy is retained.

4.3.2 Mirror Advance

The mirror terms can be handled either explicitly or implicitly, with implicit treatment being essential for simulations involving kinetic electrons. For both types of time advancement, the equation is transformed from k_y -space to y -space to accommodate the y -dependent advection speed. Importantly, the mirror operator does not require gyroaverages and perpendicular derivatives, and so there are no drawbacks in updating this term in (x, y) -space as opposed to (k_x, k_y) -space. The algorithm for updating the mirror advance parallels the one detailed in Section 3.6.3, except that the Fourier representation of the distribution function, $\hat{g}_\nu$, is replaced by its real-space representation, g_ν .

4.3.3 Intermediate Field Solve

From the discussions of gyroaveraging it is clear that y -dependent geometric terms introduce added computational challenges. One key place this arises is in quasineutrality. Here, gyroaveraging is required, and the Jacobian associated with the velocity-space integral contains y -dependent coefficients, such as the magnetic field. Given that the gyrokinetic equation is coupled with quasineutrality, it is required that they are solved consistently. The aim is to solve for the Fourier modes of ϕ given those of g_ν . However, it should be noted that the distribution function and electric potential are decomposed into harmonics at the guiding centre and particle position respectively;

$$g_\nu = \sum_{\mathbf{k}} \hat{g}_{\mathbf{k},\nu} e^{i\mathbf{k} \cdot \mathbf{R}_\nu}, \quad (4.10)$$

$$\phi = \sum_{\mathbf{k}} \hat{\phi}_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (4.11)$$

When these decompositions are used in quasineutrality, the Bessel functions naturally arise from the velocity integrals, as they are evaluated at fixed particle position:

$$\sum_{\nu} Z_{\nu} n_{\nu} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{r}} \left(\frac{2B_0}{\sqrt{\pi}} \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} d\mu_{\nu} J_0(a_{\mathbf{k},\nu}) \hat{g}_{\mathbf{k},\nu} + \frac{Z_{\nu}}{T_{\nu}} (\Gamma_0(k_{\perp}\rho_{\text{th},\nu}) - 1) \hat{\phi}_{\mathbf{k}} \right) = 0, \quad (4.12)$$

where

$$\Gamma_0(k_{\perp}\rho_{\text{th},\nu}) \doteq \frac{1}{n_{\nu}} \int d^3v J_0^2(a_{\mathbf{k},\nu}) F_{0,\nu}. \quad (4.13)$$

Recall that $k_{\perp}$ and $a_{\mathbf{k},\nu}$ are both y -dependent, so it is necessary to further expand equation (4.12). Hence, anticipating a convolution of the terms with y -dependence, as in Section 4.2.4, the quantities $Q \doteq J_0 B_0$ and $\Delta \doteq \sum_{\nu} (Z_{\nu}^2 n_{\nu} / T_{\nu}) (1 - \Gamma_0(k_{\perp}\rho_{\text{th},\nu}))$ are defined. These quantities are expanded in Fourier harmonics to obtain the (k_x, k_y) -component of quasineutrality:

$$\sum_{k'_y, \nu} \frac{2Z_{\nu} n_{\nu}}{\sqrt{\pi}} \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} d\mu_{\nu} \hat{Q}_{(k_x, k_y - k'_y), k'_y, \nu} \hat{g}_{(k_x, k_y - k'_y), \nu} = \sum_{k'_y} \hat{\Delta}_{(k_x, k_y - k'_y), k'_y} \hat{\phi}_{(k_x, k_y - k'_y)}. \quad (4.14)$$

This inclusion of y -dependent geometric coefficients converts a local response expression in k_y , appearing in the FT version of **stella**, to a matrix-vector multiplication. These operators are denoted using

$$\hat{Q}_g [\hat{g}_{\nu}] = \hat{Q}_{\phi} [\hat{\phi}], \quad (4.15)$$

where $\hat{g}_{\nu}$ and $\hat{\phi}$ are vectors with Fourier components $\{\hat{g}_{\mathbf{k},\nu}\}$ and $\{\hat{\phi}_{\mathbf{k}}\}$ respectively, for each z -location. The array $\hat{Q}_{\phi}$ is band diagonal, square matrix of dimension $(N_{k_y} N_{k_x} N_z)$, and can be inverted using band LU decomposition and back substitution. Given $\hat{g}_{\nu}^{n+1, \mathcal{M}}$ one can solve for $\hat{\phi}^{n+1, \mathcal{M}}$ to find the corresponding electric potential at an intermediate time step.

4.3.4 Parallel Streaming

If the implicit approach used in Section 3.6.4 for the original FT version of the **stella** code were applied in this case, the result would require inverting a large dense matrix at each time step due to the coupling between different k_y -values. This would be computationally costly. Therefore, the equations are structured to distinguish terms with and without y -dependent prefactors. For instance, the streaming coefficient may be decomposed into a component that is constant across field lines and a residual, with the expectation that the residual is small:

$$\hat{\mathbf{b}} \cdot \nabla z = \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y + \left[\hat{\mathbf{b}} \cdot \nabla z - \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \right], \quad (4.16)$$

where

$$\left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \doteq \frac{1}{2\pi} \int_0^{2\pi} dy \, \hat{\mathbf{b}} \cdot \nabla z, \quad (4.17)$$

is the y -averaged component, which is a constant in Fourier space. The y -independent piece can be treated implicitly using a response matrix approach [73], similar to that outlined in Section 3.6.4, and the remainder is included via an iterative scheme. The field-line coupling in φ_ν is attributed to gyroaveraging, which originates from the equilibrium scale. As a result, φ_ν is largely y -independent. Hence, we can make a similar decomposition for φ_ν in (3.32): $\varphi_\nu = \phi + (\varphi_\nu - \phi)$. Here ϕ is the non-gyroaveraged electric potential, and these terms are referred to as the ‘drift-kinetic terms’ as they are equivalent to the long-wavelength limit. It is assumed that the dominant contribution of φ_ν is given by ϕ with the residual being small, which is guaranteed at long wavelengths relative to the Larmor radius, $k_\perp \rho_i \ll 1$, where the CFL constraint due to the parallel dynamics is most stringent [9]. The parallel streaming term can hence be fully decomposed into dominant and correction contributions:

$$\begin{aligned} \frac{\partial g_\nu}{\partial t} = & -v_{\text{th},\nu} v_\parallel \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \left(\frac{\partial g_\nu}{\partial z} + \frac{Z_\nu}{T_\nu} \frac{\partial \phi}{\partial z} \langle e^{-v_\nu^2} \rangle_y \right) \\ & - v_{\text{th},\nu} v_\parallel \left(\hat{\mathbf{b}} \cdot \nabla z - \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \right) \left(\frac{\partial g_\nu}{\partial z} + \frac{Z_\nu}{T_\nu} \frac{\partial \varphi_\nu}{\partial z} e^{-v_\nu^2} \right) \\ & - v_{\text{th},\nu} v_\parallel \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \frac{Z_\nu}{T_\nu} \left(\frac{\partial \varphi_\nu}{\partial z} e^{-v_\nu^2} - \frac{\partial \phi}{\partial z} \langle e^{-v_\nu^2} \rangle_y \right). \end{aligned} \quad (4.18)$$

The parallel streaming term is treated implicitly using a centred scheme in time, as implemented in the FT version of **stella**, to enable larger time steps while maintaining minimal numerical dissipation. We find that the inclusion of terms evaluated at time level n on the right-hand side of the equation causes numerical instability when both the parallel streaming and mirror terms are present for small $k_\perp$. Thus, the parallel streaming terms are always treated as if they are fully implicit in time to ensure numerical stability.

The terms with y -dependent coefficients are treated with a pseudo-spectral approach, and all terms are advanced in Fourier space. The discretisation is consistent with the original FT version of the **stella** code [9]. The same discretisation in z is used as in equations (3.33)-(3.35). However, the equation is now fully implicit in time to ensure numerical stability, meaning u_t must now be equal to 1. Using these discretisations the following symbolic form for parallel streaming is obtained

$$\hat{\mathbf{g}}_*^{n+1} - \hat{\mathbf{g}}_*^{n+1,\mathcal{M}} = \hat{\mathcal{S}}_g^{(0)} [\hat{\mathbf{g}}^{n+1}] + \hat{\mathcal{S}}_\phi^{(0)} [\hat{\boldsymbol{\phi}}^{n+1}] + \hat{\mathcal{S}}_g^{(1)} [\hat{\mathbf{g}}^{n+1}] + \hat{\mathcal{S}}_\phi^{(1)} [\hat{\boldsymbol{\phi}}^{n+1}], \quad (4.19)$$

and similarly quasineutrality

$$\hat{\mathcal{Q}}_g^{(0)} [\hat{\mathbf{g}}^{n+1}] + \hat{\mathcal{Q}}_g^{(1)} [\hat{\mathbf{g}}^{n+1}] = \hat{\mathcal{Q}}_\phi^{(0)} [\hat{\phi}^{n+1}] + \hat{\mathcal{Q}}_\phi^{(1)} [\hat{\phi}^{n+1}]. \quad (4.20)$$

Here the operators with the superscript (0) are defined to be those which are y -independent, and those with a superscript (1) are taken to be the residual. They are defined as:

$$\hat{\mathcal{S}}_g^{(0)} [\hat{\mathbf{g}}] = -\Delta t v_{\text{th},\nu} v_{\parallel} \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \frac{\partial \hat{\mathbf{g}}}{\partial z}, \quad \hat{\mathcal{S}}_g^{(1)} = \hat{\mathcal{S}}_g - \hat{\mathcal{S}}_g^{(0)}, \quad (4.21)$$

$$\hat{\mathcal{S}}_\phi^{(0)} [\hat{\phi}] = -\Delta t v_{\text{th},\nu} v_{\parallel} \left\langle \hat{\mathbf{b}} \cdot \nabla z \right\rangle_y \left(\frac{Z_\nu}{T_\nu} \frac{\partial \hat{\phi}}{\partial z} \langle e^{-v_\nu^2} \rangle_y \right), \quad \hat{\mathcal{S}}_\phi^{(1)} = \hat{\mathcal{S}}_\phi - \hat{\mathcal{S}}_\phi^{(0)}, \quad (4.22)$$

and

$$\hat{\mathcal{Q}}_g^{(0)} [\hat{\mathbf{g}}] = \sum_\nu \frac{2Z_\nu n_\nu \langle B_0 \rangle_y}{\sqrt{\pi}} \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} d\mu_\nu \hat{\mathbf{g}}, \quad \hat{\mathcal{Q}}_g^{(1)} = \hat{\mathcal{Q}}_g - \hat{\mathcal{Q}}_g^{(0)}, \quad (4.23)$$

$$\hat{\mathcal{Q}}_\phi^{(0)} [\hat{\phi}] = \sum_\nu \frac{Z_\nu^2 n_\nu}{T_\nu} \langle 1 - \Gamma_0(k_{\perp} \rho_{\text{th},\nu}) \rangle_y \hat{\phi}, \quad \hat{\mathcal{Q}}_\phi^{(1)} = \hat{\mathcal{Q}}_\phi - \hat{\mathcal{Q}}_\phi^{(0)}. \quad (4.24)$$

This form offers no immediate improvement, as all the correction terms are still treated implicitly and would require a response matrix to solve. Consequently, an iterative implicit scheme is chosen, in which the correction terms are evaluated using the previous iteration, while the drift-kinetic terms are assessed at the current iteration step. This approach enables all terms to be treated implicitly whilst ensuring the response matrix is only $N_{z+} \times N_{z+}$ in size, where N_{z+} is the size of the extended z domain, described in Section 3.5.2. The correction terms act as a source in the equation at each iteration step, and are updated with each iteration until they converge to the implicit solution. With this adjustment, improved convergence and stability in the algorithm is achieved. The iterative scheme is split as

$$\begin{aligned} \hat{\mathbf{g}}_*^{n+1,j+1} - \hat{\mathbf{g}}_*^{n+1,\mathcal{M}} &= \hat{\mathcal{S}}_g^{(0)} [\hat{\mathbf{g}}^{n+1,j+1}] + \hat{\mathcal{S}}_\phi^{(0)} [\bar{\phi}^{n+1,j+1}] + \hat{\mathcal{S}}_g^{(1)} [\hat{\mathbf{g}}^{n+1,j}] \\ &\quad + \hat{\mathcal{S}}_\phi^{(0)} [\hat{\phi}^{n+1,j} - \bar{\phi}^{n+1,j}] + \hat{\mathcal{S}}_\phi^{(1)} [\hat{\phi}^{n+1,j}], \end{aligned} \quad (4.25)$$

where j is the iteration index. Here $\bar{\phi}^{n+1,j+1}$ is an artificial field, defined through an artificial quasineutrality:

$$\hat{\mathcal{Q}}_g^{(0)} [\hat{\mathbf{g}}^{n+1,j+1}] + \hat{\mathcal{Q}}_g^{(1)} [\hat{\mathbf{g}}^{n+1,j}] = \hat{\mathcal{Q}}_\phi^{(0)} \cdot \bar{\phi}^{n+1,j+1} + \hat{\mathcal{Q}}_\phi^{(0)} \cdot (\hat{\phi}^{n+1,j} - \bar{\phi}^{n+1,j}) + \hat{\mathcal{Q}}_\phi^{(1)} \cdot \hat{\phi}^{n+1,j}. \quad (4.26)$$

The initial guesses for $\hat{\mathbf{g}}^{n+1}$ and $\hat{\phi}^{n+1}$ are taken from the mirror advance step, $\hat{\mathbf{g}}^{n+1,0} = \hat{\mathbf{g}}^{n+1,\mathcal{M}}$ and $\hat{\phi}^{n+1,0} = \hat{\phi}^{n+1,\mathcal{M}}$. At the $(j+1)^{\text{th}}$ iteration, the values $\hat{\mathbf{g}}^{n+1,j}$ and $\hat{\phi}^{n+1,j}$

are known. The distribution function, $\hat{\mathbf{g}}^{n+1,j+1}$, is further split into a homogeneous and inhomogeneous contribution: $\hat{\mathbf{g}}^{n+1,j+1} = \hat{\mathbf{g}}_{\text{hom}}^{n+1,j+1} + \hat{\mathbf{g}}_{\text{inhom}}^{n+1,j+1}$, as in Section 3.6.4. The equations for each are

$$\hat{\mathbf{g}}_{*,\text{inhom}}^{n+1,j+1} - \hat{\mathbf{g}}_*^{n+1,\mathcal{M}} = \hat{\mathcal{S}}_g^{(0)} [\hat{\mathbf{g}}_{\text{inhom}}^{n+1,j+1}] + \hat{\mathcal{S}}_\phi^{(0)} [\hat{\phi}^{n+1,j} - \bar{\phi}^{n+1,j}] + \hat{\mathcal{S}}_g^{(1)} [\hat{\mathbf{g}}^{n+1,j}] + \hat{\mathcal{S}}_\phi^{(1)} [\hat{\phi}^{n+1,j}], \quad (4.27)$$

and

$$\hat{\mathbf{g}}_{*,\text{hom}}^{n+1,j+1} = \hat{\mathcal{S}}_g^{(0)} [\hat{\mathbf{g}}_{\text{hom}}^{n+1,j+1}] + \hat{\mathcal{S}}_\phi^{(0)} [\bar{\phi}^{n+1,j+1}]. \quad (4.28)$$

A Green's function approach is used to solve equation (4.28) for $\hat{\mathbf{g}}_{\text{hom}}^{n+1,j+1}$ by providing a unit impulse at each z -location on the extended z -domain, and solving for the response of $\hat{\mathbf{g}}_{\text{hom}}^{n+1,j+1}$. From this one obtains

$$\hat{g}_i^{n+1,j+1} = \underbrace{\sum_{p=1}^{N_{z+}} \frac{\delta \hat{g}_{i,\text{hom}}^{n+1,j+1}}{\delta \bar{\phi}_p} \bar{\phi}_p^{n+1,j+1}}_{\hat{\mathbf{g}}_{\text{hom}}^{n+1,j+1}} + \hat{g}_{i,\text{inhom}}^{n+1,j+1}, \quad (4.29)$$

where $\delta \hat{g}_{i,\text{hom}}^{n+1,j+1} / \delta \bar{\phi}_p \doteq \mathcal{R}_i$ is the response of $\hat{g}_{i,\text{hom}}^{n+1,j+1}$ at grid location z_i to a unit perturbation in $\bar{\phi}_p$ at grid location z_p . Substituting this expression for $\hat{\mathbf{g}}^{n+1,j+1}$ into quasineutrality gives

$$\left\{ \hat{\mathcal{Q}}_\phi^{(0)} - \hat{\mathcal{Q}}_g^{(0)} [\mathcal{R}] \right\} \bar{\phi}^{n+1,j+1} = \hat{\mathcal{Q}}_g^{(0)} [\hat{\mathbf{g}}_{\text{inhom}}^{n+1,j+1}] + \hat{\mathcal{S}}_\phi^{(0)} [\hat{\phi}^{n+1,j} - \bar{\phi}^{n+1,j}] + \hat{\mathcal{Q}}_g^{(1)} [\hat{\mathbf{g}}^{n+1,j}] - \hat{\mathcal{Q}}_\phi^{(1)} \cdot \hat{\phi}^{n+1,j}. \quad (4.30)$$

The iterative implicit scheme first solves equation (4.27) for $\hat{\mathbf{g}}_{\text{inhom}}^{n+1,j+1}$ given $\hat{\mathbf{g}}^{n+1,j}$ and $\hat{\phi}^{n+1,j}$. Then equation (4.30) is solved using LU decomposition and back substitution to obtain $\bar{\phi}^{n+1,j+1}$. The iterative implicit scheme employed allows the response matrix to be of dimension $N_{z+} \times N_{z+}$, where $N_{z+} = N_{k_x} N_z$. This approach circumvents the inclusion of the N_{k_y} dimension within the primary response matrix, thereby reducing its overall size. This is beneficial for both the computational, and memory cost of treating parallel streaming. Following back substitution, equation (4.25) can be used to solve for $\hat{\mathbf{g}}^{n+1,j+1}$ by performing a sweep in z . Full quasineutrality can then be used to solve for $\hat{\phi}^{n+1,j+1}$ using (4.15), and the iteration process is repeated with the newly found $\hat{\mathbf{g}}^{n+1,j+1}$ and $\hat{\phi}^{n+1,j+1}$ as the inputs for $\hat{\mathbf{g}}^{n+1,j}$ and $\hat{\phi}^{n+1,j}$ respectively. The iteration process is complete when $|\hat{\phi}^{n+1,j+1} - \hat{\phi}^{n+1,j}| < \epsilon_{rr}$, with ϵ_{rr} some specified error. In this case $\bar{\phi}^{n+1,j+1} \rightarrow \bar{\phi}^{n+1}$, $\hat{\phi}^{n+1,j+1} \rightarrow \hat{\phi}^{n+1}$, and $\hat{\mathbf{g}}^{n+1,j+1} \rightarrow \hat{\mathbf{g}}^{n+1}$ and the equations (4.26), (4.25) go to (4.20) and (4.19) respectively. A flow chart of this implicit scheme is shown in Figure 4.2.

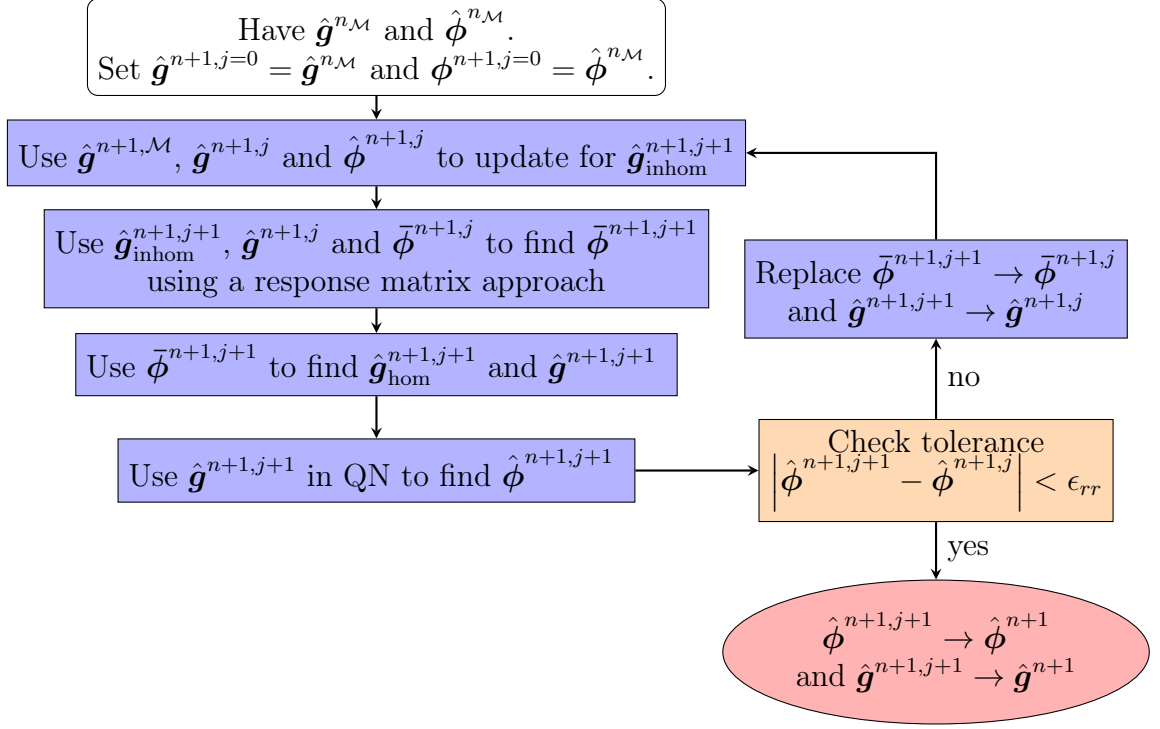


Figure 4.2: Flow chart showing the iteration scheme employed for implicitly advancing the parallel streaming term in the FA algorithm.

Figure 4.3 shows the normalised modulus of the difference between $\hat{\phi}$ values at successive iterations, relative to the previous time level. The plot demonstrates that changes in $\hat{\phi}$ decrease with each iteration, with minimal absolute difference between the sixth and seventh iterations—indicating that $\hat{\phi}$ is converging to a given value. This example, using a W7-X configuration with $N_y = 32$ y -modes, illustrates typical behaviour within a given time level using the iteration scheme. As N_y increases, more iterations are needed to achieve the same level of normalised convergence due to the growing number of couplings in the correction terms. Section 4.4 gives more information for the computation cost incurred using this iterative scheme.

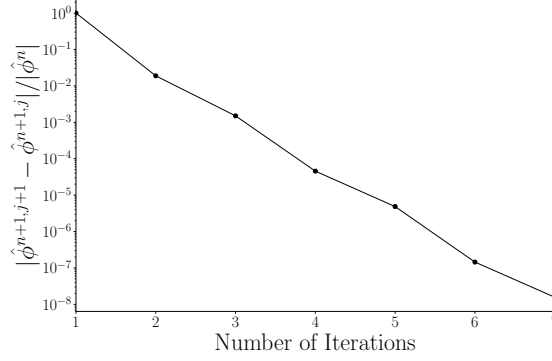


Figure 4.3: A plot showing the normalised convergence of $|\hat{\phi}^{n+1,j+1} - \hat{\phi}^{n+1,j}|$ with the number of iterations for the parallel streaming iterative-implicit scheme. As the number of iterations completed increases the difference between successive values of $\hat{\phi}$ decreases.

4.3.5 Boltzmann Electrons

In this section, the special case of Boltzmann electrons is considered, which was discussed in Section 2.3.3. By treating electrons as Boltzmann, their response is integrated into a modified quasineutrality condition rather than being treated as an independent kinetic species. Compared with the kinetic electron case, the computational cost associated with treating electrons using a Boltzmann response is significantly reduced, as it eliminates the requirement to capture the fast parallel dynamics of electrons, relaxing the CFL constraint on the time step. It is hence important to implement as an option in the code. In this case the electrons adopt a modified Boltzmann response of the form $\delta n_e = en_e(\phi - \Phi)/T_e$, where Φ indicates a flux surface average of ϕ defined by equation (2.49), which is repeated here for convenience:

$$\Phi \doteq \frac{\int dy dz \mathcal{J}(y, z) \phi}{\int dy dz \mathcal{J}(y, z)}.$$

Using this, and re-defining $\Delta \rightarrow \Delta + n_e/T_e$ in equation (4.14) to capture the Boltzmann response of the electrons, quasineutrality becomes

$$\begin{aligned} \sum_{k'_y} \sum_{\nu \neq e} Z_\nu n_\nu \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} d\mu_\nu \hat{Q}_{(k_x, k'_y), k_y - k'_y, \nu} \hat{g}_{(k_x, k'_y), \nu} \\ - \sum_{k'_y} \hat{\Delta}_{(k_x, k'_y), k_y - k'_y} \hat{\phi}_{(k_x, k'_y)} = -\delta_{k_y, 0} \hat{\Phi}_{k_x} \frac{n_e}{T_e}, \end{aligned} \quad (4.31)$$

with $\delta_{k_y,0}$ the Kronecker delta-function. To obtain an expression for $\hat{\Phi}_{k_x}$, the function $\mathcal{J}(y, z)$ is expanded in Fourier harmonics in y and Fourier-transformed in x :

$$\hat{\Phi}_{k_x} = \frac{\int dz \sum_{k'_y} \hat{\mathcal{J}}_{-k'_y}(z) \hat{\phi}_{(k_x, k'_y)}}{\int dy dz \mathcal{J}(y, z)}. \quad (4.32)$$

The linearity of the quasineutrality constraint is used to decompose

$$\hat{\phi}_{(k_x, k'_y)} = \hat{\phi}_{(k_x, k'_y)}^{\text{inhom}} + \hat{\phi}_{(k_x, k'_y)}^{\text{hom}}, \quad (4.33)$$

with $\hat{\phi}_{(k_x, k'_y)}^{\text{inhom}}$ and $\hat{\phi}_{(k_x, k'_y)}^{\text{hom}}$ referring to the Fourier components of the ϕ that satisfy the inhomogeneous and homogeneous components of quasineutrality, respectively:

$$\begin{aligned} \sum_{k'_y} \sum_{\nu \neq e} Z_\nu n_\nu \frac{2}{\pi^{1/2}} \int_{-\infty}^{\infty} dv_{\parallel} \int_0^{\infty} d\mu_{\nu} \hat{Q}_{(k_x, k'_y), k_y - k'_y, \nu} \hat{g}_{(k_x, k'_y), \nu} \\ - \sum_{k'_y} \hat{\Delta}_{(k_x, k'_y), k_y - k'_y} \hat{\phi}_{(k_x, k'_y)}^{\text{inhom}} = 0, \end{aligned} \quad (4.34)$$

and

$$\sum_{k'_y} \hat{\Delta}_{(k_x, k'_y), k_y - k'_y} \hat{\phi}_{(k_x, k'_y)}^{\text{hom}} = \delta_{k_y, 0} \hat{\Phi}_{k_x} \frac{n_e}{T_e}. \quad (4.35)$$

Note that adding together (4.34) and (4.35) recovers the complete quasineutrality equation for $\hat{\phi}_{(k_x, k'_y)}$. Equation (4.34) can be solved for $\hat{\phi}_{(k_x, k'_y)}^{\text{inhom}}$ provided $\hat{g}_{(k_x, k'_y), \nu}$. To obtain $\hat{\phi}_{(k_x, k'_y)}^{\text{hom}}$, a Green's function approach is employed in which

$$\hat{\phi}_{(k_x, k'_y)}^{\text{hom}} = \frac{\delta \hat{\phi}_{(k_x, k'_y)}^{\text{hom}}}{\delta \hat{\Phi}_{k_x}} \hat{\Phi}_{k_x}. \quad (4.36)$$

The response vector $\{\delta \hat{\phi}_{(k_x, k'_y)}^{\text{hom}} / \delta \hat{\Phi}_{k_x}\}$ is calculated once during initialisation by solving (4.35) for $\hat{\phi}_{(k_x, k'_y)}^{\text{hom}}$ for each k_x , with $\hat{\Phi}_{k_x} = \delta_{k_x} = 1$. Using equation (4.36) and the decompositions (4.33) in (4.32) gives an equation for $\hat{\Phi}_{k_x}$:

$$\hat{\Phi}_{k_x} \left(1 - \frac{\int dz \sum_{k'_y} \hat{\mathcal{J}}_{-k'_y}(z) \left(\delta \hat{\phi}_{(k_x, k'_y)}^{\text{hom}} / \delta \hat{\Phi}_{k_x} \right)}{\int dy dz \mathcal{J}(y, z)} \right) = \frac{\int dz \sum_{k'_y} \hat{\mathcal{J}}_{-k'_y}(z) \hat{\phi}_{(k_x, k'_y)}^{\text{inhom}}}{\int dy dz \mathcal{J}(y, z)}. \quad (4.37)$$

Once $\hat{\Phi}_{k_x}$ is obtained, it can be used in equation (4.31) to calculate $\hat{\phi}_{\mathbf{k}}$.

4.4 Numerical Results

In this section, we present analysis of computational results obtained with the FA code, alongside comparisons to the existing FT code version. Code agreement in

the limiting scenario of a tokamak as the parameter ρ_* approaches zero, where both codes are expected to agree, is demonstrated. In addition, we investigate a stellarator configuration with Boltzmann electrons, conducting scans across a range of ρ_* values, and further verifying against published results. We then present some novel results including kinetic electrons for a non-axisymmetric case. The definitions and normalisations of the fluxes used in this section can be found in Appendix A.1. For this section it is also useful to introduce the following simulation parameters which give the size of the domain in (x, y)

$$y_0 = 2\pi L_y \rho_r = \frac{\sqrt{s}}{\rho_*} = \frac{r}{L_r} \frac{1}{\rho_*} = \frac{1}{\Delta k_y \rho_r},$$

$$x_0 = 2\pi L_x \rho_r = \frac{1}{\Delta k_x \rho_r}.$$

It should be noted that the quantities N_x and N_y used in this section are the number of grid points in k_x and k_y respectively, and includes the zero-padding used for the ‘2/3 rule’.

The input files for the simulations in this section can be found at https://github.com/GeorgiaActon/Thesis_input_files.git.

4.4.1 Expectation

When considering a linear flux-tube simulation for a given field line, each binormal wavenumber k_y can be assigned an independent growth rate and frequency. For axisymmetric magnetic fields, the growth rate and frequency spectra are the same for every field line within a flux surface; for non-axisymmetric magnetic fields, the spectra will differ between field lines.

For an FA simulation, all binormal wavenumbers are coupled, and only a single growth rate and frequency can meaningfully be extracted from the simulation. While it is theoretically possible to derive the FT spectrum using the FA code for an axisymmetric system, doing so may prove challenging in practice, as the fastest growing mode tends to dominate the overall behavior, in an initial value code.

In Figure 4.4 the purple line represents the anticipated result from a single linear FT simulation, where the binormal wavenumbers are decoupled, allowing k_y to serve as a well-defined normal mode of the system, producing a spectrum of the linear growth rate against k_y . To produce this purple curve one single simulation would be conducted and the growth rate for each k_y value is plotted. However, reproducing the same ‘spectrum’ for the FA case is less straightforward due to mode-coupling in

the binormal direction. To effectively represent the spectrum in an FA simulation for a given ρ_* , we fix ρ_* —which in turn fixes the minimum k_y —and perform a series of linear simulations. In each simulation, the number of modes, N_y , is increased to extend the maximum value of k_y . Each simulation gives a single, poloidally-global growth rate, which is then plotted against the corresponding maximum k_y . This scanning process produces a result comparable to the black curve in Figure 4.4. Here, as the maximum k_y is increased, it is expected that a maximum poloidally-global growth rate sustained by all field lines is found and communicated to all k_y . Given that all field lines are identical in their geometry they can support the same growth rate spectrum, meaning the maximum poloidally-global growth rate will correspond to the local peak.

Up to this maximum growth rate, the results from the FA simulation and the FT simulation should correspond. Beyond this peak, the FT spectrum eventually shows a decline in growth rates, indicating that these modes are less unstable. In contrast, when mode-coupling is present, the FA results should plateau. This flattening occurs because, even with an increased maximum k_y , the most unstable mode remains present, ultimately determining the growth rate of the simulation.

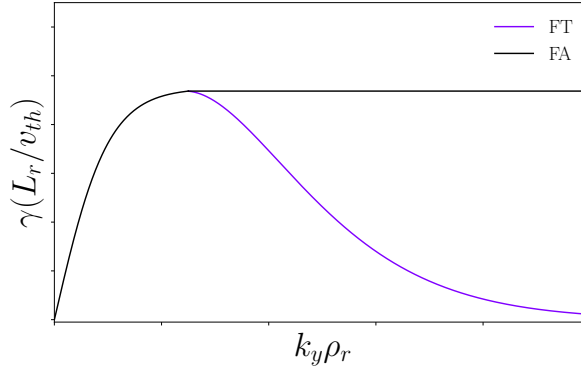


Figure 4.4: Sketch of the expectation for the growth rate spectrum versus the binormal wavenumber, k_y , for an axisymmetric device. Here, the purple line represents an example spectrum one could obtain from an FT simulation, and the black line represents the expected analogous FA-simulation where the poloidally-global growth rate is plotted against the maximum value of k_y in the linear simulation.

In stellarators, each field line, labelled by the index α , experiences distinct geometry, resulting in varied growth rate spectra, including different maximum growth rates and potentially peaking at different values of k_y . This is illustrated in Figure 4.5 by the coloured lines, with each colour corresponding to the expected result of a FT

simulation for a different value of α . In an FA simulation, the full poloidal geometry of the stellarator is taken into account for a given flux surface, enabling the identification of the most unstable growth rate that can be supported across the device. This scan can similarly be performed by fixing the minimum k_y of the simulation and scanning in N_y to increase the maximum k_y included. Plotting the global growth rate against the maximum k_y produces a plot akin to the black curve in Figure 4.5. As the maximum k_y is increased, the simulation effectively captures a weighted average of the maximum growth rates associated with each field line, reflecting the global growth rate that can be sustained across the variations in geometry. In this example, while the FT spectrum exhibits a decline in growth rates as k_y increases, the FA simulation is, again, expected to plateau, as the mode that is most unstable is still present in the simulation, and the growth rate of this mode persists as the largest growth rate in the simulation. Importantly, this plateau does not necessarily occur at the maximal growth rate across all field lines; rather, it plateaus at an averaged value that is supported by the collective behavior of all field lines.

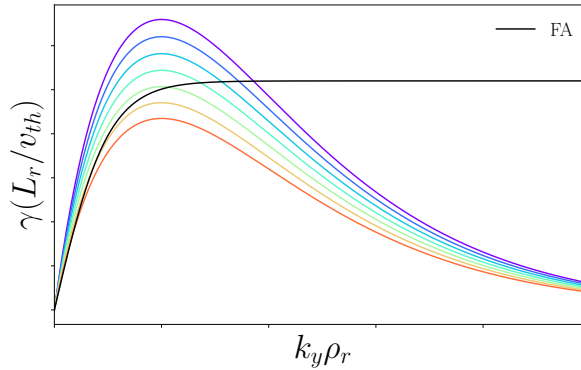


Figure 4.5: Sketch of the expectation for the growth rate spectrum versus the binormal wavenumber, k_y for a non-axisymmetric device. Here, the coloured lines represent example spectra one could obtain from an FT simulation for different field lines within a given flux surface of a device. Each field line samples different geometry, and hence will support different instabilities and exhibit different spectra. The black line represents the expected analogous FA-simulation spectrum.

4.4.2 Axisymmetric Benchmark Cases

We begin by considering an axisymmetric geometry. This is done as a benchmark, as it is expected that the FT result will capture the dynamics well, given that each field line has the same magnetic geometry. For these benchmarks the axisymmetric Cyclone Base Case (CBC) [32] geometry is used. Two different geometric implementations are

considered for the benchmark: Miller and VMEC. In the Miller case, the z -coordinate corresponds to the poloidal angle, θ . This choice ensures that, in an axisymmetric configuration, geometric quantities remain constant across all values of α at a fixed z , eliminating any mode-coupling introduced by the FA code. Conversely, the VMEC implementation uses the toroidal angle, ζ , as the field-line following coordinate z . Here, in (α, ζ) coordinates, geometric quantities vary with α at fixed z , even in the axisymmetric cases. As a result, the FA code captures non-zero mode-coupling, when treating the axisymmetric device using the full FA algorithms. Figure 4.6 illustrates the impact of these coordinate choices. The black line shows a field line at $\alpha = 0$, while the background contours represent the magnetic field strength—with red and blue indicating a stronger and weaker magnetic field respectively. In the Miller case (left), varying α does not alter the geometry as a function of z , since $z = \theta$ holds geometry constant. However, in the VMEC case (right), changes in α shift the geometry along the $z = \zeta$ direction. This is because the geometry will be constant along $\theta = \alpha + \iota\zeta$. This is evident when observing the white line in Figure 4.6: in the Miller case, the magnetic field strength along this line remains unchanged for all α , while in the VMEC case, it varies with α at fixed ζ . Thus, under VMEC geometry, the FA formulation interprets an axisymmetric device as if it were a stellarator, capturing mode-coupling.

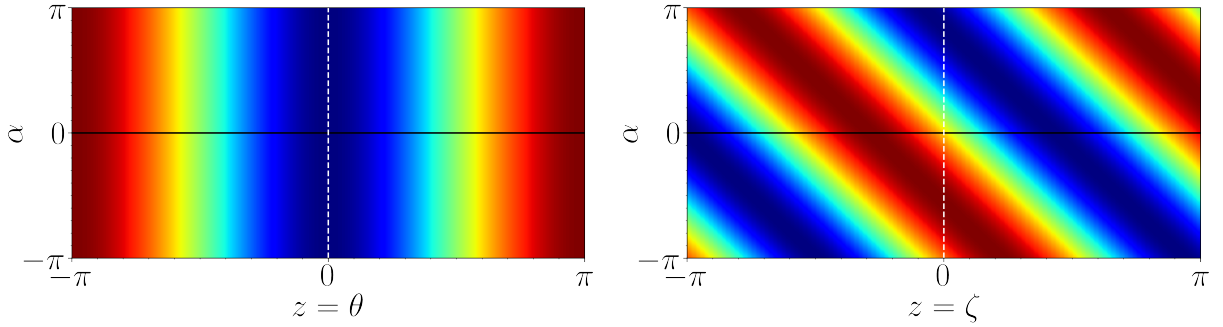


Figure 4.6: The magnetic geometry for the CBC as considered by the Miller (left) and VMEC (right) implementations. Here the black line indicates a field line at $\alpha = 0$, and the white dashed line indicates fixed $z = 0$. It can be seen that for the Miller case the white line samples uniform magnetic geometry, whilst for the VMEC case the magnetic geometry varies with α .

4.4.2.1 Miller Geometry: kinetic electron results

The initial validation is carried out using a Miller geometry based on CBC, with Boltzmann and kinetic electrons. This serves as a benchmark for evaluating the performance of the algorithms. Due to the coordinate choices, the Miller geometry

produces no mode-coupling from full surface effects. As a result, we expect good agreement between FA and FT simulations. Results from the FA simulations are compared against an equivalent FT simulation using `stella`. The resolutions and parameters for these simulations using Miller geometry are given in Table 4.1.

<i>Parameter</i>	<i>Linear Simulation</i>	<i>Nonlinear Simulation</i>
r/L_r	0.5	0.5
L_r/L_{T_i}	2.498	2.498
L_r/L_{T_e}	2.498	2.498
L_r/L_{n_i}	0.8	0.8
L_r/L_{n_e}	0.8	0.8
N_θ	3	1
N_z	32	32
N_μ	12	12
$N_{v_\parallel}$	72	72
N_x	1	192
N_y	64	64
y_0	25	10
ρ_*	0.02	0.01

Table 4.1: Simulation parameters for the case with CBC using Miller geometry.

Here r/L_r gives the normalised radial location of the simulation, L_r/L_{T_i} , L_r/L_{T_e} , L_r/L_{n_i} , L_r/L_{n_e} the species temperature and density gradients, N_θ is the number of poloidal turns, N_z , N_μ , $N_{v_\parallel}$, N_x , and N_y are the number of grid points in z , μ , $v_\parallel$, k_x and k_y respectively.

First, the spectra of the linear growth rate and real frequency are shown in Figure 4.7, as a function of $k_y\rho_r$, for Boltzmann and kinetic electrons. Due to the absence of mode-coupling in the FA simulation, true spectra for the growth rate and frequency as a function of $k_y\rho_r$ can be produced, as in this limit the different k_y -modes evolve independently.

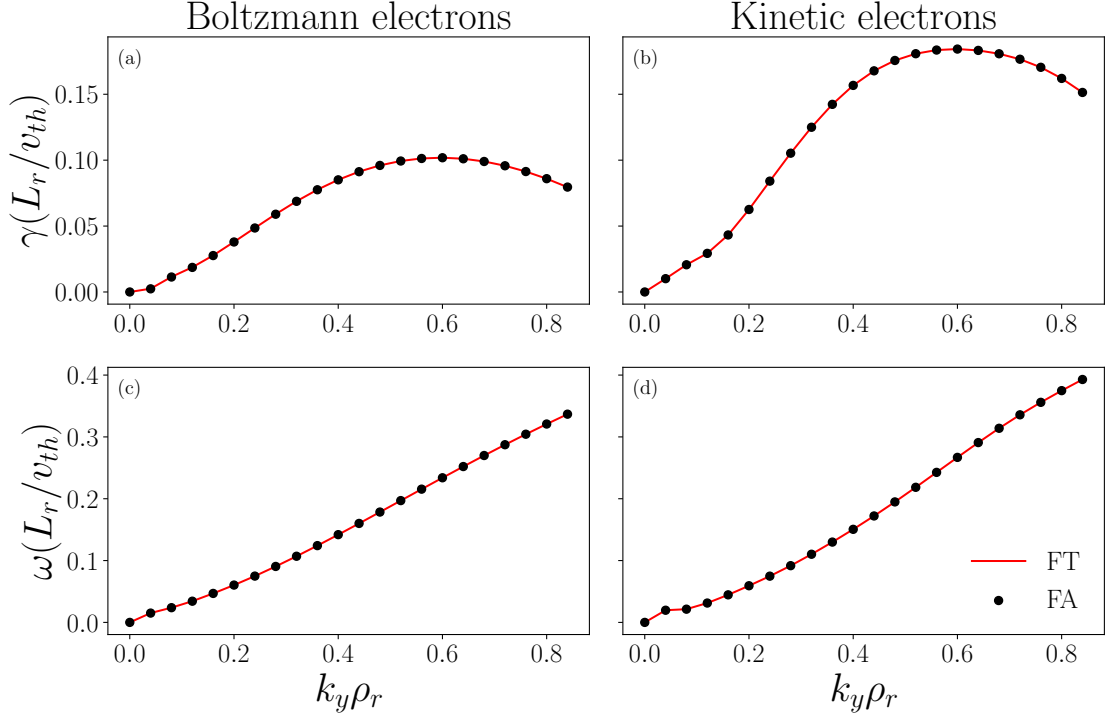


Figure 4.7: A comparison of the linear growth rate (top) and real frequency (bottom) spectra against binormal mode number for CBC using Miller geometry with Boltzmann (left) and kinetic (right) electrons. The FT (red) and FA (black) simulation results are overlaid for comparison.

This plot shows good agreement between the two codes, indicating that in the limit of no mode-coupling the algorithms used in the FT and FA simulations produce the same results.

Next, nonlinear simulations are performed using the same Miller geometry with kinetic electrons. Both the FT and FA codes are run implicitly with twist-and-shift boundary conditions. The resolutions for these simulations are given in Table 4.1, and the resulting time traces are shown in Figure 4.8. In this regime, the saturated turbulent heat flux values from the FT and FA simulations agree to within 5%, indicating strong consistency between the two approaches. The spectra of the squared electrostatic potential and heat fluxes also show close agreement, as illustrated in Figure 4.9. This consistency is expected, as the Miller geometry is axisymmetric and full surface mode-coupling is zero. Some small discrepancies occur in the nonlinear spectra at small values of $k_y \rho_r$, which may be due to the presence of a finite ρ_* , which enter in the twist-and-shift boundary conditions, as discussed in Section 3.5.1.

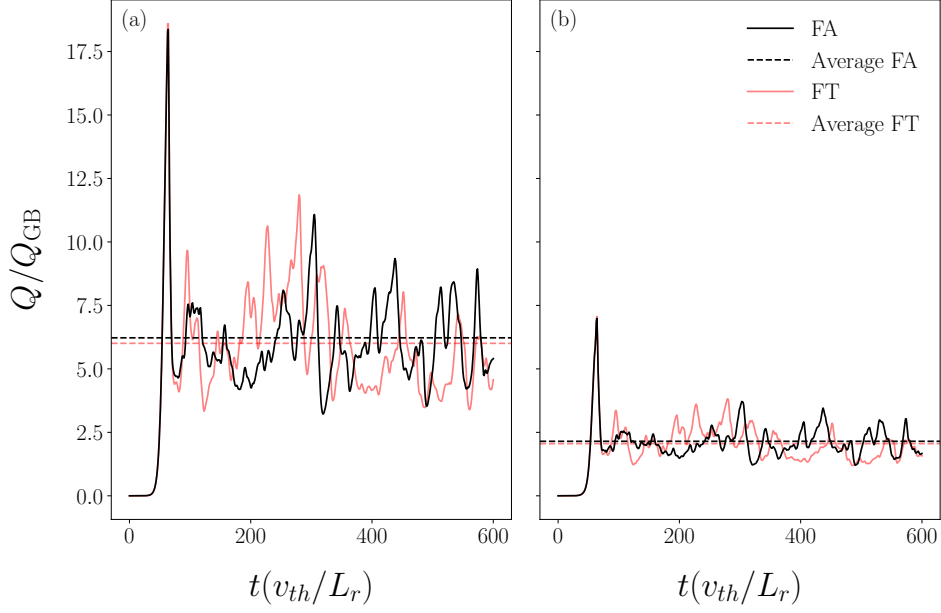


Figure 4.8: Time evolution of the nonlinear heat flux for FT (red) and FA (black) simulations for CBC using Miller geometry, including kinetic electrons. Solid lines represent the instantaneous heat flux, while dashed lines indicate the average heat flux over the final 60% of the simulation time. Panels (a) and (b) show the ion and electron heat fluxes, respectively.

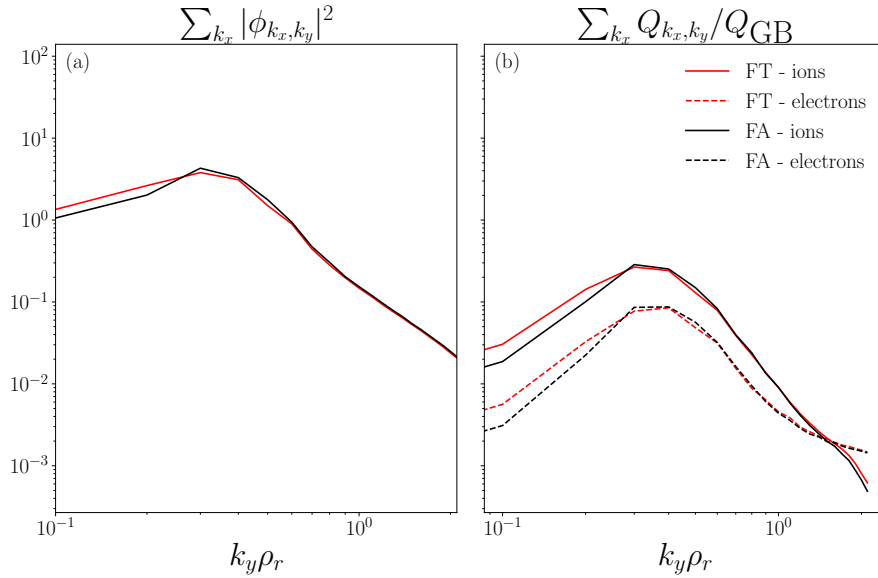


Figure 4.9: Spectra of the squared electrostatic potential (a) and heat flux (b) for CBC using Miller geometry, with kinetic electrons, corresponding to the nonlinear simulations shown in Figure 4.8.

4.4.2.2 VMEC Geometry: ρ_* convergence

In order to determine the dependence on ρ_* , when full surface effects are included, the convergence of the linear growth rate as a function of ρ_*^{-1} is considered for the axisymmetric CBC configuration generated by VMEC. These are compared to the results obtained from an FT simulation, and it is anticipated that the results should converge as $\rho_* \rightarrow 0$. Unlike the previous Miller case, which has no mode-coupling in the FA simulation, the VMEC geometry gives geometric coefficients which vary with α at a fixed z . In this case, a finite number of binormal modes N_y and a finite ρ_* result in mode-coupling that is absent in the Miller limit. Nevertheless, as $\rho_* \rightarrow 0$, this mode-coupling reduces, and we expect the FA simulation to converge to the FT result. For this, the maximum $k_y \rho_r$ value in the FA simulations is kept fixed at $k_y \rho_r = 0.4$, whilst ρ_* is decreased. The result obtained from each FA simulation is compared with the result obtained by the FT simulation.

This convergence is demonstrated in Figure 4.10, the left-hand side panels of which show results for linear simulations with Boltzmann electrons, using CBC magnetic geometry. As $\rho_* \rightarrow 0$, the linear growth rate and real frequency from the FA simulations approaches those produced in the FT limit.

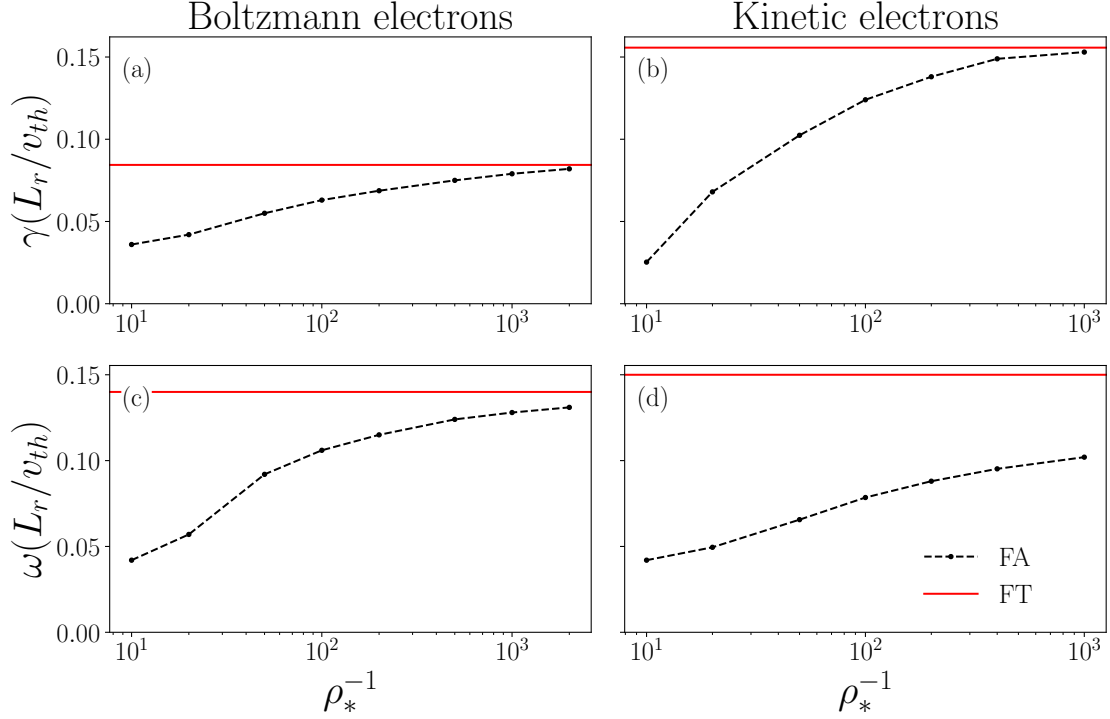


Figure 4.10: Scans showing the linear growth rate (top) and real frequency (bottom) as a function of inverse ρ_* , for CBC using VMEC geometry with a Boltzmann (left) and kinetic (right) electron response. The black points are from the FA scan, and the red line is the FT limit.

The same scan in ρ_*^{-1} for CBC geometry with kinetic electrons is shown in the right-hand side panels of Figure 4.10. In this case, the growth rate converges to the FT limit; however, the real frequency does not reach the corresponding FT value. This discrepancy suggests that achieving full convergence may require a smaller value of ρ_* , and potentially signals a stronger dependence on ρ_* and full-surface effects when kinetic electrons are included.

4.4.3 Non-Axisymmetric Benchmark Case

The following benchmarks have been carried out using a VMEC equilibrium corresponding to the high-mirror configuration of the W7-X magnetic geometry [43, 45].

4.4.3.1 Linear Results with Boltzmann Electrons

The first benchmark compares the linear growth rates and frequencies obtained from the FA simulation using the `stella` code and the FT simulations performed using the GX code, both of which are conducted using explicit algorithms. The simulation parameters used in the `stella` simulations are given in Table 4.2. For the FT results,

<i>Parameter</i>	<i>Value</i>
r/L_r	0.5
L_r/L_T	4.0
L_r/L_n	1.0
N_z/N_θ	128
N_μ	12
$N_{v_\parallel}$	72
y_0	10 and 50
ρ_*	0.05 and 0.01
N_x	1

Table 4.2: List of simulation parameters for the linear Boltzmann electron simulations using W7-X geometry.

linear growth rate spectra are computed across a set of field lines spanning one poloidal turn. These are shown as different coloured lines in Figure 4.11, with the colour corresponding to the field-line label α . In contrast, the FA results are obtained by scanning in $k_y \rho_r$: the number of binormal modes N_y is increased while keeping ρ_* fixed, effectively extending the maximum resolved k_y in the simulation. This method is discussed in more detail in Section 4.4.1.

It can be seen from Figure 4.11 that the FA results behave as expected: the growth rate represents a weighted average of the FT results and plateaus beyond the peak growth rate of the most unstable field line ($\alpha = 0$) from the FT simulations. This plateau occurs because the maximal growth rate present in the system remains constant beyond that point. Two FA cases are shown: one with $\rho_* = 0.05$ (plotted in black) and another with $\rho_* = 0.01$ (plotted in grey). The case with the smaller ρ_* value (grey) more closely follows the growth rate of the most unstable field line, $\alpha = 0$, but still falls short of reaching the peak growth rate observed in the FT simulations.

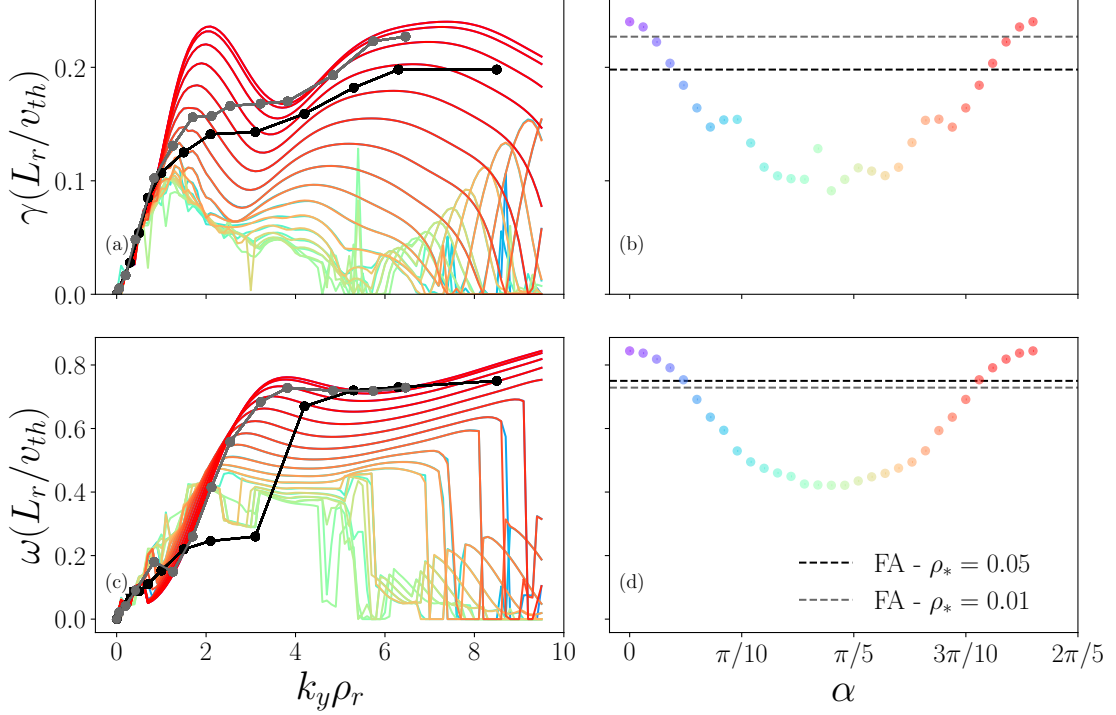


Figure 4.11: Comparison of linear growth rate (top) and real frequency (bottom) spectra as functions of binormal mode number, using the high-mirror W7-X configuration with Boltzmann electrons. Panels (a) and (c) show the FT spectra for field lines within a given flux surface at various values of α (represented by different coloured lines). The FA results for $\rho_* = 0.05$ (black) and $\rho_* = 0.01$ (grey) are overlaid. Panels (b) and (d) display the maximum growth rate and frequency, respectively, as functions of α . Dashed black and grey lines indicate the maximum global growth rates from the corresponding FA simulations. This demonstrates that the flux-annulus code produces results that differ from those of the flux-tube code, consistent with the expectations outlined in Section 4.4.1. For the smaller value of ρ_* , the flux-annulus result more closely resembles that of the most unstable flux-tube, as anticipated. In the limit $\rho_* \rightarrow 0$, the two approaches are expected to converge.

Next, the FA linear growth rate and frequency spectra are compared against results from other codes using the benchmark data presented in Sánchez et al. (2021) [106]. These comparisons are performed using the same W7-X equilibrium, and are shown in Figure 4.12. The marker shapes and colours are kept consistent with those used in the original paper. The overlaid black and grey points are from the **stella** FA simulations with $\rho_* = 0.05$ and $\rho_* = 0.01$ respectively. Good agreement is observed between the FA **stella** code and the other global-in- α codes, which are highlighted with filled markers in the figure, as none of these experience a peak at $k_y \rho_r \sim 2$.

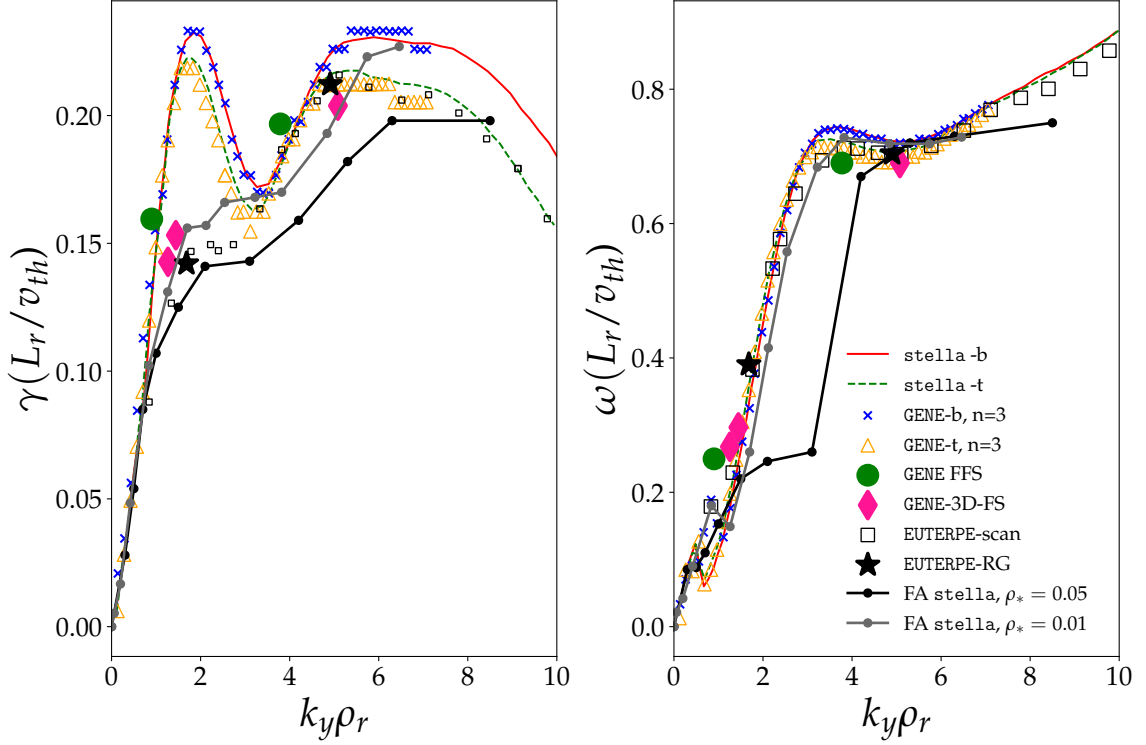


Figure 4.12: Comparison of the real and imaginary components of the growth rate for the FA code with the benchmark W7-X case performed by Sánchez et al. [106]. For the FA case the following simulation parameters are used: $L_r/L_T = 4.0$, $L_n = 1.0$, $\rho_* = 0.05$ (black) and $\rho_* = 0.01$ (grey), $N_z = 128$, $N_\mu = 12$, $N_{v_\parallel} = 36$, $r/L_r = 0.5$. For the other codes, the simulation parameters are provided in Sánchez et al. [106]. This demonstrates good agreement between the flux-annulus version of **stella** and the other global codes, as all exhibit similar behavior in comparison to the flux-tube codes. In particular, while the flux-tube codes show a clear peak in the growth rate at $k_y \rho_r$, this peak is absent in the results from all global codes.

4.4.3.2 Nonlinear Results with Boltzmann Electrons

To compare the FA code with the FT results in the nonlinear regime with Boltzmann electrons, a scan is performed over the flux-tube length, by increasing the number of poloidal turns, N_θ , thereby allowing the flux tube to sample more of the magnetic geometry. It is expected that, as $\rho_* \rightarrow 0$ and $N_\theta \rightarrow \infty$, the spectra for the potential and heat flux at high wavenumbers will converge. However, some disagreement may exist at lower wavenumbers due to the coupling into, and correct modelling of, the zonal mode by the FA code. Both the FA and FT results are obtained using **stella**. The FT simulations are performed explicitly for $\alpha = 0$, for the ‘bean-shaped’ cross-section, such that the flux tube exhibits up-down symmetry. This symmetry, and the small value of the global magnetic shear in W7-X, allows for the use of periodic

boundary conditions, described in Section 3.5.1.

In the FT scan, N_θ is increased from one to five, while maintaining a constant grid-point density in the z -direction. The regions of geometry sampled with each additional poloidal turn are illustrated in Figure 4.13, where the straight coloured lines indicate the z -domain introduced by each successive poloidal turn. The contours of the background represent the magnetic field strength for the high-mirror W7-X configuration, with red and blue indicating regions of stronger and weaker magnetic field strength respectively.

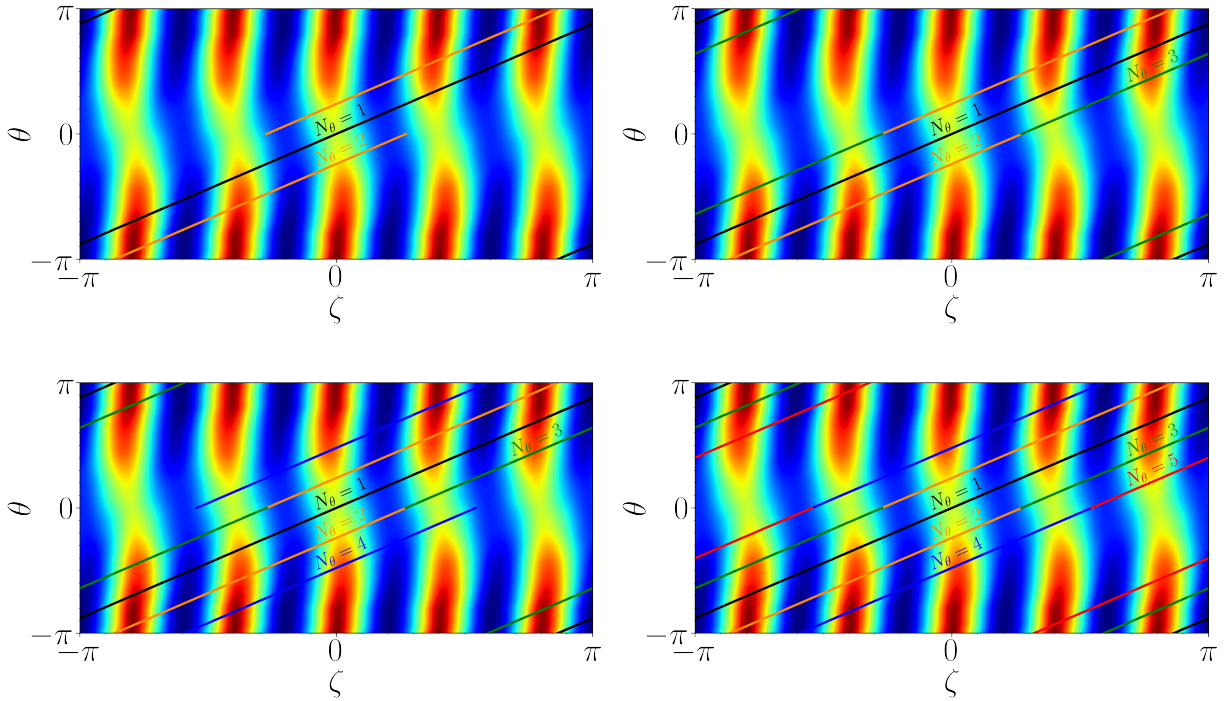


Figure 4.13: The geometry sampled by a single field line in an FT simulation, as it is extended for multiple poloidal turns. The flux tube for the ‘bean-shaped’ cross-section is centered at $(\alpha, \zeta) = (0, 0)$, and extended along the coloured lines for 2, 3, 4, and 5 poloidal turns, N_θ .

The resolutions used for the following nonlinear simulations are given in Table 4.3. Two sets of resolutions are used; ‘high’ and ‘low’ resolutions, which are used to ease computation for the FT simulations with higher numbers of poloidal turns. The time traces, and time-averages of the heat flux for the FT and FA simulations are shown in the left panel of Figure 4.14. The different coloured lines correspond to explicitly-run FT simulations with varying numbers of poloidal turns. The number of poloidal turns for each simulation is shown in the right panel of Figure 4.14. Each coloured point represents the saturated heat flux from the corresponding time trace, plotted against the number of poloidal turns. In this figure, the FA results are presented for both

<i>Parameter</i>	<i>FT - High Res.</i>	<i>FT - Low Res.</i>	<i>FA - $\rho_* = 0.025$</i>	<i>FA - $\rho_* = 0.025$</i>
r/L_r	0.5	0.5	0.5	0.5
L_r/L_T	3.0	3.0	3.0	3.0
L_r/L_n	1.0	1.0	1.0	1.0
N_θ	1, 1.5, 2, 2.5, 3	4, 5	$\iota = 0.88028$	$\iota = 0.88028$
N_z	$48N_\theta$	$36N_\theta$	48	36
N_μ	24	12	24	12
$N_{v_\parallel}$	72	72	72	72
N_x	108	81	108	81
N_y	128	128	128	324
y_0	20	20	20	60
ρ_*	0.0	0.0	0.025	0.01
x_0	30	30	30	30

Table 4.3: List of simulation parameters for the nonlinear simulations in the high-mirror W7-X configuration with Boltzmann electrons.

the explicit and implicit algorithms at $\rho_* = 0.025$, along with another explicit FA simulation conducted at $\rho_* = 0.01$. These are overlaid in the left panel of Figure 4.14, and the black horizontal lines give the corresponding saturated heat flux for the FA simulations. All simulations give saturated heat fluxes that are in reasonable agreement. The FT simulations exhibit variation in heat flux with the number of poloidal turns. It should also be noted that FT simulations with a larger number of poloidal turns ideally require longer run times to ensure confidence in the saturation level.

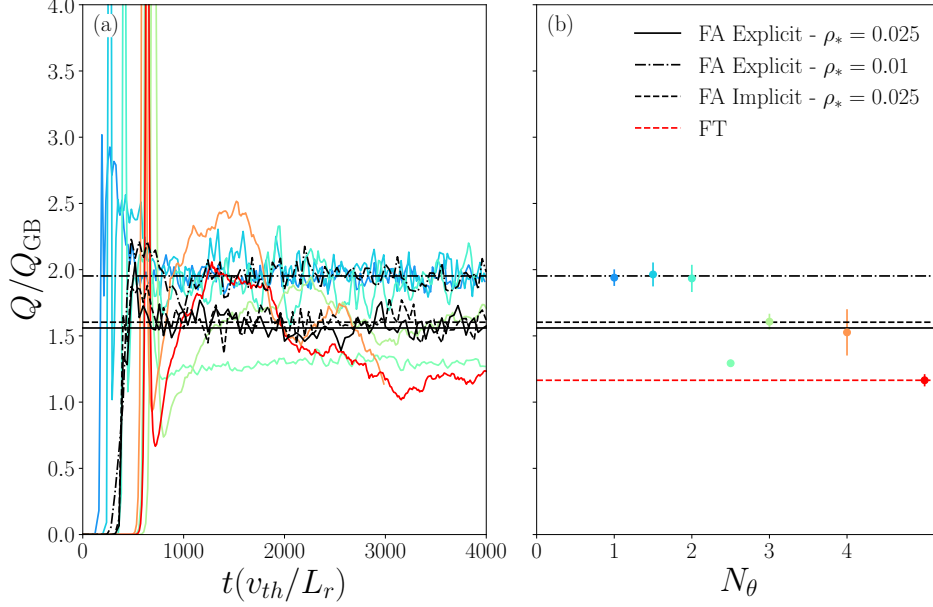


Figure 4.14: Time trace (a), and time-average (b) of the heat flux from flux-tube simulations with varying parallel domain lengths, compared with the corresponding flux-annulus result. FA results for $\rho_* = 0.025$ (solid and dashed for explicit and implicit algorithms respectively) and $\rho_* = 0.01$ (dot-dashed line), are given. The FT results are performed with $y_0 = 20$.

The corresponding potential and heat flux spectra for these simulations are provided in Figure 4.15. The explicit and implicit algorithms for the FA simulations show excellent agreement for the same value of ρ_* . In the FT simulations with fewer poloidal turns, the spectra show different behaviour at higher mode numbers and the peaks for both the potential and heat flux are at larger values of $k_y \rho_r$. As the number of poloidal turns increases, the FT spectra progressively converge towards those of the FA simulations, with the exception of the lowest $k_y \rho_r$ values, where discrepancies remain likely due to the influence of zonal flows.

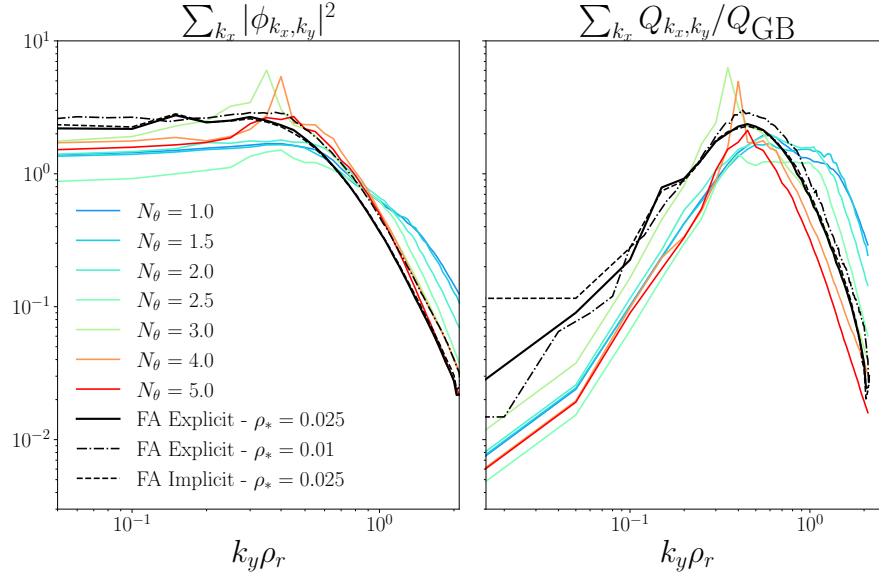
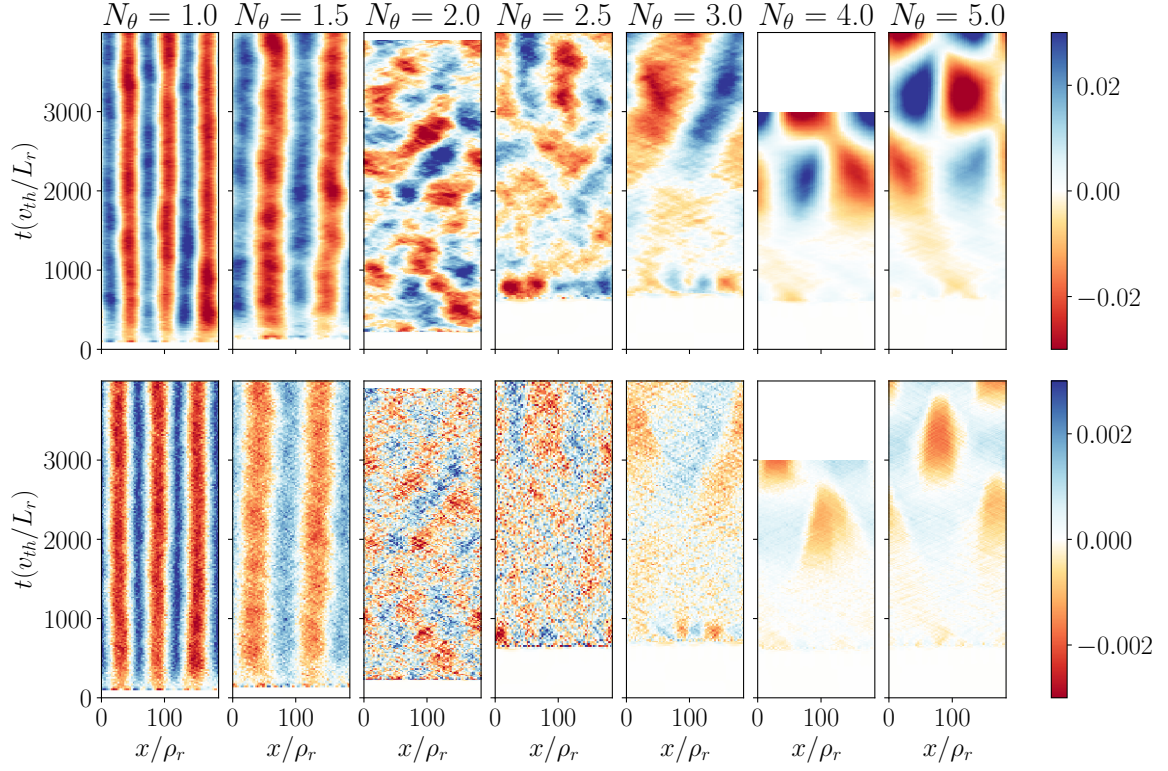
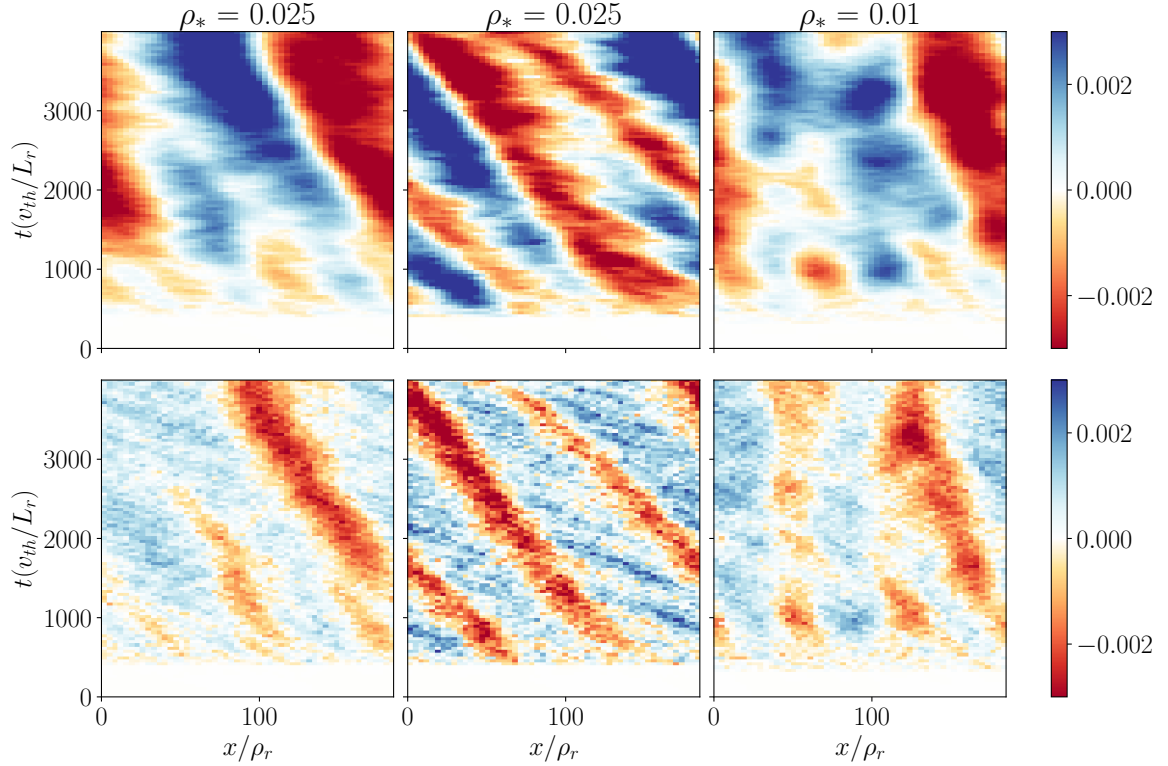


Figure 4.15: Spectra of the squared electrostatic potential (a) and heat flux (b) from nonlinear simulations. Coloured lines indicate FT simulations with varying numbers of poloidal turns corresponding to the same values as in Figure 4.14. Black lines show FA simulation results for both explicit and implicit algorithms, and for different values of ρ_* . These spectra correspond to the time-averaged quantities presented in Figure 4.14.

To see what may be causing the differences between the FT and FA codes at smaller k_y -values, the zonal potential and zonal flow are plotted in Figure 4.16 against time and the real-space radial coordinate, x . Though the saturated heat fluxes and spectra given reasonable agreement, it is clear that the evolution of the zonal components are very different in the FT and FA cases. As the number of poloidal turns increases for the FT simulations, the zonal potential and zonal flow begin to more closely resemble those of the FA result. However there are still noticeable differences between the FA and FT results even for the $N_\theta = 5$ FT simulation.



(a) Zonal potential (top) and zonal flow (bottom) for the FT simulations with different N_θ .



(b) Zonal potential (top) and zonal flow (bottom) from FA simulations. From left to right, the panels show the results of the explicit algorithm with $\rho_* = 0.025$, implicit algorithm with $\rho_* = 0.025$, explicit algorithm with $\rho_* = 0.01$.

Figure 4.16: The amplitudes of the zonal potential and zonal flow are compared between the FT and FA results. These are plotted as functions of radial position and time, normalised to the maximum value of ϕ .

4.4.3.3 Numerical Cost

The cost associated with a nonlinear **stella** FT simulation for $N_\theta = 1$ was the lowest among the FT cases considered. this simulation used 5 nodes on Leonardo, with a total of 560 CPUs, and it completed in approximately 48 hours. Consequently, the total computational cost amounted to 26,880 CPU hours. However, as the number of poloidal turns increased, the computational cost of each FT simulation increases approximately linearly. For instance, the case with $N_\theta = 5$ —which was run with lower resolutions—required around 144 hours to complete, translating to a total of 80,640 CPU hours. It should be highlighted that the $N_\theta = 5$ case was run with lower resolutions than the $N_\theta = 1$ simulations and the corresponding FA simulations. Had it been run with the same resolutions as the $N_\theta = 1$ simulation, it would have been more expensive.

In comparison, the case using the explicit FA algorithm with $\rho_* = 0.025$ required around 43,550 CPU hours, and the implicit FA simulation required a total of 51,690 CPU hours to simulate the 4000 normalised time units. It is important to note that the implicit algorithm is not expected to provide significant computational savings in cases involving Boltzmann electrons. It is included primarily as a benchmark for evaluating implicit methods. The computational benefits from implicit algorithms arise when treating kinetic electrons. In nonlinear simulations with Boltzmann electrons, the time step is constrained by nonlinearity for both the implicit and explicit schemes. Since the implicit advance is slightly more computationally intensive than its explicit counterpart, simulations using the FA method with Boltzmann electrons are actually more expensive when run implicitly with the same time step. This is also true for the FT algorithms. Although both explicit and implicit FA simulations are more expensive than the FT simulation with $N_\theta = 1$, it remains significantly less costly than the FT simulation with $N_\theta = 5$, which used lower resolutions, and is required in order to begin capturing the correct zonal dynamics. The computational cost of the various simulations, including the explicit FA simulation with $\rho_* = 0.01$, are summarised in Table 4.4.

<i>Simulation Case</i>	<i>Time Step (dt)</i>	<i>CPU Hours</i>
FT, $N_\theta = 1$	0.049666	26,880
FT, $N_\theta = 5$	0.030470	80,640
FA (explicit), $\rho_* = 0.025$	0.1	43,550
FA (implicit), $\rho_* = 0.025$	0.1	51,690
FA (explicit), $\rho_* = 0.01$	0.1	35,420

Table 4.4: Computational cost for the nonlinear simulations with Boltzmann electrons.

4.4.4 Nonlinear Results with Kinetic Electron Results

The nonlinear results incorporating kinetic electrons are now presented. The simulation parameters employed are detailed in Table 4.5. These results have been obtained using the implicit algorithms outlined earlier in this chapter, which is necessary due to the stringent CFL conditions introduced by simulating kinetic electron dynamics.

In preliminary tests with the FA version for linear simulations using an explicit scheme, it was observed that a time step on the order of 1×10^{-9} was required to ensure numerical stability. Since the timestep in nonlinear simulations is constrained by the underlying linear dynamics this timestep can be taken as an indicative upper bound for the nonlinear simulations (it is possible that the nonlinear term may restrict the time step further). By contrast, the nonlinear kinetic electron simulations presented here with the implicit scheme were numerically stable using a significantly larger timestep of approximately 10^{-3} . This clearly demonstrates the necessity for an implicit approach when treating kinetic electrons.

<i>Parameter</i>	<i>FT Simulation 1</i>	<i>FT Simulation 2</i>	<i>FA Simulation 1</i>	<i>FA Simulation 2</i>
r/L_r	0.5	0.5	0.5	0.5
L_r/L_T	3.0	3.0	3.0	3.0
L_r/L_n	1.0	1.0	1.0	1.0
N_θ	1	1	1.136	0.2272
N_z	64	64	64	16
N_μ	12	24	12	24
$N_{v_\parallel}$	72	72	72	72
N_x	96	108	96	108
N_y	96	144	96	256
y_0	10	20	10	50
ρ_*	0.0	0.0	0.05	0.01
x_0	10	20	10	50
Δt (CFL limited)	1.6577×10^{-2}	2.1195×10^{-2}	7.7464×10^{-3}	5.6368×10^{-3}
t_{final}	1500	1470	355	270
CPU hours	46,080	76,800	107,520	128,640

Table 4.5: List of simulation parameters for the two sets of nonlinear simulations with kinetic electrons.

The time trace for the first set of nonlinear simulations with kinetic electrons in the high-mirror W7-X geometry (parameter sets Simulation 1 in Table 4.5) is shown in Figure 4.17. In this figure, the red lines represent the FT results with one poloidal turn, while the black lines correspond to the FA results over one full toroidal turn. The associated spectra of the squared electrostatic potential and the heat flux are shown in Figure 4.18.

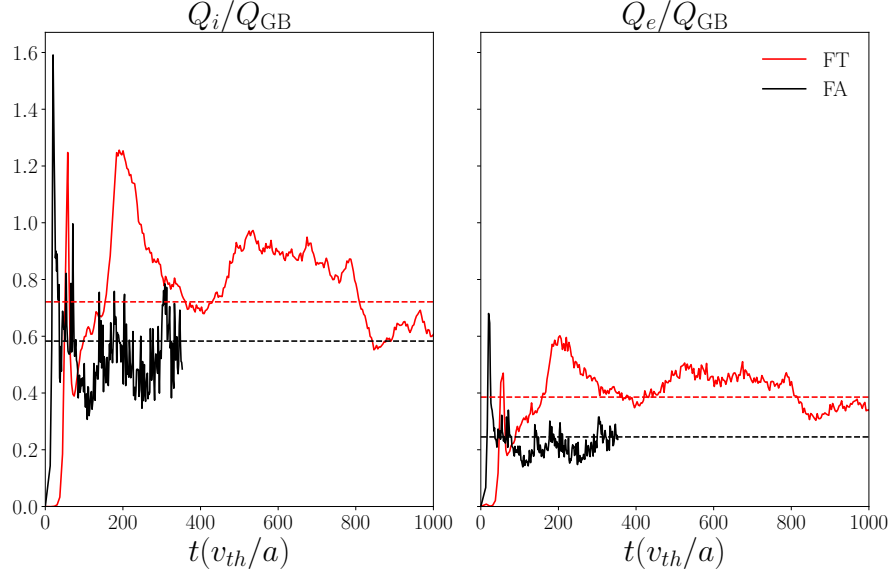


Figure 4.17: Time traces from nonlinear simulations with kinetic electrons in the high-mirror W7-X geometry (parameter sets Simulation 1 in Table 4.5). The left and right panels show the ion and electron heat fluxes, respectively. Red lines indicate the FT results, while black lines represent the corresponding FA results.

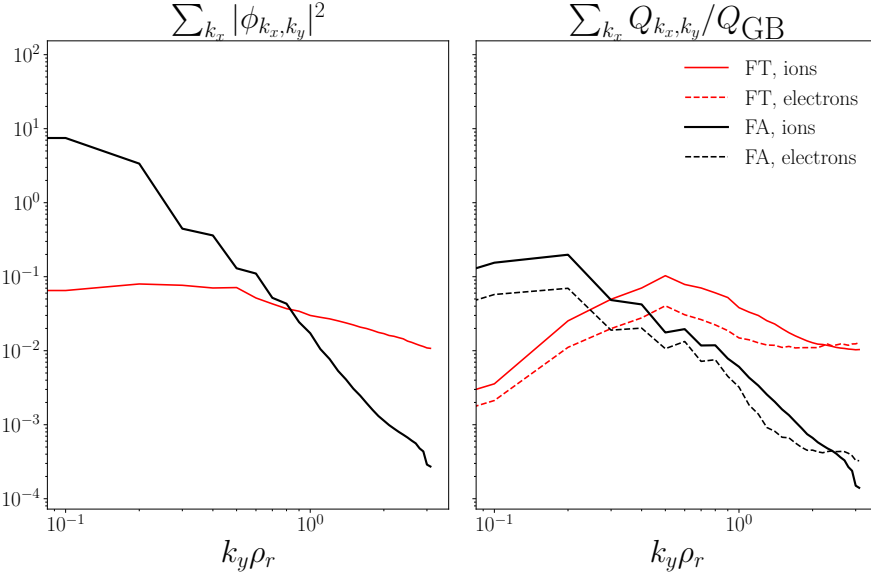


Figure 4.18: Spectra of the squared electrostatic potential and heat flux from the nonlinear simulations with kinetic electrons, corresponding to Figure 4.17.

Notable differences are observed between the FT and FA simulations. The FT heat flux peaks at a wavenumber $k_y \rho_r$ of order unity. In contrast, the FA heat flux peaks much closer to the system (box) scale. This shift may be a result of increased coupling with the zonal flows in the FA case, which tends to suppress small-scale fluctuations and drive the energy towards larger spatial scales.

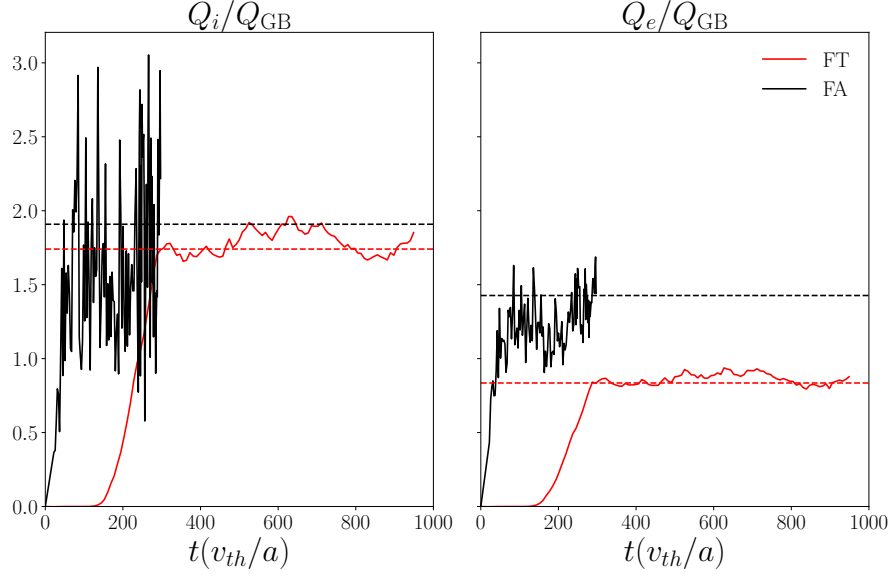


Figure 4.19: Time traces from nonlinear simulations with kinetic electrons in the high-mirror W7-X geometry (parameter sets Simulation 2 in Table 4.5). The left and right panels show the ion and electron heat fluxes, respectively. Red lines indicate the FT results, while black lines represent the corresponding FA results.

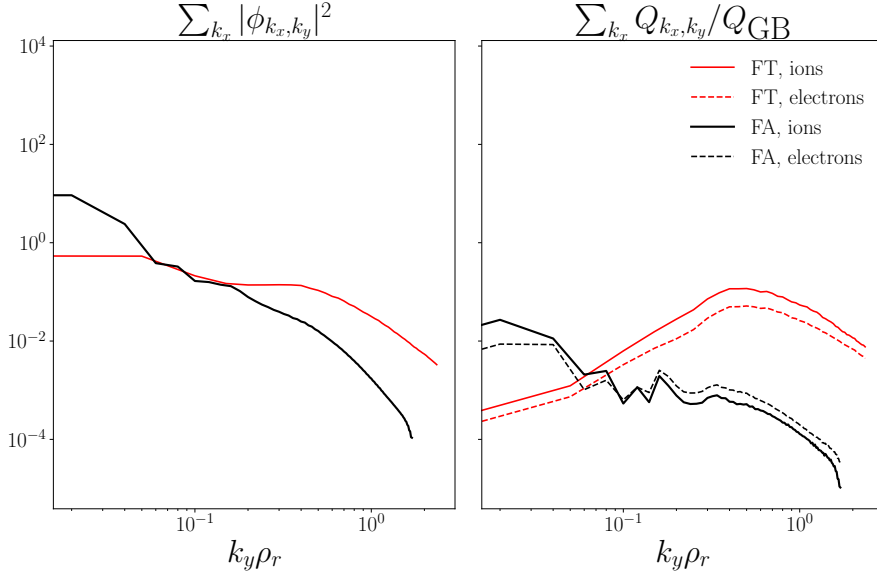


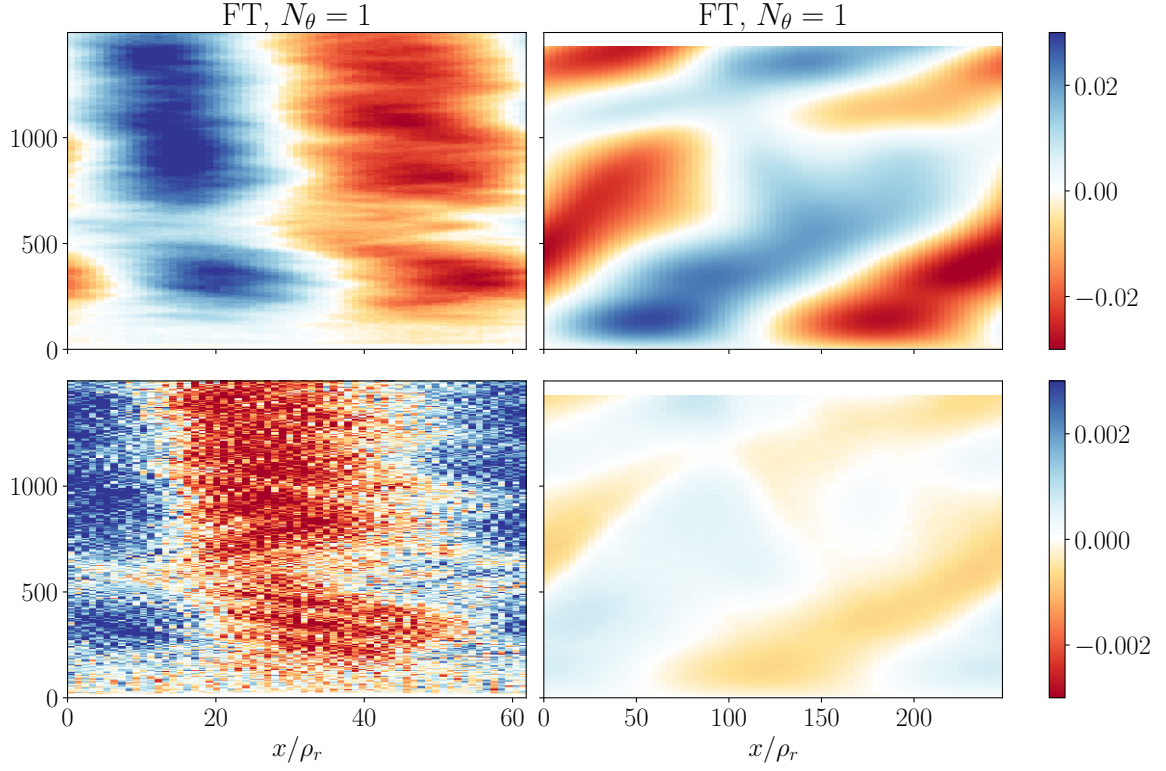
Figure 4.20: Spectra of the squared electrostatic potential and heat flux from the nonlinear simulations with kinetic electrons corresponding to Figure 4.19.

The time trace for the second set of nonlinear simulations, carried out at $\rho_* = 0.01$ with kinetic electrons, is presented in Figure 4.19 (parameter sets Simulation 2 in Table 4.5). Given the box-scale character of the turbulence observed in the previous FA example, a significantly smaller ρ_* was used for the FA simulations in this case.

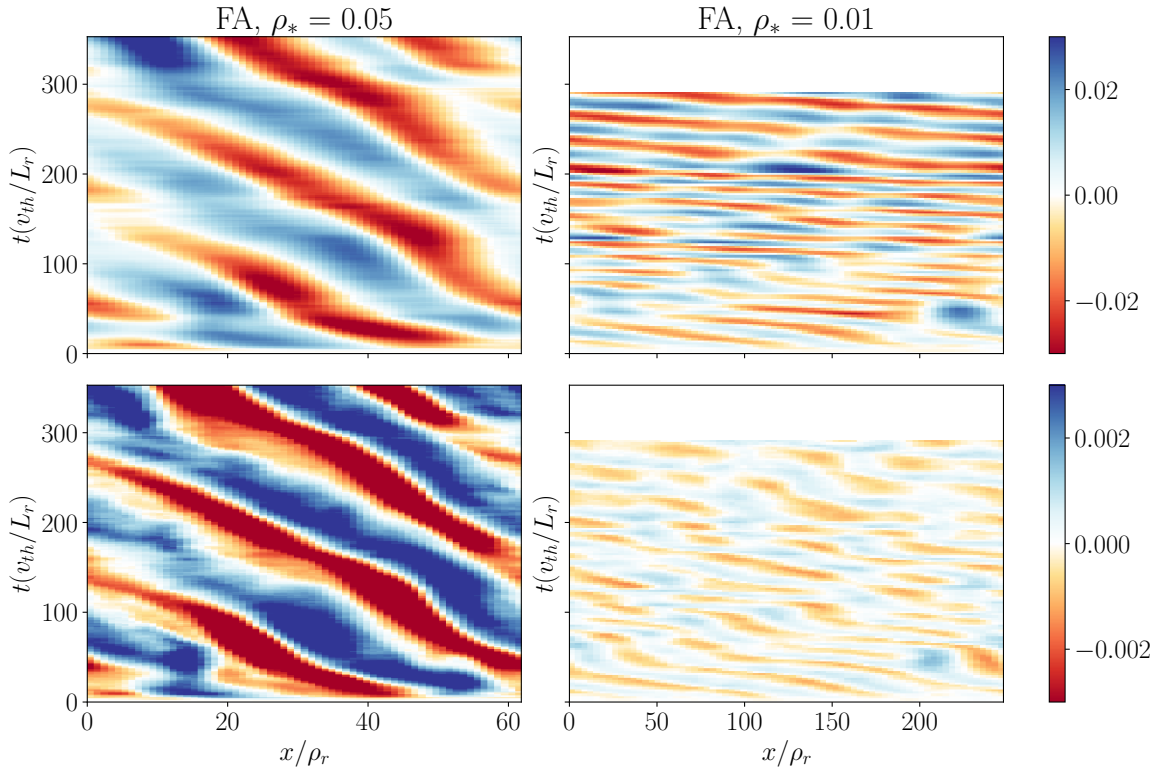
Since the previous FT simulations did not exhibit strong box-scale behaviour, only a modest increase in y_0 was required, as the previous domain already captured the peak of the turbulence. As before, the red lines represent the FT results over one poloidal turn. Here, however, the black lines correspond to FA results computed over a single field period (i.e. one fifth of the device), rather than a full toroidal turn. This reduced domain was chosen to alleviate the substantial memory constraints associated with the FA simulations. The corresponding spectra of the squared electrostatic potential and the heat flux are shown in Figure 4.20.

As with the earlier simulations, the FA results exhibit a peak at the box-scale. This again may suggest an enhanced interaction between the turbulence and the magnetic geometry in the FA simulations, particularly at lower k_y values. The influence of long-wavelength modes and zonal flows appears to be more pronounced in these configurations when using FA compared with FT approaches. However, it should be noted that there is an $\mathcal{O}(1)$ difference in the heat flux between the FT simulations with different values of x_0 and y_0 , which suggests that box-scale effects may already be influencing the FT results. In particular, the y -extent of the second FT simulation is comparable to the flux surface circumference, implying that the FA simulations are capturing geometric variations on a spatial scale similar to that of the turbulence itself. In this context, it is perhaps not surprising that the long-wavelength components of the spectra differ so markedly between the FT and FA cases, as the FA simulations are more sensitive to global geometric features than the more local FT approach may not fully resolve.

The zonal electrostatic potential and corresponding zonal flows for these simulations are shown in Figure 4.21. A clear distinction can be seen between the FA and FT cases: the FA simulations exhibit significantly finer zonal structure compared to the broader, more coherent zonal modes in the FT case. This trend is even more pronounced in the lower ρ_* FA simulation, where the zonal structures are finer still, and the zonal modes appear to propagate radially at a faster rate. While these observations are intriguing, further investigation is needed to fully understand the underlying mechanisms driving this behaviour, assess its physical realism, and explore its sensitivity to variations in ρ_* .



(a) Zonal potential (top) and zonal flow (bottom) for the FT simulations with $y_0 = 10$ (left) and $y_0 = 20$ (right).



(b) Zonal potential (top) and zonal flow (bottom) from FA simulations, with $\rho_* = 0.05$ (left), and $\rho_* = 0.01$ (right).

Figure 4.21: The amplitudes of the zonal potential and zonal flow are compared between the FT and FA results. These are plotted as functions of radial position and time, normalised to the maximum value of ϕ .

Chapter 5

Optimisation of Linear Microstability Using an Adjoint Method

We did not embark upon this calculation because it was easy. We did it because we thought it would be easy.

Michael Hardman

This chapter is adapted from Acton et al. 2024 [3].

5.1 Introduction

The growth rate of linear microinstabilities can be influenced by a large number of parameters, and their gradients, including plasma density, temperature, flow profiles, and the magnetic geometry. In an idealised situation, one would determine how such growth rates depend on all parameters governing the system, and then design MCF devices that are optimised with respect to microstability. However analytical searches are intractable, and, because the number of tunable parameters in modern MCF devices is large, full numerical parameter scans are often prohibitively expensive.

This chapter aims to address the optimisation challenge by employing an adjoint approach [96] that enables efficient calculation of the local variation of the linear growth rate with respect to all parameters of interest within the local, δf -gyrokinetic model. In contrast to a finite-difference calculation, the adjoint method is essentially independent of the dimension of the parameter space. Consequently, it can be used to optimise over a large number of variables at once, without incurring additional computational cost beyond solving the system equations; the associated cost of the adjoint method is roughly equivalent to solving the linearised gyrokinetic system of equations twice.

The application of the adjoint method to the linear gyrokinetic equation is presented and its utility in the optimisation process is demonstrated by calculating the sensitivity of the linear growth rate of a selected microinstability to geometrical parameters in a tokamak. The microinstability considered is the ITG mode introduced in Section 1.1.1, and the aim is to maximise the critical ion temperature gradient for instability onset with respect to variations of the local flux-surface shape. In general, the adjoint method presented here is agnostic to the mechanism behind the instability, and gives the gradient of the linear growth rate with respect to a set of external parameters, independent of the mode type. Given that the linear growth rate is a continuous function this approach can be applied in situations where the instability mechanism changes.

Adjoint methods are a powerful tool for efficiently calculating the numerical derivatives of an objective function for a linear system of equations. The gyrokinetic equation, along with its associated fields, can be linearised by neglecting the nonlinear terms. Because the distribution function and fields are time-dependent, one in principle needs to account for their time history when formulating an adjoint method. However, this can be obviated by decomposing the solutions into normal modes and considering the long-time limit, as we now describe. For a generic set of initial conditions, the fastest growing (or slowest decaying) mode is likely to have an initial amplitude that is comparable to, or smaller than, other normal modes in the system. This results in an initial transient period due to the superposition of these different normal modes, followed by exponential growth or decay once the fastest growing (or slowest decaying) mode has a much larger amplitude than all others. Figure 5.1 shows a schematic of the typical expected behaviour exhibited by the system electrostatic potential—the distribution function will behave similarly. A critical consideration for adjoint methods, particularly in time-dependent or high-dimensional systems, is the substantial memory requirement often imposed by the need to store the full forward solution trajectory for the backward adjoint integration. To circumvent this, we will adopt an adjoint formulation that avoids the need to store the time history of the forward problem by extracting the fastest growing mode.

Therefore one takes the following approach, decomposing the distribution function as¹ $\hat{g}_{\mathbf{k}} = \sum_j \bar{g}_{\mathbf{k},j} e^{\gamma_{\mathbf{k},j} t}$ with $\gamma_{\mathbf{k},j} \in \mathbb{C}$ the complex frequencies, and $\bar{g}_{\mathbf{k},j}$ the amplitude for the j^{th} normal mode². In the local approximation outlined in Section 2.3.2, quasineu-

¹It is worth remembering here that $\mathbf{k}$ denotes the Fourier mode from the spatial decomposition, and j is the subscript denoting the temporal normal mode.

²There are instances when the linear growth rates of two independent modes may be very similar, and a single dominant growth rate may require long run times to become evident.

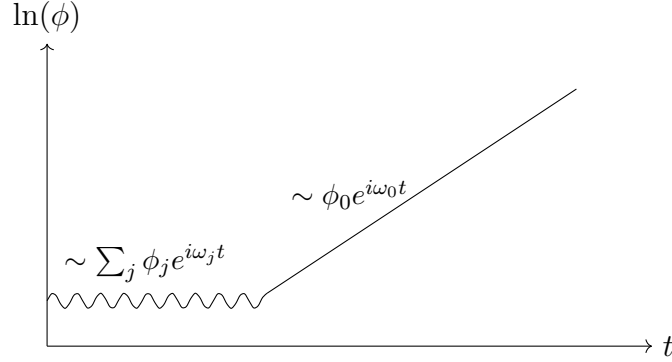


Figure 5.1: A schematic of the temporal evolution of the electric potential during a typical linear instability.

trality and Ampère’s law, along with the linearity of the gyrokinetic equation, ensure that all fluctuating quantities for a single normal mode, of a given wavenumber, share the same complex frequency and will thus exhibit the same time-dependent behaviour during the period of exponential growth or decay. In the non-local limit, as described in Section 2.3.1, the perturbations at all participating wavenumbers in field line label α will grow with a single global growth rate that is supported across the entire flux annulus. In both instances, after a sufficiently long period of time (the exact length of which will depend on both the relative growth rates and starting amplitudes of each mode), the fastest growing mode will dominate, meaning the time dependent behaviour at long times can be approximated using a single temporal mode; $\hat{g}_{\mathbf{k}} \approx \bar{g}_{\mathbf{k},0} e^{\gamma_{\mathbf{k},0} \bar{t}}$, with $\Re(\gamma_{\mathbf{k},0}) > \Re(\gamma_{\mathbf{k},j})$, $\forall j, j \neq 0$.

Here it is of interest to minimise the dominant linear growth rate in order to try and achieve microstability. The linear growth rate, $\gamma_{\mathbf{k},0}$, is minimised with respect to a set of currently unspecified parameters $\{p_i\}$, which are taken to be the components of the parameter vector $\mathbf{p} \in \mathbb{R}^{\mathcal{N}}$. This section outlines how one can take advantage of the adjoint method to efficiently obtain the gradient $\nabla_{\mathbf{p}} \gamma_{\mathbf{k},0}$ that is needed to achieve such a minimisation. In order to do this for an unspecified set of parameters, one could calculate the growth rates at nearby values of p_i and use a finite difference scheme to obtain the components of $\nabla_{\mathbf{p}} \gamma_{\mathbf{k},0}$. However, this is computationally expensive, scaling with the dimensionality of $\mathbf{p}$, denoted $\mathcal{N}_p$. In contrast, the adjoint method allows us to solve the system equations only once, and in exchange one must solve the set of adjoint equations, for which the cost is computationally roughly equivalent to the original system equations.

It should be noted that is also possible to use an eigenvalue solver for the above

decomposition and extraction of the fastest growing mode. However many gyrokinetic codes are initial-value solvers, which eases implementation and integration within existing simulation frameworks. Moreover, the computational cost of running an initial-value gyrokinetic simulation to extract the fastest growing linear eigenmode is comparable to that of using an eigenvalue solver.

It should also be noted that while focusing on a single, fastest-growing normal mode limits our analysis to established instabilities rather than their onset, this approximation offers significant computational advantages. Fully resolving the time-dependent PDE without this simplification would necessitate storing the entire system's time history, leading to prohibitive memory and computational costs for adjoint-based gradient calculations.

5.2 Adjoint Method Introduction and Procedure

The adjoint-based optimisation method is a procedure that, at its heart, efficiently calculates derivatives of a desired quantity with respect to a potentially large number of parameters. The associated computational cost depends only on the expense of solving both the original and adjoint systems of equations, and is essentially independent of the dimension of the parameter space. Adjoint methods have already been successfully applied to certain other aspects of MCF devices, such as optimising coil configurations for stellarator geometries, and divertor designs (see, e.g., [94, 44, 89, 29]). The novelty here is to apply the adjoint method to geometric optimisation and plasma microstability, which includes the complexity of the full linearised gyrokinetic system. To begin, the underlying principles of the adjoint method are introduced by considering a simplified model system.

In this section a simple model is introduced to demonstrate the implementation of the adjoint method. Consider a system defined through a constraint partial differential equation (PDE), which is the residual of the forward problem:

$$\hat{L}[\mathbf{p}; f(\mathbf{p}, \mathbf{s})] = 0. \quad (5.1)$$

Here $\hat{L}$ is some linear operator, in the domain $\Omega \subset \mathbb{R}^n$ that has explicit dependence on a parameter vector $\mathbf{p} \in \mathbb{R}^N$, and acts on a function $f(\mathbf{p}, \mathbf{s})$ which may also depend on another set of variables $\mathbf{s}$; specifically, $\hat{L}$ is taken to be some partial differential equation (PDE) constraint whose adjoint, $\hat{L}^\dagger$, exists and is well defined. The aim is to optimise an aspect of the system with respect to the set of independent parameters $\mathbf{p} = \{p_i\}$. Equation (5.1) is to be satisfied for a given $\mathbf{p}$ provided a consistent $f(\mathbf{p}, \mathbf{s})$.

The forward PDE may contain both explicit and implicit dependence on $\mathbf{p}$ through the operator itself, or through $f(\mathbf{p}, \mathbf{s})$.

Suppose the aim is to compute the gradient of the function $\hat{H} = \hat{H}[\mathbf{p}; f]$ with respect to $\mathbf{p}$, $\nabla_{\mathbf{p}} \hat{H} \in \mathbb{R}^N$. Here $\hat{H}[\mathbf{p}; f]$ is some functional operator that depends on $\mathbf{p}$ and acts on f . The function $\hat{H}$ can be chosen to be written in the following way:

$$\hat{H}[\mathbf{p}; f] = \langle \hat{h}[\mathbf{p}; f(\mathbf{p})], f \rangle, \quad (5.2)$$

with $\langle \cdot, \cdot \rangle$ some appropriate definition of inner product³, and $\hat{h}[\mathbf{p}; f]$ some linear operator whose adjoint exists and is well defined. The aim is to perturb $\mathbf{p}$ and study the response of $\hat{H}$, i.e., to compute the gradient with respect to $\mathbf{p}$. One method to calculate such derivatives would be to consider a standard finite-difference approach for each element of $\mathbf{p}$:

$$\frac{\partial \hat{H}}{\partial p_i} = \frac{\hat{H}[p_i + \delta p_i; f(p_i + \delta p_i)] - \hat{H}[p_i; f(p_i)]}{\delta p_i} + \mathcal{O}(\delta p_i^2). \quad (5.3)$$

However, this approach is computationally expensive when $\mathcal{N}_p$ is large, as equation (5.1) must be solved once to obtain the unperturbed values, and again for each $\{p_i + \delta p_i\}$ in order to compute each derivative. This process equates to solving the original PDE $\mathcal{N}_p + 1$ times. An alternative method is to use the adjoint technique.

5.3 Example: the Heat Conduction Equation

Consider the transient heat conduction equation in a domain $\Omega \subset \mathbb{R}^n$ over a time interval $t \in [0, T]$:

$$\frac{\partial f}{\partial t} - \nabla \cdot (\kappa \nabla f) = S \quad \text{in } \Omega \times (0, T), \quad (5.4)$$

with boundary and initial conditions

$$\begin{aligned} f &= f_D \quad \text{on } \partial\Omega_D \times (0, T), \\ -\kappa \nabla f \cdot \hat{\mathbf{n}} &= q \quad \text{on } \partial\Omega_N \times (0, T), \\ f(\mathbf{x}, 0) &= f_0(\mathbf{x}) \quad \text{in } \Omega. \end{aligned} \quad (5.5)$$

Here, f is the temperature field, κ is the thermal conductivity, S is a heat source term, and $\hat{\mathbf{n}}$ is the outward unit normal to the boundary. Define a cost functional H

³The form of inner product is chosen consistently and appropriately for the system that is being considered.

that we aim to minimise, for example:

$$H = \frac{1}{2} \int_0^T \int_{\Omega} (f - f_{\text{target}})^2 \, d\mathbf{x} \, dt, \quad (5.6)$$

where f_{target} is a desired temperature distribution. To derive the adjoint problem, we introduce a Lagrangian functional that incorporates the forward PDE constraint using an adjoint variable, $\lambda(\mathbf{x}, t)$:

$$\mathcal{L}[f, \lambda] = H(f) + \int_0^T \int_{\Omega} \lambda \left(\frac{\partial f}{\partial t} - \nabla \cdot (\kappa \nabla f) - S \right) \, d\mathbf{x} \, dt. \quad (5.7)$$

To find the stationary point of $\mathcal{L}$, we take the variation with respect to f and integrate by parts. Setting the first variation to zero yields the adjoint equation:

$$-\frac{\partial \lambda}{\partial t} - \nabla \cdot (\kappa \nabla \lambda) = f - f_{\text{target}} \quad \text{in } \Omega \times (0, T), \quad (5.8)$$

with final-time and boundary conditions:

$$\begin{aligned} \lambda(\mathbf{x}, T) &= 0 \quad \text{in } \Omega, \\ \lambda &= 0 \quad \text{on } \partial\Omega_D \times (0, T), \\ k \nabla \lambda \cdot \mathbf{n} &= 0 \quad \text{on } \partial\Omega_N \times (0, T). \end{aligned} \quad (5.9)$$

The adjoint variable λ evolves backward in time and provides gradient information for sensitivity analysis or optimisation. Once λ is computed, the gradient of the cost functional with respect to a control parameter (e.g., S , u_0 , or boundary conditions) can be evaluated using the chain rule involving λ .

5.3.1 Gradient Computation via the Adjoint Method

Suppose the heat conduction equation depends on the parameter vector $\mathbf{p} \in \mathbb{R}^{\mathcal{N}}$, which may represent, for example, a distributed heat source, an initial condition, or a spatially varying conductivity. The solution $f = f(\mathbf{p})$ depends on $\mathbf{p}$ through the PDE, and the objective functional $H = H[\mathbf{p}; f(\mathbf{p})]$ depends both on the state and the control.

Our goal is to compute the gradient of H with respect to $\mathbf{p}$, denoted $\nabla_{\mathbf{p}} H \in \mathbb{R}^{\mathcal{N}}$. Direct computation of this gradient via finite differences or sensitivity analysis requires solving $\mathcal{N} + 1$ forward problems, which becomes computationally intractable for large

$\mathcal{N}$. The adjoint method provides a more efficient alternative. Consider the residual of the heat conduction equation (forward problem) as a constraint:

$$\hat{L}[\mathbf{p}; f(\mathbf{p})] := \frac{\partial f}{\partial t} - \nabla \cdot (\kappa \nabla f) - S = 0 \quad \text{in } \Omega \times (0, T), \quad (5.10)$$

with appropriate boundary and initial conditions. The dependence on $\mathbf{p}$ enters through the source term $S(\mathbf{p})$, $\kappa(\mathbf{p})$ or other parts of the model. Define the Lagrangian functional:

$$\mathcal{L}[\mathbf{p}; f, \lambda] = H[\mathbf{p}; f(\mathbf{p})] + \int_0^T \int_{\Omega} \lambda \left(\frac{\partial f}{\partial t} - \nabla \cdot (\kappa \nabla f) - S \right) d\mathbf{x} dt, \quad (5.11)$$

where $\lambda = \lambda(\mathbf{x}, t)$ is the adjoint variable. We seek the gradient derivative of H with respect to $\mathbf{p}$:

$$\nabla_{\mathbf{p}} H = \frac{\partial H}{\partial f} \nabla_{\mathbf{p}} f + \frac{\partial H}{\partial p}. \quad (5.12)$$

Since f satisfies the PDE constraint, we apply the adjoint method to eliminate the need to compute $\nabla_{\mathbf{p}} f$ directly. We define the adjoint equation by requiring that the variation of the Lagrangian with respect to f vanishes:

$$\frac{\delta \mathcal{L}}{\delta f} = 0 \quad \Rightarrow \quad -\frac{\partial \lambda}{\partial t} - \nabla \cdot (\kappa \nabla \lambda) = \frac{\partial H}{\partial f} \quad \text{in } \Omega \times (0, T), \quad (5.13)$$

with final time and boundary conditions:

$$\lambda(\mathbf{x}, T) = 0, \quad (5.14)$$

$$\lambda = 0 \quad \text{on } \partial\Omega_D, \quad (5.15)$$

$$k \nabla \lambda \cdot \mathbf{n} = 0 \quad \text{on } \partial\Omega_N. \quad (5.16)$$

Once the adjoint solution λ is obtained, the gradient is given by:

$$\nabla_{\mathbf{p}} H = \frac{\partial H}{\partial p} - \int_0^T \int_{\Omega} \lambda \frac{\partial \hat{L}}{\partial p} d\mathbf{x} dt. \quad (5.17)$$

This formula expresses the gradient in terms of the partial derivative $\partial H / \partial p$, which captures the explicit dependence of H on $\mathbf{p}$, and the adjoint-weighted sensitivity of the PDE residual to $\mathbf{p}$.

5.4 Adjoint Method Introduction and Procedure

The optimisation Lagrangian must first be defined;

$$\mathcal{L}[\mathbf{p}; f, \lambda] = \hat{H}[\mathbf{p}; f(\mathbf{p})] + \left\langle \hat{L}[\mathbf{p}; f(\mathbf{p})], \lambda \right\rangle, \quad (5.18)$$

where f is part of the Hilbert space $f \in \mathcal{F}$, and $\lambda \in \mathcal{F}^*$ is an adjoint variable, residing in the dual space of $\mathcal{F}$. It should be noted that in the above expression f need not obey the original objective function defined in (5.1) when we consider variations of the Lagrangian; however, when we evaluate the Lagrangian using values of $\mathbf{p}$ and the corresponding $f(\mathbf{p})$ that satisfy the objective function, equation (5.18) reduces to $\mathcal{L} = \hat{H}$, as the original objective function vanishes, $\hat{L} = 0$, in accordance with (5.1).

By varying $\mathcal{L}$ with respect to perturbations in $\mathbf{p}$, the solutions that are stationary with respect to the functions f and λ can be found, subject to the constraint equation given by (5.1). From this a constraint equation for the adjoint variable λ is found, the solution to which will be denoted by λ . Perturbing the system given by equation (5.18) yields an expression for the gradient of $\hat{H}$, as shown in (5.19), where $\nabla_{\mathbf{p}}$ indicates the gradient derivative with respect to $\mathbf{p}$, and is evaluated by perturbing all quantities that have explicit and implicit $\mathbf{p}$ dependence⁴:

$$\nabla_{\mathbf{p}} \mathcal{L}[\mathbf{p}; f, \lambda] \Big|_{\mathbf{p}=\mathbf{p}_0, f=f(\mathbf{p}_0), \lambda=\lambda(\mathbf{p}_0)} = \nabla_{\mathbf{p}} \hat{H}[\mathbf{p}; f(\mathbf{p})] \Big|_{\mathbf{p}=\mathbf{p}_0, f=f(\mathbf{p}_0), \lambda=\lambda(\mathbf{p}_0)}, \quad (5.19)$$

with both sides being evaluated at $f = f(\mathbf{p}_0)$, $\lambda = \lambda(\mathbf{p}_0)$, and $\mathbf{p} = \mathbf{p}_0$, where $\mathbf{p}_0 \in \mathbb{R}^N$ are the initial set of parameters. Note that the term involving the adjoint variable is not present because both $\hat{L}[\mathbf{p}; f]$, and $\nabla_{\mathbf{p}} \hat{L}[\mathbf{p}; f]$ are zero when evaluated at $f = f(\mathbf{p}_0)$, in accordance with (5.1)⁵.

As the right-hand side of (5.19) is equivalent to the desired gradient of $\hat{H}$, the desired derivatives can be obtained by computing the gradient derivative of the optimisation Lagrangian given by equation (5.18). The aim of the adjoint method is to take advantage of the freedom available in choosing the adjoint variable to simplify the calculation of $\nabla_{\mathbf{p}} \hat{H}$, by eliminating the dependence on $\nabla_{\mathbf{p}} f$. The desired result is

$$\begin{aligned} \nabla_{\mathbf{p}} \hat{H}[\mathbf{p}; f(\mathbf{p})] \Big|_{\mathbf{p}_0} &= \nabla_{\mathbf{p}} \mathcal{L}[\mathbf{p}; f, \lambda] \Big|_{\mathbf{p}_0} \\ &= \partial_{\mathbf{p}} \hat{H}[\mathbf{p}; f] \Big|_{\mathbf{p}_0} + \left\langle \partial_{\mathbf{p}} \hat{L}[\mathbf{p}; f] \Big|_{\mathbf{p}_0}, \lambda \right\rangle, \end{aligned} \quad (5.20)$$

⁴This amounts to perturbing the explicit dependencies on $\mathbf{p}$ as it appears in (5.18), as well as f , and λ .

⁵ $\nabla_{\mathbf{p}} \hat{L}[\mathbf{p}; f] = 0$ as $\hat{L}[\mathbf{p}; f]$ is identically zero for each choice of $\mathbf{p}$ and $\mathbf{p} + \delta p_i$.

where equation (5.18) was used to evaluate the right-hand side. Here the notation $\partial_{\mathbf{p}}(\cdot)|_{\mathbf{p}_0}$ has been introduced for brevity to indicate that the variables f and λ are being held fixed when taking this partial derivative, and that the resulting expression has been evaluated at $\mathbf{p}_0$. It shall be assumed that everywhere $\partial_{\mathbf{p}}(\cdot)$ appears, f and λ are held constant for the partial differentiation.

The expressions derived here for $\nabla_{\mathbf{p}}\mathcal{L}$ and hence for $\nabla_{\mathbf{p}}\hat{H}$ are exact (to machine precision), as they arise from analytic differentiation using the tangent-linear and adjoint formulations. This stands in contrast to finite-difference approximations, which estimate the gradient by evaluating the model at perturbed parameter values and differencing. Such finite-difference methods incur numerical error and require multiple PDE solves across parameter space. In contrast, the adjoint-based approach used here requires only one forward and one adjoint solve and yields gradients accurate to machine precision.

The derivative of (5.18) is as follows:

$$\nabla_{\mathbf{p}}\mathcal{L}[\mathbf{p}; f, \lambda] = \nabla_{\mathbf{p}}\hat{H}[\mathbf{p}; f] + \left\langle \nabla_{\mathbf{p}}\hat{L}[\mathbf{p}; f], \lambda \right\rangle + \left\langle \hat{L}[\mathbf{p}; f], \nabla_{\mathbf{p}}\lambda \right\rangle + \partial_{\mathcal{J}} \left\langle (\nabla_{\mathbf{p}}\mathcal{J}) \hat{L}[\mathbf{p}; f], \lambda \right\rangle. \quad (5.21)$$

where the notation $\partial_{\mathcal{J}} \left\langle \nabla_{\mathbf{p}}(\mathcal{J}) \hat{L}, \lambda \right\rangle$ is used to indicate the derivative of any $\mathbf{p}$ -dependent terms appearing in the Jacobian of the inner product itself. It is now advantageous to split the derivatives into terms which explicitly act on f , and those which do not. To do this, write the explicit dependence of $\hat{H} = \hat{H}[\mathbf{p}; f]$ and $\hat{L} = \hat{L}[\mathbf{p}; f]$, to obtain the following expression:

$$\begin{aligned} \nabla_{\mathbf{p}}\mathcal{L}[\mathbf{p}; f, \lambda] &= \partial_{\mathbf{p}}\hat{H}[\mathbf{p}; f] + \left\langle \left. \frac{\delta\hat{H}}{\delta f} \right|_{\mathbf{p}; f}, \nabla_{\mathbf{p}}f \right\rangle \\ &\quad + \left\langle \partial_{\mathbf{p}}\hat{L}[\mathbf{p}; f], \lambda \right\rangle + \left\langle \left. \frac{\delta\hat{L}}{\delta f} \right|_{\mathbf{p}; f} [\nabla_{\mathbf{p}}f], \lambda \right\rangle + \left\langle \hat{L}[\mathbf{p}; f], \nabla_{\mathbf{p}}\lambda \right\rangle \\ &\quad + \partial_{\mathcal{J}} \left\langle (\nabla_{\mathbf{p}}\mathcal{J}) \hat{L}[\mathbf{p}; f], \lambda \right\rangle. \end{aligned} \quad (5.22)$$

Here $\delta\hat{H}/\delta f$ and $\delta\hat{L}/\delta f$ denote the functional derivatives of $\hat{H}$ and $\hat{L}$ with respect to f , respectively, and $[\nabla_{\mathbf{p}}f]$ indicates that the operator acts on the gradient $\nabla_{\mathbf{p}}f$. Solutions around $\mathbf{p}_0$ are sought, for which $\hat{L}[\mathbf{p}_0; f(\mathbf{p}_0)] = 0$ is satisfied. Hence, equation

(5.22) may be written as:

$$\begin{aligned} \nabla_{\mathbf{p}} \mathcal{L}[\mathbf{p}; f, \lambda] \big|_{\mathbf{p}_0, f, \lambda} &= \partial_{\mathbf{p}} \hat{H}[\mathbf{p}; f] \big|_{\mathbf{p}_0, f} + \left\langle \frac{\delta \hat{H}}{\delta f} \bigg|_{\mathbf{p}_0, f}, \nabla_{\mathbf{p}} f \right\rangle \\ &+ \left\langle \partial_{\mathbf{p}} \hat{L}[\mathbf{p}; f], \lambda \right\rangle \big|_{\mathbf{p}_0, f, \lambda} + \left\langle \frac{\delta \hat{L}}{\delta f} \bigg|_{\mathbf{p}_0, f} [\nabla_{\mathbf{p}} f], \lambda \right\rangle \bigg|_{\mathbf{p}_0, f, \lambda}. \end{aligned} \quad (5.23)$$

Note that the terms corresponding to the derivative of λ and the Jacobians have been removed, as $\hat{L}[\mathbf{p}_0; f(\mathbf{p}_0)]$ is identically zero. The computation of $\nabla_{\mathbf{p}} f$ is costly for a large number of parameters, as (5.1) necessarily needs to be solved for f at the initial coordinate, $\mathbf{p}_0$, and for each new perturbed coordinate. As an alternative, the coefficients of $\nabla_{\mathbf{p}} f$ terms are collected by performing integrations by parts to move the functional derivatives to act on f and λ :

$$\begin{aligned} \nabla_{\mathbf{p}} \mathcal{L}[\mathbf{p}; f, \lambda] \big|_{\mathbf{p}_0, f, \lambda} &= \partial_{\mathbf{p}} \hat{H}[\mathbf{p}; f] \big|_{\mathbf{p}_0, f, \lambda} + \left\langle \partial_{\mathbf{p}} \hat{L}[\mathbf{p}; f], \lambda \right\rangle \big|_{\mathbf{p}_0, f, \lambda} \\ &+ \left\langle \left(\frac{\delta \hat{H}}{\delta f} \right)^{\dagger} [\mathbf{p}; f] + \left(\frac{\delta \hat{L}}{\delta f} \right)^{\dagger} [\mathbf{p}; \lambda], \nabla_{\mathbf{p}} f \right\rangle \bigg|_{\mathbf{p}_0, f, \lambda} + \mathcal{B.C.} \end{aligned} \quad (5.24)$$

Here $\mathcal{B.C.}$ are the boundary conditions that result from the integrations by parts, which themselves contain terms involving $\nabla_{\mathbf{p}} f$. The adjoint operators $\left(\frac{\delta \hat{H}}{\delta f} \right)^{\dagger}$ and $\left(\frac{\delta \hat{L}}{\delta f} \right)^{\dagger}$ are defined such that they satisfy the integration by parts. In order to remove the requirement to find $\nabla_{\mathbf{p}} f$, the adjoint variable λ is sought such that it satisfies the adjoint equation:

$$\left(\frac{\delta \hat{H}}{\delta f} \right)^{\dagger} [\mathbf{p}; f] + \left(\frac{\delta \hat{L}}{\delta f} \right)^{\dagger} [\mathbf{p}; \lambda] = 0. \quad (5.25)$$

The boundary conditions for λ are not arbitrary choices, but are entirely determined by the forward PDE (equation 5.1) and the functional form of $\hat{H}$ through the variational principle, ensuring that the $\mathcal{B.C.}$ term vanishes. With this (5.24) reduces to:

$$\nabla_{\mathbf{p}} \mathcal{L}[\mathbf{p}; f, \lambda] \big|_{\mathbf{p}_0} = \partial_{\mathbf{p}} \hat{H}[\mathbf{p}; f] \big|_{\mathbf{p}_0} + \left\langle \partial_{\mathbf{p}} \hat{L}[\mathbf{p}; f], \lambda \right\rangle \big|_{\mathbf{p}_0}, \quad (5.26)$$

where f and λ are the solutions to (5.1) and (5.25) respectively. Hence, using the analog of (5.19), the desired result in (5.20) is recovered.

5.5 Adjoint Method for Gyrokinetics

The aim is to apply these methods to gyrokinetics to optimise for microstability. To do this, we consider the linear gyrokinetic equation in the limit of no α -coupling. We further consider the behaviour beyond the transient period, illustrated in Figure 5.1, for which there is a single dominant normal mode. In this limit the distribution function can be approximated as $\hat{g}_{\mathbf{k}} \approx \bar{g}_{\mathbf{k},0} e^{\gamma_{\mathbf{k},0} t}$, with $\Re(\gamma_{\mathbf{k},0}) > \Re(\gamma_{\mathbf{k},j})$, $\forall j, j \neq 0$. In this case the time derivative can be well approximated as $\partial \hat{g}_{\mathbf{k}} / \partial t \approx \gamma_{\mathbf{k},0} \bar{g}_{\mathbf{k},0}$, and the unnormalised gyrokinetic equation becomes

$$\begin{aligned} \hat{G}_{\mathbf{k},\nu} = & \gamma_{\mathbf{k},0} \bar{g}_{\mathbf{k},\nu,0} + v_{\parallel} \hat{\mathbf{b}} \cdot \nabla z \left[\frac{\partial \bar{g}_{\mathbf{k},\nu,0}}{\partial z} + \frac{Z_{\nu} e}{T_{\nu}} \frac{\partial \bar{\chi}_{\mathbf{k},\nu,0}}{\partial z} F_{0,\nu} \right] \\ & + i\omega_{*,\mathbf{k},\nu} F_{0,\nu} \bar{\chi}_{\mathbf{k},\nu,0} + i\omega_{d,\mathbf{k},\nu} \left[\bar{g}_{\mathbf{k},\nu,0} + \frac{Z_{\nu} e}{T_{\nu}} \bar{\chi}_{\mathbf{k},\nu,0} F_{0,\nu} \right] \\ & - \mu_{\nu} \hat{\mathbf{b}} \cdot \nabla B_0 \frac{\partial \bar{g}_{\mathbf{k},\nu,0}}{\partial v_{\parallel}} + \frac{Z_{\nu} e}{T_{\nu}} \frac{\mu_{\nu}}{m_{\nu} c} (\hat{\mathbf{b}} \cdot \nabla B_0) F_{0,\nu} J_{0,\mathbf{k},\nu} \bar{A}_{\parallel,\mathbf{k},0} - \hat{C}_{\mathbf{k},\nu} [\{\bar{g}_{\mathbf{k},\nu',0}, \bar{\chi}_{\mathbf{k},\nu',0}\}]. \end{aligned} \quad (5.27)$$

We start by considering the general case of low-flow, linear, electromagnetic, δf -gyrokinetics including collisions, with equations (5.27) and (2.45)-(2.47) as the functional operators defining the system in the long time limit.

5.5.1 General Formalism

Since we consider the gyrokinetic system in the linear regime, with no α -coupling, it is possible to consider each perpendicular wavenumber individually. Given that the post-transient limit is being considered, with only one dominant growth rate, it will henceforth be assumed that only a single perpendicular wavenumber is being considered, and the $\mathbf{k}$ subscript will be dropped, along with the subscript that denotes the dominant mode for the distribution function and field quantities. For notational simplicity the over-bars that appear on the distribution function and fields that denoted normal mode decomposition will also be dropped. Hence everywhere $\hat{g}_{\nu}$, $\hat{\phi}$, $\hat{A}_{\parallel}$, and $\delta \hat{B}_{\parallel}$ appear it shall be assumed that they contain suppressed spatial and temporal Fourier subscripts.

Consider now a set of variables that influences the system and which exists within a parameter space spanned by all possible $\mathbf{p}$. The set $\{p_i\}$, $i \in [1, \mathcal{N}_p]$ is taken to be linearly independent, with no time variation, and it shall be explored how variations within this space affect the linear growth rate. At present one need not specify which

variables are denoted by $\mathbf{p}$, and thus one derives a general set of adjoint equations for the unnormalised gyrokinetic-Maxwell system above, (5.27) and (2.45)-(2.47).

The objective functional, $\hat{G}_\nu$, is a linear functional of $\{\hat{g}_\nu, \hat{\phi}, \hat{A}_\parallel, \delta\hat{B}_\parallel\}$, which are coupled to the fields through the field equations, (2.45)-(2.47). Taking the derivatives of $\hat{G}_\nu$ invariably leads to derivatives acting on all four of these variables. This is undesirable because it requires calculation of the gradients of $\hat{g}_\nu$ and the fields, which in turn requires the gyrokinetic system to be solved $\mathcal{N}_p + 1$ times, as discussed above. In order to eliminate these four derivatives, four adjoint variables are introduced, which multiply the corresponding constraint equations. The optimisation Lagrangian is thus

$$\mathcal{L} := \left\langle \hat{G}_\nu, \lambda_\nu \right\rangle_{z,v,\nu} + \left\langle \hat{Q}, \xi \right\rangle_z + \left\langle \hat{M}, \zeta \right\rangle_z + \left\langle \hat{N}, \sigma \right\rangle_z, \quad (5.28)$$

with the angle brackets representing inner products defined through⁶:

$$\begin{aligned} \langle a, b \rangle_z &= \int \frac{dz}{B_0} \frac{1}{\hat{\mathbf{b}} \cdot \nabla z} a b^*, \quad \langle a, b \rangle_{v,\nu} = \sum_\nu \frac{2\pi B_0}{m_\nu} \int d^2v a b^*, \\ \langle a, b \rangle_{z,v,\nu} &= \sum_\nu \int \frac{dz}{B_0} \frac{1}{\hat{\mathbf{b}} \cdot \nabla z} \frac{2\pi B_0}{m_\nu} \int d^2v a b^*. \end{aligned} \quad (5.29)$$

These inner products were chosen because the velocity integrals naturally appear in the field equations, while the spatial integrals correspond to field line averages. An additional benefit of this choice is that it leads to the cancellation of certain terms later in the calculation.

The set of adjoint variables λ_ν , ξ , ζ , and σ have been introduced, with their forms to be determined⁷. The variable λ_ν is identified as the adjoint variable to the distribution function, $\hat{g}_\nu$, whereas ξ , ζ , and σ are adjoint to the field variables (referred to henceforth as adjoint fields). The quantities $\hat{G}_\nu$, $\hat{Q}$, $\hat{M}$, and $\hat{N}$ are the objective functions, and note that, since $\hat{G}_\nu = \hat{Q} = \hat{M} = \hat{N} = 0$ for a consistent set of $\{\mathbf{p}_0, \hat{g}_\nu(\mathbf{p}_0), \hat{\phi}(\mathbf{p}_0), \hat{A}_\parallel(\mathbf{p}_0), \delta\hat{B}_\parallel(\mathbf{p}_0)\}$, $\mathcal{L}|_{\mathbf{p}_0} = 0$ for all choices of adjoint variables. For later convenience the functional operators are decomposed into their components

⁶Note that these must be linear in their arguments, exhibit conjugate symmetry, and satisfy the Cauchy-Schwartz inequality to be well defined. All arguments are required to be square-integrable in order for the optimisation to be finite.

⁷Note that the adjoint variables here are defined in Fourier space, so λ_ν , ξ , ζ , and σ also contain suppressed $\mathbf{k}$ subscripts, and thus the calculation is done for each $\mathbf{k}$ independently. The derivatives $\nabla_{\mathbf{p}}$ indicate how these Fourier components respond to external changes of the parameters $\mathbf{p}$ in the system.

that act on $\hat{g}_\nu$, $\hat{\phi}$, $\hat{A}_\parallel$, and $\delta\hat{B}_\parallel$ separately:

$$\hat{G}_\nu[\mathbf{p}; \hat{g}_\nu, \hat{\phi}, \hat{A}_\parallel, \delta\hat{B}_\parallel] = \hat{G}_{g,\nu}[\mathbf{p}; \hat{g}_\nu] + \hat{G}_{\phi,\nu}[\mathbf{p}; \hat{\phi}] + \hat{G}_{A_\parallel,\nu}[\mathbf{p}; \hat{A}_\parallel] + \hat{G}_{B_\parallel,\nu}[\mathbf{p}; \delta\hat{B}_\parallel], \quad (5.30)$$

$$\hat{Q}[\mathbf{p}; \hat{g}_\nu, \hat{\phi}, \delta\hat{B}_\parallel] = \left\langle \hat{Q}_{g,\nu}[\mathbf{p}; \hat{g}_\nu], \mathbb{I} \right\rangle_{v,\nu} + \hat{Q}_\phi[\mathbf{p}; \hat{\phi}] + \hat{Q}_{B_\parallel}[\mathbf{p}; \delta\hat{B}_\parallel], \quad (5.31)$$

$$\hat{M}[\mathbf{p}; \hat{g}_\nu, \hat{A}_\parallel] = \left\langle \hat{M}_{g,\nu}[\mathbf{p}; \hat{g}_\nu], \mathbb{I} \right\rangle_{v,\nu} + \hat{M}_{A_\parallel}[\mathbf{p}; \hat{A}_\parallel], \quad (5.32)$$

$$\hat{N}[\mathbf{p}; \hat{g}_\nu, \hat{\phi}, \delta\hat{B}_\parallel] = \left\langle \hat{N}_{g,\nu}[\mathbf{p}; \hat{g}_\nu], \mathbb{I} \right\rangle_{v,\nu} + \hat{N}_\phi[\mathbf{p}; \hat{\phi}] + \hat{N}_{B_\parallel}[\mathbf{p}; \delta\hat{B}_\parallel], \quad (5.33)$$

with $\mathbb{I}$ simply equal to one. Explicit expressions for these operators in unnormalised coordinates are given in Appendix B.1.

Next, consider taking the gradient of (5.28) with respect to the variables in the parameter space, $\mathbf{p}$. By isolating all terms multiplying derivatives of $\{\hat{g}_\nu, \hat{\phi}, \hat{A}_\parallel, \delta\hat{B}_\parallel\}$ each of their coefficients can be set to zero—for this the self-adjoint property of collision operators is used [58]. We are at liberty to do this because of the freedom that exists in choosing the adjoint variables introduced in equation (5.28). The gradient of (5.28) is expanded using equations (5.30)–(5.33):

$$\begin{aligned} \nabla_{\mathbf{p}} \mathcal{L} = & \left\langle \nabla_{\mathbf{p}} \hat{g}_\nu, \hat{G}_{g,\nu}^\dagger[\mathbf{p}; \lambda_\nu] + \hat{Q}_{g,\nu}^\dagger[\mathbf{p}; \xi] + \hat{M}_{g,\nu}^\dagger[\mathbf{p}; \zeta] + \hat{N}_{g,\nu}^\dagger[\mathbf{p}; \sigma] \right\rangle_{z,v,\nu} \\ & + \left\langle \nabla_{\mathbf{p}} \hat{\phi}, \left\langle \hat{G}_{\phi,\nu}^\dagger[\mathbf{p}; \lambda_\nu] \right\rangle_{v,\nu} + \hat{Q}_\phi^\dagger[\mathbf{p}; \xi] + \hat{N}_\phi^\dagger[\mathbf{p}; \sigma] \right\rangle_z \\ & + \left\langle \nabla_{\mathbf{p}} \hat{A}_\parallel, \left\langle \hat{G}_{A_\parallel,\nu}^\dagger[\mathbf{p}; \lambda_\nu] \right\rangle_{v,\nu} + \hat{M}_{A_\parallel}^\dagger[\mathbf{p}; \zeta] \right\rangle_z \\ & + \left\langle \nabla_{\mathbf{p}} \delta\hat{B}_\parallel, \left\langle \hat{G}_{B_\parallel,\nu}^\dagger[\mathbf{p}; \lambda_\nu] \right\rangle_{v,\nu} + \hat{Q}_{B_\parallel}^\dagger[\mathbf{p}; \xi] + \hat{N}_{B_\parallel}^\dagger[\mathbf{p}; \sigma] \right\rangle_z \\ & + \mathcal{B}. \end{aligned} \quad (5.34)$$

Here the daggers that appear on the functional operators denote the adjoint of those operators with respect to the relevant inner product. The partial derivative, $\partial_{\mathbf{p}}$, is taken at fixed $\hat{g}_\nu$, $\hat{\phi}$, $\hat{A}_\parallel$, and $\delta\hat{B}_\parallel$. Note that it acts on the inner product itself, in addition to the terms within it, to account for the $\mathbf{p}$ -dependence of the Jacobians⁸ and functional operators. The term ‘ $\mathcal{B}$ ’ accounts for boundary terms that arise from the integration by parts to invert operators such as $\langle \nabla_{\mathbf{p}} \partial_z \hat{g}_\nu, \lambda_\nu \rangle_z$ onto the adjoint

⁸It is noteworthy to point out that although the Jacobians present in the integrals (as well as the Lagrange multipliers themselves) are $\mathbf{p}$ -dependent, it is anticipated that when the distribution function and fields are evaluated at $\mathbf{p}_0$ their derivatives provide zero contribution as they multiply the functional operators, which are identically zero at $\mathbf{p}_0$. Hence, it would be justified to pull the partial derivative through these inner products, and the final derivative of the growth rate obtained should be the same.

variables, producing terms of the form $\langle \nabla_{\mathbf{p}} \hat{g}_\nu, \partial_z \lambda_\nu \rangle_z$, and is given by

$$\begin{aligned} \mathcal{B} = & \sum_{\nu} \frac{2\pi B_0}{m_\nu} \int d^2v \, v_{\parallel} \, \lambda_{\nu}^* \left[\nabla_{\mathbf{p}} \hat{g}_\nu + \frac{Z_\nu e}{T_\nu} J_{0,\nu} F_{0,\nu} \nabla_{\mathbf{p}} \hat{\phi} \right. \\ & \left. - 2 \frac{Z_\nu e}{T_\nu} v_{\parallel} J_{0,\nu} F_{0,\nu} \nabla_{\mathbf{p}} \hat{A}_{\parallel} + 4\mu_\nu \frac{J_{0,\nu}}{a_\nu} F_{0,\nu} \nabla_{\mathbf{p}} \delta \hat{B}_{\parallel} \right] \Big|_{z=-\infty}^{z=\infty} \\ & - \sum_{\nu} \frac{2\pi B_0}{m_\nu} \int dz \int d\mu \, \mu_\nu \frac{\partial B_0}{\partial z} \lambda_{\nu}^* \nabla_{\mathbf{p}} \hat{g}_\nu \Big|_{v_{\parallel}=-\infty}^{v_{\parallel}=\infty}. \end{aligned} \quad (5.35)$$

Setting this term to zero determines the boundary conditions for the adjoint variables. The incoming boundary conditions along the magnetic field on $\hat{g}_\nu$ are taken to be $\hat{g}_\nu(z \rightarrow -\infty, v_{\parallel} > 0, \mu_\nu) \rightarrow 0$, and $\hat{g}_\nu(z \rightarrow \infty, v_{\parallel} < 0, \mu_\nu) \rightarrow 0$, independently of $\mathbf{p}$, such that $d_{\mathbf{p}} \hat{g}_\nu = 0$ at these limits. The boundary condition $\lambda_{\nu}^*(z \rightarrow -\infty, v_{\parallel} < 0, \mu_\nu) \rightarrow 0$, and $\lambda_{\nu}^*(z \rightarrow \infty, v_{\parallel} > 0, \mu_\nu) \rightarrow 0$ is imposed on λ_{ν}^* in order to eliminate the boundary term arising from the z integration by parts, and hence remove the need to calculate $\nabla_{\mathbf{p}} \{\hat{g}_\nu, \hat{\phi}, \hat{A}_{\parallel}, \delta \hat{B}_{\parallel}\}$ at the boundaries in z . The boundary term arising from integration by parts in $v_{\parallel}$ is automatically satisfied as it is assumed that $\hat{g}_\nu(z, v_{\parallel} \rightarrow \pm\infty, \mu_\nu) = 0$, $\forall \{z, \mu_\nu\}$ independently of $\mathbf{p}$. However it is convenient to impose that $\lambda_{\nu}^*(z, v_{\parallel} \rightarrow \pm\infty, \mu_\nu) = 0$ such that λ_{ν}^* and $\hat{g}_\nu$ satisfy similar boundary conditions, whilst also ensuring λ_{ν} is sensibly defined and normalisable⁹. The substitution $\lambda_{\nu}^{\leftrightarrow,*} = \lambda_{\nu}^*(z, -v_{\parallel}, \mu_\nu)$ is made such that the adjoint equations more closely resemble those in the original gyrokinetic system. This redefines the z -boundary condition on the adjoint variable $\lambda_{\nu}^{\leftrightarrow,*}(z \rightarrow -\infty, v_{\parallel} > 0, \mu_\nu) \rightarrow 0$, and $\lambda_{\nu}^{\leftrightarrow,*}(z \rightarrow \infty, v_{\parallel} < 0, \mu_\nu) \rightarrow 0$, which now mirrors those satisfied by $\hat{g}_\nu$.

Setting the remaining coefficients of $\nabla_{\mathbf{p}} \hat{g}_\nu$, $\nabla_{\mathbf{p}} \hat{\phi}$, $\nabla_{\mathbf{p}} \hat{A}_{\parallel}$, and $\nabla_{\mathbf{p}} \delta \hat{B}_{\parallel}$ in equation (5.34) equal to zero yields the constraint equations for the adjoint variables

$$\hat{G}_{g,\nu}^\dagger[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] + \hat{Q}_{g,\nu}^\dagger[\mathbf{p}; \xi] + \hat{M}_{g,\nu}^\dagger[\mathbf{p}; \zeta] + \hat{N}_{g,\nu}^\dagger[\mathbf{p}; \sigma] = 0, \quad (5.36)$$

$$\left\langle \hat{G}_{\phi,\nu}^\dagger[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] \right\rangle_{v,\nu} + \hat{Q}_{\phi}^\dagger[\mathbf{p}; \xi] + \hat{N}_{\phi}^\dagger[\mathbf{p}; \sigma] = 0, \quad (5.37)$$

$$\langle \hat{G}_{A_{\parallel},\nu}^\dagger[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] \rangle_{v,\nu} + \hat{M}_{A_{\parallel}}^\dagger[\mathbf{p}; \zeta] = 0, \quad (5.38)$$

$$\left\langle \hat{G}_{B_{\parallel},\nu}^\dagger[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] \right\rangle_{v,\nu} + \hat{Q}_{B_{\parallel}}^\dagger[\mathbf{p}; \xi] + \hat{N}_{B_{\parallel}}^\dagger[\mathbf{p}; \sigma] = 0. \quad (5.39)$$

The expressions for these adjoint operators are stated in Appendix B.2. The z -derivatives which appear in the $\hat{G}_\nu$ operators, under the velocity integrals, in equations (5.37)-(5.39) pose a potential difficulty; to calculate the adjoint fields, information

⁹An additional consequence is that this choice simplifies the implementation into an existing gyrokinetic code.

is required for $\lambda_\nu^{\leftrightarrow}$ at all z , for both positive and negative velocities. Given the boundary conditions on λ_ν and the propagation of information by advection this information is not readily available. This is akin to the problem faced when solving the gyrokinetic system; an incoming boundary condition is imposed on the distribution function at $z \rightarrow \pm\infty$ when $v_\parallel < 0$ and $v_\parallel > 0$ respectively. The outgoing boundary information is not known a priori but must instead be solved for. Anticipating the numerical difficulties this will create, moments of (5.36) are taken to simplify the adjoint equations. This brings them into a form more closely resembling that of the gyrokinetic field equations. It should be emphasised that a fully kinetic treatment is retained in doing so.

A summary of this calculation can be found in Appendix B.3, and the result after algebraic manipulation is:

$$\begin{aligned} \gamma^* \lambda_\nu^{\leftrightarrow} + v_\parallel \hat{\mathbf{b}} \cdot \nabla z \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z} - \frac{\mu_\nu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial v_\parallel} - i\omega_{d,\nu} \lambda_\nu^{\leftrightarrow} + Z_\nu e J_{0,\nu} \xi \\ - \frac{4\pi}{k_\perp^2} Z_\nu e J_{0,\nu} \frac{v_\parallel}{c} \zeta + 8\pi \frac{J_{1,\nu}}{a_\nu} \mu_\nu \sigma - \hat{C}_\nu[\lambda_\nu^{\leftrightarrow}] = 0, \end{aligned} \quad (5.40)$$

$$\xi + \frac{1}{\eta} \sum_\nu \frac{2\pi B_0}{m_\nu} \int d^2v \left[i\omega_{*,\nu} + \frac{Z_\nu e}{T_\nu} \gamma^* \right] J_{0,\nu} F_{0,\nu} \lambda_\nu^{\leftrightarrow} = 0, \quad (5.41)$$

$$\zeta - \frac{1}{k_\perp^2} \sum_\nu \frac{2\pi B_0}{m_\nu} \int d^2v \frac{v_\parallel}{c} \left[i\omega_{*,\nu} + \frac{Z_\nu e}{T_\nu} \gamma^* \right] J_{0,\nu} F_{0,\nu} \lambda_\nu^{\leftrightarrow} = 0, \quad (5.42)$$

$$\sigma - \sum_\nu \frac{2\pi B_0}{m_\nu} \int d^2v \left(2 \frac{\mu_\nu}{Z_\nu e} \frac{J_{1,\nu}}{a_\nu} \right) \left[i\omega_{*,\nu} + \frac{Z_\nu e}{T_\nu} \gamma^* \right] F_{0,\nu} \lambda_\nu^{\leftrightarrow} = 0, \quad (5.43)$$

with $\eta = \sum_\nu (Z_\nu e)^2 n_\nu / T_\nu$. Noting that one can also rewrite equation (5.27) as $\hat{G}_\nu[\mathbf{p}; \hat{g}_\nu, \hat{\phi}, \hat{A}_\parallel, \delta \hat{B}_\parallel] = \gamma \hat{g}_\nu + \hat{L}_\nu[\mathbf{p}; \hat{g}_\nu, \hat{\phi}, \hat{A}_\parallel, \delta \hat{B}_\parallel]$, and using $\nabla_{\mathbf{p}} \mathcal{L}|_{\mathbf{p}_0} = 0$, the equation (5.34) can be rearranged to obtain

$$\nabla_{\mathbf{p}} \gamma \langle \hat{g}_\nu, \lambda_\nu \rangle_{z,v,\nu} = - \left[\langle \partial_{\mathbf{p}} \hat{L}_\nu, \lambda_\nu \rangle_{z,v,\nu} + \langle \partial_{\mathbf{p}} \hat{Q}, \xi \rangle_z + \langle \partial_{\mathbf{p}} \hat{M}, \zeta \rangle_z + \langle \partial_{\mathbf{p}} \hat{N}, \sigma \rangle_z \right] \Big|_{\mathbf{p}_0}, \quad (5.44)$$

where the partial derivatives have been pulled inside the inner products as the contribution arising from the Jacobian derivatives is zero by virtue of $\hat{G}_\nu(\mathbf{p}_0) = \hat{Q}(\mathbf{p}_0) = \hat{M}(\mathbf{p}_0) = \hat{N}(\mathbf{p}_0) = 0$.

To solve for the derivative of the linear growth rate, the following procedure is taken: The gyrokinetic equation, (5.27), is solved in combination with the field equations, (2.45)-(2.47), to obtain $\hat{g}_\nu$ and γ , and then equations (5.40)-(5.43) are used to solve for the adjoint variables. These quantities are then all fed into equation (5.44) to compute $\nabla_{\mathbf{p}} \gamma$.

5.5.2 Electrostatic, Collisionless Limit

The numerical tests are performed in the electrostatic, collisionless regime outlined in Section 2.3.3 with $C_{\nu,\nu'} = 0$. The linear, collisionless, electrostatic gyrokinetic equation in the long time limit is:

$$\begin{aligned} \hat{G}_\nu[\mathbf{p}; \hat{g}_\nu, \hat{\phi}] = & \gamma \hat{g}_\nu + v_\parallel \hat{\mathbf{b}} \cdot \nabla z \left[\frac{\partial \hat{g}_\nu}{\partial z} + \frac{Z_\nu e}{T_\nu} \frac{\partial J_{0,\nu} \hat{\phi}}{\partial z} F_{0,\nu} \right] + i\omega_{*,\nu} J_{0,\nu} \hat{\phi} F_{0,\nu} \\ & - \frac{\mu_\nu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{\partial \hat{g}_\nu}{\partial v_\parallel} + i\omega_{d,\nu} \left[\hat{g}_\nu + \frac{Z_\nu e}{T_\nu} J_{0,\nu} \hat{\phi} F_{0,\nu} \right], \end{aligned} \quad (5.45)$$

which is closed by the electrostatic limit of quasineutrality:

$$\hat{Q}[\mathbf{p}; \hat{g}_\nu, \hat{\phi}] = \sum_\nu Z_\nu e \left[\frac{2\pi B_0}{m_\nu} \int d^2v J_{0,\nu} \hat{g}_\nu + \frac{Z_\nu n_\nu}{T_\nu} (\Gamma_{0,\nu} - 1) \hat{\phi} \right], \quad (5.46)$$

with $\hat{G}_\nu$ and $\hat{Q}$ identically zero, and $\hat{M}$, $\hat{N}$ providing no contribution in the electrostatic limit. In the above the long time behaviour of a single wavenumber is being considered, and the associated subscripts have been suppressed.

As in Section 5.5.1 the functional operators $\hat{G}_\nu[\mathbf{p}; \hat{g}_\nu, \hat{\phi}]$, and $\hat{Q}[\mathbf{p}; \hat{g}_\nu, \hat{\phi}]$ are decomposed into components that act on $\hat{g}_\nu$ and $\hat{\phi}$ separately, with all other operators in (5.30)-(5.33) set to zero. The derivation in Section 5.5.1 is unchanged, with the exception that some terms may now be omitted. The resulting derivative of the growth rate in the electrostatic, collisionless regime is:

$$\nabla_{\mathbf{p}} \gamma \langle \hat{g}_\nu, \lambda_\nu \rangle_{z,v,\nu} = - \left[\langle \partial_{\mathbf{p}} \hat{L}_\nu, \lambda_\nu \rangle_{z,v,\nu} + \langle \partial_{\mathbf{p}} \hat{Q}, \xi \rangle_z \right] \Big|_{\mathbf{p}_0}, \quad (5.47)$$

where $\hat{L}_\nu = \hat{G}_\nu - \gamma \hat{g}_\nu$ is given by equation (5.45), and the adjoint equations are

$$\gamma^* \lambda_\nu^{\leftrightarrow} + v_\parallel \hat{\mathbf{b}} \cdot \nabla z \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z} - \frac{\mu_\nu}{m_s} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial v_\parallel} - i\omega_{d,\nu} \lambda_\nu^{\leftrightarrow} + Z_\nu e J_{0,\nu} \xi = 0, \quad (5.48)$$

$$\xi + \frac{1}{\eta} \sum_\nu \frac{2\pi B_0}{m_\nu} \int d^2v \left[i\omega_{*,\nu} + \frac{Z_\nu e}{T_\nu} \gamma^* \right] J_{0,\nu} F_{0,\nu} \lambda_\nu^{\leftrightarrow} = 0, \quad (5.49)$$

with $\lambda_\nu^{\leftrightarrow}(v_\parallel) = \lambda_\nu(-v_\parallel)$, and $\eta = \sum_\nu (Z_\nu e)^2 n_\nu / T_\nu$ as before.

5.6 Normalisations and Magnetic Geometry

To evaluate (5.47) the set of variables $\{\gamma, \hat{g}_\nu, \hat{\phi}, \lambda_\nu^{\leftrightarrow}, \xi\}$ must be solved for, and evaluated at the unperturbed geometric values, $\mathbf{p}_0$. This is done by implementing the

adjoint system within the local flux-tube δf -gyrokinetic code **stella** [9] the equations, normalisations and algorithms for which are outlined in Sections 3.3-3.4. The normalisations outlined in Section 3.3 are used, with the list of normalised parameters given in Table 3.1 and Table 3.2. It has been previously shown that shaping can affect the microstability of a plasma in an MCF device [13], hence $\mathbf{p}$ will be chosen to vary the magnetic geometry. The specific choice for $\mathbf{p}$ is taken to be the Miller parametrisation of the equilibrium outlined in Section 2.2.1.2, and the application of the adjoint method is detailed in this case. For this section we reinstate the tilde notation to denote normalised variables.

5.6.1 Gyrokinetic Normalisations

Here the normalised gyrokinetic equation and field equations are presented for the long time limit. The reader is reminded that the subscripts denoting wavenumber, $\mathbf{k}$, have been suppressed, as these can be considered independently of one another. All variables are taken to the left hand side to define objective functions, $\hat{G}_\nu$, $\hat{Q}$, $\hat{M}$, and $\hat{N}$ that are identically zero. The normalised low-flow, electromagnetic gyrokinetic equation taken in the long time limit is

$$\begin{aligned} \hat{G}_\nu = & \tilde{\gamma} \tilde{g}_\nu + \tilde{v}_{\text{th},\nu} \tilde{v}_\parallel \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z} \left[\frac{\partial \tilde{g}_\nu}{\partial \tilde{z}} + \frac{Z_\nu}{\tilde{T}_\nu} \frac{\partial \langle \tilde{\chi} \rangle_\nu}{\partial \tilde{z}} e^{-\tilde{v}_\nu^2} \right] \\ & + i \tilde{\omega}_{*,\nu} e^{-\tilde{v}_\nu^2} \langle \tilde{\chi} \rangle_\nu + i \tilde{\omega}_{d,\nu} \left[\tilde{g}_\nu + \frac{Z_\nu}{\tilde{T}_\nu} \langle \tilde{\chi} \rangle_{\mathbf{k},\nu} e^{-\tilde{v}_\nu^2} \right] \\ & - \tilde{v}_{\text{th},\nu} \tilde{\mu}_\nu \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{B}_0 \frac{\partial \tilde{g}_\nu}{\partial \tilde{v}_\parallel} + 2 \frac{Z_\nu}{\tilde{m}_\nu} \tilde{\mu}_\nu \hat{\mathbf{b}} \cdot \tilde{B}_0 e^{-\tilde{v}_\nu^2} J_{0,\mathbf{k},\nu} \tilde{A}_\parallel - \hat{C}_\nu[\{\tilde{g}_{\nu'}, \tilde{\chi}_{\nu'}\}], \end{aligned} \quad (5.50)$$

The corresponding transformed, normalised field equations are given by equations (3.10)-(3.12).

5.6.2 Adjoint Normalisation

The normalisation of the adjoint variables is chosen in such a way that the optimisation Lagrangian is dimensionless. In general this is achieved by enforcing that the dimension of the adjoint variables in Section 5.5 satisfy:

$$[\lambda_\nu] = [\hat{G}_\nu]^{-1}, \quad [\xi] = [\hat{Q}]^{-1}, \quad [\zeta] = [\hat{M}]^{-1}, \quad [\sigma] = [\hat{N}]^{-1}, \quad (5.51)$$

with $[A]$ denoting the dimensionality of A . The normalised adjoint variables should then satisfy:

$$\tilde{\lambda}_\nu = \frac{\lambda_\nu}{[\lambda_\nu]}, \quad \tilde{\xi} = \frac{\xi}{[\xi]}, \quad \tilde{\zeta} = \frac{\zeta}{[\zeta]}, \quad \tilde{\sigma} = \frac{\sigma}{[\sigma]}. \quad (5.52)$$

From considering the dimensions of the distribution functions, fields, and the defined inner products, the adjoint normalisation follows:

$$\tilde{\lambda}_\nu = \lambda_\nu \frac{F_{0,\nu}}{e^{-\tilde{v}_\nu^2}} \frac{\rho_r v_{\text{th},r}}{L_r^2}, \quad \tilde{\xi} = \xi \frac{n_r e \rho_r}{L_r}, \quad \tilde{\zeta} = \zeta \frac{B_r \rho_r^2}{L_r}, \quad \tilde{\sigma} = \sigma \frac{B_r \rho_r}{L_r}. \quad (5.53)$$

The adjoint equation, (5.40) as written in terms of normalised coordinates, is multiplied by a factor of $(\rho_r/L_r)F_{0,\nu}/e^{-\tilde{v}_\nu^2}$ to obtain the electromagnetic, collisional adjoint equations in normalised units:

$$\begin{aligned} \frac{\partial \tilde{\lambda}_\nu^{\leftrightarrow}}{\partial \tilde{t}} + \tilde{\gamma}^* \tilde{\lambda}_\nu^{\leftrightarrow} + \tilde{v}_{\text{th},\nu} \tilde{v}_\parallel \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z} \frac{\partial \tilde{\lambda}_\nu^{\leftrightarrow}}{\partial \tilde{z}} - \tilde{v}_{\text{th},\nu} \tilde{\mu}_\nu \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{B}_0 \frac{\partial \tilde{\lambda}_\nu^{\leftrightarrow}}{\partial \tilde{v}_\parallel} - i \tilde{\omega}_{d,\nu} \tilde{\lambda}_\nu^{\leftrightarrow} \\ + Z_\nu \tilde{n}_\nu J_{0,\nu} \tilde{\xi} - \frac{\beta_r}{(k_\perp \rho_r)^2} Z_\nu \tilde{n}_\nu \tilde{v}_{\text{th},\nu} J_{0,\nu} \tilde{v}_\parallel \tilde{\zeta} + 2\beta_r \tilde{T}_\nu \tilde{\mu}_\nu \frac{J_{1,\nu}}{\tilde{a}_\nu} \tilde{\sigma} - \hat{C}_\nu[\lambda_\nu] = 0, \end{aligned} \quad (5.54)$$

$$\tilde{\xi} + \frac{1}{\tilde{\eta}} \sum_\nu \frac{2\tilde{B}_0}{\sqrt{\pi}} \int d^2 \tilde{v} \left[i \tilde{\omega}_{*,\nu} + \frac{Z_\nu}{\tilde{T}_\nu} \tilde{\gamma}^* \right] J_{0,\nu} e^{-\tilde{v}_\nu^2} \tilde{\lambda}_\nu^{\leftrightarrow} = 0, \quad (5.55)$$

$$\tilde{\zeta} - \sum_\nu \frac{2\tilde{B}_0}{\sqrt{\pi}} \int d^2 \tilde{v} (2\tilde{v}_{\text{th},\nu} \tilde{v}_\parallel) \left[i \tilde{\omega}_{*,\nu} + \frac{Z_\nu}{\tilde{T}_\nu} \tilde{\gamma}^* \right] J_{0,\nu} e^{-\tilde{v}_\nu^2} \tilde{\lambda}_\nu^{\leftrightarrow} = 0, \quad (5.56)$$

$$\tilde{\sigma} - \sum_\nu \frac{2\tilde{B}_0}{\sqrt{\pi}} \int d^2 \tilde{v} \left(4\tilde{\mu}_\nu \frac{\tilde{T}_\nu}{Z_\nu} \frac{J_{1,\nu}}{\tilde{a}_\nu} \right) \left[i \tilde{\omega}_{*,\nu} + \frac{Z_\nu}{\tilde{T}_\nu} \tilde{\gamma}^* \right] e^{-\tilde{v}_\nu^2} \tilde{\lambda}_\nu^{\leftrightarrow} = 0. \quad (5.57)$$

The inclusion of an artificial time derivative in equation (5.54) is a practical numerical strategy, and the steady state solution for $\tilde{\lambda}_\nu$ is obtained when $\partial_{\tilde{t}} \tilde{\lambda}_\nu = 0$. This approach transforms the (potentially stiff or non-invertible) stationary problem into a well-posed relaxation system that can be marched forward in pseudo-time until convergence. This choice is primarily motivated by our use of an initial value code, with the absence of a dedicated eigenvalue solver for direct computation of the stationary solution. While it is true that introducing an artificial time derivative can theoretically lead to non-convergent behavior, such as periodic orbits, rather than asymptoting to a stationary solution, for the class of linear stability and adjoint problems considered here, this method has consistently and reliably converged to the correct steady-state solutions in practice.

As before, one can use $\hat{G}_\nu[\mathbf{p}; \tilde{g}_\nu, \tilde{\phi}, \tilde{A}_\parallel, \delta \tilde{B}_\parallel] = \tilde{\gamma} \tilde{g}_\nu + \hat{L}_\nu[\mathbf{p}; \tilde{g}_\nu, \tilde{\phi}, \tilde{A}_\parallel, \delta \tilde{B}_\parallel]$, with the constraint $\nabla_{\mathbf{p}} \mathcal{L}|_{\mathbf{p}_0} = 0$ to rearrange the above and obtain:

$$\nabla_{\mathbf{p}} \tilde{\gamma} \langle \tilde{g}_\nu, \tilde{\lambda}_\nu \rangle_{\tilde{z}, \tilde{v}, \nu} = - \left[\langle \partial_{\mathbf{p}} \hat{L}_\nu, \tilde{\lambda}_\nu \rangle_{\tilde{z}, \tilde{v}, \nu} + \langle \partial_{\mathbf{p}} \hat{Q}, \tilde{\xi} \rangle_{\tilde{z}} + \langle \partial_{\mathbf{p}} \hat{M}, \tilde{\zeta} \rangle_{\tilde{z}} + \langle \partial_{\mathbf{p}} \hat{N}, \tilde{\sigma} \rangle_{\tilde{z}} \right] \Big|_{\mathbf{p}_0}, \quad (5.58)$$

with closure equations provided by (5.54)-(5.57). As in previous chapters, the tilde notation denoting normalised quantities will henceforth be dropped for the remainder of this chapter, unless stated otherwise.

5.6.3 Electrostatic Collisionless Limit

Finally, in the electrostatic, collisionless regime, the system of equations to solve in the normalised **stella** coordinates is given by

$$\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} + \gamma^* \lambda_\nu^{\leftrightarrow} + v_{\text{th},\nu} v_\parallel \hat{\mathbf{b}} \cdot \nabla_z \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z} - v_{\text{th},\nu} \mu_\nu \hat{\mathbf{b}} \cdot \nabla B_0 \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial v_\parallel} - i \omega_{d,\nu} \lambda_\nu^{\leftrightarrow} + Z_\nu n_\nu J_{0,\nu} e^{-v_\nu^2} \xi = 0, \quad (5.59)$$

$$\xi + \frac{1}{\eta} \sum_\nu \frac{2B_0}{\sqrt{\pi}} \int d^2v \left[i \omega_{*,\nu} + \frac{Z_\nu}{T_\nu} \gamma^* \right] J_{0,\nu} e^{-v_\nu^2} \lambda_\nu^{\leftrightarrow} = 0, \quad (5.60)$$

and

$$\nabla_{\mathbf{p}} \gamma \langle \hat{g}_\nu, \lambda_\nu \rangle_{z,v,\nu} = - \left[\langle \partial_{\mathbf{p}} \hat{L}_\nu, \lambda_\nu \rangle_{z,v,\nu} + \langle \partial_{\mathbf{p}} \hat{Q}, \xi \rangle_z \right] \Big|_{\mathbf{p}_0}. \quad (5.61)$$

5.6.4 Miller Geometry

The coefficients in equations (5.50)-(3.12), and thus the associated linear growth rates, are implicitly dependent on the magnetic geometry. For the remainder of the chapter $\mathbf{p}$ will be taken to be an appropriate set of parameters that specifies the local magnetic geometry. In particular, the Miller formalism [86] will be used to parameterise the magnetic field on the flux surface of interest, as outlined in Section 2.2.1.2.

The user specified input parameters used in the local version of **stella** are $\{r_{\psi_0}, R_{\psi_0}, \Delta_{\psi_0}, q_{\psi_0}, \hat{s}_{\psi_0}, \kappa_{\psi_0}, \kappa'_{\psi_0}, R_{\text{geo},\psi_0}, \delta_{\psi_0}, \delta'_{\psi_0}, \beta'_{\psi_0}\}$, which are the Miller parameters defined in Section 2.2.1.2 evaluated at the flux surface ψ_0 . Note that **stella** defines the triangularity as $\delta \doteq \arcsin(\bar{\delta})$. Here $R_{\text{geo},\psi_0} \equiv I(\psi_0)/B_r$ sets the reference magnetic field. Further information about how the Miller geometry is treated in **stella** is given in Appendix B.4. Henceforth the parameters $\mathbf{p}$ will be set as $\mathbf{p} := \{r_{\psi_0}, R_{\psi_0}, \Delta_{\psi_0}, q_{\psi_0}, \hat{s}_{\psi_0}, \kappa_{\psi_0}, \kappa'_{\psi_0}, R_{\text{geo},\psi_0}, \delta_{\psi_0}, \delta'_{\psi_0}, \beta'_{\psi_0}\}$. The subscript ψ_0 shall be dropped, and everywhere these Miller parameters appear it shall be assumed that they are evaluated at a given flux location, ψ_0 .

5.7 Numerical Implementation

As discussed, the aim when applying the adjoint method to a system unstable to ITG is to find the magnetic geometry that minimises the linear growth rate for the

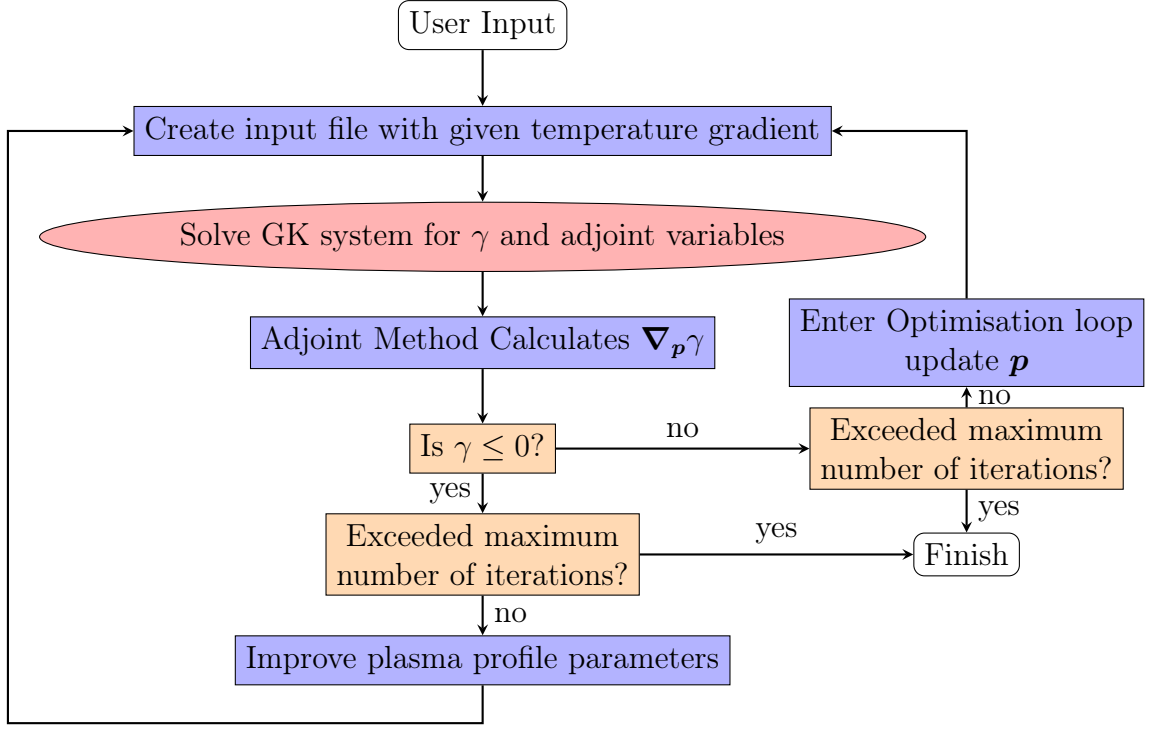


Figure 5.2: Flow chart showing a practical implementation of the adjoint method, to optimise for a plasma profile parameter by modifying the geometry coefficients, $\mathbf{p}$.

ITG instability, whilst maximising the linear critical temperature gradient across the device. This requires three distinct stages: first, the computation of $\nabla_{\mathbf{p}}\gamma$ at a fixed ion temperature gradient, T'_i ; second, its subsequent use in an optimisation algorithm to find the $\mathbf{p}$ that minimises γ for this given T'_i ; third, iteration of this procedure with variable T'_i to find the maximum temperature gradient for which $\gamma \leq 0$ for the range of $\mathbf{p}$ considered.

5.7.1 Initial Simulation

Solving, at an initial set of $\mathbf{p}_0$, for γ , $\hat{g}_\nu$ and $\hat{\phi}$ is achieved by performing a standard **stella** run for a sufficiently long time, such that the solution is dominated by a single normal mode. To determine the time at which this is satisfied, a convergence test can be employed: the growth rate is calculated at each time step and if the value of this is constant in time (within a specified tolerance) then the system is deemed to be converged. There are two components of the convergence test. The first is to check that the growth rates calculated at adjacent time steps are within a given tolerance of each other. The second is to perform a windowed average to check that the growth

rate remains consistent over a defined number of time steps. Two windowed averages are done; one over N_t time steps and one over $\mathbb{Z}(N_t/2)$ time steps. When these two window-averaged growth rates agree within a set tolerance, $\hat{g}_\nu$ and $\hat{\phi}$ are taken to have converged. The corresponding growth rate is then calculated from the windowed average¹⁰.

5.7.2 Adjoint Simulation

As previously introduced, an artificial time dependence is added to the adjoint equations to facilitate computation. The solution is found in the steady state limit, in which the time derivative appearing in the adjoint equations goes to zero.

The adjoint equations are treated in a similar way to the treatment of the usual gyrokinetic system of equations in **stella**. The main aim is to ensure that the parallel streaming term may be treated separately from the rest of the dynamics through operator splitting. This is done by discretising in time, and splitting the time derivative into a series of three steps:

$$\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} = \left(\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} \right)_1 + \left(\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} \right)_2 + \left(\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} \right)_3, \quad (5.62)$$

where

$$\left(\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} \right)_1 = -\gamma^* \lambda_\nu^{\leftrightarrow} + i\omega_{d,\nu} \lambda_\nu^{\leftrightarrow} - Z_\nu n_\nu \xi, \quad (5.62a)$$

$$\left(\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} \right)_2 = v_{\text{th},\nu} \mu_\nu \hat{\mathbf{b}} \cdot \nabla B_0 \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial v_\parallel}, \quad (5.62b)$$

$$\left(\frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial t} \right)_3 = -v_{\text{th},\nu} v_\parallel \hat{\mathbf{b}} \cdot \nabla z \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z}. \quad (5.62c)$$

Analogous to **stella**, the terms in (5.62a) are treated explicitly using a strong-stability-preserving, third order Runge–Kutta method [47]. The Runge–Kutta algorithm is second order accurate in time step size, Δt , whilst the operator splitting is only first order accurate in Δt .

The parallel streaming and mirror terms, given by equations (5.62c) and (5.62b) respectively, are treated separately, due to the presence of the prefactor $v_{\text{th},\nu}$, which

¹⁰In cases where there exist two modes of similar growth rates, it may take a long run time to isolate the dominant mode. In these instances of ambiguity, the adjoint method as demonstrated here may fail to provide the correct gradient of the linear growth rate in the desired parameter space at that point. However, it should be noted that one can move away from this point of ambiguity in the growth rate by applying a random kick in the parameter space, and as soon as you move away from this region the adjoint method can continue to be applicable, and will provide the correct gradient in the parameter space. Hence, even if the random step is taken in the wrong direction within the space, the adjoint method should correct itself to then step in the direction of stability.

increases the relative amplitude of these terms when considering electron dynamics. As a result these terms have the potential to exert a stringent Courant–Friedrichs–Lewy (CFL) condition on the simulation, and require a small time step to be taken in order to retain accuracy. Thus, these terms are treated implicitly in time to circumvent this constraint.

The mirror term, (5.62b), is a simple advection equation of $\lambda_\nu^{\leftrightarrow}$ in $v_{\parallel}$, which is treated using a semi-Lagrange method¹¹, akin to the algorithm used to advect the distribution function in $v_{\parallel}$ within **stella**. The mirror coefficients are independent of both time and $v_{\parallel}$ and hence the exact characteristics of this equation are known. The interpolation in $v_{\parallel}$ is fourth order accurate in $v_{\parallel}$ grid spacing, $\Delta v_{\parallel}$ [9].

The streaming term, (5.62c), is also an advection equation in z , which is treated using the Thompson algorithm for tri-diagonal solve. When fully centred in time the discretisation reduces to the Crank–Nicolson method, which is second order accurate in time and z grid spacing, Δt and Δz , respectively.

We are seeking a steady state solution to the adjoint equations. The same convergence test is performed on $\lambda_\nu^{\leftrightarrow}$ as that performed on the distribution function, in order to check that the complex growth rate has converged to zero within a given tolerance. When this is satisfied, the resulting $\lambda_\nu^{\leftrightarrow}$ is used to solve for λ_ν . This is then stored for use in the remainder of the calculation.

5.7.3 Optimisation Loop

Once the gradients $\nabla_{\mathbf{p}}\gamma$ are obtained they can be used inside an optimisation loop to find the $\mathbf{p}$ that minimises γ . The Levenberg-Marquardt (LM) algorithm [119] is employed to find the local minimum. This method adopts a steepest descent behaviour when the location in parameter space is considered to be far from the minimum, and progresses towards Gauss-Newton behaviour as one approaches the minimum. This is achieved by introducing a damping factor, Γ , which is updated with each iteration. The Levenberg-Marquardt algorithm is designed specifically to solve nonlinear least squares problems. Consider the least squares problem to find parameters $\mathbf{p}$ that minimise the squared residual:

$$\min_{\mathbf{p}} \frac{1}{2} \|\gamma(\mathbf{p})\|^2 = \frac{1}{2} \sum_{i=1}^m \gamma_i(\mathbf{p})^2. \quad (5.63)$$

¹¹This is an option within the FT version of **stella**. This has not been detailed in this thesis as it is not readily extendable to the FA formalism.

The algorithm does this by iteratively solving the following:

$$[\mathbf{H} + \Gamma \text{diag}(\mathbf{H})]d\mathbf{p} = -\nabla_{\mathbf{p}}\gamma, \quad (5.64)$$

with $\mathbf{H} = \nabla_{\mathbf{p}}^2\gamma \approx (\nabla_{\mathbf{p}}\gamma)^\dagger \nabla_{\mathbf{p}}\gamma$ the Hessian matrix. When Γ is large $\mathbf{p}_{\text{new}} \approx \mathbf{p}_{\text{old}} - \boldsymbol{\alpha} \cdot \nabla_{\mathbf{p}}\gamma$, with $\boldsymbol{\alpha} \doteq [\Gamma \text{diag}(\mathbf{H})]^{-1}$, which mimics the gradient descent algorithm. However, when Γ is small (5.64) reduces to $\mathbf{p}_{\text{new}} \approx \mathbf{p}_{\text{old}} - \mathbf{H}^\dagger \cdot \nabla_{\mathbf{p}}\gamma$, which matches with the Gauss-Newton algorithm.

The LM formalism is derived using the Taylor expansion, and as such a trust region is included within the optimisation loop to ensure that the updated value of $\mathbf{p}$ is close enough to the previous value, such that the Taylor approximation is valid within the limits for which the algorithm is applied. The trust region for $\mathbf{p}$ is defined via:

$$T = \frac{0.5 d\mathbf{p}^\dagger \cdot \mathbf{H} \cdot d\mathbf{p}}{d\mathbf{p} \cdot \nabla_{\mathbf{p}}\gamma} < \bar{\epsilon}, \quad (5.65)$$

where $\bar{\epsilon}$ is a chosen tolerance. If $T > \bar{\epsilon}$ for a given $d\mathbf{p}$ the algorithm rejects the output $\mathbf{p}_{\text{new}}$ and increases the weight Γ , in an attempt to improve the accuracy of the approximation. This helps ensure that the updated value of $\mathbf{p}$ is a reasonable one. It is worth noting that the LM algorithm is designed to find local minima, so there is no guarantee that the minimum obtained is the global minimum of the system.

It should be emphasised that the gradient-based optimisation algorithm is independent of the adjoint method that has been developed for gyrokinetic microstability. The optimisation loop may be itself optimised to efficiently search for regions of stability given a gradient input. Different algorithms, and indeed different parameter choices within each algorithm, will yield different efficiencies in finding stable solutions. An illustrative example of this is given later in Figure 5.3, where the step size for the optimisation loop is varied to yield two distinct paths through the parameter space using the adjoint gradient. However, this external optimisation loop incurred negligible run time cost, so will not be included in the focus of this thesis and we will not labour on enhancing the gradient-based optimisation loop.

5.8 Numerical Results

5.8.1 Initial Benchmark

The first numerical check performed was to ensure that the values of $d_{\mathbf{p}}\gamma$ obtained from the adjoint method agree with those obtained using a finite difference approach.

Following this a more extensive benchmark can be performed by conducting a parameter scan in the growth rate using `stella` for different values of the Miller parameters. An initial set of parameters is chosen and the adjoint-optimisation scheme described above is performed. The results are overlaid on those from the parameter scan for their comparison. As a proof of principle two parameters: triangularity, δ , and elongation, κ , are chosen to be varied, whilst holding the other Miller variables fixed. Given that the Miller parametrisation is local to a given flux surface, this variation is not necessarily consistent with a global solution to the Grad–Shafranov equation. The choice to vary these two parameters in isolation is driven by two considerations; first, it is only intended as a proof-of-principle check of the adjoint approach so simplicity is desirable. Second, previous research has shown that maximal shaping, with large elongation and triangularity, minimises the linear ITG instability. In contrast, the effect on the growth rate from changing parameters such as κ' and δ' is small, ordered with the inverse aspect ratio, $a/R_0 \ll 1$, where a and R_0 are the scales associated with the minor and major radii respectively [57]. Thus the that the linear growth rates are expected to be most sensitive to elongation and triangularity. Table 5.1 lists the values of input equilibrium variables in the Miller geometry. These have been chosen to coincide with values used in Beeke (2020) [10] in order to verify the qualitative behaviour found.

Given these initial values of δ and κ , the linear growth rates are determined for a grid of perpendicular wavenumbers within $k_x < 2.0$, $k_y < 2.0$, which reveals that the most unstable mode is found at $\{k_y, k_x\} = \{0.68, 0.0\}$ for mass ratio $m_e/m_i = 2.7 \times 10^{-4}$, normalised species temperatures and densities of $T_i = T_e = 1$, $n_i = n_e = 1$, and normalised species temperature and density gradients of $L_r/L_{T_i} = L_r/L_{T_e} = 2.42$, $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$.

Figure 5.3 shows a scan in the linear growth in elongation and triangularity obtained by running `stella` with the Miller parameters specified in Table 5.1, and the values of κ and δ adjusted accordingly. The contour colour indicates the magnitude of the growth rate, and the plot extends over a range of values that has been set by reasonable physical constraints on MCF devices.

Figure 5.3 shows that increasing the elongation and triangularity of the flux surface reduces the linear growth rate, and that there exists a region of stability when the shaping is maximal, in agreement with previous work. The path taken using the adjoint method is indicated in white. At the chosen starting point, located in the region of instability, the gradient $d_{\mathbf{p}}\gamma$ is calculated and the value of $\mathbf{p} = \{\delta, \kappa\}$ is updated using the previously mentioned LM method. The final point is found to be

Table 5.1: List of Miller Parameters

<i>Miller Parameter</i>	<i>Initial Value</i>
r	0.5
R_0	2.94
R_{geo}	2.94
Δ	-0.11
κ	1.52
κ'	0.10
q	2.02
$\hat{s}$	0.34
δ	0.14
δ'	0.29
β'	0.069

Equilibrium Miller parameter values used in the initial benchmark simulations in Section 5.8.1.

locally stable as the growth rate here is negative. The algorithm then checks a nearby point to determine if a small region of stability exists and, once this has been verified, outputs this as the final $\mathbf{p}$ value.

A second path is plotted on Figure 5.3 in grey. This path is taken using the same adjoint technique, but increasing the step size within the optimisation loop. When the step-size is small, as with the white path, the LM algorithm more closely resembles a gradient-descent method, however when the step size is larger, as with the grey path, the LM algorithm resembles Newton’s method for gradient optimisation. The figure illustrates how the adjoint algorithm can be combined with an optimiser to quickly and efficiently converge to a stable region of parameter space.

5.8.2 Increasing the Critical Temperature Gradient

It has been shown that the adjoint method is a powerful technique for computing stable points in a large parameter space; Figure 5.3 shows that it can efficiently be used inside an optimisation loop to find a minimum of the linear growth rate for a given temperature gradient. Once the adjoint optimisation loop locates a region in the parameter space with negative or zero growth rate, the local temperature gradient (or other plasma profile variable of interest) may be increased and the process repeated. Conversely, if a minimum positive (unstable) growth rate is found, the temperature gradient can be reduced to seek out the optimal shape that maximises the critical temperature gradient at which linear instability occurs.

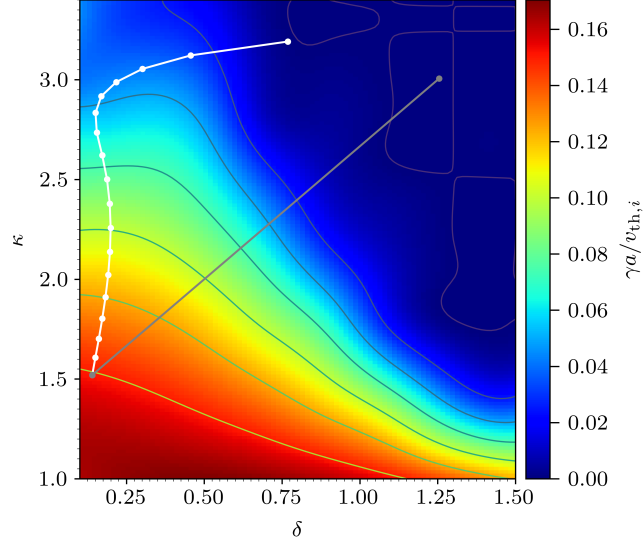


Figure 5.3: Two-dimensional parameter scan over elongation and triangularity, with the colour map indicating the amplitude of the linear growth rate. Here $k_y = 0.68$, $k_x = 0.0$, $m_e/m_i = 2.7 \times 10^{-4}$, $T_i = T_e = 1$, $n_i = n_e = 1$, $L_r/L_{T_i} = L_r/L_{T_e} = 2.42$, $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$, with a the minor radius of the last closed flux surface. The path taken by the optimisation algorithm is indicated in white, with the initial point $\kappa = 1.5$, and $\delta = 0.14$. A second path, drawn in grey, is shown indicating the adjoint optimisation with a larger step size for the optimisation loop.

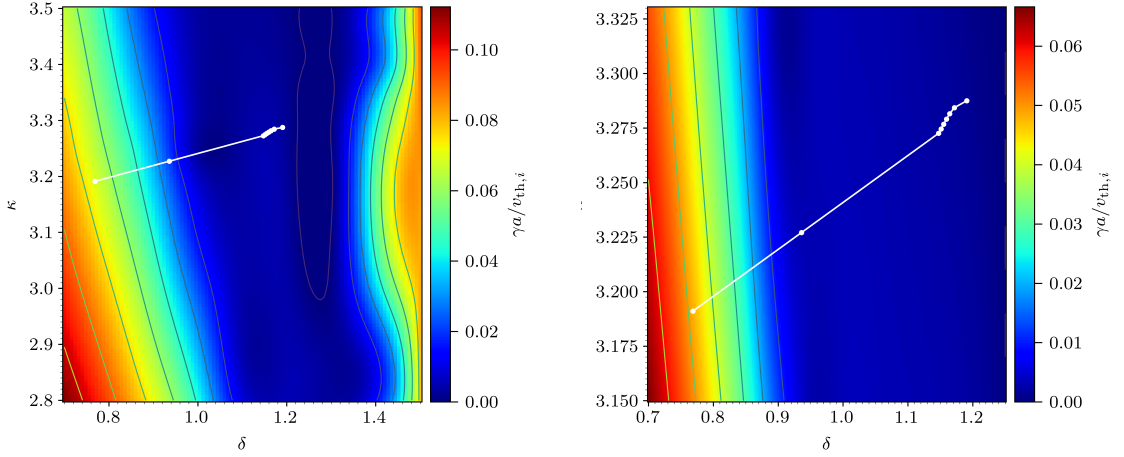


Figure 5.4: Plots showing a parameter scan in elongation and triangularity, with a temperature gradient of $L_r/L_{T_i} = 3.80$, increased from the $L_r/L_{T_i} = 2.42$ value used in Figure 5.3. The geometry of the initial point, located in the unstable region, is provided by final point in Figure 5.3 and is now unstable due to the increased temperature gradient. Here $k_y = 0.68$, $k_x = 0.0$, $m_e/m_i = 2.7 \times 10^{-4}$, $T_i = T_e = 1$, $n_i = n_e = 1$, $L_r/L_{T_i} = L_r/L_{T_e} = 3.80$, $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$. Note that the colour scales used in the figures above are different than that used in Figure 5.3. The right hand side plot is a zoomed region from the figure of the left.

Figure 5.4 demonstrates an iterative use of the adjoint optimisation. Here the temperature gradient has been increased compared to the setup used to generate Figure 5.3. Inside each temperature gradient iteration the adjoint method is used to find a stable geometric configuration. This process could be continued, increasing the temperature gradient until no region of stability is available, indicating a limiting temperature gradient that can be achieved through geometric considerations alone. Figure 5.4 is a demonstration that the adjoint method may be used to increase the temperature gradient, whilst retaining stability, using geometry. Though only two parameters have been considered in the above, with the focus on ITG for clarity, it is possible to employ the adjoint method to optimise over a large number of geometric parameters simultaneously. Such an exploration would be expensive using traditional finite difference methods. One could also extend this to include simulations with kinetic electron effects. In this case, careful consideration would be given to the individual ramping strategies for the ion temperature, electron temperature, and density gradients to systematically explore the multi-dimensional parameter space.

5.8.3 Negative Triangularity

To form a final example here, it is noted that there has been previous evidence that negative triangularity can offer improvements for microstability [97], so the benchmark scan is repeated using a geometry with negative triangularity. All values of equilibrium Miller parameters are the same as in Table 5.1, except now the initial value of triangularity is set to $\delta = -0.14$.

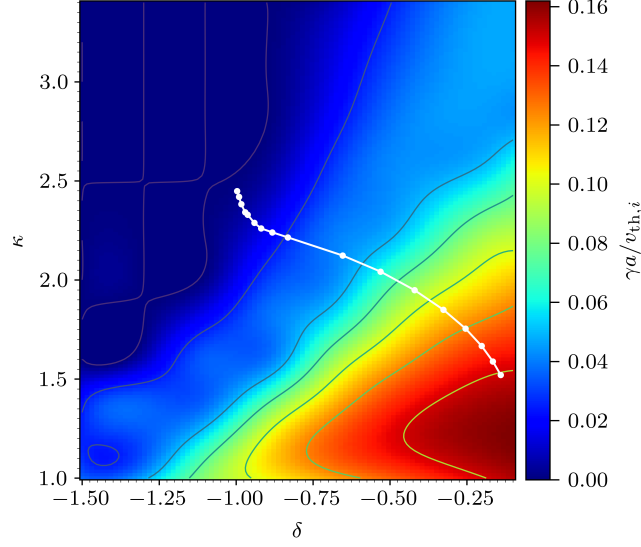


Figure 5.5: Growth rate contours for a parameter scan with negative triangularity for $k_y = 0.68$, $k_x = 0.0$ and equilibrium parameters $m_e/m_i = 2.7 \times 10^{-4}$, $T_i = T_e = 1$, $n_i = n_e = 1$, $L_r/L_{T_i} = L_r/L_{T_e} = 3.80$, $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$. The white line indicates the path taken by the optimisation algorithm. The initial values of $\{\delta, \kappa\}$ are taken to be $\{-0.14, 1.52\}$.

The path taken using the adjoint optimisation loop is again shown in Figure 5.5 by the white path, starting in the dark red region at $\{\delta, \kappa\} = \{-0.14, 1.52\}$. This highlights a key feature of the gradient-based optimisation method: the solution is not unique, and the output can depend on the starting region within parameter space. See Figure 5.6 for an illustrative example of this, which is produced by combining the results demonstrated so far. Finally, the temperature gradient for the negative triangularity case is increased, and the procedure is repeated to find a stable region of parameter space. The input parameters $\{\delta, \kappa\} = \{-0.9965, 2.4488\}$ are again taken to be the outputs from the previously optimised case at a temperature gradient of $L_r/L_{T_i} = 2.42$. The adjoint algorithm coupled with the gradient optimiser is used to look for a stable region of parameter space at an increased temperature gradient of $L_r/L_{T_i} = 3.8$. The results of this are shown in Figure 5.7.

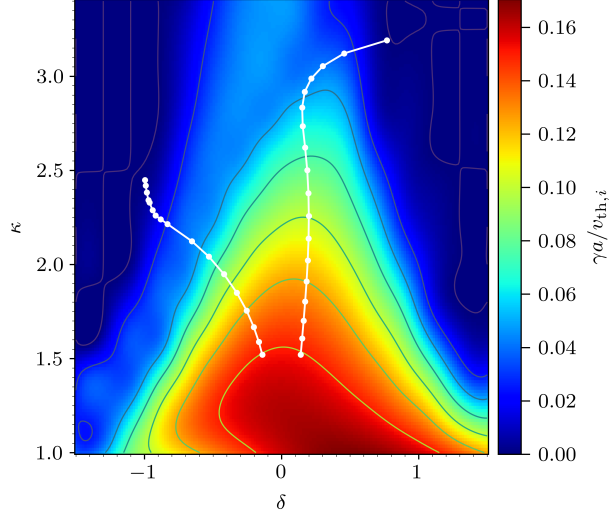


Figure 5.6: Growth rate contours for a parameter scan with both positive and negative triangularity for $k_y = 0.68$, $k_x = 0.0$ and equilibrium parameters $m_e/m_i = 2.7 \times 10^{-4}$, $T_i = T_e = 1$, $n_i = n_e = 1$, $L_r/L_{T_i} = L_r/L_{T_e} = 3.80$, $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$. The white line indicates the two paths taken by the optimisation algorithm starting in different regions in parameter space. The initial values of $\{\delta, \kappa\}$ are taken to be $\{0.14, 1.52\}$ and $\{-0.14, 1.52\}$.

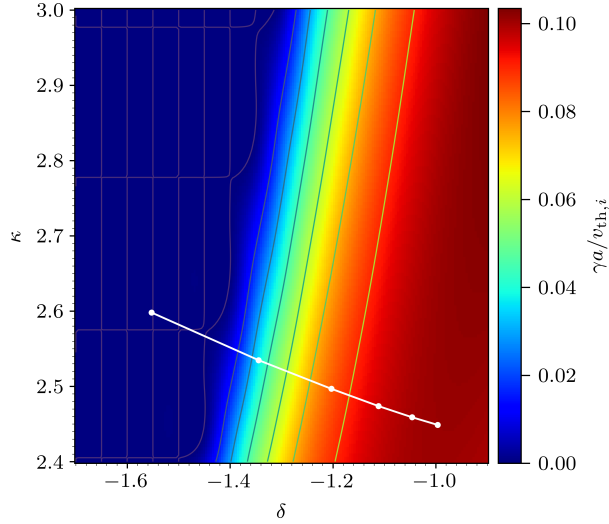


Figure 5.7: Growth rate contours for a parameter scan with negative triangularity at a temperature gradient of $L_r/L_{T_i} = 3.80$, increased from $L_r/L_{T_i} = 2.42$. The geometry of the initial point is taken as the final point in Figure 5.5, and is now unstable due to the increased temperature gradient. The scan is performed at the same parameter values as those in Figure 5.5 – $k_y = 0.68$, $k_x = 0.0$ $m_e/m_i = 2.7 \times 10^{-4}$, $T_i = T_e = 1$, $n_i = n_e = 1$, $L_r/L_{T_i} = L_r/L_{T_e} = 3.80$, $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$. The white line indicates the path taken by the optimisation algorithm. The initial values of $\{\delta, \kappa\}$ are taken to be $\{-0.9965, 2.4488\}$.

5.9 Numerical Efficiency

In the demonstrations presented above, where only a 2D parameter scan is involved, a finite difference approach requires three simulations plus one additional simulation at the next iteration to compute the growth rates. The adjoint method necessitates one simulation of the gyrokinetic equation and the solution of the adjoint equations, which have a similar computational cost. In such low-dimensional cases, the advantage of the adjoint approach is limited compared to the finite difference method. A conventional finite difference approach for a parameter vector $\mathbf{p}$ of size $\mathcal{N}_p$ scales linearly with the number of parameters, requiring $\mathcal{N}_p + 1$ simulations to compute one gradient. Conversely, the adjoint method operates independently of the number of parameters considered, incurring only a marginal cost for the calculation of essential partial derivatives, which are highly efficient for computational systems and can be executed by a single processor. This makes the adjoint approach suitable for high-dimensional spaces without significant computational cost increases.

The progressive computational improvement of using the adjoint method is illustrated by cases extending those presented in Section 5.8. Comparing to a finite difference approach demonstrates the favourable scaling with increasing $\mathcal{N}_p$. In order for the comparisons to be fair, all simulations are run to a standard time of $100L_r/v_{th,i}$.

For the case where $\mathcal{N}_p = 2$ the optimisation is performed with respect to two parameters, $\{\delta, \kappa\}$. To calculate a gradient at each point in the parameter space requires a total of ~ 4.752 CPU (Central Processing Unit) hours on 4 nodes, each with 48 cores on the supercomputer Marconi. The same calculation as performed using the adjoint method requires a total of ~ 3.168 CPU hours. Though this is only a modest improvement, it will be shown that as $\mathcal{N}_p$ increases the advantage of using the adjoint method, over a finite difference scheme, becomes increasingly apparent.

Increasing the number of parameters to $\mathcal{N}_p = 4$, optimising over the Miller parameters $\{\Delta, \kappa, q, \delta\}$, the finite difference approach requires ~ 7.920 CPU hours to compute a gradient at each point in parameter space. However, when computing the same gradient using the adjoint method the CPU time, to the precision of the CPU clock, is ~ 3.168 CPU hours.

If the number of parameters is increased further to $\mathcal{N}_p = 7$, optimising over the Miller parameters $\{R_0, R_{geo}, \Delta, \kappa, q, \hat{s}, \delta\}$, using the finite difference approach requires ~ 12.672 CPU hours to compute each gradient at a point in parameter space. When using the adjoint method to compute the same gradient, to the precision of the CPU clock, the CPU time is still ~ 3.168 CPU hours.

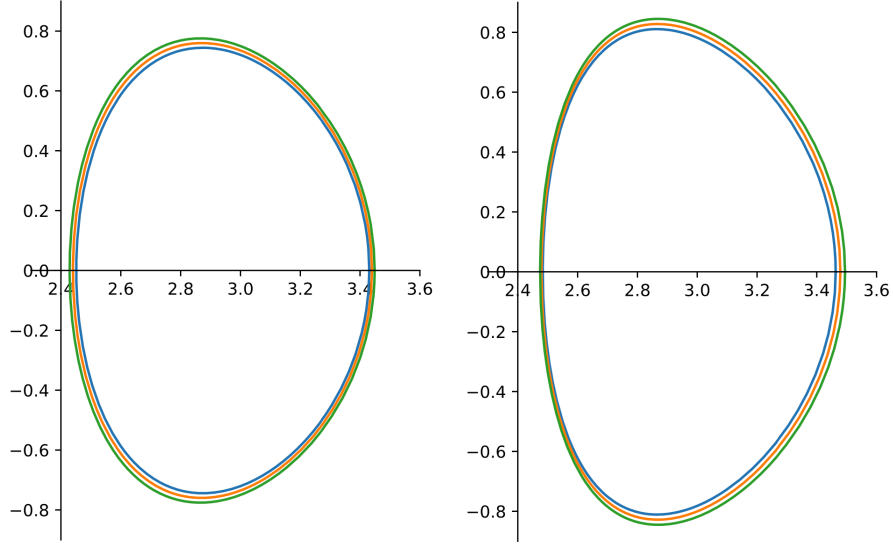


Figure 5.8: Plots of the flux surfaces in the poloidal cross section. The orange surface is the flux surface of interest, at a normalised radial location of $r/L_r = 0.5$, and the blue and green surfaces are the two adjacent flux surfaces. The image on the left is the initial unstable configuration, before optimising. The image on the right is the stable, optimised configuration found using the adjoint-LM system.

Hence, it may be concluded that the numerical speed up is significant with increasing $\mathcal{N}_p$. This allows for the potential inclusion of multiple parameters within our shaping optimisation at very little additional computing cost.

As a closing demonstration, for the example including seven parameters, the process of stepping through parameter space to find a point of stability is iterated. The set of initial parameters is taken to be those given in Table 5.1, and the simulation uses a temperature and density gradient of $L_r/L_{T_i} = L_r/L_{T_e} = 2.42$ and $L_r/L_{n_i} = L_r/L_{n_e} = 0.81$ respectively. The magnetic geometry is perturbed by varying $\{R_0, R_{\text{geo}}, \Delta, \kappa, q, \hat{s}, \delta\}$. Following iterations of the coupled adjoint-LM system a stable configuration is found, with the final set of parameter values: $\{R_0, R_{\text{geo}}, \Delta, \kappa, q, \hat{s}, \delta\} = \{2.979, 2.846, 0.562, 1.656, 2.085, 0.167, 0.225\}$. The cross sections of the initial, unstable configuration and the final, stable configuration, with their corresponding neighbouring flux surfaces, are shown in Figure 5.8. It is important to note that this process was performed for a single mode and does not account for the global equilibrium. Additionally, this example excludes engineering constraints, such as the increased complexity of coil designs. To develop this into a practical method for future machine design, the analysis must be extended to multiple mode numbers and to incorporate relevant engineering constraints.

5.10 Conclusion and Discussion

An adjoint method tailored for local, linear gyrokinetics has been derived and its numerical integration within the δf -gyrokinetic code **stella** has been demonstrated. As a proof-of-principle, the effectiveness of the adjoint method is demonstrated, as applied to a gyrokinetic system, by finding the magnetic geometry of a flux surface that is stable to microinstability. The illustrative example given has focused on increasing the temperature gradient, whilst preserving microstability; ion temperature gradient (ITG) instabilities are often prevalent in fusion devices due to the existence of large temperature gradients, and thus it is conceivable that geometric considerations could help mitigate their growth and so improve overall efficiency.

As the computational cost of the adjoint method remains independent of the number of parameters, its applicability to high-dimensional parameter spaces is readily apparent. The advantages become more pronounced with an expanding number of parameters, as the adjoint method outperforms traditional techniques for calculating gradients, where the computation cost scales with parameter count. This becomes especially beneficial when examining devices like stellarators, which have a large number of geometric parameters that can influence the microinstability of the confined plasma.

It is important to stress that although a specific example focused on increasing the ion temperature gradient has been considered, this approach can be readily extended to optimise the density gradient or other plasma properties, by adapting the overarching LM optimisation loop, without necessitating alterations to the adjoint calculation itself. Such adaptability enables the application of **stella**, and other local δf -gyrokinetic codes, to the exploration of the impact of shaping on various types of microinstabilities, and assessing how geometry can be instrumental in mitigating their growth.

While the numerical examples presented above have focused on the electrostatic, collisionless regime for optimisation with respect to the Miller geometry, equations (5.40)-(5.44) are general. They can be applied to an electromagnetic system, including collisions, and can be optimised using any appropriate set $\{p_i\}$. It should also be stressed that the adjoint method is agnostic to the underlying mechanism driving the linear instability, and provides a gradient in the desired parameter space independently of the drive. As a result the method presented here is applicable to a wide range of microinstabilities.

Note that it should also be emphasised that the adjoint method and the gradient-based optimisation presented are independent of one another. The simplified case of optimising the linear growth rate of a single wavenumber, $\mathbf{k}$, has been considered. For more practical purposes one may wish to apply the adjoint method to a range of wavenumbers, and design the external optimisation routine to take an appropriately weighted average of these, as the most unstable wavenumber may change as one moves within the parameter space.

Chapter 6

Summary and Discussion

In this thesis, the gyrokinetic formalism has been explored as a framework for modeling plasma turbulence in magnetically confined fusion devices. A key component of this work has been the development and implementation of the full flux-annulus (FA) version of the δf -gyrokinetic code **stella**. This extends the algorithms of the flux-tube (FT) version of the code to capture the mode-coupling that arises in non-axisymmetric geometries. One key motivator for this is to correctly model and capture the zonal flows, which have the potential to strongly influence the turbulent saturation levels.

To extend this capability to the kinetic electron regime, we have further developed an iterative-implicit algorithm tailored for efficient and stable time evolution. The accuracy of the FA approach has been benchmarked in the linear regime with Boltzmann electron response, where good agreement with existing flux-tube-based codes has been demonstrated. Notably, we have presented visual evidence of physical differences introduced by the FA treatment—particularly in the zonal flow dynamics—highlighting the importance of capturing global coupling in realistic magnetic configurations.

In addition to the FA code development, an adjoint-based optimisation framework was introduced for minimising the linear growth rate in the gyrokinetic system. While the derivation was performed in the electromagnetic, collisional regime without mode-coupling, the numerical implementation and demonstration was carried out in the electrostatic, collisionless limit. This approach was shown to significantly outperform finite-difference methods in terms of computational efficiency as the number of parameters included in the optimisation increased, making it particularly suitable for high-dimensional optimisation tasks. This is particularly relevant to non-axisymmetric magnetic geometries, where the parameter space is large.

6.1 Future work

Several avenues for future research naturally follow from the work presented in this thesis. One immediate direction is to further apply the FA simulations in order to gain a deeper understanding of zonal flows in non-axisymmetric geometries. Since these flows play a critical role in regulating turbulent transport, a more accurate representation within poloidally-global geometries could reveal new insights into their behaviour and impact.

The FA simulations provide a powerful tool for addressing a wide range of outstanding physics questions. Key among these are determining when the FA growth rates differ from the FT ones, as well as understanding the roles of ρ_* and magnetic geometry. Another important avenue of investigation lies in quantifying the impact of modelling only a single field period versus the full torus, particularly in terms of the role played by the connection length. While the flux-annulus code is presently designed to model turbulence over the complete 2π toroidal domain, simulating a fraction of the device would yield considerable computational savings. Thus, a critical investigation into the viability and potential constraints of this capability is warranted. Furthermore, while we have so far employed phase-shift-periodic parallel boundary conditions, it remains to be explored how non-periodic boundary conditions influence the resulting dynamics. Finally, the question of how long it takes for a global mode structure to develop, and whether this timescale is long relative to the nonlinear saturation time, will be crucial in assessing the relevance of global effects in practical scenarios.

An important extension to include within the FA capabilities will be incorporating electromagnetic effects into the framework. This would enable the simulation of a broader range of physical phenomena, particularly those relevant to high- β plasmas, where electromagnetic fluctuations become significant.

In parallel, there are several opportunities to build upon the adjoint-based optimisation framework. A logical next step is to implement the electromagnetic equations derived in this thesis directly into gyrokinetic codes, enabling adjoint-based optimisation in fully electromagnetic regimes. Applying this method to realistic three-dimensional stellarator geometries would also be a valuable advance, particularly given the high dimensionality of their configuration spaces. This could be done by coupling the adjoint method to a global equilibrium solver, such as **VMEC**. Furthermore, adapting the analytical framework to account for global effects—such as radial and poloidal mode-coupling, which involve interactions between different k_x and k_y

modes—would enhance the method’s applicability to capture more comprehensive physics and the zonal flow dynamics. Finally, the development of a time-dependent adjoint formulation could allow for the inclusion of dynamic mechanisms such as $\mathbf{E} \times \mathbf{B}$ flow shear, opening the door to optimisation strategies that target not just stability, but transport dynamics as well.

These extensions represent natural and impactful directions that build on the foundation laid in this thesis, contributing to the broader goal of improving predictive and optimisation capabilities in gyrokinetic simulations for fusion plasma research.

Appendix A: Full Flux-Annulus

This Appendix contains details of the derivations required in Chapter 4.

A.1 Flux-Surface Average

The chosen coordinate system is (α, ψ, z) , which is right-handed, as $\nabla\alpha \times \nabla\psi \cdot \nabla z = B$. The flux surface average is then

$$\langle f \rangle_x \doteq \frac{1}{A} \int_{z_{\min}}^{z_{\max}} dz \int_{y_{\min}}^{y_{\max}} dy \mathcal{J} f, \quad (\text{A.1})$$

with $\mathcal{J} = (d\alpha/dy)(d\psi/dx)/(\mathbf{B} \cdot \nabla z) = -(\psi_t/|\psi_t|)/B_0$ and

$$A = \int_{z_{\min}}^{z_{\max}} dz \int_{y_{\min}}^{y_{\max}} dy \mathcal{J}, \quad (\text{A.2})$$

and

$$\mathbf{B} \cdot \nabla \zeta = \nabla \psi_t \times \nabla \theta \cdot \nabla \zeta. \quad (\text{A.3})$$

With this the relationship between ψ_{vmec} and the toroidal flux is found

$$\psi_{\text{vmec}} = \frac{1}{2\pi} \int_0^s ds' \int_0^{2\pi} d\theta |\sqrt{g}| \mathbf{B} \cdot \nabla \zeta = \int_0^s ds' \int_0^{2\pi} d\theta |\sqrt{g}| \nabla \psi_t \times \nabla \theta \cdot \nabla \zeta, \quad (\text{A.4})$$

$$= \frac{1}{2\pi} \int_0^s ds' \int_0^{2\pi} d\theta \frac{\nabla \psi_t \times \nabla \theta \cdot \nabla \zeta}{|\nabla \theta \times \nabla \zeta \cdot \nabla s'|}, \quad (\text{A.5})$$

$$= -\frac{1}{2\pi} \int_0^s ds' \int_0^{2\pi} d\theta \frac{\nabla \theta \times \nabla \zeta \cdot \nabla \psi_t}{\nabla \theta \times \nabla \zeta \cdot \nabla s'} = -\psi_t. \quad (\text{A.6})$$

A.1.1 Particle Flux

The aim is to calculate the cross-surface particle flux. To obtain the radial components of flux quantities, the flux is projected onto the ∇x direction:

$$\Gamma_\nu = \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \int d^3\mathbf{v} \frac{\mathbf{v}_E \cdot \nabla x}{\langle |\nabla x| \rangle_x} \delta f_\nu \right\rangle_x, \quad (\text{A.7})$$

with $\delta f_\nu = g_\nu + Z_\nu e/T_\nu(\langle \phi \rangle_{\mathbf{R}_\nu} - \phi)$. Expanding the $\mathbf{E} \times \mathbf{B}$ velocity and inner product:

$$\Gamma_\nu = \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \int d^3\mathbf{v} \frac{\partial \phi}{\partial \alpha} \frac{dx}{d\psi} \frac{c}{\langle |\nabla x| \rangle_x} \delta f_\nu \right\rangle_x. \quad (\text{A.8})$$

In order to correspond with the variables in **stella** the heat flux is normalised to it's gyroBohm value. Defining $\tilde{\Gamma}_\nu = (\Gamma_\nu/n_r v_{\text{th},r})(L_r/\rho_r)^2$ yields

$$\tilde{\Gamma}_\nu = -\frac{\tilde{n}_\nu}{\langle |\nabla x| \rangle_x} \frac{\psi_t}{|\psi_t|} \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \rho_r \frac{\partial \tilde{\phi}}{\partial y} \frac{\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_\parallel \int_0^{\infty} d\tilde{\mu}_\nu \langle \delta \tilde{f}_\nu \rangle_{\mathbf{R}_\nu} \right\rangle_x. \quad (\text{A.9})$$

Recall the definition of particle density as the first moment of the gyroaveraged distribution function

$$\delta \tilde{n}_\nu \doteq \frac{\delta n_\nu}{n_r} = \tilde{n}_\nu \frac{2\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_\parallel \int_0^{\infty} d\tilde{\mu}_\nu \langle \delta \tilde{f}_\nu \rangle_{\mathbf{R}_\nu}, \quad (\text{A.10})$$

to simplify the expression for the heat flux

$$\tilde{\Gamma}_\nu = -\frac{1}{2 \langle |\nabla x| \rangle_x} \frac{\psi_t}{|\psi_t|} \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \frac{\partial \tilde{\phi}}{\partial y / \rho_r} \delta \tilde{n}_\nu \right\rangle_x. \quad (\text{A.11})$$

Because the Jacobian, $\mathcal{J}$, appearing in the flux surface average has y -dependence of the form $1/(\tilde{B}_0 \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z})$, this is factored out and the Fourier expansion of $\delta \tilde{n}_\nu / (\tilde{B}_0 \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z})$ in (x, y) is taken to get

$$\tilde{\Gamma}_\nu = -\frac{1}{2A \langle |\nabla x| \rangle_x} \int_{z_{\min}}^{z_{\max}} dz \left(\sum_{\mathbf{k}} i k_y \rho_r \hat{\phi}_{\mathbf{k}}^*(z) \delta \hat{n}_{\mathbf{k},\nu}(z) \right), \quad (\text{A.12})$$

with $\delta \hat{n}_{\mathbf{k},\nu}(z)$ the Fourier coefficient of $\delta \tilde{n}_\nu / (\tilde{B}_0 \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z})$, and $\mathbf{k} = (k_x, k_y)$. To obtain $\delta \hat{n}_{\mathbf{k},\nu}(z)$, begin by computing $\delta \hat{f}_{\mathbf{k},\nu}$ from $\hat{g}_{\mathbf{k},\nu}$ and $\hat{\phi}_{\mathbf{k}}$, then use this to evaluate the $\mathbf{k}$ -component of $\langle \tilde{f}_\nu \rangle_{\mathbf{R}_\nu}$. Perform an inverse Fourier transform to express the distribution function in (y, k_x) -space. Integrate over velocity space to obtain $\delta \tilde{n}_{k_x,\nu}(y, z)$. This transformation to y -space is required due to the y -dependent coefficient in the velocity space Jacobian. Divide $\delta \tilde{n}_{k_x,\nu}(y, z)$ by $(\tilde{B}_0 \hat{\mathbf{b}} \cdot \tilde{\nabla} \tilde{z})$ in y -space, then take the Fourier transform of the result to obtain $\delta \hat{n}_{\mathbf{k},\nu}(z)$.

A.1.2 Momentum Flux

The radial component of the angular momentum flux is

$$\Pi_\nu = \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \int d^3\mathbf{v} m_\nu R^2 \mathbf{v} \cdot \nabla \zeta \frac{\mathbf{v}_E \cdot \nabla x}{\langle |\nabla x| \rangle_x} \delta f_\nu \right\rangle_x. \quad (\text{A.13})$$

As before, the momentum flux is normalised to its gyroBohm value by defining the normalised quantity $\tilde{\Pi}_\nu \doteq (\Pi_\nu/m_\nu n_\nu L_\nu v_{\text{th},\nu}^2)(L_\nu/\rho_\nu)^2$:

$$\tilde{\Pi}_\nu = -\frac{v_{\text{th},\nu}}{v_{\text{th},r}} \frac{\tilde{n}_\nu \tilde{\Pi}_\nu}{|\langle \nabla x \rangle_x|} \frac{\psi_t}{|\psi_t|} \frac{1}{L_\nu} \int_{x_0}^{x_0+L_\nu} dx \left\langle \rho_\nu \frac{\partial \tilde{\phi}}{\partial y} \frac{\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_\parallel \int_0^\infty d\tilde{\mu}_\nu \tilde{R}^2 \langle \delta \tilde{f}_\nu \tilde{\mathbf{v}} \cdot \tilde{\nabla} \zeta \rangle_{\mathbf{r}_\nu} \right\rangle_x, \quad (\text{A.14})$$

where $\tilde{R} \doteq R/L_\nu$. Consider the gyroaveraged term $\langle \delta \tilde{f}_\nu \tilde{\mathbf{v}} \cdot \tilde{\nabla} \zeta \rangle_{\mathbf{r}_\nu}$ and decompose the toroidal speed as follows:

$$\mathbf{v} \cdot \nabla \zeta = v_\parallel \frac{d\zeta}{dz} + \mathbf{v}_\perp \cdot \nabla \zeta, \quad (\text{A.15})$$

where $d\zeta/dz$ is a quantity that can be determined from geometric coefficients. Using this the non-normalised gyroaverage can be expanded into known terms

$$\begin{aligned} \langle \delta f_\nu \mathbf{v} \cdot \nabla \zeta \rangle_{\mathbf{r}_\nu} &= \left\langle \delta f_\nu \left(\tilde{v}_\parallel \frac{d\zeta}{dz} + \mathbf{v}_\perp \cdot \nabla \zeta \right) \right\rangle_{\mathbf{r}_\nu}, \\ &= v_\parallel \frac{d\zeta}{dz} \langle \delta f_\nu \rangle_{\mathbf{r}_\nu} + \nabla \zeta \cdot \langle \mathbf{v}_\perp \delta f_\nu \rangle_{\mathbf{r}_\nu}. \end{aligned} \quad (\text{A.16})$$

The first term is the familiar result $\langle \delta f_\nu \rangle_{\mathbf{r}_\nu} = J_0(a_{\mathbf{k},\nu})$. The second term can also be manipulated into a simplified form

$$\begin{aligned} \langle \mathbf{v}_\perp \delta f_\nu \rangle_{\mathbf{r}_\nu} &= \sum_{\mathbf{k}} \delta \hat{f}_{\mathbf{k},\nu} e^{i\mathbf{k} \cdot \mathbf{r}} \langle e^{i\mathbf{v}_\perp \cdot \frac{\hat{\mathbf{b}} \times \mathbf{k}}{\Omega_\nu}} \mathbf{v}_\perp \rangle_{\mathbf{r}_\nu}, \\ &= -i \sum_{\mathbf{k}} \delta \hat{f}_{\mathbf{k},\nu} e^{i\mathbf{k} \cdot \mathbf{r}} \frac{1}{\Omega_\nu} \hat{\mathbf{b}} \times \frac{\partial}{\partial \mathbf{k}} \langle e^{i\mathbf{k} \cdot \frac{\mathbf{v}_\perp \times \hat{\mathbf{b}}}{\Omega_\nu}} \rangle_{\mathbf{r}_\nu}, \\ &= i \sum_{\mathbf{k}} \delta \hat{f}_{\mathbf{k},\nu} e^{i\mathbf{k} \cdot \mathbf{r}} \frac{\hat{\mathbf{b}} \times \mathbf{k}_\perp}{k_\perp} v_\perp J_1(a_{\mathbf{k},\nu}). \end{aligned} \quad (\text{A.17})$$

Hence, the $\mathbf{k}^{\text{th}}$ Fourier component of the second term in (A.16) can be written as:

$$\begin{aligned} [\nabla \zeta \cdot \langle \mathbf{v}_\perp \delta f_\nu \rangle_{\mathbf{r}_\nu}]_{\mathbf{k}} &= i \nabla \zeta \cdot \left(\hat{\mathbf{b}} \times \mathbf{k}_\perp \right) \frac{v_\perp^2}{v_{\text{th},\nu}} \frac{J_1(a_{\mathbf{k},\nu})}{a_{\mathbf{k},\nu}} \rho_\nu \delta \hat{f}_{\mathbf{k},\nu}, \\ &= -i \frac{1}{B} \frac{d\alpha}{dy} \frac{d\psi}{dx} \nabla \zeta \cdot ((\nabla x \times \nabla y) \times (k_x \nabla x + k_y \nabla y)) \frac{v_\perp^2}{v_{\text{th},\nu}} \frac{J_1(a_{\mathbf{k},\nu})}{a_{\mathbf{k},\nu}} \rho_\nu \delta \hat{f}_{\mathbf{k},\nu}, \\ &= i \frac{1}{B} \frac{d\alpha}{dy} \frac{d\psi}{dx} \frac{\partial \zeta}{\partial y} [\nabla \zeta \cdot \nabla x (k_x \nabla x \cdot \nabla y + k_y |\nabla y|^2) \\ &\quad - \nabla \zeta \cdot \nabla y (k_x |\nabla x|^2 + k_y \nabla x \cdot \nabla y)] \frac{v_\perp^2}{v_{\text{th},\nu}} \frac{J_1(a_{\mathbf{k},\nu})}{a_{\mathbf{k},\nu}} \rho_\nu \delta \hat{f}_{\mathbf{k},\nu}. \end{aligned} \quad (\text{A.18})$$

Coefficients such as $\nabla \zeta \cdot \nabla x$ and $\nabla \zeta \cdot \nabla y$ can be obtained from geometry files provided by e.g. VMEC geometries. The gyroaveraged toroidal flow (A.17) can thus be related

to the normalised parallel and perpendicular flows,

$$\delta \tilde{u}_{\parallel, \nu} \doteq \frac{v_{\text{th}, \nu}}{v_{\text{th}, r}} \frac{2\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_{\parallel} \int_0^{\infty} d\tilde{\mu}_{\nu} \langle \delta \tilde{f}_{\nu} \rangle_{\mathbf{r}_{\nu}} \tilde{v}_{\parallel}, \quad (\text{A.19})$$

$$\delta \tilde{\mathbf{u}}_{\perp, \nu} \doteq \frac{v_{\text{th}, \nu}}{v_{\text{th}, r}} \frac{2\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_{\parallel} \int_0^{\infty} d\tilde{\mu}_{\nu} \langle \tilde{\mathbf{v}}_{\perp} \delta \tilde{f}_{\nu} \rangle_{\mathbf{r}_{\nu}}. \quad (\text{A.20})$$

Combining the above, the normalised toroidal momentum flux is

$$\tilde{\Pi}_{\nu} = -\frac{\tilde{n}_{\nu} \tilde{m}_{\nu} \tilde{n}_{\nu}}{2 \langle |\nabla x| \rangle_x} \frac{\psi_t}{|\psi_t|} \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \frac{\partial \tilde{\phi}}{\partial y / \rho_r} \left(\delta u_{\parallel, \nu} \frac{d\zeta}{dz} + \delta \tilde{\mathbf{u}}_{\perp, \nu} \cdot \tilde{\nabla} \zeta \right) \delta \tilde{p}_{\nu} \right\rangle_x. \quad (\text{A.21})$$

A.1.3 Heat Flux

The definition for the radial heat flux is

$$Q_{\nu} = \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \int d^3 \mathbf{v} \frac{m_{\nu} v^2}{2} \frac{\mathbf{v}_E \cdot \nabla x}{\langle |\nabla x| \rangle_x} \delta f_{\nu} \right\rangle_x. \quad (\text{A.22})$$

Normalising this to its gyroBohm value $\tilde{Q}_{\nu} \doteq (Q_{\nu} / n_r T_r v_{\text{th}, r}) (L_r / \rho_r)^2$ gives:

$$\tilde{Q}_{\nu} = -\frac{\tilde{n}_{\nu} \tilde{T}_{\nu}}{\langle |\nabla x| \rangle_x} \frac{\psi_t}{|\psi_t|} \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \rho_r \frac{\partial \tilde{\phi}}{\partial y} \frac{\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_{\parallel} \int_0^{\infty} d\tilde{\mu}_{\nu} \langle \delta \tilde{f}_{\nu} \rangle_{\mathbf{r}_{\nu}} \tilde{v}^2 \right\rangle_x. \quad (\text{A.23})$$

To relate this to the pressure perturbations, recall the normalised definition

$$\delta \tilde{p}_{\nu} = \frac{2\tilde{B}_0}{\pi^{1/2}} \int_{-\infty}^{\infty} d\tilde{v}_{\parallel} \int_0^{\infty} d\tilde{\mu}_{\nu} \langle \delta \tilde{f}_{\nu} \rangle_{\mathbf{r}_{\nu}} \tilde{v}^2. \quad (\text{A.24})$$

Combining the above, a definition for the normalised heat flux can be obtained

$$\tilde{Q}_{\nu} = -\frac{1}{2 \langle |\nabla x| \rangle_x} \frac{\psi_t}{|\psi_t|} \frac{1}{L_x} \int_{x_0}^{x_0+L_x} dx \left\langle \frac{\partial \tilde{\phi}}{\partial y / \rho_r} \delta \tilde{p}_{\nu} \right\rangle_x. \quad (\text{A.25})$$

Appendix B: Adjoint

This Appendix contains details of the derivations required in Chapter 5.

B.1 Decomposition of Operators

The definitions of the operators introduced in Section 5.5 in (5.30)-(5.31) are

$$\begin{aligned}
\hat{G}_{g,\nu}[\mathbf{p}; \hat{g}_\nu] &= \gamma \hat{g}_\nu + v_\parallel \hat{\mathbf{b}} \cdot \nabla z \frac{\partial \hat{g}_\nu}{\partial z} - \frac{\mu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{\partial \hat{g}_\nu}{\partial v_\parallel} + i\omega_{d,\nu} \hat{g}_\nu - \hat{C}_\nu[\mathbf{p}; \hat{g}_{\nu'}], \\
\hat{G}_{\phi,\nu}[\mathbf{p}; \hat{\phi}] &= v_\parallel \hat{\mathbf{b}} \cdot \nabla z \frac{Z_\nu e}{T_\nu} F_{0,\nu} \frac{\partial J_{0,\nu} \hat{\phi}}{\partial z} + i \frac{Z_\nu e}{T_\nu} \omega_{d,\nu} J_{0,\nu} F_{0,\nu} \hat{\phi} + i\omega_{*,\nu} J_{0,\nu} F_{0,\nu} \hat{\phi} \\
&\quad - \hat{C}_\nu \left[\mathbf{p}; \frac{Z_{\nu'} e}{T_{\nu'}} J_{0,\nu'} F_{0,\nu'} \hat{\phi} \right], \\
\hat{G}_{A_\parallel,\nu}[\mathbf{p}; \hat{A}_\parallel] &= -\frac{v_\parallel^2}{c} \hat{\mathbf{b}} \cdot \nabla z \frac{Z_\nu e}{T_\nu} F_{0,\nu} \frac{\partial J_{0,\nu} \hat{A}_\parallel}{\partial z} - i \frac{Z_\nu e}{T_\nu} \omega_{d,\nu} J_{0,\nu} F_{0,\nu} \frac{v_\parallel}{c} \hat{A}_\parallel - i\omega_{*,\nu} J_{0,\nu} F_{0,\nu} \frac{v_\parallel}{c} \hat{A}_\parallel \\
&\quad + \frac{Z_\nu e}{T_\nu} \frac{\mu_\nu}{m_\nu c} \hat{\mathbf{b}} \cdot \nabla B_0 F_{0,\nu} J_{0,\nu} \hat{A}_\parallel + \hat{C}_\nu \left[\mathbf{p}; \frac{Z_{\nu'} e}{T_{\nu'}} \frac{v_\parallel}{c} J_{0,\nu'} F_{0,\nu'} \hat{A}_\parallel \right], \\
\hat{G}_{B_\parallel,\nu}[\mathbf{p}; \delta \hat{B}_\parallel] &= 2v_\parallel \hat{\mathbf{b}} \cdot \nabla z F_{0,\nu} \frac{\partial}{\partial z} \left(\frac{\mu_\nu}{T_\nu} \frac{J_{1,\nu}}{a_\nu} \delta \hat{B}_\parallel \right) + 2i\omega_{d,\nu} \frac{\mu_\nu}{T_\nu} \frac{J_{1,\nu}}{a_\nu} F_{0,\nu} \delta \hat{B}_\parallel \\
&\quad - \hat{C}_\nu \left[\mathbf{p}; 2 \frac{J_{1,\nu'}}{a_{\nu'}} \frac{\mu_{\nu'}}{T_{\nu'}} F_{0,\nu'} \delta \hat{B}_\parallel \right] + 2i\omega_{*,\nu} \frac{\mu_\nu}{Z_\nu e} \frac{J_{1,\nu}}{a_\nu} F_{0,\nu} \delta \hat{B}_\parallel, \\
\hat{Q}_{g,\nu}[\mathbf{p}; \hat{g}_\nu] &= Z_\nu e J_{0,\nu} \hat{g}_\nu, \\
\hat{Q}_\phi[\mathbf{p}; \hat{\phi}] &= \sum_\nu \frac{(Z_\nu e)^2 n_\nu}{T_\nu} (\Gamma_{0,\nu} - 1) \hat{\phi}, \\
\hat{Q}_{B_\parallel}[\mathbf{p}; \delta \hat{B}_\parallel] &= \sum_\nu Z_\nu e n_\nu \frac{\Gamma_{1,\nu}}{B_0} \delta \hat{B}_\parallel, \\
\hat{M}_{g,\nu}[\mathbf{p}; \hat{g}_\nu] &= -\frac{4\pi}{k_\perp^2} \frac{v_\parallel}{c} Z_\nu e J_{0,\nu} \hat{g}_\nu, \\
\hat{M}_{A_\parallel}[\mathbf{p}; \hat{A}_\parallel] &= \left[1 + \frac{4\pi}{k_\perp^2 c^2} \sum_\nu \frac{(Z_\nu e)^2 n_\nu}{m_\nu} \Gamma_{0,\nu} \right] \hat{A}_\parallel, \\
\hat{N}_{g,\nu}[\mathbf{p}; \hat{g}_\nu] &= 8\pi \frac{J_{1,\nu}}{a_\nu} \mu_\nu \hat{g}_\nu, \\
\hat{N}_\phi[\mathbf{p}; \hat{\phi}] &= \left[4\pi \sum_\nu \frac{Z_\nu e n_\nu}{B_0} \Gamma_{1,\nu} \right] \hat{\phi},
\end{aligned}$$

$$\hat{N}_{B_{\parallel}}[\mathbf{p}; \delta \hat{B}_{\parallel}] = \left[1 + 16\pi \sum_{\nu} \frac{n_{\nu} T_{\nu}}{B_0^2} \Gamma_{2,\nu} \right] \delta \hat{B}_{\parallel}, \quad (\text{B.26})$$

and $\hat{C}_{\nu}$ is the appropriate collision operator, which can be thought of as acting on $\hat{g}_{\nu'}$, $\hat{\phi}$, $\hat{A}_{\parallel}$, $\hat{B}_{\parallel}$ separately:

$$\begin{aligned} \hat{C}_{\nu}[\mathbf{p}; \{\hat{g}_{\nu'}, \hat{\chi}_{\nu'}\}] = & \hat{C}_{\nu}[\mathbf{p}; \hat{g}_{\nu'}] + \hat{C}_{\phi,\nu} \left[\mathbf{p}; \frac{Z_{\nu}}{T_{\nu}} J_{0,\nu'} \hat{\phi} \right] + \hat{C}_{A_{\parallel},\nu} \left[\mathbf{p}; J_{0,\nu'} \hat{A}_{\parallel} \right] \\ & + \hat{C}_{B_{\parallel},\nu} \left[\mathbf{p}; \frac{\mu_{\nu'}}{Z_{\nu'}} \frac{J_{1,\nu'}}{a_{\nu'}} \hat{B}_{\parallel} \right]. \end{aligned} \quad (\text{B.27})$$

B.2 Adjoints of Operators

The adjoint operators appearing in equations (5.36)-(5.39) are obtained by performing integration by parts wherever a derivative acts on the distribution function or a field variable. Following the change of variables $\tilde{v}_{\parallel} \rightarrow -\tilde{v}_{\parallel}$ these operators take the form:

$$\begin{aligned} \hat{G}_{g,\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] &= \gamma^* \lambda_{\nu}^{\leftrightarrow} + v_{\parallel} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial \lambda_{\nu}^{\leftrightarrow}}{\partial z} - \frac{\mu_{\nu}}{m_{\nu}} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{\partial \lambda_{\nu}^{\leftrightarrow}}{\partial v_{\parallel}} - i\omega_{d,\nu} \lambda_{\nu}^{\leftrightarrow} - \hat{C}_{\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu'}], \\ \hat{G}_{\phi,\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] &= \frac{Z_{\nu} e}{T_{\nu}} J_{0,\nu} F_{0,\nu} \hat{S}_{\nu} - \frac{Z_{\nu} e}{T_{\nu}} J_{0,\nu} F_{0,\nu} \hat{C}_{\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu'}], \\ \hat{G}_{A_{\parallel},\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] &= \frac{Z_{\nu} e}{T_{\nu}} J_{0,\nu} F_{0,\nu} \frac{v_{\parallel}}{c} \hat{S}_{\nu} + \frac{Z_{\nu} e}{T_{\nu}} \frac{\mu_{\nu}}{m_{\nu} c} \hat{\mathbf{b}} \cdot \nabla B_0 F_{0,\nu} J_{0,\nu} \lambda_{\nu}^{\leftrightarrow} + \frac{Z_{\nu} e}{T_{\nu}} \frac{v_{\parallel}}{c} J_{0,\nu} F_{0,\nu} \hat{C}_{\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu'}], \\ \hat{G}_{B_{\parallel},\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu}^{\leftrightarrow}] &= 2 \frac{\mu_{\nu}}{T_{\nu}} \frac{J_{1,\nu}}{\tilde{a}_{\nu}} F_{0,\nu} \hat{S}_{\nu} - 2 \frac{J_{1,\nu}}{a_{\nu}} \frac{\mu_{\nu}}{T_{\nu}} F_{0,\nu} \hat{C}_{\nu}^{\dagger}[\mathbf{p}; \lambda_{\nu'}], \\ \hat{Q}_{g,\nu}^{\dagger}[\mathbf{p}; \xi] &= Z_{\nu} e J_{0,\nu} F_{0,\nu} \xi, \\ \hat{Q}_{\phi}^{\dagger}[\mathbf{p}; \xi] &= \sum_{\nu} \frac{(Z_{\nu} e)^2 n_{\nu}}{T_{\nu}} (\Gamma_{0,\nu} - 1) \xi, \\ \hat{Q}_{B_{\parallel}}^{\dagger}[\mathbf{p}; \xi] &= 4\pi \sum_{\nu} \frac{Z_{\nu} e n_{\nu}}{B_0} \Gamma_{1,\nu} \xi, \\ \hat{M}_{g,\nu}^{\dagger}[\mathbf{p}; \zeta] &= -\frac{4\pi}{k_{\perp}^2} \frac{v_{\parallel}}{c} Z_{\nu} e J_{0,\nu} \zeta, \\ \hat{M}_{A_{\parallel}}^{\dagger}[\mathbf{p}; \zeta] &= \left[1 + \frac{4\pi}{k_{\perp}^2 c^2} \sum_{\nu} \frac{(Z_{\nu} e)^2 n_{\nu}}{m_{\nu}} \Gamma_{0,\nu} \right] \zeta, \\ \hat{N}_{g,\nu}^{\dagger}[\mathbf{p}; \sigma] &= 8\pi \frac{J_{1,\nu}}{a_{\nu}} \mu_{\nu} \sigma, \\ \hat{N}_{\phi}^{\dagger}[\mathbf{p}; \sigma] &= \left[4\pi \sum_{\nu} \frac{Z_{\nu} e n_{\nu}}{B_0} \Gamma_{1,\nu} \right] \sigma, \\ \hat{N}_{B_{\parallel}}^{\dagger}[\mathbf{p}; \sigma] &= \left[1 + 16\pi \sum_{\nu} \frac{n_{\nu} T_{\nu}}{B_0^2} \Gamma_{2,\nu} \right] \sigma, \end{aligned} \quad (\text{B.28})$$

and $\hat{C}_\nu^\dagger[\mathbf{p}; \cdot] = \hat{C}_\nu[\mathbf{p}; \cdot]$ is the self-adjoint collision operator. Here, the following definitions have been made

$$\hat{S}[\mathbf{p}; \lambda_\nu^{\leftrightarrow}] = v_\parallel \hat{\mathbf{b}} \cdot \nabla_z \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z} - i\omega_{d,\nu} \lambda_\nu^{\leftrightarrow} - i\omega_{*,\nu} \lambda_\nu^{\leftrightarrow}. \quad (\text{B.29})$$

B.3 Simplifying Adjoint Equations

Consider taking the following moments of (5.36):

$$\left\langle \frac{Z_\nu e}{T_\nu} J_{0,\nu} F_{0,\nu}, \cdot \right\rangle_{v,\nu}, \quad \left\langle \frac{Z_\nu e}{T_\nu} J_{0,\nu} F_{0,\nu} \frac{v_\parallel}{c}, \cdot \right\rangle_{v,\nu}, \quad \left\langle 2 \frac{J_{1,\nu}}{a_\nu} F_{0,\nu} \frac{\mu_\nu}{T_\nu}, \cdot \right\rangle_{v,\nu}, \quad (\text{B.30})$$

giving

$$\begin{aligned} 0 = & \sum_\nu \frac{2\pi B_0}{m_\nu} \int dv_\parallel \int d\mu_\nu \alpha_\nu(z, v_\parallel, \mu_\nu) F_{0,\nu} \left\{ \gamma^* \lambda_\nu^{\leftrightarrow} + v_\parallel \hat{\mathbf{b}} \cdot \nabla_z \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z} \right. \\ & - \frac{\mu_\nu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla_z \frac{\partial B_0}{\partial z} \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial v_\parallel} - i\omega_{d,\nu} \lambda_\nu^{\leftrightarrow} + Z_\nu e J_{0,\nu} \xi - \frac{4\pi}{k_\perp^2} Z_\nu e J_{0,\nu} \frac{v_\parallel}{c} \zeta + 8\pi \frac{J_{1,\nu}}{a_\nu} \mu_\nu \sigma \\ & \left. - \hat{C}_\nu[\lambda_\nu^{\leftrightarrow}] \right\}, \end{aligned} \quad (\text{B.31})$$

where α_ν can take the forms:

$$\alpha_\nu = \begin{cases} \frac{Z_\nu e}{T_\nu} J_{0,\nu} \\ \frac{Z_\nu e}{T_\nu} J_{0,\nu} \frac{v_\parallel}{c} \\ 2 \frac{J_{1,\nu}}{a_\nu} \frac{\mu_\nu}{T_\nu} \end{cases}. \quad (\text{B.32})$$

We now identify different terms in (B.31) for each α_ν which are odd in $v_\parallel$ so evaluate to zero when integrated over the domain $\{-\infty, \infty\}$:

$$\begin{aligned} 0 = & \sum_\nu \frac{2\pi B_0}{m_\nu} \int d^2v \frac{Z_\nu e}{T_\nu} F_{0,\nu} \left\{ \gamma^* J_{0,\nu} \lambda_\nu^{\leftrightarrow} + v_\parallel \hat{\mathbf{b}} \cdot \nabla_z J_{0,\nu} \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial z} - i\omega_{d,\nu} J_{0,\nu} \lambda_\nu^{\leftrightarrow} \right. \\ & - \frac{\mu_\nu}{m_\nu} \hat{\mathbf{b}} \cdot \nabla_z \frac{\partial B_0}{\partial z} \frac{\partial \lambda_\nu^{\leftrightarrow}}{\partial v_\parallel} + Z_\nu e J_{0,\nu}^2 \xi + \underbrace{\frac{4\pi}{k_\perp^2} Z_\nu e J_{0,\nu}^2 \frac{v_\parallel}{c} \zeta}_{\text{odd in } v_\parallel} + 8\pi \frac{J_{1,\nu} J_{0,\nu}}{a_\nu} \mu_\nu \sigma - \frac{Z_\nu e}{T_\nu} J_{0,\nu} \hat{C}_\nu[\lambda_\nu^{\leftrightarrow}] \left. \right\}, \end{aligned} \quad (\text{B.33})$$

$$\begin{aligned}
0 = \sum_{\nu} \frac{2\pi B_0}{m_{\nu}} \int d^2v \frac{Z_{\nu} e}{T_{\nu}} F_{0,\nu} & \left\{ \gamma^* J_{0,\nu} \frac{v_{\parallel}}{c} \lambda_{\nu}^{\leftrightarrow} + \frac{v_{\parallel}^2}{c} \hat{\mathbf{b}} \cdot \nabla z J_{0,\nu} \frac{\partial \lambda_{\nu}^{\leftrightarrow}}{\partial z} \right. \\
& - \frac{\mu_{\nu}}{m_{\nu}} \frac{v_{\parallel}}{c} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{\partial \lambda_{\nu}^{\leftrightarrow}}{\partial v_{\parallel}} - i\omega_{d,\nu} \frac{v_{\parallel}}{c} J_{0,\nu} \lambda_{\nu}^{\leftrightarrow} \\
& + \underbrace{Z_{\nu} e J_{0,\nu}^2 \frac{v_{\parallel}}{c} \xi + \frac{4\pi}{k_{\perp}^2} Z_{\nu} e J_{0,\nu}^2 \left(\frac{v_{\parallel}}{c} \right)^2 \zeta}_{\text{odd in } v_{\parallel}} + \underbrace{8\pi \frac{J_{1,\nu} J_{0,\nu}}{a_{\nu}} \frac{v_{\parallel}}{c} \mu_{\nu} \sigma - \frac{Z_{\nu} e}{T_{\nu}} J_{0,\nu} \frac{v_{\parallel}}{c} \hat{C}_{\nu}[\lambda_{\nu}^{\leftrightarrow}]}_{\text{odd in } v_{\parallel}} \left. \right\}, \tag{B.34}
\end{aligned}$$

$$\begin{aligned}
0 = 2 \sum_{\nu} \frac{2\pi B_0}{m_{\nu}} \int d^2v F_{0,\nu} & \left\{ \frac{J_{1,\nu} \mu_{\nu}}{a_{\nu}} \frac{\mu_{\nu}}{T_{\nu}} \gamma^* \lambda_{\nu}^{\leftrightarrow} + v_{\parallel} \frac{J_{1,\nu} \mu_{\nu}}{a_{\nu}} \frac{\mu_{\nu}}{T_{\nu}} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial \lambda_{\nu}^{\leftrightarrow}}{\partial z} \right. \\
& - \frac{\mu_{\nu}}{m_{\nu}} \hat{\mathbf{b}} \cdot \nabla z \frac{\partial B_0}{\partial z} \frac{J_{1,\nu} \mu_{\nu}}{a_{\nu}} \frac{\mu_{\nu}}{T_{\nu}} \frac{\partial \lambda_{\nu}^{\leftrightarrow}}{\partial v_{\parallel}} - i\omega_{d,\nu} \frac{J_{1,\nu} \mu_{\nu}}{a_{\nu}} \frac{\mu_{\nu}}{T_{\nu}} \lambda_{\nu}^{\leftrightarrow} + \frac{Z_{\nu} e \mu_{\nu}}{T_{\nu}} \frac{J_{0,\nu}}{a_{\nu}} \frac{J_{1,\nu}}{a_{\nu}} \xi \\
& + \underbrace{\frac{4\pi}{k_{\perp}^2} Z_{\nu} e \frac{J_{0,\nu} J_{1,\nu}}{a_{\nu}} \frac{v_{\parallel}}{c} \frac{\mu_{\nu}}{T_{\nu}} \zeta}_{\text{odd in } v_{\parallel}} + 8\pi \left(\frac{J_{1,\nu}}{a_{\nu}} \right)^2 \frac{\mu_{\nu}^2}{T_{\nu}} \sigma - 2 \frac{J_{1,\nu} \mu_{\nu}}{a_{\nu}} \frac{\mu_{\nu}}{T_{\nu}} \hat{C}_{\nu}[\lambda_{\nu}^{\leftrightarrow}] \left. \right\}. \tag{B.35}
\end{aligned}$$

Integration by parts can be performed on the remaining $v_{\parallel}$ derivative, using velocity-independence of the fields. Using equations (5.37)-(5.39) it is then possible to simplify the above equations to produce the results (5.41)-(5.43).

B.4 Geometry Implementation

The code `stella` has an input option to use the Miller parametrisation of a flux surface; it takes a set of input variables to describe the local geometry of a specified flux surface along with the two adjacent flux surfaces on either side. Equation (2.12) is taken in the form

$$R(r, \theta) = R_0(r) + r_0(r) \cos[\theta + \sin(\theta)\delta(r)], \tag{B.36}$$

with the triangularity redefined as $\delta(r) \doteq \arcsin[\bar{\delta}(r)]$. Consider Taylor expanding in r about $r = r_{\psi_0}$

$$R_0(r) = R_0(r_{\psi_0}) + \left. \frac{dR_0}{dr} \right|_{r_{\psi_0}} (r - r_{\psi_0}) + \dots \approx R_{\psi_0} + \Delta_{\psi_0} dr + \mathcal{O}(dr^2), \tag{B.37}$$

with $dr = r - r_{\psi_0}$, $R_{\psi_0} = R_0(r_{\psi_0})$, and $\Delta_{\psi_0} = \left. \frac{dR_0}{dr} \right|_{r_{\psi_0}}$. Similarly,

$$\delta(r) = \delta(r_{\psi_0}) + \left. \frac{d\delta}{dr} \right|_{r_{\psi_0}} (r - r_{\psi_0}) + \dots \approx \delta_{\psi_0} + \delta'_{\psi_0} dr + \mathcal{O}(dr^2), \tag{B.38}$$

with $\delta_{\psi_0} = \delta(r_{\psi_0})$, and $\delta'_{\psi_0} = d\delta/dr|_{r_{\psi_0}}$. We also define $r_0(r_{\psi_0}) = r_{0,\psi_0}$, such that $r_0(r) = r_{0,\psi_0} + dr$. Combining this and equations (B.37) and (B.38) gives

$$R(r, \theta) \approx R_{\psi_0} + r_{0,\psi_0} \cos[\theta + \sin(\theta)\delta_{\psi_0}] + \left\{ \Delta_{\psi_0} + \cos[\theta + \sin(\theta)\delta_{\psi_0}] - r_{0,\psi_0} \sin[\theta + \sin(\theta)\delta_{\psi_0}] \sin(\theta)\delta'_{\psi_0} \right\} dr, \quad (\text{B.39})$$

such that the above definition holds on any given flux surface, ψ_0 . Equivalently, (2.13) can also be expanded about $r = r_{\psi_0}$ by first expanding the elongation

$$\kappa(r) = \kappa(r_{\psi_0}) + \left. \frac{d\kappa}{dr} \right|_{r_{\psi_0}} (r - r_{\psi_0}) + \dots \approx \kappa_{\psi_0} + \kappa'_{\psi_0} dr + \mathcal{O}(dr^2), \quad (\text{B.40})$$

with $\kappa_{\psi_0} = \kappa(r_{\psi_0})$, and $\kappa'_{\psi_0} = d\kappa/dr|_{r_{\psi_0}}$ to then write

$$Z(r, \theta) \approx r_{0,\psi_0} \kappa_{\psi_0} \sin(\theta) + [\kappa_{\psi_0} + r_{0,\psi_0} \kappa'_{\psi_0}] \sin(\theta) dr. \quad (\text{B.41})$$

These functions are used to describe the geometry of the flux surface of interest and the two adjacent flux surfaces by setting $dr = \{0, \pm\Lambda\}$, with $\Lambda \ll 1$ a constant, in order to evaluate their radial derivatives. These quantities are then used to compute the Jacobian, magnetic field strength and configuration, along with other functions defined on the flux surface.

Bibliography

- [1] I. G. Abel, G. G. Plunk, E. Wang, M. Barnes, S. C. Cowley, W. Dorland, and A. A. Schekochihin. Multiscale gyrokinetics for rotating tokamak plasmas: fluctuations, transport and energy flows. *Reports on Progress in Physics*, 76(11):116201, 2013.
- [2] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York, NY, first edition, 1965.
- [3] G. O. Acton, M. Barnes, S. Newton, and H. Thienpondt. Optimisation of gyrokinetic microstability using adjoint methods. *Journal of Plasma Physics*, 90(4):905900406, 2024.
- [4] J. A. Alonso, E. Sánchez, I. Calvo, J. L. Velasco, K. J. McCarthy, A. Chmyga, L. G. Eliseev, T. Estrada, R. Kleiber, L. I. Krupnik, et al. Observation of oscillatory radial electric field relaxation in a helical plasma. *Physical Review Letters*, 118(18):185002, 2017.
- [5] T. M. Antonsen Jr and B. Lane. Kinetic equations for low frequency instabilities in inhomogeneous plasmas. *The Physics of Fluids*, 23(6):1205–1214, 1980.
- [6] A. Bader, M. Drevlak, D. T. Anderson, B. J. Faber, C. C. Hegna, K. M. Likin, J. C. Schmitt, and J. N. Talmadge. Stellarator equilibria with reactor relevant energetic particle losses. *Journal of Plasma Physics*, 85(5):905850508, 2019.
- [7] M Barnes, I. G. Abel, W. Dorland, T. Görler, G. W. Hammett, and F. Jenko. Direct multiscale coupling of a transport code to gyrokinetic turbulence codes. *Physics of Plasmas*, 17(5), 2010.
- [8] M. Barnes, W. Dorland, and T. Tatsuno. Resolving velocity space dynamics in continuum gyrokinetics. *Physics of Plasmas*, 17(3), 2010.

- [9] M. Barnes, F. I. Parra, and M. Landreman. stella: An operator-split, implicit-explicit δf -gyrokinetic code for general magnetic field configurations. *Journal of Computational Physics*, 391:365–380, 2019.
- [10] O. Beeke, M. Barnes, M. Romanelli, M. Nakata, and M. Yoshida. Impact of shaping on microstability in high-performance tokamak plasmas. *Nuclear Fusion*, 61(6):066020, 2021.
- [11] M. A. Beer, S. C. Cowley, and G. W. Hammett. Field-aligned coordinates for nonlinear simulations of tokamak turbulence. *Physics of Plasmas*, 2(7):2687–2700, 1995.
- [12] Michael A. Beer. *Gyrofluid Models of Turbulent Transport in Tokamaks*. Ph.d. thesis, Princeton University, Princeton, NJ, 1995.
- [13] Emily A. Belli. *Studies of Numerical Algorithms for Gyrokinetics and the Effects of Shaping on Plasma Turbulence*. PhD thesis, Princeton University, Princeton, NJ, 2006.
- [14] H. Biglari, P. H. Diamond, and P. W. Terry. Influence of sheared poloidal rotation on edge turbulence. *Physics of Fluids B: Plasma Physics*, 2(1):1–4, 1990.
- [15] A. H. Boozer. Transport and isomorphic equilibria. *Physics of Fluids*, 26(2):496–499, 1983.
- [16] A. H. Boozer. Stellarators as a fast path to fusion energy. *arXiv preprint arXiv:1912.06289*, 2019.
- [17] Allen H. Boozer. Guiding center drift equations. *Physics of Fluids*, 23(5):904–908, 1980.
- [18] Allen H. Boozer. Plasma equilibrium with rational magnetic surfaces. *Physics of Fluids*, 24(11):1999–2003, 1981.
- [19] John P. Boyd. *Chebyshev and Fourier Spectral Methods*. Dover Publications, 2nd edition, 2001.
- [20] K. H. Burrell. Turbulence and sheared flow. *Science*, 281(5384):1816–1817, 1998.

- [21] J. Candy, C. Holland, R. E. Waltz, M. R. Fahey, and E. A. Belli. Tokamak profile prediction using direct gyrokinetic and neoclassical simulation. *Physics of Plasmas*, 16(6):060704, 2009.
- [22] J. Candy and R. E. Waltz. An Eulerian gyrokinetic-Maxwell solver. *Journal of Computational Physics*, 186(2):545–581, 2003.
- [23] J. R. Cary and S. G. Shasharina. Omnigenity and quasihelicity in helical plasma confinement systems. *Physics of Plasmas*, 4(9):3323–3333, 1997.
- [24] P. J. Catto. Linearized gyro-kinetics. *Plasma Physics*, 20(7):719, 1978.
- [25] P. J. Catto. Practical gyrokinetics. *Journal of Plasma Physics*, 85(3):925850301, 2019.
- [26] F. F. Chen. *Introduction to Plasma Physics and Controlled Fusion*, volume 1. Springer, 1984.
- [27] B. Coppi, M. N. Rosenbluth, and R. Z. Sagdeev. Instabilities due to temperature gradients in complex magnetic field configurations. *Physics of Fluids*, 10(3):582–588, 1967.
- [28] S. C. Cowley, R. M. Kulsrud, and R. Sudan. Considerations of ion-temperature-gradient-driven turbulence. *Physics of Fluids B*, 3(10):2767–2782, 1991.
- [29] W. Dekeyser. *Optimal plasma edge configurations for next-step fusion reactors (optimale plasmarand-configuraties voor nieuwe generatie fusiereactoren)*. PhD thesis, Ghent University, 2014.
- [30] A. Di Siena, A. B. Navarro, T. Luda, G. Merlo, M. Bergmann, L. Leppin, T. Görler, J. B. Parker, L. LoDestro, T. Dannert, et al. Global gyrokinetic simulations of asdex upgrade up to the transport timescale with gene-tango. *Nuclear Fusion*, 62(10):106025, 2022.
- [31] P. H. Diamond, S. I. Itoh, K. Itoh, and T. S. Hahm. Zonal flows in plasma—a review. *Plasma Physics and Controlled Fusion*, 47(5):R35, 2005.
- [32] 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, et al. Comparisons and physics basis of tokamak transport models and turbulence simulations. *Physics of Plasmas*, 7(3):969–983, 2000.

- [33] A. M. Dimits, J. A. Byers, T. J. Williams, B. I. Cohen, W. Q. Xu, R. H. Cohen, J. A. Crotinger, and A. I. Shestakov. Gyrokinetic and global fluid simulations of tokamak microturbulence and transport. Technical Report UCRL-ID-118764, Lawrence Livermore National Laboratory, Livermore, CA, USA, 1994.
- [34] W. Dorland, F. Jenko, M. Kotschenreuther, and B. N. Rogers. Electron temperature gradient turbulence. *Physical Review Letters*, 85(26):5579, 2000.
- [35] E. J. Doyle, W. A. Houlberg, Y. Kamada, and et al. Chapter 2: Plasma confinement and transport [progress in the iter physics basis]. *Nuclear Fusion*, 47(6):S18–S127, 2007.
- [36] J. F. Drake, N. T. Gladd, C. S. Liu, and C. L. Chang. Microtearing modes and anomalous transport in tokamaks. *Physical Review Letters*, 44(15):994, 1980.
- [37] J. F. Drake and Y. C. Lee. Kinetic theory of tearing instabilities. *The Physics of Fluids*, 20(8):1341–1353, 1977.
- [38] B. J. Faber, M. J. Pueschel, P. W. Terry, C. C. Hegna, and J. E. Roman. Stellarator microinstabilities and turbulence at low magnetic shear. *Journal of Plasma Physics*, 84(5):905840503, 2018.
- [39] M. E. Fenstermacher, L. R. Baylor, E. de la Luna, M. G. Dunne, G. T. A. Huijsmans, A. Kirk, F. M. Laggner, T. H. Osborne, C. Paz-Soldan, S. Saarelma, et al. Progress in pedestal and edge physics: Chapter 3 of the special issue: on the path to tokamak burning plasma operation. *Nuclear Fusion*, 65(5):053001, 2025.
- [40] J. P. Freidberg. *Ideal MHD*. Cambridge University Press, 2014.
- [41] E. A. Frieman and L. Chen. Nonlinear gyrokinetic equations for low-frequency electromagnetic waves in general plasma equilibria. *The Physics of Fluids*, 25(3):502–508, 1982.
- [42] D. A. Gates and L. Delgado-Aparicio. Origin of tokamak density limit scalings. *Physical Review Letters*, 108(16):165004, 2012.
- [43] J. Geiger, C. D. Beidler, Y. Feng, H. Maaßberg, N. B. Marushchenko, and Y. Turkin. Physics in the magnetic configuration space of w7-x. *Plasma Physics and Controlled Fusion*, 57(1):014004, 2014.

- [44] A. Geraldini, M. Landreman, and E. J. Paul. An adjoint method for determining the sensitivity of island size to magnetic field variations. *Journal of Plasma Physics*, 87(3):905870302, 2021.
- [45] A. González-Jerez, P. Xanthopoulos, J. M. García-Regaña, I. Calvo, J. Alcusón, A. B. Navarro, M. Barnes, F. I. Parra, and J. Geiger. Electrostatic gyrokinetic simulations in wendelstein 7-x geometry: benchmark between the codes stella and gene. *Journal of Plasma Physics*, 88(3):905880310, 2022.
- [46] A. G. Goodman, K. C. Mata, S. A. Henneberg, R. Jorge, M. Landreman, G. G. Plunk, H. M. Smith, R. J. J. Mackenbach, C. D. Beidler, and P. Helander. Constructing precisely quasi-isodynamic magnetic fields. *Journal of Plasma Physics*, 89(5):905890504, 2023.
- [47] S. Gottlieb, C. W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43(1):89–112, 2001.
- [48] 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. American Institute of Physics, 2006.
- [49] M. Greenwald. Density limits in toroidal plasmas. *Plasma Physics and Controlled Fusion*, 44(8):R27, 2002.
- [50] T. S. Hahm and W. M. Tang. Properties of ion temperature gradient drift instabilities in h-mode plasmas. Technical report, Princeton Plasma Physics Lab., 1988.
- [51] G. W. Hammett, M. A. Beer, W. Dorland, S. C. Cowley, and S. A. Smith. Developments in the gyrofluid approach to tokamak turbulence simulations. *Plasma Physics and Controlled Fusion*, 35(8):973, 1993.
- [52] P. Helander. Theory of plasma confinement in non-axisymmetric magnetic fields. *Reports on Progress in Physics*, 77(8):087001, 2014.
- [53] P. Helander, A. Mishchenko, R. Kleiber, and P. Xanthopoulos. Oscillations of zonal flows in stellarators. *Plasma Physics and Controlled Fusion*, 53(5):054006, 2011.

- [54] S. A. Henneberg, M. Drevlak, and P. Helander. Improving fast-particle confinement in quasi-axisymmetric stellarator optimization. *Plasma Physics and Controlled Fusion*, 62(1):014023, 2019.
- [55] S. A. Henneberg, M. Drevlak, C. Nührenberg, C. D. Beidler, Y. Turkin, J. Loizu, and P. Helander. Properties of a new quasi-axisymmetric configuration. *Nuclear Fusion*, 59(2):026014, 2019.
- [56] E. G. Highcock, M. Barnes, A. A. Schekochihin, F. I. Parra, C. M. Roach, and S. C. Cowley. Transport bifurcation in a rotating tokamak plasma. *Physical Review Letters*, 105(21):215003, 2010.
- [57] E. G. Highcock, N. R. Mandell, M. Barnes, and W. Dorland. Optimisation of confinement in a fusion reactor using a nonlinear turbulence model. *Journal of Plasma Physics*, 84(2):905840208, 2018.
- [58] F. L. Hinton. Collisional transport in plasma. *Handbook of Plasma Physics*, 1(147):331, 1983.
- [59] M. Hirsch, J. Baldzuhn, C. Beidler, R. Brakel, R. Burhenn, A. Dinklage, H. Ehmler, M. Endler, V. Erckmann, Y. Feng, et al. Major results from the stellarator wendelstein 7-as. *Plasma Physics and Controlled Fusion*, 50(5):053001, 2008.
- [60] S. P. Hirshman and D. K. Lee. Momcon: A spectral code for obtaining three-dimensional magnetohydrodynamic equilibria. *Computer Physics Communications*, 39(2):161–172, 1986.
- [61] S. P. Hirshman, P. Merkel, et al. Three-dimensional free boundary calculations using a spectral green’s function method. *Computer Physics Communications*, 43(1):143–155, 1986.
- [62] S. P. Hirshman and J. C. Whitson. Steepest descent moment method for three-dimensional magnetohydrodynamic equilibria. Technical report, Oak Ridge National Lab., 1983.
- [63] D. D-M Ho and R. M. Kulsrud. Neoclassical transport in stellarators. *The Physics of Fluids*, 30(2):442–461, 1987.

- [64] J. V. Hofmann, J. Baldzuhn, R. Brakel, Y. Feng, S. Fiedler, J. Geiger, P. Grigull, G. Herre, R. Jaenicke, M. Kick, et al. Stellarator optimization studies in w7-as. *Plasma Physics and Controlled Fusion*, 38(12A):A193, 1996.
- [65] G. G. Howes, S. C. Cowley, W. Dorland, G. W. Hammett, E. Quataert, and A. A. Schekochihin. Astrophysical gyrokinetics: basic equations and linear theory. *The Astrophysical Journal*, 651(1):590, 2006.
- [66] L.-M. Imbert-Gérard, E. J Paul, and A. M. Wright. *An introduction to stellarators: From magnetic fields to symmetries and optimization*. SIAM, 2024.
- [67] P. G. Ivanov, A. A. Schekochihin, and W. Dorland. Dimits transition in three-dimensional ion-temperature-gradient turbulence. *Journal of Plasma Physics*, 88(5):905880506, 2022.
- [68] P. G. Ivanov, A. A. Schekochihin, W. Dorland, A. R. Field, and F. I. Parra. Zonally dominated dynamics and dimits threshold in curvature-driven itg turbulence. *Journal of Plasma Physics*, 86(5):855860502, 2020.
- [69] F. Jenko. Electron-temperature-gradient-driven turbulence. In *APS Division of Plasma Physics Meeting Abstracts*, volume 41, pages FI1–06. American Physical Society, 1999.
- [70] F. Jenko. Massively parallel vlasov simulation of electromagnetic drift-wave turbulence. *Computer Physics Communications*, 125(1-3):196–209, 2000.
- [71] S. Jolliet and Y. Idomura. Simulating plasma turbulence with the global eulerian gyrokinetic code gt5d. *Progress in Nuclear Science and Technology*, 2:85–89, 2011.
- [72] R. Kleiber, M. Borchardt, R. Hatzky, A. Könies, H. Leyh, A. Mishchenko, J. Riemann, C. Slaby, J. M. García-Regaña, E. Sanchez, et al. Euterpe: A global gyrokinetic code for stellarator geometry. *Computer Physics Communications*, 295:109013, 2024.
- [73] M. Kotschenreuther, G. Rewoldt, and W. M. Tang. Comparison of initial value and eigenvalue codes for kinetic toroidal plasma instabilities. *Computer Physics Communications*, 88(2-3):128–140, 1995.

- [74] J. A. Krommes and G. Hu. The role of dissipation in the theory and simulations of homogeneous plasma turbulence, and resolution of the entropy paradox. *Physics of Plasmas*, 1(10):3211–3238, 1994.
- [75] M. D. Kruskal and R. M. Kulsrud. Equilibrium of a magnetically confined plasma in a toroid. Technical report, Princeton University Plasma Physics Laboratory, 1958.
- [76] S. H. Ku, R. Hager, A. Scheinberg, J. Dominski, A. Sharma, M. Churchill, J. Choi, B. Sturdevant, A. Mollén, G. Wilkie, C. S. Chang, E. Yoon, M. Adams, J. Seo, S. Koh, E. D’Azevedo, S. Abbott, P. H. Worley, S. Ethier, G. Park, J. Lang, B. MacKie-Mason, K. Germaschewski, E. Suchyta, V. Carey, M. Cole, P. Trivedi, and J. Chowdhury. Xgc, Jun 2018. Computer Software.
- [77] M. Landreman and W. Sengupta. Constructing stellarators with quasisymmetry to high order. *Journal of Plasma Physics*, 85(6):815850601, 2019.
- [78] E. Lanti, N. Ohana, N. Tronko, T. Hayward-Schneider, A. Bottino, B. F. McMillan, A. Mishchenko, A. Scheinberg, A. Biancalani, P. Angelino, et al. Orb5: a global electromagnetic gyrokinetic code using the pic approach in toroidal geometry. *Computer Physics Communications*, 251:107072, 2020.
- [79] B. F. Lee, E. J. Paul, G. Stadler, and M. Landreman. Stellarator coil optimization supporting multiple magnetic configurations. *Nuclear Fusion*, 63(1):014002, 2022.
- [80] W. W. Lee. Gyrokinetic particle simulation model. *Journal of Computational Physics*, 72(1):243–269, 1987.
- [81] Z. Lin, T. A. Hahm, W. W. Lee, W. M. Tang, and R. B. White. Turbulent transport reduction by zonal flows: Massively parallel simulations. *Science*, 281(5384):1835–1837, 1998.
- [82] J. F. Lobsien, M. Drevlak, T. S. Pedersen, et al. Stellarator coil optimization towards higher engineering tolerances. *Nuclear Fusion*, 58(10):106013, 2018.
- [83] N. R. Mandell, W. Dorland, I. Abel, R. Gaur, P. Kim, M. Martin, and T. Qian. Gx: a gpu-native gyrokinetic turbulence code for tokamak and stellarator design. *Journal of Plasma Physics*, 90(4):905900402, 2024.

- [84] P. Mantica, C. Angioni, B. Baiocchi, M. Baruzzo, M. N. A. Beurskens, J. P. S. Bizarro, R. V. Budny, P. Buratti, A. Casati, C. Challis, J. Citrin, G. Colyer, F. Crisanti, A. C. A. Figueiredo, L. Frassinetti, C. Giroud, N. Hawkes, J. Hobbirk, E. Joffrin, T. Johnson, E. Lerche, P. Migliano, V. Naulin, A. G. Peeters, G. Rewoldt, F. Ryter, A. Salmi, R. Sartori, C. Sozzi, G. Staebler, D. Strintzi, T. Tala, M. Tsalas, D. Van Eester, T. Versloot, P. C. deVries, J. Weiland, and JET. Ion heat transport studies in jet. *Plasma Physics and Controlled Fusion*, 53(12):124033, 2011.
- [85] M. Maurer, A. B. Navarro, T. Dannert, M. Restelli, F. Hindenlang, T. Görler, D. Told, D. Jarema, G. Merlo, and F. Jenko. Gene-3d: a global gyrokinetic turbulence code for stellarators. *Journal of Computational Physics*, 420:109694, 2020.
- [86] R. L. Miller, M. S. Chu, J. M. Greene, Y. R. Lin-Liu, and R. E. Waltz. Noncircular, finite aspect ratio, local equilibrium model. *Physics of Plasmas*, 5(4):973–978, 1998.
- [87] S. L. Newton, S. C. Cowley, and N. F. Loureiro. Understanding the effect of sheared flow on microinstabilities. *Plasma Physics and Controlled Fusion*, 52(12):125001, 2010.
- [88] R. Nies, F. I. Parra, M. Barnes, N. R. Mandell, and W. Dorland. Saturation of magnetised plasma turbulence by propagating zonal flows. arXiv:2409.02283, 2024.
- [89] R. Nies, E. J. Paul, S. R. Hudson, and A. Bhattacharjee. Adjoint methods for quasi-symmetry of vacuum fields on a surface. *Journal of Plasma Physics*, 88(1):905880106, 2022.
- [90] S. A. Orszag. On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *Journal of Atmospheric Sciences*, 28(6):1074–1074, 1971.
- [91] J. B. Parker, L. L. LoDestro, D. Told, G. Merlo, L. F. Ricketson, A. Campos, F. Jenko, and J. A. F. Hittinger. Bringing global gyrokinetic turbulence simulations to the transport timescale using a multiscale approach. *Nuclear Fusion*, 58(5):054004, 2018.

- [92] S. E. Parker, Y. Chen, W. Wan, B. I. Cohen, and W. M. Nevins. Electromagnetic gyrokinetic simulations. *Physics of Plasmas*, 11(5):2594–2599, 2004.
- [93] S. E. Parker, W. W. Lee, and R. A. Santoro. Gyrokinetic simulation of ion temperature gradient driven turbulence in 3d toroidal geometry. *Physical Review Letters*, 71(13):2042, 1993.
- [94] E. J. Paul, M. Landreman, A. Bader, and W. Dorland. An adjoint method for gradient-based optimization of stellarator coil shapes. *Nuclear Fusion*, 58(7):076015, 2018.
- [95] A. G. Peeters, Y. Camenen, F. J. Casson, W. A. Hornsby, A. P. Snodin, D. Strintzi, and G. Szepesi. The nonlinear gyro-kinetic flux tube code gkw. *Computer Physics Communications*, 180(12):2650–2672, 2009.
- [96] O. Pironneau. On optimum design in fluid mechanics. *Journal of Fluid Mechanics*, 64:97–110, 1974.
- [97] A. Pochelon, T. P. Goodman, M. Henderson, C. Angioni, R. Behn, S. Coda, F. Hofmann, J.-P. Hogge, N. Kirneva, A. A. Martynov, J.-M. Moret, Z. A. Pietrzyk, F. Porcelli, H. Reimerdes, J. Rommers, E. Rossi, O. Sauter, M. Q. Tran, H. Weisen, S. Alberti, S. Barry, P. Blanchard, P. Bosshard, R. Chavan, B. P. Duval, Y. V. Esipchuck, D. Fasel, A. Favre, S. Franke, I. Furno, P. Gorguerat, P.-F. Isoz, B. Joye, J. B. Lister, X. Llobet, J.-C. Magnin, P. Mandrin, A. Manini, B. Marlétaz, P. Marmillod, Y. Martin, J.-M. Mayor, J. Mlynar, C. Nieswand, P. J. Paris, A. Perez, R. A. Pitts, K. A. Razumova, A. Refke, E. Scavino, A. Sushkov, G. Tonetti, F. Troyon, W. Van Toledo, and P. Vyas. Energy confinement and mhd activity in shaped tcv plasmas with localized electron cyclotron heating. *Nuclear Fusion*, 39(11Y):1807–1818, 1999.
- [98] O. P. Pogutse. Magnetic drift instability in a collisionless plasma. *Plasma Physics*, 10(7):649, 1968.
- [99] T. Qian, B. Buck, R. Gaur, N. Mandell, P. Kim, and W. Dorland. Stellarator profile predictions using trinity3d and gx. In *APS Division of Plasma Physics Meeting Abstracts*, volume 2022, pages BO03–006. American Physical Society, 2022.

- [100] B. Rao, G. Wang, Y. H. Ding, K. X. Yu, Q. L. Li, N. C. Wang, B. Yi, J. Y. Nan, Y. S. Cen, Q. M. Hu, et al. Introduction to resonant magnetic perturbation coils of the j-text tokamak. *Fusion engineering and design*, 89(4):378–384, 2014.
- [101] E. Rodriguez, E. J. Paul, and A. Bhattacharjee. Measures of quasisymmetry for stellarators. *Journal of Plasma Physics*, 88(1):905880109, 2022.
- [102] B. N. Rogers, W. Dorland, and M. Kotschenreuther. Generation and stability of zonal flows in ion-temperature-gradient mode turbulence. *Phys. Rev. Lett.*, 85(25):5336–5339, 2000.
- [103] F. Romanelli. Ion temperature-gradient-driven modes and anomalous ion transport in tokamaks. *Physics of Fluids B: Plasma Physics*, 1(5):1018–1025, 1989.
- [104] M. N. Rosenbluth and F. L. Hinton. Poloidal flow driven by ion-temperature-gradient turbulence in tokamaks. *Physical Review Letters*, 80(4):724, 1998.
- [105] L. I. Rudakov and R. Z. Sagdeev. On the instability of a nonuniform rarefied plasma in a strong magnetic field. *Soviet Physics Doklady*, 6:415, 1961.
- [106] E. Sánchez, J. M. García-Regaña, A. B. Navarro, J. H. E. Proll, C. M. Moreno, A. González-Jerez, I. Calvo, R. Kleiber, J. Riemann, J. Smoniewski, et al. Gyrokinetic simulations in stellarators using different computational domains. *Nuclear Fusion*, 61(11):116074, 2021.
- [107] E. Sánchez, A. Mishchenko, J. M. García-Regaña, R. Kleiber, A. Bottino, L. Villard, et al. Nonlinear gyrokinetic pic simulations in stellarators with the code euterpe. *Journal of Plasma Physics*, 86(5):855860501, 2020.
- [108] V. D. Shafranov. Plasma equilibrium in a magnetic field. *Reviews of Plasma Physics*, 2:103, 1966.
- [109] D. A. St-Onge. A novel approach to radially global gyrokinetic simulation using the flux-tube code stella. *Journal of Computational Physics*, 468:111498, 2022.
- [110] D. A. St-Onge, M. Barnes, and F. I. Parra. A phase-shift-periodic parallel boundary condition for low-magnetic-shear scenarios. *Plasma Physics and Controlled Fusion*, 65(1):015016, 2022.
- [111] G. M. Staebler, J. Candy, N. T. Howard, and C. Holland. The role of zonal flows in the saturation of multi-scale gyrokinetic turbulence. *Physics of Plasmas*, 23(6), 2016.

- [112] A. Staniforth and J. Côté. Semi-lagrangian integration schemes for atmospheric models—a review. *Monthly Weather Review*.
- [113] Gilbert Strang. On the construction and comparison of difference schemes. *SIAM Journal on Numerical Analysis*, 5(3):506–517, 1968.
- [114] W. M. Tang, J. W. Connor, and R. J. Hastie. Kinetic-ballooning-mode theory in general geometry. *Nuclear Fusion*, 20(11):1439, 1980.
- [115] H Thienpondt, J. M. García-Regaña, G. Acton, and A González-Jerez. Stel-lapy: Python package for the gyrokinetic code stella. https://github.com/stellaGK/stella/tree/master/POST_PROCESSING/stellapy, 2022.
- [116] H. Thienpondt, J. M. García-Regaña, I. Calvo, G. Acton, and M. Barnes. Influence of the density gradient on turbulent heat transport at ion-scales: an inter-machine study with the gyrokinetic code stella. *Nuclear Fusion*, 65(1):016062, 2024.
- [117] L. H. Thomas. Elliptic problems in linear difference equations over a network. Technical report, Watson Scientific Computing Laboratory, Columbia University, 1949.
- [118] W. P. Thurston. *Three-Dimensional Geometry and Topology, Volume 1*. Princeton University Press, 1997.
- [119] M. K. Transtrum and J. P. Sethna. Improvements to the levenberg-marquardt algorithm for nonlinear least-squares minimization. arXiv:1201.5885, 2012.
- [120] R. E. Waltz. Three-dimensional global numerical simulation of ion temperature gradient mode turbulence. *Physics of Fluids*, 31(7), 1988.
- [121] W. X. Wang, Z. T. 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), 2006.
- [122] T. H. Watanabe and H. Sugama. Velocity-space structures of distribution function in toroidal ion temperature gradient turbulence. *Nuclear Fusion*, 46(1):24, 2005.

- [123] T. H. Watanabe, H. Sugama, A. Ishizawa, and M. Nunami. Flux tube train model for local turbulence simulation of toroidal plasmas. *Physics of Plasmas*, 22(2):022507, 2015.
- [124] A. Weller, S. Sakakibara, K. Y. Watanabe, K. Toi, J. Geiger, M. C. Zarnstorff, S. R. Hudson, A. Reiman, A. Werner, C. Nührenberg, et al. Significance of mhd effects in stellarator confinement. *Fusion science and technology*, 50(2):158–170, 2006.
- [125] A. Zocco, P. Xanthopoulos, H. Doerk, J. W. Connor, and P. Helander. Threshold for the destabilisation of the ion-temperature-gradient mode in magnetically confined toroidal plasmas. *Journal of Plasma Physics*, 84(1):715840101, 2018.