ABSTRACT

Title of dissertation: GYROFLUID MODELING OF TURBULENT, KINETIC PHYSICS

Kate Marie Despain, Doctor of Philosophy, 2011

Dissertation directed by: Dr William Dorland
Department of Physics

Gyrofluid models to describe plasma turbulence combine the advantages of fluid models, such as lower dimensionality and well-developed intuition, with those of gyrokinetics models, such as finite Larmor radius (FLR) effects. This allows gyrofluid models to be more tractable computationally while still capturing much of the physics related to the FLR of the particles.

We present a gyrofluid model derived to capture the behavior of slow solar wind turbulence and describe the computer code developed to implement the model. In addition, we describe the modifications we made to a gyrofluid model and code that simulate plasma turbulence in tokamak geometries. Specifically, we describe a nonlinear phase mixing phenomenon, part of the $\mathbf{E} \times \mathbf{B}$ term, that was previously missing from the model. An inherently FLR effect, it plays an important role in predicting turbulent heat flux and diffusivity levels for the plasma. We demonstrate this importance by comparing results from the updated code to studies done previously by gyrofluid and gyrokinetic codes. We further explain what would be

necessary to couple the updated gyrofluid code, `gryffin`, to a turbulent transport code, thus allowing `gryffin` to play a role in predicting profiles for fusion devices such as ITER and to explore novel fusion configurations. Such a coupling would require the use of Graphical Processing Units (GPUs) to make the modeling process fast enough to be viable. Consequently, we also describe our experience with GPU computing and demonstrate that we are poised to complete a `gryffin` port to this innovative architecture.

Gyrofluid Modeling of Turbulent, Kinetic Physics

by

Kate Marie Despain

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2011

Advisory Committee:
Dr. William Dorland, Chair/Advisor
Dr. Ramani Duraiswami, Dean's Representative
Dr. Thomas Antonsen
Dr. Adil Hassam
Dr. Edward Ott

Dedication

To my parents for years of support and love; I miss you, Dad

Table of Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation

Turbulence in plasmas is an ubiquitous phenomenon. It is observed in the largest scales man can probe, on the order of galaxies, and can be observed in terrestrial fusion devices in laboratories. In the heavens, an understanding of turbulence can potentially lead us to an understanding of high energy cosmic rays, of the process through which stars accrete, and of how stellar clusters form. On earth, a growing comprehension of turbulence has already aided mankind in the design of better fusion devices and configurations and can potentially lead us to the goal of sustained fusion energy.

Since turbulence extends through so many scales, several sets of approximations are needed to describe the entire system. When turbulent eddies are much larger than the Larmor radius of the particles that make up the plasma, a model that treats the plasma as a magnetized fluid, such as Magnetohydrodynamics (MHD), is an adequate description of the dynamics of the system. However, as turbulent structures approach the Larmor radius of the particles in question, MHD 's assumptions begin to break down. A model such as gyrokinetics, which accounts for the finite Larmor radius (FLR) effects of the particles, becomes necessary to accurately capture the behavior of the plasma.

Gyrokinetics was developed as a model to describe low frequency fluctuations in plasmas with gentle equilibrium gradient scale lengths compared to the particles' Larmor radii [32] [35] [25] [16]. It has been quite successful over the years in predicting turbulent behavior in both astrophysical and laboratory contexts [19] [14]. Gyrokinetics starts with a description of the evolution of the particle distribution function through the Fokker-Plank equation and adds assumptions to create a tractable equation. The result is five dimensional and retains much of the physics needed to describe turbulent systems. Despite this reduction, nonlinear gyrokinetic simulations can still take tens of thousands of computer hours to investigate a single parameter set or a single plasma discharge.

Gyrofluid models further simplify the gyrokinetic equation by considering only moment expansions of the distribution function. With carefully and cleverly chosen closures, they are able to capture FLR effects within a fluid framework, allowing theorists to maintain much of the intuition developed through working with MHD. At the same time, they are more computationally tractable as they have two fewer dimensions to evolve than the corresponding gyrokinetic model.

In the context of astrophysical plasmas, computational codes based on gyrofluid models are able to probe a larger span of spatial scales than their gyrokinetic counterparts for the same computational cost. This allows them to capture gross features, such as the Kolmogorov spectra found in the solar wind and interstellar medium, over a wider range of spatial scales. This also makes broad parameter studies of Kolmogorov spectra, including transition regions, more feasible. While the details of what causes transitions and shifts in astrophysical Kolmogorov spec-

tra is better explored by a gyrokinetic code, a gyrofluid code can be used to narrow the parameter space of interest, allowing for a more efficient use of computational resources.

In the context of fusion plasmas, gyrofluid codes have the potential to probe more of parameter space in less time. Again, the fine details of the calculation can be investigated with a gyrokinetic code, but a gyrofluid code can be used to find parameter ranges of potential interest or potential disaster, making the overall process of searching through parameter space more efficient.

In the past, gyrofluid models have not always agreed with turbulent flux predictions made by gyrokinetic models. There had been some concern that these models did not capture enough of the relevant physics to make them useful tools [12]. However, with the inclusion of closures that account for both zonal flows and nonlinear phase mixing by $\mathbf{E} \times \mathbf{B}$ drifts, we believe that these models can be resurrected and used to aid in the study of viable fusion configurations.

Fortunately, there are several studies that have been performed by multiple gyrofluid and gyrokinetic codes or multiple gyrokinetic codes that can be used as a benchmark for updated gyrofluid codes. One widely used example was written up by Dimits, et al and involved parameters from the Cyclone base case [12]. This study involved eight codes and found relatively good agreement for predicted heat flux at experimentally relevant parameters among the gyrokinetic codes used. Such wide spread benchmarking exercises are rare, but there exist other studies that can be used to validate a gyrofluid code against a gyrokinetic one.

The state of the art in gyrokinetic modeling goes one step further than predict-

ing heat fluxes and involves using gyrokinetic turbulence simulations in conjunction with transport equations to model and predict equilibrium density and temperature profiles for fusion devices. One such framework, `TRINITY`, employs several turbulence simulations for each point on a coarse space-time grid [2]. The latest reported transport simulation required roughly 25,000 computer hours and must be run on supercomputers. Such cost makes it impractical to run sensitivity studies or explore a large parameter space looking for innovative new fusion device designs.

Recently, however, GPUs have started to compete with supercomputers in the realm of scientific computing. State of the art GPUs have on the order of 500 cores designed to be used in parallel. If an algorithm is employed which requires little or no communication between GPUs, speed-ups have the potential to be significant - on the order of ten to a hundred times their CPU counterparts. Due to their five dimensional nature, well-resolved gyrokinetic simulations are too large to fit on a single GPU. However, a three dimensional gyrofluid code should be able to do so, allowing an entire `TRINITY` transport/turbulence calculation to be run on a single GPU cluster. With the appropriate combination of physics model, computational algorithm and hardware, a simulation of an entire tokamak (minus the edge) could be performed on an inexpensive, local CPU/GPU cluster.

1.2 Outline of Thesis

In Chapter 2, we give an overview of Kolomogorov energy spectra and describe the gyrofluid model we have developed to study plasma behavior in conditions sim-

ilar to those found in the solar wind. In Chapter 3, we describe the computer code we have developed to solve the nonlinear equations of the solar wind model and present our results. In Chapter 4, we transition to the toroidal, fusion perspective and describe the process of nonlinear phase mixing in the context of a well established gyrofluid model. In Chapter 5, we give an overview of the gyrofluid code `gryffin` used to study nonlinear phase mixing and present the results of several test cases where nonlinear phase mixing should bring gyrofluid turbulence flux predictions into better agreement with gyrokinetic predictions. In Chapter 6, we give an overview of the turbulent transport solver `TRINITY` and an overview of what would be required to incorporate `gryffin` into the `TRINITY` framework. Finally, in Chapter 7, we discuss our experience using GPUs for scientific computing and our plans to use that experience to port `gryffin` to the GPU as part of the `TRINITY` framework.

Chapter 2

Gyrofluid Model for the Slow Solar Wind

2.1 Introduction

In this chapter we seek to develop a model that accurately captures the turbulent physics of the slow solar wind. We begin by describing some of the properties of the slow solar wind. In Section 2.2 we describe the Kolmogorov power law, a common framework used to understand turbulent systems. In Sections 2.3 and 2.4 we introduce the model and closures we have developed to study the Kolmogorov Power spectrum of the slow solar wind. In Section 2.6 we describe the properties of the linear dispersion relation of our model. Finally, in Section 2.7 we compare our model to the Chew-Goldberger-Low (CGL) model.

The slow solar wind is a weakly collisional system. This is mainly due to the long mean free path of its particles - close to 1 AU. Ions and electrons have a nearly constant temperature that is approximately equal. The constant temperatures suggest an adiabatic system with at most weak heat fluxes. (For a detailed discussion of solar wind parameters, see [33] and references therein.) There is some evidence that Landau damping plays a role in solar wind dynamics. However, the effect appears to be roughly 10% or less [20].

Turbulence in the solar wind is anisotropic. That is, energy is not distributed evenly between the parallel and perpendicular directions. Instead, more energy can

be found in turbulent structures perpendicular to the background magnetic field. This suggests that the turbulence is the result of interacting Alfvén waves whose nonlinear interactions lead to an anisotropic distribution of energy, as explained in the next section.

2.2 Kolmogorov Power Laws

Due to its complexity, turbulence is difficult to characterize. Fortunately, in the early 1940s, Kolmogorov noticed a general relationship between turbulent energy and eddy scale size, often referred to as a power law spectrum or a Kolmogorov spectrum [26]. This relationship depends on the rate of energy transfer ε and the viscosity or damping mechanism. Assuming that the value of ε is set at some scale larger than that of interest and that damping only occurs at some scale smaller than that of interest, an "inertial range" can be defined where the details of the large-scale energy injection or stirring and the small-scale viscous damping mechanism do not affect the dynamics. The quantity ε is assumed to be constant in time, and energy in this inertial range is assumed to be only transferred locally, between eddies of similar sizes.

If in addition to the above assumptions we add homogeneity and isotropy, a power law spectrum for the turbulence can be found to within a constant using dimensional analysis. To make this easier, the relationship is considered in k-space, the inverse of the eddy scale size, and in logarithm space. Energy is binned at each k value. In log space, the width of each bin is given as $\Delta k \approx k$, so the total energy

in each bin can be expressed as kE_k.

In hydrodynamic flows, turbulent energy comes from the shearing of eddies within the flows. For an eddy being torn apart by oppositely moving shear flows, the difference in velocity Δu between a fluid element on one side of the eddy and the other side is related to the turbulent energy as $(\Delta u)^2 \sim kE_k$. The amount of time it takes for this eddy to break up depends on the velocity difference and the size of the eddy, $\tau \sim 1/(k\Delta u)$. Assuming that all the energy injected into the system becomes turbulent energy, we can write $\varepsilon \sim kE_k/\tau$, and solve for E_k.

$$E_k \sim \varepsilon^{2/3} k^{-5/3} \tag{2.1}$$

This relation gives us our power law spectrum.

Magnetohydrodynamic turbulent energy is equivalent to the energy found in perturbations to the magnetic fields (magnetic energy) and flows (kinetic energy.) Because all Alfvénic eddies travel along the magnetic field with the same phase and group velocities, Alfvénic turbulence is characterized by interacting, counter-propagating Alfvén waves with a dispersion relation $\omega = k_\parallel v_A$. Here, $k_\parallel$ is the parallel wave number and v_A is the Alfvén speed [27] [23]. [1]

As the Alfvén waves travel along the magnetic field lines and pass through each other, they interact nonlinearly in the direction perpendicular to the magnetic field. This nonlinear behavior acts as the "shear" that creates perpendicular perturbations - allowing energy to transfer to smaller scales in that direction. Eventually, the perpendicular interactions grow to the same scale as the parallel interactions, $k_\perp \delta u \sim$

[1] Since the Alfvén waves are an exact solution to the nonlinear system along the field lines, co-propagating Alfvén waves never have the opportunity to interact along the field lines.

$k_\parallel v_A$, where $k_\perp$ is the perpendicular wave number and δu is a typical perturbation to the flow. This state is described by Goldreich and Sridhar as "critical balance." [17] In critical balance, the turbulent interactions drive cascades of energy parallel and perpendicular to the background magnetic field, though $k_\parallel \ll k_\perp$.

In this critically balanced state, one can find the relationship between energy and perpendicular wave number by following an argument similar to the one in hydrodynamics. The magnetic and flow perturbations are assumed to be of the same order and are related to the turbulent energy as $(\delta u)^2 \sim k_\perp E_{k_\perp}$. Using the same assumption that our injected energy becomes turbulent energy, we are able to write

$$E_{k_\perp} \sim \varepsilon^{2/3} k_\perp^{-5/3}. \tag{2.2}$$

(Because $k_\parallel \ll k_\perp$, it is experimentally difficult to observe the steeper parallel fluctuation spectrum $E_{k_\parallel}$ and we choose not to focus on this quantity.)

At scales near the ion Larmor radius, the plasma no longer can be represented by a set of propagating Alfvén waves. At this scale, perpendicular structures affect the linear dynamics which can be described by a set of kinetic Alfvén waves (KAWs). They have a linear dispersion relation of the form, $\omega = k_\parallel v_A k_\perp \rho_i$ which allows both counter- and co-propagating KAWs to interact, since the group velocities of different wave packets now vary. The energy in the perturbed fields are no longer equivalent, $\delta u \sim k_\perp \delta b \sim \delta e$ where δb denotes a magnetic field perturbation and δe an electric field one. This dependence on $k_\perp$ affects the energy spectra as well, so that

$$E_{b,k_\perp} \sim \varepsilon^{2/3} k_\perp^{-7/3} \tag{2.3}$$

$$E_{e,k_\perp} \sim \varepsilon^{2/3} k_\perp^{-1/3} \tag{2.4}$$

This splitting between the magnetic and electric field energy spectra has been observed in the Solar Wind [1]. Nonlinear gyrokinetic simulations of the Alfvén/KAW transition have been performed [19]. We have developed a fluid description of this turbulence to allow larger and more convincing simulation-based studies (*e.g.,* with a wider intertial range).

2.3 Gyrofluid Model for Solar Wind Parameters

Gyrofluid equations are obtained by taking velocity moments of the gyrokinetic equation. The gyrokinetic equation is derived from the Fokker-Planck equation with the following assumptions: first, there is a strong background magnetic field, so that $\delta B/B_0 \ll 1$; second, the strong magnetic field guarantees that quantities of interest fluctuate more slowly than the ion gyroperiod; and, third, fluctuations with parallel wavelengths are much larger than the ion gyroradius. No such length assumption is made for the perpendicular direction. An ordering parameter, ϵ, is defined such that

$$\epsilon = \frac{\rho}{L} \tag{2.5}$$

where ρ is the gyroradius of the species of interest and L is a typical parallel wavelength of the system. A derivation of the gyrokinetic equation can be found in Appendix A. Our model includes equations for both ions and electrons. For both

species we start with Equation A.61.

$$\frac{\partial \langle \delta f_{1,s} \rangle}{\partial t} + \frac{q_s}{T_s} F_{0,s} v_{\parallel,s} \nabla_\parallel \langle \phi \rangle + \frac{q_s v_{\parallel,s}}{T_s B_0} F_{0,s} \left\langle \left[\phi, A_\parallel \right] \right\rangle - \left\langle \frac{q_s \delta B_\parallel}{T_s B_0} F_{0,s} \mathbf{v}_{\perp,s} \cdot \nabla_\perp \phi \right\rangle + v_{\parallel,s} \nabla_\parallel \langle \delta f_{1,s} \rangle$$
$$- \frac{c}{B_0} \langle [\delta f_{1,s}, \phi] \rangle + \frac{v_{\parallel,s}}{B_0} \left\langle \left[\delta f_{1,s}, A_\parallel \right] \right\rangle - \left\langle \frac{\delta B_\parallel}{B_0} \mathbf{v}_{\perp,s} \cdot \nabla_\perp \delta f_{1,s} \right\rangle + \frac{q_s}{c T_s} F_{0,s} v_{\parallel,s} \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t}$$
$$+ \frac{q_s}{c T_s} F_{0,s} \left\langle \mathbf{v}_{\perp,s} \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle + \left\langle \Omega_{c,s} \frac{\partial \delta f_{2,s}}{\partial \theta} \right\rangle = \langle C \left(\delta f_{1,s}, F_{0,s} \right) \rangle + \langle C \left(F_{0,s}, \delta f_{1,s} \right) \rangle \tag{2.6}$$

where s is the species label and the angle brackets, $\langle\rangle$, denote a gyroaverage, defined in Equation A.31. Square brackets, $[\,]$, indicate a Possion bracket, $[f, g] = \partial_x f \partial_y g - \partial_y f \partial_x g$. F_0 is the zeroth order pertrubation to the particle distribution function(pdf) and δf_1 is the first order perturbation to the pdf. The $\parallel$ and $\perp$ subscripts are used to denote directions with respect to the background magnetic field, B_0. The quantity q is charge, T is temperature (a constant), ϕ is the electrostatic potential, $\mathbf{v}$ is velocity, and ∇ denotes a derivative. Here, c is the speed of light, and $\mathbf{A}$ is the magnetic vector potential. The collisions are represented on the right hand side of the equation.

We start with the electrons and assume that $k_\perp \rho_e \ll 1$. This allows us to treat the gyro-radius of the electrons as a small perturbation, and we can Taylor expand around it.

$$\phi \left(\mathbf{R} - \frac{\mathbf{v} \times \hat{\mathbf{z}}}{\Omega_e}, t \right) = \phi \left(\mathbf{R}, t \right) - \frac{\partial \phi \left(\mathbf{R}, t \right)}{\partial \mathbf{R}} \cdot \left(\frac{\mathbf{v} \times \hat{\mathbf{z}}}{\Omega_e} \right) + \cdots \tag{2.7}$$

where ϕ is representative of any of the fields or the pdf. The first term in the expansion is independent of gyroangle. The second term depends on gyroangle through a single power $\mathbf{v}_\perp$. For terms in the equation with no additional θ dependence, the contribution from the second term in the expansion is zero. Applying this expansion

and the gyroaverages leaves

$$\frac{\partial \delta f_1}{\partial t} + v_\parallel \nabla_\parallel \delta f_1 + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \phi - \frac{c}{B_0}\left[\delta f_1, \phi\right] + \frac{v_\parallel}{B_0}\left[\delta f_1, A_\parallel\right] + \frac{q v_\parallel}{T B_0} F_0 \left[\phi, A_\parallel\right] + \frac{q}{cT} F_0 v_\parallel \frac{\partial A_\parallel}{\partial t}$$
$$- \frac{m v_\perp^2}{2 T B_0} F_0 \frac{\partial \delta B_\parallel}{\partial t} + \frac{m_e c v_\perp^2}{2 e B_0^2}\left[\delta f_1, \delta B_\parallel\right] - \frac{m_e c v_\perp^2}{2 T B_0^2} F_0 \left[\phi, \delta B_\parallel\right] = \left\langle C\left(\delta f_1, F_0\right)\right\rangle + \left\langle C\left(F_0, \delta f_1\right)\right\rangle \tag{2.8}$$

This version of the gyrokinetic equation is used to find the electron velocity moments.

We keep four electron moments for our model.

$$\frac{\partial}{\partial t}\left(\frac{n_e}{n_0} - \frac{\delta B_\parallel}{B_0}\right) + \frac{c}{B_0}\left[\phi, \frac{n_e}{n_0} - \frac{\delta B_\parallel}{B_0}\right] + \nabla_\parallel u_{\parallel,e} + \frac{1}{B_0}\left[u_{\parallel,e}, A_\parallel\right] + \frac{c}{e B_0}\left[\frac{p_{\perp,e}}{n_0}, \frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.9}$$

$$\frac{\partial}{\partial t}\left(u_{\parallel,e} - \frac{e}{c m_e} A_\parallel\right) + \frac{c}{B_0}\left[\phi, u_{\parallel,e} - \frac{e}{c m_e} A_\parallel\right] - \frac{e}{m_e}\nabla_\parallel \phi + \frac{1}{m_e}\nabla_\parallel \frac{p_{\parallel,e}}{n_0} + \frac{1}{m_e B_0}\left[\frac{p_{\parallel,e}}{n_0}, A_\parallel\right]$$
$$+ \frac{c}{e B_0}\left[\frac{q_{\perp,e}}{n_0}, \frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.10}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\parallel,e}}{n_0 T_{0,e}} - \frac{\delta B_\parallel}{B_0}\right) + \frac{c}{B_0}\left[\phi, \frac{p_{\parallel,e}}{n_0 T_{0,e}} - \frac{\delta B_\parallel}{B_0}\right] + \nabla_\parallel \frac{q_{\parallel,e}}{n_0 T_{0,e}}$$
$$+ \frac{1}{B_0}\left[\frac{q_{\parallel,e}}{n_0 T_{0,e}}, A_\parallel\right] + \frac{m_e c}{e B_0}\left[\frac{r_{\parallel,\perp,e}}{n_0 T_{0,e}}, \frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.11}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\perp,e}}{n_0 T_{0,e}} - 2\frac{\delta B_\parallel}{B_0}\right) + \frac{c}{B_0}\left[\phi, \frac{p_{\perp,e}}{n_0 T_{0,e}} - 2\frac{\delta B_\parallel}{B_0}\right] + \nabla_\parallel \frac{q_{\perp,e}}{n_0 T_{0,e}} + \frac{1}{B_0}\left[\frac{q_{\perp,e}}{n_0 T_{0,e}}, A_\parallel\right]$$
$$+ \frac{m_e c}{e B_0}\left[\frac{r_{\perp,\perp,e}}{n_0 T_{0,e}}, \frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.12}$$

For the ions, we again start with Equation A.61. We do not, however, use the Taylor expansion, since $k_\perp \rho_i$ is not always small for the regime in which we are interested. Instead, we can use Equation A.46 to replace δf_1 inside the nonlinear terms.

$$\frac{\partial \left\langle \delta f_1 \right\rangle}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \left\langle \phi \right\rangle + \frac{q v_\parallel}{T B_0} F_0 \left\langle \left[\phi, A_\parallel\right] \right\rangle - \left\langle \frac{q \delta B_\parallel}{T B_0} F_0 \mathbf{v}_\perp \cdot \nabla_\perp \phi \right\rangle + v_\parallel \nabla_\parallel \left\langle \delta f_1 \right\rangle$$
$$- \frac{c}{B_0}\left\langle \left[\delta f_{1,h} - \frac{q}{T} F_0 \phi, \phi\right] \right\rangle + \frac{v_\parallel}{B_0}\left\langle \left[\delta f_{1,h} - \frac{q}{T} F_0 \phi, A_\parallel\right] \right\rangle - \left\langle \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \left(\delta f_{1,h} - \frac{q}{T} F_0 \phi\right) \right\rangle$$
$$+ \frac{q}{cT} F_0 v_\parallel \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t} + \frac{q}{cT} F_0 \left\langle \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle + \left\langle \Omega_c \frac{\partial \delta f_2}{\partial \theta} \right\rangle = \left\langle C\left(\delta f_1, F_0\right)\right\rangle + \left\langle C\left(F_0, \delta f_1\right)\right\rangle \tag{2.13}$$

We assume that the gyroaverage of the last term on the right hand side also evaluates to zero as it has for the previous orders. If we remove terms that cancel out and terms that evaluate to zero, we are left with

$$\frac{\partial \langle \delta f_1 \rangle}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \langle \phi \rangle + v_\parallel \nabla_\parallel \langle \delta f_1 \rangle - \frac{c}{B_0} \langle [\delta f_{1,h}, \phi] \rangle + \frac{v_\parallel}{B_0} \left\langle \left[\delta f_{1,h}, A_\parallel \right] \right\rangle + \frac{1}{B_0} \langle [\delta f_{1,h}, \mathbf{v}_\perp \cdot \mathbf{A}_\perp] \rangle$$
$$+ \frac{q}{cT} F_0 v_\parallel \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t} + \frac{q}{cT} F_0 \left\langle \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle = \langle C(\delta f_1, F_0) \rangle + \langle C(F_0, \delta f_1) \rangle \tag{2.14}$$

We can now exploit the fact that $\delta f_{1,h}$ is independent of gyroangle and move the gyroaveraging operator in the nonlinear terms to just the fields. If we then use Equation A.48 to replace $\delta f_{1,h}$ (introduced by using Equation A.46), and take the appropriate gyroaverages, we can write

$$\frac{\partial \langle \delta f_1 \rangle}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel J_0 \phi + v_\parallel \nabla_\parallel \langle \delta f_1 \rangle - \frac{c}{B_0} \left[\langle \delta f_1 \rangle, J_0 \phi \right] + \frac{v_\parallel}{B_0} \left[\langle \delta f_1 \rangle, J_0 A_\parallel \right]$$
$$+ \frac{q v_\parallel}{T B_0} F_0 \left[J_0 \phi, J_0 A_\parallel \right] - \left[\langle \delta f_1 \rangle, \frac{v_\perp^2}{\Omega_i} \frac{J_1}{k_\perp v_\perp / \Omega_i} \frac{\delta B_\parallel}{B_0} \right] - \frac{q}{T} F_0 \left[J_0 \phi, \frac{v_\perp^2}{\Omega_i} \frac{J_1}{k_\perp v_\perp / \Omega_i} \frac{\delta B_\parallel}{B_0} \right]$$
$$+ \frac{q}{cT} F_0 v_\parallel J_0 \frac{\partial A_\parallel}{\partial t} - \frac{q}{cT} F_0 \frac{v_\perp^2}{\Omega_i} \frac{J_1}{k_\perp v_\perp / \Omega_i} \frac{\partial \delta B_\parallel}{\partial t} = \langle C(\delta f_1, F_0) \rangle + \langle C(F_0, \delta f_1) \rangle \tag{2.15}$$

where the operator J_0 is the zeroth order Bessel function of the first kind in k-space, and J_1 is the first order Bessel function of the first kind. Both have an argument of $k_\perp v_\perp / \Omega_i$.

This dependence on both space and velocity coordinates makes it impossible to integrate the nonlinear terms without knowing expressions for the fields and the pdf. In other words, the Bessel function operators do not commute with the Poisson bracket or the velocity integral. Instead of trying to integrate this equation exactly, we approximate the effect of the Bessel functions. We replace J_0 and $J_1 / (k_\perp v_\perp / \Omega_i)$ with appropriate approximations denoted $\tilde{J}_0$ (Equation 2.30) and $\tilde{J}_1$ (Equation 2.31),

that are independent of velocity but not space. This allows us to integrate over velocity.

Since this equation expresses the evolution of the *gyroaveraged* ion pdf, the corresponding moments are for *gyroaveraged* density, fluid flow, pressure, etc. We initially keep four moments for the ions as well.

$$\frac{\partial}{\partial t}\left(\frac{\bar{n}_i}{n_0} - 2\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right) + \frac{c}{B_0}\left[\tilde{J}_0\phi, \frac{\bar{n}_i}{n_0} - 2\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] + \nabla_\parallel \bar{u}_{\parallel,i} + \frac{1}{B_0}\left[\bar{u}_{\parallel,i}, \tilde{J}_0 A_\parallel\right] \\ -2\frac{c}{eB_0}\left[\frac{\bar{p}_{\perp,i}}{n_0}, \tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.16}$$

$$\frac{\partial}{\partial t}\left(\bar{u}_{\parallel,i} + \frac{e}{m_i c}\tilde{J}_0 A_\parallel\right) + \frac{c}{B_0}\left[\tilde{J}_0\phi, \bar{u}_{\parallel,i} + \frac{e}{m_i c}\tilde{J}_0 A_\parallel\right] + \frac{e}{m_i}\nabla_\parallel \tilde{J}_0\phi + \frac{1}{m_i}\nabla_\parallel\frac{\bar{p}_{\parallel,i}}{n_0} + \frac{1}{m_i B_0}\left[\frac{\bar{p}_{\parallel,i}}{n_0}, \tilde{J}_0 A_\parallel\right] \\ -2\frac{c}{eB_0}\left[\bar{q}_{\perp,i}, \tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.17}$$

$$\frac{\partial}{\partial t}\left(\frac{\bar{p}_{\parallel,i}}{n_0 T_{0,i}} - 2\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right) + \frac{c}{B_0}\left[\tilde{J}_0\phi, \frac{\bar{p}_{\parallel,i}}{n_0 T_{0,i}} - 2\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] + \nabla_\parallel\frac{\bar{q}_{\parallel,i}}{n_0 T_{0,i}} + \frac{1}{B_0}\left[\frac{\bar{q}_{\parallel,i}}{n_0 T_{0,i}}, \tilde{J}_0 A_\parallel\right] \\ -2\frac{m_i c}{eB_0}\left[\frac{\bar{r}_{\parallel,\perp,i}}{n_0 T_{0,i}}, \tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.18}$$

$$\frac{\partial}{\partial t}\left(\frac{\bar{p}_{\perp,i}}{n_0 T_{0,i}} - 4\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right) + \frac{c}{B_0}\left[\tilde{J}_0\phi, \frac{\bar{p}_{\perp,i}}{n_0 T_{0,i}} - 4\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] + \nabla_\parallel\frac{\bar{q}_{\perp,i}}{n_0 T_{0,i}} + \frac{1}{B_0}\left[\frac{\bar{q}_{\perp,i}}{n_0 T_{0,i}}, \tilde{J}_0 A_\parallel\right] \\ -2\frac{m_i c}{eB_0}\left[\frac{r_{\perp,\perp,i}}{n_0 T_{0,i}}, \tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] = 0 \tag{2.19}$$

where the overbar indicates a gyroaveraged moment. For example,

$$\bar{n}_i = \int_{-\infty}^{\infty} \langle \delta f_1 \rangle \, d^3 v \tag{2.20}$$

where $\langle \delta f_1 \rangle$ is defined in Equation A.48. In addition to the particle moments, we will need to use field equations to close our system. We start with the gyrokinetic field equations (see Section A.5) and use our definitions of the moments to evaluate the velocity integrals.

We start with the gyrokinetic version of Poisson's equation, Equation A.70.

$$\int_{-\infty}^{\infty} \left(q_e F_{0,e} + q_e \delta f_{1,e}\right) d^3 v_e + \int_{-\infty}^{\infty} \left(q_i F_{0,i} + q_i \left\langle \delta f_{1,i} \right\rangle + \frac{q_i^2}{T_{0,i}} F_0 \left\langle \phi \right\rangle - \frac{q_i^2}{T_{0,i}} F_0 \phi \right) d^3 v_i = 0 \tag{2.21}$$

Evaluating this in particle position coordinates gives

$$-e n_{0,e} - e n_e + e n_{0,i} + e \tilde{J}_0 \bar{n}_i + \frac{e^2 n_{0,i}}{T_{0,i}} \left(\tilde{J}_0^2 - 1 \right) \phi = 0 \tag{2.22}$$

This must be satisfied order by order to maintain quasineutrality. We assume quasineutrality and get $n_{0,e} = n_{0,i} = n_0$. And

$$-\frac{n_e}{n_0} + \tilde{J}_0 \frac{\bar{n}_i}{n_0} + \frac{e}{T_{0,i}} \left(\tilde{J}_0^2 - 1 \right) \phi = 0 \tag{2.23}$$

For the parallel part of Ampere's Law, we start with Equation A.83. For our model, this becomes

$$\begin{aligned} -\frac{c}{4\pi} \nabla_\perp^2 A_\parallel &= \int_{-\infty}^{\infty} \left(q_e v_{\parallel,e} F_{0,e} + q_e v_{\parallel,e} \delta f_{1,e} \right) d^3 v_e \\ &\quad + \int_{-\infty}^{\infty} \left(q_i v_{\parallel,i} F_{0,i} + q_i v_{\parallel,i} \left\langle \delta f_{1,i} \right\rangle + \frac{q_i^2}{T_{0,i}} v_{\parallel,i} F_{0,i} \left\langle \phi \right\rangle - \frac{q^2}{T_{0,i}} v_{\parallel,i} F_{0,i} \phi \right) d^3 v \\ &= -e n_0 u_{\parallel,e} + e n_0 \tilde{J}_0 \bar{u}_{\parallel,i} \end{aligned}$$

or more succinctly

$$-\frac{c}{4\pi e n_0} \nabla_\perp^2 A_\parallel = \tilde{J}_0 \bar{u}_{\parallel,i} - u_{\parallel,e} \tag{2.24}$$

And for the perpendicular part of Ampere's Law, we start with Equation A.85.

$$\begin{aligned}
\frac{c}{4\pi}\nabla_\perp^2 \delta B_\parallel &= \nabla_\perp \cdot \int_{-\infty}^{\infty} q_e \left\langle \hat{\mathbf{z}} \times \mathbf{v}_{\perp,e} \left(F_{0,e} + \delta f_{1,e}\right)\right\rangle d^3 v_e \\
&+ \nabla_\perp \cdot \int_{-\infty}^{\infty} q_i \left\langle \hat{\mathbf{z}} \times \mathbf{v}_{\perp,i} \left(F_{0,i} + \langle \delta f_{1,i}\rangle + \frac{q_i}{T_{0,i}} F_{0,i} \langle \phi \rangle - \frac{q_i}{T_{0,i}} F_{0,i} \phi\right)\right\rangle d^3 v_i \\
&= \frac{q_e k_\perp^2}{2\Omega_e} \int_{-\infty}^{\infty} v_{\perp,e}^2 \left(\frac{q_e}{T_{0,e}} F_{0,e}\phi + \delta f_{1,e}\right) d^3 v_e \\
&+ \frac{q_i k_\perp^2}{\Omega_i} \int_{-\infty}^{\infty} v_{\perp,i}^2 \frac{J_1\left(\frac{k_\perp v_\perp}{\Omega_i}\right)}{k_\perp v_\perp/\Omega_i} \left(\langle \delta f_{1,i}\rangle + \frac{q_i}{T_{0,i}} F_{0,i} \langle \phi \rangle\right) d^3 v_i \\
&= \frac{q_e k_\perp^2}{m_e \Omega_e} \left(-n_0 e\phi + p_{\perp,e}\right) + \frac{q_i k_\perp^2}{\Omega_i} \tilde{J}_1 \left(\frac{2\bar{p}_{\perp,i}}{m_i} + \frac{2 n_0 q_i T_{0,i}}{T_{0,i} m_i} \tilde{J}_0 \phi\right) \\
&= -\nabla_\perp^2 \frac{c}{B_0} \left(p_{\perp,e} - n_0 e\phi\right) - \nabla_\perp^2 \frac{m_i c}{B_0} \tilde{J}_1 \left(\frac{2\bar{p}_{\perp,i}}{m_i} + \frac{2 n_0 q_i}{m_i} \tilde{J}_0 \phi\right)
\end{aligned}$$

again succinctly as

$$\frac{\delta B_\parallel}{B_0} = -\frac{4\pi}{B_0^2}\left(p_{\perp,e} - n_0 e\phi\right) - \frac{8\pi}{B_0^2} \tilde{J}_1 \left(\bar{p}_{\perp,i} + n_0 e \tilde{J}_0 \phi\right) \tag{2.25}$$

2.4 Closures

Fluid models are based on the idea that the bulk behavior of a group of particles, which can be described by a particle distribution function, can be captured as meaningful quantities by taking the average behavior of the whole group [34]. This bulk behavior is broken into an infinite number of quantities - density, mean velocity, pressure, heat flux, etc. If taken all together, these moments would exactly capture the dynamics of the particles as described by the Fokker-Planck equation. However, to avoid keeping track of a potentially infinite number of moments, higher order moments are often assumed to describe only small to negligible corrections to the overall behavior of the system.

On examination of the evolution equations for the moments, one finds that each moment needs information from a higher order moment to accurately capture its behavior. In order to close the system, one must determine how many moments to keep and how to approximate the effect of moments not evolved.

Here we are considering the case of the slow solar wind which is nearly adiabatic. This allows us to close the system of equations with the assumption that $q = 0$. This closure is somewhat brutal and does not allow us to capture Landau damping. However, it does allow us to capture the correct linear kinetic response in the fluid limit of large ξ_s where $\xi_s = \frac{\omega}{\sqrt{2}k_\parallel v_{t,s}}$. Particularly at low β, this is a good approximation since $v_A > v_{ti}, v_{te}$. For higher values of β the long mean free path closures of Hammett and co-workers could be employed to model the small parallel heat fluxes [18] [13]. Here, we are focused on understanding the consequences of ion gyration on fluctuations with perpendicular wavelengths in the range of the thermal ion Larmor radius.

In addition, we assume that $\bar{u}_{\parallel,i} = 0$. If we were to combine the $\bar{u}_{\parallel,i}$ evolution equation with the $u_{\parallel,e}$ evolution equation to investigate how $A_\parallel$ evolves in time, we find that the contribution from the $\bar{u}_{\parallel,i}$ equation would be smaller than that of the $u_{\parallel,e}$ by a factor of the electron to ion mass ratio. We choose not to keep the $\bar{u}_{\parallel,i}$ evolution equation and simply allow $\bar{u}_{\parallel,i} = 0$.

We choose $\tilde{J}_0$ and $\tilde{J}_1$ so that the ion moments capture the corresponding linear kinetic behavior from the gyrokinetic model. We solve for the linear version of the field equations in the gyrokinetics framework in Section A.5. We can compare those results to the linear version of our field equations in the gyrofluid framework to

choose appropriate approximations.

If we linearize equations 2.9 - 2.12 and 2.16 - 2.19, we can express each of the moments in terms of fields. This allows us to eliminate the moments from our field equations in preparation for comparing them to results from gyrokinetics. Poisson's equation becomes

$$\begin{aligned}&\left[1-\tilde{J}_0^2\left(1+\frac{1}{2\xi_i^2}\right)-\frac{T_{0,i}}{T_{0,e}}\frac{1}{2\xi_e^2}\right]\phi+\frac{\omega}{ck_\parallel}\left[\frac{1}{2\xi_i^2}\tilde{J}_0^2+\frac{T_{0,i}}{T_{0,e}}\frac{1}{2\xi_e^2}\right]A_\parallel\\&-\frac{T_{0,i}}{e}\left[2\left(1+\frac{1}{2\xi_i^2}\right)\tilde{J}_1\tilde{J}_0-\left(1+\frac{1}{2\xi_e^2}\right)\right]\frac{\delta B_\parallel}{B_0}=0\end{aligned}\tag{2.26}$$

Because we have made the assumption that $\bar{u}_{\parallel,i}=0$, we do not need to consider the parallel part of Ampere's Law. However, the perpendicular part tells us

$$\begin{aligned}&-\left[2\tilde{J}_1\tilde{J}_0\left(1+\frac{1}{2\xi_i^2}\right)-\left(1+\frac{1}{2\xi_e^2}\right)\right]\phi+\frac{\omega}{ck_\parallel}\left[2\tilde{J}_1\tilde{J}_0\frac{1}{2\xi_i^2}-\frac{1}{2\xi_e^2}\right]A_\parallel\\&-\frac{2T_{0,i}}{e}\left[4\tilde{J}_1^2\left(1+\frac{1}{2\xi_i^2}\right)+\frac{T_{0,e}}{T_{0,i}}\left(1+\frac{1}{2\xi_e^2}\right)+\frac{1}{\beta_i}\right]\frac{\delta B_\parallel}{B_0}=0\end{aligned}\tag{2.27}$$

In order to compare equations 2.26 and 2.27 to gyrokinetic results, we need to make a few additional assumptions for the gyrokinetic equations. First, we are in a limit where $k_\perp^2\rho_e^2\ll 1$. This allows us to write $\Gamma_{0,e}\approx 1-k_\perp^2\rho_e^2$ and $\Gamma_{1,e}\approx 1-3/2k_\perp^2\rho_e^2$ (see Equations A.66 and A.67). We also take the large argument limit of the plasma dispersion function, $Z\left(\xi\gg 1\right)\approx -\xi^{-1}-2^{-1}\xi^{-3}$. This is consistent with the long mean free path assumption - any short scale perturbations in the parallel direction are eliminated quickly by free streaming particles. Using these assumptions, we can write the gyrokinetic version of the equations as

$$\begin{aligned}&\left[\left(1-\Gamma_{0,i}\left(1+\frac{1}{2\xi_i^2}\right)\right)-\frac{T_{0,i}}{T_{0,e}}\frac{1}{2\xi_e^2}\right]\phi+\frac{\omega}{ck_\parallel}\left[\Gamma_{0,i}\frac{1}{2\xi_i^2}+\frac{T_{0,i}}{T_{0,e}}\frac{1}{2\xi_e^2}\right]A_\parallel\\&+\frac{T_{0,i}}{e}\left[-\Gamma_{1,i}\left(1+\frac{1}{2\xi_i^2}\right)+\left(1+\frac{1}{2\xi_e^2}\right)\right]\frac{\delta B_\parallel}{B_0}=0\end{aligned}\tag{2.28}$$

$$
\begin{aligned}
&-\left[\Gamma_{1,i}\left(1+\frac{1}{2\xi_i^2}\right)-\left(1+\frac{1}{2\xi_e^2}\right)\right]\phi+\frac{\omega}{ck_\parallel}\left[\Gamma_{1,i}\frac{1}{2\xi_i^2}-\frac{1}{2\xi_e^2}\right]A_\parallel \\
&-\frac{2T_{0,i}}{e}\left[\Gamma_{1,i}\left(1+\frac{1}{2\xi_i^2}\right)+\frac{T_{0,e}}{T_{0,i}}\left(1+\frac{1}{2\xi_e^2}\right)+\frac{1}{\beta_i}\right]\frac{\delta B_\parallel}{B_0}=0
\end{aligned}
\tag{2.29}
$$

Comparing the two sets of equations allows us to fix approximations for $\tilde{J}_0$ and $\tilde{J}_1$:

$$\tilde{J}_0^2=\Gamma_{0,i}\left(k_\perp^2\rho_i^2\right)=e^{k_\perp^2\rho_i^2}I_0\left(k_\perp^2\rho_i^2\right) \tag{2.30}$$

$$\tilde{J}_1^2=\frac{1}{4}\Gamma_{1,i}\left(k_\perp^2\rho_i^2\right)=e^{k_\perp^2\rho_i^2}\left(I_0\left(k_\perp^2\rho_i^2\right)-I_1\left(k_\perp^2\rho_i^2\right)\right) \tag{2.31}$$

where I_0 and I_1 are the zeroth and first order modified Bessel functions of the first kind, respectively.

2.5 Final Equations

We use our closure assumptions from the previous section to write a final form for our equations. In addition, we note that the $u_{\parallel,e}$ terms found in the second and third electron moments will be smaller than the rest of the terms in their equations by a factor of the electron to ion mass ratio. We will drop these terms as well.

$$\frac{d}{dt}\left(\frac{n_e}{n_0}-\frac{\delta B_\parallel}{B_0}\right)=-\nabla_\parallel u_{\parallel,e}-\frac{1}{B_0}\left[u_{\parallel,e},A_\parallel\right]-\frac{c}{eB_0}\left[\frac{p_{\perp,e}}{n_0},\frac{\delta B_\parallel}{B_0}\right] \tag{2.32}$$

$$\frac{dA_\parallel}{dt}=-c\nabla_\parallel\phi+\frac{c}{e}\nabla_\parallel\frac{p_{\parallel,e}}{n_0}+\frac{c}{eB_0}\left[\frac{p_{\parallel,e}}{n_0},A_\parallel\right] \tag{2.33}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\parallel,e}}{n_0T_{0,e}}-\frac{\delta B_\parallel}{B_0}\right)=0 \tag{2.34}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\perp,e}}{n_0T_{0,e}}-2\frac{\delta B_\parallel}{B_0}\right)=0 \tag{2.35}$$

$$\frac{d}{dt}\left(\frac{\bar{n}_i}{n_0}-2\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right)=2\frac{c}{eB_0}\left[\frac{\bar{p}_{\perp,i}}{n_0},\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right] \tag{2.36}$$

$$\frac{d}{dt}\left(\frac{\bar{p}_{\perp,i}}{n_0T_{0,i}}-4\tilde{J}_1\frac{\delta B_\parallel}{B_0}\right)=0 \tag{2.37}$$

where for the electrons, $\frac{d}{dt} = \frac{\partial}{\partial t} + [\phi,]$ and for the ions, $\frac{d}{dt} = \frac{\partial}{\partial t} + \left[\tilde{J}_0\phi, \right]$.

$$\frac{e}{T_{0,i}}\left(\tilde{J}_0^2 - 1\right)\phi = \frac{n_e}{n_0} - \tilde{J}_0\frac{\bar{n}_i}{n_0} \tag{2.38}$$

$$\frac{c}{4\pi e n_0}\nabla_\perp^2 A_\parallel = u_{\parallel,e} \tag{2.39}$$

$$\frac{\delta B_\parallel}{B_0} = -\frac{4\pi}{B_0^2}\left(p_{\perp,e} - n_0 e\phi\right) - \frac{8\pi}{B_0^2}\tilde{J}_1\left(\bar{p}_{\perp,i} + n_0 e\tilde{J}_0\phi\right) \tag{2.40}$$

2.6 Linear Dispersion Relation

In ensure that our model captures the dynamics of Alfvén waves and kinetic Alfvén waves, we examine the linear dispersion relation of our system of equations. When linearized, they can be written as

$$\begin{pmatrix} \omega\left(\tilde{J}_0^2 - 1\right) & k_\parallel v_A^2 k_\perp^2 \rho_i^2 & \omega\left(2\tilde{J}_0\tilde{J}_1 - 1\right) \\ k_\parallel & -\omega & -\frac{1}{\tau}k_\parallel \\ 2\tilde{J}_1\tilde{J}_0 - 1 & 0 & \frac{2}{\beta_i} + \frac{2}{\tau} + 8\tilde{J}_1^2 \end{pmatrix} \begin{pmatrix} \frac{e}{T_{0,i}}\phi \\ \frac{e}{cT_{0,i}}A_\parallel \\ \frac{\delta B_\parallel}{B_0} \end{pmatrix} = 0 \tag{2.41}$$

This leads to a dispersion relation of the form

$$\omega^2 = k_\parallel^2 v_A^2 k_\perp^2 \rho_i^2 \frac{\left(2\left(\tau + \beta_i\tau\Gamma_{1,i}\right) + \beta_i\left(1 + \Gamma_{1,i}\right)\right)}{\beta_i\tau\left(1 - \Gamma_{1,i}\right)^2 - 2\left(\Gamma_{0,i} - 1\right)\left(\tau + \beta_i + \beta_i\tau\Gamma_{1,i}\right)} \tag{2.42}$$

where we have replaced $\tilde{J}_0$ and $\tilde{J}_1$ with their approximations. In the limit where $k_\perp^2\rho_i^2 \ll 1$, $\Gamma_0 \to 1 - 1/2k_\perp^2\rho_i^2$ and $\Gamma_1 \to 1 - 3/2k_\perp^2\rho_i^2$. We expect Alfvén waves to dominate in this regime. Our dispersion relation becomes

$$\omega^2 = k_\parallel^2 v_A^2 \tag{2.43}$$

which recovers the Alfvén wave dispersion relation exactly. In the $k_\perp^2\rho_i^2 \gg 1$ limit, $\Gamma_0, \Gamma_1 \to 0$. Our dispersion relation becomes

$$\omega^2 = k_\parallel^2 v_A^2 k_\perp^2 \rho_i^2 \frac{2\tau + \beta_i}{2\tau + 2\beta_i + 2\tau\beta_i} \tag{2.44}$$

This recovers the kinetic Alfvén wave dispersion relation for small β_i and $\tau = 1$. Though it does not exactly recover the KAW behavior, it does allow for a qualitative model that can begin to capture finite $\delta B_\parallel$ effects.

2.7 Chew-Goldberger-Low Closure

The assumption that $q = 0$ is reminiscent of the closure used by Chew, Goldberger, and Low for their fluid equations, commonly referred to as the CGL equations [9]. They argue that a consequence of setting $q = 0$ is that the pressure obeys the following relations

$$\frac{d}{dt}\frac{P_\parallel B^2}{n^3} = 0 \tag{2.45}$$

$$\frac{d}{dt}\frac{P_\perp}{nB} = 0 \tag{2.46}$$

where n is density, B is the magnitude of the magnetic field, and P is the pressure.

If we consider the ions, remove the assumption that their fluid velocity is zero, and take the $k_\perp \rho_i \ll 1$ limit, we can write their continuity and pressure equations as

$$\frac{\partial}{\partial t}\left(\frac{n_i}{n_0} - \frac{\delta B_\parallel}{B_0}\right) + \nabla_\parallel u_{\parallel,i} = 0 \tag{2.47}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\parallel,i}}{n_0 T_{0,i}} - \frac{\delta B_\parallel}{B_0}\right) + 3\nabla_\parallel u_{\parallel,i} = 0 \tag{2.48}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\perp,i}}{n_0 T_{0,i}} - 2\frac{\delta B_\parallel}{B_0}\right) + \nabla_\parallel u_{\parallel,i} = 0 \tag{2.49}$$

Using the continuity equation to replace the $\nabla u_{\parallel,i}$ terms in the pressure equations

allows us to write them as

$$\frac{\partial}{\partial t}\left(\frac{p_{\parallel,i}}{n_0 T_{0,i}} + 2\frac{\delta B_\parallel}{B_0} - 3\frac{n_i}{n_0}\right) = 0 \tag{2.50}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\perp,i}}{n_0 T_{0,i}} - \frac{\delta B_\parallel}{B_0} - \frac{n_i}{n_0}\right) = 0 \tag{2.51}$$

This can be further simplified to

$$\frac{\partial}{\partial t}\left(\frac{p_{\parallel,i}\delta B_\parallel^2}{n_i^3}\right) = 0 \tag{2.52}$$

$$\frac{\partial}{\partial t}\left(\frac{p_{\perp,i}}{\delta n_i B_\parallel}\right) = 0 \tag{2.53}$$

which is consistent with the CGL approximation.

Chapter 3

Numerics and Computational Results

3.1 Introduction

In this chapter, we describe the code we developed to study the kinetic physics in the slow solar wind. In Section 3.2, we report the normalizations used, and in Section 3.3 we describe the numerical algorithms employed. In Section 3.4, we discuss our verification procedures, and in Section 3.5 we report results from the initial nonlinear runs.

3.2 Normalizations

Because we are interested in studying the behavior of Alfvén and kinetic Alfvén waves, we have chosen to normalize our equations to the Alfvén speed, v_A, and the ion Larmor radius, ρ_i. Relevant quantities can be expressed as

$$\begin{aligned} \hat{n}_s = \frac{n_s}{n_0} \quad ; \quad \hat{u}_{\parallel,s} = \frac{u_{\parallel,s}}{v_A} \quad ; \quad \hat{p}_s = \frac{p_s}{n_0 T_{0,i}} \quad ; \quad \hat{\nabla} = \rho_i \nabla \\ \hat{\phi} = \frac{c}{v_A B_0 \rho_i}\phi \quad ; \quad \hat{A}_\parallel = \frac{1}{B_0 \rho_i} A_\parallel \quad ; \quad \delta\hat{B}_\parallel = \frac{\delta B_\parallel}{B_0} \quad ; \quad \hat{t} = \frac{v_A}{\rho_i} t \end{aligned} \tag{3.1}$$

In our equations, we also include a source term that serves to drive the system at large scales. This is modeled as a virtual antenna.

$$\nabla_\perp^2 A_\parallel = J_{\parallel,plasma} + J_{\parallel,antenna} \tag{3.2}$$

where $J_{\parallel,plasma} = u_{\parallel,e}$. We choose to write the antenna current in terms of a magnetic

potential as well, $J_{\parallel,antenna} = \nabla_\perp^2 A_A$.

The slow solar wind is a nearly collisionless system. However, a few collisions are necessary to dissipate the energy at high $k_\perp$'s. Our model is not designed to accurately capture the details of this damping mechanism, but we need some form of dissipation at small scales to balance the driving term. We express our dissipation terms in the form

$$\nu_d \left(-\hat{\nabla}_\perp^2\right)^m f \tag{3.3}$$

where $\nu_d = \nu$ and $f = \hat{n}_{e,i}$ for the virtual viscosity term and $\nu_d = \eta$ and $f = \hat{A}_\parallel$ for the virtual resistivity term. m is a parameter that allows us to control how much dissipation is included in the system [7]. To guarantee that the dissipation scale is removed from the inertial range scale, we set $m = 3$. Our normalized set of equations with driving and damping becomes

$$\frac{d}{d\hat{t}}\left(\hat{n}_e - \delta\hat{B}_\parallel\right) = -\hat{\nabla}_\parallel \hat{u}_{\parallel,e} - \left[\hat{u}_{\parallel,e}, \hat{A}_\parallel\right] - \sqrt{\frac{\beta_i}{2}}\left[\hat{p}_{\perp,e}, \delta\hat{B}_\parallel\right] + \nu\hat{\nabla}_\perp^{2m}\hat{n}_e \tag{3.4}$$

$$\frac{d\hat{A}_\parallel}{d\hat{t}} = -\hat{\nabla}_\parallel\hat{\phi} + \sqrt{\frac{\beta_i}{2}}\hat{\nabla}_\parallel\hat{p}_{\parallel,e} + \sqrt{\frac{\beta_i}{2}}\left[\hat{p}_{\parallel,e}, \hat{A}_\parallel\right] + \eta\hat{\nabla}_\perp^{2m}\hat{A}_\parallel \tag{3.5}$$

$$\frac{d}{d\hat{t}}\left(\tau\hat{p}_{\parallel,e} - \delta\hat{B}_\parallel\right) = 0 \tag{3.6}$$

$$\frac{d}{d\hat{t}}\left(\tau\hat{p}_{\perp,e} - 2\delta\hat{B}_\parallel\right) = 0 \tag{3.7}$$

$$\frac{d}{d\hat{t}}\left(\hat{\bar{n}}_i - 2\tilde{J}_1\delta\hat{B}_\parallel\right) = 2\sqrt{\frac{\beta_i}{2}}\left[\hat{\bar{p}}_{\perp,i}, \tilde{J}_1\delta\hat{B}_\parallel\right] + \nu\hat{\nabla}_\perp^{2m}\hat{\bar{n}}_i \tag{3.8}$$

$$\frac{d}{d\hat{t}}\left(\hat{\bar{p}}_{\perp,i} - 4\tilde{J}_1\delta\hat{B}_\parallel\right) = 0 \tag{3.9}$$

$$\sqrt{\frac{2}{\beta_i}}\left(\Gamma_{0,i}-1\right)\hat{\phi}=\hat{n}_e-\tilde{J}_0\hat{\bar{n}}_i \tag{3.10}$$

$$\sqrt{\frac{2}{\beta_i}}\hat{\nabla}_\perp^2\left(\hat{A}_\parallel-\hat{A}_A\right)=\hat{u}_{\parallel,e} \tag{3.11}$$

$$\frac{2}{\beta_i}\delta\hat{B}_\parallel=-\left(\hat{p}_{\perp,e}-\sqrt{\frac{2}{\beta_i}}\hat{\phi}\right)-2\tilde{J}_1\left(\hat{\bar{p}}_{\perp,i}+\sqrt{\frac{2}{\beta_i}}\tilde{J}_0\hat{\phi}\right) \tag{3.12}$$

3.3 Numerical Algorithms

We have employed a 3rd order Adams-Bashforth/2nd order Backwards Difference Formula (AB3/BDF2) algorithm to advance our equations in time. This method was first suggested by Hulsen [22] as part of a study of hybrid algorithms for stiff systems. This hybrid algorithm allows us to take stable large time steps early on while the contributions from the nonlinear terms are small. The time step adjusts itself to guarantee that if the shortest wavelength mode is traveling at the fastest speed of any in the simulation domain, the time step can capture this behavior. The algorithm can be expressed as

$$\frac{3}{2}u_{n+1}-2u_n+\frac{1}{2}u_{n-1}=\frac{8\Delta t}{3}N\left(u_n\right)-\frac{7\Delta t}{3}N\left(u_{n-1}\right)+\frac{2\Delta t}{3}N\left(u_{n-2}\right)+L\left(u_{n+1}\right) \tag{3.13}$$

where u is the state vector, $N\left(\right)$ is the nonlinear operator, and $L\left(\right)$ is the linear operator.

Typically, implicit schemes such as this require a matrix inversion every time step to solve for u_{n+1}. However, we have analytically performed the inversion for our six time-stepping equations, thus negating the need for the costly operation each time step. (See Appendix B.) In addition, we have treated Δt as an analytic parameter, so no inversions are necessary when the time step changes.

The spatial grid for the code employs Fourier modes ($e^{-i\mathbf{k}\cdot\mathbf{x}}$) as the basis set in a pseudospectral scheme. Most of the calculations are done in three-dimensional k-space, taking advantage of the fact that spatial derivatives in real space become simple multiplications in k-space. This minimizes the memory needed for a given accuracy since the error of calculating a derivative goes like $(1/N)^N$.

Spectral methods have the advantage that they use the whole grid to calculate a derivative.

$$\frac{\partial f}{\partial x} = i\,k\,f \quad \Rightarrow \sim O\left(h^N\right) \tag{3.14}$$

and since $h = \frac{L}{N}$, this ultimately becomes an error estimate of order $(1/N)^N$.

The non-linear, Poisson bracket terms are calculated in real space where they can be treated as pure multiplication. The Fourier transforms necessary to carry this out require de-aliasing, which mandates that we retain only the k-modes that satisfy $k \leq \frac{2}{3}k_{max}$, where k_{max} is the highest k mode our basis set can resolve. This is due to the fact that the transform will try to resolve wave numbers higher than k_{max} and only capture data at the discretization points.

For example, imagine a three point grid in space as in Figure 3.1. The red mode would describe a wave that has zero value at the end points and a finite value at the middle point. The green mode would have a zero value at all three grid points. The blue mode would have zero value at the end points and a finite value at the middle point - making it indistinguishable from the red mode! The grid has no way to account for the extra wiggles in the blue mode, and thus cannot resolve it. For these higher k modes, we set the amplitudes to zero so that their values do

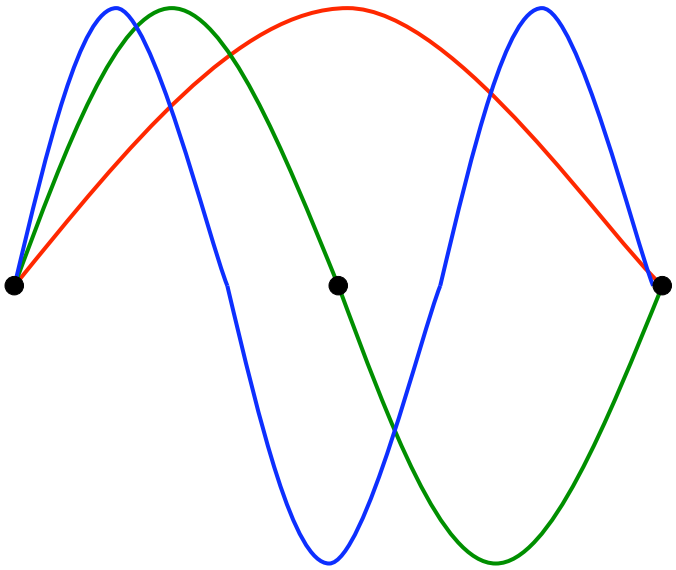

Figure 3.1: A three point grid would be unable to distinguish the red mode from the blue mode. They are both zero at the end points and have some finite value at the middle point. Dealiasing is used to remove higher modes that cannot be fully resolved by the grid.

not "alias" to a lower mode when transformed.

3.4 Verification

In computation, verification is the process by which one demonstrates that a given algorithm can solve a desired set of equations accurately. As one moves from the continuous ideal of pure mathematics to the discrete approximation of mathematical operators, there are a wealth of potential errors and instabilities into which one might fall. As a result, an entire branch of computational science and mathematics has dedicated itself to the study of such potential problems.

Numerical analysis has developed, and continues to develop, tests that allow a programmer to assess the strengths and weaknesses of a given algorithm. For example, through numerical analysis, we can determine the size of a time step that will allow the maintenance of a given level of accuracy. Through numerical analysis, we know that dealiasing is a necessary part of an algorithm employing discrete Fourier transforms. Numerical analysis helps us to determine if a given algorithm will be able to capture solutions to a given set of equations.

The difficulty comes when we try to study a system of equations of which we do not know all of the properties. Often, we use computers to probe the boundaries of what we do not know or understand. In these cases, it is crucial that the algorithms employed can accurately solve problems which have known solutions and are similar to or part of larger, more complicated problems.

In order to verify our algorithms, we looked at two different cases. The first was to test the linear part of the algorithm, including the subroutine to drive the system and the treatment of the artificial damping terms. The second was to test the Poisson bracket subroutine.

3.4.1 Linear Test Case

For the linear test case, we used Fourier and Laplace transforms to find the response of $\hat{A}_{\|}$ to a virtual antenna perturbing the magnetic field, represented as

$\hat{A}_A$. Analytically, this can be written as

$$\hat{A}_\parallel = -\hat{\omega}_L^2 \hat{A}_A \left(\frac{e^{-i\hat{\omega}_d t}}{(\hat{\omega}_d - \hat{\omega}_-)(\hat{\omega}_d - \hat{\omega}_+)} + \frac{e^{-i\hat{\omega}_+ t}}{(\hat{\omega}_+ - \hat{\omega}_d)(\hat{\omega}_+ - \hat{\omega}_-)} + \frac{e^{-i\hat{\omega}_- t}}{(\hat{\omega}_- - \hat{\omega}_d)(\hat{\omega}_- - \hat{\omega}_+)} \right) \tag{3.15}$$

where $\hat{\omega}_d$ is the normalized driving frequency of the antenna, $\hat{\omega}_L$ is the linear dispersion relation without damping or driving (the normalized version of Equation 2.42), and $\hat{\omega}_\pm$ are the normalized frequencies from the linear dispersion relation without driving, but with damping. The frequencies are

$$\begin{aligned} \hat{\omega}_\pm = & -i\hat{k}_\perp^{2m} \frac{2\nu b + \eta(2b - \beta_i c^2 f g \tau)}{2(2b - \beta_i c^2 f g \tau)} \\ & \pm \frac{i}{2(2b - \beta_i c^2 f g \tau)} \sqrt{\hat{k}_\perp^{4m} (2\nu b - \eta(2b - \beta_i c^2 f g \tau))^2 + 4q\hat{k}_\parallel^2 \hat{k}_\perp^2 (2b - \beta_i c^2 f g \tau)} \end{aligned} \tag{3.16}$$

where b, c, f, g, and q are functions of β_i, τ, and the Bessel approximations. They are defined in Appendix B. Figure 3.2 shows the analytic response to the driving and damping, as well as what the code calculated. Figure 3.3 show the absolute value of the relative error for the calculation. It is consistent with a second order accurate scheme.

3.4.2 Nonlinear Test Case

For the Poisson bracket test case, we used a set of reduced MHD equations to run the Orszag-Tang MHD benchmark [30]. This system consists of two equations

$$\frac{\partial}{\partial t} \nabla_\perp^2 \phi = -\left[\phi, \nabla_\perp^2 \phi\right] + \left[A_\parallel, \nabla_\perp^2 A_\parallel\right] + \nu \nabla_\perp^2 \phi \tag{3.17}$$

$$\frac{\partial}{\partial t} A_\parallel = -\left[\phi, A_\parallel\right] + \eta \nabla_\perp^2 A_\parallel \tag{3.18}$$

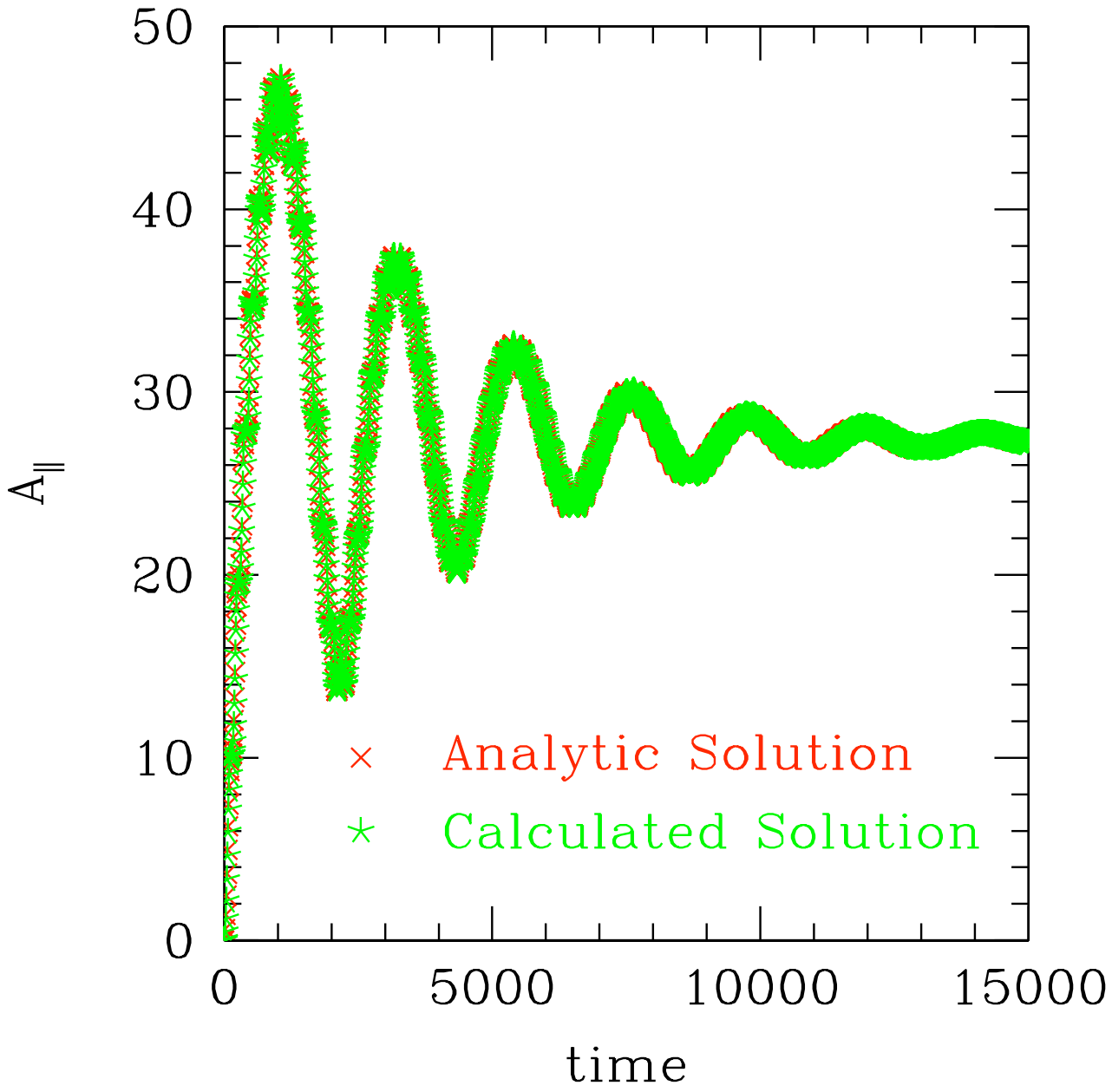

Figure 3.2: Verification of the driven, damped linear system. The predicted and calculated responses agree to within less than 1%.

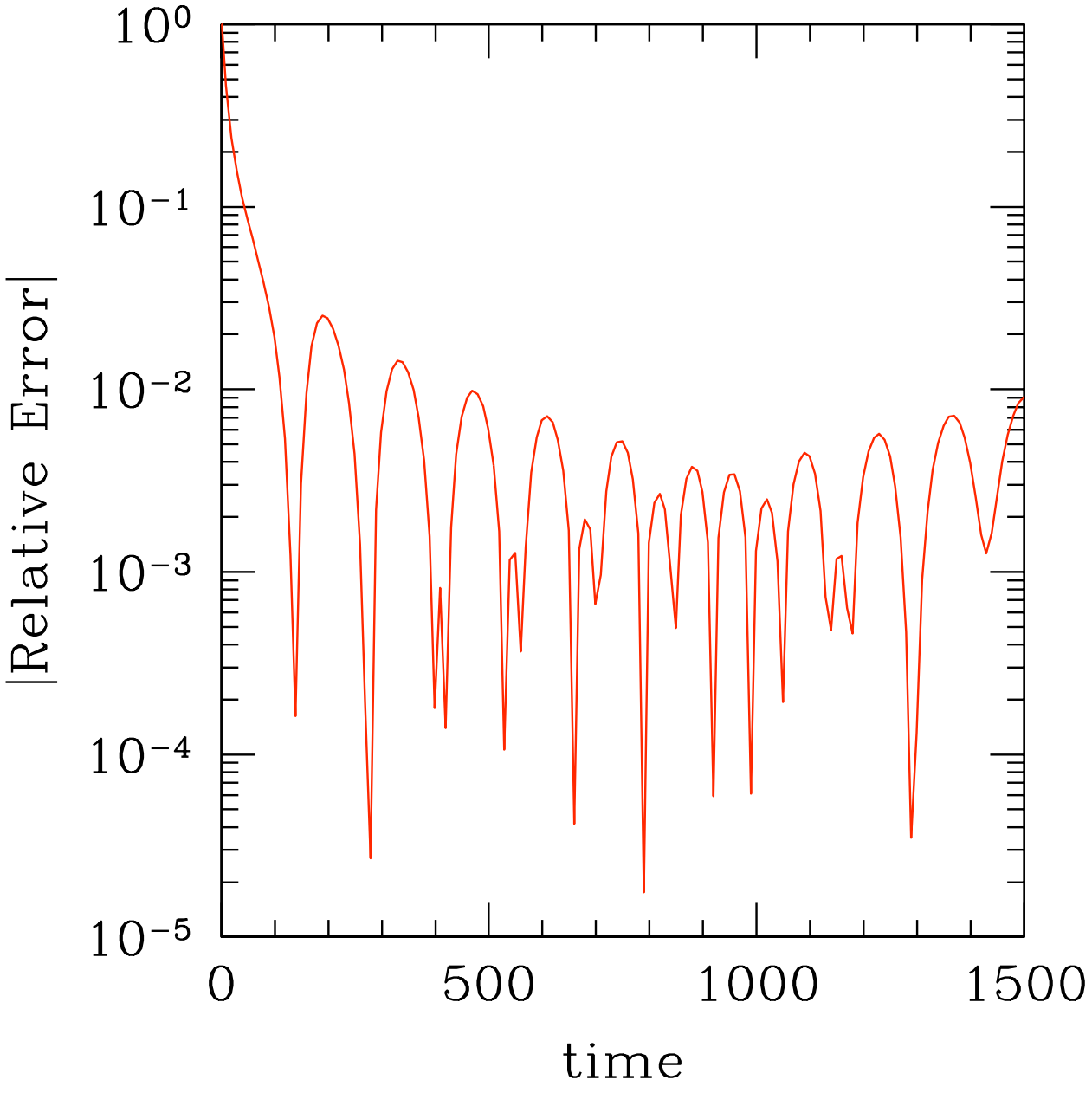

Figure 3.3: The absolute value of the relative error between the predicted and calculated responses for the driven, damped linear system. The high initial error is due to the error by the computer in calculating the exponential of a tiny number.

and initial conditions

$$\phi = -2\left(\cos x + \cos y\right); \quad A_{\parallel} = \cos y + \frac{1}{2}\cos 2x \tag{3.19}$$

We note that the initial conditions in the original paper contain a sign error or that the plots in the original paper are oriented incorrectly. These initial conditions describe two vortices in a hyperbolic tangent type configuration. The calculation follows them as they collide and form a thin current sheet. Figures 3.4 can be compared to Figure 5 in the original paper. It shows the total, magnetic, and kinetic energies as a function of time. Figure 3.5 can be compared to Figure 7 of the original paper. It shows a contour plot of the current at $t = 1$. The current sheet in the middle of the plot is clearly visible.

3.5 Results

The driving mechanism for turbulence in the solar wind is still an open question. However, observational evidence suggests that an inertial region with the appropriate Kolmogorov scaling does exist [1]. We use a virtual antenna to introduce energy into our system at the largest wavelengths, simulating the arrival of energy at our system size from larger, presumably turbulent scale sizes.

We drove our antenna at three low $k_{\perp}$ modes corresponding to $(k_x, k_y, k_z) = (1,2,1), (1,1,2), (2,1,1)$ at $\omega_A = 1.2 \times 10^{-4} - i0.6 \times 10^{-4}$ and amplitude of 500.0 in normalized units. Dissipation was determined by $\eta = \nu = 0.5$. Our system encompassed $k_{\perp}\rho_i$ values of 0.015 through 1.4 in a box that in real space is composed of 240^3 grid points. We used a low $\beta = 0.01$ value and $\tau = 1.0$.

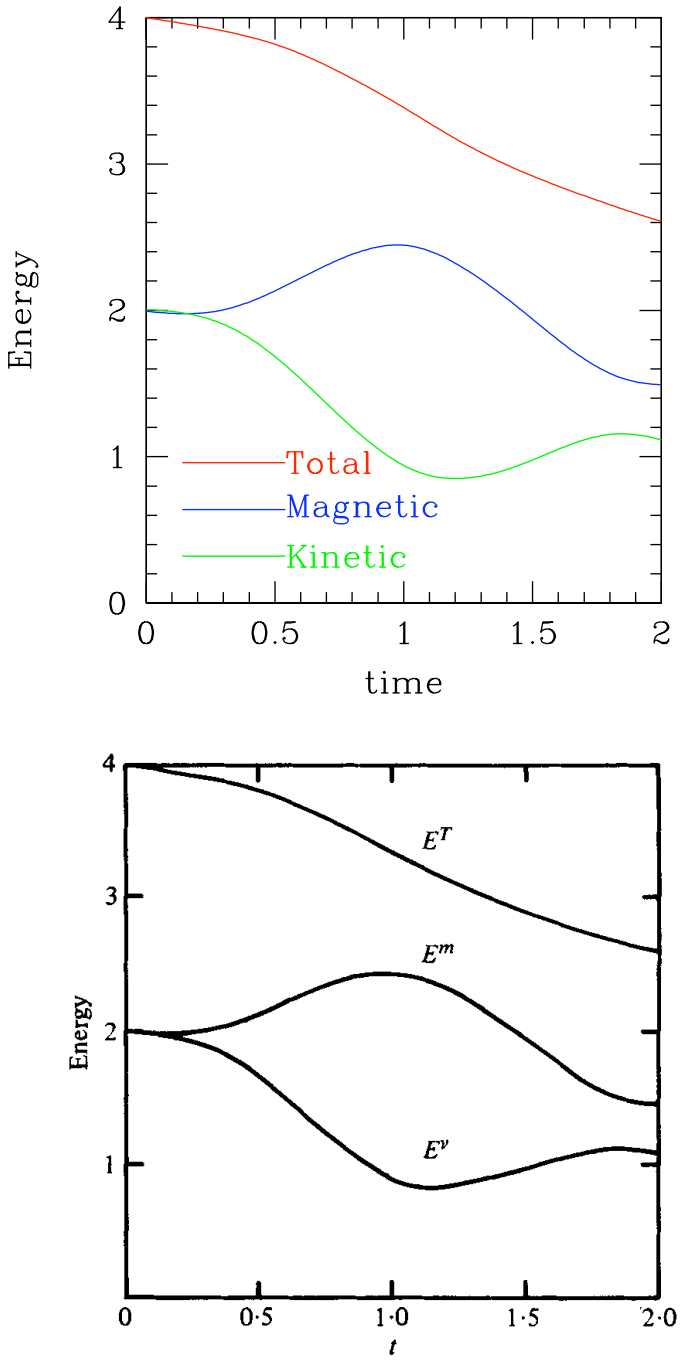

Figure 3.4: Verification of the coding that solves for the nonlinear, Poisson bracket terms using the Orszag-Tang problem. These graph shows the evolution of the energy as a function of time. The top graph is the result from our code. The bottom graph is Figure 5 of Orszag and Tang's paper [30].

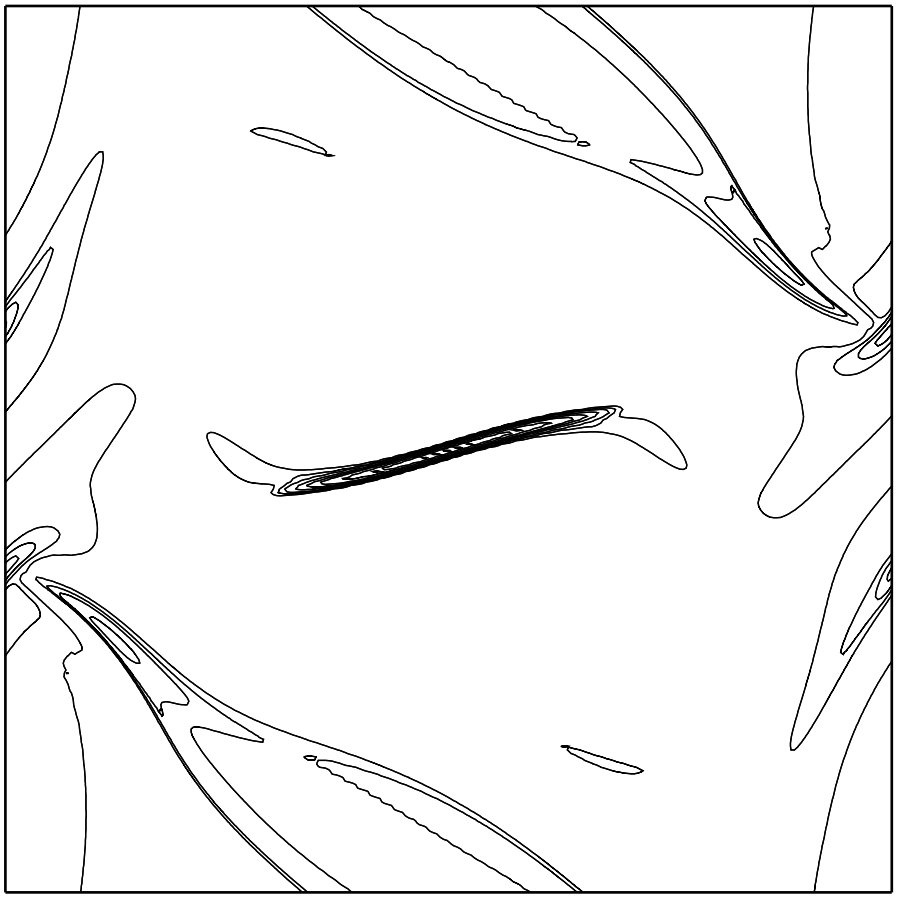

Figure 3.5: Verification of the coding that solves for the nonlinear, Poisson bracket terms. This shows a contour of the current at $t = 1$. A current sheet is clearing forming in the center of the simulation domain. This can be compared to Figure 7 of Orszag and Tang's paper [30].

The resulting energy spectrum is plotted in Figure 3.6. The magnetic and electric field energies are aligning themselves with the expected slope of $(k_\perp \rho_i)^{-5/3}$. A slope of $(k_\perp \rho_i)^{-3/2}$ fits the data better if we were to include the modes that are being driven. However, we do not wish to diagnose our virtual antenna, and conclude that $(k_\perp \rho_i)^{-5/3}$ is a better fit for the turbulent system.

Density perturbations are not expected to be seen in purely Alfvénic turbulence, but KAW turbulence would predict the energy in density perturbations to have a slope of $(k_\perp \rho_i)^{1/3}$ at long wavelengths (small $k_\perp \rho_i$ values). The $\delta B_\parallel$ spectrum appears to be following the same trend.

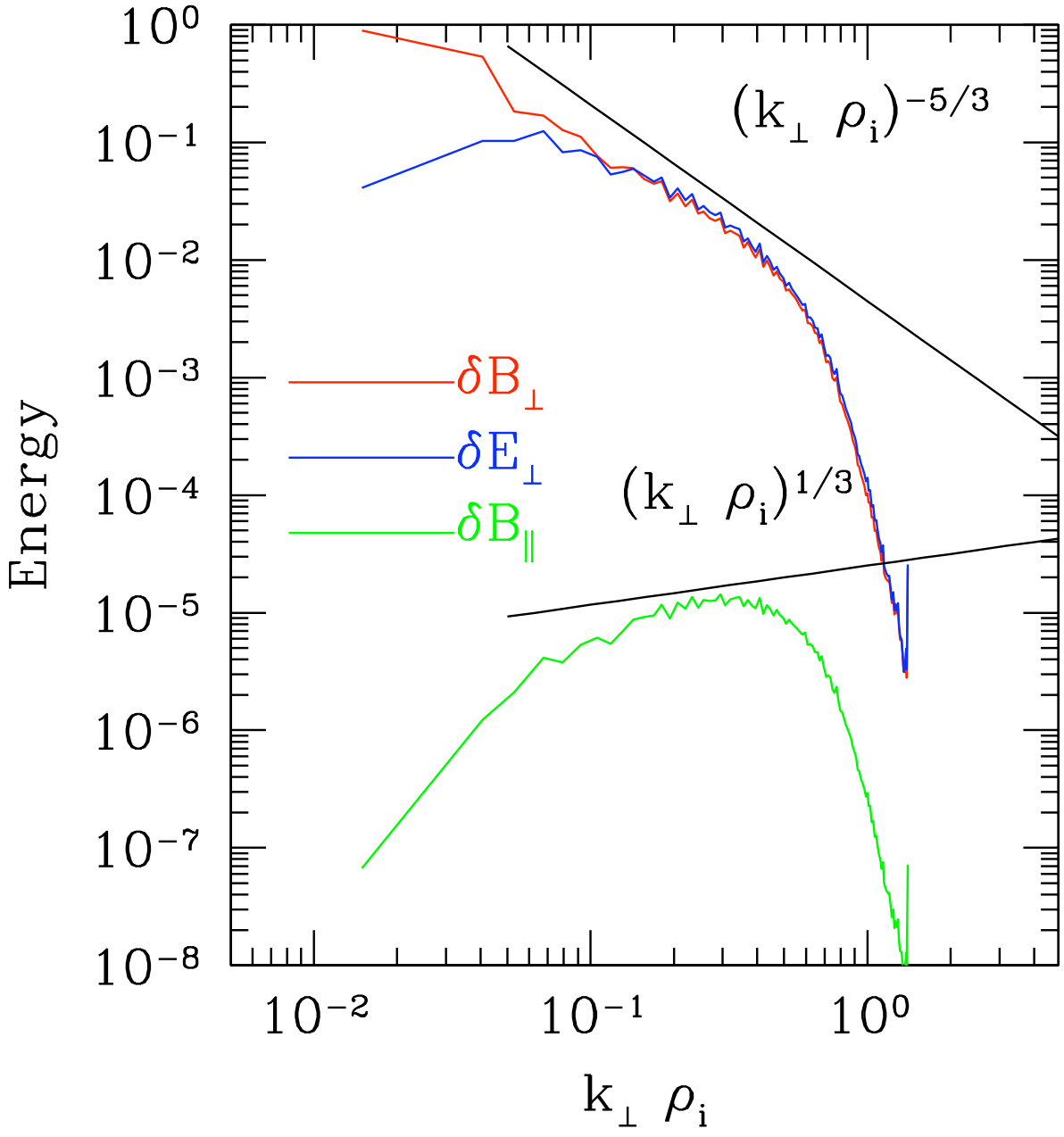

Figure 3.6: Energy in the perpendicular magnetic and electric fields are aligned with a $(k_\perp \rho_i)^{-5/3}$ slope as predicted by a Goldreich and Sridhar Kolmogorov-type argument [17]. Energy in the parallel magnetic field appears to be following the trend predicted for density fluctuations in kinetic Alfvén waves in the long wavelength limit corresponding to a $(k_\perp \rho_i)^{1/3}$ slope.

Chapter 4

Gyrofluid Modeling of Fusion Plasmas and Nonlinear Phase Mixing

4.1 Introduction

Since Alfvén introduced MHD, fluid models have had a rich history of providing physicists with a greater understanding of plasma systems, their waves, and instabilities. Using the MHD framework, the plasma community has been able to create MHD stable fusion devices and to probe the next level of plasma complexity - their kinetic nature. This has led most notably to the development of gyrokinetic theory [25] [16] which provides an even richer understanding of a plasma system.

Gyrokinetic models can be expensive to investigate computationally and, as they involve a fully time, space, and velocity-dependent particle distribution function, the results can sometimes be difficult to conceptualize. Gyrofluid models seek to find a compromise between the physics-rich kinetic models and the more intuitive, less expensive fluid models. The fluid model we describe in this chapter is obtained by taking moments of the gyrokinetic equation (see Appendix A), thus reducing the complexity of the equations from five dimensions to three. Closures are chosen to capture the linear response of the kinetic system - thus capturing kinetic behavior such as Landau damping and FLR effects.

The model we describe in Section 4.2 was originally developed by Beer et al, [4]. There has been concern in the past, however, that this model is not sufficiently

rich enough to capture some of the more important features of turbulence in fusion devices. In Section 4.3, we propose a potential remedy to this situation, nonlinear phase mixing by $\mathbf{E} \times \mathbf{B}$ drifts. We describe the physical mechanism involved and introduce a mathematical representation that can be incorporated into the original equations.

4.2 Equations

For our modeling of kinetic fusion systems, we use the equations developed by Beer, et al. [4]. These equations were derived to model low frequency fluctuations consistent with gyrokinetic orderings (see Equations A.1 and A.2), and maintain toroidal effects such as curvature and ∇B drifts, toroidal finite Larmor radius effects, and the mirroring force.

This introduces additional terms into the gyrokinetic equation beyond what is discussed in Chapter 2 and Appendix A. Curvature and ∇B drifts are expressed as

$$\mathbf{v}_d = \frac{v_\parallel^2}{\Omega}\hat{\mathbf{b}} \times \left(\hat{\mathbf{b}} \cdot \nabla\hat{\mathbf{b}}\right) + \frac{v_\perp^2}{2B\Omega}\hat{\mathbf{b}} \times \nabla B \tag{4.1}$$

and are included in addition to the parallel velocity and the $\mathbf{E} \times \mathbf{B}$ drift. We also keep track of equilibrium gradients for the density and temperatures.

This model evolves six gyrofluid moments - density, parallel velocity, parallel and perpendicular temperature, and parallel and perpendicular heat flux. As in the slab geometry case described in Chapter 2, the gyroaveraging operator, J_0, needs to be approximated in order to evaluate the integrals. For this model, the gyro average of J_0 is approximated as $\Gamma_0^{1/2}$ where Γ_0 is defined in Equation A.66. Gyroaverages

of more complicated combinations of J_0's and velocities are treated as explained in [4]. The equations can be written as

$$\frac{\partial n}{\partial t} + \mathbf{v}_\Psi \cdot \nabla n + \left[\frac{1}{2}\hat{\nabla}_\perp^2 \mathbf{v}_\Psi\right] \cdot \nabla T_\perp + B\nabla_\parallel \frac{u_\parallel}{B} - \left(1 + \frac{\eta_\perp}{2}\hat{\nabla}_\perp^2\right) i\omega_\star \Psi + \left(2 + \frac{1}{2}\hat{\nabla}_\perp^2\right) i\omega_d \Psi$$
$$+ i\omega_d \left(T_\parallel + T_\perp + 2n\right) = 0$$

$$\frac{\partial u_\parallel}{\partial t} + \mathbf{v}_\Psi \cdot \nabla u_\parallel + \left[\frac{1}{2}\hat{\nabla}_\perp^2 \mathbf{v}_\Psi\right] \cdot \nabla q_\perp + B\nabla_\parallel \frac{T_\parallel}{B} + B\nabla_\parallel \frac{n}{B} + \nabla_\parallel \Psi + \left(T_\perp + n + \frac{1}{2}\hat{\nabla}_\perp^2 \Psi\right) \nabla_\parallel \ln B$$
$$+ i\omega_d \left(q_\parallel + q_\perp + 4u_\parallel\right) = 0$$

$$\frac{\partial T_\parallel}{\partial t} + \mathbf{v}_\Psi \cdot \nabla T_\parallel + B\nabla_\parallel \frac{q_\parallel + 2u_\parallel}{B} + 2\left(q_\perp + u_\parallel\right) \nabla_\parallel \ln B - \eta_\parallel i\omega_\star \Psi + 2\, i\omega_d \Psi + i\omega_d \left(6T_\parallel + 2n\right)$$
$$+ 2\left|\omega_d\right| \left(\nu_1 T_\parallel + \nu_2 T_\perp\right) = -\frac{2}{3}\nu_{ii} \left(T_\parallel - T_\perp\right)$$

$$\frac{\partial T_\perp}{\partial t} + \mathbf{v}_\Psi \cdot \nabla T_\perp + \left[\frac{1}{2}\hat{\nabla}_\perp^2 \mathbf{v}_\Psi\right] \cdot \nabla n + \left[\hat{\hat{\nabla}}_\perp{}^2 \mathbf{v}_\Psi\right] \cdot \nabla T_\perp - B\nabla_\parallel \frac{u_\parallel}{B} + B^2 \nabla_\parallel \frac{q_\perp + u_\parallel}{B^2}$$
$$- \left[\frac{1}{2}\hat{\nabla}_\perp^2 + \eta_\perp \left(1 + \hat{\hat{\nabla}}_\perp{}^2\right)\right] i\omega_\star \Psi + \left(1 + \frac{1}{2}\hat{\nabla}_\perp^2 + \hat{\hat{\nabla}}_\perp{}^2\right) i\omega_d \Psi + i\omega_d \left(4T_\perp + n\right)$$
$$+ 2\left|\omega_d\right| \left(\nu_3 T_\parallel + \nu_4 T_\perp\right) = \frac{1}{3}\nu_{ii} \left(T_\parallel - T_\perp\right)$$

$$\frac{\partial q_\parallel}{\partial t} + \mathbf{v}_\Psi \cdot \nabla q_\parallel + \left(3 + \beta_\parallel\right) \nabla_\parallel T_\parallel + \sqrt{2} D_\parallel \left|k_\parallel\right| \left(q_\parallel - q_\parallel^{(0)}\right) + i\omega_d \left(-3q_\parallel - 3q_\perp + 6u_\parallel\right)$$
$$+ \left|\omega_d\right| \left(\nu_5 u_\parallel + \nu_6 q_\parallel + \nu_7 q_\perp\right) = -\nu_{ii} q_\parallel$$

$$\frac{\partial q_\perp}{\partial t} + \mathbf{v}_\Psi \cdot \nabla q_\perp + \left[\frac{1}{2}\hat{\nabla}_\perp^2 \mathbf{v}_\Psi\right] \cdot \nabla u_\parallel + \left[\hat{\hat{\nabla}}_\perp{}^2 \mathbf{v}_\Psi\right] \cdot \nabla q_\perp + \nabla_\parallel \left(T_\perp + \frac{1}{2}\hat{\nabla}_\perp^2 \Psi\right)$$
$$+ \sqrt{2} D_\perp \left|k_\parallel\right| \left(q_\perp - q_\perp^{(0)}\right) + \left(T_\perp - T_\parallel + \hat{\hat{\nabla}}_\perp{}^2 \Psi - \frac{1}{2}\hat{\nabla}_\perp^2 \Psi\right) \nabla_\parallel \ln B$$
$$+ i\omega_d \left(-q_\parallel - q_\perp + u_\parallel\right) + \left|\omega_d\right| \left(\nu_8 u_\parallel + \nu_9 q_\parallel + \nu_{10} q_\perp\right) = -\nu_{ii} q_\perp$$

where

$$\begin{aligned}
&\Psi = \Gamma_0^{1/2}(b)\,\Phi && ;\mathbf{v}_\Psi = \frac{c}{B}\hat{\mathbf{b}} \times \Gamma_0^{1/2}\Phi \\
&\frac{1}{2}\hat{\nabla}_\perp^2 = b\frac{\partial \Gamma_0^{1/2}}{\partial b} && ;\hat{\hat{\nabla}}_\perp{}^2 \Psi = b\frac{\partial^2}{\partial b^2}\left(b\Gamma_0^{1/2}\right)\Phi \\
&i\omega_\star = \frac{cT_{i0}}{eB^2 n_0}\mathbf{B} \times \nabla \mathbf{n_0} \cdot \nabla && ;i\omega_d = \frac{cT_{i0}}{eB^3}\mathbf{B} \times \nabla B \cdot \nabla
\end{aligned} \tag{4.2}$$

$\eta_\parallel$ is the ratio of the density gradient scale length to the parallel temperature gradient scale length, and $\eta_\perp$ is the ratio of the density gradient scale length to the perpendicular temperature gradient scale length. ν_{ii} is the ion -ion collision frequency. $\beta_\parallel$, $D_\parallel$, $D_\perp$, and ν_{1-10} are coefficients that are set by the closure assumptions.

This set of equations is ultimately closed using a technique developed by Hammett and Perkins [18] and further expanded by Dorland and Hammett [13] and Beer and Hammett [4]. Higher moments are expressed as linear functions of lower moments. The linear functions are then compared to their kinetic counterparts and the coefficients of the lower moments are adjusted to match the kinetic response of the system. In this way, the fluid model is able to capture kinetic effects such as Landau damping and toroidal phase mixing.

In addition to the closure coefficients, $q_\parallel^{(0)}$ and $q_\perp^{(0)}$ are closure approximations that allow the gyrofluid equations to capture the linearly-undamped, persistent, zonal flow behavior [5]. These terms are written as

$$q_\parallel^{(0)} = 3ik_r\rho_i\frac{qB_0}{\epsilon B}T_\parallel \quad ; q_\perp^{(0)} = ik_r\rho_i\frac{qB_0}{\epsilon B}T_\perp \tag{4.3}$$

The superscript indicates that these terms are only included when the toroidal wave number is zero. The q (without subscript) on the right hand side of the equation is the safety factor and ϵ is a normalized distance from the magnetic field axis.

4.3 Nonlinear Phase Mixing

The model described in the previous section is able to capture much of the relevant physics needed to describe tokamak turbulence. However, in a survey of

models investigating transport by Ion Temperature Gradient(ITG) turbulence in tokamaks, this gyrofluid model overestimated the flux by a factor of roughly 2 compared to more inclusive gyrokinetic models [12]. We believe the discrepancy can be remedied by the inclusion of nonlinear phase mixing by $\mathbf{E} \times \mathbf{B}$ drifts.

Imagine that there existed an approximately sinusoidal potential, a simplified version of the potential one might expect in a system with large zonal flows. If there existed a density perturbation with a Maxwellian distribution of velocities at the point were the gradient of the potential was the greatest, as represented by the first half of the cartoon in Figure 4.1, particles with lower energies and smaller gyroradii would be accelerated more than particles with greater energies and larger gyroradii. This is due to the difference in effective gradient of the gyroaveraged potential for the different sized gyroradii. Larger gyroradii would feel a weaker gradient. This effect would smear or phase mix away the density perturbation.

This effect was originally introduced by Dorland [13] for the case of slab geometry, but the belief was that it would not be important in toroidal geometry. Specifically, simulations did not exhibit large fluctuations at high $k_\perp \rho_i$'s where this effect would be strongest. In fact, we expect the strength of this effect to depend on the gradient of the potential and the relative size of the gyroradii compared to that gradient. Such that, gradient scale lengths that are long compared to the gyroradii of the perturbation would take longer to phase mix away the perturbation as in Figure 4.2. So while the effect might take longer at smaller at $k_\perp \rho_i$, it can still be important.

In order to find a mathematical representation of this phenomenon, we start

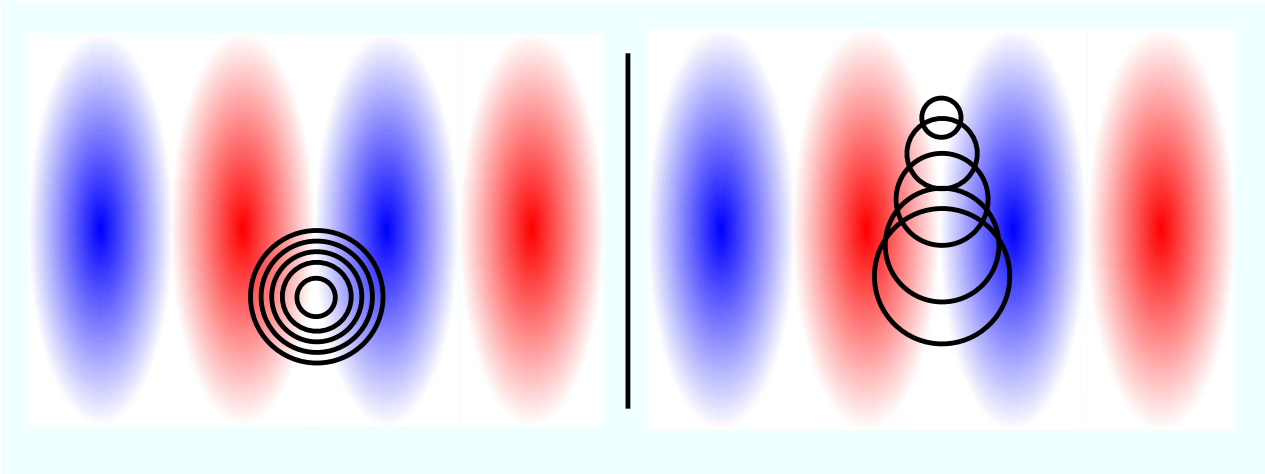

Figure 4.1: A cartoon describing nonlinear phase mixing by the $\mathbf{E} \times \mathbf{B}$ drift. The rings represent a density perturbation with a Maxwellian distribution of velocities. The colored ovals represent contours of the electric potential. Particles with larger gyroradii average over more variation in the potential's gradient and feel a weaker electric field than particles with smaller gyroradii. This creates a mechanism for dissipating the density perturbation.

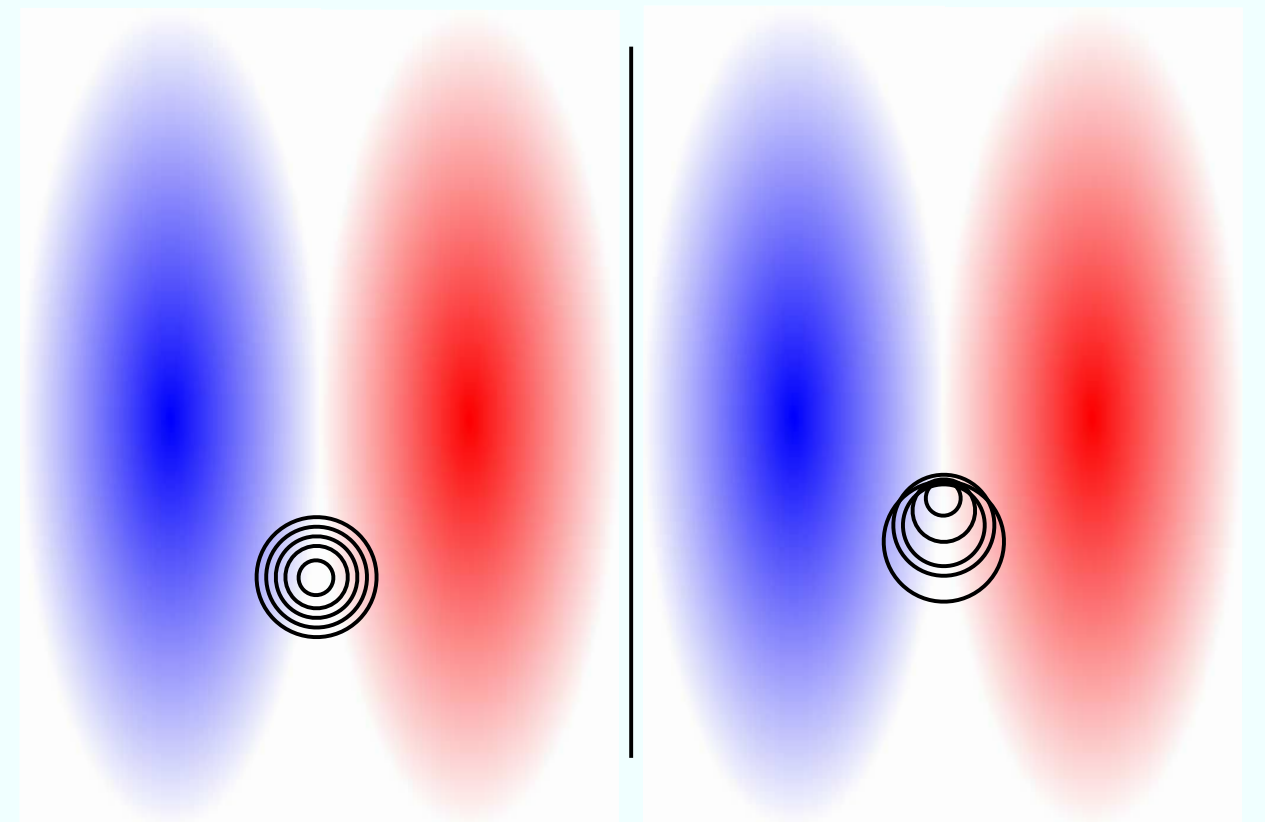

Figure 4.2: Like the previous figure, particles with different gyroradii experience a different acceleration. However, the distribution of accelerations depends on the size of the gradient compared to the size of the gyroradii of the particles.

with the $\mathbf{E} \times \mathbf{B}$ nonlinearity. In a convenient conservative form, we can write it as

$$\nabla \cdot \left(F \frac{c}{B} J_0 \left(\frac{k_\perp v_\perp}{\Omega} \right) \hat{\mathbf{b}} \times \nabla \Phi \right) \tag{4.4}$$

where J_0 only acts on the potential. To approximate the Bessel function, we Taylor expand to

$$\nabla \cdot \left(F \frac{c}{B} \hat{\mathbf{b}} \times \nabla \Phi \right) - \nabla \cdot \left(F \frac{c}{B} \hat{\mathbf{b}} \times \nabla \frac{k_\perp^2 v_\perp^2}{4\Omega_i^2} \Phi \right) + \nabla \cdot \left(F \frac{c}{B} \hat{\mathbf{b}} \times \nabla \frac{k_\perp^4 v_\perp^4}{64\Omega_i^4} \Phi \right) + \cdots \tag{4.5}$$

To take care of the closure for this sequence of terms, we again assume that higher moments can be expressed as linear combinations of lower moments. We will use the parallel pressure moment to demonstrate the closure technique. First, we integrate over velocity space.

$$\nabla \cdot \left(\frac{c}{B} p_\parallel \hat{\mathbf{b}} \times \nabla \Phi \right) - \nabla \cdot \left(r_{\parallel,\perp} \frac{c}{B} \hat{\mathbf{b}} \times \nabla \frac{k_\perp^2}{4\Omega_i^2} \Phi \right) + \cdots \tag{4.6}$$

We don't want to evolve $r_{\parallel,\perp}$ so we choose a closure such that $r_{\parallel,\perp} = \alpha_1 n + \alpha_2 T_\parallel + \alpha_3 T_\perp$. If our perturbed distribution function were a Maxwellian, $\alpha_1 = \alpha_2 = \alpha_3 = 1$. This technique of assuming the perturbation is itself a Maxwellian is known as the cumulant discard approximation, because the difference between the actual value of α and the value of α pertaining to a Maxwellian is known as the cumulant. However, this is a poor model for describing turbulence [28]. Instead, we choose α_1 and α_3 such that all the density and perpendicular temperature terms in the expansion can be recombined and expressed as $\mathbf{v}_\Psi \cdot \nabla n + \left[\frac{1}{2} \hat{\nabla}_\perp^2 \mathbf{v}_\Psi \right] \cdot \nabla T_\perp$ where any error is assumed to be in α_2 and higher order $k_\perp$ terms. The convective derivative term $\mathbf{v}_\Psi \cdot \nabla T_\parallel$ also comes from this expansion and is included as part of the full time derivative. At this point, we neglect any terms of order $k_\perp^4$ or higher. Taking all of

this into account, we are left with

$$-\left(\alpha_2 - 1\right) \nabla \cdot \left(T_{\parallel} \frac{c}{B} \hat{\mathbf{b}} \times \nabla \frac{k_{\perp}^2}{4\Omega_i^2} \Phi \right) \tag{4.7}$$

This term of order $k_{\perp}^2$ is what is missing from our gyrofluid model and represents the nonlinear phase mixing behavior. It depends on the gradient of the gyroaverage of Φ as well as the wave number of the perturbation itself.

We can get a better sense of the effect of this term by considering how it alone effects the particle distribution function in a kinetic formulation. We use a simplified model for the potential, $\Phi\left(x\right) = \Phi_0 e^{ik_x x}$ which approximates zonal flow behavior. Since J_0 only acts on Φ, the argument of J_0 will only have k_x dependence, too.

$$\frac{\partial f}{\partial t} + \hat{\mathbf{b}} \times J_0 \left(\frac{k_x v_x}{\Omega} \right) \nabla \Phi_0 e^{ik_x x} \cdot \nabla f = 0 \tag{4.8}$$

Now we consider a simplified pdf f that has a Maxwellian velocity distribution and sinusoidal variation in the y-direction. If we introduce it at $t = 0$, we find that

$$f = F_M e^{ik_y y} e^{-J_0\left(\frac{k_x v_\perp}{\Omega}\right) ik_x \Phi k_y t} \tag{4.9}$$

If we Taylor expand the J_0 function in order to take moments, we will find that the moments will depend on time as

$$m\left(t\right) \propto \frac{e^{-ik_y (k_x \Phi) t}}{1 - k_y k_x^2 \rho_i^2 \left(k_x \Phi\right) t} \tag{4.10}$$

where the k_x's come from the derivatives of Φ and the k_y from derivatives of the moment. Based on the form of the solution, we would expect the nonlinear phase mixing term to not only effect the phase of the oscillations, but also provide damping.

For this reason, our nonlinear phase mixing term will be approximated by

$$\nu_{pm} \left| \frac{1}{2} \hat{\nabla}_{\perp}^{2} \mathbf{v}_{\Psi} \cdot \nabla \right| T_{\parallel} \tag{4.11}$$

where the absolute value guarantees damping. We further simplify this expression by summing over the contribution from the $\mathbf{E} \times \mathbf{B}$ flow so that

$$\nu'_{pm} = \nu_{pm} \sum_{k_x} \frac{1}{2} \left| \hat{\nabla}_{\perp}^{2} v_y \right| \tag{4.12}$$

This form retains the dependence on the magnitude of v_y (the zonal flow), and on $k_{\perp}^{4}$, though in a more simplified fashion. Like our other closure coefficients, we are free to choose ν_{pm} so as to match the kinetic response of the system. This term can be included in the appropriate moment equations by making the transformation

$$\frac{\partial}{\partial t} \rightarrow \frac{\partial}{\partial t} + \nu'_{pm} \left| k_y \right| \tag{4.13}$$

It is important to note that this term only represents the damping by the zonal flows on the other fluctuations and not the reciprocal damping of the other fluctuations on the zonal flow potential. Because zonal flow structures are typically large amplitude and persist in time, the last term in the denominator of Equation 4.10, proportional to Φ, is larger for longer. For the smaller amplitude, shorter-lived perturbations, this term is much smaller, and consequently, so is the damping.

In the next chapter, we describe how we have incorporated this term into a gyrofluid code and its effect on predicted turbulent heat fluxes.

Chapter 5

Gyrofluid Models - Results

5.1 Introduction

Gyrofluid models are able to capture trends in turbulent plasma behavior without the computational and algebraic expense of a complete treatment of all of the physics, whether the physics is relevant to the behavior or not. Clearly, we would like to ultimately understand all of the relevant physics, but such an exhaustive study is currently not practical given time and monetary constraints. To find a cost effective way forward requires us to make judicious approximations and find tools that allow us to use what resources we have more efficiently. This is the niche for gyrofluid models. They are useful for parameter scans and directing attention to broad features that merit further investigation. They also allow us to extend our fluid intuition into kinetic regimes.

The gyrofluid model described in Chapter 4 is implemented in the gyrofluid code `gryffin`. The equations are solved using local flux tube geometry. The simulation domain is elongated along the field lines and follows them as they twist and shear. However, this behavior is mapped to a rectangular representation [31] [11] that allows for the use of periodic boundary conditions and a Fourier representation [3]. The domain is large enough to accommodate several turbulent correlation lengths both parallel and perpendicular to the magnetic field.

`gryffin` was originally developed to study ITG turbulence. In 2000 it was included in a study of gyrokinetic and gyrofluid ITG codes. Its predicted turbulent diffusivity was a factor of roughly 3.5 too high compared to the other codes for an experimentally relevant set of parameters [12]. With a view to remedy this discrepancy, we have introduced the nonlinear phase mixing term described in Chapter 4. In this chapter, we present a wide study of the effect of the nonlinear phase mixing term on `gryffin`-predicted turbulence levels.

`gryffin` was not been used actively as a research code in the past decade. As such, it needed to be updated to work with present day versions of mathematical, FFT, and data storage libraries. With the changes in the code, we wanted to be sure that it recovered the same results as previously reported. We have run the code in linear mode and recovered the predicted growth rates and frequencies originally reported by Beer, et al. in Figure 6 of [4] and in Figure 2 of [6]. We have also recovered the results reported in Dimits, et al. [12] which will be described below.

In Section 5.2 we present results from the updated `gryffin` code and compare them to studies performed previously that included gyrofluid codes. Specifically, we look at the Cyclone base case in Section 5.2.1. In Section 5.2.2 we reproduce a trapped particle scan used to investigate if the inclusion of persistent zonal flow behavior remedied discrepencies between gyrokinetic and gyrofluid predictions. In Section 5.2.3 we present the results of a study of Electron Temperature Gradient turbulence. In Section 5.3 we investigate a situation where general geometry is used. In Section 5.4, we compare results from `gryffin` to results from two gyrokinetic codes showing the applicability of the local flux tube approximation in systems for

which the gyroradius of the dominant species is small compared to the size of the device.

5.2 Comparisons to Previous Gyrofluid Studies

5.2.1 Cyclone Base Case

The Cyclone base case is a set of parameters developed from DIII-D H-mode shot #81499. This case was used to compare various gyrofluid and gyrokinetic codes. In addition to closely matching experimental parameters, assumptions include ion and electron densities and temperatures that are equivalent, a ratio of density to temperature gradient scale length that is $L_n/L_T = \eta_i = 3.114$, a safety factor $q = 1.4$, perfectly toroidal geometry, a shear $\hat{s} \approx 0.786$, a ratio of the major radius to the temperture gradient scale length $R/L_T = 6.92$, and $\epsilon = r/R = 0.18$ where R is the major radius, and r is the distance from the center of the plasma to the flux surface being studied. In addition, only electrostatic fluctuations are modeled. Electrons are assumed to be adiabatic, and only one ion species is evolved.

Figure 5.1 is a reproduction of Figure 3 from the comparison paper by Dimits, et al.[12]. The plus signs represent `gryffin`'s original predicted diffusivity which was too high by roughly a factor of 3.5 at the experimentally relevant value of $R/L_T = 3.114$ [12]. With the inclusion of closures that account for the long-time zonal flows [1], represented by the black diamonds, the discrepancy at that temperature

[1]Zonal flow behavior is already included in the gyrofluid model. However, it is damped away over time. The zonal flow closures allow the system to relax to a non-zero zonal flow state.

gradient drops to a factor of 2. This suggests that while the long-time zonal flows are important to include, they do not fully account for the discrepancy between the predictions from gyrofluid and gyrokinetic codes.

We ran a similar temperature gradient scale scan to the one found in the paper, with and without nonlinear phase mixing. We find that including nonlinear phase mixing decreases the predicted diffusivity by a factor of roughly 2.5 as can be seen in Figure 5.2. Thus, the inclusion of nonlinear phase mixing brings the gyrofluid results into the regime of the gyrokinetic ones.

5.2.2 Trapped Particle Scan

The discrepancy between gyrofluid and gyrokinetic results was originally hypothesized to be the suppression of turbulent flux by long term persistence of large amplitude zonal flows. Zonal flows were ultimately damped out completely in the gyrofluid treatment. In order to test this idea, cases were run at three different minor to major radius ratios, $\epsilon = r/R$. In the limit $\epsilon \to 0$, the number of trapped particles also goes to zero and the zonal flows disappear. Consequently, any discrepancies due to zonal flows should disappear as well. The initial comparison between gyrofluid and gyrokinetic results failed to show an ϵ dependence, maintaining a discrepancy of roughly a factor of 2 across the scan as shown in Figure 5.3.

The parameters for this test case come from an L-mode shot (#41309) of the Tokamak Fusion Test Reactor(TFTR) that was originally part of the Numerical Tokamak Project [10]. For this case, again the ion and electron densities and tem-

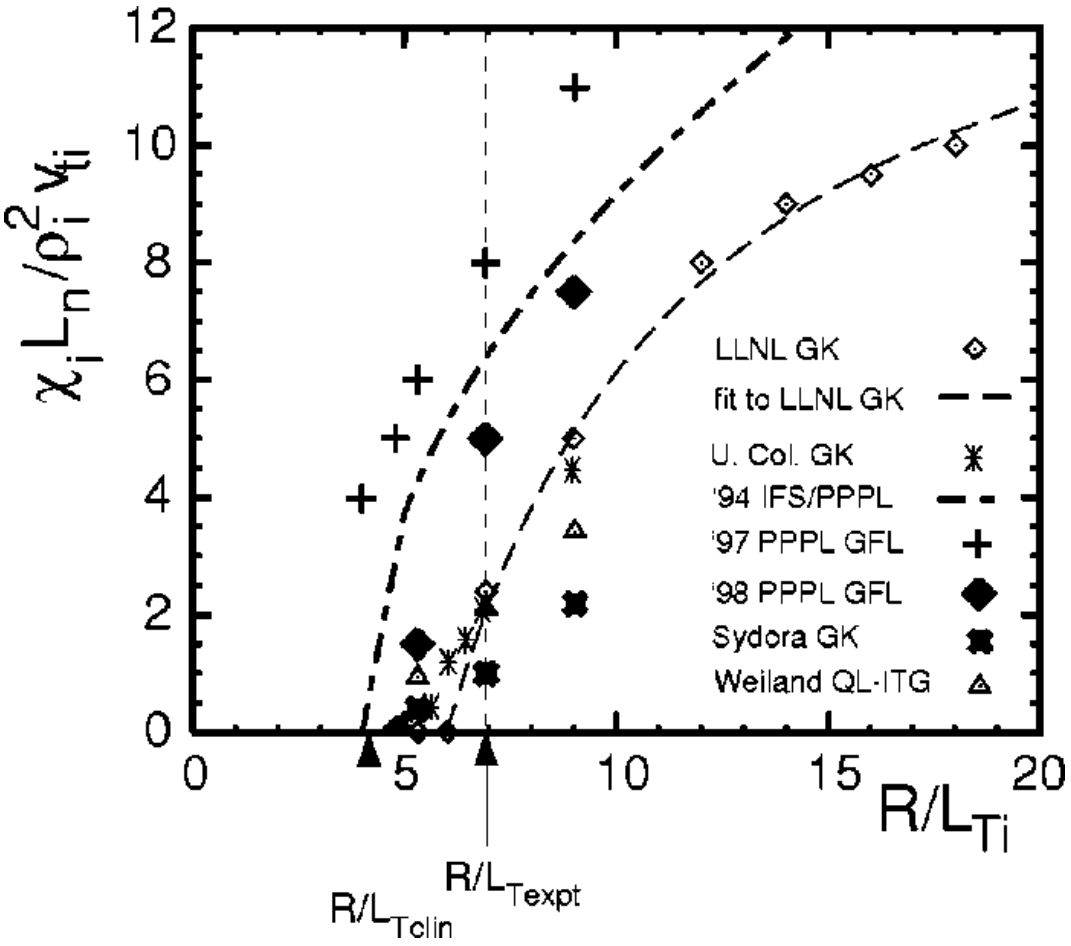

Figure 5.1: Figure 3 from the comparison paper by Dimits, et al. [12]. `gryffin`'s results originally differed from the gyrokinetic codes by a factor of 3.5 at the experimentally relevant temperature gradient of 3.114. With the inclusion of a closure that accounts for long-time zonal flows, the factor drops to roughly 2.

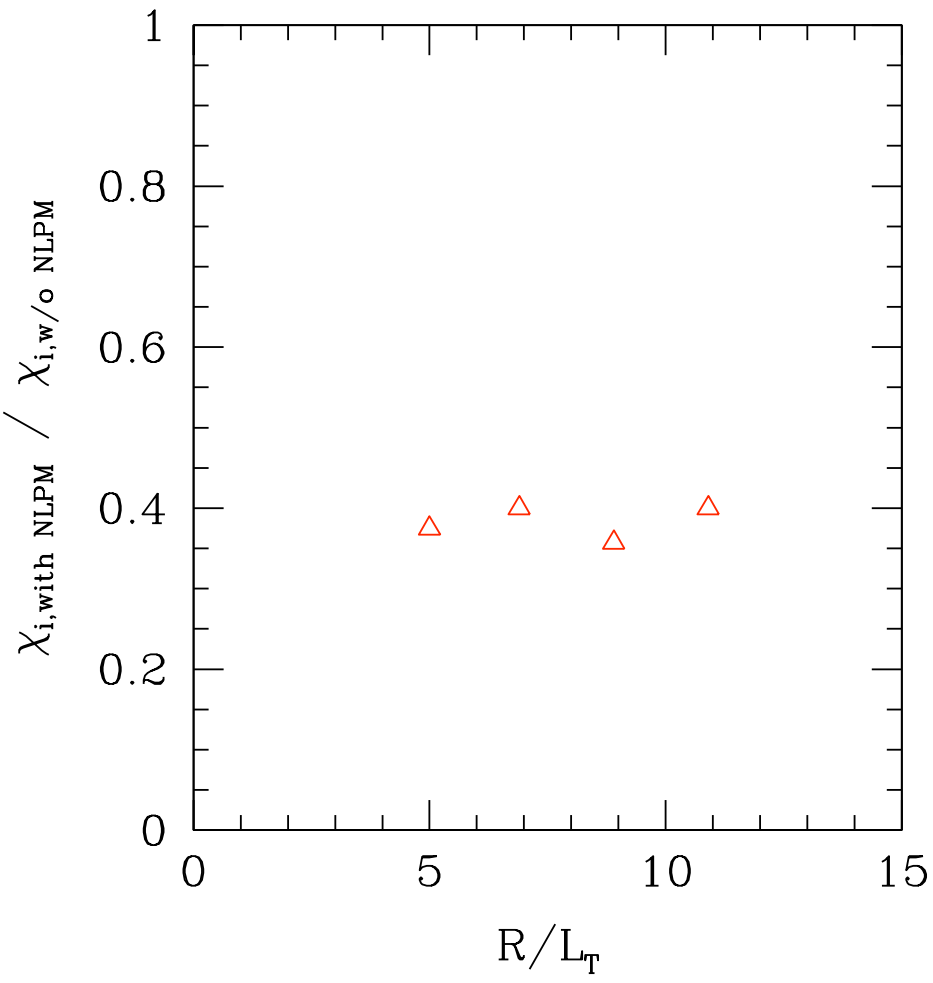

Figure 5.2: We ran a similar temperature gradient scan with and without nonlinear phase mixing. We found that including nonlinear phase mixing decreases the predicted diffusivity by a factor of roughly 2.5, bringing it closer to the gyrokinetic code predictions.

peratures are assumed equal. Electrostatic fluctuations, adiabatic electrons, one ion species, and perfectly toroidal geometry are again assumed. This time, however, $\eta_i = 4.0$, $q = 2.4$, $\hat{s} = 1.55$, and $R/L_T = 10.0$. The initial case has $\epsilon = 0.2$; a case with more trapped particles and a higher $\epsilon = 0.4$ as well as a no trapped particle, $\epsilon = 0.0$ case were also considered. Once again, we replicated the scan and found that including nonlinear phase mixing decreases predicted diffusivities by roughly 2.5, even in the $\epsilon = 0$ case. (See Figure 5.4.) Thus, the nonlinear phase mixing is able to suppress flux even for cases where there is little to no persistent zonal flows.

5.2.3 Electron Temperature Gradient Turbulence

Unlike ITG turbulence, Electron Temperature Gradient (ETG) Turbulence is not susceptible to large secondary instabilities that create zonal flows. As a result, we would not expect that including our nonlinear phase mixing term would have as much of an effect on an ETG simulation. Using the Cyclone Base Case parameters with gyrokinetic codes, Jenko and Dorland [24] showed that the relative heat transport for an ETG mode is significantly higher than that of an ITG mode. Figure 5.5 is taken from their paper and shows an ETG run (top) and an ITG run (bottom). Figure 5.6 shows two `gryffin` runs, again with an ETG run (top) and an ITG run (bottom) on a log scale. Nonlinear phase mixing is included in both `gryffin` runs. The results suggest that the effect of nonlinear phase mixing is enhanced by persistent zonal flow behavior.

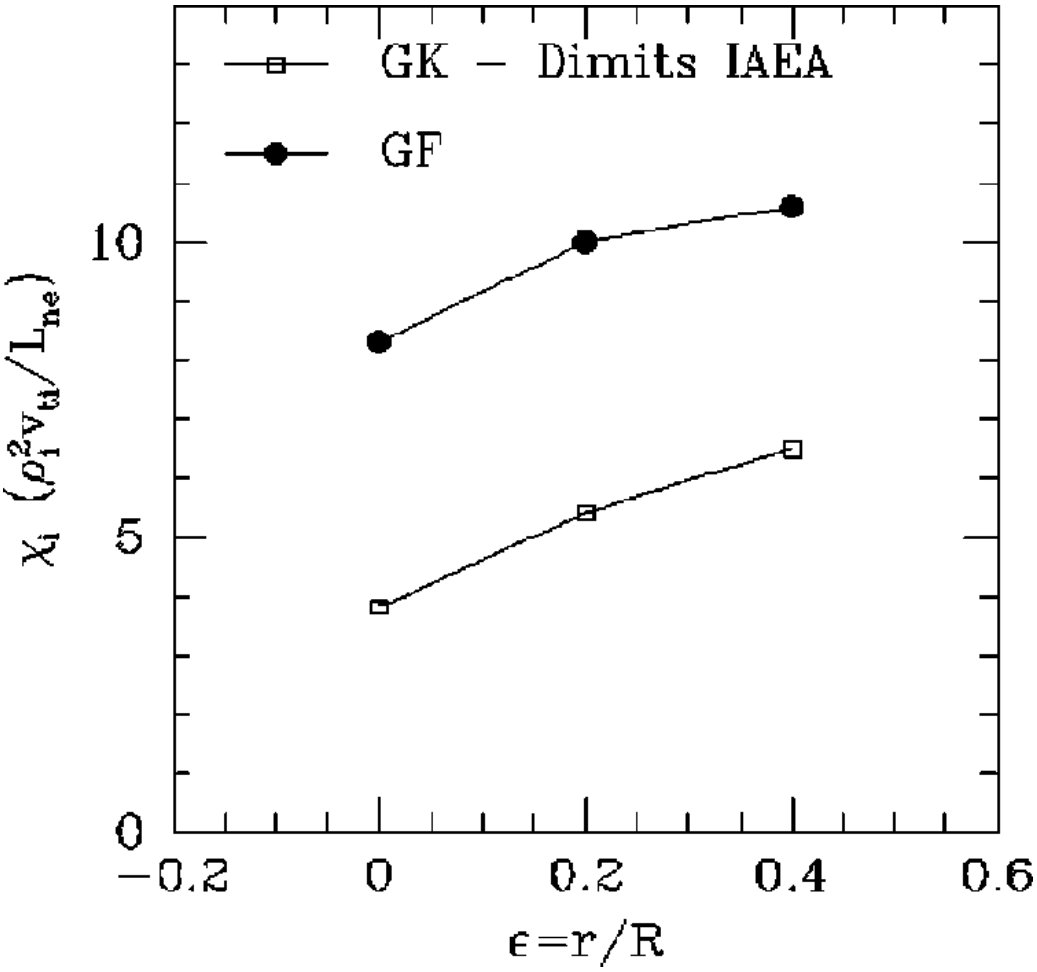

Figure 5.3: Figure 4 from the comparison paper by Dimits, et al. [12]. `gryffin`'s results originally differed from the gyrokinetic codes by a factor of 2 for this parameter set. If the discrepancy were due to zonal flows, the gyrofluid and gyrokinetic results should agree in the $\epsilon = 0$ limit.

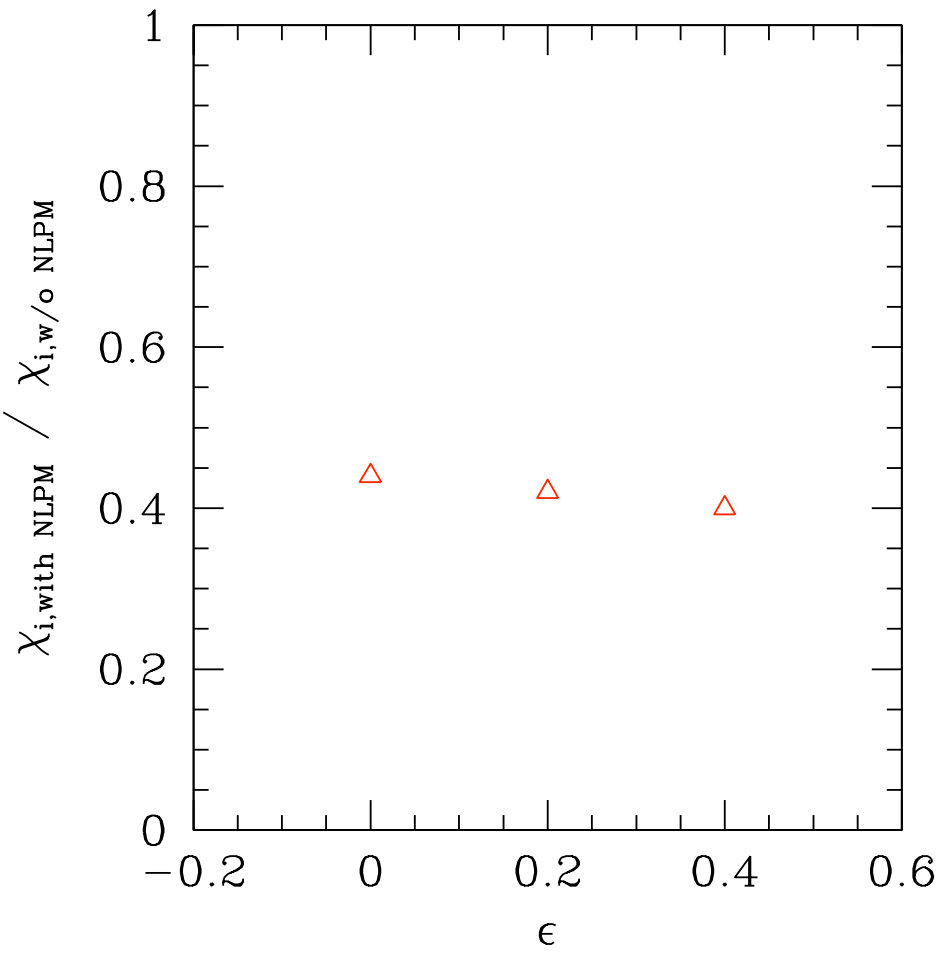

Figure 5.4: Ratio of ion thermal diffusivities for models with and without nonlinear phase mixing for TFTR shot # 41309 parameters. The nonlinear phase mixing brings the predicted diffusivities into better agreement with gyrokinetic predictions for all values of ϵ.

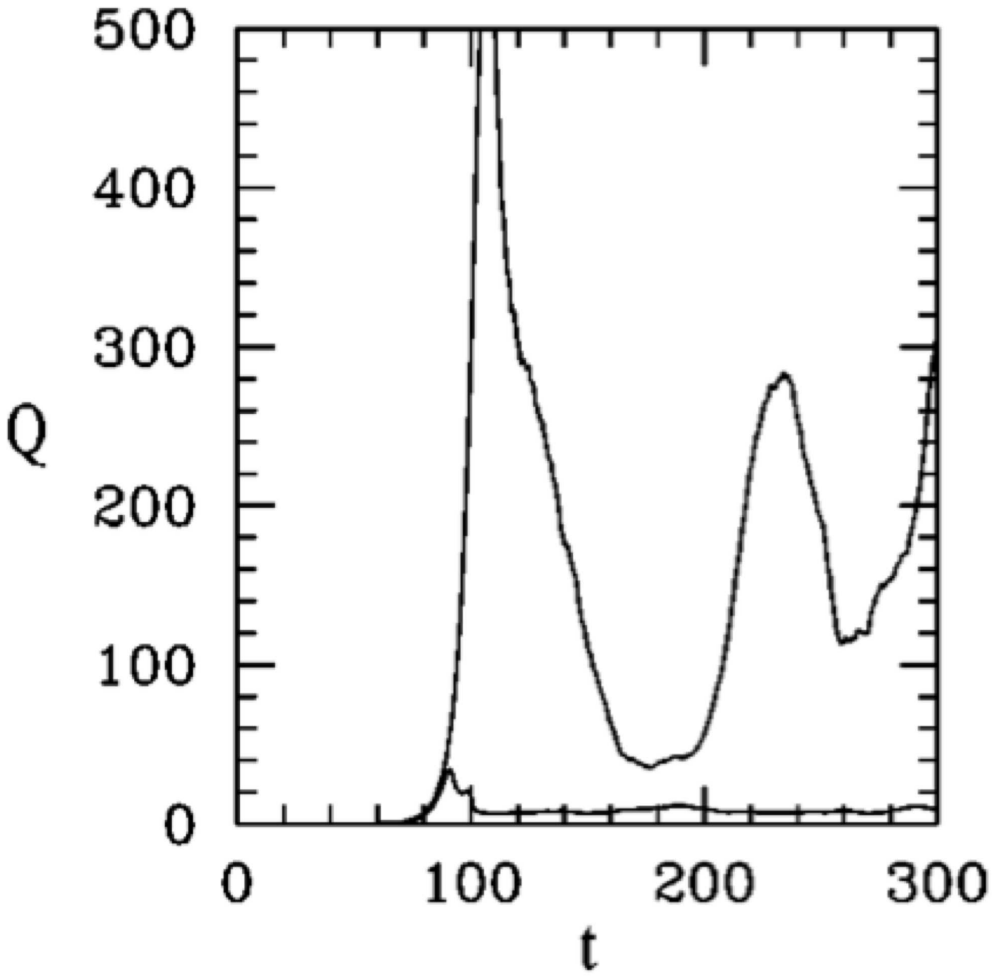

Figure 5.5: Figure 1 from the paper by Jenko and Dorland [24]. The top curve represents flux from ETG turbulence and the lower curve heat flux from ITG turbulence. ETG turbulence has a higher relative heat flux than ITG turbulence because secondary instabilities in ETG turbulence are weaker and do not lead to the formation of strong zonal flows.

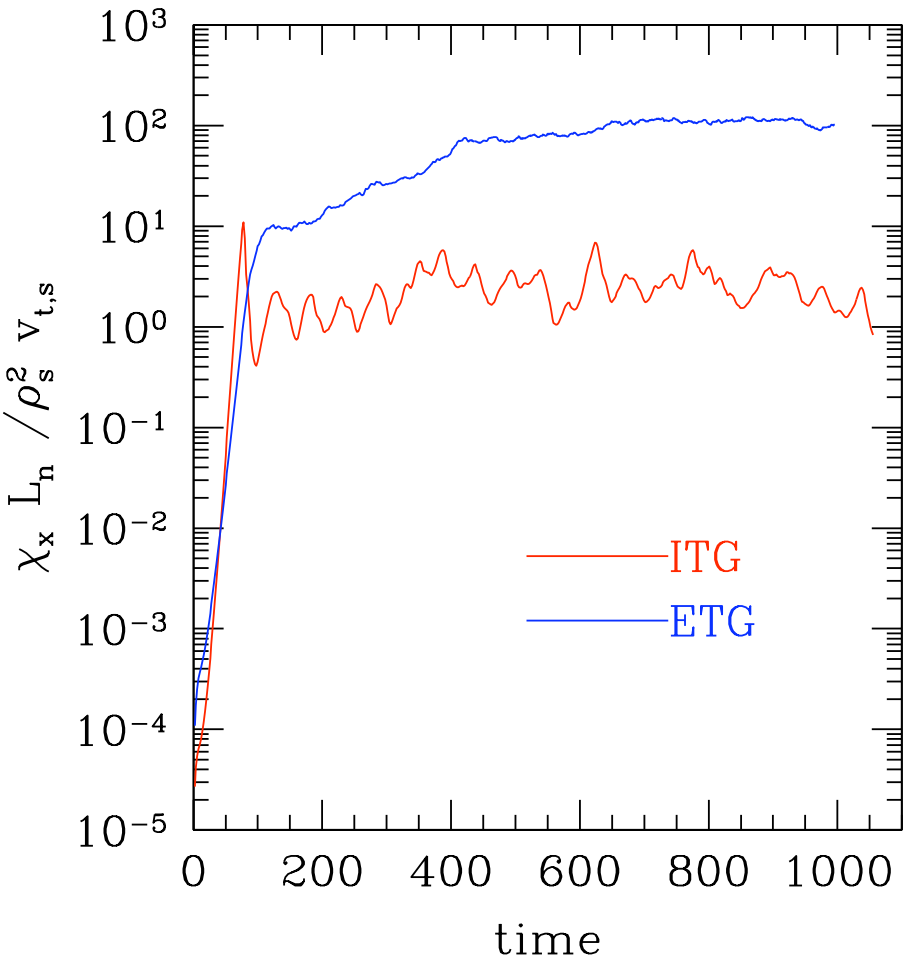

Figure 5.6: ETG turbulence is not squelched by the nonlinear phase mixing model. The larger relative flux compared to ITG turbulence is maintained. This suggests that the effect of nonlinear phase mixing does have some dependence on zonal flows as expected. Note that this figure uses a log scale along the ordinate.

5.3 General Geometry

The Cyclone base case assumes ideal toroidal geometry. However, we would like to ensure that both `gryffin` and our nonlinear phase mixing model work in cases of general geometry. We choose to run a study similar to one included in a Mikkelsen and Dorland paper on the effect of collisions on zonal flows in a realistic geometry [29]. Since zonal flows suppress turbulent transport, the fear was that too high of collisionality would weaken the effect of the zonal flows. The model they used, and that we use here, is based on the H-mode Alcator C-Mod shot # 960116027 that is included in the ITER Profile Database [15]. We assume adiabatic electrons and use a collisionality consistent with experimental parameters.

Figure 5.7 shows the temperature gradient scan from the original paper. Figure 5.8 shows the results from `gryffin` with the nonlinear phase mixing term. We find that we over predict the heat flux at lower temperature gradients - near marginal stability, and under predict at higher temperature gradients, thus missing the steep dependence on temperature gradient scale length originally reported by Mikkelsen and Dorland. This suggests that our current model of nonlinear phase mixing is too crude to capture an accurate dependence of flux on L_T in a model that also includes additional complexities such as collisionallity.

5.4 Local Limit

The final test case we consider comes from a study of the radial profile of the turbulent diffusivity as a function of $\rho_\star$, the ratio of the ion sound gyroradius to

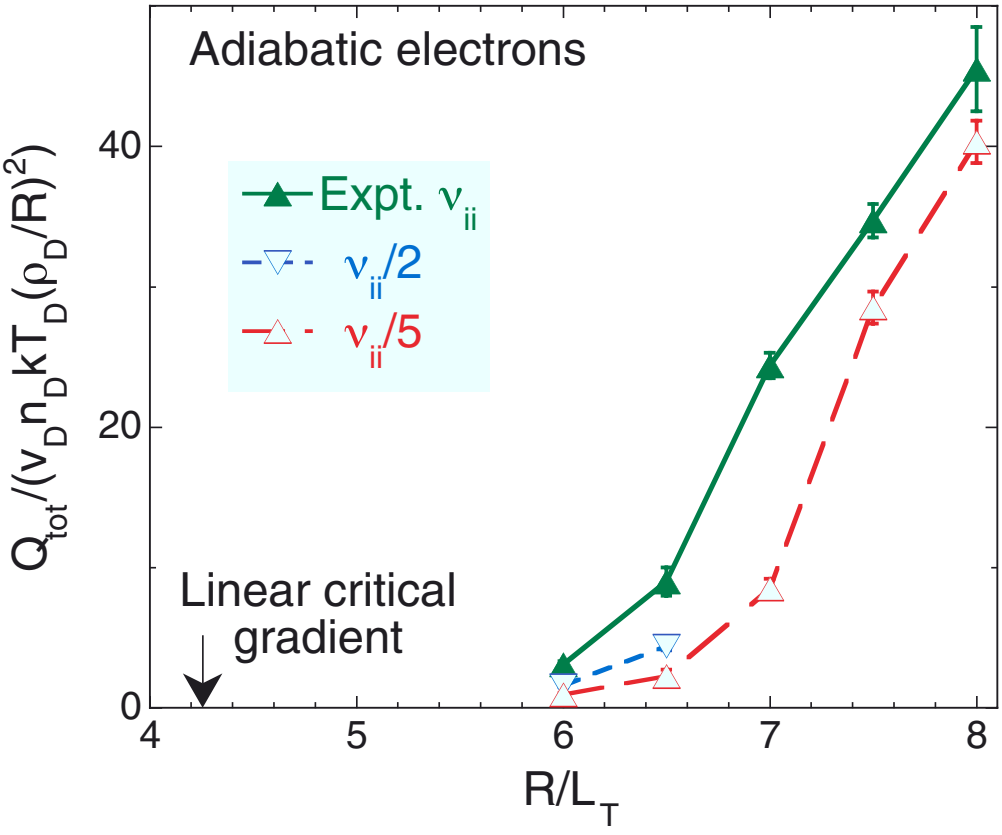

Figure 5.7: Figure 2 from the Mikkelsen and Dorland paper [29] showing the heat flux as a function of temperature gradient scale length (L_T) using adiabatic electrons and realistic geometry. The heat flux shows a stiff dependence on L_T.

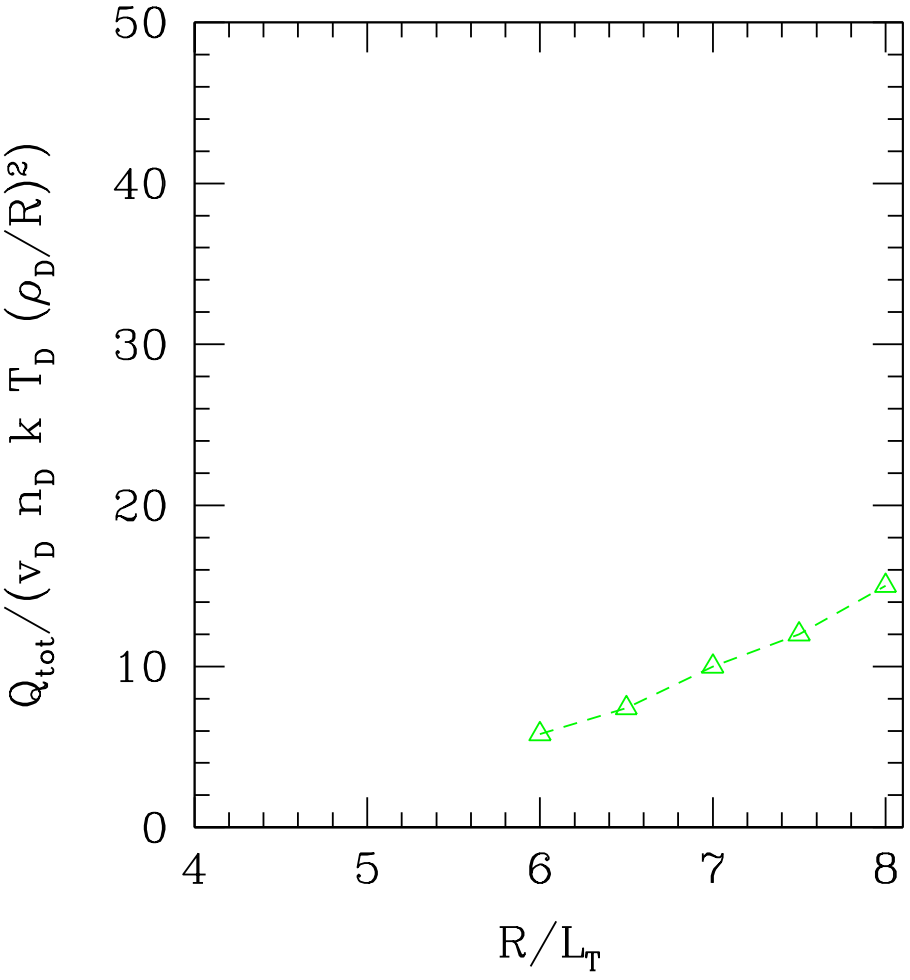

Figure 5.8: Results from `gryffin` show underdamping at low L_T, and overdamping at high L_T, thus missing the stiff dependence of the heat flux on the temperature gradient scale length.

the minor radius. In the limit of small $\rho_\star$, gyrokinetic codes based on a local, flux tube geometry should predict the same diffusivities as global gyrokinetic codes. [2] Cyclone base case parameters are again employed. Figure 5.9 is from the original paper by Candy, Waltz, and Dorland [8] and shows that this is indeed the case.

We ran the gyrofluid code for the same r/a values using a varying safety factor profile and a varying temperature gradient profile given by

$$q\left(r\right) = 0.854 + 2.184r^2 \tag{5.1}$$

$$\frac{dT}{dr} = \left(\frac{dT}{dR}\right)_0 \left[1 + 0.3\left(r/a - 0.5\right)\right] \tag{5.2}$$

as in the original paper. Our results are presented in Table 5.1. The second and third column are the numerical values read from Figure 5.9. The fourth and fifth columns have converted the numbers to units consistent with `gryffin`'s output. Column six reports the values from `gryffin`. `gryffin` is able to capture the trend of the diffusivity radial profile.

5.5 Conclusions

Overall, the inclusion of nonlinear phase mixing brings gyrofluid-predicted, turbulent heat flux and diffusivity into better agreement with gyrokinetic predictions. Our current model of nonlinear phase mixing appears to suppress turbulent flux even in the case of little to no persistent zonal flows as shown in the trapped particle example. At the same time, it does appear to have move of an effect when

[2] The local, flux tube geometry is explicitly derived in the small $\rho_\star$ limit.

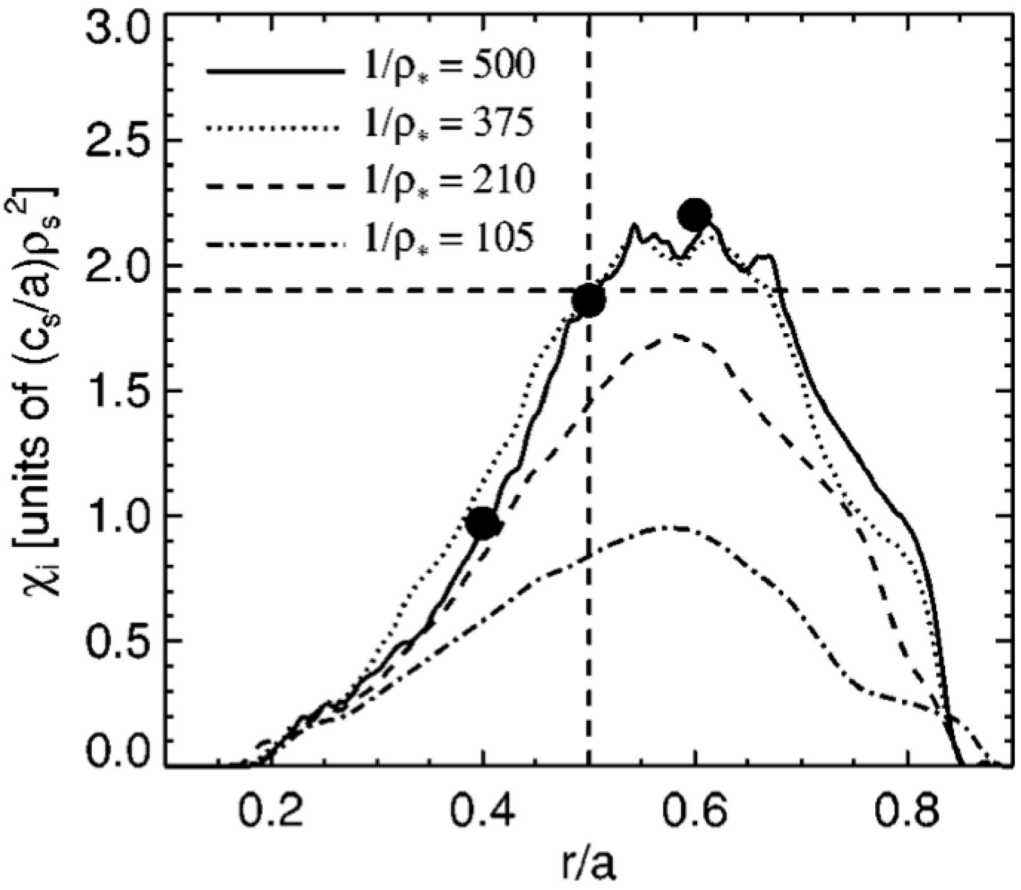

Figure 5.9: Time average diffusivities ξ or Cyclone base case parameters and varying values of $\rho_\star$, the ratio of the ion sound gyroradius to the minor radius. The lines are from a global gyrokinetic code, `GYRO`, and the dots are from the local gyrokinetic code `gs2`. The results from the local gyrokinetic code agree with the global code in the limit of small $\rho_\star$, where the local approximations are valid. [8]

Table 5.1: Comparisons of Diffusivity

r/a	Global GK ($\frac{\rho^2 v_t}{a}$)	Local GK ($\frac{\rho^2 v_t}{a}$)	Glob GK ($\frac{\rho^2 v_t}{L_n}$)	Loc GK ($\frac{\rho^2 v_t}{L_n}$)	Loc GF ($\frac{\rho^2 v_t}{L_n}$)
0.4	1.0	1.0	1.2	1.2	1.8
0.5	1.9	1.9	2.4	2.4	2.3
0.6	2.1	2.2	2.6	2.8	3.0

persistent zonal flows are present, as demonstrated in the ETG examples. However, we find that the current model for the nonlinear phase mixing tends to overdamp the flux at higher temperature gradient scale lengths and underdamp at lower temperature gradient scale lengths. This trend is consistent for the first two and the fourth studies presented here, but is particularly evident in the case where realistic geometry is used and collisionality is included.

From these results, we see that our current model is too crude to accurately track the effect of nonlinear phase mixing for higher temperature gradient scale lengths particularly when richer physics is included in the model as well. In the future, we hope to improve the crude model presented here by retaining the spatial dependence of the $\mathbf{E} \times \mathbf{B}$ drift's contribution.

Even without the improved model for nonlinear phase mixing, the predicted fluxes from `gryfifn` are now close enough to those predicted by gyrokinetics codes to be used in a turbulent transport solver where the details of turbulent flux calcu-

lations are less important than the overall flux predicitons. In the next chapter, we discuss plans for such a coupling.

Chapter 6

Coupling `gryffin` to `TRINITY`

6.1 Introduction

`TRINITY` is a turbulent transport solver designed to predict and model density and temperature profiles in fusion devices. Fusion devices often have equilibrium scale gradients on the order of the size of the device (meters) in addition to the turbulent structures on the order of the ion Larmor radius (millimeters). Also, transport time scales are on the order of the lifetime of the discharge (seconds) while turbulent fluctuations vary on the order of milliseconds.

To resolve behavior at these disparate scales using a single algorithm or grid, the most direct approach, would require simulation sizes and computational time far beyond the capacity and lifetime of current supercomputers (see Table 6.1). Instead, carefully constructed algorithms can be used to create density and temperature equilibrium profiles based on underlying turbulence values. `TRINITY` is designed to do precisely that. In Section 6.2, we describe the transport equations that `TRINITY` solves. In Section 6.3, we describe the algorithm employed.

6.2 `TRINITY` Equations

`TRINITY` evolves the density, toroidal angular momentum (summed over species),

Table 6.1: Computational Cost of Simulations

Model	Spatial Resolution	Spatial Grid Points	Velocity Resolution	Velocity Grid Points	Time Resolution	Time Grid Points	Total
Direct Approach	10^{-3} cm	10^{15}	10^{7} cm/s	10^{6}	10^{-7} s	10^{7}	10^{28}
Full - f Gyrokinetics	$10^{-3}/10$ cm	10^{11}	10^{7} cm/s	10^{4}	10^{-6} s	10^{6}	10^{21}
Coupled Gyrokinetic	$10^{-3}/10$ cm	10^{11}	10^{7} cm/s	10^{4}	$10^{-6}/10^{-2}$	10^{5}	10^{20}
Coupled Gyrofluid	$10^{-3}/10$ cm	10^{11}	N/A	N/A	$10^{-6}/\ 10^{-2}$	10^{5}	10^{16}

and pressure moments of the large scale, slowly varying equilibrium of the plasma. Like the other fluid models described herein, the derivation begins with the Fokker-Planck equation, Equation A.32, and the gyrokinetic orderings are applied (see Equations A.1 and A.2). The slow variation of the background distribution function enters at order ϵ^2, the same order as the time and space variations of δf_2.

In order to avoid solving for δf_2, we assume that there are no intermediate scale dynamics that are important to either the equilibrium evolution or the turbulence. This allows us to average over the intermediate scales in both space and time for which we expect the variations in δf_2 to average to zero. We choose to ignore time variations of the magnetic field since it evolves on a slower resistive time scale. The resultant equations can be written in a simplified form as

$$\frac{\partial n}{\partial \tau} = -\frac{\partial \psi}{\partial V}\frac{\partial}{\partial \psi}\left(\frac{\partial V}{\partial \psi}\left\langle\left\langle \mathbf{\Gamma}\cdot\nabla\psi\right\rangle\right\rangle\right) + C + \left\langle\left\langle S_n\right\rangle\right\rangle \tag{6.1}$$

$$\frac{\partial \left\langle L\right\rangle_t}{\partial t} = -\sum_s \frac{\partial \psi}{\partial V}\frac{\partial}{\partial \psi}\left(\frac{\partial V}{\partial \psi}\left\langle\left\langle \mathbf{\Pi}\cdot\nabla\psi\right\rangle\right\rangle\right) + \sum_s \left\langle\left\langle S_{L_s}\right\rangle\right\rangle \tag{6.2}$$

$$\frac{3}{2}\frac{\partial p}{\partial \tau} = -\frac{\partial \psi}{\partial V}\frac{\partial}{\partial \psi}\left[\frac{\partial V}{\partial \psi}\left\langle\left\langle \mathbf{Q}\cdot\nabla\psi\right\rangle\right\rangle\right] + C + S_p + \frac{3}{2}n\sum_u \nu_u^{\varepsilon}\left(T - T_u\right) \tag{6.3}$$

The double brackets represent the spatial and temporal averaging operators. C represents the contribution from collisions at this order that are associated with neoclassical transport. $S_{n,L_s,p}$ represent source terms for density, species-specific toroidal angular momentum, and temperature, respectively. The last term in the heat flux equation represents heat exchange between species. The flux functions in the gyrokinetic formulation are written as

$$\mathbf{\Gamma} = \int_{-\infty}^{\infty}\left[\mathbf{v}_\chi \delta f_{1,h}\right] d^3v \tag{6.4}$$

$$\mathbf{\Pi} = \int_{-\infty}^{\infty}\left[\left(mR^2\mathbf{v}\cdot\nabla\phi\right)\mathbf{v}_\chi \delta f_{1,h}\right] d^3v \tag{6.5}$$

$$\mathbf{Q} = \int_{-\infty}^{\infty}\left[\frac{mv^2}{2}\mathbf{v}_\chi \delta f_{1,h}\right] d^3v \tag{6.6}$$

where R is the major radius and ϕ is the toroidal angle. $\mathbf{v}_\chi$ represents the drifts from the perturbed fields $\frac{c}{B_0}\hat{\mathbf{b}}\times\nabla\chi$ where $\chi = \Phi + \frac{1}{c}\mathbf{v}\cdot\mathbf{A}$. For the gyrofluid model, we integrate over velocity. Using the moments from the gyrofluid model described in Chapter 4 these are written as

$$\mathbf{\Gamma} = \left[\Gamma_0^{1/2} n + \frac{1}{2}\hat{\nabla}_\perp^2 \Gamma_0^{1/2} T_\perp - (1-\Gamma_0)\,\Phi\right] \nabla\Phi \tag{6.7}$$

$$\mathbf{\Pi}_s = m_s R^2 \left[\frac{B_\phi v_{t,s}}{RB_0}\left(\Gamma_0^{1/2} u_\parallel + \frac{1}{2}\hat{\nabla}_\perp^2 \Gamma_0^{1/2} q_\perp\right) + i\frac{\mathbf{k}\cdot\nabla\psi}{R^2 B_0}\frac{v_{t,s}^2}{\Omega_{0,s}}\left(\Gamma_0^{1/2}\left(n + T_\perp\right) + \frac{1}{2}\hat{\nabla}_\perp^2 \Gamma_0^{1/2} T_\perp\right)\right] \nabla\Phi \tag{6.8}$$

$$\mathbf{Q} = \left[\Gamma_0^{1/2}\left(T_\perp + \frac{1}{2}T_\parallel + \frac{3}{2}n\right) + \frac{1}{2}\hat{\nabla}_\perp^2 \Gamma_0^{1/2}\left(n + T_\perp\right)\right] \nabla\Phi \tag{6.9}$$

where $v_{t,s}$ is the thermal speed for a given species s and $\Omega_{0,s}$ is the gyrofrequency for a given species s. The transport equations are closed by prescribing boundary conditions. At the outer edge, boundary conditions can be set by experimental values or by an analytic form that is appropriately well behaved. The inner boundary condition is assumed to be the magnetic axis.

6.3 TRINITY Algorithm

TRINITY has the arduous task of trying to resolve short turbulent time scales while simultaneously capturing behavior on much longer transport scales in fusion devices. Its approach is to run copies of a sophisticated turbulence code for various points on a rough spatial and temporal grid on the scale of the full device and the lifetime of the discharge. Inputs for the turbulence code are dependent on their position within the grid. The turbulence calculations are allowed to run until the turbulence reaches a state where the time-averaged flux does not vary much - a sort of quasi steady state. [1] The flux values are then used at the coarse grid scale and at each coarse time step, Newton's method is used to iterate to a converged solution.

[1] True steady state is not practical in a turbulent system. However, turbulence can settle into a

Fluxes from the turbulence code at each coarse grid point depend nonlinearly on the local values of density, temperature, and angular momentum of the overall system. In order to take large, implicit time steps on the transport level (to solve Equations 6.1 - 6.3), the nonlinear terms involving the flux functions need to be linearized. This is done using a Taylor expansion about the grid point, where spatial dependence is assumed to come through the moments.

In fact, empirical evidence from experiment and numerical simulations suggests that the fluxes are stiffly dependent on the *gradients* of the moments at each grid point. `TRINITY`'s algorithm is designed to take advantage of this stiffness by only considering the local, partial derivative of the fluxes with respect to the gradient scale lengths. To find the local, partial derivative, four turbulence simulations are run at each grid point: the original plus one each for a slightly different value of one of the gradient scale lengths - density, electron temperature, and ion temperature.

Fortunately, `TRINITY` only requires a single number per turbulence simulation per species for each of the fluxes. This significantly reduces communication from the turbulence simulations to the transport calculation. Likewise, communication of the evolution of the profile density and temperature gradients to the turbulence calculation requires minimal communication.

The minimal communication coupled with the independent turbulence simulations makes `TRINITY` a good candidate for improved efficiency through a CPU/GPU hybrid implementation. The next chapter describes GPU computing and ends with

state where it appears to be fluctuating around a nearly constant value. This constant value can be considered the flux of the quasi steady state.

a discussion of how TRINITY can be ported to a CPU/GPU cluster.

Chapter 7

GPU Computing

7.1 Introduction

In the last couple of years, GPU computing has established itself as a likely candidate to join the ranks of High Performance Scientific Computing options. Scientific research groups have reported code speed-ups of anywhere between 10 and 150 times their CPU versions. Graphical Processing Units (GPUs) were originally developed to speed up graphics rendering, in particular to allow computer games to include more visual details without slowing down the pace of the game. As a result, designers developed chips with many arithmetic logic units (ALUs), making it more efficient to calculate and extrapolate image data than to store and retrieve it. ALUs are where most of the numerical calculations on a chip occur and as their number increases, so does the potential to perform more calculations per clock cycle. It also creates the potential for these calculations to happen in parallel, making GPUs particularly attractive for scientific computing applications, and allows for the code accelerations reported.

In the past, however, it has only been possible to program GPUs using an application programming interface (API) specifically designed for graphics. This required not only specialized knowledge of graphics computer languages, but also the ability to cast one's problem in a graphics framework. This programming model

had a reputation for a steep learning curve and was generally avoided.

In 2007, nVIDIA released a new approach to programming on GPUs, particularly for non-graphics programmers. Compute Unified Device Architecture (CUDA) was designed as a hybrid software and hardware architecture for utilizing GPUs with its own APIs. CUDA as a language is a natural extension of C for working on a parallel architecture, making it readily accessible to C programmers. Unfortunately, the hybrid nature of the approach does mean that CUDA only works on nVIDIA hardware. [1]

CUDA was designed with the C/C++ programmer in mind. However, it is not impossible for FORTRAN programmers and codes to take advantage of CUDA. In this chapter, we describe the middleware library we developed to allow FORTRAN to take advantage of CUDA and GPUs. We also describe some of our successes and failures in adapting CUDA to our scientific applications.

7.2 FLAGON

FLAGON (Fortran-9x Library for GPU Numerics) is an interface between C-based nVIDIA CUDA and scientific codes written in FORTRAN-9x. The goal of FLAGON is to allow a user to program on a GPU in a framework that is native to FORTRAN. In order to do this, FLAGON allows the user to call functions and subroutines that allocate data on the GPU, transfer data between the GPU and

[1] An open source language known as OpenCL has been developed which is based very closely on CUDA and can be used on a wide variety of proprietary architectures. Its main focus is creating a language that can be used across heterogenous hardware.

CPU, and call the CUDA FFT library, `CUFFT`, the CUDA linear algebra package, `CUBLAS`, and the open source scanning package, `cudpp`. FLAGON also includes a consistent framework to call kernels written on the GPU. All of this is done through wrappers which hide the C/C++ specific details.

FLAGON's framework to call GPU kernels makes it particularly powerful and easy to use. Loops on CPUs are easily ported to GPU kernels. For example, consider the following simple loop:

```
do i = 1, N
   a(i) = b * c(i)
end do
```

where `a` and `c` are arrays of rank 1, and `b` is a scalar.

The CPU executes this code in serial, meaning that the middle line is executed sequentially `N` number of times. On a GPU, the number of serial trips through the loop can be replaced by many threads doing the same line of code at the same time, but with different `i` values. The corresponding code in a CUDA kernel would look like

```
int tid = blockIdx.x*blockDim.x + threadIdx.x;
a[tid] = b * c[tid]
```

where `blockIdx.x`, `blockDim.x`, and `threadIdx.x` are CUDA intrinsic variables unique to each thread.

FLAGON provides a way to call GPU kernels from FORTRAN using the CUDA Driver API. This API assigns objects, such as kernels, modules, or processes,

"handles" or unique identifiers that are assigned when the object is created. These handles are then passed to functions that assign it various attributes or act on it some way. For example, a handle belonging to a kernel would be sent to a function that assigns it a certain number of threads organized into a user-specified format. The handle might then be sent to a function that tells it what parameters the kernel will need for execution - including pointers to any arrays. Ultimately, the handle would be sent to a function that actually executes the kernel.

FLAGON hides all of this handle passing, written in C, in wrapper functions. These wrapper functions accept as input the name of the kernel, the parameters needed by the kernel (organized into arrays), and the block size and shapes desired for execution. Block size and shape determinations are left to the user since these can be factors in optimization that can vary from application to application.

A typical program might be structured in the following way:

- FLAGON is opened and some initial parameters are set with a call to `open_devObjects`
- The file containing specialized kernels is loaded onto the GPU with a call to `fc_LoadDevFunc`.
- Memory on the GPU is allocated by defining a data object of type `devVar`. This is done through use of the function `allocate_dv`. This data type stores the pointer to the GPU memory as well as information about the data size, shape, and type.
- Data is initialized on the CPU and transfered to the GPU through the use of the `transfer` subroutines.

- Data on the GPU is manipulated through the use of custom-made kernels. Kernels are launched through calls to `devf_explicit_execute`. Arguments passed to these execute subroutines require their arguments in a specific order: first, the number of `devVar`'s needed by the kernel; second, the number of integer parameters needed by the kernel; third, the number of real parameters needed by the kernel; fourth, an array of pointers to the `devVar`'s; fifth, an array of the integer parameters; sixth, an array of the real parameters; seventh, the number of bytes of shared memory needed by the kernel; eighth - tenth, the number of thread blocks desired in the x-,y-,and z-directions; eleventh and twelfth, the number of grids desired in the x- and y-directions.

- Data on the GPU is transfered back to the CPU for diagnostics, plotting, etc.

- FLAGON is closed with a call to `close_devObjects`

7.3 Test Cases

7.3.1 Orszag-Tang

As a proof of principle exercise, we took the code developed to solve the Orszag-Tang reduced MHD equations (see Section 3.4) and ported it to the GPU. This version of this code only used the explicit part of the algorithm described in Section 3.3 and is structured such that each mode can be evolved in time almost in parallel. Loops, like those used in the time-stepping algorithm, were replaced by kernel calls. The calculation of the nonlinear terms requires the use of FFTs, a

process that requires information from all of the modes. These can be solved using CUDA's FFT library, `CUFFT`.

This code gave us the opportunity to test FLAGON's ability to transfer data to and from the GPU, to execute custom-made kernels, and to interact with the CUDA libraries. Table 7.1 compares the amount of time the Orszag-Tang code took to run on a GPU, a dual core CPU, and on eight cores of a CPU. The GPU used was a GeForce 9800 GTX with 128 cores and 512 MB of memory. The dual core CPU was an Intel® Pentium® 4. The eight cores were part of the Supercomputer Bassi at NERSC, and the processors are IBM Power 5 cores.

Table 7.1: Comparison of run times for an Orszag-Tang Reduced MHD code. N^2 is the problem size. The next three columns are total run times reported in seconds. Numbers in parenthesis indicate the number of processors on the CPU. The last two columns are ratios. Run times were calculated to include initializations and I/O time.

$\mathbf{N^2}$	**CPU (2)**	**CPU (16)**	**GPU**	**CPU (16)/GPU**	**CPU (2)/GPU**
256^2	209	209	26	8	8
1024^2	40,991	19,107	1,455	13	28

7.3.2 MPI and FLAGON

In addition to the two-dimensional Orszag-Tang Reduced MHD code, we ported a three-dimensional Reduced MHD code with twice as many nonlinear terms to the GPU. On the GeForce 9800 GTX GPU, we were able to run on a grid of 512 x 512 x 4 at a rate of approximately 6 timesteps/sec. Like the Orszag-Tang code, this includes 2 FFTs each timestep.

However, in order to run the more sophisticated six-moment model with a fully hybrid algorithm, we would need more memory. As a result, we added the capability of using MPI with FLAGON and developed a dual GPU version of the original Orszag-Tang code.

This dual GPU version splits the calculation such that one of the two variables (stream function or parallel magnetic potential) resides on each card. Since the equations are coupled, copies of the variables are exchanged once each time step. The calculation of kinetic energy is performed on one card, magnetic energy on the other. Both are sent back to the CPU and then to a single thread to be written to output.

We compared the dual GPU version to the single GPU version, to a serial CPU version, and to an OpenMP version of the code. The OpenMP version of the code is similar to the serial version of the code, but the compiler parallelizes the do-loops. For the machine and architecture used, four threads were spawned to perform the work.

We timed the four versions of the code for five different problem sizes. The

three smallest sizes were run with a time step of 10^{-3}. The two larger sizes were run with a time step of 10^{-4}. The results can be found in Table 7.2.

Table 7.2: Time for a single run of the Orszag-Tang code in seconds.

Size	**1 GPU**	**2 GPUs**	**CPU**	**OpenMP**
64^2	17.2	16.2	1.4	2.3
128^2	19.1	18.2	5.7	7.8
256^2	27.2	28.3	31.4	34.4
512^2	612.0	796.5	2122.4	2249.4
1024^2	2123.4	3135.4	12075.0	12458.2

For the smallest two problem sizes, the serial CPU version runs faster than the other code versions. We assume that the overhead to spawn OpenMP threads is greater than the advantage of having multiple threads perform the calculations for so small a problem size. However, this does not explain why the OpenMP version runs slower even for larger problem sizes.

Assuming the overhead is the same whether four threads or eight threads are spawned, we ran the 512 x 512 case with eight threads. However, this calculation took nearly twice as long - 4120.6 sec. If we decreased the number of threads to two, the calculation took 2182.4 sec. This implies that whatever is causing the OpenMP version to run longer is dependent on the number of threads.

For the smaller problem sizes, we assume that the time it takes to transfer

data between the CPU and GPU dominates the run time for the GPU versions of the code causing them to take longer than the CPU versions.

For the larger problem sizes, the single GPU code runs more quickly than the MPI dual GPU code. This implies that the time it takes to transfer data from one GPU to another is greater than the time saved by splitting up the calculation. We timed how long it took transfer data between the CPU and GPU and the time it took to transfer data between two MPI threads. The results are found in Table 7.3. The larger problem sizes require a smaller time step, thus requiring more data transfers per run.

Our Orszag-Tang code does not fit well into the distributed memory, embarrassingly parallel regime that one might expect to benefit the most from a combined GPU, MPI approach. However, it does serve to show that MPI can be used together with FLAGON to generate correct results. In addition, this exercise allowed the development of functions required for such an approach.

7.4 Conclusions and Future Work

GPU computing is a powerful tool for solving complicated physical systems. As long as the code maintains an 'embarassingly' parallel structure, a single problem can be distributed over several GPUs. However, if communication between GPUs is necessary, the technology does not yet exist to make this efficient.

nVIDIA, however, has taken a different approach and has released a new line of GPUs known as the "Fermi" or "Tesla" series. The newest cards boast 6.0 GB of

Table 7.3: Amount of time for data transfers in the Orszag-Tang code. Time values are in seconds. The GPU/CPU time reflects the total time it takes to transfer data to the GPU and back to the CPU. This dual transfer occurs once per GPU per time step. The CPU/CPU time reflects the total time it takes for each MPI thread to receive and send data. This operation also occurs once per time step.

Size	**GPU / CPU**	**CPU / CPU**
64 x 64	7.8×10^{-5}	6.3×10^{-5}
128 x 128	1.6×10^{-4}	1.6×10^{-4}
256 x 256	3.9×10^{-4}	5.9×10^{-4}
512 x 512	1.6×10^{-3}	3.4×10^{-3}
1024 x 1024	6.0×10^{-3}	1.8×10^{-2}

memory and 448 cores. Even our more sophisticated model and algorithm described in Chapters 2 and 3 should be able to fit on a card this size.

The high parallelism of `TRINITY`'s algorithm makes it an ideal candidate to be run on a heterogeneous architecture where the more sparse, equilibrium level grid can live on the CPU and the individual turbulence calculations can be run each on its own GPU. Fortunately, `gryffin` is a small enough code that it should easily fit onto the state of the art GPUs. While `gryffin` is not in a position to capture the details of accurate turbulence and transport calculations, it can certainly be used to direct the fusion community's focus to potentially interesting regions of parameter space. In this way, we hope to create a tool for understanding the vast parameter space for fusion devices that have as of yet gone unexplored due to limited resources.

Appendix A

Derivation of the Gyrokinetic Equations

This appendix contains a derivation of the gyrokinetic equation that is used as a basis for taking moments to develop the gyrofluid equations. I follow similar derivations worked out by Greg Howes, et al.[21] and Alex Schekochihin, et al. [33].

A.1 Initial Assumptions

These equations are derived in a slab geometry with a background magnetic field, $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, and a spatially uniform equilibrium distribution function, $\nabla F_0 = 0$. Weak coupling, strong magnetization, low frequencies, and small fluctuations are assumed. Weak coupling is inherent to the definition of a plasma. Small fluctuations mean that the ion Larmour radius is much smaller than the macroscopic scale length(ie the system size), and low frequencies mean that structures of interest are evolving much more slowly than the ion cyclotron frequency. In symbols,

$$\rho_i \ll L \quad , \quad \omega \ll \Omega_i \tag{A.1}$$

Our ordering parameter is defined based on length scales so that

$$\epsilon = \frac{\rho_i}{l_0} \ll 1 \tag{A.2}$$

where l_0 is a typical parallel wavelength of the fluctuations of the system. We also assume a separation of three time scales:

- A large ion cyclotron frequency, Ω_i
- A smaller turbulence/fluctuation frequency, $\omega \sim \frac{v_{ti}}{l_0} \sim \mathcal{O}(\epsilon\Omega_i)$
- And a transport rate, $\frac{1}{t_{heat}} \sim \epsilon^2\omega \sim \mathcal{O}(\epsilon^3\Omega_i)$

Collisions are on the order of the turbulence frequency.

There are only four quantities of which we keep track: the ion distribution function, the electron distribution function, the magnetic field, and the electric field. These quantities can be expanded in orders of epsilon. The equilibrium portion of all four quantities is assumed to vary on the slowest time scale, that of the transport rate. We choose to work in a frame where the electric field has no equilibrium component, and there are no initial equilibrium flows. Our quantities can be written as follows:

$$f_s(\mathbf{r},\mathbf{v},t) = F_{0,s}(\mathbf{v},t) + \delta f_{1,s}(\mathbf{r},\mathbf{v},t) + \delta f_{2,s}(\mathbf{r},\mathbf{v},t) + \ldots \tag{A.3}$$

$$\mathbf{B}(\mathbf{r},t) = \mathbf{B}_0 + \delta\mathbf{B}(\mathbf{r},t) = B_0\hat{\mathbf{z}} + \delta B_{\parallel}(\mathbf{r},t)\,\hat{\mathbf{z}} + \nabla \times A_{\parallel}(\mathbf{r},t)\,\hat{\mathbf{z}} \tag{A.4}$$

$$\mathbf{E}(\mathbf{r},t) = \delta\mathbf{E}(\mathbf{r},t) = -\nabla\phi(\mathbf{r},t) - \frac{1}{c}\frac{\partial}{\partial t}\mathbf{A}(\mathbf{r},t) \tag{A.5}$$

where f_s is the particle distribution function with the subscript s denoting the species, and ϕ and $\mathbf{A}$ are the electric and magnetic potentials, respectively.

Perturbed fields are assumed to order as follows:

$$\frac{\delta B}{B_0} \sim \epsilon \quad ; \quad \frac{\delta E}{B_0 v_t/c} \sim \epsilon \tag{A.6}$$

Spatial scales are anisotropic so that fluctuations perpendicular to the background magnetic field are on the order of the ion Larmor radius, and the fluctuations

parallel to the background magnetic field are on the order of the system size.

$$k_\perp \sim \frac{\hat{\mathbf{z}} \times \nabla \delta f}{\delta f} \sim \frac{\hat{\mathbf{z}} \times \nabla \delta \mathbf{B}}{|\delta \mathbf{B}|} \sim \frac{\hat{\mathbf{z}} \times \nabla \delta \mathbf{E}}{|\delta E|} \sim \mathcal{O}\left(\frac{1}{\rho_i}\right) \tag{A.7}$$

$$k_\parallel \sim \frac{\hat{\mathbf{z}} \cdot \nabla \delta f}{\delta f} \sim \frac{\hat{\mathbf{z}} \cdot \nabla \delta \mathbf{B}}{|\delta \mathbf{B}|} \sim \frac{\hat{\mathbf{z}} \cdot \nabla \delta \mathbf{E}}{|\delta E|} \sim \mathcal{O}\left(\frac{1}{l_0}\right) \tag{A.8}$$

A.2 Useful Mathematical Definitions

We can describe the particles' position as:

$$\mathbf{r} = \mathbf{R}_{gc} - \frac{\mathbf{v} \times \hat{\mathbf{z}}}{\Omega_i} \tag{A.9}$$

where $\mathbf{R}_{gc}$ is the position of the guiding center. The velocity can be expressed (in cylindrical coordinates) as

$$\mathbf{v} = v_\parallel \hat{\mathbf{z}} + v_\perp \left(\cos\theta \hat{\mathbf{x}} + \sin\theta \hat{\mathbf{y}}\right) \tag{A.10}$$

where $v_\parallel$ and $v_\perp$ are constants.

We will sometimes use the following coordinates to describe the velocity

$$\varepsilon = \frac{mv^2}{2} + q\phi \quad ; \quad \mu = \frac{v_\perp^2}{2B_0} \quad ; \quad \theta = \tan^{-1}\left(\frac{v_y}{v_x}\right) \tag{A.11}$$

where ε is the energy of the particle, μ is the magnetic moment, and θ is the gyroangle. The velocity derivative in these coordinates is

$$\frac{\partial}{\partial \mathbf{v}} = \mathbf{v}\frac{\partial}{\partial \varepsilon} + \frac{\mathbf{v}_\perp}{B_0}\frac{\partial}{\partial \mu} - \frac{\mathbf{v}_\perp \times \hat{\mathbf{z}}}{v_\perp^2}\frac{\partial}{\partial \theta} \tag{A.12}$$

The differential in the integral becomes

$$d^3\mathbf{v} = \frac{B_0}{v_\parallel}\partial\varepsilon\partial\mu\partial\theta \tag{A.13}$$

The derivative w.r.t. time of the particle position can be expressed in guiding center coordinates as follows:

$$\left(\frac{\partial}{\partial t}\right)_{\mathbf{r}} = \left(\frac{\partial}{\partial t}\right)_{\mathbf{R}_{gc}} + \frac{d\mathbf{R}_{gc}}{dt}\cdot\frac{\partial}{\partial \mathbf{R}_{gc}} \tag{A.14}$$

The derivative w.r.t. time in guiding center coordinates can be written as follows

$$\begin{aligned}
\frac{d\mathbf{R}_{gc}}{dt} &= \mathbf{v} + \frac{1}{\Omega_c}\mathbf{a}\times\hat{\mathbf{z}} && \text{(A.15)}\\
&= \mathbf{v} + \frac{1}{\Omega_c}\frac{q}{m}\left(\delta\mathbf{E} + \frac{1}{c}\mathbf{v}\times B_0\hat{\mathbf{z}} + \frac{1}{c}\mathbf{v}\times\delta B\right)\times\hat{\mathbf{z}} && \text{(A.16)}\\
&= \mathbf{v} + \frac{c}{B_0}\delta\mathbf{E}\times\hat{\mathbf{z}} - \mathbf{v}_\perp + v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0} - \mathbf{v}_\perp\frac{\delta B_\parallel}{B_0} && \text{(A.17)}\\
&= v_\parallel\hat{\mathbf{z}} + \frac{c}{B_0}\delta\mathbf{E}\times\hat{\mathbf{z}} + v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0} - \mathbf{v}_\perp\frac{\delta B_\parallel}{B_0} && \text{(A.18)}\\
&= v_\parallel\hat{\mathbf{z}} + \frac{c}{B_0}\delta\mathbf{E}\times\hat{\mathbf{z}} + v_\parallel\frac{\delta\mathbf{B}_\perp}{B_0} - \mathbf{v}_\perp\frac{\delta B_\parallel}{B_0} && \text{(A.19)}\\
& && \text{(A.20)}
\end{aligned}$$

Expressing this in terms of potentials:

$$\begin{array}{lllll}
\dfrac{d\mathbf{R}_{gc}}{dt} = v_\parallel\hat{\mathbf{z}} & -\dfrac{c}{B_0}\nabla_\perp\phi\times\hat{\mathbf{z}} & -\dfrac{1}{B_0}\dfrac{\partial\mathbf{A}_\perp}{\partial t}\times\hat{\mathbf{z}} & +\dfrac{v_\parallel}{B_0}\delta\mathbf{B}_\perp & -\dfrac{\delta B_\parallel}{B_0}\mathbf{v}_\perp \\
v_{th} & \frac{c}{B_0}\frac{1}{\rho_s}\epsilon B_0\frac{v_{th}}{c}\rho & \frac{1}{B_0}\omega\epsilon B_0\rho & \frac{v_{th}}{B_0}\epsilon B_0 & \epsilon\frac{B_0}{B_0}v_{th}
\end{array} \tag{A.21}$$

The top line contains the terms in the equation, and the bottom line contains

various constants and the expansion parameter. I will be using this method to express the order of various terms throughout the rest of this derivation.

$$\begin{array}{lllll} \frac{d\mathbf{R}_{gc}}{dt} & = v_{\parallel}\hat{\mathbf{z}} & -\frac{c}{B_0}\nabla_{\perp}\phi\times\hat{\mathbf{z}} & -\frac{1}{B_0}\frac{\partial\mathbf{A}_{\perp}}{\partial t}\times\hat{\mathbf{z}} & +\frac{v_{\parallel}}{B_0}\delta\mathbf{B}_{\perp} \quad -\frac{\delta B_{\parallel}}{B_0}\mathbf{v}_{\perp} \\ & v_{th} & \epsilon v_{th} & \epsilon^2 v_{th} & \epsilon v_{th} \quad\quad \epsilon v_{th} \end{array} \tag{A.22}$$

Keeping terms through $\mathcal{O}(\epsilon)$

$$\frac{d\mathbf{R}_{gc}}{dt} = v_{\parallel}\hat{\mathbf{z}} - \frac{c}{B_0}\nabla_{\perp}\phi\times\hat{\mathbf{z}} + \frac{v_{\parallel}}{B_0}\delta\mathbf{B}_{\perp} - \frac{\delta B_{\parallel}}{B_0}\mathbf{v}_{\perp} \tag{A.23}$$

We will also use the second derivative w.r.t. time

$$\frac{d^2\mathbf{R}}{dt^2} = \frac{1}{v_{\parallel}}\frac{d\varepsilon}{dt}\hat{\mathbf{z}} - \frac{B_0}{2v_{\parallel}}\frac{\partial\mu}{\partial t}\hat{\mathbf{z}} + \left[\frac{B_0}{v_{\perp}}\cos\theta\frac{\partial\mu}{\partial t} - v_{\perp}\sin\theta\frac{\partial\theta}{\partial t}\right]\hat{\mathbf{x}} + \left[\frac{B_0}{v_{\perp}}\sin\theta\frac{\partial\mu}{\partial t} + v_{\perp}\cos\theta\frac{\partial\theta}{\partial t}\right]\hat{\mathbf{y}} \tag{A.24}$$

Using $\frac{\partial\mu}{\partial t} = 0$ and $\frac{\partial\theta}{\partial t} = \Omega_c$, we can write

$$\frac{d^2\mathbf{R}}{dt^2} = \frac{1}{v_{\parallel}}\frac{d\varepsilon}{dt}\hat{\mathbf{z}} - v_{\perp}\sin\theta\Omega_c\hat{\mathbf{x}} + v_{\perp}\cos\theta\Omega_c\hat{\mathbf{y}} \tag{A.25}$$

The energy derivative can be expressed as

$$\frac{d\varepsilon}{dt} = m\mathbf{v}\cdot\frac{d\mathbf{v}}{dt} + q\frac{\partial\phi}{\partial t} + qv_{\parallel}\nabla_{\parallel}\phi + q\mathbf{v}_{\perp}\cdot\nabla_{\perp}\phi \tag{A.26}$$

$$= q\mathbf{v}\cdot\mathbf{E} + q\mathbf{v}\cdot\mathbf{v}\times\mathbf{B} + q\frac{\partial\phi}{\partial t} + qv_{\parallel}\nabla_{\parallel}\phi + q\mathbf{v}_{\perp}\cdot\nabla_{\perp}\phi \tag{A.27}$$

$$= -qv_{\parallel}\nabla_{\parallel}\phi - \frac{q}{c}v_{\parallel}\frac{\partial A_{\parallel}}{\partial t} - q\mathbf{v}_{\perp}\cdot\nabla_{\perp}\phi - \frac{q}{c}\frac{\partial\mathbf{A}_{\perp}}{\partial t} + q\frac{\partial\phi}{\partial t} + qv_{\parallel}\nabla_{\parallel}\phi + q\mathbf{v}_{\perp}\cdot\nabla_{\perp}\phi \tag{A.28}$$

$$= -\frac{q}{c}v_{\parallel}\frac{\partial A_{\parallel}}{\partial t} - \frac{q}{c}\frac{\partial\mathbf{A}_{\perp}}{\partial t} + q\frac{\partial\phi}{\partial t} \tag{A.29}$$

We use two types of gyroaverages, one keeps $\mathbf{R}_{gc}$ fixed and the other keeps $\mathbf{r}$ fixed. The first is the ring average at fixed guiding center:

$$\langle a\left(\mathbf{r},\mathbf{v},t\right)\rangle_{\mathbf{R}_{gc}} = \frac{1}{2\pi}\oint d\theta\; a\left(\mathbf{R}_{gc} - \frac{\mathbf{v}\times\hat{\mathbf{z}}}{\Omega_i},\mathbf{v},t\right) \tag{A.30}$$

The second is the ring average at fixed position:

$$\langle a\left(\mathbf{R},\mathbf{v},t\right)\rangle_{\mathbf{r}} = \frac{1}{2\pi}\oint d\theta\; a\left(\mathbf{r} + \frac{\mathbf{v}\times\hat{\mathbf{z}}}{\Omega_i},\mathbf{v},t\right) \tag{A.31}$$

both integrations are done keeping $\mathbf{v}$ constant.

A.3 The Fokker-Planck Equation

We begin with the Fokker-Planck equation.

$$\frac{\partial f_s}{\partial t} + \mathbf{v}\cdot\nabla f_s + \frac{q_s}{m_s}\left(-\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} + \frac{1}{c}\mathbf{v}\times\mathbf{B}\right)\cdot\frac{\partial f_s}{\partial \mathbf{v}} = \left(\frac{\partial f_s}{\partial t}\right)_c \tag{A.32}$$

where q_s is the charge for a given species, m_s is the mass of a given species, c is the speed of light, and $()_c$ represents collisions.

We consider only one species and drop the subscript s. Including perturbations and explicitly separating parallel and perpendicular directions, we can write

$$
\begin{aligned}
&\frac{\partial F_0}{\partial t}+\frac{\partial \delta f_1}{\partial t}+\frac{\partial \delta f_2}{\partial t}+\mathbf{v}_\perp\cdot\nabla_\perp\delta f_1+\mathbf{v}_\perp\cdot\nabla_\perp\delta f_2+v_\parallel\nabla_\parallel\delta f_1+v_\parallel\nabla_\parallel\delta f_2-\frac{q}{m}\nabla_\perp\phi\cdot\frac{\partial F_0}{\partial \mathbf{v}_\perp}\\
&-\frac{q}{m}\nabla_\perp\phi\cdot\frac{\partial \delta f_1}{\partial \mathbf{v}_\perp}-\frac{q}{m}\nabla_\perp\phi\cdot\frac{\partial \delta f_2}{\partial \mathbf{v}_\perp}-\frac{q}{m}\nabla_\parallel\phi\frac{\partial F_0}{\partial v_\parallel}-\frac{q}{m}\nabla_\parallel\phi\frac{\partial \delta f_1}{\partial v_\parallel}-\frac{q}{m}\nabla_\parallel\phi\frac{\partial \delta f_2}{\partial v_\parallel}-\frac{q}{m}\frac{\partial \mathbf{A}_\perp}{\partial t}\cdot\frac{\partial F_0}{\partial \mathbf{v}_\perp}\\
&-\frac{q}{m}\frac{\partial \mathbf{A}_\perp}{\partial t}\cdot\frac{\partial \delta f_1}{\partial \mathbf{v}_\perp}-\frac{q}{m}\frac{\partial \mathbf{A}_\perp}{\partial t}\cdot\frac{\partial \delta f_2}{\partial \mathbf{v}_\perp}-\frac{q}{m}\frac{\partial A_\parallel}{\partial t}\frac{\partial F_0}{\partial v_\parallel}-\frac{q}{m}\frac{\partial A_\parallel}{\partial t}\frac{\partial \delta f_1}{\partial v_\parallel}-\frac{q}{m}\frac{\partial A_\parallel}{\partial t}\frac{\partial \delta f_2}{\partial v_\parallel}\\
&+\frac{q}{mc}\mathbf{v}_\perp\times\mathbf{B}_0\cdot\frac{\partial F_0}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\perp\times\mathbf{B}_0\cdot\frac{\partial \delta f_1}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\perp\times\mathbf{B}_0\cdot\frac{\partial \delta f_2}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\perp\times\delta\mathbf{B}_\parallel\cdot\frac{\partial F_0}{\partial \mathbf{v}_\perp}\\
&+\frac{q}{mc}\mathbf{v}_\perp\times\delta\mathbf{B}_\parallel\cdot\frac{\partial \delta f_1}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\perp\times\delta\mathbf{B}_\parallel\cdot\frac{\partial \delta f_2}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\perp\times\mathbf{B}_\perp\cdot\frac{\partial F_0}{\partial \mathbf{v}_\parallel}+\frac{q}{mc}\mathbf{v}_\perp\times\mathbf{B}_\perp\cdot\frac{\partial \delta f_1}{\partial \mathbf{v}_\parallel}\\
&+\frac{q}{mc}\mathbf{v}_\perp\times\mathbf{B}_\perp\cdot\frac{\partial \delta f_2}{\partial \mathbf{v}_\parallel}+\frac{q}{mc}\mathbf{v}_\parallel\times\mathbf{B}_\perp\cdot\frac{\partial F_0}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\parallel\times\mathbf{B}_\perp\cdot\frac{\partial \delta f_1}{\partial \mathbf{v}_\perp}+\frac{q}{mc}\mathbf{v}_\parallel\times\mathbf{B}_\perp\cdot\frac{\partial \delta f_2}{\partial \mathbf{v}_\perp}\\
&=C\left(F_0,F_0\right)+C\left(\delta f_1,F_0\right)+C\left(\delta f_2,F_0\right)+C\left(F_0,\delta f_1\right)+C\left(F_0,\delta f_2\right)+C\left(\delta f_1,\delta f_1\right)+C\left(\delta f_1,\delta f_2\right)\\
&+C\left(\delta f_2,\delta f_1\right)+C\left(\delta f_2,\delta f_2\right)
\end{aligned}
\tag{A.33}
$$

We order these terms based on ω, remembering that the equilibrium quantities change on the order of the transport rate, $\epsilon^2\omega$.

$$
\begin{array}{llll}
\dfrac{\partial F_0}{\partial t} & +\dfrac{\partial \delta f_1}{\partial t} & +\dfrac{\partial \delta f_2}{\partial t} & +\mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 \\
\epsilon^2 \omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 & \omega F_0 \\
+\mathbf{v}_\perp \cdot \nabla_\perp \delta f_2 & +v_\parallel \nabla_\parallel \delta f_1 & +v_\parallel \nabla_\parallel \delta f_2 & -\dfrac{q}{m}\nabla_\perp \phi \cdot \dfrac{\partial F_0}{\partial \mathbf{v}_\perp} \\
\epsilon\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 & \omega F_0 \\
-\dfrac{q}{m}\nabla_\perp \phi \cdot \dfrac{\partial \delta f_1}{\partial \mathbf{v}_\perp} & -\dfrac{q}{m}\nabla_\perp \phi \cdot \dfrac{\partial \delta f_2}{\partial \mathbf{v}_\perp} & -\dfrac{q}{m}\nabla_\parallel \phi \dfrac{\partial F_0}{\partial v_\parallel} & -\dfrac{q}{m}\nabla_\parallel \phi \dfrac{\partial \delta f_1}{\partial v_\parallel} \\
\epsilon\omega F_0 & \epsilon^2\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 \\
-\dfrac{q}{m}\nabla_\parallel \phi \dfrac{\partial \delta f_2}{\partial v_\parallel} & -\dfrac{q}{mc}\dfrac{\partial \mathbf{A}_\perp}{\partial t} \cdot \dfrac{\partial F_0}{\partial \mathbf{v}_\perp} & -\dfrac{q}{mc}\dfrac{\partial \mathbf{A}_\perp}{\partial t} \cdot \dfrac{\partial \delta f_1}{\partial \mathbf{v}_\perp} & -\dfrac{q}{mc}\dfrac{\partial \mathbf{A}_\perp}{\partial t} \cdot \dfrac{\partial \delta f_2}{\partial \mathbf{v}_\perp} \\
\epsilon^3\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 & \epsilon^3\omega F_0 \\
-\dfrac{q}{mc}\dfrac{\partial A_\parallel}{\partial t}\dfrac{\partial F_0}{\partial v_\parallel} & -\dfrac{q}{mc}\dfrac{\partial A_\parallel}{\partial t}\dfrac{\partial \delta f_1}{\partial v_\parallel} & -\dfrac{q}{m}c\dfrac{\partial A_\parallel}{\partial t}\dfrac{\partial \delta f_2}{\partial v_\parallel} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_0 \cdot \dfrac{\partial F_0}{\partial \mathbf{v}_\perp} \\
\epsilon\omega F_0 & \epsilon^2\omega F_0 & \epsilon^3\omega F_0 & \frac{1}{\epsilon}\omega F_0 \\
+\dfrac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_0 \cdot \dfrac{\partial \delta f_1}{\partial \mathbf{v}_\perp} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_0 \cdot \dfrac{\partial \delta f_2}{\partial \mathbf{v}_\perp} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \delta\mathbf{B}_\parallel \cdot \dfrac{\partial F_0}{\partial \mathbf{v}_\perp} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \delta\mathbf{B}_\parallel \cdot \dfrac{\partial \delta f_1}{\partial \mathbf{v}_\perp} \\
\omega F_0 & \epsilon\omega F_0 & \omega F_0 & \epsilon\omega F_0
\end{array}
\tag{A.34}
$$

$$
\begin{array}{llll}
+\dfrac{q}{mc}\mathbf{v}_\perp \times \delta\mathbf{B}_\parallel \cdot \dfrac{\partial \delta f_2}{\partial \mathbf{v}_\perp} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_\perp \cdot \dfrac{\partial F_0}{\partial \mathbf{v}_\parallel} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_\perp \cdot \dfrac{\partial \delta f_1}{\partial \mathbf{v}_\parallel} & +\dfrac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_\perp \cdot \dfrac{\partial \delta f_2}{\partial \mathbf{v}_\parallel} \\
\epsilon^2 \omega F_0 & \omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 \\
+\dfrac{q}{mc}\mathbf{v}_\parallel \times \mathbf{B}_\perp \cdot \dfrac{\partial F_0}{\partial \mathbf{v}_\perp} & +\dfrac{q}{mc}\mathbf{v}_\parallel \times \mathbf{B}_\perp \cdot \dfrac{\partial \delta f_1}{\partial \mathbf{v}_\perp} & +\dfrac{q}{mc}\mathbf{v}_\parallel \times \mathbf{B}_\perp \cdot \dfrac{\partial \delta f_2}{\partial \mathbf{v}_\perp} & = C\left(F_0, F_0\right) \\
\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 & \omega F_0 \\
+C\left(\delta f_1, F_0\right) & +C\left(\delta f_2, F_0\right) & +C\left(F_0, \delta f_1\right) & +C\left(F_0, \delta f_2\right) \\
\epsilon\omega F_0 & \epsilon^2\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 \\
+C\left(\delta f_1, \delta f_1\right) & +C\left(\delta f_1, \delta f_2\right) & +C\left(\delta f_2, \delta f_1\right) & +C\left(\delta f_2, \delta f_2\right) \\
\epsilon^2\omega F_0 & \epsilon^3\omega F_0 & \epsilon^3\omega F_0 & \epsilon^4\omega F_0
\end{array}
\tag{A.35}
$$

A.3.1 Lowest Order: Constraints on F_0

To lowest order, $\mathcal{O}\left(1/\epsilon\right)$, this gives

$$
\frac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_0 \cdot \frac{\partial F_0}{\partial \mathbf{v}_\perp} = 0 \tag{A.36}
$$

which simplifies down to

$$
\frac{\partial F_0}{\partial \theta} = 0 \tag{A.37}
$$

This means that F_0 can not depend on gyro-angle.

A.3.2 The next lowest order: F_0, a particular solution for δf_1, and constraints on the particular solution of δf_1

The next order, $\mathcal{O}(1)$, equation gives,

$$\begin{aligned}&\mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 - \frac{q}{m}\nabla_\perp \phi \cdot \frac{\partial F_0}{\partial \mathbf{v}_\perp} + \frac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_0 \cdot \frac{\partial \delta f_1}{\partial \mathbf{v}_\perp} + \frac{q}{mc}\mathbf{v}_\perp \times \delta\mathbf{B}_\parallel \cdot \frac{\partial F_0}{\partial \mathbf{v}_\perp} \\ &+\frac{q}{mc}\mathbf{v}_\perp \times \mathbf{B}_\perp \cdot \frac{\partial F_0}{\partial \mathbf{v}_\parallel} + \frac{q}{mc}\mathbf{v}_\parallel \times \mathbf{B}_\perp \cdot \frac{\partial F_0}{\partial \mathbf{v}_\perp} = C\left(F_0, F_0\right)\end{aligned} \tag{A.38}$$

The fourth term doesn't survive due to the constraint on F_0 from the previous order. At this point, we can multiply the above equation by $1 + \ln F_0$ and integrate over all of phase space. We argue that the perturbed quantities average to zero in physical space. This assumption eliminates all terms except for the collisional term.

$$\int d^3\mathbf{x} \int d^3v \left(\ln F_0 C\left(F_0, F_0\right)\right) = 0 \tag{A.39}$$

Boltzmann's H-theorem tells us that F_0 must be Maxwellian in order for F_0 to satisfy the above equation while keeping entropy constant.

$$F_0 = \frac{n_0}{\left(2\pi\right)^{3/2} v_t^3} \exp\left(-\frac{v^2}{2v_t^2}\right) \tag{A.40}$$

If we plug this solution back into Equation A.38, the fifth and sixth terms cancel out. Also, because F_0 is a Maxwellian, $C\left(F_0, F_0\right) = 0$.

$$\mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 - \Omega_i \frac{\partial \delta f_1}{\partial \theta} + \frac{q}{m}\mathbf{v}_\perp \cdot \nabla_\perp \phi \frac{1}{v_t^2} F_0 = 0 \tag{A.41}$$

Assuming the temperature is a constant, we can write this as

$$\mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 - \Omega_i \frac{\partial \delta f_1}{\partial \theta} = -\mathbf{v}_\perp \cdot \nabla_\perp \left(\frac{q\phi}{T_i} \right) F_0 \tag{A.42}$$

The particular solution can be chosen to be

$$\begin{aligned} \delta f_{1,p} &= -\frac{q}{T}\phi F_0 & \text{(A.43)} \\ &= -\frac{q}{T}\phi \frac{n_0}{\pi^{3/2} v_t^3} \exp\left(-\frac{v^2}{2v_t^2} \right) & \text{(A.44)} \end{aligned}$$

The homogenous solution must then satisfy

$$\mathbf{v}_\perp \cdot \nabla_\perp \delta f_{1,h} - \Omega_i \frac{\partial \delta f_{1,h}}{\partial \theta} = \Omega_i \left(\frac{\partial \delta f_{1,h}}{\partial \theta} \right)_{\mathbf{R}_{gc}} = 0 \tag{A.45}$$

We see that $\delta f_{1,h}$ is independent of gyrophase as well. Altogether, we have

$$\delta f_1 = \delta f_{1,h} - \frac{q}{T}\phi F_0 \tag{A.46}$$

We can express δf_1 in terms of its gyrophase dependent and independent parts. In order to do this, let's take

$$\langle \delta f_1 \rangle_{\mathbf{R}_{gc}} = \langle \delta f_{1,h} \rangle_{\mathbf{R}_{gc}} - \left\langle \frac{q}{T}\phi F_0 \right\rangle_{\mathbf{R}_{gc}} \tag{A.47}$$

We've already shown that $\delta f_{1,h}$ and F_0 are independent of θ at fixed guiding center, $\mathbf{R}_{gc}$, and q and T are constants.

$$\langle \delta f_1 \rangle_{\mathbf{R}_{gc}} = \delta f_{1,h} - \frac{q}{T} F_0 \langle \phi \rangle_{\mathbf{R}_{gc}} \tag{A.48}$$

Eliminating $\delta f_{1,h}$ allows us to write,

$$\delta f_1 = \langle \delta f_1 \rangle_{\mathbf{R}_{gc}} + \frac{q}{T} F_0 \left(\langle \phi \rangle_{\mathbf{R}_{gc}} - \phi \right) \tag{A.49}$$

A.4 Transformation to Guiding Center Coordinates and Alternate Velocity Coordinates

Before we proceed, we choose to write out a new Fokker-Planck equation for δf_1 in guiding center coordinates and the alternate velocity coordinates. We note that in the alternate velocity coordinates

$$F_0 = \frac{n_0}{(2\pi)^{3/2} v_t^3} \exp\left(-\frac{\varepsilon}{T} + \frac{q\phi}{T} \right) \tag{A.50}$$

The Fokker-Planck equation becomes

$$\begin{aligned}
&\frac{\partial F_0}{\partial t} + \frac{\partial \delta f_1}{\partial t} + \frac{\partial \delta f_2}{\partial t} + v_\parallel \nabla_\parallel F_0 - \frac{c}{B_0} \nabla_\perp \phi \times \hat{\mathbf{z}} \cdot \nabla_\perp F_0 + \frac{v_\parallel}{B_0} \nabla_\perp \times A_\parallel \hat{\mathbf{z}} \cdot \nabla_\perp F_0 - \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp F_0 \\
&+ v_\parallel \nabla_\parallel \delta f_1 - \frac{c}{B_0} \nabla_\perp \phi \times \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_1 + \frac{v_\parallel}{B_0} \nabla_\perp \times A_\parallel \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_1 - \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 \\
&+ v_\parallel \nabla_\parallel \delta f_2 - \frac{c}{B_0} \nabla_\perp \phi \times \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_2 + \frac{v_\parallel}{B_0} \nabla_\perp \times A_\parallel \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_2 - \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_2 \\
&- \frac{q}{c} v_\parallel \frac{\partial A_\parallel}{\partial t} \frac{\partial F_0}{\partial \varepsilon} - \frac{q}{c} \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \frac{\partial F_0}{\partial \varepsilon} + q \frac{\partial \phi}{\partial t} \frac{\partial F_0}{\partial \varepsilon} + \Omega_c \frac{\partial F_0}{\partial \theta} \\
&- \frac{q}{c} v_\parallel \frac{\partial A_\parallel}{\partial t} \frac{\partial \delta f_1}{\partial \varepsilon} - \frac{q}{c} \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \frac{\partial \delta f_1}{\partial \varepsilon} + q \frac{\partial \phi}{\partial t} \frac{\partial \delta f_1}{\partial \varepsilon} + \Omega_c \frac{\partial \delta f_1}{\partial \theta} \\
&- \frac{q}{c} v_\parallel \frac{\partial A_\parallel}{\partial t} \frac{\partial \delta f_2}{\partial \varepsilon} - \frac{q}{c} \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \frac{\partial \delta f_2}{\partial \varepsilon} + q \frac{\partial \phi}{\partial t} \frac{\partial \delta f_2}{\partial \varepsilon} + \Omega_c \frac{\partial \delta f_2}{\partial \theta} \\
&= \text{Collisions}
\end{aligned} \tag{A.51}$$

The F_0 derivatives can be expressed as

$$\frac{\partial F_0}{\partial t} = \frac{q}{T} F_0 \frac{\partial \phi}{\partial t} \tag{A.52}$$

$$\nabla_\parallel F_0 = \frac{q}{T} F_0 \nabla_\parallel \phi \tag{A.53}$$

$$\nabla_\perp F_0 = \frac{q}{T} F_0 \nabla_\perp \phi \tag{A.54}$$

$$\frac{\partial F_0}{\partial \varepsilon} = -\frac{1}{T} F_0 \tag{A.55}$$

$$\frac{\partial F_0}{\partial \theta} = 0 \tag{A.56}$$

Plugging these back in to the full equation leaves

$$\begin{aligned}
&\frac{q}{T} F_0 \frac{\partial \phi}{\partial t} + \frac{\partial \delta f_1}{\partial t} + \frac{\partial \delta f_2}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \phi - \frac{cq}{TB_0} F_0 \nabla_\perp \phi \times \hat{\mathbf{z}} \cdot \nabla_\perp \phi + \frac{q v_\parallel}{TB_0} F_0 \nabla_\perp \times A_\parallel \hat{\mathbf{z}} \cdot \nabla_\perp \phi \\
&- \frac{q \delta B_\parallel}{TB_0} F_0 \mathbf{v}_\perp \cdot \nabla_\perp \phi + v_\parallel \nabla_\parallel \delta f_1 - \frac{c}{B_0} \nabla_\perp \phi \times \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_1 + \frac{v_\parallel}{B_0} \nabla_\perp \times A_\parallel \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_1 - \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 \\
&+ v_\parallel \nabla_\parallel \delta f_2 - \frac{c}{B_0} \nabla_\perp \phi \times \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_2 + \frac{v_\parallel}{B_0} \nabla_\perp \times A_\parallel \hat{\mathbf{z}} \cdot \nabla_\perp \delta f_2 - \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_2 \\
&+ \frac{q}{cT} F_0 v_\parallel \frac{\partial A_\parallel}{\partial t} + \frac{q}{cT} F_0 \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} - \frac{q}{T} F_0 \frac{\partial \phi}{\partial t} + \Omega_c \frac{\partial F_0}{\partial \theta} - \frac{q}{c} v_\parallel \frac{\partial A_\parallel}{\partial t} \frac{\partial \delta f_1}{\partial \varepsilon} - \frac{q}{c} \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \frac{\partial \delta f_1}{\partial \varepsilon} \\
&+ q \frac{\partial \phi}{\partial t} \frac{\partial \delta f_1}{\partial \varepsilon} + \Omega_c \frac{\partial \delta f_1}{\partial \theta} - \frac{q}{c} v_\parallel \frac{\partial A_\parallel}{\partial t} \frac{\partial \delta f_2}{\partial \varepsilon} - \frac{q}{c} \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \frac{\partial \delta f_2}{\partial \varepsilon} + q \frac{\partial \phi}{\partial t} \frac{\partial \delta f_2}{\partial \varepsilon} + \Omega_c \frac{\partial \delta f_2}{\partial \theta} \\
&= \text{Collisions}
\end{aligned} \tag{A.57}$$

We can rewrite terms of the form $\nabla_\perp g \times \hat{\mathbf{z}} \cdot \nabla_\perp h$ as $[h, g]$ and note that $[g, g] = 0$.

$$
\begin{aligned}
&\frac{q}{T}F_0\frac{\partial\phi}{\partial t}+\frac{\partial\delta f_1}{\partial t}+\frac{\partial\delta f_2}{\partial t}+\frac{q}{T}F_0v_\parallel\nabla_\parallel\phi+\frac{qv_\parallel}{TB_0}F_0\left[\phi,A_\parallel\right]-\frac{q\delta B_\parallel}{TB_0}F_0\mathbf{v}_\perp\cdot\nabla_\perp\phi+v_\parallel\nabla_\parallel\delta f_1\\
&-\frac{c}{B_0}\left[\delta f_1,\phi\right]+\frac{v_\parallel}{B_0}\left[\delta f_1,A_\parallel\right]-\frac{\delta B_\parallel}{B_0}\mathbf{v}_\perp\cdot\nabla_\perp\delta f_1+v_\parallel\nabla_\parallel\delta f_2-\frac{c}{B_0}\left[\delta f_2,\phi\right]+\frac{v_\parallel}{B_0}\left[\delta f_2,A_\parallel\right]\\
&-\frac{\delta B_\parallel}{B_0}\mathbf{v}_\perp\cdot\nabla_\perp\delta f_2+\frac{q}{cT}F_0v_\parallel\frac{\partial A_\parallel}{\partial t}+\frac{q}{cT}F_0\mathbf{v}_\perp\cdot\frac{\partial\mathbf{A}_\perp}{\partial t}-\frac{q}{T}F_0\frac{\partial\phi}{\partial t}-\frac{q}{c}v_\parallel\frac{\partial A_\parallel}{\partial t}\frac{\partial\delta f_1}{\partial\varepsilon}-\frac{q}{c}\mathbf{v}_\perp\cdot\frac{\partial\mathbf{A}_\perp}{\partial t}\frac{\partial\delta f_1}{\partial\varepsilon}\\
&+q\frac{\partial\phi}{\partial t}\frac{\partial\delta f_1}{\partial\varepsilon}+\Omega_c\frac{\partial\delta f_1}{\partial\theta}-\frac{q}{c}v_\parallel\frac{\partial A_\parallel}{\partial t}\frac{\partial\delta f_2}{\partial\varepsilon}-\frac{q}{c}\mathbf{v}_\perp\cdot\frac{\partial\mathbf{A}_\perp}{\partial t}\frac{\partial\delta f_2}{\partial\varepsilon}+q\frac{\partial\phi}{\partial t}\frac{\partial\delta f_2}{\partial\varepsilon}+\Omega_c\frac{\partial\delta f_2}{\partial\theta}\\
&=C\left(\delta f_1,F_0\right)+C\left(F_0,\delta f_1\right)
\end{aligned}
\tag{A.58}
$$

Writing this with the ordering parameter, we have

$$
\begin{array}{lllll}
\dfrac{q}{T}F_0\dfrac{\partial\phi}{\partial t} & +\dfrac{\partial\delta f_1}{\partial t} & +\dfrac{\partial\delta f_2}{\partial t} & +\dfrac{q}{T}F_0v_\parallel\nabla_\parallel\phi & +\dfrac{qv_\parallel}{TB_0}F_0\left[\phi,A_\parallel\right]\\
\epsilon\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 & \epsilon\omega F_0 & \epsilon\omega F_0\\
-\dfrac{q\delta B_\parallel}{TB_0}F_0\mathbf{v}_\perp\cdot\nabla_\perp\phi & +v_\parallel\nabla_\parallel\delta f_1 & -\dfrac{c}{B_0}\left[\delta f_1,\phi\right] & +\dfrac{v_\parallel}{B_0}\left[\delta f_1,A_\parallel\right] & -\dfrac{\delta B_\parallel}{B_0}\mathbf{v}_\perp\cdot\nabla_\perp\delta f_1\\
\epsilon\omega F_0 & \epsilon\omega F_0 & \epsilon\omega F_0 & \epsilon\omega F_0 & \epsilon\omega F_0\\
+v_\parallel\nabla_\parallel\delta f_2 & -\dfrac{c}{B_0}\left[\delta f_2,\phi\right] & +\dfrac{v_\parallel}{B_0}\left[\delta f_2,A_\parallel\right] & -\dfrac{\delta B_\parallel}{B_0}\mathbf{v}_\perp\cdot\nabla_\perp\delta f_2 & +\dfrac{q}{cT}F_0v_\parallel\dfrac{\partial A_\parallel}{\partial t}\\
\epsilon^2\omega F_0 & \epsilon^2\omega F_0 & \epsilon^2\omega F_0 & \epsilon^2\omega F_0 & \epsilon\omega F_0\\
+\dfrac{q}{cT}F_0\mathbf{v}_\perp\cdot\dfrac{\partial\mathbf{A}_\perp}{\partial t} & -\dfrac{q}{T}F_0\dfrac{\partial\phi}{\partial t} & -\dfrac{q}{c}v_\parallel\dfrac{\partial A_\parallel}{\partial t}\dfrac{\partial\delta f_1}{\partial\varepsilon} & -\dfrac{q}{c}\mathbf{v}_\perp\cdot\dfrac{\partial\mathbf{A}_\perp}{\partial t}\dfrac{\partial\delta f_1}{\partial\varepsilon} & +q\dfrac{\partial\phi}{\partial t}\dfrac{\partial\delta f_1}{\partial\varepsilon}\\
\epsilon\omega F_0 & \epsilon\omega F_0 & \epsilon^2\omega F_0 & \epsilon^2\omega F_0 & \epsilon^2\omega F_0\\
+\Omega_c\dfrac{\partial\delta f_1}{\partial\theta} & -\dfrac{q}{c}v_\parallel\dfrac{\partial A_\parallel}{\partial t}\dfrac{\partial\delta f_2}{\partial\varepsilon} & -\dfrac{q}{c}\mathbf{v}_\perp\cdot\dfrac{\partial\mathbf{A}_\perp}{\partial t}\dfrac{\partial\delta f_2}{\partial\varepsilon} & +q\dfrac{\partial\phi}{\partial t}\dfrac{\partial\delta f_2}{\partial\varepsilon} & +\Omega_c\dfrac{\partial\delta f_2}{\partial\theta}\\
\omega F_0 & \epsilon^3\omega F_0 & \epsilon^3\omega F_0 & \epsilon^3\omega F_0 & \epsilon\omega F_0\\
=C\left(\delta f_1,F_0\right) & +C\left(F_0,\delta f_1\right) & & & \\
\epsilon\omega F_0 & \epsilon\omega F_0 & & &
\end{array}
\tag{A.59}
$$

The $\mathcal{O}\left(\epsilon\right)$ equation is

$$\frac{\partial \delta f_1}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \phi + \frac{q v_\parallel}{T B_0} F_0 \left[\phi, A_\parallel\right] - \frac{q \delta B_\parallel}{T B_0} F_0 \mathbf{v}_\perp \cdot \nabla_\perp \phi + v_\parallel \nabla_\parallel \delta f_1 - \frac{c}{B_0} \left[\delta f_1, \phi\right] + \frac{v_\parallel}{B_0} \left[\delta f_1, A_\parallel\right]$$
$$- \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 + \frac{q}{cT} F_0 v_\parallel \frac{\partial A_\parallel}{\partial t} + \frac{q}{cT} F_0 \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} + \Omega_c \frac{\partial \delta f_2}{\partial \theta} = C\left(\delta f_1, F_0\right) + C\left(F_0, \delta f_1\right) \tag{A.60}$$

To get rid of the δf_2 term, we gyro-average.

$$\frac{\partial \left\langle \delta f_1 \right\rangle}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \left\langle \phi \right\rangle + \frac{q v_\parallel}{T B_0} F_0 \left\langle \left[\phi, A_\parallel\right] \right\rangle - \left\langle \frac{q \delta B_\parallel}{T B_0} F_0 \mathbf{v}_\perp \cdot \nabla_\perp \phi \right\rangle + v_\parallel \nabla_\parallel \left\langle \delta f_1 \right\rangle - \frac{c}{B_0} \left\langle \left[\delta f_1, \phi\right] \right\rangle$$
$$+ \frac{v_\parallel}{B_0} \left\langle \left[\delta f_1, A_\parallel\right] \right\rangle - \left\langle \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_1 \right\rangle + \frac{q}{cT} F_0 v_\parallel \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t} + \frac{q}{cT} F_0 \left\langle \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle$$
$$+ \left\langle \Omega_c \frac{\partial \delta f_2}{\partial \theta} \right\rangle = \left\langle C\left(\delta f_1, F_0\right) \right\rangle + \left\langle C\left(F_0, \delta f_1\right) \right\rangle \tag{A.61}$$

Using Equation A.49, we can write this in several different ways. If we want an expression for the full δf_1, we can write

$$\frac{\partial \delta f_1}{\partial t} - \frac{q}{T} F_0 \frac{\partial \left\langle \phi \right\rangle}{\partial t} + \frac{q}{T} F_0 \frac{\partial \phi}{\partial t} + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \left\langle \phi \right\rangle + \frac{q v_\parallel}{T B_0} F_0 \left\langle \left[\phi, A_\parallel\right] \right\rangle - \left\langle \frac{q \delta B_\parallel}{T B_0} F_0 \mathbf{v}_\perp \cdot \nabla_\perp \phi \right\rangle$$
$$+ v_\parallel \nabla_\parallel \delta f_1 - \frac{q}{T} F_0 v_\parallel \nabla_\parallel \left\langle \phi \right\rangle + \frac{q}{T} F_0 v_\parallel \nabla_\parallel \phi - \frac{c}{B_0} \left[\delta f_{1,h}, \left\langle \phi \right\rangle\right] + \frac{v_\parallel}{B_0} \left[\delta f_{1,h}, \left\langle A_\parallel \right\rangle\right]$$
$$- \frac{q v_\parallel}{T B_0} F_0 \left\langle \left[\phi, A_\parallel\right] \right\rangle - \left\langle \frac{\delta B_\parallel}{B_0} \mathbf{v}_\perp \cdot \nabla_\perp \delta f_{1,h} \right\rangle + \left\langle \frac{q \delta B_\parallel}{T B_0} F_0 \mathbf{v}_\perp \cdot \nabla_\perp \phi \right\rangle + \frac{q}{cT} F_0 v_\parallel \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t}$$
$$+ \frac{q}{cT} F_0 \left\langle \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle = \left\langle C\left(\delta f_1, F_0\right) \right\rangle + \left\langle C\left(F_0, \delta f_1\right) \right\rangle \tag{A.62}$$

The fourth and fifth terms of the first line cancel with the second term of the second line and the first term of the third line, respectively. The last term of the first line cancels with the third term of the third line.

$$\frac{\partial \delta f_1}{\partial t} - \frac{q}{T}F_0\frac{\partial \langle\phi\rangle}{\partial t} + \frac{q}{T}F_0\frac{\partial \phi}{\partial t} + v_\parallel \nabla_\parallel \delta f_1 + \frac{q}{T}F_0 v_\parallel \nabla_\parallel \phi - \frac{c}{B_0}\left[\delta f_{1,h}, \langle\phi\rangle\right] + \frac{v_\parallel}{B_0}\left[\delta f_{1,h}, \left\langle A_\parallel \right\rangle\right]$$
$$- \left\langle \frac{\delta B_\parallel}{B_0}\mathbf{v}_\perp \cdot \nabla_\perp \delta f_{1,h} \right\rangle + \frac{q}{cT}F_0 v_\parallel \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t} + \frac{q}{cT}F_0 \left\langle \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle = \langle C\left(\delta f_1, F_0\right)\rangle + \langle C\left(F_0, \delta f_1\right)\rangle \tag{A.63}$$

If we replace $\delta f_{1,h}$ in all but the last term of the right hand side, we have

$$\frac{\partial \delta f_1}{\partial t} - \frac{q}{T}F_0\frac{\partial \langle\phi\rangle}{\partial t} + \frac{q}{T}F_0\frac{\partial \phi}{\partial t} + v_\parallel \nabla_\parallel \delta f_1 + \frac{q}{T}F_0 v_\parallel \nabla_\parallel \phi - \frac{c}{B_0}\left[\delta f_1, \langle\phi\rangle\right] - \frac{cq}{TB_0}F_0\left[\phi, \langle\phi\rangle\right]$$
$$+ \frac{v_\parallel}{B_0}\left[\delta f_1, \left\langle A_\parallel \right\rangle\right] + \frac{qv_\parallel}{TB_0}F_0\left[\phi, \left\langle A_\parallel \right\rangle\right] + \frac{q}{cT}F_0 v_\parallel \frac{\partial \left\langle A_\parallel \right\rangle}{\partial t} + \frac{q}{cT}F_0 \left\langle \mathbf{v}_\perp \cdot \frac{\partial \mathbf{A}_\perp}{\partial t} \right\rangle$$
$$- \left\langle \frac{\delta B_\parallel}{B_0}\mathbf{v}_\perp \cdot \nabla_\perp \delta f_{1,h} \right\rangle = \langle C\left(\delta f_1, F_0\right)\rangle + \langle C\left(F_0, \delta f_1\right)\rangle \tag{A.64}$$

A.5 Gyrokinetic Maxwell's Equations

In Sec. A.4, we transformed to guiding center coordinates. However, Maxwell's equations need to be solved in particle coordinates. We can use Eqn. A.31 to transform back. We are only concerned with matching the linear response of our model to the linear part of Maxwell's equations, so we drop the nonlinear terms for this derivation.

We assume that our fields can be written in the form $\phi = \phi e^{i(\mathbf{k}\cdot\mathbf{R} - \omega t)} = \phi e^{i\mathbf{k}\cdot\left(\mathbf{r} + \frac{\mathbf{v}\times\hat{\mathbf{z}}}{\Omega_s}\right) - i\omega t}$. If this is the case, the gyroaverage back to particle coordinates will introduce Bessel functions. Applying this to Eqn. A.64 leaves

$$\delta f_1 = \frac{q}{T}F_0\left(-\left(1 + \frac{\omega}{k_\parallel v_\parallel - \omega}J_0^2\right)\phi + \frac{\omega}{ck_\parallel}J_0^2\left(1 + \frac{\omega}{k_\parallel v_\parallel - \omega}\right)A_\parallel + \frac{mv_\perp^2}{q}\frac{\omega}{k_\parallel v_\parallel - \omega}\frac{J_1}{k_\perp v_\perp/\Omega}J_0\frac{\delta B_\parallel}{B_0}\right) \tag{A.65}$$

This form can be used in the definitions of Maxwell's equations below. The following integrals will also be helpful as we proceed:

$$\int_0^\infty e^{-v_\perp^2/2v_t^2} J_0^2\left(\frac{k_\perp v_\perp}{\Omega}\right) v_\perp \, dv_\perp = v_t^2 e^{-k_\perp^2\rho^2} I_0\left(k_\perp^2\rho^2\right) = v_t^2 \Gamma_0\left(k_\perp^2\rho^2\right) \tag{A.66}$$

$$\int_0^\infty v_\perp^2 e^{-v_\perp^2/2v_t^2} J_0\left(\frac{k_\perp v_\perp}{\Omega}\right) \frac{J_1\left(\frac{k_\perp v_\perp}{\Omega}\right)}{k_\perp v_\perp/\Omega} v_\perp \, dv_\perp = v_t^4 e^{-k_\perp^2\rho^2} \left(I_0\left(k_\perp^2\rho^2\right) - I_1\left(k_\perp^2\rho^2\right)\right) = v_t^4 \Gamma_1\left(k_\perp^2\rho^2\right) \tag{A.67}$$

where I_0 is the zeroth order modified Bessel function, and I_1 is the first order modified Bessel function. In the parallel direction, we will want to use the plasma dispersion function:

$$Z\left(\xi\right) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{e^{-t^2}}{t-\xi} \, dt \tag{A.68}$$

where $\xi = \frac{\omega}{\sqrt{2}k_\parallel v_t}$and its derivative w.r.t. its argument:

$$Z'\left(\xi\right) = -2\left(1+\xi Z\right) = -2\frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \frac{te^{-t^2}}{t-\xi} \, dt \tag{A.69}$$

A.5.1 Poisson's Equation

Poisson's equation is normally written as

$$\nabla \cdot \mathbf{E} = 4\pi \sum_s q_s n_s \tag{A.70}$$

In terms of potentials, this becomes

$$\begin{array}{cccccc} -\nabla_\perp^2\phi & -\nabla_\parallel^2\phi & -\frac{1}{c}\frac{\partial\nabla_\parallel A_\parallel}{\partial t} & -\frac{1}{c}\frac{\partial\nabla_\perp\cdot\mathbf{A}_\perp}{\partial t} & = 4\pi\sum_s q_s n_s \\ \omega & \omega & \epsilon^2\omega & \epsilon\omega & \frac{1}{\epsilon}\omega\frac{c^2}{v_A^2} \end{array} \tag{A.71}$$

The term on the right hand side of the equation is bigger than the other terms in the equation by a factor of $1/\epsilon$ as well as c^2/v_A^2. This allows us to assume quasineutrality and write Poisson's Equation in gyrokinetics as

$$\sum_s \int_{-\infty}^{\infty} q_s f_s \, d^3v = 0 \tag{A.72}$$

where f_s is the full pdf. Substituting in through first order gives,

$$\sum_s \int_{-\infty}^{\infty} d^3v \left(q_s F_{0,s} - \frac{q_s^2}{T_{0,s}} F_{0,s}\phi + \frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\frac{q_s^2}{T_{0,s}} F_{0,s} J_0^2\phi + \frac{\omega}{k_\parallel}\left(1 - \frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\right)\frac{q_s^2}{cT_{0,s}} F_{0,s} J_0^2 A_\parallel \right.$$
$$\left. - \frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\frac{eq_s}{cT_{0,s}} F_{0,s}\frac{v_{\perp,s}}{k_\perp} J_1 J_0 \delta B_\parallel \right) = 0 \tag{A.73}$$

Using the definitions above, we integrate over velocity.

$$\sum_s \left(q_s n_{0,s} - \frac{q_s^2}{T_{0,s}} n_{0,s}\left(1 + \Gamma_{0,s}\xi_s Z_s\right)\phi + \frac{q_s^2}{T_{0,s}}\frac{n_{0,s}}{c}\frac{\omega}{k_\parallel}\Gamma_{0,s}\left(1 + \xi_s Z_s\right) A_\parallel \right.$$
$$\left. + \frac{n_{0,s} q_s^2}{q_s B_0}\Gamma_{1,s}\xi_s Z_s \delta B_\parallel \right) = 0 \tag{A.74}$$

We assume that Poisson's equation holds at each order. Zeroth order tells us

$$n_{0,e} = n_{0,i} = n_0 \tag{A.75}$$

The next order gives

$$
\begin{aligned}
&-e^2 n_0 \left(\frac{1}{T_{0,e}} \left(1 + \Gamma_{0,e} \xi_e Z_e\right) + \frac{1}{T_{0,i}} \left(1 + \Gamma_{0,i} \xi_i Z_i\right) \right) \phi \\
&+ \frac{e^2 n_0}{c} \frac{\omega}{k_\parallel} \left(\frac{1}{T_{0,e}} \Gamma_{0,e} \left(1 + \xi_e Z_e\right) + \frac{1}{T_{0,i}} \Gamma_{0,i} \left(1 + \xi_i Z_i\right) \right) A_\parallel \\
&+ \frac{n_0 e}{B_0} \left(-\Gamma_{1,e} \xi_e Z_e + \Gamma_{1,i} \xi_i Z_i\right) \delta B_\parallel = 0
\end{aligned}
\tag{A.76}
$$

A.5.2 Parallel Ampere's Law

The parallel part of Ampere's Law can be written

$$
\left(\nabla \times \mathbf{B}\right)_\parallel = \frac{\partial \mathbf{E}_\parallel}{\partial t} + \frac{4\pi}{c} \mathbf{J}_\parallel \tag{A.77}
$$

In terms of potentials, this can be written as

$$
\begin{array}{llll}
-\nabla_\perp^2 A_\parallel & = -\frac{1}{c} \frac{\partial \nabla_\parallel \phi}{\partial t} & -\frac{1}{c^2} \frac{\partial^2 A_\parallel}{\partial t^2} + & \frac{4\pi}{c} J_\parallel \\
\epsilon \frac{v_A^2}{v_{ti}^2} & \frac{v_A^2}{c^2} & \frac{1}{\epsilon} \frac{v_A^2}{c^2} & \epsilon
\end{array}
\tag{A.78}
$$

The two time derivative terms are smaller than the others by a factor of v_A^2/c^2 which is much less than 1 in the non-relativistic limit as is used in gyrokinetics. Ampere's Law in the parallel direction in gyrokinetics can be written as

$$
-\nabla_\perp^2 A_\parallel = \int_{-\infty}^{\infty} \left(\frac{4\pi}{c} \sum_s q_s v_{\parallel,s} f_s \right) d^3 v \tag{A.79}
$$

Substituting in through first order gives

$$
\begin{aligned}
-\nabla_\perp^2 A_\parallel = \frac{4\pi}{c} \sum_s \int_{-\infty}^{\infty} d^3 v q_s v_{\parallel,s} \Bigg(& F_{0,s} - \frac{q_s}{T_{0,s}} F_{0,s} \phi + \frac{\omega}{\omega - v_{\parallel,s} k_\parallel} \frac{q_s}{T_{0,s}} F_{0,s} J_0^2 \phi \\
& + \frac{\omega}{k_\parallel} \left(1 - \frac{\omega}{\omega - v_{\parallel,s} k_\parallel}\right) \frac{q_s}{c T_{0,s}} F_{0,s} J_0^2 A_\parallel - \frac{\omega}{\omega - v_{\parallel,s} k_\parallel} \frac{e}{c T_{0,s}} F_{0,s} \frac{v_{\perp,s}}{k_\perp} J_1 J_0 \delta B_\parallel \Bigg)
\end{aligned}
\tag{A.80}
$$

and integrating leaves

$$-\nabla_{\perp}^{2}A_{\parallel}=\frac{4\pi}{c}\sum_{s}\left(-\frac{n_{0,s}q_{s}^{2}}{T_{0,s}}\frac{\omega}{k_{\parallel}}\Gamma_{0,s}\left(1+\xi_{s}Z_{s}\right)\phi+\frac{n_{0,s}q_{s}^{2}}{cT_{0,s}}\Gamma_{0,s}\frac{\omega^{2}}{k_{\parallel}^{2}}\left(1+\xi_{s}Z_{s}\right)A_{\parallel}\right.$$
$$\left.+\frac{n_{0,s}q_{s}}{B_{0}}\frac{\omega}{k_{\parallel}}\Gamma_{1,s}\left(1+\xi_{s}Z_{s}\right)\delta B_{\parallel}\right) \tag{A.81}$$

We note that the integrals over the zeroth order p.d.f are odd w.r.t. parallel velocity, and thus evaluate to zero. We are left only with the first order relation.

$$\begin{aligned}&-\frac{4\pi n_{0}e^{2}}{c}\left(\frac{1}{T_{0,e}}\frac{\omega}{k_{\parallel}}\Gamma_{0,e}\left(1+\xi_{e}Z_{e}\right)+\frac{1}{T_{0,i}}\frac{\omega}{k_{\parallel}}\Gamma_{0,i}\left(1+\xi_{i}Z_{i}\right)\right)\phi\\&\frac{4\pi n_{0}e^{2}}{c^{2}}\left(-\frac{c^{2}k_{\perp}^{2}}{4\pi n_{0}e^{2}}+\frac{1}{T_{0,e}}\Gamma_{0,e}\frac{\omega^{2}}{k_{\parallel}^{2}}\left(1+\xi_{e}Z_{e}\right)+\frac{1}{T_{0,i}}\Gamma_{0,i}\frac{\omega^{2}}{k_{\parallel}^{2}}\left(1+\xi_{i}Z_{i}\right)\right)A_{\parallel}\\&+\frac{4\pi n_{0}e}{cB_{0}}\frac{\omega}{k_{\parallel}}\left(-\Gamma_{1,e}\left(1+\xi_{e}Z_{e}\right)+\Gamma_{1,i}\left(1+\xi_{i}Z_{i}\right)\right)\delta B_{\parallel}=0\end{aligned} \tag{A.82}$$

A.5.3 Perpendicular Ampere's Law

In order to isolate the perpendicular part of Ampere's Law, we cross it with the parallel unit vector.

$$\hat{\mathbf{z}}\times\nabla\times\mathbf{B}=\hat{\mathbf{z}}\times\frac{\partial\mathbf{E}_{\perp}}{\partial t}+\frac{4\pi}{c}\hat{\mathbf{z}}\times\mathbf{J}_{\perp} \tag{A.83}$$

It is often easier to work with the divergence of the above equation. This can be written as

$$\begin{array}{llll}-\nabla_{\perp}^{2}\delta B_{\parallel} & =-\frac{1}{c^{2}}\frac{\partial^{2}\delta B_{\parallel}}{\partial t^{2}}+ & \frac{4\pi}{c}\nabla_{\perp}\cdot\hat{\mathbf{z}}\times\mathbf{J}_{\perp}\\ \epsilon\frac{v_{A}^{2}}{v_{ti}^{2}} & \frac{1}{\epsilon}\frac{v_{A}^{2}}{c^{2}} & \epsilon\end{array} \tag{A.84}$$

Once again, the time derivative is smaller than the other terms by a factor of v_{A}^{2}/c^{2}. We can write our field equation for $\delta B_{\parallel}$ as

$$\nabla_\perp^2 \delta B_\parallel = \nabla_\perp \cdot \int_{-\infty}^{\infty} \frac{4\pi}{c} \sum_s q_s \left\langle \hat{\mathbf{z}} \times \mathbf{v}_{\perp,s} f_s \right\rangle d^3v \tag{A.85}$$

where f_s is still in guiding center coordinates, before we have used the gyroaverage defined in Eqn. A.31. This can be written as

$$f_s = F_{0,s} - \frac{q_s}{T_{0,s}} F_{0,s}\phi + \frac{\omega}{\omega - v_{\parallel,s}k_\parallel} \frac{q_s}{T_{0,s}} F_{0,s} J_0 \phi + \frac{\omega}{k_\parallel} \left(1 - \frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\right) \frac{q_s}{cT_{0,s}} F_{0,s} J_0 A_\parallel$$
$$- \frac{\omega}{\omega - v_{\parallel,s}k_\parallel} \frac{e}{cT_{0,s}} F_{0,s} \frac{v_{\perp,s}}{k_\perp} J_1 \delta B_\parallel \tag{A.86}$$

Plugging it in gives us

$$-k_\perp^2 \delta B_\parallel = \frac{4\pi}{c} \int_{-\infty}^{\infty} d^3v \sum_s q_s i\mathbf{k}_\perp \cdot \hat{\mathbf{z}} \times \mathbf{v}_{\perp,s} \left(F_{0,s} - \frac{q_s}{T_{0,s}} F_{0,s}\phi + \frac{\omega}{\omega - v_{\parallel,s}k_\parallel} \frac{q_s}{T_{0,s}} F_{0,s} J_0 \phi \right.$$
$$\left. + \frac{\omega}{k_\parallel} \left(1 - \frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\right) \frac{q_s}{cT_{0,s}} F_{0,s} J_0 A_\parallel - \frac{\omega}{\omega - v_{\parallel,s}k_\parallel} \frac{e}{cT_{0,s}} F_{0,s} \frac{v_{\perp,s}}{k_\perp} J_1 \delta B_\parallel \right) \tag{A.87}$$

The first two terms are already in particle positions, so they end up evaluating to zero.

The gyro-average for a field requires the following integral:

$$\frac{1}{2\pi} \int_0^{2\pi} i\mathbf{k}_\perp \cdot \hat{\mathbf{z}} \times \mathbf{v}_{\perp,s} e^{i\mathbf{k}_\perp \cdot \frac{\mathbf{v}_{\perp,s} \times \hat{\mathbf{z}}}{\Omega_s}} d\theta = \frac{1}{2\pi} \int_0^{2\pi} ik_\perp v_{\perp,s} \sin\theta e^{\mp i \frac{k_\perp v_{\perp,s}}{\Omega_i} \sin\theta} d\theta \tag{A.88}$$

$$= \mp k_\perp v_{\perp,s} J_1 \left(\frac{k_\perp v_{\perp,s}}{\Omega_i} \right) \tag{A.89}$$

$$= \mp \frac{k_\perp^2 v_{\perp,s}^2}{\Omega_s} \frac{J_1}{k_\perp v_{\perp,s}/\Omega_s} \tag{A.90}$$

NB: The ions have a negative sign out front and the electrons have a positive sign. Plugging that back in, we can write

$$-k_\perp^2 \delta B_\parallel = -\frac{4\pi}{c}\int_{-\infty}^{\infty} d^3v \sum_s e\frac{k_\perp^2 v_{\perp,s}^2}{\Omega_s}\frac{J_1}{k_\perp v_{\perp,s}/\Omega_s}\left(\frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\frac{q_s}{T_{0,s}}F_{0,s}J_0\phi\right.$$
$$\left.+\frac{\omega}{k_\parallel}\left(1-\frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\right)\frac{q_s}{cT_{0,s}}F_{0,s}J_0A_\parallel - \frac{\omega}{\omega - v_{\parallel,s}k_\parallel}\frac{e}{cT_{0,s}}F_{0,s}\frac{v_{\perp,s}}{k_\perp}J_1\delta B_\parallel\right) \tag{A.91}$$

Evaluating the velocity integrals leaves

$$-\delta B_\parallel = -\frac{4\pi}{c}\sum_s\left(-\frac{n_0 q_s c}{B_0}\Gamma_{1,s}\xi_s Z_s\phi + \frac{n_0 q_s}{B_0}\frac{\omega}{k_\parallel}\Gamma_{1,s}\left(1+\xi_s Z_s\right)A_\parallel + \frac{2n_0 cT_{0,s}}{B_0^2}\Gamma_{1,s}\xi_s Z_s\delta B_\parallel\right) \tag{A.92}$$

which expands to

$$-\delta B_\parallel = \left(-\frac{4\pi n_0 e}{B_0}\Gamma_{1,e}\xi_e Z_e + \frac{4\pi n_0 e}{B_0}\Gamma_{1,i}\xi_i Z_i\right)\phi$$
$$-\left(-\frac{4\pi n_0 e}{cB_0}\frac{\omega}{k_\parallel}\Gamma_{1,e}\left(1+\xi_e Z_e\right) + \frac{4\pi n_0 e}{cB_0}\frac{\omega}{k_\parallel}\Gamma_{1,i}\left(1+\xi_i Z_i\right)\right)A_\parallel \tag{A.93}$$
$$-\left(\frac{8\pi n_0 T_{0,e}}{B_0^2}\Gamma_{1,e}\xi_e Z_e + \frac{8\pi n_0 T_{0,i}}{B_0^2}\Gamma_{1,i}\xi_i Z_i\right)\delta B_\parallel$$

Appendix B

Inverted Matrix Equations

In this appendix we note the results of the matrix inversion needed for the implicit scheme described in Sec. 3.3, and define the functions that are used therein.

$$a = \sqrt{\frac{\beta_i}{2}}\Delta t\hat{k}_{\parallel} \qquad o = 6dh\tau + l = 1 + 3dhq\tau$$

$$b = \left(\langle J_0\rangle^2 - 1\right) \qquad q = 2 + \beta_i cfg$$

$$c = \left(2\langle J_1\rangle_a \langle J_0\rangle - 1\right) \qquad r = 2 - q\left(l - 6bh\right) = 2\left(1 - 2bhq\nu\Delta t\hat{k}_{\perp}^{2m}\right)$$

$$d = 2\langle J_1\rangle_a \langle J_0\rangle \qquad s = \frac{1}{3\left(2b + \beta_i cfgl\tau\right) + 4b\nu\Delta t\hat{k}_{\perp}^{2m} - 4lmq\left(\Delta t\right)^2\hat{k}_{\parallel}^2\hat{k}_{\perp}^2}$$

$$e = 4\langle J_1\rangle_a \langle J_1\rangle_b \qquad v = 3l\tau - 4mo\left(\Delta t\right)^2\hat{k}_{\parallel}^2\hat{k}_{\perp}^2$$

$$f = \frac{1}{1 + \beta_i e} \qquad w = 3\beta_i cfg\tau - 4mq\left(\Delta t\right)^2\hat{k}_{\parallel}^2\hat{k}_{\perp}^2$$

$$g = \frac{1}{\tau + \beta_i f} \qquad x = 6b + \beta_i cfgv$$

$$h = \frac{1}{2b\left(3 + 2\nu\Delta t\hat{k}_{\perp}^{2m}\right) - 3\beta_i cdfg\tau} \qquad y = 3\beta_i bcfgh\tau\nu\Delta t\hat{k}_{\perp}^{2m} + mr\left(\Delta t\right)^2\hat{k}_{\parallel}^2\hat{k}_{\perp}^2$$

$$l = 1 + 3\beta_i cdfgh\tau = 2bh\left(3 + 2\nu\Delta t\hat{k}_{\perp}^{2m}\right) \qquad z = 1 + 4lmqs\left(\Delta t\right)^2\hat{k}_{\parallel}^2\hat{k}_{\perp}^2$$

$$m = \frac{1}{3 + 2\eta\Delta t\hat{k}_{\perp}^{2m}}$$

$$\hat{n}_e^{n+1} = 4bs\hat{n}_{e,rhs} + i\frac{16abms}{\beta_i}\hat{k}_\perp^2\hat{A}_{\parallel,rhs} - \frac{16bms}{3}\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{p}_{\parallel,e,rhs} - \frac{2\beta_i bfgsv}{3}\hat{p}_{\perp,e,rhs}$$
$$+4bhsw\left\langle J_0\right\rangle\hat{\bar{n}}_{i,rhs} - \frac{2\beta_i bfgsv}{3}2\left\langle J_1\right\rangle_b\hat{\bar{p}}_{\perp,i,rhs} + \frac{4sx}{3}\hat{n}_e^n + i\frac{32abms}{\beta_i}\hat{k}_\perp^2\hat{A}_\parallel^n - \frac{32bms}{3}\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{p}_{\parallel,e}^n$$
$$-\frac{16sy}{3}\left\langle J_0\right\rangle\hat{\bar{n}}_i^n - \frac{sx}{3}\hat{n}_e^{n-1} - i\frac{8abms}{\beta_i}\hat{k}_\perp^2\hat{A}_\parallel^{n-1} + \frac{8bms}{3}\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{p}_{\parallel,e}^{n-1} + \frac{4sy}{3}\left\langle J_0\right\rangle\hat{\bar{n}}_i^{n-1}$$
$$-i\frac{8abs}{\beta_i}\hat{k}_\perp^2\hat{A}_A^{n+1} \tag{B.1}$$

$$\hat{A}_\parallel^{n+1} = -i4almqs\hat{n}_{e,rhs} + 2mz\hat{A}_{\parallel,rhs} + i\frac{4amz}{3}\hat{p}_{\parallel,e,rhs} - i\frac{2\beta_i afgm}{3}\left(o - lqsv\right)\hat{p}_{\perp,e,rhs}$$
$$+i4almqs\left\langle J_0\right\rangle\hat{\bar{n}}_{i,rhs} - i\frac{2\beta_i afgm}{3}\left(o - lqsv\right)2\left\langle J_1\right\rangle_b\hat{\bar{p}}_{\perp,i,rhs}$$
$$+i\frac{4am}{3b}\left(\beta_i cfgo - lqsx\right)\hat{n}_e^n + 4mz\hat{A}_\parallel^n + i\frac{8amz}{3}\hat{p}_{\parallel,e}^n + i\frac{4am}{3b}\left(r + 4lqsy\right)\left\langle J_0\right\rangle\hat{\bar{n}}_i^n$$
$$-i\frac{am}{3b}\left(\beta_i cfgo - lqsx\right)\hat{n}_e^{n-1} - mz\hat{A}_\parallel^{n-1} - i\frac{2amz}{3}\hat{p}_{\parallel,e}^{n-1} - i\frac{am}{3b}\left(r + 4lqsy\right)\left\langle J_0\right\rangle\hat{\bar{n}}_i^{n-1}$$
$$-4lmqs\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{A}_A^{n+1} \tag{B.2}$$

$$\hat{p}_{\parallel,e}^{n+1} = -2\beta_i cfgls\hat{n}_{e,rhs} - i8acfglms\hat{k}_\perp^2\hat{A}_{\parallel,rhs} + \frac{2}{3}\left(1 + 4\beta_i cfglms\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\right)\hat{p}_{\parallel,e,rhs}$$
$$-\frac{\beta_i fgl}{3}\left(1 - \beta_i cfgsv\right)\hat{p}_{\perp,e,rhs} + 2\beta_i cfgls\left\langle J_0\right\rangle\hat{\bar{n}}_{i,rhs} - \frac{\beta_i fgl}{3}\left(1 - \beta_i cfgsv\right)2\left\langle J_1\right\rangle_b\hat{\bar{p}}_{\perp,i,rhs}$$
$$+\frac{2\beta_i cfgl}{3b}\left(1 - sx\right)\hat{n}_e^n - i16acfglms\hat{k}_\perp^2\hat{A}_\parallel^n + \frac{4}{3}\left(1 + 4\beta_i cfglms\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\right)\hat{p}_{\parallel,e}^n$$
$$+\frac{8\beta_i cfg}{3b}\left(lsy - bh\nu\Delta t\hat{k}_\perp^{2m}\right)\left\langle J_0\right\rangle\hat{\bar{n}}_i^n - \frac{\beta_i cfgl}{6b}\left(1 - sx\right)\hat{n}_e^{n-1} + i4acfglms\hat{k}_\perp^2\hat{A}_\parallel^{n-1}$$
$$-\frac{1}{3}\left(1 + 4\beta_i cfglms\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\right)\hat{p}_{\parallel,e}^{n-1} - \frac{2\beta_i cfg}{3b}\left(lsy - bh\nu\Delta t\hat{k}_\perp^{2m}\right)\left\langle J_0\right\rangle\hat{\bar{n}}_i^{n-1} + i4acfgls\hat{k}_\perp^2\hat{A}_A^{n+1} \tag{B.3}$$

$$\hat{p}_{\perp,e}^{n+1} = -4\beta_i cfgls\hat{n}_{e,rhs} - i16acfglms\hat{k}_\perp^2\hat{A}_{\parallel,rhs} + \frac{16\beta_i cfglms}{3}\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{p}_{\parallel,e,rhs}$$
$$+\frac{2}{3}\left(1-\beta_i fgl\left(1-\beta_i cfgsv\right)\right)\hat{p}_{\perp,e,rhs} + 4\beta_i cfgls\left\langle J_0\right\rangle\hat{\bar{n}}_{i,rhs}$$
$$-\frac{2\beta_i fgl}{3}\left(1-\beta_i cfgsv\right)2\left\langle J_1\right\rangle_b\hat{\bar{p}}_{\perp,i,rhs} + \frac{4\beta_i cfgl}{3b}\left(1-sx\right)\hat{n}_e^n - i32acfglms\hat{k}_\perp^2\hat{A}_\parallel^n$$
$$+\frac{32\beta_i cfglms}{3}\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{p}_{\parallel,e}^n + \frac{4}{3}\hat{p}_{\perp,e}^n + \frac{16\beta_i cfg}{3b}\left(lsy - bh\nu\Delta t\hat{k}_\perp^{2m}\right)\left\langle J_0\right\rangle\hat{\bar{n}}_i^n - \frac{\beta_i cfgl}{3b}\left(1-sx\right)\hat{n}_e^{n-1}$$
$$+i8acfglms\hat{k}_\perp^2\hat{A}_\parallel^{n-1} - \frac{8\beta_i cfglms}{3}\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2\hat{p}_{\parallel,e}^{n-1} - \frac{1}{3}\hat{p}_{\perp,e}^{n-1} - \frac{4\beta_i cfg}{3b}\left(lsy - bh\nu\Delta t\hat{k}_\perp^{2m}\right)\left\langle J_0\right\rangle\hat{\bar{n}}_i^{n-1}$$
$$+i8acfglsk_\perp^2\hat{A}_A^{n+1} \tag{B.4}$$

$$\hat{\bar{n}}_i^{n+1} = -12\beta_i bcfghs\tau 2\left\langle J_1\right\rangle_a\hat{n}_{e,rhs} - i48abcfghms\tau\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{A}_{\parallel,rhs}$$
$$+16\beta_i bcfghms\tau\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{p}_{\parallel,e,rhs} - 2\beta_i bfgh\tau\left(1-\beta_i cfgsv\right)2\left\langle J_1\right\rangle_a\hat{p}_{\perp,e,rhs}$$
$$+4bh\left(1-3\beta_i cdfghsw\tau\right)\hat{\bar{n}}_{i,rhs} - 2\beta_i befgh\tau\left(1-\beta_i cfgsv\right)\hat{\bar{p}}_{\perp,i,rhs} + 4\beta_i cfgh\tau\left(1-sx\right)2\left\langle J_1\right\rangle_a\hat{n}_e^n$$
$$-i96abcfghms\tau\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{A}_\parallel^n + 32\beta_i bcfghms\tau\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{p}_{\parallel,e}^n$$
$$+4h\left(2b-\beta_i cdfg\tau\left(1-4sy\right)\right)\hat{\bar{n}}_i^n - \beta_i cfgh\tau\left(1-sx\right)2\left\langle J_1\right\rangle_a\hat{n}_e^{n-1} + i24abcfghms\tau\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{A}_\parallel^{n-1}$$
$$-8\beta_i bcfghms\tau\left(\Delta t\right)^2\hat{k}_\parallel^2\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{p}_{\parallel,e}^{n-1} - h\left(2b-\beta_i cdfg\tau\left(1-4sy\right)\right)\hat{\bar{n}}_i^{n-1}$$
$$+i24abcfghs\tau\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a\hat{A}_A^{n+1} \tag{B.5}$$

$$\hat{\bar{p}}_{\perp,i}^{n+1} = -4\beta_i cfglsr2\left\langle J_1\right\rangle_a \hat{n}_{e,rhs} - i16acfglmsr\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{A}_{\parallel,rhs}$$
$$+\frac{16\beta_i cfglmsr}{3}\left(\Delta t\right)^2 \hat{k}_\parallel^2\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{p}_{\parallel,e,rhs} - \frac{2\beta_i fgl\tau}{3}\left(1-\beta_i cfgsv\right)2\left\langle J_1\right\rangle_a \hat{p}_{\perp,e,rhs}$$
$$+4\beta_i cdfglsr\hat{\bar{n}}_{i,rhs} + \frac{2}{3}\left(1-\beta_i efgl\tau\left(1-\beta_i cfgsv\right)\right)\hat{\bar{p}}_{\perp,i,rhs} + \frac{4\beta_i cfgl\tau}{3b}\left(1-sx\right)2\left\langle J_1\right\rangle_a \hat{n}_e^n$$
$$-i32acfglmsr\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{A}_\parallel^n + \frac{32\beta_i cfglmsr}{3}\left(\Delta t\right)^2 \hat{k}_\parallel^2\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{p}_{\parallel,e}^n$$
$$+\frac{16\beta_i cdfg\tau}{3b}\left(lsy - bh\nu\Delta t\hat{k}_\perp^{2m}\right)\hat{\bar{n}}_i^n + \frac{4}{3}\hat{\bar{p}}_{\perp,i}^n - \frac{\beta_i cfgl\tau}{3b}\left(1-sx\right)2\left\langle J_1\right\rangle_a \hat{n}_e^{n-1}$$
$$+i8acfglmsr\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{A}_\parallel^{n-1} - \frac{8\beta_i cfglmsr}{3}\left(\Delta t\right)^2 \hat{k}_\parallel^2\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{p}_{\parallel,e}^{n-1}$$
$$-\frac{4\beta_i cdfg\tau}{3b}\left(lsy - bh\nu\Delta t\hat{k}_\perp^{2m}\right)\hat{\bar{n}}_i^{n-1} - \frac{1}{3}\hat{\bar{p}}_{\perp,i}^{n-1} + i8acfglsr\hat{k}_\perp^2 2\left\langle J_1\right\rangle_a \hat{A}_A^{n+1} \tag{B.6}$$

Bibliography

[1] S. D Bale, P. J Kellogg, F. S Mozer, T. S Horbury, and H Reme. Measurement of the electric fluctuation spectrum of magnetohydrodynamic turbulence. *Phys. Rev. Lett.*, 94(21):1–4, Jun 2005.

[2] 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. *Phys. Plasmas*, 17(5):056109, Jan 2010.

[3] M Beer, S Cowley, and G Hammett. Field-aligned coordinates for nonlinear simulations of tokamak turbulence. *Phys. Plasmas*, Jan 1995.

[4] M. A Beer and G. W Hammett. Toroidal gyrofluid equations for simulations of tokamak turbulence. *Phys. Plasmas*, 3:4046, Nov 1996.

[5] MA Beer and GA Hammett. The dynamics of small-scale turbulence-driven flows. *Varenna Proceedings*, pages 1–11, Sep 1998.

[6] MA Beer, GW Hammett, Princeton University, and Dept. of Energy. Bounce averaged trapped electron fluid equations for plasma turbulence. *Phys. Plasmas*, 3(11):4018–4022, 1996.

[7] D Biskamp, E Schwarz, A Zeiler, A Celani, and JF Drake. Electron magnetohydrodynamic turbulence. *Phys. Plasmas*, 6:751, 1999.

[8] J Candy, R. E Waltz, and W Dorland. The local limit of global gyrokinetic simulations. *Phys. Plasmas*, 11:L25, May 2004.

[9] G. F Chew, M. L Goldberger, and F. E Low. The boltzmann equation and the one-fluid hydromagnetic equations in the absence of particle collisions. *Proceedings of the Royal Society of London. Series A*, 236:112, Jul 1956.

[10] B. I Cohen, D. C Barnes, J. M Dawson, G. W Hammett, W. W Lee, G. D Kerbel, J.-N Leboeuf, P. C Liewer, T Tajima, and R. E Waltz. The numerical tokamak project: simulation of turbulent transport. *Computer Physics Communications*, 87:1, May 1995.

[11] S. C Cowley, R. M Kulsrud, and R Sudan. Considerations of ion-temperature-gradient-driven turbulence. *Physics of Fluids B: Plasma Physics*, 3:2767, Oct 1991.

[12] A. M Dimits, G Bateman, M. A Beer, B. I Cohen, W Dorland, G. W Hammett, C Kim, J. E Kinsey, M Kotschenreuther, A. H Kritz, L. L Lao, J Mandrekas, W. M Nevins, S. E Parker, A. J Redd, D. E Shumaker, R Sydora, and J Weiland. Comparisons and physics basis of tokamak transport models and turbulence simulations. *Phys. Plasmas*, 7:969, Mar 2000.

[13] W Dorland and GW Hammett. Gyrofluid turbulence models with kinetic effects. *Physics of Fluids B Plasma Physics*, 5:812–812, 1993.

[14] W Dorland, F Jenko, M Kotschenreuther, and BN Rogers. Electron temperature gradient turbulence. *Phys. Rev. Lett.*, 85(26):5579–5582, 2000.

[15] Culham Centre for Fusion Energy. The international multi-tokamak confinement profile database. `http://tokamak-profiledb.ccfe.ac.uk/`, July 2010.

[16] EA Frieman and L Chen. Nonlinear gyrokinetic equations for lowfrequency electromagnetic waves in general plasma equilibria. *Physics of fluids*, 25:502, 1982.

[17] P Goldreich and S Sridhar. Toward a theory of interstellar turbulence. 2: Strong alfvenic turbulence. *The Astrophysical Journal*, 438:763, 1995.

[18] Gregory W Hammett and Francis W Perkins. Fluid moment models for landau damping with application to the ion-temperature-gradient instability. *Phys. Rev. Lett.*, 64:3019, Jun 1990.

[19] G Howes, W Dorland, S Cowley, G Hammett, E Quataert, A A Schekochihin, and T Tatsuno. Kinetic simulations of magnetized turbulence in astrophysical plasmas. *Phys. Rev. Lett.*, Jan 2008.

[20] G. G Howes, S. C Cowley, W Dorland, G. W Hammett, E Quataert, and A. A Schekochihin. A model of turbulence in magnetized plasmas: Implications for the dissipation range in the solar wind. *J. Geophys. Res.*, 113(A5):1–24, May 2008.

[21] GG Howes, SC Cowley, W Dorland, GW Hammett, E Quataert, and AA Schekochihin. Astrophysical gyrokinetics: Basic equations and linear theory. *The Astrophysical Journal*, 651(1):590–614, 2006.

[22] MA Hulsen. Stability of the implicit/explicit extension of the stiffly-stable schemes of gear. *MEAH - Internal Reports for the Section of Fluid Mechanics at Delft University of Technology*, 138, 1996.

[23] P. S Iroshnikov. Turbulence of a conducting fluid in a strong magnetic field. *Soviet Astronomy*, 7:566, Feb 1964.

[24] F Jenko and W Dorland. Prediction of significant tokamak turbulence at electron gyroradius scales. *Phys. Rev. Lett.*, 89:225001, Nov 2002.

[25] TM Antonsen Jr and B Lane. Kinetic equations for low frequency instabilities in inhomogeneous plasmas. *Physics of Fluids*, 23:1205, 1980.

[26] A. N Kolmogorov. The local structure of turbulence in incompressible viscous fluid for very large reynolds numbers. *Royal Society (London)*, 434:9, Jul 1991.

[27] RH Kraichnan. Inertial-range spectrum of hydromagnetic turbulence. *Physics of Fluids*, 8:1385, 1965.

[28] JA Krommes. Fundamental statistical descriptions of plasma turbulence in magnetic fields. *Physics Reports*, 360(1):1, 2002.

[29] D Mikkelsen and W Dorland. Dimits shift in realistic gyrokinetic plasma-turbulence simulations. *Phys. Rev. Lett.*, 101(13):135003, Sep 2008.

[30] S. A Orszag and C.-M Tang. Small-scale structure of two-dimensional magnetohydrodynamic turbulence. *Journal of Fluid Mechanics*, 90:129, Jan 1979.

[31] K. V Roberts and J. B Taylor. Gravitational resistive instability of an incompressible plasma in a sheared magnetic field. *Physics of fluids*, 8:315, Feb 1965.

[32] PH Rutherford and EA Frieman. Drift instabilities in general magnetic field configurations. *Physics of Fluids*, 11:569, 1968.

[33] A. A Schekochihin, S. C Cowley, W Dorland, G. W Hammett, G. G Howes, E Quataert, and T Tatsuno. Astrophysical gyrokinetics: Kinetic and fluid turbulent cascades in magnetized weakly collisional plasmas. *The Astrophysical Journal Supplement*, 182:310, May 2009.

[34] GI Taylor. Diffusion by continuous movements. *Proc. London Math. Soc*, 20(2):196–211, 1921.

[35] J. B Taylor and R. J Hastie. Stability of general plasma equilibria - i formal theory. *Plasma Physics*, 10:479, Jan 1968.