A Multiscale Approach to Turbulence and Transport in Magnetically Confined Plasmas

Student Number: 394365

University of Oxford

Submitted as one unit in partial completion of the

Mathematical and Theoretical Physics Master's Degree

Trinity 2017

Acknowledgements

First and foremost, I would like to thank Michael Barnes for his dedicated supervision this year. He has willingly given many hours to help me navigate my way through gyrokinetics for the first time, and has been a crucial source of assistance throughout the course of this work. I would also like to thank Felix Parra and Alex Schekochihin for their excellent plasma courses and tutoring, within both the MMathPhys and at Worcester College, and for aiding my understanding with many useful discussions. Lastly, for making my year in the Oxford Plasma Theory group a fun and memorable one, I would like to thank the current plasma DPhil students: Valerian Chen, Nicolas Christen, Plamen Ivanov, and Jason Parisi; special thanks are owed to Alessandro Geraldini, who also tutored the two MMathPhys plasma courses, and Michael Hardman, whose notes on the gyrokinetic equation were invaluable in the first two Chapters of this work.

Throughout my work on this dissertation, I have drawn on many resources in the well-established field of gyrokinetics. In places where a result has been directly taken from elsewhere, I have included a citation. However, since this dissertation presents gyrokinetics and the transport equations from the ground up, I have often taken inspiration from multiple sources and chosen to piece together my own account of how it all fits together, not quite matching any one source. Consequently, in order to give credit where it is due, a list of all the resources I have drawn on at some point in this dissertation is included below.

To get a feel for the subject, I began by studying the linear theory in a slab geometry featured in [1], before extending to the magnetic geometry featured in this work. In Chapter 2, the standard gyrokinetic ordering in Section 2.2 also features in many other papers, including [1–4]. The single-particle motion calculations were influenced by [5], and the development of the flux coordinate system in Section 2.4 was based on [4, 6]. The averages defined in Section 2.5 are based on those featured in [2, 4, 6, 7]. For the asymptotic expansion in Chapter 3, I have drawn on [2, 4, 5]. Chapter 4 was influenced by [2, 4], and the particle transport equation which we derive matches the equation in [3].

Abstract

In this dissertation, we present from first principles the equations which describe how turbulence evolves on the microscale, and how this can subsequently influence the large-scale properties of a fusion plasma. To begin our derivation, we first take advantage of the well-separated spatial scales in a tokamak plasma to define an expansion parameter $\epsilon \equiv \rho_i/L$, the ratio of the ion gyroradius to equilibrium scale length. We proceed to show that there is also a separation of timescales between the equilibrium, fluctuation and gyrofrequencies, $t_{eq}^{-1} \sim \epsilon^2\omega \sim \epsilon^3\Omega_{0_i}$. This, along with anisotropic turbulence and the low-flow ordering of the electric field, forms the basis of the low-flow gyrokinetic ordering which we use throughout. Separating the distribution function f into an equilibrium-scale part F and a fluctuating part δf and expanding each in ϵ, we perform an asymptotic expansion of the Fokker-Planck equation, introducing an average over gyroangle at each order to close our equations by eliminating all f_2 dependence. Examining terms order-by-order, the lowest-order distribution function F_0 is found to be a Maxwellian, and the first-order corrections F_1 and δf_1 can be decomposed: F_1 is split into a gyrotropic part $\widetilde{F_1}$ and a finite-gyroradius correction $\boldsymbol{\rho} \cdot \nabla F_0$, where $\boldsymbol{\rho}$ is the gyroradius vector, and δf_1 is split into a gyrotropic part h and a Boltzmann response to the fluctuating electric potential $\delta\phi$. The equilibrium density n_0 and temperature T_0 are also shown to be constant on a flux surface, and the equations which describe the evolution of the neoclassical $(\widetilde{F_1})$ and turbulent (h) parts of f are presented. At third order in ϵ, we finally present the transport equations for the evolution of n_0 and p_0, illustrating the contributions from neoclassical, collisional and turbulent fluxes. We demonstrate the need to average both over a flux surface, to leave only the cross-flux-surface contribution, and intermediate time and space scales, such that only the statistical average of the turbulence contributes to the flux. An explicit derivation of the n_0 evolution equation is shown.

Contents

The future is green energy, sustainability, renewable energy.

— Arnold Schwarzenegger

1 Preface

The move to renewable energy sources is a vital step for humanity over the coming decades. One particular method of obtaining almost limitless clean energy has presented itself as an ideal solution to this global problem: fusion of the lighter elements, as in the core of the Sun. A plentiful supply of reactants, non-radioactive products, no geographical constraints and no risk of nuclear meltdown make terrestrial nuclear fusion experiments the best candidate for securing a clean energy future.

Magnetic confinement fusion (MCF), in which large doughnut-shaped devices known as tokamaks confine a plasma with strong magnetic fields while it is heated to around 150 million °C, has achieved a high fusion rate. However, in current experiments, energetic breakeven is yet to be reached. Vital to understanding the large-scale heat and particle transport in these devices is the coupling of the large and small scales in the system: turbulent transport on the microscale causes heat to be lost from the central plasma in the tokamak, leading to reduced fusion performance.

The goal of this dissertation is to present, from first principles, the equations which describe how turbulence evolves on the microscale, and how turbulence can influence the large-scale properties of a fusion plasma. We will derive properties of the particle distribution function at both large (system) scales, and small (gyroradius) scales, using a systematic order-by-order expansion of the Fokker-Planck equation. We will then use the gyrokinetic ordering to take moments of the Fokker-Planck equation, and show how equations for the density and temperature evolution inside a tokamak can be derived.

The beginning is the most important part of the work.

— Plato

2

Background

2.1 Introduction

We begin this dissertation with a discussion of some important ideas concerning how particles move in the strong magnetic fields inside a tokamak. In such a situation, we find that we can motivate a separation of space and time scales based on physical arguments and experimental data, which we can use to simplify the equations of single-particle motion. Treatment of the magnetic geometry inside a tokamak is made significantly more tractable by the use of flux coordinates. The separation of spacetime scales allows us to later simplify the evolution equations for the equilibrium quantities by averaging over the intermediate scales – such an average will coarse-grain the high-frequency, short-wavelength variation on the gyroscale. We define these intermediate spacetime averages, along with averages over the particle's gyromotion and over a flux surface, at the end of this Chapter.

The rest of this Chapter is organised as follows. First, we discuss the orderings and assumptions which are built into the gyrokinetic model, justifying the relative sizes of all parameters. Then, we will calculate the motion of a single particle in a strong magnetic field, in order to choose an appropriate set of variables in which to describe the problem. Following this, we discuss in detail the magnetic geometry of a tokamak plasma. Finally, we will use our results to define a set of averages which will be extensively referred back to throughout this dissertation.

2.2 Multiple Scales and the Gyrokinetic Ordering

In a magnetised plasma, a charged particle will gyrate around its guiding centre with a lowest-order frequency

$$\Omega_{0_s} = \frac{Z_s e B_0}{m_s c}, \tag{2.1}$$

known as the *gyrofrequency* of species s. Here, B_0 is the magnitude of the equilibrium magnetic field, c is the speed of light, e is the electron charge, and Z_s and m_s are the atomic number and mass of species s respectively. The radius of gyration is

$$\rho_s = \frac{m_s c v_\perp}{Z_s e B_0} \sim \frac{v_{Ts}}{\Omega_{0_s}}, \tag{2.2}$$

known as the *gyroradius* of species s. Here, v_{T_s} is the thermal speed of species s.

The plasma beta for a tokamak plasma satisfies $\beta_i \equiv p_{0_i}/(B_0^2/8\pi) \ll 1$. For example, inside the Joint European Torus (JET), which has achieved the highest ratio of energy out to energy in to date, the conditions are such that $\beta_i \sim 10^{-2}$ [4]. We are therefore experimentally justified in making the assumption of *strong magnetisation*, which leads to the ion gyroradius[1] being much smaller than the system scale L. We define an expansion parameter ϵ in the following way:

$$\boxed{\epsilon \equiv \frac{\rho_i}{L} \ll 1.} \tag{2.3}$$

This value has been measured inside JET: taking L to be of order the tokamak minor radius, the plasma inside JET satisfies $\epsilon \sim 10^{-3}$ [4]. If we assume that the characteristic fluctuation frequency ω is such that[2] $\omega \sim v_{Ti}/L$, we can see that

$$\omega \sim \epsilon \Omega_{0_i} \quad \Rightarrow \quad \boxed{\frac{\omega}{\Omega_{0_i}} \sim \epsilon \ll 1.} \tag{2.4}$$

It is clear that the turbulent and equilibrium spatial scales, and the turbulent and gyroperiod timescales, are both well-separated by a factor of $\epsilon \ll 1$.

[1] Inside a tokamak, there are electrons present as well as ions. We will choose to order relative to the ion parameters, since the electron gyrofrequency is such that $\Omega_{0_e} = (m_i/Zm_e)\Omega_{0_i} \gg \Omega_{0_i}$, and the electron gyroradius is such that $\rho_e \sim Z\sqrt{m_e/m_i}\rho_i \ll \rho_i$ as long as $T_i \sim T_e$. This means that we can treat Ω_{0_i} as the fast timescale, and ρ_i as the small length scale, since the electron equivalents will be even faster and smaller and so will be negligible in comparison.

[2] The fluctuation frequency is given by the typical speed-distance ratio for a turbulent eddy, which is v_E/ρ_i. Hence, in the low-flow regime, $\omega \sim \epsilon v_{T_i}/\rho_i \sim v_{T_i}/L$.

A further separation of timescales can be found by turning to a diffusive treatment of turbulent fluctuations. Let $\xi_0 = \xi_0(x,t)$ be some quantity that varies on the equilibrium spacetime scale due to diffusive transport on the microscale. The cross-field diffusion equation in one dimension is then

$$\frac{\partial \xi_0}{\partial t} \sim \frac{(\Delta x)^2}{\Delta t}\frac{\partial^2 \xi_0}{\partial x^2}, \tag{2.5}$$

where Δx and Δt are the diffusive step length and timescale. We can define t_{eq} and L as the time and space scales for the equilibrium variation through derivatives of ξ_0, i.e.

$$\frac{\partial \xi_0}{\partial t} \sim \frac{\xi_0}{t_{eq}}, \qquad \frac{\partial^2 \xi_0}{\partial x^2} \sim \frac{\xi_0}{L^2}. \tag{2.6}$$

We will assume that turbulent eddies are of the length scale ρ_i in the cross-field direction, and of length scale L along $\hat{\mathbf{b}}$, where $\hat{\mathbf{b}} = \mathbf{B}_0/B_0$ is the unit direction of the equilibrium magnetic field.[3] Therefore, the scales for the turbulent cross-field transport are $\Delta t \sim \omega^{-1}$ and $\Delta x \sim \rho_i$ respectively, and so we may order terms in the diffusion equation to obtain the size of the equilibrium timescale,

$$\frac{1}{t_{eq}} \sim \rho_i^2 \omega \frac{1}{L^2} \sim \epsilon^3 \Omega_{0_i}. \tag{2.7}$$

We can clearly see that there are three well-separated timescales (t_{eq}^{-1}, ω, Ω_{0_i}) and two well-separated spatial scales (ρ_i, L). Their relative sizes are summarised:

$$\boxed{\frac{1}{t_{eq}} \sim \epsilon^2 \omega \sim \epsilon^3 \Omega_{0_i}, \qquad \rho_i \sim \epsilon L.} \tag{2.8}$$

The scales for spatial and temporal derivatives are also summarised:

$$\boxed{\nabla \xi_0 \sim \frac{\xi_0}{L}, \qquad \hat{\mathbf{b}} \cdot \nabla \delta\xi \sim \frac{\delta\xi}{L}, \qquad \nabla_\perp \delta\xi \sim \frac{\delta\xi}{\rho_i},} \tag{2.9}$$

$$\boxed{\frac{\partial \xi_0}{\partial t} \sim \epsilon^3 \Omega_{0_i} \xi_0, \qquad \frac{\partial \delta\xi}{\partial t} \sim \epsilon \Omega_{0_i} \delta\xi,} \tag{2.10}$$

where ξ_0 is any equilibrium quantity and $\delta\xi$ is any fluctuating quantity.

[3] An intuitive argument as to why this should be the case is as follows: perpendicular to $\hat{\mathbf{b}}$, transport is suppressed due to the tight gyration of particles around the field lines, which reduces the step length for diffusive transport in this direction. In contrast, particles can stream freely along $\hat{\mathbf{b}}$, resulting in a shearing and stretching of turbulent eddies in this direction. This is also in good agreement with experiment [8–10].

In our treatment, the electric and magnetic fields inside the tokamak are split into equilibrium and fluctuating parts as follows:

$$\boxed{\mathbf{B}(\mathbf{r},t) = \mathbf{B}_0(\mathbf{r}) + \delta\mathbf{B}(\mathbf{r},t),} \qquad \boxed{\mathbf{E}(\mathbf{r},t) = \delta\mathbf{E}(\mathbf{r},t),} \tag{2.11}$$

with $|\delta\mathbf{B}|/|\mathbf{B}_0| \sim \epsilon$. Here, we have set $\mathbf{E}_0$ to zero since we will consider only the low-flow regime, in which the $E \times B$ velocity is an order smaller than the thermal velocity,

$$\mathbf{v}_E = \frac{c}{B_0}\delta\mathbf{E} \times \hat{\mathbf{b}} \sim \epsilon v_{Ti}, \tag{2.12}$$

with $B_0 \equiv |\mathbf{B}_0|$. We have also chosen not to include explicit time dependence in the equilibrium magnetic field.[4] From (2.12), we can deduce the size of $\delta\mathbf{E}$:

$$|\delta\mathbf{E}| \sim \epsilon\frac{v_{Ti}}{c}B_0. \tag{2.13}$$

We choose to write the fluctuating fields in terms of a scalar and vector potential,

$$\delta\mathbf{B}(\mathbf{r},t) = \nabla \times \delta\mathbf{A}(\mathbf{r},t), \qquad \delta\mathbf{E}(\mathbf{r},t) = -\nabla\delta\phi(\mathbf{r},t) - \frac{1}{c}\frac{\partial\delta\mathbf{A}(\mathbf{r},t)}{\partial t}. \tag{2.14}$$

Using the fact that $\delta\mathbf{A}$ and $\delta\phi$ vary on the microscale, we can use (2.13) and (2.14) to find that

$$|\delta\mathbf{A}| \sim \epsilon\rho_i B_0, \qquad \delta\phi \sim \epsilon\rho_i\frac{v_{Ti}}{c}B_0. \tag{2.15}$$

In a similar fashion, we find that

$$\left|\frac{1}{c}\frac{\partial\delta\mathbf{A}}{\partial t}\right| \sim \frac{1}{c}\omega(\epsilon\rho_i B_0) \sim \epsilon^2\frac{v_{Ti}}{c}B_0, \tag{2.16}$$

meaning that the electric field is *electrostatic* to leading order,

$$\delta\mathbf{E} = -\nabla\delta\phi + \mathcal{O}(\epsilon^2\frac{v_{Ti}}{c}B_0). \tag{2.17}$$

The final assumptions we make are that the collision frequency ν is of order the fluctuation frequency ω, and that the variation of the distribution function in velocity space is of order the thermal velocity:

$$\boxed{\nu \sim \omega,} \qquad \boxed{\left|\frac{\partial f_s}{\partial\mathbf{v}}\right| \sim \frac{f_s}{v_T}.} \tag{2.18}$$

[4]It can be shown that the timescale for variation of the equilibrium magnetic field is the *resistive timescale*, which is much longer than the timescale for the equilibrium particle and energy transport. Hence, in deriving the transport equations, the equilibrium field can be treated as constant. (More information can be found in [4].)

The first assumption is made in order to keep our derivation as general as possible; if we wish to, we may later take either limit $\nu \ll \omega$ or $\nu \gg \omega$ as subsidiary expansions. The second assumption follows from the first, and the condition that the turbulent collision terms balance with the evolution of the fluctuating part of the distribution function δf:

$$\frac{\partial \delta f}{\partial t} \sim C[\delta f_s, F_{0'_s}] \quad \Rightarrow \quad \omega \delta f \sim \nu v_{T_i}^2 \frac{\partial^2 \delta f}{\partial v^2} \quad \Rightarrow \quad \delta v \sim v_{T_i}. \tag{2.19}$$

2.3 Single Particle Motion

The motion of a single particle of mass m and charge Ze in the electric and magnetic fields described above is governed by the equations

$$\frac{d\mathbf{v}}{dt} = \frac{Ze}{m}\left(\delta \mathbf{E} + \frac{\mathbf{v} \times (\mathbf{B}_0 + \delta \mathbf{B})}{c}\right), \qquad \frac{d\mathbf{r}}{dt} = \mathbf{v}, \tag{2.20}$$

with the Vlasov operator d/dt defined as

$$\frac{d}{dt} \equiv \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla + \frac{d\mathbf{v}}{dt} \cdot \frac{\partial}{\partial \mathbf{v}}. \tag{2.21}$$

We define the gyrophase φ by decomposing the velocity $\mathbf{v}$ into components parallel and perpendicular to the local equilibrium magnetic field direction $\hat{\mathbf{b}}$,

$$\mathbf{v} = v_{\parallel} \hat{\mathbf{b}} + v_{\perp}(\cos\varphi \hat{\mathbf{e}}_1 + \sin\varphi \hat{\mathbf{e}}_2), \tag{2.22}$$

where we have chosen our basis vectors $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{b}}\}$ such that they form an instantaneous, local, right-handed orthonormal set, obeying $\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2 = \hat{\mathbf{b}}$ and $(\hat{\mathbf{e}}_1 \times \hat{\mathbf{e}}_2) \cdot \hat{\mathbf{b}} = 1$.

We wish to find a set of phase space variables which isolates the fast-timescale variation. *A priori*, we would expect the kinetic energy and the magnetic moment of an individual particle,

$$\varepsilon = \frac{1}{2} m \mathbf{v} \cdot \mathbf{v}, \qquad \mu = \frac{m}{2B_0} \mathbf{v}_{\perp} \cdot \mathbf{v}_{\perp} \tag{2.23}$$

to vary on a longer timescale than the gyrophase φ, since for a free particle in a magnetic field they are conserved, and the electric field and collision frequency are both small. We proceed to calculate the rate of change of these variables to show that this is indeed the case.

First, we calculate the rate of change of μ. Writing the perpendicular velocity $\mathbf{v}_\perp$ as $(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \cdot \mathbf{v}$, we find that

$$\boxed{\frac{d\mu}{dt} = -\frac{\mu}{B_0}\mathbf{v} \cdot \nabla B_0 + \frac{m}{B_0}\mathbf{v}_\perp \cdot \left(\frac{Ze}{m}\left(\delta\mathbf{E} + \frac{\mathbf{v} \times \delta\mathbf{B}}{c}\right) - v_\parallel \mathbf{v} \cdot \nabla\hat{\mathbf{b}}\right),} \tag{2.24}$$

using 2.21, with (2.20) for $d\mathbf{v}/dt$, and $\mathbf{v}_\perp \cdot (\mathbf{v} \times \mathbf{B}_0) = 0$. By using (2.13) for the order of $\delta\mathbf{E}$, we can see that the magnetic moment μ varies on a timescale $\sim \epsilon\Omega_0$.[5]

Now turning to the rate of change of kinetic energy, we can see immediately that

$$\boxed{\frac{d\varepsilon}{dt} = m\mathbf{v} \cdot \frac{d\mathbf{v}}{dt} = Ze\mathbf{v} \cdot \delta\mathbf{E},} \tag{2.25}$$

where again we have substituted for $d\mathbf{v}/dt$ using (2.20). Referring to (2.13) again, we can see that ε also varies on a timescale $\sim \epsilon\Omega_0$.

The rate of change of φ may be calculated as follows. We begin with taking the two perpendicular components of the velocity, $\mathbf{v} \cdot \hat{\mathbf{e}}_1 = v_\perp \cos\varphi$ and $\mathbf{v} \cdot \hat{\mathbf{e}}_2 = v_\perp \sin\varphi$, and taking the time derivative of each equation. We can manipulate the resulting equations to give

$$v_\perp \frac{d\varphi}{dt} = \cos\varphi \frac{d}{dt}(\mathbf{v} \cdot \hat{\mathbf{e}}_2) - \sin\varphi \frac{d}{dt}(\mathbf{v} \cdot \hat{\mathbf{e}}_1) = -\frac{\mathbf{v} \times \hat{\mathbf{b}}}{v_\perp} \cdot \frac{d\mathbf{v}}{dt} + v_\perp \hat{\mathbf{e}}_1 \cdot \frac{d\hat{\mathbf{e}}_2}{dt}, \tag{2.26}$$

where we have used the standard properties of a right-handed orthonormal basis. Hence, using (2.20) to substitute for $d\mathbf{v}/dt$, we find that

$$\boxed{\frac{d\varphi}{dt} = -\Omega_0 - \frac{\mathbf{v} \times \hat{\mathbf{b}}}{v_\perp^2} \cdot \frac{Ze}{m}\left(\delta\mathbf{E} + \frac{\mathbf{v} \times \delta\mathbf{B}}{c}\right) + \hat{\mathbf{e}}_1 \cdot \frac{d\hat{\mathbf{e}}_2}{dt},} \tag{2.27}$$

where we have identified

$$-\frac{\mathbf{v} \times \hat{\mathbf{b}}}{v_\perp^2} \cdot \frac{Ze}{m}\frac{\mathbf{v} \times \mathbf{B}_0}{c} = -\Omega_0. \tag{2.28}$$

(2.27) tells us that the leading-order timescale for the rate of change of the gyrophase is Ω_0, as predicted, which is an order faster than the variation of ε and μ. This means we have found a set of velocity space variables which isolates the fast gyromotion of the particles. To complete the set of variables, we must include the sign of the

[5]We omit the subscript i from now on. All further references to Ω_0 and ρ will be referring to ion quantities.

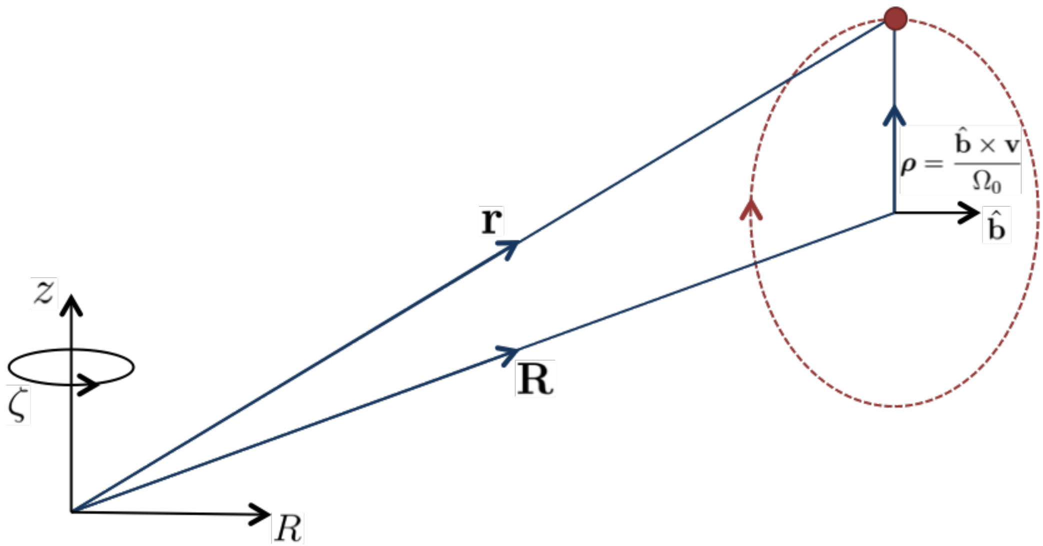

Figure 2.1: The guiding centre transformation expresses $\mathbf{r}$ in terms of the guiding centre $\mathbf{R}$ and the gyroradius vector $\boldsymbol{\rho} = \hat{\mathbf{b}} \times \mathbf{v}/\Omega_0$. The ion is shown as a red dot, and its instantaneous gyromotion is shown as a red dotted line.

parallel velocity $\sigma = v_\parallel/|v_\parallel|$, meaning that our final set of velocity-space variables is $(\varphi, \varepsilon, \mu, \sigma)$. We will interchange between these *gyrokinetic variables* and the standard velocity space variables $\mathbf{v}$ later on, as is convenient.

We now turn to the position variable $\mathbf{r}$. Its rate of change is given by (2.22), and therefore it clearly still varies with φ on the fast timescale Ω_0. In order to separate scales, we make the *guiding centre transformation* by expressing $\mathbf{r}$ in terms of the guiding centre position $\mathbf{R}$ and the gyroradius vector $\boldsymbol{\rho}$:

$$\boxed{\mathbf{R} \equiv \mathbf{r} - \boldsymbol{\rho},} \qquad \boldsymbol{\rho} = \frac{\hat{\mathbf{b}} \times \mathbf{v}}{\Omega_0}. \tag{2.29}$$

A visualisation of this transformation is given in Figure 2.1.[6]

The rate of change of $\mathbf{R}$ is given by

$$\frac{d\mathbf{R}}{dt} = v_\parallel \hat{\mathbf{b}} + \mathbf{v} \times \left(\mathbf{v} \cdot \nabla \left(\frac{\hat{\mathbf{b}}}{\Omega_0} \right) \right) + \frac{Ze}{m} \left(\delta\mathbf{E} + \frac{\mathbf{v} \times \delta\mathbf{B}}{c} \right) \times \left(\frac{\hat{\mathbf{b}}}{\Omega_0} \right) + \mathcal{O}(\epsilon^3 v_T), \tag{2.30}$$

where we have used (2.20), and the fact that $\hat{\mathbf{b}}/\Omega_0$ is a function of position only. It will prove convenient to define the velocities

$$\mathbf{v}_M \equiv \mathbf{v} \times \left(\mathbf{v} \cdot \nabla \left(\frac{\hat{\mathbf{b}}}{\Omega_0} \right) \right), \qquad \mathbf{v}_E \equiv \frac{Ze}{m} \left(\delta\mathbf{E} + \frac{\mathbf{v} \times \delta\mathbf{B}}{c} \right) \times \left(\frac{\hat{\mathbf{b}}}{\Omega_0} \right), \tag{2.31}$$

[6] This transformation can also be derived using the systematic approach for higher-order corrections to the gyrokinetic variable $\mathbf{R}$ detailed in [11].

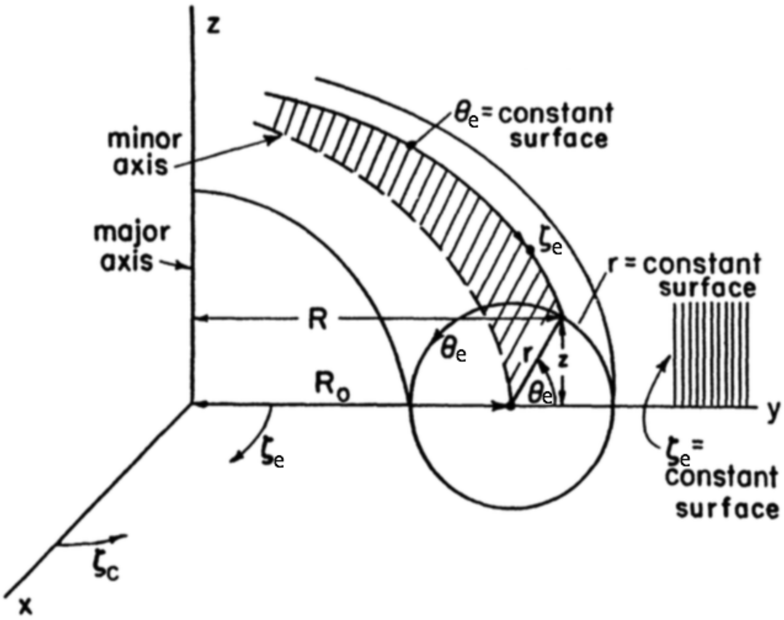

Figure 2.2: The elementary toroidal system: an ideal torus, defined such that both the handle of the torus and its poloidal cross-section are circular. The major and minor axes, along with the two coordinate systems: cylindrical (R, ζ_c, Z) and elementary toroidal (r, θ_e, ζ_e), are shown. Note that ζ_c and ζ_e differ by a sign, to ensure both systems are right-handed. Figure adapted from [6].

meaning that we can write the succinct expression

$$\boxed{\frac{d\mathbf{R}}{dt} = v_{\|}\hat{\mathbf{b}} + \mathbf{v}_M + \mathbf{v}_E + \mathcal{O}(\epsilon^3 v_T).} \tag{2.32}$$

2.4 Flux Coordinates and Magnetic Geometry

Flux coordinates

A tokamak is topologically equivalent to a solid torus, and so we first consider the *elementary* toroidal system shown in Figure 2.2. We can choose either conventional cylindrical coordinates, or elementary toroidal coordinates: r, a variable which characterises the length outwards from the minor axis, and two angle variables θ_e and ζ_e, known as the *poloidal* and *toroidal angles* respectively.

For a magnetic field configuration which is not an ideal torus, but is topologically equivalent, we can define a useful coordinate system by making use of the fact that our plasma is magnetised. Instead of the minor axis, we define the *magnetic axis* to be the field line where the equilibrium poloidal field B_{0_θ} is zero, meaning that it is a

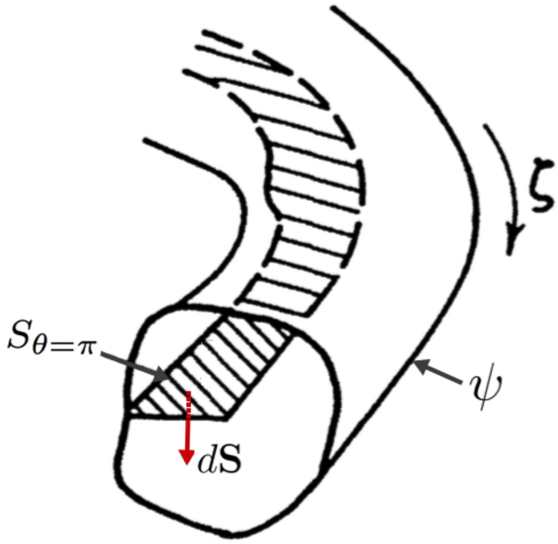

Figure 2.3: An illustration of the $S_{\theta=\pi}$ surface used in defining the poloidal ribbon flux (lined surface). It is defined as the intersection of a surface of constant ψ with the $\theta = \pi$ surface. This flux is clearly monotonically increasing as we move away from the magnetic axis (dashed curve), with a value of zero on the magnetic axis. Figure adapted from [6].

closed curve around the major axis. A *flux surface* is a closed toroidal surface on which $\hat{\mathbf{b}} \cdot \hat{\mathbf{n}} = 0$ everywhere, where $\hat{\mathbf{n}}$ is the surface normal. This notion allows us to define a new way of characterising distance outwards from the magnetic axis by the use of *flux surface labels*: these are constant on a flux surface, meaning $\hat{\mathbf{b}} \cdot \nabla\Gamma = 0$ for a general flux surface label Γ. There is some freedom in choosing Γ, but it is convenient to define it in such a way that $\Gamma = 0$ on the magnetic axis, and that it is monotonically increasing as we move outwards from the magnetic axis (similar to r in the elementary case). We therefore choose Γ to be the poloidal ribbon flux ψ, defined as the flux through a circular ribbon stretched between the flux surface and the magnetic axis (Figure 2.3):[7]

$$\boxed{\psi = \frac{1}{2\pi} \int_{S_{\theta=\pi}} d\mathbf{S} \cdot \mathbf{B}_0.} \tag{2.33}$$

Note that the form of ψ as shown above is not dependent on the particular forms of θ and ζ we choose; all that is required is that they are 2π-periodic and are poloidal and toroidal variables. We have therefore found a good set of *flux coordinates*: (ψ, θ, ζ).

[7] The 2π normalisation is included according to convention.

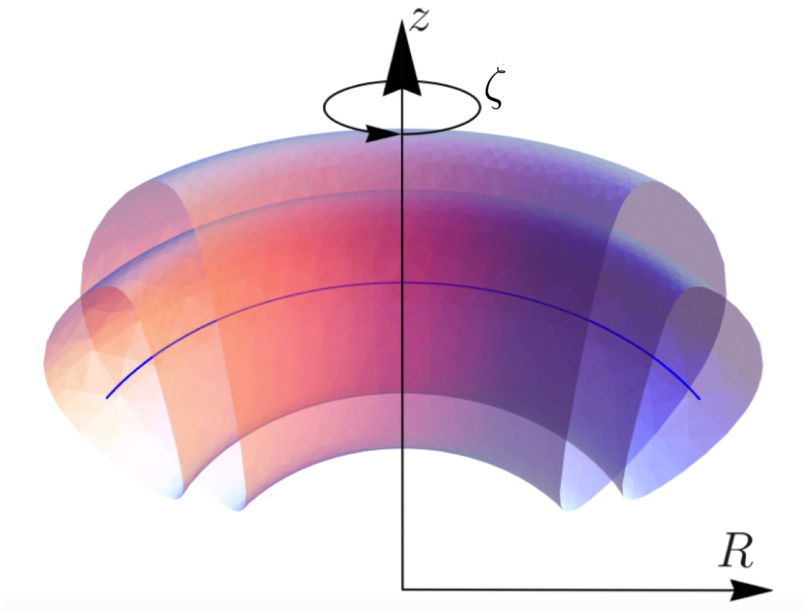

Figure 2.4: The cylindrical coordinate system employed in a tokamak. The axisymmetric flux surfaces are shown as concentric surfaces about the magnetic axis (blue curve). Figure adapted from [4].

Axisymmetric magnetic field

In the case of an axisymmetric equilibrium magnetic field[8], we can write $\mathbf{B}_0$ in a particularly elegant form, which we derive here. For an axisymmetric magnetic geometry, in cylindrical coordinates (R, ζ, z) as shown in Figure 2.4, we can write the field as

$$\mathbf{B}_0 = \nabla \times \mathbf{A}_0 = \underbrace{\left(\frac{\partial A_{0_R}}{\partial z} - \frac{\partial A_{0_z}}{\partial R} \right) R}_{\equiv I(R,z)} \nabla \zeta + \frac{1}{R} \frac{\partial (R A_{0_\zeta})}{\partial R} \nabla z - \frac{\partial A_{0_\zeta}}{\partial z} \nabla R, \tag{2.34}$$

where $I(R, z) = R^2 \mathbf{B}_0 \cdot \nabla \zeta$ is a measure of the toroidal magnetic field. Recalling our definition of the poloidal ribbon flux (cf. (2.33)), we see that

$$\begin{aligned} \psi(R, z) \equiv \frac{1}{2\pi} \int_{S_{\theta=\pi}} \mathbf{B}_0(R, z) \cdot d\mathbf{S} &= \frac{1}{2\pi} \int_{\partial S_{\theta=\pi}} \mathbf{A}_0(R, z) \cdot d\mathbf{l} \\ &= \frac{1}{2\pi} \int_0^{2\pi} R d\zeta A_{0_\zeta}(R, z) \\ &= R A_{0_\zeta}(R, z). \end{aligned} \tag{2.35}$$

Taking the gradient, we find that

$$\nabla \psi = \frac{\partial (R A_{0_\zeta})}{\partial R} \nabla R + R \frac{\partial A_{0_\zeta}}{\partial z} \nabla z, \tag{2.36}$$

[8] In modern tokamaks, the deviation from this condition is much smaller than ϵ (cf. [4]) and so we are safe to make this assumption for $\mathbf{B}_0$, since all deviations are small enough to be absorbed into the $\mathcal{O}(\epsilon^2 B_0)$ (or higher) corrections.

and hence that

$$\nabla\psi \times \nabla\zeta = \frac{1}{R}\frac{\partial(RA_{0_\zeta})}{\partial R}\nabla z - \frac{\partial A_{0_\zeta}}{\partial z}\nabla R, \tag{2.37}$$

which exactly matches the part of $\mathbf{B}_0$ that we didn't define as $I(R, z)$. We can therefore write an axisymmetric equilibrium magnetic field as[9]

$$\boxed{\mathbf{B}_0 = I(R, z)\nabla\zeta + \nabla\psi \times \nabla\zeta.} \tag{2.38}$$

2.5 Useful Averages

Gyroaverage

The first average which we will introduce is the *gyroaverage* $\langle\alpha\rangle_{\mathbf{R}}$ of any function of gyroangle $\alpha(\varphi, \ldots)$ at fixed guiding centre $\mathbf{R}$, defined by

$$\boxed{\langle\alpha\rangle_{\mathbf{R}} \equiv \frac{1}{2\pi}\int_0^{2\pi} d\varphi\alpha(\varphi, \ldots)\bigg|_{\mathbf{R}}.} \tag{2.39}$$

This average eliminates the fast, Ω-timescale variation of the function $\alpha(\varphi, \ldots)$, reducing the dimension of phase space from 6D to 5D.

Intermediate spatial average

Utilising the separation of spatial scales, we employ an *intermediate spatial average* (ISA) over length scales λ such that $\rho \ll \lambda \ll L$. We can formally define this length as $\lambda \sim \epsilon^{\frac{1}{2}}L$, since ϵ is sufficiently small as to make $\epsilon^{\frac{1}{2}} \ll 1$, meaning λ is still well-separated from both the gyroradius and equilibrium scales.

Armed with the information about the magnetic geometry of our system from Section 2.4, we are define the ISA of some quantity X as a volume average over the toroidal annulus between two flux surfaces characterised by ψ and $\psi + \Delta\psi$, where the characteristic length of the annulus in the radial direction $d \sim \lambda$:

$$\boxed{\langle X\rangle_\lambda \equiv \frac{1}{\Delta V}\int_{\Delta V} X\, d^3\mathbf{r} = \frac{1}{\Delta V}\int_0^{2\pi} d\zeta \int_{-\pi}^{\pi} d\theta \int_{\psi}^{\psi+\Delta\psi} d\psi' JX.} \tag{2.40}$$

[9]It can also be shown that $I = I(\psi)$ is a flux function; however, we will not use that property in this work.

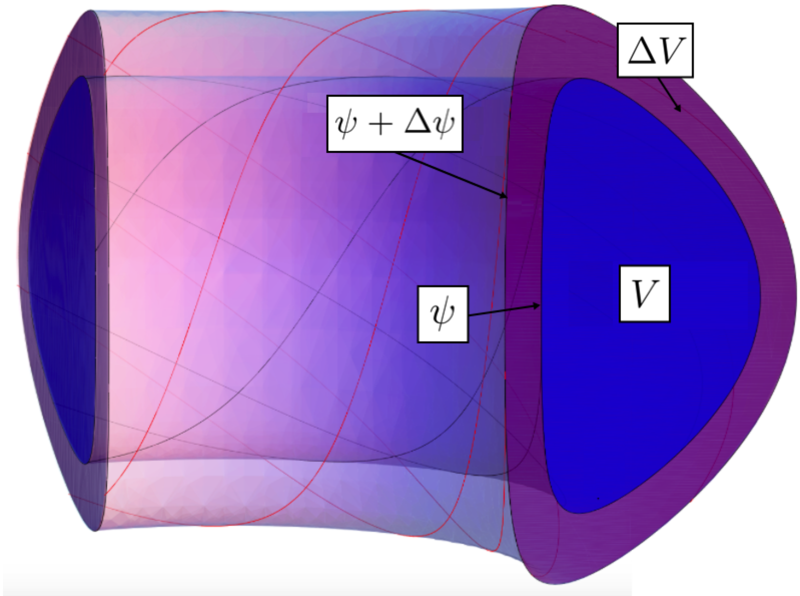

Figure 2.5: The quantities involved in the intermediate spatial average are labelled in a tokamak. Figure adapted from [4].

Here, ΔV is the volume of the toroidal annulus,

$$\Delta V = \int_0^{2\pi} d\zeta \int_{-\pi}^{\pi} d\theta \int_{\psi}^{\psi+\Delta\psi} d\psi' J, \tag{2.41}$$

and J is the Jacobian of the coordinate transformation $(\mathbf{r}) \to (\psi', \theta, \zeta)$,[10]

$$J \equiv \left| \det \frac{\partial(\mathbf{r})}{\partial(\psi', \theta, \zeta)} \right| = \frac{\partial \mathbf{r}}{\partial \zeta} \cdot \left(\frac{\partial \mathbf{r}}{\partial \psi'} \times \frac{\partial \mathbf{r}}{\partial \theta} \right) = \frac{\partial x_i}{\partial \zeta} \left(\varepsilon_{ijk} \frac{\partial x_j}{\partial \psi'} \frac{\partial x_k}{\partial \theta} \right) = \Big(\nabla\zeta \cdot (\nabla\psi' \times \nabla\theta) \Big)^{-1}. \tag{2.42}$$

Note that since (ψ', θ, ζ) forms a right-handed coordinate system, the quantity above is always positive and so we no longer require the modulus. A cartoon of the structure of the flux surfaces relevant for the ISA is shown in Figure 2.5. Importantly, the ISA is taken at a fixed location in space characterised by some ψ, and since $\mathbf{B}_0$ does not vary with time, the ISA commutes with all time averages.

Flux surface average

We define the *flux surface average* as the limit of the ISA as $\Delta\psi$ and ΔV tend to zero.[11] The result is an average over the flux surface characterised by ψ. The volume

[10] Note that ψ' is used as an integration variable instead of ψ since we wish to keep the ISA as a function of the flux surface we are averaging around, which we have chosen to denote by ψ.

[11] These limits can clearly be taken concurrently, since $\Delta\psi \to 0$ implies vanishing ΔV.

element $d^3\mathbf{r}$ in the ISA can be decomposed by definition into

$$d^3\mathbf{r} \equiv |d\mathbf{r} \cdot d\mathbf{S}(\psi)|, \tag{2.43}$$

where $d\mathbf{r}$ is the infinitesimal change in the position vector along the ψ coordinate curve, and $d\mathbf{S}(\psi)$ is the vector surface element on the flux surface labelled by ψ. If we align $d\mathbf{S}$ such that when $d\mathbf{r}$ is in the positive-ψ direction, the quantity in (2.43) is positive, we may omit the absolute value bars. The outward normal to a flux surface labelled by ψ is $\nabla\psi/|\nabla\psi|$, since ψ is increasing away from the magnetic axis, and so $d\mathbf{S}(\psi)$ can be written as $\nabla\psi dS/|\nabla\psi|$. Using that $d\mathbf{r} \cdot \nabla\psi \equiv d\psi$, we are left with $d^3\mathbf{r} = d\psi dS/|\nabla\psi|$, and hence we may rewrite the ISA as

$$\langle X\rangle_\lambda = \frac{1}{\Delta V}\int_{\Delta V} X\ d^3\mathbf{r} = \frac{1}{\Delta V}\int_{\psi}^{\psi+\Delta\psi} d\psi' \int_{\partial V(\psi')} \frac{XdS}{|\nabla\psi'|}, \tag{2.44}$$

where $\partial V(\psi')$ is the flux surface labelled by ψ'.

Taking the limits $\Delta V \to 0$, $\Delta\psi \to 0$, we may formally substitute $\partial V(\psi') \to \partial V(\psi)$ and $|\nabla\psi'| \to |\nabla\psi|$ in (2.44). Doing so, we find that

$$\lim_{\substack{\Delta V\to 0\\ \Delta\psi\to 0}} \langle X\rangle_\lambda = \lim_{\substack{\Delta V\to 0\\ \Delta\psi\to 0}} \frac{1}{\Delta V}\int_{\psi}^{\psi+\Delta\psi} d\psi' \int_{\partial V(\psi')} \frac{XdS}{|\nabla\psi'|} = \lim_{\substack{\Delta V\to 0\\ \Delta\psi\to 0}} \frac{\Delta\psi}{\Delta V}\int_{\partial V(\psi)} \frac{XdS}{|\nabla\psi|}. \tag{2.45}$$

Substituting $V' = dV/d\psi$ in place of $\Delta V/\Delta\psi$, we arrive at an expression for the FSA:

$$\boxed{\langle X\rangle_\psi \equiv \lim_{\substack{\Delta V\to 0\\ \Delta\psi\to 0}} \langle X\rangle_\lambda = \frac{1}{V'}\int_{\partial V(\psi)} \frac{XdS}{|\nabla\psi|}.} \tag{2.46}$$

It is easy to see that the FSA can also be written as an integral over angles,

$$\langle X\rangle_\psi = \frac{1}{V'}\int_{-\pi}^{\pi} d\theta \int_0^{2\pi} d\zeta JX, \tag{2.47}$$

by implementing the following result in (2.46):

$$Jd\theta d\zeta = \frac{1}{|\nabla\psi|}\frac{d\theta d\zeta}{\hat{\mathbf{e}}^\psi \cdot (\nabla\theta \times \nabla\zeta)} = \frac{dS}{|\nabla\psi|}, \tag{2.48}$$

where $\hat{\mathbf{e}}^\psi = \nabla\psi/|\nabla\psi|$ is the unit contravariant basis vector in the ψ' direction.

Flux surface average of a divergence

A quantity it will later prove useful to have calculated is the FSA of a divergence $\nabla \cdot \mathbf{X}$, for some general vector $\mathbf{X}$. We do so as follows:

$$\begin{aligned}\langle \nabla \cdot \mathbf{X} \rangle_\psi &= \lim_{\substack{\Delta V \to 0 \\ \Delta\psi \to 0}} \frac{1}{\Delta V} \int_{\Delta V} \nabla \cdot \mathbf{X} d^3\mathbf{r} \\ &= \lim_{\substack{\Delta V \to 0 \\ \Delta\psi \to 0}} \frac{1}{\Delta V} \int_{\partial \Delta V} \mathbf{X} \cdot \frac{\nabla\psi'}{|\nabla\psi'|} dS \\ &= \lim_{\substack{\Delta V \to 0 \\ \Delta\psi \to 0}} \frac{1}{\Delta V} \left(\int_{\partial V(\psi+\Delta\psi)} \mathbf{X} \cdot \frac{\nabla\psi'}{|\nabla\psi'|}\bigg|_{\psi'=\psi+\Delta\psi} dS - \int_{\partial V(\psi)} \mathbf{X} \cdot \frac{\nabla\psi'}{|\nabla\psi'|}\bigg|_{\psi'=\psi} dS \right).\end{aligned} \tag{2.49}$$

Here, we have split $\partial\Delta V$ into its inner and outer bounding flux surfaces, and used that $\nabla\psi'/|\nabla\psi'|$ may be substituted for the flux surface normal $\hat{\mathbf{n}}$ up to a minus sign; indeed, the minus sign in the second term is due to the surface normal of the inner flux surface pointing anti-parallel to $\nabla\psi'$, towards the magnetic axis. Using (2.46), we can see that the first quantity in the brackets is just $V'\langle \mathbf{X} \cdot \nabla\psi \rangle_{\psi+\Delta\psi}$, and the second quantity is $V'\langle \mathbf{X} \cdot \nabla\psi \rangle_\psi$. We therefore find that

$$\langle \nabla \cdot \mathbf{X} \rangle_\psi = \lim_{\substack{\Delta V \to 0 \\ \Delta\psi \to 0}} \frac{1}{\Delta V} \Big(V'\langle \mathbf{X} \cdot \nabla\psi \rangle_{\psi+\Delta\psi} - V'\langle \mathbf{X} \cdot \nabla\psi \rangle_\psi \Big) = \frac{1}{V'} \frac{\partial}{\partial\psi} \Big(V'\langle \mathbf{X} \cdot \nabla\psi \rangle_\psi \Big), \tag{2.50}$$

where we have used the chain rule to obtain the final equality. We therefore have

$$\boxed{\langle \nabla \cdot \mathbf{X} \rangle_\psi = \frac{1}{V'} \frac{\partial}{\partial\psi} \Big(V'\langle \mathbf{X} \cdot \nabla\psi \rangle_\psi \Big).} \tag{2.51}$$

Intermediate time average

Analogously to the intermediate spatial average, we can define an intermediate time average (ITA) of some time-dependent quantity $\xi(t)$:

$$\boxed{\overline{\xi(t)} \equiv \frac{1}{\tau} \int_{t-\tau/2}^{t+\tau/2} \xi(t') dt',} \tag{2.52}$$

where τ is defined to be in between and well-separated from the turbulent and equilibrium timescales, $t_{eq}^{-1} \ll \tau^{-1} \ll \omega$. We may formally define its size as $\tau^{-1} \sim \epsilon\omega \sim \epsilon^2\Omega$. We proceed to see how this time average acts on various quantities.

This average coarse-grains timescales which are at least an order smaller than the equilibrium timescale. We may therefore treat all equilibrium quantities $\xi_0(t)$ as constant on this timescale, and therefore approximate $\xi_0(t') \approx \xi_0(t)$ in the integrand, meaning that $\overline{\xi_0(t)} \approx \xi_0(t)$. To see what kind of error we can expect, we can Taylor expand at the bounds of our integral: $\xi_0(t + \mathcal{O}(\tau)) = \xi_0(t) + \mathcal{O}(\tau)\partial\xi_0(t)/\partial t$. It can be seen that the last term is $\mathcal{O}(\epsilon\xi_0(t))$ by using that equilibrium quantities vary on the $t_{eq}^{-1} \sim \epsilon^3\Omega$ timescale, and then making use of the ordering stated above to state $\Omega\tau \sim \epsilon^{-2}$. We may therefore conclude that $\overline{\xi_0(t)} = \xi_0(t) + \mathcal{O}(\epsilon\xi_0(t))$ – hence, the ITA leaves equilibrium quantities unchanged to leading order in ϵ.

Since a fluctuating quantity $\delta\xi(t)$ oscillates many times during the period the ITA is averaging over, we use statistical periodicity of the turbulence in time to assert that $\overline{\delta\xi(t)} \approx 0$. To estimate the error here, we observe that the only contribution to the integral will be from the parts of the fluctuations which don't cancel, and so after cancellation the leftover piece will be of order the product of the typical width of a fluctuation period with the typical height of a fluctuation. We use that the typical amplitude of a fluctuation is $\delta\xi(t)$, and the fluctuation period $\omega^{-1} \sim \epsilon\tau$, to find that $\overline{\delta\xi(t)} = \mathcal{O}(\epsilon\delta\xi(t))$ – hence, the ITA drops each fluctuating term by an order in ϵ.

Be grateful of small things, because it is in them that your strength lies.

— Mother Teresa

3 Gyrokinetic Expansion

3.1 Introduction

Here we perform an asymptotic expansion of the Fokker-Planck equation. In doing so, we will derive many key results, including the gyrokinetic equation which describes the evolution of h, the gyrophase-independent, first-order piece of the fluctuating distribution function δf.

The remainder of this Chapter is organised as follows. We start with the Fokker-Planck equation for species s,

$$\boxed{\frac{df}{dt} = \frac{\partial f_s}{\partial t} + \mathbf{v} \cdot \frac{\partial f_s}{\partial \mathbf{r}} + \frac{Z_s e}{m_s}\left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c}\right) \cdot \frac{\partial f_s}{\partial \mathbf{v}} = \sum_{s'} C[f_s, f_{s'}],} \tag{3.1}$$

where f_s is the distribution function of species s, $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic field, c is the speed of light, $Z_s e$ and m_s are the charge and mass of species s, and $C[f_s, f_{s'}]$ is the collision operator describing collisions of species s with another species s'. We formally separate this equation into equilibrium and fluctuating parts, and then systematically expand in the small parameter ϵ (cf. (2.3)). Assuming all relevant dynamics occur on timescales much greater than a gyroperiod, the gyroaverage defined in (2.39) allows us to close our equations by eliminating the need to calculate the second-order correction f_2. We examine different orders of the gyroaveraged Fokker-Planck equation to derive a hierarchical set of equations for the components of the distribution function f: namely, the neoclassical ($\widetilde{F_1}$) and turbulent (h) first-order corrections to the Maxwell-Boltzmann equilibrium ($F_0 \exp(-Ze\delta\phi/T)$).

3.2 Fokker-Planck Equation

We begin our manipulation of the Fokker-Planck equation in the original $(\mathbf{r}, \mathbf{v})$ coordinates. Starting with (3.1) and substituting for the electric field in terms of potentials using (2.14), we find

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \frac{Ze}{m}\left(-\nabla\delta\phi - \frac{1}{c}\frac{\partial \delta\mathbf{A}}{\partial t} + \frac{\mathbf{v}\times\delta\mathbf{B}}{c} + \frac{\mathbf{v}\times\mathbf{B}_0}{c} \right) \cdot \frac{\partial f}{\partial \mathbf{v}} = \sum_{s'} C[f, f_{s'}], \quad (3.2)$$

where we have dropped the species label s. We split our distribution function f into an equilibrium part F and a fluctuating part δf and expand each in ϵ,

$$f = F_0 + F_1 + F_2 + \ldots + \delta f_1 + \delta f_2 + \ldots, \quad (3.3)$$

where the subscript indicates the order of each term in ϵ.[1]

We will take the collision operator to be ordered as

$$C[f, g] \sim \nu \frac{fg}{F_0} \sim \epsilon\Omega_0 \frac{fg}{F_0} \quad (3.4)$$

in accordance with [2, 4], since the collision frequency ν is taken to be of order the fluctuation frequency ω.

We proceed to examine the Fokker-Planck equation order-by-order.

$\mathcal{O}(\Omega_0 F_0)$: F_0 is gyrotropic

At lowest order in ϵ, we find that

$$\frac{Ze}{m}\left(\frac{\mathbf{v}\times\mathbf{B}_0}{c}\right) \cdot \frac{\partial F_0}{\partial \mathbf{v}} = 0 \qquad \Rightarrow \qquad \left.\frac{\partial F_0}{\partial \varphi}\right|_{\mathbf{r}} = 0, \quad (3.5)$$

where we have used that $\mathbf{v} \times \hat{\mathbf{b}} = -\partial \mathbf{v}/\partial\varphi$, making use of (2.22) and the properties of a right-handed orthonormal basis. This tells us that F_0 is independent of gyroangle at constant $\mathbf{r}$. We can use the chain rule to find

$$\left.\frac{\partial}{\partial \varphi}\right|_{\mathbf{r}} = \left.\frac{\partial}{\partial \varphi}\right|_{\mathbf{R}} + \left.\frac{\partial \mathbf{R}}{\partial \varphi}\right|_{\mathbf{r}} \cdot \left.\frac{\partial}{\partial \mathbf{R}}\right|_{\varphi} \quad (3.6)$$

where, using (2.22) and (2.29), we can calculate

$$\left.\frac{\partial \mathbf{R}}{\partial \varphi}\right|_{\mathbf{r}} = \frac{\partial}{\partial \varphi}\left.\left(\mathbf{r} - \frac{v_\perp}{\Omega_0}\hat{\mathbf{b}} \times (\cos\varphi\hat{\mathbf{e}}_1 + \sin\varphi\hat{\mathbf{e}}_2)\right)\right|_{\mathbf{r}} = \frac{1}{\Omega_0}\mathbf{v}_\perp \quad (3.7)$$

[1] Note that $\delta f_0 = 0$ (there is no fluctuation of the distribution function to lowest order).

and

$$\left.\frac{\partial}{\partial \mathbf{r}}\right|_{\varphi} = \frac{d\mathbf{R}}{d\mathbf{r}} \cdot \left.\frac{\partial}{\partial \mathbf{R}}\right|_{\varphi} = \left(\mathbf{I} - \nabla\left(\frac{\hat{\mathbf{b}}}{\Omega_0}\right) \times \mathbf{v}\right) \cdot \left.\frac{\partial}{\partial \mathbf{R}}\right|_{\varphi} = (1 + \mathcal{O}(\epsilon)) \left.\frac{\partial}{\partial \mathbf{R}}\right|_{\varphi}, \tag{3.8}$$

using the fact that $\hat{\mathbf{b}}/\Omega_0$ is an equilibrium quantity and so varies on the length scale L. (3.8) tells us that partial derivatives taken with respect to $\mathbf{R}$ are equivalent to partial derivatives taken with respect to $\mathbf{r}$ plus an $\mathcal{O}(\epsilon)$ correction, and so at any given order we can freely interchange the two.[2] We do so in (3.6), and obtain a useful relationship between the two angle derivatives:

$$\boxed{\left.\frac{\partial}{\partial \varphi}\right|_{\mathbf{r}} = \left.\frac{\partial}{\partial \varphi}\right|_{\mathbf{R}} + \frac{1}{\Omega_0}\mathbf{v}_\perp \cdot \nabla + \mathcal{O}(\epsilon).} \tag{3.9}$$

Applying (3.9) in (4.13), we can finally write

$$\boxed{\left.\frac{\partial F_0}{\partial \varphi}\right|_{\mathbf{R}} = 0} \tag{3.10}$$

since F_0 is an equilibrium quantity, meaning the $(1/\Omega_0)\mathbf{v}_\perp \cdot \nabla F_0$ term is $\mathcal{O}(\epsilon F_0)$. Thus, to leading order, $F_0(\mathbf{r}, \mathbf{v}, t)$ is also independent of gyrophase at fixed guiding centre: we say that F_0 is *gyrotropic* in $(\mathbf{r}, \mathbf{v})$ coordinates to leading order. We will use (4.13) in the following Section, where we stay in $(\mathbf{r}, \mathbf{v})$ coordinates.

$\mathcal{O}(\epsilon\Omega_0 F_0)$: F_0 is Maxwellian, n and T are flux functions, and δf_1 and F_1 can be decomposed

At $\mathcal{O}(\epsilon)$, our equation is

$$\begin{aligned}\mathbf{v} \cdot \nabla F_0 + \mathbf{v}_\perp \cdot \nabla \delta f_1 + \frac{Ze}{m}\left(-\nabla \delta\phi + \frac{\mathbf{v} \times \delta\mathbf{B}}{c}\right) \cdot \frac{\partial F_0}{\partial \mathbf{v}} \\ - \Omega_0 \left.\frac{\partial F_1}{\partial \varphi}\right|_{\mathbf{r}} - \Omega_0 \left.\frac{\partial \delta f_1}{\partial \varphi}\right|_{\mathbf{r}} = \sum_{s'} C[F_0, F_{0s'}],\end{aligned} \tag{3.11}$$

where we have used again that $\mathbf{v} \times \hat{\mathbf{b}} = -\partial \mathbf{v}/\partial \varphi$ to substitute for the two terms involving $\mathbf{B}_0$. Multiplying this equation by $(1 + \log F_0)$ and manipulating the

[2] We hereafter denote both $\partial/\partial \mathbf{r}$ and $\partial/\partial \mathbf{R}$ as ∇.

result, we find that

$$\mathbf{v}\cdot\nabla(F_0\log F_0) + (1+\log F_0)\mathbf{v}_\perp\cdot\nabla\delta f_1 + \frac{\partial}{\partial\mathbf{v}}\cdot\left[\frac{Ze}{m}\left(-\nabla\delta\phi + \frac{\mathbf{v}\times\delta\mathbf{B}}{c}\right)F_0\log F_0\right]$$
$$-(1+\log F_0)\left(\Omega_0\frac{\partial F_1}{\partial\varphi}\bigg|_{\mathbf{r}} + \Omega_0\frac{\partial\delta f_1}{\partial\varphi}\bigg|_{\mathbf{r}}\right) = (1+\log F_0)\sum_{s'}C[F_0,F_{0s'}]. \tag{3.12}$$

On integration in $d^3\mathbf{v}$, keeping $\mathbf{r}$ constant and using that F_0 goes to zero at infinity and all quantities are single-valued in φ, we dispose of the exact divergence in $\partial/\partial\mathbf{v}$ and the two derivatives in φ. Collisions between any two species conserve particle number,

$$\int d^3\mathbf{v}\sum_{s'}C[F_0,F_{0s'}] = 0, \tag{3.13}$$

and so we are left with

$$\int d^3\mathbf{v}\Big(\mathbf{v}\cdot\nabla(F_0\log F_0) + (1+\log F_0)\mathbf{v}_\perp\cdot\nabla\delta f_1\Big) = \int d^3\mathbf{v}\log F_0\sum_{s'}C[F_0,F_{0s'}]. \tag{3.14}$$

We immediately see that the $\mathbf{v}_\perp\cdot\nabla(F_0\log F_0)$ component of the first term integrates to zero, since F_0 is independent of φ, and $\int_0^{2\pi}\mathbf{v}_\perp d\varphi = 0$ using (2.22). Further, we may use (3.9) to rewrite the second term as

$$\int d^3\mathbf{v}(1+\log F_0)\mathbf{v}_\perp\cdot\nabla\delta f_1 = \int d^3\mathbf{v}(1+\log F_0)\Omega_0\left(\frac{\partial}{\partial\varphi}\bigg|_{\mathbf{r}} - \frac{\partial}{\partial\varphi}\bigg|_{\mathbf{R}}\right)\delta f_1 = 0, \tag{3.15}$$

where for the last equality we use again that F_0 is independent of φ, and δf_1 is independent of φ at fixed $\mathbf{r}$, and single-valued in φ at fixed $\mathbf{R}$. Hence, (3.14) becomes

$$\int d^3\mathbf{v}\; v_\parallel\hat{\mathbf{b}}\cdot\nabla(F_0\log F_0) = \int d^3\mathbf{v}\log F_0\sum_{s'}C[F_0,F_{0s'}]. \tag{3.16}$$

Applying the intermediate spatial average (ISA) (cf. Section 2.5), we have

$$\left\langle\int d^3\mathbf{v}\; v_\parallel\hat{\mathbf{b}}\cdot\nabla(F_0\log F_0)\right\rangle_\lambda = \left\langle\int d^3\mathbf{v}\log F_0\sum_{s'}C[F_0,F_{0s'}]\right\rangle_\lambda. \tag{3.17}$$

To make progress, we follow [4] and transform the integral on the left hand side into an integral in the gyrokinetic variables $(\mu,\varepsilon,\varphi,\sigma)$. The volume element in velocity space is $d^3\mathbf{v} = v_\perp dv_\parallel dv_\perp d\varphi = v_\perp J' d\varepsilon d\mu d\varphi$, where J' is the Jacobian of the transformation $(v_\parallel, v_\perp, \varphi) \to (\varepsilon,\mu,\varphi)$ given by

$$J' = \left|\det\begin{bmatrix}\partial v_\parallel/\partial\varepsilon & \partial v_\perp/\partial\varepsilon\\ \partial v_\parallel/\partial\mu & \partial v_\perp/\partial\mu\end{bmatrix}\right| = \frac{B_0}{m^2|v_\parallel|v_\perp}, \tag{3.18}$$

and so the left-hand side of (3.17) becomes

$$\begin{aligned}\left\langle \int d^3\mathbf{v}\ v_\parallel \hat{\mathbf{b}} \cdot \nabla(F_0 \log F_0) \right\rangle_\lambda &= \left\langle \sum_\sigma \int \frac{d\varepsilon d\mu d\varphi}{m^2|v_\parallel|}\ v_\parallel \mathbf{B}_0 \cdot \nabla(F_0 \log F_0) \right\rangle_\lambda \\ &= \left\langle \nabla \cdot \left(\sum_\sigma \int \frac{d\varepsilon d\mu d\varphi}{m^2|v_\parallel|}\ v_\parallel \mathbf{B}_0 (F_0 \log F_0) \right) \right\rangle_\lambda \\ &= \left\langle \nabla \cdot \left(\int d^3\mathbf{v}\ v_\parallel \hat{\mathbf{b}} (F_0 \log F_0) \right) \right\rangle_\lambda \\ &= \frac{1}{\Delta V} \int d^3\mathbf{v} \int_{\partial \Delta V} d\mathbf{S} \cdot (v_\parallel \hat{\mathbf{b}} F_0 \log F_0),\end{aligned} \tag{3.19}$$

where $\partial\Delta V$ is the surface which bounds the annulus volume ΔV between the flux surfaces, and $d\mathbf{S}$ is the area element pointing outwardly normal the bounding surface. Since we are integrating over flux surfaces, $\hat{\mathbf{b}} \cdot d\mathbf{S} = 0$ over the whole region of integration, meaning that the left-hand side of (3.17) disappears and we are left with

$$\left\langle \int d^3\mathbf{v} \log F_0 \sum_{s'} C[F_0, F_{0s'}] \right\rangle_\lambda = 0. \tag{3.20}$$

Recalling that the ISA was defined as an average over scales λ such that $\rho \ll \lambda \ll L$, and that F_0 only varies over the length scale L, the ISA leaves all of the quantities in the argument of (3.20) unchanged. We may therefore conclude that

$$\int d^3\mathbf{v} \log F_0 \sum_{s'} C[F_0, F_{0s'}] = 0, \tag{3.21}$$

and therefore, by Boltzmann's H-Theorem, we find that the lowest-order distribution function F_0 is a local Maxwellian:[3]

$$\boxed{F_0 = n_0(\mathbf{r},t)\left(\frac{m}{2\pi T_0(\mathbf{r},t)}\right)^{3/2} \exp\left(-\frac{mv^2}{2T_0(\mathbf{r},t)}\right) \quad \Rightarrow \quad \frac{\partial F_0}{\partial \mathbf{v}} = -\frac{m\mathbf{v}}{T_0(\mathbf{r},t)}F_0.} \tag{3.22}$$

We have included the spatial and temporal dependence in T_0 and n_0 here, noting that they are both equilibrium quantities. We also note that the H-Theorem enforces equal temperatures across all species[4], and therefore $\sum_{s'} C[F_0, F_{0s'}] = 0$. Using this in

[3] We omit a mean flow in our Maxwellian, since here we consider a subsonically rotating plasma, in which $\mathrm{Ma} \equiv u/v_T \ll 1$.

[4] However, we later allow for *ad hoc* collisional temperature equilibriation in the transport equations.

(3.11), along with (3.9) to substitute for the φ derivatives of F_1 and δf_1 at constant $\mathbf{r}$ (noting that the $\mathbf{v}_\perp \cdot \nabla$ term is an order higher when acting on F_1), we obtain

$$\mathbf{v} \cdot \nabla F_0 + \frac{ZeF_0}{T_0(\mathbf{r},t)} \mathbf{v}_\perp \cdot \nabla \delta\phi = \Omega_0 \left.\frac{\partial F_1}{\partial \varphi}\right|_{\mathbf{R}} + \Omega_0 \left.\frac{\partial \delta f_1}{\partial \varphi}\right|_{\mathbf{R}}, \tag{3.23}$$

where we have also used that the $v_\parallel \hat{\mathbf{b}} \cdot \nabla \delta\phi$ term is an order smaller. Using (3.9), we can rewrite

$$\mathbf{v}_\perp \cdot \nabla F_0(\mathbf{r}) = -\Omega_0 \frac{\partial}{\partial \varphi}\Big(F_0(\mathbf{r})\Big)\bigg|_{\mathbf{R}} = -\Omega_0 \frac{\partial}{\partial \varphi}\Big(\cancelto{0}{F_0(\mathbf{R})} + \boldsymbol{\rho} \cdot \nabla F_0(\mathbf{r})\Big)\bigg|_{\mathbf{R}}, \tag{3.24}$$

where we have Taylor expanded in the gyroradius. Using this, and noting that $\delta\phi$ is independent of φ at fixed $\mathbf{r}$, and F_0 and T are gyrotropic to leading order and so can be brought inside $\partial/\partial\varphi|_{\mathbf{R}}$, (3.23) becomes

$$v_\parallel \hat{\mathbf{b}} \cdot \nabla F_0 - \Omega_0 \frac{\partial}{\partial \varphi}\Big(\boldsymbol{\rho} \cdot \nabla F_0\Big)\bigg|_{\mathbf{R}} - \Omega_0 \frac{\partial}{\partial \varphi}\left(\frac{Ze\delta\phi}{T_0} F_0\right)\bigg|_{\mathbf{R}} = \Omega_0 \left.\frac{\partial F_1}{\partial \varphi}\right|_{\mathbf{R}} + \Omega_0 \left.\frac{\partial \delta f_1}{\partial \varphi}\right|_{\mathbf{R}}. \tag{3.25}$$

Gyroaveraging this equation (cf. (2.39)), and using that all functions of φ are single-valued in φ at fixed $\mathbf{R}$, we find that

$$\boxed{\hat{\mathbf{b}} \cdot \nabla F_0 = 0,} \tag{3.26}$$

which says that to leading order, our distribution function does not vary along magnetic field lines. Since the flux surfaces labelled by ψ also satisfy $\hat{\mathbf{b}} \cdot \nabla \psi = 0$ by definition, we conclude that $F_0 = F_0(\psi, v, t)$, i.e. F_0 is a *flux function* in position space. Inserting the Maxwellian form of F_0 from (3.22) into (3.26) and dividing through[5] by F_0, we find

$$\hat{\mathbf{b}} \cdot \left(\frac{1}{n_0}\nabla n_0 - \frac{3}{2}\frac{1}{T_0}\nabla T_0 + \frac{mv^2}{2{T_0}^2}\nabla T_0\right) = 0 \quad \forall \ \mathbf{v}. \tag{3.27}$$

Since (3.27) holds for all $\mathbf{v}$, and n_0 and T_0 are independent of velocity, we conclude $\hat{\mathbf{b}} \cdot \nabla n_0 = 0$ and $\hat{\mathbf{b}} \cdot \nabla T_0 = 0$ separately, i.e. density and temperature are also flux functions:

$$\boxed{n_0 = n_0(\psi, t), \qquad T_0 = T_0(\psi, t).} \tag{3.28}$$

[5]This is allowed because F_0 is Maxwellian, and therefore non-zero everywhere except at infinity.

We now substitute (3.26) into (3.25), and average over perturbations, to find that F_1 obeys the equation

$$\Omega_0 \frac{\partial F_1}{\partial \varphi}\bigg|_{\mathbf{R}} = -\Omega_0 \frac{\partial}{\partial \varphi}\Big(\boldsymbol{\rho}\cdot\nabla F_0(\mathbf{r})\Big)\bigg|_{\mathbf{R}}. \tag{3.29}$$

This clearly has the solution

$$\boxed{F_1 = -\boldsymbol{\rho}\cdot\nabla F_0 + \widetilde{F_1}(\mathbf{R},\varepsilon,\mu,\sigma,t),} \tag{3.30}$$

where the constant of gyrophase integration $\widetilde{F_1}$ is the gyrotropic part of F_1. We have used gyrokinetic variables in $\widetilde{F_1}$, in order to illustrate the lack of φ dependence. The fact it was constructed to be independent of φ at constant $\mathbf{R}$ also implies that it is a function of $\mathbf{R}$ only in position space. Subtracting (3.29) from (3.25), we are left with an equation featuring only fluctuating terms:

$$\frac{\partial \delta f_1}{\partial \varphi}\bigg|_{\mathbf{R}} = -\frac{\partial}{\partial \varphi}\left(\frac{Ze\delta\phi}{T_0}F_0\right)\bigg|_{\mathbf{R}}, \tag{3.31}$$

and so we can clearly see that δf_1 has the solution

$$\boxed{\delta f_1 = -\frac{Ze\delta\phi}{T_0}F_0 + h(\mathbf{R},\varepsilon,\mu,\sigma,t),} \tag{3.32}$$

where the constant of gyrophase integration h is the gyrotropic part of δf_1. Since it is also a function of $\mathbf{R}$ only in position space, we conclude that it must correspond to a distribution of charged rings. The first term reveals that, as might be expected, there is a Boltzmann correction to the Maxwellian equilibrium due to the electrostatic part of $\delta\mathbf{E}$:

$$F_0 \to F_0 \exp\left(-\frac{Ze\delta\phi}{T_0}\right). \tag{3.33}$$

A summary of the results we have derived in this Section can be found below:

$$f = F_0 + \widetilde{F_1} - \boldsymbol{\rho}\cdot\nabla F_0 - \frac{Ze\delta\phi}{T_0}F_0 + h + F_2 + \delta f_2 + \mathcal{O}(\epsilon^3 F_0), \tag{3.34}$$

$$\frac{\partial F_0}{\partial \varphi}\bigg|_{\mathbf{R}} = \frac{\partial \widetilde{F_1}}{\partial \varphi}\bigg|_{\mathbf{R}} = \frac{\partial h}{\partial \varphi}\bigg|_{\mathbf{R}} = 0, \tag{3.35}$$

$$F_0 = n_0(\psi,t)\left(\frac{m}{2\pi T_0(\psi,t)}\right)^{3/2}\exp\left(-\frac{mv^2}{2T_0(\psi,t)}\right). \tag{3.36}$$

$\mathcal{O}(\epsilon^2\Omega_0 F_0)$: Gyrokinetic equation and neoclassical theory

At this order, it is convenient to transform to the gyrokinetic variables derived in Section 2.3. The gyroaveraged Fokker-Planck equation in terms of these new variables is

$$\left\langle \frac{df}{dt} \right\rangle_{\mathbf{R}} = \left\langle \frac{\partial f}{\partial t} + \frac{d\mathbf{R}}{dt}\cdot\frac{\partial f}{\partial \mathbf{R}} + \frac{d\mu}{dt}\frac{\partial f}{\partial \mu} + \frac{d\varepsilon}{dt}\frac{\partial f}{\partial \varepsilon} + \frac{d\varphi}{dt}\frac{\partial f}{\partial \varphi} \right\rangle_{\mathbf{R}} = \left\langle \sum_{s'} C[f, f_{s'}] \right\rangle_{\mathbf{R}}. \quad (3.37)$$

We will absorb terms into an equilibrium distribution as a function of the guiding centre variable $\mathbf{R}$, by recognising the Taylor expansion of $F_0(\mathbf{R})$ as follows:

$$F_0(\mathbf{r}) - \boldsymbol{\rho}\cdot\nabla F_0 = F_0(\mathbf{R}) + \mathcal{O}(\epsilon^2 F_0), \quad (3.38)$$

where the correction can be absorbed into f_2. We reiterate that F_0 being a function of the gyrokinetic variables $(\mathbf{R}, \varepsilon, t)$ means that it is unaffected by the gyroaverage, which is taken at constant $\mathbf{R}$. The fluctuating potentials, $\delta\phi$ and $\delta\mathbf{A}$, remain functions of $(\mathbf{r}, \mathbf{v}, t)$.[6] For clarity, we will separately consider each term in the expansion of $\langle df/dt\rangle_{\mathbf{R}}$:

$$\left\langle \frac{df}{dt} \right\rangle_{\mathbf{R}} = \left\langle \frac{dF_0}{dt} + \frac{d\widetilde{F_1}}{dt} - \frac{d}{dt}\left(\frac{Ze\delta\phi}{T_0}F_0\right) + \frac{dh}{dt} + \frac{d}{dt}\Big(F_2 + \delta f_2\Big) \right\rangle_{\mathbf{R}} + \mathcal{O}(\epsilon^3\Omega_0 F_0). \quad (3.39)$$

Since we balanced terms at lower orders in the previous Sections, we will be left with only the $\mathcal{O}(\epsilon^2\Omega_0 F_0)$ terms on the left-hand side of (3.37). The collision terms on the right-hand side can also be expanded to find that they contribute at this order. By averaging over perturbations, we will discover the evolution equations for h and $\widetilde{F_1}$.

Beginning with F_0, we see that its full form in the gyrokinetic variables is

$$F_0(\psi(\mathbf{R}), \varepsilon, t) = n(\psi(\mathbf{R}), t)\left(\frac{m}{2\pi T_0(\psi(\mathbf{R}), t)}\right)^{3/2} \exp\left(-\frac{\varepsilon}{T_0(\psi(\mathbf{R}), t)}\right), \quad (3.40)$$

and so using (3.26), (2.25) for $d\varepsilon/dt$, and (2.32) for $d\mathbf{R}/dt$, we find that

$$\frac{dF_0}{dt} = (\mathbf{v}_M + \mathbf{v}_E)\cdot\nabla F_0 + \frac{ZeF_0}{T_0}\left(\mathbf{v}\cdot\nabla\delta\phi + \frac{\mathbf{v}}{c}\cdot\frac{\partial\delta\mathbf{A}}{\partial t}\right) + \mathcal{O}(\epsilon^3\Omega_0 F_0). \quad (3.41)$$

Gyroaveraging this equation, and using that $\langle F_0(\mathbf{R}, \varepsilon, t)\rangle_{\mathbf{R}} = F_0$, we see that

$$\left\langle \frac{dF_0}{dt} \right\rangle_{\mathbf{R}} = \Big(\langle\mathbf{v}_M\rangle_{\mathbf{R}} + \langle\mathbf{v}_E\rangle_{\mathbf{R}}\Big)\cdot\nabla F_0 + \frac{ZeF_0}{T_0}\left\langle v_\parallel\hat{\mathbf{b}}\cdot\nabla\delta\phi + \frac{\mathbf{v}}{c}\cdot\frac{\partial\delta\mathbf{A}}{\partial t}\right\rangle_{\mathbf{R}} + \mathcal{O}(\epsilon^3\Omega_0 F_0), \quad (3.42)$$

[6]They obviously cannot be written as a Taylor expansion in the gyroradius, since they vary on this scale, and so must be kept functions of the *exact* position.

where we have used (3.9) to eliminate $\mathbf{v}_\perp \cdot \nabla\delta\phi$ as usual.

Calculating $\langle \mathbf{v}_E \rangle_\mathbf{R}$ using (2.31), we find that

$$\begin{aligned} \langle \mathbf{v}_E \rangle_\mathbf{R} &= \frac{c}{B}\left\langle \hat{\mathbf{b}} \times \left(\nabla\delta\phi + \frac{1}{c}\frac{\partial \delta\mathbf{A}}{\partial t} - \nabla\left(\frac{\mathbf{v}\cdot\delta\mathbf{A}}{c}\right) + \frac{1}{c}(\mathbf{v}\cdot\nabla)\delta\mathbf{A}\right)\right\rangle_\mathbf{R} \\ &= \frac{c}{B}\hat{\mathbf{b}} \times \langle\nabla\chi\rangle_\mathbf{R} + \mathcal{O}(\epsilon^2 v_T), \end{aligned} \tag{3.43}$$

defining $\chi \equiv \delta\phi - \mathbf{v}\cdot\delta\mathbf{A}/c$ as the *gyrokinetic potential.* We have also used (3.9) to eliminate $\langle \mathbf{v}_\perp \cdot \nabla\delta\mathbf{A}\rangle_\mathbf{R}$, and that the $v_\parallel \hat{\mathbf{b}}\cdot\nabla\delta\mathbf{A}$ term is higher-order. We define the generalised $E \times B$ velocity to be

$$\mathbf{v}_\chi \equiv \frac{c}{B_0}\hat{\mathbf{b}} \times \langle\nabla\chi\rangle_\mathbf{R}. \tag{3.44}$$

Calculating $\langle \mathbf{v}_B \rangle_\mathbf{R}$ using (2.31), we find that

$$\langle \mathbf{v}_B \rangle_\mathbf{R} = v_\parallel^2 \hat{\mathbf{b}} \times \left(\hat{\mathbf{b}}\cdot\nabla\left(\frac{\hat{\mathbf{b}}}{\Omega_0}\right)\right) + \frac{v_\perp^2}{2}\hat{\mathbf{e}}_1 \times \left(\hat{\mathbf{e}}_1\cdot\nabla\left(\frac{\hat{\mathbf{b}}}{\Omega_0}\right)\right) + \frac{v_\perp^2}{2}\hat{\mathbf{e}}_2 \times \left(\hat{\mathbf{e}}_2\cdot\nabla\left(\frac{\hat{\mathbf{b}}}{\Omega_0}\right)\right) \tag{3.45}$$

using (2.22), and noting that all cross terms involving one $\mathbf{v}_\perp$ will gyroaverage to zero, all cross terms within $\mathbf{v}_\perp\mathbf{v}_\perp$ involving just one sine or cosine will also gyroaverage to zero, and squares of sines and cosines each gyroaverage to 1/2. Using simple vector calculus and rewriting $\hat{\mathbf{e}}_1\hat{\mathbf{e}}_1 + \hat{\mathbf{e}}_2\hat{\mathbf{e}}_2 = \mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}$, this can be written in the more physically illuminating form

$$\langle \mathbf{v}_B \rangle_\mathbf{R} = \frac{v_\parallel^2}{\Omega_0}\hat{\mathbf{b}} \times (\hat{\mathbf{b}}\cdot\nabla\hat{\mathbf{b}}) + \frac{v_\perp^2}{2\Omega_0}\hat{\mathbf{b}}\cdot(\nabla\times\hat{\mathbf{b}})\hat{\mathbf{b}} + \frac{v_\perp^2}{2\Omega_0^2}\hat{\mathbf{b}}\times\nabla\Omega_0, \tag{3.46}$$

in which we can identify the well-known Baños, ∇B and curvature drifts

$$\mathbf{v}_{Ba} \equiv \frac{v_\perp^2}{2\Omega_0}\hat{\mathbf{b}}\cdot(\nabla\times\hat{\mathbf{b}})\hat{\mathbf{b}}, \tag{3.47}$$

$$\mathbf{v}_{\nabla B} \equiv \frac{v_\perp^2}{2\Omega_0^2}\hat{\mathbf{b}}\times\nabla\Omega_0, \tag{3.48}$$

$$\mathbf{v}_\kappa \equiv \frac{v_\parallel^2}{\Omega_0}\hat{\mathbf{b}}\times(\hat{\mathbf{b}}\cdot\nabla\hat{\mathbf{b}}) \tag{3.49}$$

respectively. Our final equation for $\langle dF_0/dt\rangle_\mathbf{R}$ is then

$$\boxed{\begin{aligned} \left\langle\frac{dF_0}{dt}\right\rangle_\mathbf{R} &= (\mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi)\cdot\nabla F_0 \\ &\quad + \frac{ZeF_0}{T_0}\left\langle\left(v_\parallel\hat{\mathbf{b}}\cdot\nabla\delta\phi + \frac{\mathbf{v}}{c}\cdot\frac{\partial\delta\mathbf{A}}{\partial t}\right)\right\rangle_\mathbf{R} + \mathcal{O}(\epsilon^3\Omega_0F_0), \end{aligned}} \tag{3.50}$$

where we have used that $\mathbf{v}_{Ba} \cdot \nabla F_0 \propto \hat{\mathbf{b}} \cdot \nabla F_0 = 0$.

On to $\widetilde{F_1}$, and using (2.32) for $d\mathbf{R}/dt$ and that $\widetilde{F_1}$ is gyrotropic, we find

$$\frac{d\widetilde{F_1}}{dt} = v_\parallel \hat{\mathbf{b}} \cdot \nabla \widetilde{F_1} + \frac{d\mu}{dt}\frac{\partial \widetilde{F_1}}{\partial \mu} + \frac{d\varepsilon}{dt}\frac{\partial \widetilde{F_1}}{\partial \varepsilon} + \mathcal{O}(\epsilon^3 \Omega_0 F_0). \tag{3.51}$$

Again performing a gyroaverage,

$$\left\langle \frac{d\widetilde{F_1}}{dt} \right\rangle_{\mathbf{R}} = v_\parallel \hat{\mathbf{b}} \cdot \nabla \widetilde{F_1} + \left\langle \frac{d\mu}{dt} \right\rangle_{\mathbf{R}} \frac{\partial \widetilde{F_1}}{\partial \mu} + \left\langle \frac{d\varepsilon}{dt} \right\rangle_{\mathbf{R}} \frac{\partial \widetilde{F_1}}{\partial \varepsilon} + \mathcal{O}(\epsilon^3 \Omega_0 F_0). \tag{3.52}$$

Using (2.24) for $d\mu/dt$, we calculate its gyroaverage and find[7]

$$\left\langle \frac{d\mu}{dt} \right\rangle_{\mathbf{R}} = -\frac{\mu}{B_0} v_\parallel \hat{\mathbf{b}} \cdot \nabla B_0 + \frac{Ze}{B_0} \left\langle \mathbf{v}_\perp \cdot \left(\delta \mathbf{E} + \frac{\mathbf{v} \times \delta \mathbf{B}}{c} \right) \right\rangle_{\mathbf{R}} - \mu v_\parallel (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) : \nabla \hat{\mathbf{b}}, \tag{3.53}$$

where we have decomposed $\mathbf{v}$ using (2.22), and used that $\langle \hat{\mathbf{v}}_\perp \hat{\mathbf{v}}_\perp \rangle_{\mathbf{R}} = (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}})/2$ and $\langle \hat{\mathbf{v}}_\perp \rangle_{\mathbf{R}} = 0$. Rewriting the second term in terms of potentials using (2.14), and replacing $\mathbf{v} \times \delta \mathbf{B} \to v_\parallel \hat{\mathbf{b}} \times \delta \mathbf{B}$ since it appears in a triple product with $\mathbf{v}_\perp$, we find that it becomes

$$\frac{Ze}{B_0} \left\langle \mathbf{v}_\perp \cdot \left(-\nabla \delta\phi - \frac{1}{c}\frac{\partial \delta \mathbf{A}}{\partial t} + \frac{1}{c}\Big(\nabla (v_\parallel \hat{\mathbf{b}} \cdot \delta \mathbf{A}) - v_\parallel \hat{\mathbf{b}} \cdot \nabla \delta \mathbf{A}\Big) \right) \right\rangle_{\mathbf{R}} = -\frac{Ze}{B_0} \left\langle \frac{\mathbf{v}_\perp}{c} \cdot \frac{\partial \delta \mathbf{A}}{\partial t} \right\rangle_{\mathbf{R}} \tag{3.54}$$

using (3.9) to eliminate both $\langle \mathbf{v}_\perp \cdot \nabla (\dots) \rangle_{\mathbf{R}}$ terms, and noting that $v_\parallel \hat{\mathbf{b}} \cdot \nabla \delta \mathbf{A}$ is higher order. We can also simplify the third term as follows:

$$-\mu v_\parallel (\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) : \nabla \hat{\mathbf{b}} = -\mu v_\parallel \Big(\partial_j b_j - \frac{1}{2} \cancelto{0}{b_i \partial_i (b_j b_j)} \Big) = -\mu v_\parallel \nabla \cdot \hat{\mathbf{b}}. \tag{3.55}$$

Substituting these into (3.53), we find that

$$\left\langle \frac{d\mu}{dt} \right\rangle_{\mathbf{R}} = -\mu v_\parallel \cancelto{0}{\frac{\nabla \cdot \mathbf{B}_0}{B_0}} - \frac{Ze}{B_0} \left\langle \frac{\mathbf{v}_\perp}{c} \cdot \frac{\partial \delta \mathbf{A}}{\partial t} \right\rangle_{\mathbf{R}} = \mathcal{O}(\epsilon^2 \Omega_0 \mu), \tag{3.56}$$

where we have used that $\nabla \cdot \hat{\mathbf{b}} + (1/B_0) \hat{\mathbf{b}} \cdot \nabla B_0 = (1/B_0) \nabla \cdot \mathbf{B}_0$. Using (2.25) for $d\varepsilon/dt$, we calculate its gyroaverage in a similar fashion:

$$\left\langle \frac{d\varepsilon}{dt} \right\rangle_{\mathbf{R}} = -Ze \left\langle v_\parallel \hat{\mathbf{b}} \cdot \nabla \delta\phi + \frac{\mathbf{v}}{c} \cdot \frac{\partial \delta \mathbf{A}}{\partial t} \right\rangle_{\mathbf{R}} = \mathcal{O}(\epsilon^2 \Omega_0 \varepsilon). \tag{3.57}$$

[7]The double contraction operation is defined as the trace of the matrix obtained by standard matrix multiplication of the arguments, $\underline{\underline{\mathbf{A}}} : \underline{\underline{\mathbf{B}}} \equiv \sum_{i,j=1}^{3} A_{ji} B_{ij}$.

Inserting these into (3.52), the second and third terms are moved to $\mathcal{O}(\epsilon^3\Omega_0 F_0)$, and so

$$\boxed{\left\langle \frac{d\widetilde{F_1}}{dt} \right\rangle_{\mathbf{R}} = v_\parallel \hat{\mathbf{b}} \cdot \nabla \widetilde{F_1} + \mathcal{O}(\epsilon^3\Omega_0 F_0).} \tag{3.58}$$

The gyroaveraged derivative of the Boltzmann term is

$$\left\langle \frac{d}{dt}\left(\frac{-Ze\delta\phi}{T_0}F_0\right)\right\rangle_{\mathbf{R}} = -\left\langle \frac{Ze\delta\phi}{T_0}\frac{dF_0}{dt} + \frac{ZeF_0}{T_0}\frac{d\delta\phi}{dt} - \frac{Ze\delta\phi F_0}{T_0^2}\frac{dT_0}{dt}\right\rangle_{\mathbf{R}}. \tag{3.59}$$

Using (3.41) to substitute for dF_0/dt and ordering terms, we see that the first term is

$$-\left\langle \frac{Ze\delta\phi}{T_0}\frac{dF_0}{dt}\right\rangle_{\mathbf{R}} = -\left(\frac{Ze}{T_0}\right)^2 F_0 \left\langle \frac{1}{2}\mathbf{v}_\perp \cdot \nabla(\delta\phi)^2 \right\rangle_{\mathbf{R}} + \mathcal{O}(\epsilon^3\Omega_0 F_0) = \mathcal{O}(\epsilon^3\Omega_0 F_0), \tag{3.60}$$

where we have used that $\langle \delta\phi \mathbf{v}_\perp \cdot \nabla\delta\phi\rangle_{\mathbf{R}} = \langle (1/2)\mathbf{v}_\perp \cdot \nabla(\delta\phi)^2\rangle_{\mathbf{R}}$, which we may then eliminate using (3.9) as usual. Therefore, we may neglect this term. Moving to the second term in (3.59), d/dt becomes the Vlasov operator (2.21) since $\delta\phi = \delta\phi(\mathbf{r},t)$, and so

$$-\left\langle \frac{ZeF_0}{T_0}\frac{d\delta\phi}{dt}\right\rangle_{\mathbf{R}} = -\frac{ZeF_0}{T_0}\left\langle \frac{\partial\delta\phi}{\partial t} + v_\parallel \hat{\mathbf{b}}\cdot\nabla\delta\phi \right\rangle_{\mathbf{R}} + \mathcal{O}(\epsilon^3\Omega_0 F_0), \tag{3.61}$$

again using (3.9) to eliminate $\langle \mathbf{v}_\perp \cdot \nabla\delta\phi\rangle_{\mathbf{R}}$. For the third and final term in (3.59), we can replace $\mathbf{R} \to \mathbf{r}$ in T_0[8] and again use the Vlasov operator to find

$$\begin{aligned}\left\langle \frac{Ze\delta\phi F_0}{T_0^2}\frac{dT_0}{dt}\right\rangle_{\mathbf{R}} &= \frac{ZeF_0}{T_0^2}\left\langle \bigg(\mathbf{v}_\perp \cdot \nabla(\delta\phi T_0) - T_0 \mathbf{v}_\perp \cdot \nabla\delta\phi \bigg)\right\rangle_{\mathbf{R}} + \mathcal{O}(\epsilon^3\Omega_0 F_0) \\ &= \mathcal{O}(\epsilon^3\Omega_0 F_0),\end{aligned} \tag{3.62}$$

where we have written $\langle \delta\phi\mathbf{v}_\perp \cdot \nabla T_0\rangle_{\mathbf{R}} = \langle \mathbf{v}_\perp \cdot \nabla(\delta\phi T_0) - T_0\mathbf{v}_\perp \cdot \nabla\delta\phi\rangle_{\mathbf{R}}$ and again used (3.9) to eliminate each term separately to leading order. We have also used that T_0 varies on the equilibrium timescale to say that $\partial T_0/\partial t \sim \epsilon^3\Omega T_0$, and that T_0 is a flux function to cancel the $v_\parallel \hat{\mathbf{b}} \cdot \nabla T_0$ term. Therefore, we may also neglect this term in (3.59) to $\mathcal{O}(\epsilon^2\Omega_0 F_0)$. Consequently, the total Boltzmann term becomes

$$\boxed{\left\langle \frac{d}{dt}\left(\frac{-Ze\delta\phi}{T_0}F_0\right)\right\rangle_{\mathbf{R}} = -\frac{ZeF_0}{T_0}\left\langle \frac{\partial\delta\phi}{\partial t} + v_\parallel \hat{\mathbf{b}}\cdot\nabla\delta\phi \right\rangle_{\mathbf{R}} + \mathcal{O}(\epsilon^3\Omega_0 F_0).} \tag{3.63}$$

For dh/dt, we can use (3.56) and (3.57) to readily find that

$$\boxed{\left\langle \frac{dh}{dt}\right\rangle_{\mathbf{R}} = \frac{\partial h}{\partial t} + \left(v_\parallel \hat{\mathbf{b}} + \mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi\right)\cdot\nabla h + \mathcal{O}(\epsilon^3\Omega_0 F_0),} \tag{3.64}$$

[8] Since T_0 is an equilibrium quantity, the correction term in the Taylor expansion is higher-order.

where we have dropped the parallel Baños drift $\mathbf{v}_{Ba}$ since $\hat{\mathbf{b}} \cdot \nabla h \sim \epsilon F_0/L$, making this term $\mathcal{O}(\epsilon^3 \Omega_0 F_0)$.

Finally, we turn to δf_2 and F_2. Since these terms are already $\mathcal{O}(\epsilon^2 F_0)$, we need only the $\mathcal{O}(\Omega_0)$ part of the Vlasov operator to $\mathcal{O}(\epsilon^2 \Omega_0 F_0)$, namely $-\Omega_0 \partial/\partial\varphi$. Hence,

$$\boxed{\left\langle \frac{d(F_2+\delta f_2)}{dt} \right\rangle_{\mathbf{R}} = \left\langle -\Omega_0 \frac{\partial(F_2+\delta f_2)}{\partial\varphi}\bigg|_{\mathbf{R}} + \mathcal{O}(\epsilon^3 \Omega_0 F_0) \right\rangle_{\mathbf{R}} = \mathcal{O}(\epsilon^3 \Omega_0 F_0),} \tag{3.65}$$

since δf_2 and F_2 are both single-valued in φ at fixed $\mathbf{R}$. Consequently, under the gyrokinetic approximation, we do not need to know anything about the second-order parts of the distribution function in order to see how the first-order parts evolve.

We are now in a position to bring together all of the terms in the $\mathcal{O}(\epsilon^2 \Omega_0 F_0)$ Fokker-Planck equation. Using (3.50), (3.58), (3.64), (3.63) and (3.65), we find that

$$\begin{aligned}\left\langle \frac{df}{dt} \right\rangle_{\mathbf{R}} = \frac{\partial h}{\partial t} + (v_\parallel \hat{\mathbf{b}} + \mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi) \cdot \nabla h - \frac{ZeF_0}{T_0} \frac{\partial \langle \chi \rangle_{\mathbf{R}}}{\partial t} \\ + (\mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi) \cdot \nabla F_0 + v_\parallel \hat{\mathbf{b}} \cdot \nabla \widetilde{F_1} + \mathcal{O}(\epsilon^3 \Omega_0 F_0),\end{aligned} \tag{3.66}$$

noting that the two $(ZeF_0/T)\langle v_\parallel \hat{\mathbf{b}} \cdot \nabla \delta\phi \rangle_{\mathbf{R}}$ terms have exactly cancelled. We have also substituted in the gyrokinetic potential $\chi \equiv \delta\phi - \mathbf{v} \cdot \delta\mathbf{A}/c$.

We now consider the collision terms. Recalling that $\delta\phi$ and T_0 are independent of velocity and so can be brought outside of the collision operator, and that a Maxwellian F_0 enforces $\sum_{s'} C[F_0, F_{0'_s}] = 0$, we see that the only the h and $\widetilde{F_1}$ terms survive inside the $\mathcal{O}(\epsilon^2 \Omega_0 F_0)$ collision operator. Hence, we are finally left with an equation for the gyroaveraged $\mathcal{O}(\epsilon^2 \Omega_0 F_0)$ part of the Fokker-Planck equation:

$$\boxed{\begin{aligned}\frac{\partial h}{\partial t} + (v_\parallel \hat{\mathbf{b}} + \mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi) \cdot \nabla h - \frac{ZeF_0}{T_0} \frac{\partial \langle \chi \rangle_{\mathbf{R}}}{\partial t} + (\mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi) \cdot \nabla F_0 \\ + v_\parallel \hat{\mathbf{b}} \cdot \nabla \widetilde{F_1} = \left\langle \sum_{s'} \Big(C[F_0, (h+\widetilde{F_1})_{s'}] + C[(h+\widetilde{F_1}), F_{0s'}] \Big) \right\rangle_{\mathbf{R}}.\end{aligned}} \tag{3.67}$$

We proceed to average this equation over intermediate spatial scales (cf. Section 2.5), under which all the terms linear in fluctuations vanish. The quadratic term $\mathbf{v}_\chi \cdot \nabla h$ also vanishes:

$$\langle \mathbf{v}_\chi \cdot \nabla h \rangle_\lambda = \langle \nabla \cdot (\mathbf{v}_\chi h) \rangle_\lambda = \frac{1}{\Delta V} \int_{\partial \Delta V} d\mathbf{S} \cdot \mathbf{v}_\chi h = 0, \tag{3.68}$$

where we have used that $(\nabla \cdot \mathbf{v}_\chi)h = \mathcal{O}(\epsilon^3 \Omega_0 F_0)$ for the first equality, and *statistical periodicity* of the turbulence in the final equality.[9] The terms which survive the ISA determine the *neoclassical* part of the first-order distribution function, $\widetilde{F_1}$:

$$\boxed{(\mathbf{v}_{\nabla B} + \mathbf{v}_\kappa) \cdot \nabla F_0 + v_\parallel \hat{\mathbf{b}} \cdot \nabla \widetilde{F_1} = C_{nc}^{(l)}[\widetilde{F_1}],} \tag{3.69}$$

where $C_{nc}^{(l)}[\widetilde{F_1}] \equiv \sum_{s'}(C[F_0, \widetilde{F_{1s'}}] + C[(\widetilde{F_1}, F_{0s'}])$ is the gyrotropic part of the linearised neoclassical collision operator. This equation determines how the distribution function is affected by the magnetic geometry, and would be the only correction to the classical Maxwellian distribution function if we did not consider turbulent fluctuations.

Subtracting the neoclassical equation from (3.67), we are left with an equation featuring only fluctuating terms, known as the *gyrokinetic equation*:

$$\boxed{\frac{\partial h}{\partial t} + (v_\parallel \hat{\mathbf{b}} + \mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi) \cdot \nabla h - \frac{ZeF_0}{T_0} \frac{\partial \langle \chi \rangle_\mathbf{R}}{\partial t} + \mathbf{v}_\chi \cdot \nabla F_0 = \left\langle C_t^{(l)}[h] \right\rangle_\mathbf{R},} \tag{3.70}$$

where $C_t^{(l)}[h] \equiv \sum_{s'}(C[F_0, h_{s'}] + C[h, F_{0s'}])$ is the linearised turbulent collision operator. This is an evolution equation for h, the turbulent distribution function. We reiterate that this equation is in terms of gyrokinetic variables $(\mathbf{R}, \varepsilon, \mu, \varphi, \sigma)$, with $v_\parallel = \sigma |v_\parallel|$, and

$$|v_\parallel| = \sqrt{\frac{2(\varepsilon - \mu B)}{m}}, \qquad \sigma = \frac{v_\parallel}{|v_\parallel|} = \pm 1. \tag{3.71}$$

Following [2], the evolution of the turbulent distribution function h chacterised by (3.70) can be split into three distinct categories. First, we have the source terms[10] $\partial \langle \chi \rangle_\mathbf{R} / \partial t$ and $\mathbf{v}_\chi \cdot \nabla F_0$, which tell us that turbulence is driven by both fluctuations in the potentials, and macroscopic gradients. These terms typically drive large-scale structure in velocity space. There are also the advection terms, $(v_\parallel \hat{\mathbf{b}} + \mathbf{v}_{\nabla B} + \mathbf{v}_\kappa + \mathbf{v}_\chi) \cdot \nabla h$, which lead to phase-mixing and the development of small-scale structure in velocity space. The nonlinearity in the final term is responsible for the coupling of different modes, leading to chaotic behaviour within the turbulence. Finally, we have the collision terms $\langle \sum_{s'}(C[F_0, h_{s'}] + C[h, F_{0s'}]) \rangle_\mathbf{R}$, which lead to dissipation and drive the system towards a Maxwellian velocity distribution.

[9] This says that because the turbulent correlation length is much smaller than λ, the two boundaries are decorrelated and hence the fluctuating quantities are statistically the same on each boundary.

[10] These are not source terms in the 'external' sense, since the potentials are themselves dependent on h.

The proportion of ingredients is important, but the final result is also a matter of how you put them together. Equilibrium is key.

— Alain Ducasse, Michelin-star chef

4 Transport Equations

4.1 Introduction

In this Chapter, we present the evolution equations for the macroscopic quantities within a plasma, namely the equilibrium density n_0, and pressure p_0. These equations dictate how particles and kinetic energy are distributed throughout the plasma, and how they evolve on macroscopic timescales. In terrestrial fusion experiments, a high pressure and particle number must be maintained over the confinement time at the centre of the device to achieve a high fusion rate. The equations we develop here will describe the transport of these quantities away from the magnetic axis, and are consequently a vital tool for informing the design of present and future fusion devices. The transport equations naturally require taking moments of the Fokker-Planck equation correct to $\mathcal{O}(\epsilon^3\Omega_0F_0)$, since n_0 and p_0 are defined as moments of the equilibrium distribution function F_0, which evolves on the timescale $t_{eq}^{-1} \sim \epsilon^3\Omega_0$.

The rest of this Chapter is organised as follows. First, we will start again with the Fokker-Planck equation in $(\mathbf{r}, \mathbf{v})$ coordinates. To derive the evolution equation for n_0, we proceed to take the density moment. We will then average the resulting equation over intermediate scales and over a flux surface, in order to obtain the density transport equation. The p_0 derivation is omitted for lack of space, and is instead quoted from [3].

4.2 Evolution of the Density Profile

To begin our calculation of the density profile evolution, we return to (3.1), the exact Fokker-Planck equation in $(\mathbf{r}, \mathbf{v})$ variables. Since $\mathbf{r}$ and $\mathbf{v}$ are independent variables, we can write the acceleration term as an exact divergence in velocity space, and the velocity term as a divergence in position space,

$$\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{v} f) + \frac{Ze}{m}\frac{\partial}{\partial \mathbf{v}} \cdot \left[\left(\mathbf{E} + \frac{\mathbf{v} \times \mathbf{B}}{c}\right) f\right] = \sum_{s'} C[f, f_{s'}]. \tag{4.1}$$

Taking the density moment by integrating over all velocities, and using that the terms under the exact divergence vanish at infinity and collisions conserve particle number (cf. (3.13)), we arrive at

$$\int d^3\mathbf{v} \left(\frac{\partial f}{\partial t} + \nabla \cdot (\mathbf{v} f)\right) = 0 \tag{4.2}$$

which is the usual continuity equation for particle number.

We first want to average (4.2) over a flux surface, so that we are left with only the *cross-flux-surface* flux, since this is the relevant flux for fusion experiments. Performing the FSA (cf. Section 2.46) on (4.2) and using (2.51), we see that the second term becomes

$$\left\langle \int d^3\mathbf{v} \nabla \cdot (\mathbf{v} f) \right\rangle_\psi = \frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \int d^3\mathbf{v} f \mathbf{v} \cdot \nabla \psi \right\rangle_\psi \right). \tag{4.3}$$

We now take advantage of the axisymmetry of the magnetic field configuration within a tokamak, and proceed by using the results of Section 2.4. Taking the cross product of (2.38) with $\nabla \zeta$ and using simple vector identities, and that $\nabla \phi \cdot \nabla \psi = 0$ in an axisymmetric configuration, we see that we can write $\nabla \psi = -R^2 \mathbf{B}_0 \times \nabla \zeta$. Inserting this into (4.3), we find

$$\begin{aligned}
\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \int d^3\mathbf{v} f \mathbf{v} \cdot \nabla \psi \right\rangle_\psi \right) &= \frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \int d^3\mathbf{v}\ (-R^2 f) \nabla \zeta \cdot (\mathbf{v} \times \mathbf{B}_0) \right\rangle_\psi \right) \\
&= -\frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \int d^3\mathbf{v}\ (R^2 \nabla \zeta \cdot \mathbf{v})\ B_0 \frac{\partial f}{\partial \varphi} \right\rangle_\psi \right) \\
&= \frac{1}{V'}\frac{\partial}{\partial \psi}\left(V' \left\langle \int d^3\mathbf{v}\ (R^2 \nabla \zeta \cdot \mathbf{v}) \left((\mathbf{v} \times \mathbf{B}_0) \cdot \frac{\partial f}{\partial \mathbf{v}}\right) \right\rangle_\psi \right),
\end{aligned} \tag{4.4}$$

where we have used that $\mathbf{v}\times\mathbf{B}_0 = -B_0\partial\mathbf{v}/\partial\varphi$ and integrated by parts in φ, using that f is single-valued in φ and so the boundary term vanishes. Using the Fokker-Planck equation again, we can substitute for $(\mathbf{v}\times\mathbf{B}_0)\cdot\partial f/\partial\mathbf{v}$ in the expression above to find that our flux-surface-averaged (4.2) becomes

$$\left\langle \int d^3\mathbf{v}\left(\frac{\partial f}{\partial t}\right)\right\rangle_\psi = \frac{1}{V'}\frac{\partial}{\partial\psi}\left(V'\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\,\mathbf{v}\left[\frac{B_0}{\Omega_0}\left(\frac{\partial f}{\partial t}+\mathbf{v}\cdot\nabla f-\sum_{s'}C[f,f_{s'}]\right)\right.\right.\right.$$
$$\left.\left.\left.+\left(-c\nabla\delta\phi-\frac{\partial\delta\mathbf{A}}{\partial t}+\mathbf{v}\times\delta\mathbf{B}\right)\cdot\frac{\partial f}{\partial\mathbf{v}}\right]\right\rangle_\psi\right). \tag{4.5}$$

Now, we wish to average this equation over intermediate temporal and spatial[1] scales, such that only the statistical average of the turbulence will contribute to the macroscopic flux[2]. Since $\mathbf{B}_0$, and therefore ψ, is time-independent, the ITA and the FSA clearly commute. A spatial average of (4.5) becomes purely an average over ψ, since all dependence on θ and ζ has vanished due to the FSA. We then take advantage of the separation of scales in ψ to say that the $\partial/\partial\psi$ acting in (4.5) is over scales ψ_0 much larger than $\Delta\psi$, and hence our average over $\Delta\psi$ may be brought inside $\partial/\partial\psi$. By doing so, we see that the FSA in (4.5) becomes an ISA, by using the relation

$$\frac{1}{\Delta\psi}\int_{\psi_0}^{\psi_0+\Delta\psi} d\psi\langle X\rangle_\psi = \frac{1}{V'\Delta\psi}\int_{\psi_0}^{\psi_0+\Delta\psi} d\psi\int_{-\pi}^{\pi} d\theta\int_0^{2\pi} d\zeta JX \approx \langle X\rangle_\lambda, \tag{4.6}$$

using (2.47) and (2.5) and approximating $V'\Delta\psi\approx V$, since V varies on the large-scale ψ_0. We will hereafter rename $\partial/\partial\psi\to\partial/\partial\psi_0$ to signify that this derivative is taken over large scales in ψ.

All that remains is to insert (3.34) for the decomposition of f,[3]

$$f = F_0 + F_1 - \frac{Ze\delta\phi}{T}F_0 + h + F_2 + \delta f_2 + \mathcal{O}(\epsilon^3 F_0), \tag{4.7}$$

into this equation, average the result over intermediate timescales using the ITA, and examine the remaining terms up to $\mathcal{O}(\epsilon^3\Omega_0 F_0)$ under the velocity integral. An important result to remember is that the ITA drops the order of each fluctuating

[1] The average over intermediate spatial scales is required because there may be persistent structures in the turbulence constant on flux surfaces, known as *zonal flows*, which vary over small ψ.

[2] This can be thought of as an ensemble average over perturbations, via the *ergodic principle*.

[3] For now, we keep $F_0 = F_0(\mathbf{r})$ and $F_1 = \widetilde{F_1} - \boldsymbol{\rho}\cdot\nabla F_0$.

term by ϵ, so if we can show that a term linear in fluctuations is $\mathcal{O}(\epsilon^3\Omega_0 F_0)$ before the ITA, it will not contribute at this order.

Carrying out this procedure, we see that the left-hand side just becomes

$$\overline{\left\langle \int d^3\mathbf{v}\left(\frac{\partial F_0}{\partial t} - \frac{\partial}{\partial t}\left(\frac{Ze\delta\phi}{T}F_0\right) + \frac{\partial h}{\partial t} + \frac{\partial F_1}{\partial t} + \frac{\partial \delta f_2}{\partial t} + \frac{\partial F_2}{\partial t}\right)\right\rangle}_\lambda = \frac{\partial n_0}{\partial t}, \tag{4.8}$$

where we have defined the equilibrium number density $n_0 \equiv \langle \int d^3\mathbf{v} F_0\rangle_\lambda$. Here, we have used that the partial time derivatives of F_1 and F_2 are higher order, that the ITA drops the order of the time derivative of δf_2, and the combination of the ITA and ISA drops the order of h and $\delta\phi$.[4]

Before treating the right-hand side of (4.5), it is useful to first calculate the order of the prefactor of the square bracket. From the definition of ψ_0 as the poloidal ribbon flux, we can see that $\psi_0 \sim B_0 L^2$. Using this, along with maximally ordering $R \sim L$ and using $v_T/L \sim \epsilon\Omega_0$, we see that $(1/V')\partial/\partial\psi_0(V'(R^2\nabla\zeta\cdot\mathbf{v})) \sim \epsilon\Omega_0/B_0$.

Immediately, we see that the first terms in the square bracket are multiplied by an overall factor of ϵ, meaning that we only need to keep them up to $\mathcal{O}(\epsilon^2\Omega_0 F_0)$. Performing the same treatment as before on $\partial f/\partial t$, we see again that only the F_0 contribution survives, but may now be neglected since it is an order higher in ϵ.

Turning our attention to the $\mathbf{v}\cdot\nabla f$ term in (4.5), we insert the full expansion of f and perform the ITA to find[5]

$$\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle \frac{R^2 B_0}{\Omega_0}\nabla\zeta\cdot\int d^3\mathbf{v}\ \mathbf{v}\mathbf{v}\cdot\nabla\left(F_0(\mathbf{R}) + \widetilde{F_1} - \frac{Ze\delta\phi}{T}F_0 + h + F_2 + \delta f_2\right)\right\rangle}_\lambda\right)$$
$$= \frac{B_0}{\Omega_0}\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\ \mathbf{v}\mathbf{v}\cdot\nabla\widetilde{F_1}\right\rangle_\lambda\right), \tag{4.9}$$

where the F_2 term is higher order, and the δf_2 term has been eliminated by the ITA. The h and $\delta\phi$ terms can be integrated by parts in space, using the divergence theorem to eliminate the boundary term, since contributions from the ψ and $\psi+\Delta\psi$ flux surfaces are statistically equal and opposite: the surface normals are antiparallel to leading order, and statistical periodicity enforces that fluctuating quantities are

[4] The two averages combine to form the turbulence average used in [4] to eliminate these terms. One could also substitute for these terms using the gyrokinetic equation, and show that each term vanishes separately.

[5] Remember that B_0/Ω_0 is just mc/Ze, a constant.

statistically the same on both surfaces. The remaining term is higher order, since the gradient is now applied to equilibrium quantities. The F_0 term vanishes by employing[6]

$$\int d^3\mathbf{r} \int_{\mathbf{r}} d^3\mathbf{v} = \int d^3\mathbf{R} \int_{\mathbf{R}} d^3\mathbf{v} \tag{4.10}$$

(where the subscript denotes the variable to be held fixed during integration), and using that F_0 is a function of v only in velocity space, meaning the angle integrals can be performed first to leave $\int d^3\mathbf{v} F_0 \mathbf{v}\mathbf{v} \propto \mathbf{I}$. We then use $\nabla F_0 = F_0'(\psi)\nabla\psi$ and $\nabla\psi \cdot \nabla\zeta = 0$ to eliminate the term. Consequently, we are left with just the $\widetilde{F_1}$ term, as shown. Following [2], we define the *neoclassical pressure tensor* as

$$\underline{\underline{\mathbf{P}}} \equiv \int d^3\mathbf{v}\ \mathbf{v}\mathbf{v}\widetilde{F_1} = \int d^3\mathbf{v} \left(v_\parallel^2 \hat{\mathbf{b}}\hat{\mathbf{b}} + \frac{v_\perp^2}{2}(\mathbf{I} - \hat{\mathbf{b}}\hat{\mathbf{b}}) \right)\widetilde{F_1} + \underline{\underline{\widetilde{\mathbf{P}}}}, \tag{4.11}$$

where $\underline{\underline{\widetilde{\mathbf{P}}}} \equiv \int d^3\mathbf{v}(\hat{\mathbf{b}}\mathbf{v}_\perp + \mathbf{v}_\perp\hat{\mathbf{b}})v_\parallel \widetilde{F_1}$ is antisymmetric.[7] Using this tensor, (4.9) becomes

$$\frac{B_0}{\Omega_0}\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\left\langle R^2\nabla\zeta\cdot\left(\nabla\cdot\underline{\underline{\mathbf{P}}}\right)\right\rangle_\lambda\right) = \frac{B_0}{\Omega_0}\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\left\langle \nabla\cdot\left(R^2\nabla\zeta\cdot\underline{\underline{\mathbf{P}}}\right) - \underline{\underline{\mathbf{P}}} : \nabla\left(R^2\nabla\zeta\right)\right\rangle_\lambda\right). \tag{4.12}$$

The first term vanishes using the divergence theorem, and employing periodicity in the toroidal and poloidal angles. The part of the second term involving the symmetric part of $\underline{\underline{\mathbf{P}}}$ disappears because $\nabla(R^2\nabla\zeta) = \hat{\mathbf{e}}_R\hat{\boldsymbol{e}}_\zeta - \hat{\mathbf{e}}_\zeta\hat{\boldsymbol{e}}_R$ is an antisymmetric tensor, which when double-contracted with the symmetric part of $\underline{\underline{\mathbf{P}}}$ goes to zero. The $\underline{\underline{\widetilde{\mathbf{P}}}}$ part of the second term also vanishes: using (4.10) and that $\widetilde{F_1}$ is gyrotropic, we are left with a term inside the brackets proportional to $\int_{\mathbf{R}} d\varphi\ \underline{\underline{\widetilde{\mathbf{P}}}} = 0$. Hence, we conclude that the entire $\mathbf{v}\cdot\nabla f$ term vanishes at this order.

We now consider the electromagnetic terms in (4.5). Inserting (4.7) and performing the ITA, we find for the F_0 term

$$\begin{aligned}
&\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\ \mathbf{v}\left(-c\nabla\delta\phi - \frac{\partial\delta\mathbf{A}}{\partial t} + \mathbf{v}\times\delta\mathbf{B}\right)\cdot\left(-\frac{m\mathbf{v}}{T_0}\right)F_0\right\rangle}_\lambda\right) \\
&\qquad = \frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\ \mathbf{v}\left(\mathbf{v}_\perp\cdot\nabla\delta\phi\right)\frac{mc}{T_0}F_0\right\rangle}_\lambda\right) \\
&\qquad = -\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\ \mathbf{v}\left(\left.\frac{\partial\delta\phi}{\partial\varphi}\right|_{\mathbf{R}}\right)\frac{mc}{T_0}F_0\right\rangle}_\lambda\right) = 0,
\end{aligned} \tag{4.13}$$

[6] This relation, featured in [2], can be understood by considering that each particle has a unique guiding centre, and so when integrating over the whole phase space, we can choose to count particle positions or gyrocentre positions.

[7] This can be seen by performing the φ integration, recalling $\widetilde{F_1}$ is gyrotropic.

where for the first equality we use that the ITA moves all $\mathcal{O}(\epsilon^3\Omega_0 F_0)$ terms which are linear in fluctuations to higher order, and for the second equality we have used (3.9) and (4.10), and the fact that $\delta\phi$ is independent of φ at fixed $\mathbf{r}$ and single-valued in φ at fixed $\mathbf{R}$, to eliminate the remaining term – hence, the entire F_0 term vanishes.

The Boltzmann term can be treated in an exactly analogous manner: since $\delta\phi$ and T_0 are independent of $\mathbf{v}$, the velocity derivative becomes

$$\frac{\partial}{\partial \mathbf{v}}\left(-\frac{Ze\delta\phi}{T_0}F_0\right) = \frac{Ze\delta\phi}{T_0}\frac{m\mathbf{v}}{T_0}F_0. \tag{4.14}$$

We can therefore see that we have the same form as in (4.13), except all terms have been shifted up one order in ϵ. At $\mathcal{O}(\epsilon^3\Omega_0 F_0)$, and we are instead left with the quadratic term

$$\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\;\mathbf{v}\Big(\mathbf{v}_\perp\cdot\nabla\delta\phi\Big)\delta\phi\frac{Zemc}{T_0^2}F_0\right\rangle_\lambda}\right). \tag{4.15}$$

Rewriting $\delta\phi\mathbf{v}_\perp\cdot\nabla\delta\phi = (1/2)\mathbf{v}_\perp\cdot\nabla(\delta\phi)^2$, we can perform the same treatment using (3.9) and (4.10) to see that the entire Boltzmann term also goes to zero.

To deal with the term involving h, we first integrate by parts in velocity space to find

$$\begin{aligned}\frac{1}{V'}\frac{\partial}{\partial\psi_0}&\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\;\mathbf{v}\left(-c\nabla\delta\phi-\frac{\partial\delta\mathbf{A}}{\partial t}+\mathbf{v}\times\delta\mathbf{B}\right)\cdot\left(\frac{\partial h}{\partial\mathbf{v}}\right)\right\rangle_\lambda}\right)\\ &= -\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\;\left(-c\nabla\delta\phi-\frac{\partial\delta\mathbf{A}}{\partial t}+\mathbf{v}\times\delta\mathbf{B}\right)h\right\rangle_\lambda}\right).\end{aligned} \tag{4.16}$$

Expanding $\mathbf{v}\times\delta\mathbf{B} = \nabla(\mathbf{v}\cdot\delta\mathbf{A}) - \mathbf{v}\cdot\nabla\delta\mathbf{A}$ and eliminating terms of higher order than $\mathcal{O}(\epsilon^3\Omega_0 F_0)$, we find that this can be written

$$\begin{aligned}-\frac{1}{V'}\frac{\partial}{\partial\psi_0}&\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\;\Big(-c\nabla\delta\phi+\nabla(\mathbf{v}\cdot\delta\mathbf{A})-\mathbf{v}_\perp\cdot\nabla\delta\mathbf{A}\Big)h\right\rangle_\lambda}\right)\\ &= \frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle\int d^3\mathbf{v}\;\Big(cR^2\nabla\zeta\cdot\nabla\chi\Big)h\right\rangle_\lambda}\right),\end{aligned} \tag{4.17}$$

where we have used (3.9), (4.10) and the fact that h is gyrotropic to eliminate the $\mathbf{v}_\perp\cdot\nabla\delta\mathbf{A}$ term with the φ integration, and substituted for the gyrokinetic potential χ. We can write this in a more physically illuminating form, following [2]. Recalling the definition of the generalised $E\times B$ velocity $\mathbf{v}_\chi$ from (3.44), we see that we can write

$$\mathbf{v}_\chi\times\mathbf{B}_0 = c(\langle\nabla\chi\rangle_\mathbf{R}-\hat{\mathbf{b}}\Big(\hat{\mathbf{b}}\cdot\langle\nabla\chi\rangle_\mathbf{R}\Big) \quad\Rightarrow\quad R^2\nabla\zeta\cdot\Big(\mathbf{v}_\chi\times\mathbf{B}_0\Big) = cR^2\nabla\zeta\cdot\langle\nabla\chi\rangle_\mathbf{R}+\mathcal{O}(c\chi), \tag{4.18}$$

where we have used that the second term is an order higher in ϵ. Using the relation $\nabla\psi = -R^2\mathbf{B}_0 \times \nabla\zeta$ to find that $cR^2\nabla\zeta \cdot \langle\nabla\chi\rangle_{\mathbf{R}} = -\nabla\psi \cdot \mathbf{v}_\chi$ to leading order, and noticing that the velocity integral already includes an integral over φ at constant $\mathbf{R}$ (using (4.10)) so we are free to replace $\nabla\chi$ with $\langle\nabla\chi\rangle_{\mathbf{R}}$, we finally see that the only contribution to the right-hand side of (4.5) from the turbulent fluctuations is the radial component of the generalised $E \times B$ drift:

$$\boxed{-\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle \nabla\psi \cdot \int d^3\mathbf{v}\ \mathbf{v}_\chi h \right\rangle}_\lambda\right).} \tag{4.19}$$

On to the F_1 term,[8] and we can again integrate by parts in velocity space and eliminate higher-order terms to obtain

$$\begin{aligned} \frac{1}{V'}\frac{\partial}{\partial\psi_0}&\left(V'\overline{\left\langle R^2\nabla\zeta \cdot \int d^3\mathbf{v}\ \mathbf{v}\left(-c\nabla\delta\phi - \frac{\partial\delta\mathbf{A}}{\partial t} + \mathbf{v}\times\delta\mathbf{B}\right)\cdot\left(\frac{\partial F_1}{\partial\mathbf{v}}\right)\right\rangle}_\lambda\right) \\ &= -\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta \cdot \int d^3\mathbf{v}\ \left(-c\nabla\chi - \mathbf{v}_\perp\cdot\nabla\delta\mathbf{A}\right)F_1\right\rangle}_\lambda\right). \end{aligned} \tag{4.20}$$

Again, (3.9) and (4.10) can be used to eliminate the $\mathbf{v}_\perp \cdot \nabla\delta\mathbf{A}$ term, and we can use the trick of integration by parts and employing statistical periodicity to transfer the gradient operators to equilibrium quantities as follows:

$$\begin{aligned} -\frac{1}{V'}\frac{\partial}{\partial\psi_0}&\left(V'\overline{\left\langle R^2\nabla\zeta \cdot \int d^3\mathbf{v}\ \left(-c\nabla\chi\right)F_1\right\rangle}_\lambda\right) \\ &= \frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle \int d^3\mathbf{v}\ \left(-c\chi\right)\nabla\cdot\left(F_1R^2\nabla\zeta\right)\right\rangle}_\lambda\right), \end{aligned} \tag{4.21}$$

where the boundary term vanishes as before. This term is $\mathcal{O}(\epsilon^4\Omega_0F_0)$, and hence the entire F_1 term also does not contribute at this order.

Turning our attention to the collision terms, we see they contribute at $\mathcal{O}(\epsilon^3\Omega_0F_0)$ as follows. Recalling the separation of F_1 into a gyrotropic part $\widetilde{F_1}$ and the gyroradius correction $\boldsymbol{\rho}\cdot\nabla F_0$ in (3.30), the total collision term is

$$\begin{aligned} \sum_{s'}&\Big(C[F_0,(h+\widetilde{F_1}-\boldsymbol{\rho}\cdot\nabla F_0)_{s'}] + C[(h+\widetilde{F_1}-\boldsymbol{\rho}\cdot\nabla F_0),F_{0s'}]\Big) \\ &= C_t^{(l)}[h] + C_{nc}^{(l)}[\widetilde{F_1}] - C_\rho^{(l)}[\boldsymbol{\rho}\cdot\nabla F_0], \end{aligned} \tag{4.22}$$

[8]We can treat this all together as one term, since the only property we need is that $\widetilde{F_1}$ and $\boldsymbol{\rho}\cdot\nabla F_0$ are first-order terms which vary on the length scale L.

defining $C^{(l)}_\rho[\boldsymbol{\rho}\cdot\nabla F_0] \equiv \sum_{s'}(C[F_0,(\boldsymbol{\rho}\cdot\nabla F_0)_{s'}] + C[\boldsymbol{\rho}\cdot\nabla F_0, F_{0s'}])$. The $C^{(l)}_t[h]$ term is $\mathcal{O}(\epsilon^3\Omega_0 F_0)$ and therefore vanishes under the ITA, so the total contribution to the right-hand side of (4.5) is

$$\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\;\mathbf{v}\left(-\frac{B_0}{\Omega_0}\left(C^{(l)}_{nc}[\widetilde{F_1}] - C^{(l)}_\rho[\boldsymbol{\rho}\cdot\nabla F_0]\right)\right)\right\rangle}_\lambda\right). \tag{4.23}$$

Dealing with the second term first, we can use $\nabla\psi = -R^2\mathbf{B}_0\times\nabla\zeta$ and (2.29) to find

$$R^2\frac{B_0}{\Omega_0}\mathbf{v}\cdot\nabla\zeta = R^2\frac{B_0}{\Omega_0}v_\parallel\hat{\mathbf{b}}\cdot\nabla\zeta - \boldsymbol{\rho}\cdot\nabla\psi, \tag{4.24}$$

and hence the second term can be written

$$\boxed{-\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle\nabla\psi\cdot\int d^3\mathbf{v}\;\left(\boldsymbol{\rho}C^{(l)}_\rho[\boldsymbol{\rho}\cdot\nabla F_0]\right)\right\rangle}_\lambda\right),} \tag{4.25}$$

where the first term in (4.24) vanishes since it is odd in velocity space. Substituting in the neoclassical equation (3.69) to simplify the first term in (4.23), we find it becomes

$$\begin{aligned}
&\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle R^2\nabla\zeta\cdot\int d^3\mathbf{v}\;\mathbf{v}\left(-\frac{B_0}{\Omega_0}\left(\cancelto{0}{(\mathbf{v}_{\nabla B}+\mathbf{v}_\kappa)\cdot\nabla F_0} + v_\parallel\hat{\mathbf{b}}\cdot\nabla\widetilde{F_1}\right)\right)\right\rangle}_\lambda\right)\\
&\quad= -\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle\int d^3\mathbf{v}\;\frac{Iv_\parallel}{\Omega_0}\left(v_\parallel\hat{\mathbf{b}}\cdot\nabla\widetilde{F_1}\right)\right\rangle}_\lambda\right)\\
&\quad= \frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle\int d^3\mathbf{v}\;\widetilde{F_1}v_\parallel\hat{\mathbf{b}}\cdot\nabla\left(\frac{Iv_\parallel}{\Omega_0}\right)\right\rangle}_\lambda\right),
\end{aligned} \tag{4.26}$$

where the first term in the first line is odd in velocity, and we have integrated by parts in velocity space in the same way we did in (3.19) to obtain the result. We now employ a relation, proven in [4], which states that

$$v_\parallel\hat{\mathbf{b}}\cdot\nabla\left(\frac{Iv_\parallel}{\Omega_0}\right) = -(\mathbf{v}_{\nabla B}+\mathbf{v}_\kappa)\cdot\nabla\psi \tag{4.27}$$

in order to find that the final contribution from the second term of (4.23) to the right-hand side of (4.5) is

$$\boxed{-\frac{1}{V'}\frac{\partial}{\partial\psi_0}\left(V'\overline{\left\langle\nabla\psi\cdot\int d^3\mathbf{v}\;(\mathbf{v}_{\nabla B}+\mathbf{v}_\kappa)\widetilde{F_1}\right\rangle}_\lambda\right)} \tag{4.28}$$

Lastly, we can clearly see that the f_2 terms will both be $\mathcal{O}(\epsilon^4\Omega F_0)$, and so they can both be neglected to leading order. An external particle source term S_n may also be included *ad hoc* at this order, to account for equilibrium-scale input of particles into the system, such as via neutral beam injection.

Bringing together all of the terms which contribute to the right-hand side of (4.5), and moving them to the left-hand side, we realise we have arrived at an equation for the equilibrium-scale cross-flux-surface particle transport which is correct to $\mathcal{O}(\epsilon^3 \Omega F_0)$:

$$\frac{\partial n_0}{\partial t} + \frac{1}{V'}\frac{\partial}{\partial \psi_0}\left(V' \overline{\langle \Gamma \rangle}_\lambda\right) = S_n \tag{4.29}$$

$$\Gamma \equiv \nabla\psi \cdot \int d^3\mathbf{v}\, \left(\mathbf{v}_\chi h + (\mathbf{v}_{\nabla B} + \mathbf{v}_\kappa)\widetilde{F_1} + \boldsymbol{\rho} C_\rho^{(l)}[\boldsymbol{\rho}\cdot\nabla F_0]\right). \tag{4.30}$$

This equation manifestly features only the cross-flux-surface flux, and describes the net flux due to the combined effect of turbulent, neoclassical and collisional transport.

4.3 Evolution of the Pressure Profile

A similar derivation (which we omit for lack of space) shows that the pressure evolves according to (cf. [3])

$$\frac{3}{2}\frac{\partial p_0}{\partial t} + \frac{1}{V'}\frac{\partial}{\partial \psi_0}\left(V' \overline{\langle Q \rangle}_\lambda\right) = -\overline{\langle H \rangle}_\lambda + \frac{3}{2} n_0 \sum_{s'} \nu^\varepsilon_{ss'}(T_{s'} - T_s) + S_p \tag{4.31}$$

$$Q \equiv \nabla\psi \cdot \int d^3\mathbf{v}\, \frac{mv^2}{2}\left(\mathbf{v}_\chi h + (\mathbf{v}_{\nabla B} + \mathbf{v}_\kappa)\widetilde{F_1} + \boldsymbol{\rho} C_\rho^{(l)}[\boldsymbol{\rho}\cdot\nabla F_0]\right) \tag{4.32}$$

$$H \equiv \int d^3\mathbf{v}\; e\left(\frac{\partial \chi}{\partial t}\right). \tag{4.33}$$

Here, Q is the cross-flux-surface heat flux, H is the rate of heating by the fluctuating fields, S_p is the external heat source, and $\nu^\varepsilon_{ss'}$ is the collisional energy exchange frequency given by

$$\nu^\varepsilon_{ss'} \equiv \frac{6.88 (m_s m_{s'})^{1/2} Z_s^2 Z_{s'}^2 e^4 n_{0_{s'}} \log\Lambda_{ss'}}{(m_s T_{s'} + m_{s'} T_s)^{3/2}}, \tag{4.34}$$

where $\log\Lambda_{ss'}$ is the Coulomb logarithm. It can be seen, rather intuitively, that the mechanisms responsible for particle transport also allow kinetic energy to be transported, by comparing Q and Γ. Notable additions in the equation for heat transport are H, accounting for the work done on the particles by the fluctuating potentials, and $(3/2)n_0 \sum_{s'} \nu^\varepsilon_{ss'}(T_{s'} - T_s)$, accounting for collisional heating by other species.[9]

[9] This is the *ad hoc* collisional energy exchange mentioned previously.

A conclusion is the place where you got tired thinking.

— Martin H. Fischer

5 Conclusions

In this work, we have taken advantage of the large separation of scales to systematically expand the Fokker-Planck equation in a small parameter ϵ, and therefore derive equations for the turbulent (h) and neoclassical ($\widetilde{F_1}$) first-order corrections to the Maxwell-Boltzmann equilibrium distribution function for a tokamak plasma in a steady equilibrium magnetic field. Assuming that all relevant dynamics occured on a timescale much larger than that of a gyroperiod, we averaged our equations over gyroangle at each order, meaning that the second-order correction to the distribution function was not required in our equations for the first-order pieces. We also found that the turbulent distribution function physically describes a distribution of charged rings in gyrokinetic coordinates over a 5D phase space. In order to simplify our treatment of the magnetic geometry of the system, we developed the flux-coordinate formalism and employed axisymmetry, which we utilised to define averages both over intermediate spatial scales, and over a flux surface. We also defined an average over timescales well-separated from both the turbulent and equilibrium scales, and used this, along with the spatial equivalent, to average moments of the distribution function over intermediate scales. In so doing, we kept only the statistical average of the turbulent contributions, and by averaging these equations over a flux surface we ensured that we kept only terms which contributed to the cross-flux-surface flux. The end result is a set of equations which describe the mean density and temperature evolution inside a tokamak. Manifest in these equations are the contributions to the fluxes from collisions and neoclassical transport, as well as from turbulent fluctuations of the electromagnetic potentials; the latter is responsible for both transporting particles and kinetic energy outwards via the generalised $E \times B$ velocity, and for doing work on the particles via the time derivative of the gyrokinetic potential in the heating term of the energy transport equation. We have therefore shown that micro- and macro-scale phenomena in a tokamak are intrinsically coupled – the gyrokinetic equation demonstrates that the turbulent distribution function evolves in a way which depends on macroscopic gradients, which are themselves influenced by the turbulence as shown by the transport equations.

In addition to elucidating the physics of coupling across scales in a tokamak, the approach developed in this dissertation is used numerically in `Trinity` [2, 3], a code developed to simulate the effect of turbulence on the radial profiles of equilibrium quantities by coupling a transport solver to continuum gyrokinetic codes, such as `GS2` [12]. This is much more computationally efficient than full-f 6D simulations; in addition, the latter encounters difficulties due to the large separation of scales demanding that the distribution function and fields must be calculated to very high order. Gyrokinetic theory and simulations therefore provide a way of routinely studying cross-flux-surface tokamak transport, an understanding of which is key to designing future fusion experiments.

References

[1] G. G. Howes et al. “Astrophysical Gyrokinetics: Basic Equations and Linear Theory”. In: *Astrophysical Journal* 651.506102 (2006).

[2] M. Barnes. “Trinity: A Unified Treatment of Turbulence, Transport, and Heating in Magnetized Plasmas”. PhD thesis. University of Maryland, 2008.

[3] M. Barnes et al. “Direct multiscale coupling of a transport code to gyrokinetic turbulence codes”. In: *Physics of Plasmas* 17.056109 (2009).

[4] I. G. Abel et al. “Multiscale Gyrokinetics for Rotating Tokamak Plasmas: Fluctuations, Transport and Energy Flows”. In: *Rep. Prog. Phys.* 76.116201 (2013).

[5] M. Hardman. “Notes on the derivation of the Gyrokinetic Equation”. Personal notes. 2017.

[6] W. D. D’haeseleer et al. *Flux Coordinates and Magnetic Field Structure: A Guide to a Fundamental Tool of Plasma Theory*. Springer Series in Computational Physics. Springer-Verlag, 1991.

[7] *Article on Flux Coordinates*. URL: `http://fusionwiki.ciemat.es/wiki/Flux_coordinates`.

[8] G. D. Conway. “Turbulence measurements in fusion plasmas”. In: *Plasma Physics and Controlled Fusion* 50.124026 (2008).

[9] R. J. Fonck et al. “Long-wavelength density turbulence in the TFTR tokamak”. In: *Physical Review Letters* 70.3736 (1993).

[10] G. R. McKee et al. “Non-dimensional scaling of turbulence characteristics and turbulent diffusivity”. In: *Nuclear Fusion* 41.1235 (2001).

[11] F. I. Parra and P. J. Catto. “Limitations of gyrokinetics on transport time scales”. In: *Plasma Physics and Controlled Fusion* 50.065014 (2008).

[12] *GS2*. URL: `http://gs2.sourceforge.net`.