

Exciting Simulations in AC Tokamak Ramp Downs

Thomas Malcolm

Oct 2023

A thesis submitted for the degree of
Bachelor of Mathematical Sciences (Honours)
of the Australian National University



Australian
National
University

For George, Paul, Ringo and John.



Declaration

The work in this thesis is my own except where otherwise stated.

Thomas Malcolm.

Acknowledgements

First and foremost I'd like to thank Matthew Hole, my supervisor. I came into this year a little daunted at the prospect of tackling a physics-based project with a physics background not extending past "A Brief History of Time". Your patience and willingness to explain concepts to me I'll forever appreciate, and your energetic approach to problem solving made me feel both at home and excited for the work we have done in this thesis. Thank you for always taking the time to talk to me, and for being a fantastic supervisor.

In a similar vein, thank you to the rest of the Plasma Theory and Modelling group for always being willing to answers questions when I had them. You guys were definitely my saving grace at times, and I only wish I spent more time getting to know you all this year. I should give a special shout out to Nick, Sandra, Dean and Josh specifically for being so accommodating, helpful, and kind in answering my questions when I would pop around.

I'd like to thank the other Honours students for helping make this a fantastic year. Our little nook of MSI proved very distracting, but in the best possible way. Also Joe gets a special shout out purely to spite everyone else.

Last but certainly not least, to Kirsten and Graeme. Thank you for your support these last couple years we've known each other, and the infinite kindness you give to all those around you. However silly it is, your encouragement genuinely helped motivate me at times, and it was always my pleasure having the opportunity to rant to you about whatever problem I was working on at the time, and how insurmountable it seemed! Yet each time you had confidence it would all work out fine, and then somehow it always would. I can only assume this is some wizardry I am not yet privy to, but am nevertheless indebted to your kindness and hospitality.

Abstract

One of the biggest unsolved problems of the modern day is that of plasma confinement in fusion reactors; a challenge undertaken with the expressed intent of providing a safe, clean, reliable source of renewable energy. Various approaches to tackling this problem have emerged since the field's inception in 1946, but the most prominent design that has persisted through to modern day is the familiar donut-shaped Tokamak fusion reactor design. As equally as familiar as the quest for fusion power, are the challenges that face its achievement. The problem at the heart of plasma science is that of confinement, and one of the larger antagonists to this goal is the presence of runaway electron (RE) populations. These, if managed, can wreak havoc on the reactor, and disrupt the plasma - not conducive to confinement, and thus stability.

While most fusion reactors operate in what's known as "DC" mode, where the plasma current is oriented in a single direction, we focus our investigations instead on "AC" Tokamak operating modes, where the plasma current is allowed to oscillate back and forth. This, however, introduces its own set of problems - namely that of the observed presence of "anti-parallel" currents, characterised by excess populations of runaway electrons.

Governing the behaviour of a plasma within the confines of a Tokamak's geometry is the Grad-Shafranov equation; a nonlinear, elliptic PDE derived from the principles of magnetohydrodynamics (MHD) - the union of Maxwell's equations and fluid dynamics. Solutions to this PDE provide magnetic flux information, and are instrumental to understanding the dynamics of a plasma, and thus, help inform answers to many of the challenging questions we face in fusion science. In this thesis we investigate a variation of the Grad-Shafranov equation specifically catering to current reversals, and seek to test whether we can simulate the ramp down phase of a current cycle, with the intent of identifying behaviour that might support the existence of anti-parallel currents, and/or provide explanation for the increased population of runaway electrons. We initially simulate this with

emulated data, and then attempt to apply our model to data provided by the ISTTOK reactor.

In attempting to use our model on data from ISTTOK we found that, while there are some difficulties in fitting our parameters to the data that is available, initial results show promise that the model can be used to make observations about the problems described above. We conclude with some potential avenues for further research that could see the model improved.

Contents

Acknowledgements	vii
Abstract	ix
Notation and terminology	xiii
1 Introduction	1
1.1 Plasma Science	1
1.1.1 I'm a Mathematician... what is "plasma"?	1
1.1.2 Fusion Reactor Design 101	3
1.2 What problem does this thesis address?	7
1.3 Project Progression	10
2 Background	11
2.1 Magnetohydrodynamics	11
2.1.1 MHD Theory	12
2.2 Grad-Shafranov Equation	19
2.2.1 Derivation	24
2.3 Runaway Electrons	28
2.3.1 Regular Electron Movement	29
2.3.2 Generation	31
2.3.3 Detection	32
3 Current Reversal Theory	35
3.1 Perturbation Methods	35
3.1.1 What is a Perturbation Treatment?	35
3.1.2 Regular Perturbation Theory	38
3.1.3 Ideal MHD Perturbation	41
3.2 Grad-Shafranov-Helmholtz Equation	47

3.2.1	Equation and Derivation	47
3.2.2	Current Density and Pressure Density	51
3.3	Simulations	52
4	Numerical Model Fitting	55
4.1	Non-Linear Optimisation	55
4.1.1	Least Squares	56
4.1.2	Optimisation Algorithms	57
4.2	GSH Parameter Fitting	59
4.2.1	Optimisation Problem	60
4.2.2	Parameter Space	62
4.3	Simulated Current Reversals	66
4.3.1	Method of Reversal	66
4.3.2	Results and Explanations	66
5	Comparison with Experiment	67
5.1	ISTTOK	67
5.1.1	Heavy Ion Beam Diagnostics	67
5.1.2	Reactor Specification and Data	69
5.2	Parameter Fitting	72
5.2.1	Results	72
6	Blue Skies and Horizons	81
6.1	Conclusions	81
6.2	Further Work	82
6.2.1	Simulated Electric Field via Fake Solenoids	82
6.2.2	Comparison with ISTTOK v_{loop} data	82
6.2.3	Grid Based Initial Parameter Guessing	83
A	Appendix	85
	Bibliography	89

Notation and terminology

Notation

$L^2(\mathbb{R}^n)$	Space of L^2 integrable functions on \mathbb{R}^n
B_0	On-axis magnetic field
R_0	Major radius (Tokamak)
a	Minor radius (Tokamak)
Ψ	Poloidal magnetic flux function
j_ϕ	Toroidal current density
J_ϕ	Normalised toroidal current density (normalised w.r.t. B_0)
p	Plasma pressure density

Terminology

GS / GSE	Grad-Shafranov equation
GSH	Grad-Shafranov-Helmholtz equation
MHD	Magnetohydrodynamics

Chapter 1

Introduction

1.1 Plasma Science

This thesis is multidisciplinary by nature, incorporating aspects of mathematics, physics and computer science throughout the various challenges faced in reaching the conclusions we have.

Before burying ourselves in the thick of our results we will cover the requisite knowledge for working in the plasma science simulation space. First we'll discuss what the physical object of our attention ("plasma") actually is, with a brief discussion on the design of fusion reactors, where we will emphasise the structures that specifically relate to our interests. In the background chapter we will expound on this by building the mathematical theory underpinning physical processes within the fusion reactor, providing us a way to reason about the behaviour of plasma (and related processes) inside a fusion reactor, or, more accurately, approximate their behaviour via simulation.

1.1.1 I'm a Mathematician... what is "plasma"?

While an initially daunting topic, fear not fearful mathematician, for many of the inherently physical behaviours in plasma science we require can be expressed in terms of our dearest mathematical expressions – differential equations! But first, what actually is "plasma"? Webster's dictionary defines plasma to be "a green faintly translucent quartz" [22]. While I'm sure there isn't no relation between crystals and our investigations, this is unfortunately, not the recipient of our affection.

When plasma is referred to in everyday conversation it is often noted as be-

ing the “4th state of matter”. To introduce slightly more rigour, plasma is an extension of the gaseous state of matter, where its energy (read: temperature) is increased sufficiently high that the electrons are no longer bound by the electromagnetic force to the atom’s nucleus [6]. The resulting substance is a “pool” of cations (the positively charged nuclei), and electrons (negatively charged), that exhibits interesting properties. It is these properties that we seek to exploit in the search for controlled, sustainable fusion reactions.

Plasma is abundant in nature – just not in many places that we as Humans commonly look. Stars are the most immediate example of matter in a plasma state, and are readily viewable (at least for half the day). Lightning strikes are paths through the atmosphere which are ionised, and neon signs work by heating Neon gas within a tube to ionise it [8].

The question then is, what is “fusion”, and how does this relate to plasma? The answer comes back to energy. Analogously to a fission reaction, where energy is released through the division of atoms, one can also fuse two separate atoms together and have large amounts of energy emitted as a bi-product of that fusion. It is (essentially) this extra energy that we wish to harness when harvesting energy from a fusion reactor.

While there are no shortages of elements that can theoretically be used to fuel a fusion reactor, the one most commonly associated with fusion is Deuterium – a stable isotope of Hydrogen that has a neutron in its nucleus (whereas a ‘standard’ Hydrogen atom contains only a single proton). Analogously, Tritium is an isotope of Hydrogen that has two neutrons in its nucleus (though is much less stable). Consider the fusion of a Deuterium atom with a Tritium atom:



Here two Hydrogen isotopes fuse together to form a single Helium-4 atom, and in the process of doing so release a single neutron and 17.59 MeV of energy. This excess energy is what is so attractive about fusion processes as a sustainable energy source – for such little input we receive a substantial amount of energy, and at that, using one of the most abundant resources available on Earth; water.

The question then becomes, how do we drive this fusion process? If two atoms can fuse as such, why do we not see atoms fusing everywhere around us accompanied by violent explosions destroying all that we’ve come to know and love? The answer is that we kind of do. In fact, we see this happening everyday – for those of us fortunate enough to be able to see the sun that is. Our sun is a

large ball of plasma where an estimated 9.3×10^{37} fusion reactions are expected to occur every second [23], and is one of the easiest examples of both plasma as a state of matter, and of a self-sustaining fusion reaction.

How then do two hydrogen atoms fuse? The nucleus of an atom (consisting of positively charged protons, and neutral neutrons) is positively charged, and so two atoms' nuclei will repel each other due to the Coulomb force when pushed together. This is the force we have to overcome to enable a fusion reaction to take place (and what stops the world around us burning!). To overcome this the process is relatively “simple” – we just increase the energy of our atoms so that when they collide they collide with enough energy to overcome this force, allowing the strong force to become dominant, fusing the two atoms. When we energise a mass (take a gas here) of atoms enough, they become ionised however, which is exactly the state of plasma we described earlier. In other words, if we want to reason about the creation, sustainment, and effects of fusion reactions, we need first to understand the dynamics of plasma, the medium in which the fusion reactions take place. From this comes a plethora of questions, ranging from “how do we generate such a plasma?” to “how do we reliably control such a plasma” and “how do we harvest energy from a plasma”? Alas, we digress however, as we do not seek to solve the big problems in fusion science in this here mere mortal thesis. Instead, now equipped with at least a passing knowledge of what constitutes a “plasma” state and what it means for a fusion reaction to take place, we will humbly delve into the inner workings of fusion reactor terminology and design.

1.1.2 Fusion Reactor Design 101

Here we will discuss the important structural aspects of a tokamak fusion reactor. The term “tokamak” is a Russian acronym which translates as “toroidal chamber with magnetic coils” [15]. Aptly, a tokamak is a toroidal object which is used as a vessel for plasma which is driven via external magnetic coils. There are two main angular components to a tokamak - components working in the “toroidal” (around the major radius, R , of the tokamak), and those working in the “poloidal” direction (around the minor radius, a , of a cross section of the tokamak). Thus, we will often reason about various physical concepts (such as magnetic fields, current densities) in terms of one of these. Though, in fact, we will often ignore the toroidal component - these are often abstracted away under the guise of an axisymmetry assumption, which assumes that any partials with respect to the

toroidal angle are diminishing, i.e., $\frac{\partial}{\partial \phi} = 0$. We let θ denote poloidal angles, and ϕ the toroidal angles. When we refer to (x, z) , we will be referring to offsets within a poloidal cross section. Refer to figure 1.1 for a graphic distinguishing toroidal from poloidal components.

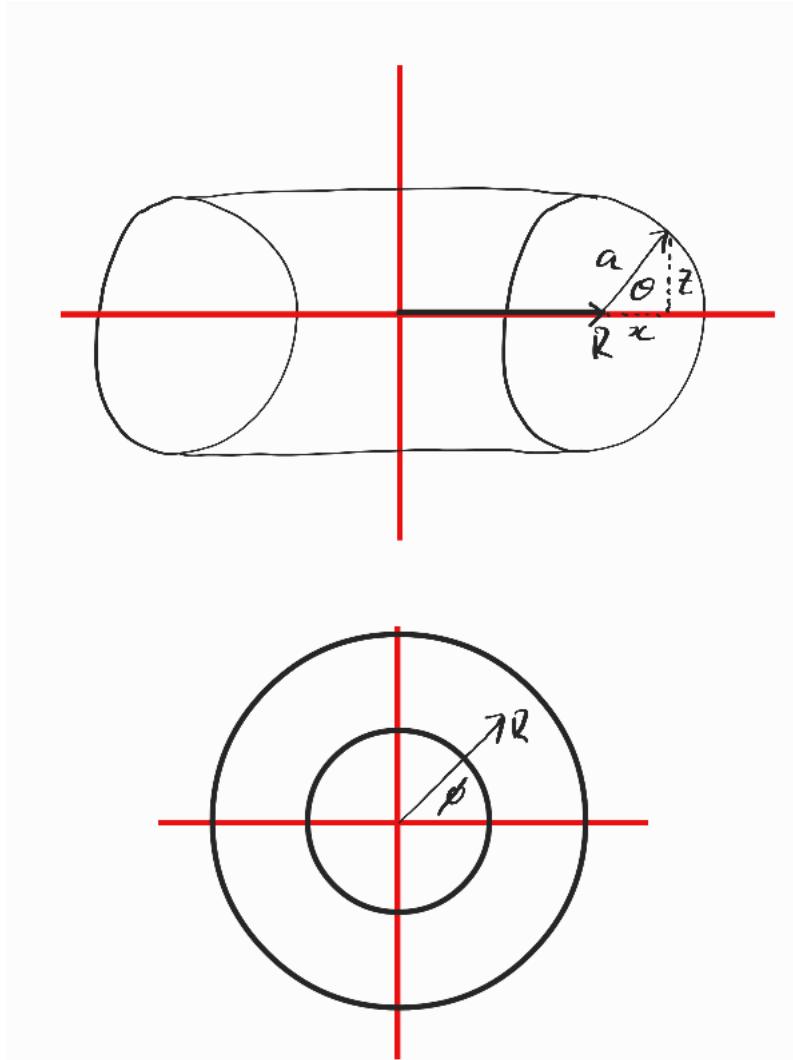


Figure 1.1: The top half depicts poloidal cross section factors, whereas the bottom shows toroidal components.

We draw specific attention to the difference between poloidal and toroidal magnetic flux. In this thesis we will concern ourselves primarily with poloidal magnetic flux, which is what controls lateral “movement” of plasma within the tokamak - which is central to the problem of confinement (pun intended). See figure 1.2 for a comparison.

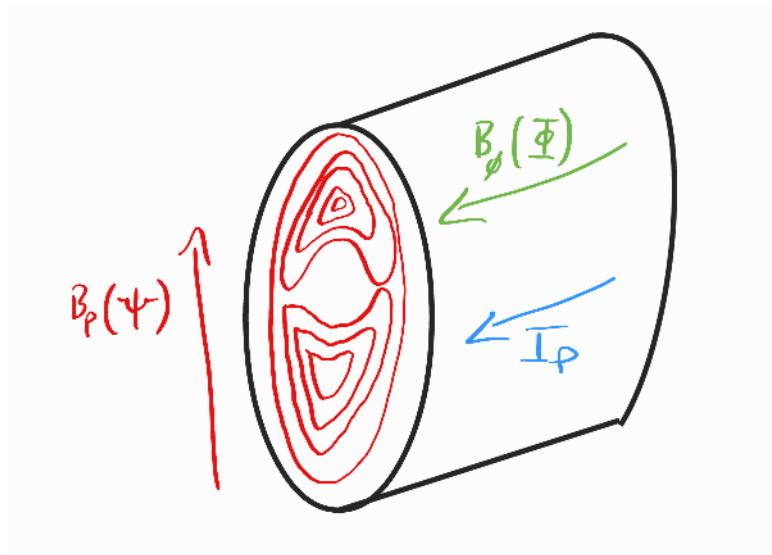


Figure 1.2: Emphasising the difference between the poloidal and magnetic fields. The flux function symbols associated with each are also provided. Alongside this we show the plasma current flow for good measure.

We control fusion reactors via a series of (meticulously!) placed solenoids around the exterior of the reactor. These are positioned such that certain solenoids can be used to shape either the toroidal or poloidal magnetic field. Current is passed through these, the strength of which drives the strength of the resultant magnetic field, and the direction of which is determined by the “right hand rule” if you recall high school physics exams with everyone staring at their hands. We see a depiction of these solenoids and their effect on the poloidal and toroidal magnetic fields in figure 1.3. Note that we have ignored one set of solenoids which accompany these, which are placed following the contours of the toroidal direction, and are used for more fine grained control of the poloidal magnetic flux - these are largely irrelevant for our purposes however, and are annoying to draw, and so are omitted from the graphic.

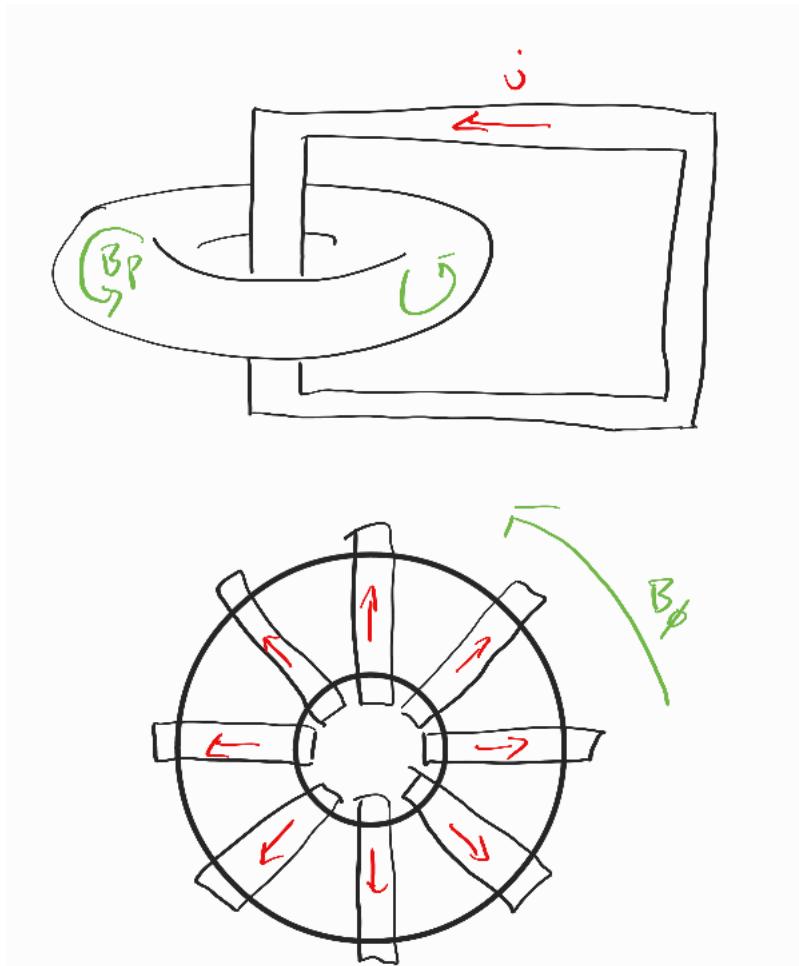


Figure 1.3: Solenoids which control the poloidal and toroidal magnetic fields.

Reactor Geometry

In reality, reactors do not have such nice geometry as the above diagrams (if we take shaky hand-drawn pictures to have “nice” geometry). Figure 1.4 shows a 3D model of the EAST reactor, showing the irregularities in the shape of the cross section that appears in a real reactor. These engineering decisions made here are interesting and could be (and have been) the subject of their own research! For our purposes however, we will assume that our tokamak is either a torus with an elliptic cross section, or even simpler, a cylinder with elliptic cross section. These act as approximations to a real tokamak’s design to make simulation simpler - it

is the “spherical cow” model for fusion reactors.

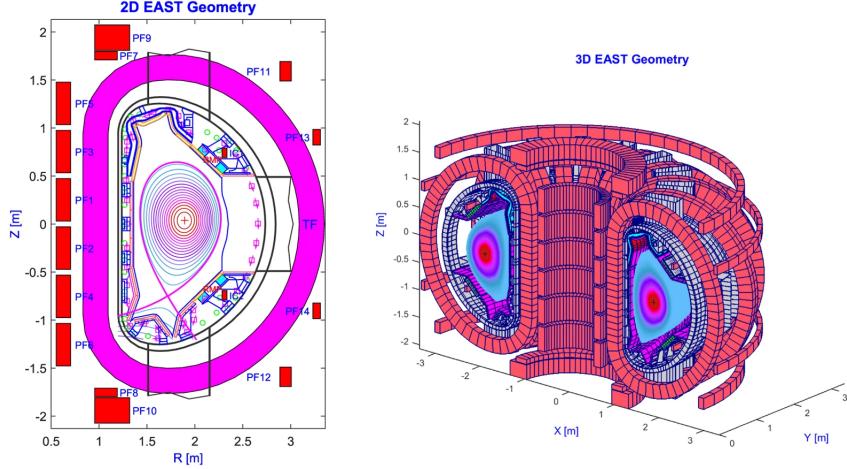


Figure 1.4: EAST fusion reactor geometry. This is a 3D model of its cross section, highlighting the irregular shape [4]

Runaway Electrons

We make one last note on runaway electrons. If a tokamak were Alexander Hamilton, then runaway electrons would be Aaron Burr. Or Hamilton. Or time? Whatever your interpretation of Lin Manuel-Miranda’s musical epic suggests is the enemy. These high energy particles have the potential to be highly damaging to a reactor. They have the ability to directly damage equipment within the reactor (such as instruments that are plasma-facing), can erode the inside protective layers of the tokamak, can disrupt the dynamics of the plasma, and more. In fact it is runaway electron mitigation techniques which are at least part of the reasoning for having such an irregularly shaped cross section as that adopted by the EAST reactor (and many more “modern” reactors). We will talk more about runaway electrons in chapter 2.

Now that we are experts in fusion reactor design (who needs a physics degree, right?), we can proceed to talk about the afflictions that modern day fusion reactors face and what this humble thesis attempts to tackle.

1.2 What problem does this thesis address?

The core problem that this thesis is built around is introduced by Hole and Malaquias in [18]. In a tokamak, external solenoids along with heating elements

within the reactor are what drive the primary current within a plasma. In a reactor operating in the DC regime, this is characterised by being a constant, unidirectional current. However, when operating in AC mode, this naturally becomes oscillatory. At every current reversal there is necessarily a quiescent phase for the plasma current, where $I_p = 0$. At this point we would anticipate there to be no plasma density as a result of this lack of a driving current, however, experiment in ISTTOK disagrees with this sentiment. Notably, reactors such as the CT-6B tokamak identify through experiment the presence of coexisting anti-parallel currents in the ramp down phase - directly contradicting ISTTOK's observations [13]. Summarily, there is no consensus in the community surrounding three central issues:

- Why does there exist a residual plasma density when there is seemingly nothing to drive it? (i.e. $I_p = 0$)
- Why do we see an increase in runaway electron generation during the ramp down phase (and partly during ramp up, though we will note on this missing data in chapter 2)
- There is experimental evidence that both supports and discounts the existence of anti-parallel current formation during the ramp down phase. What could drive this, and is there theoretical support for either side?

It is hypothesised by Hole and Malaquias that this residual plasma density is the result of persistent runaway electron populations in the plasma. It is further suggested that these populations are driven by some mechanism resulting from the ramp down phase - we discuss runaway electron generation mechanisms in chapter 2. As we noted however, experiment is inconsistent on the question of the existence of anti-parallel current channels in the plasma (both positive and negative current densities in the plasma density). In ISTTOK we note that the potential difference of the plasma is vanishing as the plasma current approaches $I_p = 0$. This combined with the observation that the pressure density profile, as measured by the ISTTOK heavy ion beam diagnostics team, show non-vanishing pressures within the plasma, indicating the presence of a residual plasma even as the plasma current vanishes. This same pressure density profile also shows no evidence of the coexistence of anti-parallel currents.

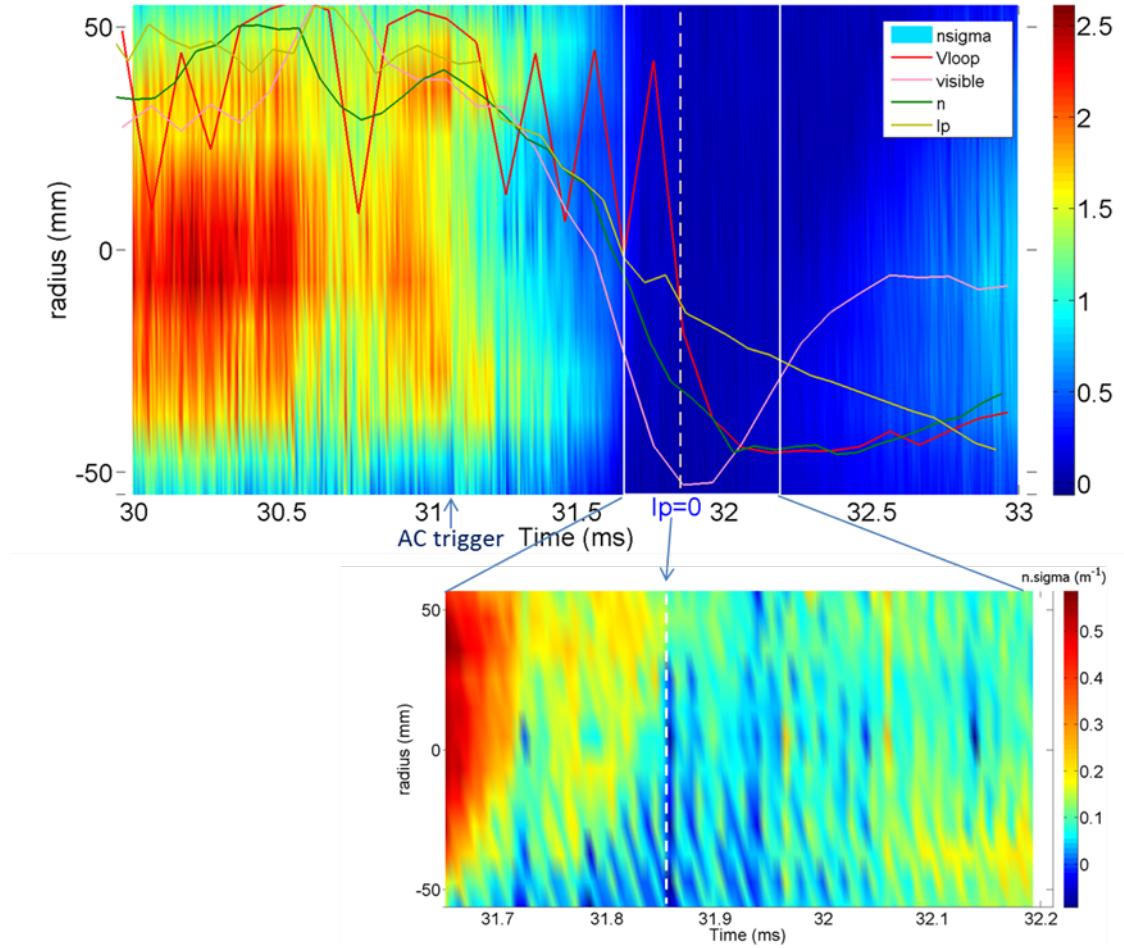


Figure 1.5: ISTTOK data showing the existence of a residual plasma density despite the plasma current I_p and v_{loop} being negligible [18].

What results do we seek?

Our broad scope goal is to simulate the behaviour of a current reversal within a tokamak, with the intent of identifying indicators of runaway electron generation, or support for the presence of anti-parallel current formation - in effect, to provide some insight into the answers behind any of the three points raised above. More specifically, however, we will develop theory around simulating the behaviour of a ramp down phase, necessarily a time evolution, when the tools we have available to us are characteristically time independent. From this we will develop a model which targets specifically the scenario that we have access to current density profile data, and from this derive information about the pressure density profile and the poloidal magnetic flux. We will run simulations for a contrived tokamak to observe some theoretical results, before applying what we've learned to data

made available by the ISTTOK project. From this we will attempt to draw inference about the behaviours listed above, and note any limitations that may exist on our model.

1.3 Project Progression

This project being inspired by work done by Hole and Malaquias and the scope for the project defined explicitly, we first begun with a review of existing theoretical work done in simulating current reversals in tokamaks. These efforts led to the discovery of Wang’s paper [30], which provided an analytic solution to a variation of the Grad-Shafranov equation constructed specifically for the modelling of equilibrium in current reversal tokamak configurations. While the presence of equilibria for AC regime tokamaks is the subject of ongoing research (with theoretical work supporting their existence, but limited experimental support yet), we can nevertheless sought to use their model to simulate a ramp down phase. We first showed that equilibrium solutions to the Grad-Shafranov equation could be perturbed by some small time dependent ε term to emulate the effect of a time evolution. For some small (undetermined) perturbation this would still be an accurate model of the physical effects within a reactor. We were able to simulate the ramp down phase for a contrived reactor using the Helmholtz variation of the Grad-Shafranov equation as proposed by Wang. However, when applying our model to data available from ISTTOK, we see difficulties in achieving an accurate model. Initial efforts suggest some resistance of the system to be accurately represented by our model, however there is a question on whether enough information exists to make statements about the poloidal magnetic field topology, which itself can provide insight into anti-parallel current channels and some runaway electron generation mechanisms. There are a plethora of ways to extend upon the work we’ve done in this thesis, some of which are suggested in chapter 6.

Without further adieu, we will begin.

Chapter 2

Background

2.1 Magnetohydrodynamics

The term “magnetohydrodynamics” (MHD) is a portmanteau of two physical concepts which are used to model plasmas inside fusion reactors (among other things): the “magneto” term comes from “magnetic field”, and “hydrodynamics” indicates a fluid dynamics component. Put together, magnetohydrodynamics is the study of electrically conductive materials that behave like fluids. Essentially it provides a way to model the behaviour of a (considerably volumous!) mass of conductive particles, and their electrodynamic forces, as if it were a fluid, as opposed to having to model individual particle interactions. This idea was first introduced by Hannes Alfvén in 1970, for which he earned the Nobel prize [1]!

Much modern research in Tokamak plasma simulation uses some variation of an MHD model (and the derived Grad-Shafranov Equation, which we will soon be introduced to) for a number of reasons, not least being its comparative computational efficiency. One primary benefit of treating our plasma as a fluid being that we avoid modelling the behaviour of each individual particle in said plasma, a simplification which becomes especially important when we consider the order of number of particles we would have to simulate is of order $\sim 10^{20}$ - far too much for the author’s ThinkPad T440p to even contemplate!

Here we will build the relevant MHD background for this thesis, deriving the MHD equations from first principles, and explaining the assumptions we make to reduce them to a simplified state known as “ideal MHD”. From there we will look at the PDE which models the behaviour of a plasma inside a Tokamak, the “Grad-Shafranov Equation”. We will then note some pitfalls of using this model to describe AC configuration Tokamaks (as we are investigating), and finish with

a discussion on runaway electrons (RE).

2.1.1 MHD Theory

Given MHD is the marriage of fluid dynamics with electrodynamics, it is only natural to begin our study looking at the equations which describe electrodynamic behaviour — Maxwell's equations describe the interaction between magnetic fields $\vec{B}(\vec{r}, t)$, electric fields $\vec{E}(\vec{r}, t)$, and the current density $j(\vec{r}, t)$ which induces them, where $\vec{r} \in \mathbb{R}^3$ is a position vector and $t \in \mathbb{R}$ describes time. Thus, we introduce Maxwell's equations:

Definition 2.1. Maxwell's equations are given [31]:

$$\nabla \times \vec{B} = \mu_0 j + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \quad (2.1)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (2.2)$$

$$\nabla \cdot \vec{B} = 0 \quad (2.3)$$

$$\nabla \cdot \vec{E} = \frac{\rho_c}{\varepsilon_0} \quad (2.4)$$

The functions driving change in this system are $\rho_c(\vec{r}, t)$, the electric charge density, and $j(\vec{r}, t)$, the electric current density. We also have μ_0 , the free-space magnetic permeability (in henry m^{-1}); ε_0 , the free-space permittivity; and c , the speed of light.

These equations give us a way to reason about the electric and magnetic fields if we're given some descriptor for the current we're passing through some medium. We are now to introduce the fluid dynamics component to our system. A fluid's mass density can be given by summing over the effects of individual "species" of particles (e.g. electrons) in the fluid:

$$\rho_c = \sum_{\sigma} m_{\sigma} n_{\sigma}$$

and its current density similarly:

$$\vec{j}(\vec{r}, t) = \sum_{\sigma} n_{\sigma} q_{\sigma} \vec{u}_{\sigma}$$

where σ describes a particle species, m_{σ} describes its mass, n_{σ} its number density (a measurement of concentration for the given particle species in a pre-defined

volume – akin to Avogadro’s constant), q_σ describes its electric charge, and \vec{u}_σ the mean velocity of this species of particle in the fluid.

Fluid Dynamics

Notation 2.2. A common simplification in notation for fluids is made in using the *Lagrangian derivative*, given:

$$\frac{D}{Dt} = \left(\frac{\partial}{\partial t} + \vec{u} \cdot \nabla \right)$$

It describes the total change in a volume within a fluid as it moves throughout said fluid. It is essentially a change in reference frame for a derivative - where a regular derivative might describe, for example, how a particle moves with respect to time in its surroundings, its Lagrangian will take into account the motion of the fluid the particle is immersed in as well.

We begin with conservation of mass, also known as the “continuity equation”.

Definition 2.3 (The continuity equation). The below relates how mass density, ρ , changes with respect to the motion of a fluid element.

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \vec{u}) \quad (2.5)$$

The derivation of the above comes from a surface integral over a volume with an outward and inward flux, and an application of Gauss’ flux law. For a full derivation, see pg. 19 - 21 of [25].

Remark 2.4. The above is a PDE with four variables: ρ , the mass density of the medium, and \vec{u} , the velocity of the fluid. This renders the system not closed, and thus too general for an analytic solution - we have more unknowns than we have equations [25]. Later we will introduce other equations to our system to apply more restrictions, and make assumptions about the physicality of the system which will reduce these dependencies, and make it determined (“closed”).

Remark 2.5. We can rewrite (2.5) with a Lagrangian frame of reference, as

$$\frac{D}{Dt} \rho = -\rho \nabla \cdot \vec{u} \quad (2.6)$$

$$\iff \frac{D}{Dt} \rho + \rho \nabla \cdot \vec{u} = 0 \quad (2.7)$$

Equation (2.5) (and equivalently (2.7)) tells us that the mass of our fluid is conserved for motion of a volume element of our fluid - one assumption we make for our model. Next we'll discuss fluid motion as described by Newton for a fluid element:

Definition 2.6 (Newtonian Fluid Motion). Newton's law for a fluid specifies:

$$\rho \frac{D}{Dt} \vec{u} = \vec{F} \quad (2.8)$$

where $\rho(\vec{r}, t)$ is the mass density of the fluid, $\vec{u}(\vec{r}, t)$ describes the velocity of the fluid element, and $\vec{F}(\vec{r}, t)$ describes the force per unit volume acting on the fluid element [25].

The forces acting on particles within a fluid can be split into two types

- Gravitational

Here, $\vec{F}_g = \rho \vec{g}$, where \vec{g} is the gravitational acceleration. This should hark back to high school physics, though note this is a vector here as we care about the direction gravity accelerates the fluid element in, and relativistic effects can be an important consideration for high mass systems. This is more relevant for cases that you are using the MHD equations to describe the dynamics of large systems, such as a star. Unsurprisingly, this is less relevant for our case of plasmas within relatively minuscule Tokamaks.

- Electromagnetic

This is the interesting part for us. As we assume our fluid is capable of conducting electricity (it's a plasma after all), there are electromagnetic forces operating within the fluid that affect the behaviour of the particles that the fluid consists of. The electromagnetic forces themselves can be split into two types, the **electric force** given by $\vec{F}_q = \rho_c \vec{E}$, and the **Lorentz force**, given $\vec{F}_L = \vec{j} \times \vec{B}$ (where $\vec{B}(\vec{r}, t)$ describes the magnetic field)

Taking these forces into account, we can describe the motion of an element of our fluid moving with velocity \vec{u} via:

$$\rho \frac{D}{Dt} \vec{u} = \vec{j} \times \vec{B} + \rho_c \vec{E} - \nabla p + \rho \vec{g} \quad (2.9)$$

Remark 2.7. Here the $\rho \vec{g}$ term could be abstracted further away into a stress tensor, as described by equation 4.20 of [25]. These pressures are largely negligible

when dealing with the scale we do in Tokamak plasmas however, and are thus ignored. We will soon drop the gravitational consideration as well anyway, but include it here now for completeness.

Remark 2.8. The equation we have introduced is a function of six variables in its complete form (with stress tensor included), though even in this form we still have more variables than we do equations (1). Similar to before, this is thus not constrained sufficiently to consider it a closed system.

We next relate a plasma's pressure to its motion. We will simply present it here, though important notes in its derivation are that we assume the plasma behaves as an ideal gas (which is to say the only interaction between particles within the plasma are via elastic collisions with each other, or the boundaries of the container it is contained within). This is equivalent to saying that energy in the system depends only on the pressure. Thus, the energy equation is given:

Definition 2.9 (Energy Equation). Where $\vec{p}(\vec{r}, t)$ describes the pressure of our fluid:

$$\frac{D}{Dt} p = -\gamma p \nabla \cdot \vec{u} + (\gamma - 1) \left[-\nabla \cdot \vec{q} + \vec{\Pi} : \nabla \vec{u} + \eta \vec{J}^2 \right] \quad (2.10)$$

where γ describes “adiabatic index” (a known constant for plasmas), \vec{q} is the heat flux through the boundary of the volume; η is the electrical resistivity of the fluid; and $\vec{\Pi}$ is the viscous stress tensor (the component which we replaced with $\rho \vec{g}$ earlier), and will soon ignore again.

The equations we've looked at constitute what are known as the fluid equations:

Definition 2.10 (Fluid Equations).

$$\frac{D}{Dt} \rho + \nabla \cdot \rho \vec{u} = 0 \quad (2.11)$$

$$\rho \frac{D}{Dt} \vec{u} = \vec{j} \times \vec{B} + \rho_c \vec{E} - \nabla p + \rho \vec{g} \quad (2.12)$$

$$\frac{D}{Dt} p = -\gamma p \nabla \cdot \vec{u} + (\gamma - 1) \left[-\nabla \cdot \vec{q} + \vec{\Pi} : \nabla \vec{u} + \eta \vec{j}^2 \right] \quad (2.13)$$

As reiterated a couple times now, these equations form an unclosed system, and are thus undetermined. To resolve this we introduce some constraints that come from electrodynamic forces, and a couple other relations that lead to a closed system.

Electrodynamics

As it currently stands, the input variables for the fluid equations are $\rho(\vec{r}, t)$ and $p(\vec{r}, t)$. We note that the electric field \vec{E} and the magnetic field \vec{B} are generated by the electric charge density, ρ_c , and the current density \vec{j} . This is where Maxwell's equations come into play. By combining Maxwell's equations with the fluid equation given above, we achieve the MHD model. The only piece to our puzzle missing is to tie the motion of the fluid (through \vec{u}) to the behaviour of the electric and magnetic fields. This is done via *Ohm's law*:

$$\vec{E} + \vec{u} \times \vec{B} = \eta \vec{j} \quad (2.14)$$

Remark 2.11. Note that the above is technically a lie, as it does not take into account relativistic effects, though for simplicity our MHD model ignores these.

Definition 2.12 (MHD Equations). The MHD equations can then be summarised:

$$\frac{D}{Dt} \rho = -\rho \nabla \cdot \vec{u} \quad (2.15)$$

$$\rho \frac{D}{Dt} \vec{u} = -\nabla p + \vec{j} \times \vec{B} + \nabla \cdot \Pi \quad (2.16)$$

$$\frac{D}{Dt} p = -\gamma p \nabla \cdot \vec{u} + (\gamma - 1) \left[-\nabla \cdot \vec{q} + \vec{\Pi} : \nabla \vec{u} + \eta \vec{j}^2 \right] \quad (2.17)$$

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \quad (2.18)$$

$$\mu_0 \vec{j} = \nabla \times \vec{B} \quad (2.19)$$

$$\vec{E} + \vec{u} \times \vec{B} = \eta \vec{j} \quad (2.20)$$

Remark 2.13. The MHD equations as presented above constitute 14 equations with 27 unknowns. The breakdown is as such:

- ρ : 1 unknown
- \vec{u} : 3 unknowns
- p : 1 unknown
- $\vec{\Pi}$: 9 unknowns
- \vec{j} : 3 unknowns
- \vec{B} : 3 unknowns

- \vec{E} : 3 unknowns
- \vec{q} : 3 unknowns
- η : 1 unknown

The above is obviously insufficiently constrained for purposes of identifying a solution. We will skip a large amount of the work required to reduce the above to a closed system, though for details see lecture 7 of [25]. For now, we will comment on two reduced MHD models:

Resistive MHD

The resistive MHD model comes about by setting $\vec{q} = 0$ and $\vec{\Pi} = 0$, in other words saying that we have no external source for current, and . Here, we have $\eta \neq 0$ notably. The model is given:

Definition 2.14 (Resistive MHD).

$$\frac{D}{Dt}\rho = -\rho \nabla \cdot \vec{u} \quad (2.21)$$

$$\rho \frac{D}{Dt}\vec{u} = -\nabla p + \vec{j} \times \vec{B} \quad (2.22)$$

$$\frac{D}{Dt}p = -\gamma p \nabla \cdot \vec{u} \quad (2.23)$$

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \quad (2.24)$$

$$\mu_0 \vec{j} = \nabla \times \vec{B} \quad (2.25)$$

$$\vec{E} + \vec{u} \times \vec{B} = \eta \vec{j} \quad (2.26)$$

Remark 2.15. The most notable effect of resistive MHD is that allowing for electrons to diffuse allows the resulting magnetic field lines to reconnect, which leads to breaks in the magnetic field line topology. This can lead to the generation of fast particles, i.e., runaway electrons.

Ideal MHD

The ideal MHD equations are one step removed from the resistive MHD model — in fact they are equivalent, only for the ideal MHD case we also ignore resistivity. Thus, set $\eta = 0$, and we obtain the ideal MHD equations:

Definition 2.16 (Ideal MHD).

$$\frac{D}{Dt}\rho = -\rho \nabla \cdot \vec{u} \quad (2.27)$$

$$\rho \frac{D}{Dt}\vec{u} = -\nabla p + \vec{j} \times \vec{B} \quad (2.28)$$

$$\frac{D}{Dt}p = -\gamma p \nabla \cdot \vec{u} \quad (2.29)$$

$$\frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E} \quad (2.30)$$

$$\mu_0 \vec{j} = \nabla \times \vec{B} \quad (2.31)$$

$$\vec{E} + \vec{u} \times \vec{B} = 0 \quad (2.32)$$

Remark 2.17. In doing this, we have removed the dependency on $\vec{\Pi}$, η and \vec{q} , which accounts for 13 unknowns. This brings the total number of equations to 14 (or 8 if you make a substitution for Ohm's law) with 14 (8) unknowns. Thus, under the ideal MHD model, the system is closed.

Remark 2.18. Equation (2.32) is the reduced Ohm's law equation. This presentation of it is a result known as the *frozen flux condition*. Nominally, it tells us that $\frac{\partial \Psi}{\partial t} = 0$ (we will introduce Ψ shortly, but for now take it to be the “magnetic flux function” (which it is) without any context). This is to say that the magnetic flux through any co-moving closed circuit is constant, in other words, field lines are to some extent “attached” to the fluid they are in (read: plasma), and similarly the plasma cannot move across the magnetic field (but the fluid can move along the field lines).

We'll present the ideal MHD equations here again but with their associated names for ease of reference:

Continuity Equation $\frac{D}{Dt}\rho = -\rho \nabla \cdot \vec{u}$	Momentum Equation $\rho \frac{D}{Dt}\vec{u} = -\nabla p + \vec{j} \times \vec{B}$	Energy Equation $\frac{D}{Dt}p = -\gamma p \nabla \cdot \vec{u}$
Maxwell-Faraday Equation $\frac{\partial \vec{B}}{\partial t} = -\nabla \times \vec{E}$	Ampere's Law $\mu_0 \vec{j} = \nabla \times \vec{B}$	Ohm's Law $\vec{E} + \vec{u} \times \vec{B} = 0$

Resistivity Note: There is one specific assumption made in ideal MHD that we wish to highlight, and that is the assumption of negligible resistivity within the plasma. In a real plasma we would expect imperfections in the way it conducts, and would expect collisions between particles in the plasma to generate a type of “friction” between each other that inhibits conductivity. When this

occurs, the magnetic field lines that are driven by the current will perturb with respect to some diffusion law, leading to changes in magnetic field line topology. These effects occur over some period of time that is dependent on the plasma's composition and geometry of the plasma - thus any plasma model which wishes to accurately represent the dynamics of the electric and magnetic field as a function of time will do well to take these effects into account, especially when it is over a "sufficient" time scale (with special consideration for what "sufficient" actually entails here). It is possible that the timescales on which a simulation is run are too small for resistive effects to meaningfully change field line topology, though this is more a question of acceptable error and initial conditions. In this thesis (as we'll see) we seek to approximate a time evolution, but do so with a series of equilibria, which as such do not take into account resistive effects. We will make note on this again when we come to discuss of our perturbation approach, and similarly when we present our simulated current reversals.

The ideal MHD equations as we stated are what govern many approximations to behaviour of plasmas in general. These in their current form of course make no assumptions about the geometry of the medium in which a plasma exists, nor any properties which may be associated with such a scenario. Next we will take these equations and confine them to the geometry of an axisymmetric Tokamak "reactor", and derive the Grad-Shafranov equation, a model which will let us reason about the poloidal magnetic field structure given some prescribed geometry and conditions.

2.2 Grad-Shafranov Equation

The Grad-Shafranov Equation (GSE) concerns a two-dimensional, axisymmetric toroidal plasma under the assumptions of ideal MHD. It is a nonlinear, elliptic partial differential equation (PDE) which presents with a cylindrical coordinate system [26]. We will first begin with some basic definitions surrounding PDEs, partly because the author is not formally educated in them and wishes to demonstrate the amount they have learned over this last year, but largely, realistically, for the obligatory Evans citation. We will then introduce the GSE, and explore some of its properties.

Definition 2.19 (Partial Differential Equation). An expression of the form

$$F(D^k u(x), D^{k-1} u(x), \dots, Du(x), u(x), x) = 0 \quad (x \in U) \quad (2.33)$$

is called a k^{th} -order Partial Differential Equation where

$$F : \mathbb{R}^{n^k} \times \mathbb{R}^{n^{k-1}} \times \cdots \times \mathbb{R}^n \times \mathbb{R} \times U \rightarrow \mathbb{R}$$

is provided, and the function

$$u : U \rightarrow \mathbb{R}$$

is an unknown [10].

To “solve” a PDE is to find such a function u , subject to any boundary conditions Γ in ∂U that may exist. A common exercise in the study of PDEs is to simply show existence and uniqueness of a solution to a PDE, with no intention of actually identifying the function itself. In our case, luckily, the GSE has identified solutions for certain assumptions (boundary conditions) imposed on it, one of which we will exploit later in the Grad-Shafranov-Helmholtz equation. Before that however, we’ll look at definitions for PDE properties that are associated with the GSE:

Notation 2.20. Evans uses the notation $D^k u$ to mean the $k \times k$ matrix of k^{th} order partial derivatives of u . While we use this when presenting the definitions below, we do not follow this convention for the rest of the thesis, reserving D for the Lagrangian derivative instead.

Definition 2.21 (PDE Linearity Classifications). Let a_α be a function such that $|\alpha| \leq k$, where k is the order of a system of PDEs. Let f be another function. Then:

1. “Linear”

A PDE as given by (2.33) is said to be “linear” if it is of the form

$$\sum_{|\alpha| \leq k} a_\alpha(x) D^\alpha u = f(x)$$

2. “Semilinear”

A PDE is said to be “semilinear” if it is of the form

$$\sum_{|\alpha|=k} a_\alpha(x) D^\alpha u + a_0(D^{k-1}u, \dots, Du, u, x) = 0$$

3. “Quasilinear”

A PDE is said to be “quasilinear” if it is of the form

$$\sum_{|\alpha|=k} a_\alpha (D^{k-1}u, \dots, Du, u, x) D^\alpha u + a_0 (D^{k-1}u, \dots, Du, u, x) = 0$$

4. “Nonlinear”

A PDE is said to be “nonlinear” if it depends nonlinearly on the highest order derivatives, i.e. it is of the form

$$\sum_{|\alpha|=k} a_\alpha (D^\alpha u, D^{k-1}u, \dots, Du, x) = 0$$

where the D^α term introduces nonlinearity by a_α .

Definition 2.22 (Elliptic PDE). A second-order PDE with presentation

$$F[u] = F(\cdot, u, Du, D^2u)$$

where $\Gamma = U \times \mathbb{R} \times \mathbb{R}^n \times \mathbb{S}^n$, $F : \Gamma \rightarrow \mathbb{R}$, and u is a solution to $F[u] = 0$ is said to be elliptic at the point $\gamma = (x, z, p, r) \in \Gamma$ if the matrix

$$F_r(\gamma) = [F_{ij}(\gamma)] := [F_{r_{ij}}] > 0$$

is true, i.e., that the matrix $F_r(\gamma)$ is positive definite for $r \in U$ [28]. Furthermore, the operator F is said to be *elliptic* if this holds true for the set

$$\Gamma_u = \{(x, u(x), Du(x), D^2u(x)) : x \in U\}$$

Remark 2.23. Note that the above representation of a second-order PDE, while not how Evans presents it, is a more abstract way of depicting the same thing. Here we can define an operator F to be in terms of our solution u and its first and second order derivatives, instead of as a decoupled statement of otherwise disconnected expressions. This lets us reason about the PDE as a whole more directly, as we have above by using the positive definiteness of the resulting matrix at a point to define ellipticity, instead of having to separately define the form of a second-order PDE and reason about a matrix explicitly in terms of the solution.

Often in our efforts we will deal not with a single PDE, but rather a system of PDEs, and so we formalise this concept as well:

Definition 2.24 (System of PDEs). An expression of the form

$$\mathbf{F}(D^k u(x), D^{k-1} u(x), \dots, Du(x), x) = 0 \quad (x \in U) \quad (2.34)$$

is called a k^{th} -order system of Partial Differential Equations where

$$\mathbf{F} : \mathbb{R}^{mn^k} \times \mathbb{R}^{mn^{k-1}} \times \dots \times \mathbb{R}^{mn} \times \mathbb{R}^m \times U \rightarrow \mathbb{R}^m$$

is provided, and the function

$$\mathbf{u} : U \rightarrow \mathbb{R}^m, \quad \mathbf{u} = (u_1, \dots, u_m)$$

In fact we have already seen an example of the above in our derivation of the MHD equations ((2.15) - (2.20)), the result of which was a system of PDEs; a system which was determined by virtue of there being as many unknown functions (u_i) as there were PDEs in the system.

We are now armed with the necessary tools to introduce the Grad-Shafranov Equation. Here we present the definition of the GSE as given by Wesson [31].

Definition 2.25 (Grad-Shafranov Equation). Let (R, ϕ, z) be the cylindrical coordinate system we take, where R is the major radial offset, ϕ is the toroidal angle, and z is the poloidal vertical offset from the centre of the cross section. $p(\Psi)$ and $f(\Psi)$ are two flux functions of Ψ , the poloidal magnetic flux function, and μ_0 is the vacuum magnetic permeability constant. Then the Grad-Shafranov Equation is given:

$$R \frac{\partial}{\partial R} \frac{1}{R} \frac{\partial \Psi}{\partial R} + \frac{\partial^2 \Psi}{\partial z^2} = -\mu_0 R^2 \frac{\partial p(\Psi)}{\partial \Psi} - \mu_0^2 f(\Psi) \frac{\partial f(\Psi)}{\partial \Psi} \quad (2.35)$$

Remark 2.26. The equation can be restated more concisely in terms of the elliptic operator

$$\Delta^* \Psi = R \nabla \cdot \left(\frac{1}{R} \nabla \Psi \right)$$

as:

$$\Delta^* \Psi = -\mu_0 R^2 p'(\Psi) - f(\Psi) f'(\Psi)$$

where prime notation denotes partial derivative with respect to Ψ .

Corollary 2.27. Associated with the GSE are explicit forms for the magnetic

field \vec{B} and current \vec{j} :

$$\vec{B} = \frac{1}{R} (\nabla\Psi \times \hat{e}_\phi) + \frac{f(\Psi)}{R} \hat{e}_\phi \quad (2.36)$$

$$\mu_0 \vec{j} = \frac{1}{R} \frac{\partial f(\Psi)}{\partial \Psi} (\nabla\Psi \times \hat{e}_\phi) - \left[\frac{\partial}{\partial R} \left(\frac{1}{R} \frac{\partial \Psi}{\partial R} + \frac{1}{R} \frac{\partial^2 \Psi}{\partial z^2} \right) \right] \hat{e}_\phi \quad (2.37)$$

where here \hat{e}_d is the unit vector in the direction d , as is convention. In their poloidal (subscript p) and toroidal (subscript ϕ) components, these are then explicitly just:

$$\begin{aligned} \vec{B}_p &= \frac{1}{R} (\nabla\Psi \times \hat{e}_\phi) \\ \vec{B}_\phi &= \frac{f(\Psi)}{R} \\ \vec{j}_p &= \frac{1}{\mu_0 R} (\nabla f(\Psi) \times \hat{e}_\phi) \\ \vec{j}_\phi &= -\frac{1}{\mu_0 R} \Delta^* \Psi \end{aligned}$$

Remark 2.28. The flux function $f(\Psi)$ describes the poloidal flux per radian in ϕ . Of note, for an axisymmetric torus the total flux function would then simply be $2\pi\Psi$. The other flux function here, $p(\Psi)$, simply describes pressure inside the Tokamak [9].

Remark 2.29. We noted in our introduction of the PDE theory that the GSE was an elliptic, nonlinear PDE. We can see the ellipticity immediately by virtue of there being an elliptic operator present in Δ^* , and the nonlinearity we can gather from the fact there are non-linear coefficients of partial derivatives with respect to Ψ , for example, the $-f(\Psi)f'(\Psi)$ in the elliptic operator form of the GSE.

The GSE equation is not without its faults, and is, in large, a considerable simplification of the actual physical behaviour of a plasma within a Tokamak – as my supervisor Matthew would say, it is very much so a spherical cow model. To explore why this is so, we'll explain the components of the GSE and the assumptions that are made in reaching it through a derivation with accompanying annotations below.

2.2.1 Derivation

Here we present a derivation of the Grad-Shafranov Equation. We begin by assuming the ideal MHD conditions, which immediately gives us that, in our plasma model:

- The plasma particle distribution approximately abides a Maxwellian distribution
- Resistivity is negligible
- There is no external (to the system) current source
- The topology of the magnetic field is fixed relative to the fluid, i.e., the magnetic field and fluid move together (the frozen-in flux theorem)
- Any velocities considered are non-relativistic (notably, this includes electron velocities)

Remark 2.30. We will largely follow Wesson's derivation of the Grad-Shafranov equation given in [31], however will correlate this with other sources too. Additionally, we will deviate from some standard notations (such as that used for major radius offset, substituting x for R) as this will be more consistent with the presentation of the Grad-Shafranov-Helmholtz given by Wang which will be the subject of our attention in the next section [30].

We make a few extra assumptions too. First, we take our fluid to be incompressible, which gives us the relation $\nabla \cdot (\rho \vec{u}) = 0$. As the Grad-Shafranov equation is also one describing an equilibrium state of the plasma, we take our fluid to be stationary, i.e., $\vec{u} = 0$. From this we gather that $\rho \frac{D}{Dt} \vec{u} = 0$, which transforms out momentum equation from

$$\rho \frac{D}{Dt} \vec{u} = -\nabla p + \vec{j} \times \vec{B} \quad (2.38)$$

into a force-balance equation relating pressure, current, and the magnetic field.

$$\nabla p = \vec{j} \times \vec{B} \quad (2.39)$$

We also add the restriction that the magnetic field is divergence free, i.e.

$$\nabla \cdot \vec{B} = 0 \quad (2.40)$$

which is to say there are no sources or sinks of magnetism within the field itself (to not make this assumption would be to suggest the existence of a magnetic monopole within our plasma). We can combine these two relations with Ampere's law, which remains unchanged, to have the set of equations

$$\nabla \cdot \vec{B} = 0 \quad (2.41)$$

$$\mu_0 \vec{j} = \nabla \times \vec{B} \quad (2.42)$$

$$\nabla p = \vec{j} \times \vec{B} \quad (2.43)$$

Remark 2.31. Equation 2.41, with 2.43, also tells us that both $\vec{B} \cdot \nabla p = 0$ and $\vec{j} \cdot \nabla p = 0$. Thus, there is no pressure gradient across magnetic field lines, and current lines also lie within magnetic surfaces.

We then introduce the poloidal magnetic flux function, Ψ , which describes the amount of poloidal magnetic flux per radian in the ϕ (toroidal) direction. At the same time we introduce our cylindrical coordinate system. We place specific emphasis on the fact we are simplifying our torus geometry to that of a cylinder here - we do this under the assumption of axisymmetry, which says that our system is invariant with respect to ϕ . We take x to be the major radius offset, with R_0 being the major radius of the torus. We also let z describe the vertical offset within our poloidal cross section, which is bound by a , the minor radius of the torus. The poloidal magnetic field can then be related to our flux function by components in the x and z direction of our cross section

$$\vec{B}_x = -\frac{1}{x} \frac{\partial \Psi}{\partial z} \quad (2.44)$$

$$\vec{B}_z = \frac{1}{x} \frac{\partial \Psi}{\partial x} \quad (2.45)$$

which, combined, lets us write:

$$\vec{B}_p := \vec{B}_x \hat{e}_x + \vec{B}_z \hat{e}_z \quad (2.46)$$

$$= -\frac{1}{x} \frac{\partial \Psi}{\partial z} \hat{e}_x + \frac{1}{x} \frac{\partial \Psi}{\partial x} \hat{e}_z \quad (2.47)$$

which is equivalent to writing

$$\vec{B}_p = \frac{1}{x} (\nabla \Psi \times e_\phi) \quad (2.48)$$

By a similar argument, we can decompose the poloidal current in terms of a

function f , the poloidal current flux function.

$$\vec{j}_x = -\frac{1}{x} \frac{\partial f}{\partial z} \quad (2.49)$$

$$\vec{j}_z = \frac{1}{x} \frac{\partial f}{\partial x} \quad (2.50)$$

By comparison, Ampere's equation (equation 2.42), combined with 2.48, gives us the equivalent relations

$$\vec{j}_x = -\frac{1}{\mu_0} \frac{\partial \vec{B}_\phi}{\partial z} \quad (2.51)$$

$$\vec{j}_z = \frac{1}{\mu_0} \frac{1}{x} \frac{\partial}{\partial x} \left(x \vec{B}_\phi \right) \quad (2.52)$$

which gives us an expression for f :

$$f = \frac{x \vec{B}_\phi}{\mu_0} \quad (2.53)$$

At this point, we are able to describe the force balance equation (2.43) in terms of f, Ψ and p . Next, we want to show that, in actuality, our poloidal current density flux function f and our pressure density function p are both functions of Ψ . Recalling we have $\vec{j} \cdot \nabla p = 0$, we can substitute our relations for the poloidal current density given in equations (2.49, 2.50) to write our pressure in terms of poloidal current:

$$\begin{aligned} \vec{j} \cdot \nabla p &= 0 \\ \left(-\frac{1}{x} \frac{\partial f}{\partial z} \hat{e}_x + \frac{1}{x} \frac{\partial f}{\partial x} \hat{e}_z \right) \cdot \nabla p &= 0 \\ \left(-\frac{1}{x} \frac{\partial f}{\partial z} \hat{e}_x + \frac{1}{x} \frac{\partial f}{\partial x} \hat{e}_z \right) \cdot \left(\frac{\partial p}{\partial x} \hat{e}_x + \frac{\partial p}{\partial z} \hat{e}_z \right) &= 0 \\ \frac{1}{x} \left(\frac{\partial f}{\partial x} \frac{\partial p}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial p}{\partial x} \right) &= 0 \\ \frac{\partial f}{\partial x} \frac{\partial p}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial p}{\partial x} &= 0 \end{aligned}$$

The last line of the above is equivalent to the statement

$$\nabla f \times \nabla p = 0$$

which tells us that f is constant with respect to p . This tells us that f is a

function of p , i.e., $f = f(p)$. We already had the result earlier that p is constant with respect to poloidal magnetic field lines, which tells us additionally that $p = p(\Psi)$, which by transitivity gives us that $f = f(\Psi)$. Thus, $p(\Psi)$ and $f(\Psi)$ are functions of the poloidal magnetic flux. This gives us explicitly a way to describe the force balance function in terms of flux functions. Thus, if we expand:

$$\vec{j} \times \vec{B} = \nabla p$$

$$\vec{j}_p \times (\vec{B}_\phi \hat{e}_\phi) + (\vec{j}_\phi \hat{e}_\phi) \times \vec{B}_p = \nabla p$$

we can substitute the expressions for \vec{B}_p and \vec{j}_p

$$\vec{B}_p = \frac{1}{x} (\nabla \Psi \times \hat{e}_\phi)$$

$$\vec{j}_p = \frac{1}{x} (\nabla f \times \hat{e}_\phi)$$

into the expanded force balance equation:

$$\vec{j}_p \times (\vec{B}_\phi \hat{e}_\phi) + (\vec{j}_\phi \hat{e}_\phi) \times \vec{B}_p = \nabla p$$

$$\left(\frac{1}{x} (\nabla f \times \hat{e}_\phi) \right) \times (\vec{B}_\phi \hat{e}_\phi) + (\vec{j}_\phi \hat{e}_\phi) \times \left(\frac{1}{x} (\nabla \Psi \times \hat{e}_\phi) \right) = \nabla p$$

Noting that the toroidal components of the flux functions are non-existent, i.e. that $\hat{e}_\phi \cdot \nabla \Psi = 0$ and $\hat{e}_\phi \cdot \nabla f = 0$

$$-\frac{\vec{B}_\phi}{x} \nabla f + \frac{\vec{j}_\phi}{x} \nabla \Psi = \nabla p \quad (2.54)$$

Recalling that f and p are functions of our poloidal magnetic flux function Ψ , we have the relations

$$\nabla f(\Psi) = \frac{\partial f}{\partial \Psi} \nabla \Psi$$

$$\nabla p(\Psi) = \frac{\partial p}{\partial \Psi} \nabla \Psi$$

Substituting these into 2.54 gives us

$$\vec{j}_\phi = x \frac{\partial p}{\partial \Psi} + \vec{B}_\phi \frac{\partial f}{\partial \Psi} \quad (2.55)$$

then, recalling that $f = (x \vec{B}_\phi)/\mu_0$, we can remove the toroidal magnetic field

component

$$\vec{j}_\phi = x \frac{\partial p}{\partial \Psi} + \frac{\mu_0}{x} f(\Psi) \frac{\partial f}{\partial \Psi} \quad (2.56)$$

Lastly, we want to remove the toroidal current component. We can do this by expanding Ampere's law again, where we also use our expressions for \vec{B}_p and \vec{j}_p . To save using more space in this derivation we'll leave this as an exercise, but the result is the expansion

$$-\mu_0 x \vec{j}_\phi = x \frac{\partial}{\partial x} \left(\frac{1}{x} \frac{\partial \Psi}{\partial x} \right) + \frac{\partial^2 \Psi}{\partial z^2} \quad (2.57)$$

Substituting 2.57 into 2.56 gives us the Grad-Shafranov equation:

$$x \frac{\partial}{\partial x} \left(\frac{1}{x} \frac{\partial \Psi}{\partial x} \right) + \frac{\partial^2 \Psi}{\partial z^2} = -\mu_0 x^2 \frac{\partial p}{\partial \Psi} - \mu_0^2 f(\Psi) \frac{\partial f}{\partial \Psi} \quad (2.58)$$

which is in the form we first saw it in in equation 2.35.

Remark 2.32. An important result here is that the total magnetic field can be expressed as

$$\vec{B} = \vec{B}_p \hat{e}_\phi + \vec{B}_\phi \hat{e}_\phi = \nabla \Psi \times \nabla \phi + f(\Psi) \nabla \phi$$

which tells us that to describe the magnetic field, the only information we require is the poloidal magnetic flux function, and the poloidal current density function (or equivalently, poloidal pressure density).

2.3 Runaway Electrons

Runaway electrons describe electrons with energy such that they are “relativistic” - when their velocity is comparable to the speed of light. At such high speeds, these electrons move freely from the confines of the magnetic field, making them difficult to control, and unwieldy. They can have destructive effects on our ability to confine a plasma, both in terms of the dynamics that are introduced by their presence, and in the destruction of the medium we rely on to perform such confinement - the reactor itself. One of the problems this thesis seeks to address is the presence of runaway electrons in AC mode tokamaks, which is conducive to understanding how to increase confinement time.

In ISTTOK, it is noted that runaway electron populations can contribute a substantial amount to the plasma we have present. It is estimated that up to 30% of the plasma current can be attributed to runaway electron current, and that up

to 1% of the plasma's density itself can be composed of runaway electrons [24].

2.3.1 Regular Electron Movement

In the presence of a magnetic field, a charged particle experiences a force and thus moves [31]. Specifically, for a particle of mass m_j and charge e_j moving through a magnetic field \vec{B} , its motion can be described

$$m_j \frac{d\vec{v}}{dt} = e_j \vec{v} \times \vec{B}$$

In a uniform magnetic field with no electric field, a charged particle will move through its medium with an oscillatory, but consistent, motion. The radius of this oscillatory movement is referred to as the “Lamor radius”, r_L . However, any change in the uniformity of the magnetic field will induce forces on the electron which remove this force balance, and will cause the electron to start to drift.

Remark 2.33. Note that the above statement for particle drift is not just applicable to electrons, but generally to charged particles in the presence of a magnetic field. This thus also affects the positively charged ions in our plasma, which also experience a drifting effect. Here, we are more concerned about these effects on electrons however.

There are two primary drift types that concern us: that purely with respect to the divergence of the magnetic field, and that for which we have perpendicular electric and magnetic fields.

Electric Field drift

Here, we have our motion equation (similar to above)

$$m_j \frac{d\vec{v}}{dt} = e_j (\vec{E} + \vec{v} \times \vec{B})$$

if we make the choice that the z is parallel to the magnetic field, and the y axis parallel to the electric field, then we can separate this into

$$\begin{aligned} m_j \frac{dv_x}{dt} &= e_j v_y \vec{B} \\ m_j \frac{dv_y}{dt} &= e_j (\vec{E} - v_x \vec{B}) \end{aligned}$$

This set of equations have readily available solutions, telling us

$$v_x = v_{\perp} \sin(\omega_{cj} t) + \frac{E}{B}$$

$$v_y = v_{\perp} \cos(\omega_{cj} t)$$

From this we gather that there is a drift component to our electron, and that it still travels with a circular motion.

$\nabla \vec{B}$ drift

Here, a changing magnetic field induces a drift velocity on charged particles within its effect. It accelerates a particle in the direction perpendicular to both the magnetic field itself, and the direction in which the magnetic field is drifting. Thus, as a charged particle moves through a magnetic field that is not uniform (i.e. $\nabla \vec{B} \neq 0$), then its motion will be affected as the strength of the magnetic field changes [12]. The general drift due to this force is commonly given:

$$\vec{v}_d = \pm \frac{1}{2} v_{\perp} r_L \frac{\vec{B} \times (\nabla \vec{B})}{B^2}$$

where r_L describes the Larmor radius, and the sign out front is dependent on the charge of the particle (as the drift effect is opposite for electrons and cations).

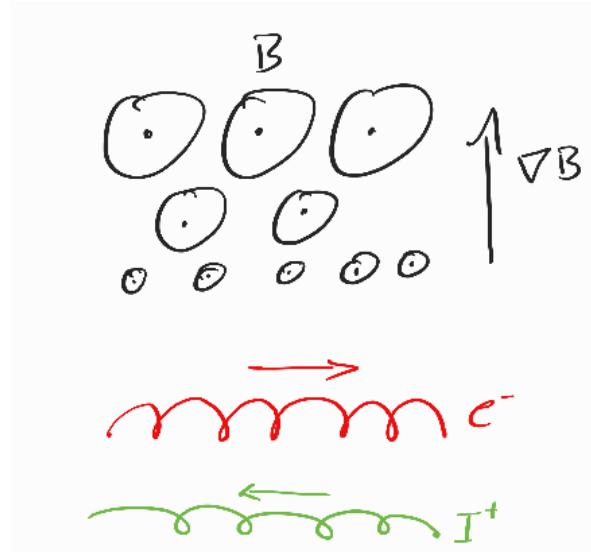


Figure 2.1: Circles with dots indicate the magnetic field \vec{B} is orientated into the page. This shows the effect of $\nabla \vec{B}$ drift on oppositely charged particles.

2.3.2 Generation

There are a number of mechanisms by which runaway electrons can be generated in a tokamak.

The simplest case is one we induce ourselves - in the warm up routine of a reactor such as ISTTOK a heating element is applied to the gas, which ionises it. The heating of the gas follows a Maxwellian distribution, which gives probability that some portion of the resulting plasma will consist of electrons which have been energised to relativistic states [2]. The Maxwellian distribution for electron energy densities can be given:

$$f_M(v) = \frac{n}{\sqrt{\pi}v_{\text{th}}} e^{-\frac{v^2}{v_{\text{th}}^2}}$$

where v_{th} describes the barrier for an electron being described thermal or not, and n is the number of electrons present. Velocities past this barrier are referred to as being in the “hot tail”, and are considered candidates for being runaway. Using this we can define an electric field for which all thermal electrons will be accelerated into becoming runaway electrons:

Definition 2.34 (Dreicer field). A Dreicer field is an electric field for which all thermal electrons, defined to have at least velocity v_{th} , will be accelerated into runaway electrons.

$$E_D = \frac{4\pi e^3 n \ln(\Lambda)}{mv_{\text{th}}^2}$$

The heating element which is the cause of this is affectionately titled a “heating lamp”, though one should likely hesitate to leave it on after you go to bed. As such, no matter our efforts in dampening REs, and in eliminating other processes which generate them, we will necessarily have some background population of runaway electrons. Notably, this will also increase in magnitude for reactors which have longer

A connected idea to this is that of runaway avalanches. If a collision takes place within the plasma between a high energy electron and particles in the gas, then a “Townsend avalanche” can occur if that collision results in the production of other high energy electrons. This avalanche effect is characterised by an exponential growth, which can be given:

$$\ln \frac{n_{\text{re}}}{n_{\text{seed}}} = \frac{e}{m_e \ln \Lambda} \frac{\Delta \Psi / (2\pi R)}{\sqrt{5 + Z_{\text{eff}}}}$$

where n_{re} is the number of resulting relativistic electrons, n_{seed} describes the number of incident runaway electrons, and Z_{eff} describes the effective background ion charge. Note that ISTTOK observes a presence of these “seed” runaway electrons after a ramp-down [18].

Another contribution is that of magnetic field topology. When there are “breaks” in the magnetic field line topology that form, electrons within the influence of these breaks can be accelerated to relativistic speeds. While the dynamics of this are more complicated than simply “topology changes, runaway electrons generated”, the occurrence of such a change in topology can be considered a type of runaway electron generation, and can be used to infer / support their generation if observed [2]. This change in topology could, for example, be characterised by the formation of an “X-point”, which is a point at which two previously distinct magnetic field regions join together to share a single point of flux density in common, which mathematically presents as a singularity, though physically acts as an unstable accelerator. The paper cited above by Boozer provides a good overview of the effects of magnetic field line topology on the generation of runaway electrons.

2.3.3 Detection

It’s all well and good to be know that runaway electrons are a possibility, but equally as important is being able to identify if you actually have runaway electrons. Unfortunately, we cannot just ask the plasma. As such, we must resort to a series of clever tricks to infer the existence of REs.

The first and most obvious of these tricks is to recall that, being electrons that are in motion, RE populations form a current. As we are able to measure current density profiles in our reactors, if we do so and observe there to be a significant, unexpected, increase in current density, we could infer that this is the consequence of the presence of an RE population. In a similar manner, we can also for fluctuations in our pressure density profile, as a current channel opening up (as happens with a stream of REs) ionises the gas it is travelling through, creating a density of plasma where it passes, which increases pressure.

Another characteristic behaviour observed for runaway electron populations is Cherenkov radiation, which presents when a burst of “hard X-rays” (HXR) appears [33]. We can directly observe and measure this phenomenon, and so it serves as a useful metric for measuring runaway electron presence, and population size.

Remark 2.35. Quick science fawning moment: Cherenkov radiation is an observ-

able phenomena presenting as a blue glow, and is the result of photons travelling faster than the speed of light! But this is only possible in certain situations. For example, in water, light waves propagate slower than the speed of light in a vacuum (the traditional value for c we're accustomed to seeing) due to the refractive nature of water. However, radioactive materials can generate beta particles which are emitted at close to the velocity of light in a vacuum, c , and thus travel faster than the light in the medium around it! [5]

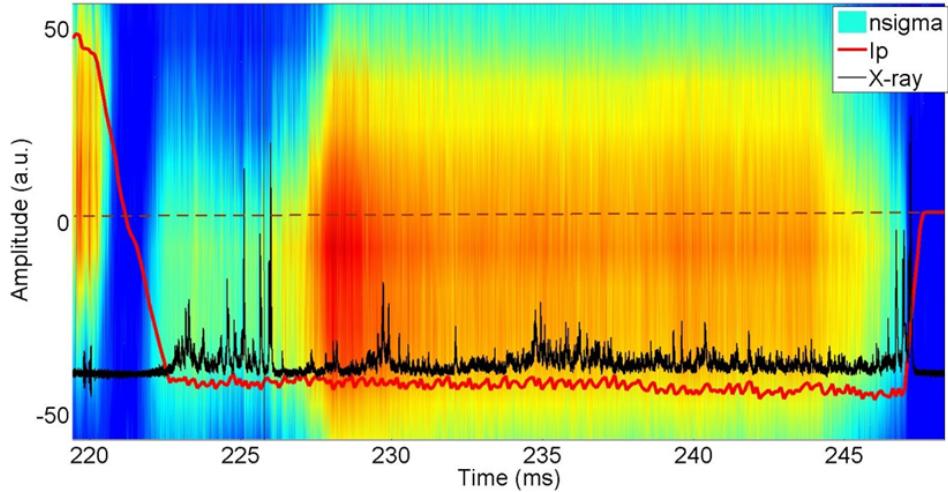


Figure 2.2: ISTTOK plasma pressure density profile with plasma current (I_p) and X-rays detected. As X-rays are an indicator of runaway electrons, this indicates a jump in REs during the ramp down phase (from about 220ms - 227ms), and also at the start of the next ramp up phase (from about 245ms) [18].

We should note that, despite the various options available to us for identifying runaway electrons, measurements can be somewhat fraught, and some populations may be mischaracterised in terms of strength, or missed entirely. For example, some instruments that are plasma-facing in a tokamak are directional by nature. This would be ok if your plasma current was oriented in a single direction, but as is in our case with a reactor operating in AC mode, the plasma current changes direction. As such, the HXR detector for example will miss radiation caused by RE populations travelling in a direction it can't detect.

Chapter 3

Current Reversal Theory

3.1 Perturbation Methods

3.1.1 What is a Perturbation Treatment?

When it comes to solving differential equations, the ideal scenario is that you find an analytic solution, ψ . This, however, is generally a difficult problem, and especially so in the case of differential equations. A great deal of research goes into trying to find analytic solutions to PDEs, which is no less true for the Grad-Shafranov equation and its variations. Though such efforts are often futile, or require making assumptions about properties of your solution which may not be representative of what you're trying to show, other means of obtaining the $\Psi(\vec{x})$ for $\vec{x} \in \Omega$ have been developed.

A traditional approach to this is to use some numerical code to approximate solution values. There are various codes which exist for this purpose specifically with regards to the GSE equation, such as SPEC [14]. These codes come in many flavours - some implement a particle simulation, which can provide precision with simulation, though are computationally costly. Others exploit various properties of the variant of GSE they target and may be computationally more feasible, but are restricted in scope, application, and/or reliability because of their assumptions.

A common problem that numerical methods seek to solve is that of time evolution. Simulations which evolve a system through time, for obvious reasons, do not discard the time dependency component in their derivation of the GSE. A common problem that numerical methods are employed to solve is that of time evolution. Simulations which evolve a system through time however often require

time dependence in the system they are trying to solve – something, if you recall from our derivation of the GSE, is not present in our model. This introduces a difficulty for us, as solutions to the GSE explicitly represent plasma equilibria, and so we expect no variation of the system with respect to time.

For this thesis, given the availability of an analytic solution to a variation of the GSE that specifically relates to current reversals, we employ a hybrid approach. In our simulation we wish to observe changes in the system through a current ramp down - this necessitates a time dependency component to our solution. However, there are a couple points to note:

- Time evolution simulations are expensive
- Existing literature on time evolution GSE do not take into account the possibility for current reversal
- The GSH variant we utilise (which has yet to be introduced) is not time dependent, i.e., was not derived taking into account a time component

There are a couple possible avenues we could explore from here. We could implement a numerical code for the time evolution, for example a particle simulation, however this is for all intents and purposes computationally infeasible given the hardware available (a beloved but dilapidated Thinkpad T440p). We could attempt to find an analytic solution to the GSE with time dependence (or an approximation to one), but given the mountain of research existing in this space and the relative inexperience of the author, this would likely be a futile task. The third, and most lucrative option, is to use an existing analytic solution to a variation of the GSE that accounts for current reversal, and explore if small changes to its state are sufficiently “good enough” approximations of our system that it can be used to make statements about experimental data.

This is where perturbation theory comes in. Intuitively we would expect a plasma to vary “smoothly” - nature rarely behaves in instantaneous ways. The analytic solution we will soon have describes the instantaneous state (a slice in time) of our plasma’s state. We may anticipate then that if we were to change something in our system by a “sufficiently small amount” (we will later comment on what exactly is meant by sufficiently small), that our system may react to that change of state in a way that approximates how it would if it were really varying through time.

Resistivity Note

When we derived the ideal MHD equations we compared them to the resistive MHD model, noting that the removal of resistivity would be an important simplification to note. Here we see its importance, as we are now allowing our system to vary with respect to time, which means that there exists some time scale for which resistive effects will begin to impact the behaviour of our plasma.

Our model assumes the ideal MHD assumptions however, which includes dropping resistivity, and so the effect of these resistive instabilities is lost in our model. The question of whether these effects are significant enough to affect the accuracy of our time evolution is one we hope to answer when we come to reviewing our simulations in chapter 5.

A perhaps more intuitive analogy is that of swimming. Imagine you are stationary sitting at the bottom of a pool, and you are looking at your arm extended out in front of you. Focus on the feeling of the water moving around your arm - this will emulate the resistivity that electrons would feel. If you move your arm abruptly, perhaps in a cutting motion, then you will feel the pressure of the water push against your arm, and that might affect how quickly you can move your arm, or if the current of the water is particularly strong, perhaps you notice your arm move not in the direction you intended. Perhaps you also have to use more energy to move your arm than you intended, so it doesn't go as far. Now imagine instead that you don't abruptly cut, but instead, very slowly, almost imperceptibly, move your arm through the water. In this case you may not be aware of any resistance at all - there is no sensation of resistance, your arm moves exactly as fast as you expect it to, and you use exactly the amount of energy to move it as you expected. However, if you do this for long enough, perhaps you will begin to notice the effects more and more. Or maybe by now you'll realise you're out of breath and need to resurface. An instructive graphic is below.



Figure 3.1: Demonstrative graphic for conducting a resistivity experiment at the bottom of a pool.

We will have the ability to determine the state of our system for some given parameters, and so the question becomes: can we vary our input in a way such that the behaviour of the analytic solution to the GSH is an approximation of a time evolution of the plasma. We are able to compare our results to experimental data, and then exploit our work to answer further questions.

3.1.2 Regular Perturbation Theory

At the heart of perturbation theory is the manipulation of functions by some small amount to achieve a goal. A familiar perturbation is the taylor expansion - if we have some function $f(x)$, then we can use n^{th} order derivatives to approximate the value of the function for some perturbation a

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x - a)^n$$

In a similar spirit, we can expand a function $f(x)$ around a point $x = a$ by some ε amount

$$f(a + \varepsilon) = f(a) + \varepsilon \frac{df}{dx} + \frac{1}{2} \varepsilon^2 \frac{d^2 f}{dx^2} + \dots$$

and it is this sentiment that drives perturbation theory - that some small change in input of a function can result in small changes in the output of that function. This is obviously not the case for all functions - immediately we see that this

fails for anything that is not differentiable, and in general if we wish to make an approximation to the n^{th} degree then we need our function to at least be $f \in C^n$ [32].

When it comes to PDEs, regular perturbation theory can be used to find solutions to unknown functions. The general process is [32]:

1. Set $\varepsilon = 0$, solve the resulting system (i.e., a solution for f_0 effectively)
2. Perturb the system by ε (i.e., go back and expand the system in terms of some $\varepsilon > 0$)
3. Let the (unknown) solution to the new perturbed system be given as

$$f_0 + \varepsilon f_1 + \varepsilon^2 f_2 + \dots$$

4. Expand the PDE in terms of ε . Collect like powers of ε and solve these as their own systems
5. Collate solutions for each f_i as found for their respective powers of ε - this is your approximation to the analytic solution to the PDE

Formally then we can state that, given a differential equation

$$F(D^k u(x), D^{k-1} u(x), \dots, Du(x), x) = 0$$

an expansion for u can be given

$$u(x) = u_0(x) + \varepsilon u_1(x) + \varepsilon^2 u_2(x) + \dots$$

If we were able to reason about all the derivatives of u (and they exist), then in the ideal case we can retrieve an analytic solution to the PDE. However, if instead the function is not C^∞ , or perhaps becomes too difficult to analytically solve, we could instead solve for an approximate solution by instead expanding, for example, to the second order:

$$u(x) = u_0(x) + \varepsilon u_1(x) + \varepsilon^2 u_2(x) + \mathcal{O}(\varepsilon^2)$$

Example

Here we will follow an example provided by [17], which emphasizes the utility of perturbation methods in solving differential equations.

Let a body have mass m with initial velocity v_0 . It moves in a straight line, however there is a resistive force opposing its motion with magnitude $av - bv^2$, where $v(\tau)$ describes the velocity of the object. We assume $b \ll a$ are constants, and Newton's law provides

$$m \frac{dv}{d\tau} = -av + bv^2$$

with the boundary condition

$$v(0) = v_0$$

We can make this system dimensionless by making a change of variables with $y = v/v_0$ and $t = \tau/(m/a)$, giving us the new system

$$\frac{dy}{dt} = -y + \varepsilon y^2$$

with boundary condition

$$y(0) = 1$$

where $\varepsilon \equiv (bv_0)/a \ll 1$. Now with our dimensionless model we can begin our perturbations. Instead of opting for an analytic solution, we'll use an approximation, so we will expand as such:

$$y(t) = y_0(t) + \varepsilon y_1(t) + \varepsilon^2 y_2(t)$$

and ignore any $\mathcal{O}(\varepsilon^3)$ terms. Substituting this into our dimensionless model gives

$$\begin{aligned} \frac{dy}{dt} &= -y + \varepsilon y^2 \\ y'_0 + \varepsilon y'_1 + \varepsilon^2 y'_2 &= -(y_0 + \varepsilon y_1 + \varepsilon^2 y_2) + \varepsilon (y_0 + \varepsilon y_1 + \varepsilon^2 y_2)^2 \end{aligned}$$

Then, we next collect similar powers of ε (representing different orders of approximation), which we observe is a series of linear ODEs:

$$\begin{aligned} y'_0 &= -y_0 \\ y'_1 &= -y_1 + y_0^2 \\ y'_2 &= -y_2 + 2y_0 y_1 \end{aligned}$$

The initial condition also tells us that $y_0(0) + \varepsilon y_1(0) + \varepsilon^2 y_2(0) = 1$, but, as we know $y_0(0) = 1$, we also get that $y_1(0) = y_2(0) = 0$. Thus, we have a series of

ODEs with initial conditions specified for each, which harks back to first year calculus courses. We readily have solutions available for each of the ODEs:

$$\begin{aligned}y_0(t) &= e^{-t} \\y_1(t) &= e^{-t} - e^{-2t} \\y_2(t) &= e^{-t} - e^{-2t} + e^{-3t}\end{aligned}$$

and thus, our approximate solution to y can be given

$$y(t) \approx e^{-t} + \varepsilon(e^{-t} - e^{-2t}) + \varepsilon^2(e^{-t} - e^{-2t} + e^{-3t})$$

As a form of validation, with omnipotent knowledge we can state that the actual solution to the original equation can be given as

$$y(t) = \frac{e^{-t}}{1 + \varepsilon(e^{-t} - 1)}$$

which has the Taylor series expansion (in ε):

$$y(t) = e^{-t} + \varepsilon(e^{-t} - e^{-2t}) + \varepsilon^2(e^{-t} - e^{-2t} + e^{-3t}) + \mathcal{O}(\varepsilon^3)$$

which is exactly what our perturbation approximation determined [17]. A cool demonstration, if we can say so ourselves!

While we won't be performing such an in depth perturbation treatment of the MHD equations themselves, it is this spirit of approximation via perturbation which we will carry through our work, and keep in mind when it comes to our simulations.

3.1.3 Ideal MHD Perturbation

At the beginning of this thesis we introduced it as being interdisciplinary in nature. Keeping in that spirit, at times here we will put on our physicists cap and seemingly arbitrarily remove terms we no longer wish to have. We kindly ask that any mathematicians reading these sections avert their gaze in such times so as to maintain sanity.

In chapter 2 we derived the ideal MHD equations, and their subsequent re-

duction used in GSE, given:

$$\vec{j} \times \vec{B} = \nabla p \quad (3.1)$$

$$\mu_0 \vec{j} = \nabla \times \vec{B} \quad (3.2)$$

$$\nabla \cdot \vec{B} = 0 \quad (3.3)$$

These by note are time independent. We wish to look at the effect of some small ($\varepsilon > 0$) time perturbation to the original system however, and so will return to their derivation, this time without discarding the time component. Consider the pre-GSE ideal MHD equations again, as given in ((2.27) - (2.32))

$$\begin{aligned} \frac{D}{Dt} \rho &= -\rho \nabla \cdot \vec{u} \\ \rho \frac{D}{Dt} \vec{u} &= -\nabla p + \vec{j} \times \vec{B} \\ \frac{D}{Dt} p &= -\gamma p \nabla \cdot \vec{u} \end{aligned}$$

with the assumptions

$$\begin{aligned} \frac{\partial \vec{B}}{\partial t} &= -\nabla \times \vec{E} \\ \mu_0 \vec{j} &= \nabla \times \vec{B} \\ \vec{E} + \vec{u} \times \vec{B} &= 0 \end{aligned}$$

We will apply our perturbation treatment to the first two equations of these (the continuity equation, momentum equation), using the assumptions to aid us. We will begin with the continuity equation.

Continuity Equation Time Perturbation

We restate the continuity equation:

$$\frac{D}{Dt} \rho = -\rho \nabla \cdot \vec{u}$$

Here we have two functions, ρ and \vec{u} , which are both functions of three dimension in space and one of time, i.e.

$$\begin{aligned} \rho &:= \rho(\vec{r}, t) \\ \vec{u} &:= \vec{u}(\vec{r}, t) \end{aligned}$$

where $\vec{r} \in \mathbb{R}^3, t \in \mathbb{R}$. Here (and throughout these perturbation treatments) we will decouple the time dependency by assuming it to be some epsilon perturbation from an initial state. For example, for mass density (ρ), we can let ρ_0 describe some initial state that is only dependent on space, and ρ_1 a perturbation component similarly dependent only on space. Let $\varepsilon_\rho > 0$ be some small perturbation factor for mass density, and t be the time evolved for. Then we prescribe a time-linear perturbation expansion as such:

$$\rho(\vec{r}, t) = \rho_0(\vec{r}) + \varepsilon_\rho t \rho_1(\vec{r}) + \mathcal{O}(\varepsilon_\rho^2)$$

and analogously:

$$\vec{u}(\vec{r}, t) = \vec{u}_0(\vec{r}) + \varepsilon_u t \vec{u}_1(\vec{r}) + \mathcal{O}(\varepsilon_u^2)$$

Here the t term effectively acts as a scaling factor on top of ε . We have also dropped any ε terms of order greater than 1 in our approximation, that is to say, if we were to be accurate our expansion would actually be:

$$\begin{aligned}\rho(\vec{r}, t) &= \rho_0(\vec{r}) + \varepsilon_\rho t \rho_1(\vec{r}) + \varepsilon_\rho^2 t \rho_2(\vec{r}) + \dots \\ \vec{u}(\vec{r}, t) &= \vec{u}_0(\vec{r}) + \varepsilon_u t \vec{u}_1(\vec{r}) + \varepsilon_u^2 t \vec{u}_2(\vec{r}) + \dots\end{aligned}$$

As we said at the start of this chapter however, this is an interdisciplinary thesis, and the above is an example of us putting on our physicist cap and deciding to remove “unnecessary” complexity. Physically we can justify this dropping of terms as we would expect terms of greater order to contribute diminishingly to our system. That is to say, on the time scales we are considering, an approximation of up to ε should be sufficient for experimental comparisons. As such, we need only to reason about that, and drop any terms on order of $\mathcal{O}(\varepsilon^2)$. We substitute these perturbation expansions into our equation. For brevity we will drop function parameters in our notation, and first begin with an expansion of the ρ term:

LHS:

$$\begin{aligned}\frac{D}{Dt} \rho &= \left(\frac{\partial}{\partial t} + \vec{u} \cdot \nabla \right) \rho \\ &= \frac{\partial}{\partial t} \rho + \vec{u} \cdot (\nabla \rho) \\ &= \frac{\partial}{\partial t} (\rho_0 + \varepsilon_\rho t \rho_1) + \vec{u} \cdot (\nabla(\rho_0 + \varepsilon_\rho t \rho_1)) \\ &= \frac{\partial}{\partial t} \rho_0 + \varepsilon_\rho \rho_1 + \vec{u} \cdot (\nabla \rho_0) + \varepsilon_\rho t \vec{u} \cdot (\nabla \rho_1)\end{aligned}$$

noting that ρ_0 is not a function of time

$$= \varepsilon_\rho \rho_1 + \vec{u} \cdot (\nabla \rho_0) + \varepsilon_\rho t \vec{u} \cdot (\nabla \rho_1)$$

RHS:

$$\begin{aligned} -\rho \nabla \cdot \vec{u} &= -(\rho_0 + \varepsilon_\rho t \rho_1)(\nabla \cdot \vec{u}) \\ &= -\rho_0(\nabla \cdot \vec{u}) - \varepsilon_\rho t \rho_1(\nabla \cdot \vec{u}) \end{aligned}$$

Combined:

$$\begin{aligned} \frac{D}{Dt} \rho &= -\rho \nabla \cdot \vec{u} \\ \varepsilon_\rho \rho_1 + \vec{u} \cdot (\nabla \rho_0) + \varepsilon_\rho t \vec{u} \cdot (\nabla \rho_1) &= -\rho_0(\nabla \cdot \vec{u}) - \varepsilon_\rho t \rho_1(\nabla \cdot \vec{u}) \\ \varepsilon_\rho \rho_1 + \varepsilon_\rho t [\vec{u} \cdot (\nabla \rho_1) + \rho_1(\nabla \cdot \vec{u})] + [\rho_0(\nabla \cdot \vec{u}) + \vec{u} \cdot (\nabla \rho_0)] &= 0 \\ \varepsilon_\rho \rho_1 + \nabla \cdot (\rho_0 \vec{u}) + \varepsilon_\rho t [\nabla \cdot (\rho_1 \vec{u})] &= 0 \end{aligned}$$

Here we have used the divergence property $\nabla \cdot (f \vec{v}) = (\nabla f) \cdot \vec{v} + f(\nabla \cdot \vec{v})$. Note that for the Grad-Shafranov equation we usually (and we shall) assume an incompressible fluid, which mathematically is significant as it gives us the property that $\nabla \cdot (\rho \vec{u}) = 0$. From the above we immediately get that $\varepsilon_\rho \rho_1 = 0$, which just tells us that the time component we introduced has no effect on the system, given the assumptions we make. This is excellent news for us, as it means that small perturbations in our system in a manner that emulates a time evolution should not affect the accuracy of our model, at least insofar as the effects of the continuity equation are concerned, and up to a term of $\mathcal{O}(\varepsilon^2)$. To be confident that we can do this for the Grad-Shafranov equation on a whole, we need to repeat this process for the other ideal MHD equations. Next we'll look at the momentum equation.

Momentum Equation

We restate the momentum equation here:

$$\rho \frac{D}{Dt} \vec{u} = -\nabla \rho + \vec{j} \times \vec{B}$$

We will follow the same process as we did for the continuity equation. Here we have four equations, however, by introducing one of our GSE assumptions we can reduce this to three immediately. For GSE we assume that our fluid has no velocity, i.e. $\vec{u} = 0$, which gives us that $\frac{D}{Dt} \vec{u} = 0$. Thus the equation we actually

have to expand is:

$$0 = -\nabla\rho + \vec{j} \times \vec{B}$$

The three equations we have are each functions of $\vec{r} \in \mathbb{R}^3$ and $t \in \mathbb{R}$, and can be linearly perturbed in time as we did earlier. As such:

$$\begin{aligned}\rho(\vec{r}, t) &:= \rho_0(\vec{r}) + \varepsilon_\rho t \rho_1(\vec{r}) + \mathcal{O}(\varepsilon_\rho^2) \\ \vec{B}(\vec{r}, t) &:= \vec{B}_0(\vec{r}) + \varepsilon_B t \vec{B}_1(\vec{r}) + \mathcal{O}(\varepsilon_B^2) \\ \vec{j}(\vec{r}, t) &:= \vec{j}_0(\vec{r}) + \varepsilon_j t \vec{j}_1(\vec{r}) + \mathcal{O}(\varepsilon_j^2)\end{aligned}$$

Making these substitutions into the momentum equation:

$$\begin{aligned}\nabla\rho &= \vec{j} \times \vec{B} \\ \nabla(\rho_0 + \varepsilon_\rho t \rho_1) &= \vec{j} \times (\vec{B}_0 + \varepsilon_B t \vec{B}_1) \\ (\nabla\rho_0) + \varepsilon_\rho t (\nabla\rho_1) &= (\vec{j} \times \vec{B}_0) + \varepsilon_B t (\vec{j} \times \vec{B}_1) \\ (\nabla\rho_0) + \varepsilon_\rho t (\nabla\rho_1) &= [(\vec{j}_0 \times \vec{B}_0) + \varepsilon_j t (\vec{j}_1 \times \vec{B}_0)] + \varepsilon_B t [(\vec{j}_0 \times \vec{B}_1) + \varepsilon_j t (\vec{j}_1 \times \vec{B}_1)] \\ (\nabla\rho_0) + \varepsilon_\rho t (\nabla\rho_1) &= \vec{j}_0 \times \vec{B}_0 + \varepsilon_j \varepsilon_B t^2 (\vec{j}_1 \times \vec{B}_1)\end{aligned}$$

If we collate terms of like order, we get

$$\begin{aligned}\nabla\rho_0 &= \vec{j}_0 \times \vec{B}_0 \\ \nabla\rho_1 &= \frac{\varepsilon_j \varepsilon_B}{\varepsilon_\rho} t (\vec{j}_1 \times \vec{B}_1)\end{aligned}$$

both of which look remarkably like the original momentum equation; an observation which is not to be mistaken for coincidence. Notably, when we take our evolution over “small enough” time scales, the perturbation components here disappear entirely, and we are left with the version of the momentum equation which is used in the regular equilibrium-state derivation of the Grad-Shafranov equation. As such, we would anticipate that small time perturbations associated with the effects of the momentum equation will introduce “small” errors in the resulting GSE solution, though (and as we hope to observe through experiment), hopefully imperceptibly so - as the above expansion seems to suggest.

The Other Ideal MHD Equations

Isaac Newton is quoted as saying “it causeth my head to ache”, which he is claimed to have said in reference to perturbation expansions in trying to determine

the orbit of the Moon [7]. Hopefully understandably, we ask that the reader excuse our omitting the working for the more involved of the ideal MHD equations, as to write it all down here would be to commit an expositionary crime. The expansion of the remaining equations however leads to a similar outcome as we found with the continuity equation and the momentum equation - that, up to an ε factor, the time perturbed ideal MHD equations do not show significant deviation from their equilibrium counterparts.

Significance

The purpose of this exercise was to provide some theoretical justification for the simulations we will later run. By showing that we can approximate a linear time evolution of the ideal MHD equations using states of equilibrium, we have shown that the GSE will be resistant to such a treatment itself. What we have not done is show that GSE itself can be evolved through time in an accurate manner. We are, knowingly, introducing errors into our system – but the hope is that the errors introduced will be insignificant with respect to the measurements we seek to make. This is the nature of mathematical modelling.

It is important to remark that the GSE is an equilibrium model - it is constructed under the assumptions of plasma equilibrium, and as we've already remarked on for resistivity, essentially “wishes away” many of the physical effects that would otherwise be introduced by some plasma which actually does change with respect to time. To that extent, it could be considered tomfoolery (which, for what it's worth, does befit the author) to assume that the work we've done is a guarantee that we can simply take equilibrium solutions to the GSE and assume that one is an evolution of the other in time, while still maintaining physicality.

While this may seem rather bleak / unpromising (after all, we are conceding that the model we have is not an entirely accurate physical representation of what's going on), the question we concern ourselves with is whether it is good enough to explain the observed runaway electron phenomena. To that extent, we will next introduce a variation of the Grad-Shafranov equation, solutions for which describe current reversal equilibrium configurations (CREC). We posit then that, if we take a solution to this system and perturb the current density profile by some small ε in time, then the resulting change in the GSH solution will be a sufficiently accurate representation of a physical system so as to explain the presence of excess runaway electrons.

3.2 Grad-Shafranov-Helmholtz Equation

The Grad-Shafranov equation, in its default state (as we have derived it), does not account for the possibility of current reversal equilibrium configurations (CRECs). That is to say, it does not permit anti-parallel current channels in an equilibrium state. This is problematic, as this is one of the behaviours we are seeking to identify in simulation. In the last 20 years work has been done on modelling such configurations, and analytic results have become more abundant [30], [21], [11]. For our purposes, we are interested in work done by Wang and Yu in deriving an analytic solution to a variant of the Grad-Shafranov equation known as the Grad-Shafranov-Helmholtz equation, which is capable of supporting CRECs. In this section we'll present their work, provide a brief derivation, and highlight the components that will be important for our simulations (the current density and pressure density profiles). Then we'll demonstrate an ability to recreate Wang's results.

3.2.1 Equation and Derivation

The work here will largely follow the paper by Wang and Yu [30], with some key notes and modifications. We are by now familiar with the GSE:

$$x \frac{\partial}{\partial x} \left(\frac{1}{x} \frac{\partial \Psi}{\partial x} \right) + \frac{\partial^2 \Psi}{\partial z^2} = -\mu_0 x^2 \frac{\partial p}{\partial \Psi} - \mu_0^2 f(\Psi) \frac{\partial f}{\partial \Psi} \quad (3.4)$$

Wang provides a slightly different (though equivalent) formulation:

Proposition 3.1. *Let R_0 and a describe the major and minor radius. B_0 is then the magnetic axis, which Wang takes to be the strength of the magnetic field at $R = R_0$. Additionally, we let $\psi = \Psi/(B_0 a^2)$ be the normalised poloidal magnetic flux, and we let $x = R/a$ and $z = Z/a$. For flux functions, we let $\beta(\psi) = (2\mu_0 p(\psi))/B_0^2$, $g(\psi) = F(\psi)/(B_0 a)$ and $j_\phi = (J_\phi \mu_0 a)/B_0$, where J_ϕ is the toroidal current density. The GSE can then equivalently be stated:*

$$\left(x \frac{\partial}{\partial x} \frac{1}{x} \frac{\partial}{\partial x} + \frac{\partial^2}{\partial z^2} \right) \psi = -\frac{1}{2} x^2 \frac{d\beta}{d\psi} - \frac{1}{2} \frac{dg^2}{d\psi} = -x j_\phi \quad (3.5)$$

where

$$a_1 = -\frac{1}{2} \frac{d\beta}{d\psi} \quad (3.6)$$

$$-a_2 - \alpha^2 \psi = -\frac{1}{2} \frac{dg^2}{d\psi} \quad (3.7)$$

Additionally we impose the boundary condition $\psi|_b = 0$.

Remark 3.2. We highlight here the parameter tuple (a_1, a_2, α) . We will see soon that these parameters can be used to determine our system. For now we simply note their significance so that the reader (that's you!) may pay them special attention through the rest of the following working.

Definition 3.3 (Helmholtz Equation). A Helmholtz equation is a problem associated with the below form

$$\nabla^2 f = -k^2 f$$

where ∇^2 is the Laplace operator, k^2 is the eigenvalue, and f is an eigenfunction.

The equation we have in 3.5 is in fact just the Grad-Shafranov equation. If, instead of a flux function ψ , we consider a toroidal vector potential factor $A = \psi/x$, then the GSE formulation given above reduces to a separate system, which can be identified as being Helmholtz:

$$\left(\frac{1}{x} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} + \frac{\partial^2}{\partial z^2} - \frac{1}{x^2} \right) A + \alpha^2 A = a_1 x - a_2 \frac{1}{x} \quad (3.8)$$

$$A|_b = 0 \quad (3.9)$$

We then seek a solution $\psi(x, z)$ for equation 3.5, which satisfies the above conditions in 3.8. Wang provides one:

Proposition 3.4. *The below is a solution to the GSH as given above*

$$\psi(x, z) = x \sum_{n=1}^{\infty} \sum_{l=0}^{\infty} \frac{(-1)^l 2a_n^u}{kv_l a_n^d (\alpha^2 - \lambda_{n,l}^2)} [c_n J_1(\mu_n x) + N_1(\mu_n x)] \cos(v_l z) \quad (3.10)$$

where here J_α are Bessel functions of the first kind, and N_α are Bessel functions of the second kind. Some of these other terms we will for now leave unexplained, as we will cover them in the derivation.

Notation 3.5. Note that we use the notation N_α to mean second-kind Bessel functions. This is an old notation, with a more modern notation being Y_α . N_α

nevertheless often appears in older physics texts, and is what we use in this thesis for Bessel functions of the second kind.

Derivation

We will assume a cylindrical coordinate system, and consider our boundary conditions to be a rectangular cross section. We will take $x \in [x_0 - a, x_0 + a]$ and $z \in [-k, k]$, where k is an elongation constant which determines the ellipticity of the cross section. With this in mind, the equivalent problem to 3.8 is the system

$$\left(\frac{1}{x} \frac{\partial}{\partial x} x \frac{\partial}{\partial x} + \frac{\partial^2}{\partial z^2} - \frac{1}{x^2} \right) U + \lambda^2 U = 0 \quad (3.11)$$

$$U(x_0 - a, z) = U(x_0 + a, z) = 0 \quad (3.12)$$

$$U(x, -k) = U(x, k) = 0 \quad (3.13)$$

If we begin by assuming that a solution is of the form $U(x, z) = \mathcal{R}(x)\mathcal{Z}(z)$, then we can describe it as an orthogonal system of eigenfunctions, and make a Fourier expansion as such:

$$A(x, z) = \sum_{n=1}^{\infty} \sum_{l=0}^{\infty} A_{n,l} \mathcal{R}_n(x) \mathcal{Z}_l(z) \quad (3.14)$$

$$a_1 x - a_2 \frac{1}{x} = \sum_{n=1}^{\infty} \sum_{l=0}^{\infty} f_{n,l} \mathcal{R}_n(x) \mathcal{Z}_l(z) \quad (3.15)$$

Equations 3.14 and 3.15 can be substituted into 3.8, which Wang uses to find an explicit form for $A_{n,l}$:

$$A_{n,l} = \frac{f_{n,l}}{\alpha^2 - \lambda_{n,l}^2} = \frac{(-1)^l 2a_n^u}{kv_l a_n^d (\alpha^2 - \lambda_{n,l}^2)} \quad (3.16)$$

Thus, we need to reason about the values $\mathcal{R}_n(x)$ and $\mathcal{Z}_l(z)$. After introducing constants v^2 and μ^2 such that $\lambda^2 = v^2 + \mu^2$ (a result of \mathcal{Z} and \mathcal{R} being orthogonal),

Wang found the relations

$$\frac{d^2\mathcal{Z}}{dx^2} + v^2 \mathcal{Z} = 0 \quad (3.17)$$

$$\mathcal{Z}(-k) = \mathcal{Z}(k) = 0 \quad (3.18)$$

$$\frac{d^2\mathcal{R}}{dx^2} + \frac{1}{2} \frac{d\mathcal{R}}{dx} + \left(\mu^2 - \frac{1}{x^2} \right) \mathcal{R} = 0 \quad (3.19)$$

$$\mathcal{R}(x_0 - a) = \mathcal{R}(x_0 + a) = 0 \quad (3.20)$$

Equations 3.17 and 3.18 are an eigenvalue problem, and provide the following eigenvalues and eigenfunctions

$$v_l = \frac{\pi}{k} \left(l + \frac{1}{2} \right) \quad (3.21)$$

$$\mathcal{Z}_l(z) = \cos(v_l z) \quad (3.22)$$

with $l \in \mathbb{Z}_{\geq 0}$. Similarly, equation 3.19 with 3.20 represent another eigenvalue problem. The eigenfunctions are given as

$$\mathcal{R}_n(x) = c_n J_1(\mu_n x) + N_1(\mu_n x) \quad (3.23)$$

where $n \in \mathbb{Z}_{\geq 1}$. Here, however, we don't have an explicit form for the eigenvalues μ_n . Instead, μ_n is described as being the zeros of the function

$$-J_1(\mu_n(x_0 + a)) \frac{N_1(\mu_n(x_0 - a))}{J_1(\mu_n(x_0 - a))} + N_1(\mu_n(x_0 + 1)) = 0 \quad (3.24)$$

where

$$c_n = -\frac{N_1(\mu_n(x_0 - a))}{J_1(\mu_n(x_0 - a))}$$

Thus we can summarise the eigenvalues and eigenfunctions of 3.11 - 3.13 are given by

$$\lambda_{n,l}^2 = v_l^2 + \mu_n^2 \quad (3.25)$$

$$U_{n,l}(x, z) = \mathcal{R}_n(x) \mathcal{Z}_l(z) \quad (3.26)$$

Thus we have an explicit form for the eigenfunctions we required in 3.16. We can also now provide expressions for the, until now, unknown variables in 3.16, the derivations for which are given by Wang in an earlier paper [29]. We provide

them en masse here:

$$e_n^u(x) = \frac{1}{\mu_n} [a_1 x^2 (c_n J_2(\mu_n x) + N_2(\mu_n x)) + a_2 (c_n J_0(\mu_n x) + N_0(\mu_n x))] \quad (3.27)$$

$$a_n^u = [e_n^u(x)]_{R_0-a}^{R_0+a} \quad (3.28)$$

$$e_n^d(x) = -\frac{1}{2} x^2 N_0(\mu_n x) N_2(\mu_n x) - \frac{1}{2} c_n^2 x^2 J_0(\mu_n x) J_2(\mu_n x) \quad (3.29)$$

$$+ c_n \left[\frac{1}{2} x^2 (J_0(\mu_n x) - J_2(\mu_n x)) N_0(\mu_n x) - \frac{1}{\mu_n} x J_0(\mu_n x) N_1(\mu_n x) \right]$$

$$a_n^d = [e_n^d(x)]_{R_0-a}^{R_0+a} \quad (3.30)$$

Note here that the superscripts u and d are not indices, but are just used to distinguish between variables. Thus, to show that $\psi(x, z)$ as provided in 3.10 is a solution, we simply substitute equations 3.22, 3.23 and 3.16 into 3.15.

3.2.2 Current Density and Pressure Density

We have an explicit formula for the poloidal magnetic flux function ψ (albeit normalised for our case), but we remain to have an expression for our current density and pressure density profiles.

Pressure Density Profile

In a small note below equation (16c) of [30], an expression for the plasma pressure is given as

$$\beta(x, z) = \beta_0 - 2a_1 \psi(x, z) \quad (3.31)$$

where β_0 is chosen such that the minimum of the expression is zero for the domain (this just ensures that it's not possible to have negative pressures).

Current Density Profile

Equation 3.5 relates the toroidal current density to our parameters a_1, a_2 and α . As we know all other components here, we can draw the toroidal current density

from this. In doing so, recall notably equations 3.6 and 3.7.

$$\begin{aligned} -xj_\phi(x, z) &= -\frac{1}{2}x^2 \frac{d\beta}{d\psi} - \frac{1}{2} \frac{dg^2}{d\psi} \\ -xj_\phi(x, z) &= a_1x^2 - a_2 - \alpha^2\psi(x, z) \\ \implies j_\phi(x, z) &= -a_1x + \frac{1}{x}a_2 + \frac{\alpha^2}{x}\psi(x, z) \end{aligned} \quad (3.32)$$

Remark 3.6. As we stated toward the start of this chapter, three parameters appear crucial to this system: (a_1, a_2, α) . We see that, in fact, the entirety of our system can be determined by them, as any other variables are configuration variables for the tokamak. Thus, if we are provided the values (a_1, a_2, α) , we can calculate the poloidal magnetic flux, the current density profile, and the pressure density profile.

3.3 Simulations

At this point, we are able to describe the poloidal magnetic field flux function, current density profile, and pressure density profile for a tokamak under the restrictions of the GSH model. However, we are only able to do this if we are given some set of parameter, (a_1, a_2, α) . In the next chapter we will deal with the question of deriving these parameters for a given current density profile, but for now we can simply reproduce the results given by Wang in section 3 of [30].

Remark 3.7. Note that in our simulation, we cannot have infinite precision when it comes to our eigenfunctions (i.e., how far the indices n, l are permitted to go). In our computations we must restrict these - however, luckily for us, the Bessel functions mean that the strength of the contribution of higher order terms is vanishingly small. In most of our simulations we have capped these at $M = 8$ levels, i.e.

$$\dots \sum_{n=1}^M \sum_{l=0}^{M-1} \dots$$

where the rest of the equation is omitted for purposes of laziness. Additionally, we set a grid size of approximately 200 steps for partitioning the x and z domains.

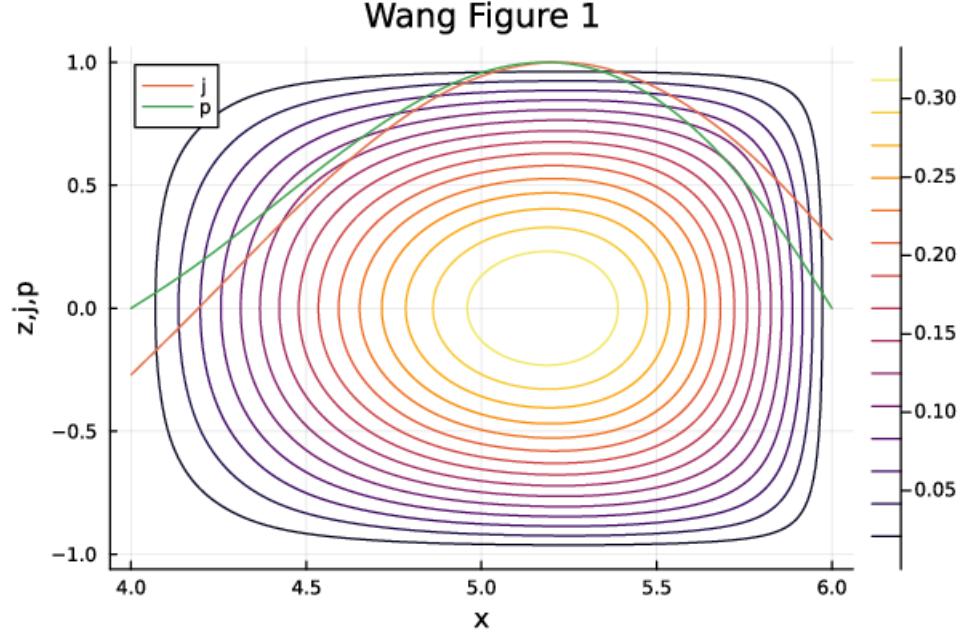


Figure 3.2: Figure 1 of [30]. Here, the parameters are $(a_1, a_2, \alpha) = (-0.04531, -1.0808, 2.1683)$, where the minor radius $a = 1$, the elongation is $k = 1$, and the major radius is taken to be $R = 5$.

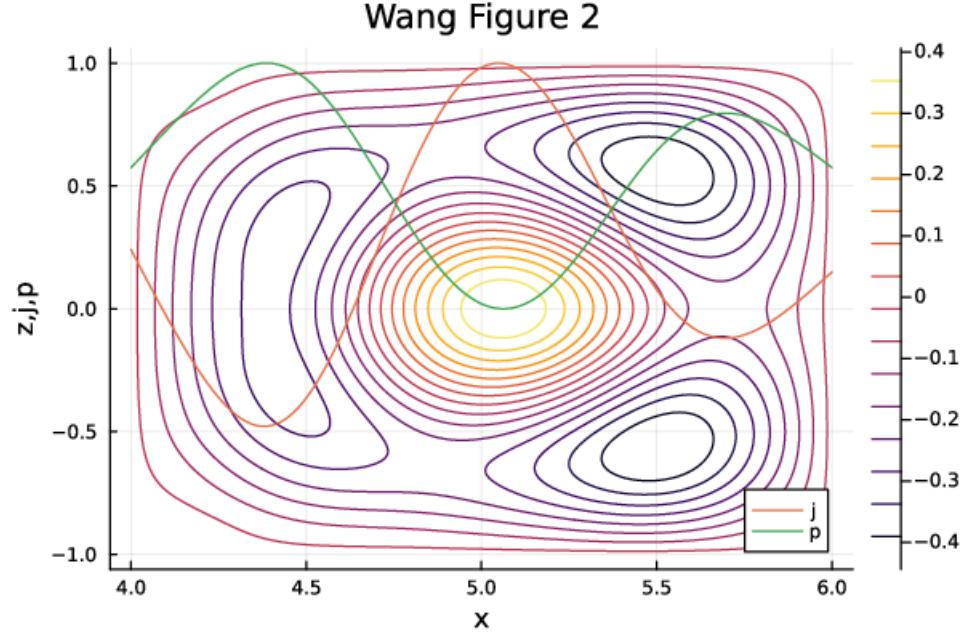


Figure 3.3: Figure 2 of [30]. Here, the parameters are $(a_1, a_2, \alpha) = (0.01, 3.1, 5.566)$, where the minor radius $a = 1$, the elongation is $k = 1$, and the major radius is taken to be $R = 5$. Note that the characteristic x-point is seemingly missing from our version - this is simply an effect of Julia's plot smoothing, and the data representing it is still available to us.

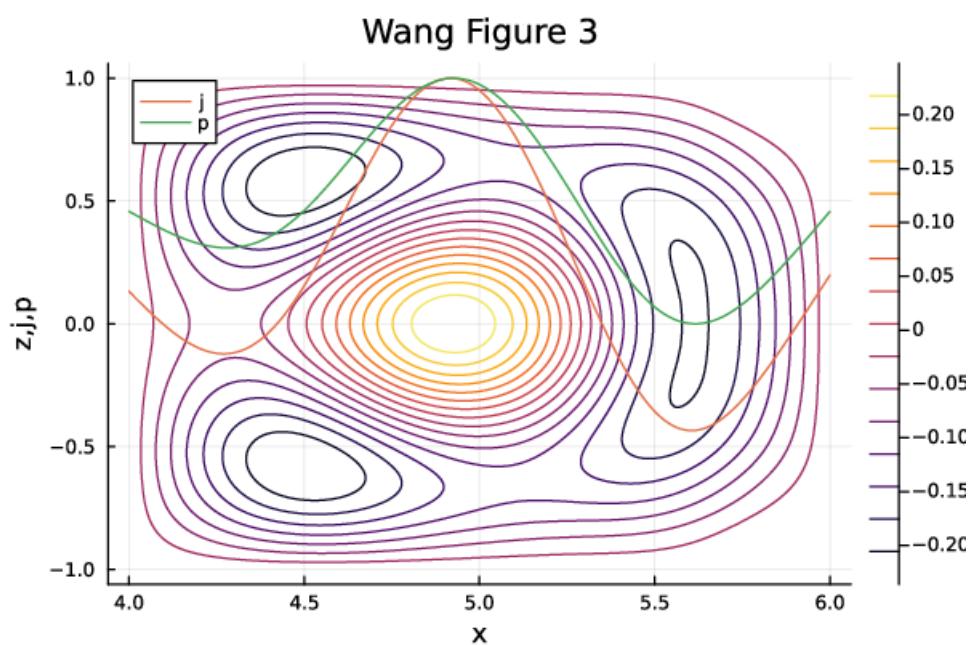


Figure 3.4: Figure 3 of [30]. Here, the parameters are $(a_1, a_2, \alpha) = (-0.06, -0.0006, 5.50)$, where the minor radius $a = 1$, the elongation is $k = 1$, and the major radius is taken to be $R = 5$.

Chapter 4

Numerical Model Fitting

With the ability to generate an equilibrium solution for the GSH equation from chapter 3, our next goal is to be able **TODO**In chapter 3 we developed the ability

4.1 Non-Linear Optimisation

Optimisation methods are, at their heart, minimisation problems (and dually maximisation problems, though the two are equivalent). We have some function $f(x)$, called the *objective function*, and some domain $\Omega \subset \mathbb{R}^n$, and we wish to determine

$$\arg \min_{x \in \Omega} f(x)$$

Often the objective function is accompanied by a set of constraints, which more or less are used to define the set Ω . If we let $c_i(x) \geq 0$ describe one of k constraints that f is subject to, then we can describe Ω as such:

$$\Omega := \{x \in \mathbb{R}^n : c_1(x) \geq 0 \wedge c_2(x) \geq 0 \wedge \dots \wedge c_k(x) \geq 0\}$$

Remark 4.1. If a given problem has no constraints (i.e. you are only presented with the task of minimising a function), then this is known as an **unconstrained optimisation** problem. As soon as you add constraints however, you have a **constrained optimisation** problem. Algorithms may have better or worse performance in reaching a local minimum depending whether the problem is constrained or not.

There are a number of difficulties that come packaged with optimisation problems. Firstly, the method by which you go about minimising your function is

dependent on the properties of your function - you will have a much easier time minimising a smooth, convex function than you would a discontinuous, non-convex function, and different algorithms will net you varying results for each. Another factor that affects the choice of optimisation methods available, is the linearity of your system. Often linear systems can have “nice” methods of solving them. However, the behaviour of non-linear systems can often make them unwieldy when it comes to finding local and/or global minima.

Definition 4.2 (Non-linear Optimisation Problem). Let $x \in \Omega \subseteq \mathbb{R}^n$. Also let $f : \Omega \rightarrow \mathbb{R}$ be the objective function, and $c_i : \Omega \rightarrow \mathbb{R}$ and $d_j : \Omega \rightarrow \mathbb{R}$ be constraints. Then a non-linear optimisation problem is a problem

$$\begin{aligned} & \min_{x \in \Omega} f(x) \\ & c_i(x) \geq 0 \quad \forall i \in \{1, \dots, m\} \\ & d_j(x) = 0 \quad \forall j \in \{1, \dots, p\} \end{aligned}$$

where one of f , c_i and/or d_j are non-linear.

4.1.1 Least Squares

Least squares problems are a subset of optimisation problems, which are largely concerned with fitting some function to a set of data. Let a set of data be given (x_i, y_i) , where x_i are independent variables, and y_i are dependent variables in the data set. We wish to fit some function $f(x, \beta)$ to fit our data as tightly as possible, where x is the independent variable as for the data, and β are parameters in the function f which can be varied. Let r_i be the *residual* error between a given data entry (x_i, y_i) and our corresponding value for it, $f(x_i, \beta)$ for a given β . Then our goal (in wanting to have a function f which matches the provided data) is to reduce the amount of residual error we have - to minimise it. We can specify this as an optimisation problem [3]

$$\min_{\beta \in \Omega} \sum_{i=1}^n \|r_i(\beta)\|^2 = \min_{\beta \in \Omega} \sum_{i=1}^n \|y_i - f(x_i, \beta)\|^2$$

In this case, our objective function is the equation $g(\beta) = \sum_{i=1}^n \|y_i - f(x_i, \beta)\|^2$.

Remark 4.3. While the above norm is left general, often optimisation algorithms will take this to be the $L^2(\mathbb{R}^n)$ norm.

Remark 4.4. The above is, by default, an unconstrained optimisation problem. When we come to formulate our own least squares problem, we will introduce our own constraints however. Additionally, the above is not specified to be linear or non-linear. In our efforts, we will deal with a non-linear system.

Thus, if you are provided with a set of data, and a function you wish to use to approximate the behaviour of the data, you can employ the use of a least squares optimisation method to manipulate that function into being as close a fit to that data as it can be. How we implement that optimisation is a question of algorithms, which we will now cover.

4.1.2 Optimisation Algorithms

Notation 4.5. When we talk about running algorithms to solve our optimisation problems, we will use a few terms. We define those here:

Feasible	A problem is feasible if the algorithm is able to find an $x^* \in \Omega$ that is a local minimum and satisfies any given conditions. Otherwise, it is infeasible .
Objective Value	The value of the objective function for the identified local minimum $x^* \in \Omega$, $f(x^*)$.
ftol	If the difference between consecutive objective function values is less than this amount, stop the optimisation algorithm and return converged
xtol	If the difference between consecutive $x \in \Omega$ values is less than this amount, stop the optimisation algorithm and return converged

The main questions that an optimisation algorithm asks at each of its iterations is: which direction should I travel, and how far should I travel in that direction? Algorithms that use this form of logic are called *descent methods*, and a general approach for them is given [3]:

-
- 1: Given initial guess x
 - 2: **while** Convergence criteria not satisfied **do**
 - 3: Choose (unit) descent direction $\delta_x \in \mathbb{R}^n$
 - 4: Choose step size $\gamma \in \mathbb{R}$
 - 5: $\delta \leftarrow \gamma \cdot \delta_x$
 - 6: Update variables $x = x + \delta$
-

Thus, descent algorithms are distinguished by the mechanism by which they choose a descent direction, and the step size.

While the algorithms we will talk about here can be applied to optimisation problems generally, we should keep in mind their application to non-linear least squares problems, as this is what we will eventually deal with. Here we'll provide an overview of two common algorithms used for finding local minima of a given function. In our attempts to fit a model to some data in section 4.2, we first attempted to use these methods. However, due to the nature of the objective function we have, these efforts were in vain, and we turned to a different algorithm (MMA). Nevertheless, it is useful for intuition to see examples of other algorithms.

Newton's method

Newton's optimisation method exploits a Taylor series expansion of a given objective function. Assume that we want to determine $\min f(x)$, where $x, \delta \in \mathbb{R}^n$. We know that an expansion about $x + \delta$ is:

$$f(x + \delta) \approx h(x + \delta) := f(x) + \nabla f(x)^T \delta + \frac{1}{2}(\delta)^T \nabla^2 f(t) \delta$$

where ∇^2 denotes the Hessian. We want to find a turning point (minimum) of the above, so we can simply set its derivative to be 0. This gives us

$$\nabla f(x) + \nabla^2 f(x) \delta = 0$$

This tells us that the optimal descent direction (and step size) is

$$\delta = -\gamma [(\nabla^2 f(x))^{-1} (\nabla f(x))]$$

where $0 < \gamma \leq 1$ is a chosen step size, which can be tailored to a given problem.

Remark 4.6. As Newton's method only necessitates that the objective function be twice differentiable, it is applicable to optimisation problems generally (i.e. it is not restricted to constrained nor unconstrained problems, nor does it differentiate between linear and non-linear problems).

Gradient descent

Gradient descent is an approach which has strong intuitive roots. If you were at the top of a hill, and wanted to get to the bottom as fast as possible, then

you might choose to go down the steepest point you see around you. This is essentially what gradient descent does. For an objective function $f(x)$, assuming that it is differentiable for some neighbourhood of a point $a \in \mathbb{R}^n$, then the optimal descent direction is $-\nabla f(a)$. This can, similar to Newton's method, be scaled by a step size γ to slow / speed up this process, though the value of this is dependent on the optimisation problem you're investigating (note that, similar to Newton's method, the gradient descent method can be applied to a large class of problems).

MMA

One of the problems we encountered in our data fitting attempts (which we will see in the next section), is that of asymptotic behaviour which affects convergence of our algorithms. Thus, we sought to employ the use of an algorithm which was more resistant to such behaviour in an objective function. This led to the discovery of the Method of Moving Asymptotes algorithm, developed by Svanberg [27]. While its implementation details are a bit volumous to repeat here, in essence, it works by creating convex sub-problems around potential local minima, and attempts to minimise these individually. It repeats this process until it is no longer decreasing, and returns the most recent potential local minima it found. This algorithm is useful to us two main reasons: 1. it is still as robust an optimisation method as Newton's method and gradient descent, in that it can handle a host of optimisation problems (including non-linear problems); and 2, it is more resistant to the asymptotic behaviour.

In our code we use an implementation provided by the NLOpt package, a non-linear optimisation library for Julia [16]. The NLOpt library, in initialising an optimisation problem for MMA, requires information about the objective function, its partial derivatives, and any constraints, if they exist (which they will for us, but we'll note on that shortly). Thus, when it comes to constructing our system, we must provide the objective function naturally, but also its partial derivatives with respect to our parameters.

4.2 GSH Parameter Fitting

In the previous chapter we were able to, given $P := (a_1, a_2, \alpha)$, reproduce results from Wang, including the poloidal magnetic flux, current density profile, and pressure density profile. However, in reality, we will not know these parameters

P for a given system - yet we are entirely dependent on them for describing our system. We then face a problem, which is that if we wish to vary our system, or if we wish to describe a real tokamak's dynamics, we need some way to determine what P should represent the system we're investigating.

Looking forward a bit, we know that from the ISTTOK project we have time series current density profile and pressure density profile data available to us. Thus we may wish to utilise at least one of these in determining P - and this is exactly what we will do. In the previous section we built the theory for, given a set of data, fitting a function f with variables β which can be varied, which is exactly the challenge we now face.

We could seek to use both the current density profile and pressure density profile, and could build an objective function around this. However, we will also note that we wish to verify our results - as such, it may be prudent to leave one of these data sets available for comparison instead. With a bit of prescient knowledge we may realise that derivatives are much easier to calculate for the current density profile, and so choose to build our objective function around this, reserving pressure density data for comparisons instead. In the spirit of laziness, this is exactly what we do.

In this chapter we will not concern ourselves with experimental data however - that is the task of chapter 5. Here we will work with contrived data. With our ability to specify parameters P and derive the equilibrium solution from that, we can generate our own simulated current density profile data, and then attempt to work in the opposite direction - to fit parameters P' to that data, which we can then compare to P for accuracy.

4.2.1 Optimisation Problem

Objective Function

Let (x_i, d_i) denote a data entry for the current density profile. We will perform a least squares optimisation, and so need to specify $f(x, \beta)$. In our case, $\beta = P = (a_1, a_2, \alpha) \in \mathbb{R}^3$, and

$$f(x, \beta) = j_\phi(x, z) = -a_1 x + \frac{1}{x} a_2 + \frac{\alpha^2}{x} \psi(x, z)$$

Note here though that the current density profile is taken to be for $z = 0$. Additionally, we can transform our current density profile function \vec{j}_ϕ to also be in terms of P by specifying them as independent variables, and so the above is more

accurately stated

$$f(x, \beta) = j_\phi(x, P) = -a_1 x + \frac{1}{x} a_2 + \frac{\alpha^2}{x} \psi(x, 0, P)$$

Assume that we have N data entries for the current density profile available to us. Then, our least squares optimisation problem is given:

$$\begin{aligned} & \min_{\beta \in \Omega} \sum_{i=1}^N \|d_i - f(x_i, \beta)\|^2 \\ \iff & \min_{P \in \Omega} \sum_{i=1}^N \|d_i - j_\phi(x_i, P)\|^2 \end{aligned}$$

for now we will take $\Omega = \mathbb{R}^3$, however this will change shortly. This, however, is the objective function we seek to minimise - in the hopes that the resulting parameter P will let $j_\phi(P)$ accurately describe the current density profile we expect, and thus the rest of the system will follow suit.

Partials

We require $\nabla g(P)$, where $g(P) = \sum_{i=1}^N \|d_i - j_\phi(x_i, P)\|^2$. We'll calculate each individually:

$$\begin{aligned} \frac{\partial g(P)}{\partial a_1} &= \frac{\partial}{\partial a_1} \sum_{i=1}^N \|d_i - j_\phi(x_i, P)\|^2 \\ &= \frac{\partial}{\partial a_1} \sum_{i=1}^N (d_i - j_\phi(x_i, P))^2 \\ &= 2 \sum_{i=1}^N \left(\frac{\partial}{\partial a_1} (d_i - j_\phi(x_i, P)) \right) \\ &= -2 \sum_{i=1}^N \left(d_i - \frac{\partial}{\partial a_1} j_\phi(x_i, P) \right) j_\phi(x_i, P) \end{aligned}$$

By similar calculation, we get the partials for a_2 and α to be:

$$\begin{aligned}\frac{\partial g(P)}{\partial a_2} &= -2 \sum_{i=1}^N \left(d_i - \frac{\partial}{\partial a_2} j_\phi(x, P) \right) j_\phi(x, P) \\ \frac{\partial g(P)}{\partial \alpha} &= -2 \sum_{i=1}^N \left(d_i - \frac{\partial}{\partial \alpha} j_\phi(x, P) \right) j_\phi(x, P)\end{aligned}$$

Thus, we now only need the partials of the toroidal current density with respect to each element of P , which we provide below

$$\begin{aligned}\frac{\partial j_\phi(x, P)}{\partial a_1} &= \frac{\partial}{\partial a_1} \left(-a_1 x + \frac{1}{x} a_2 + \frac{\alpha^2}{x} \psi(x, z, P) \right) \\ &= -x + \frac{\partial}{\partial a_1} \left(\frac{\alpha^2}{x} \psi(x, z, P) \right) \\ &= -x + \alpha^2 \sum_{n=1}^M \sum_{l=0}^{M-1} \frac{2(-1)^l [c_n J_1(\mu_n x) + N_1(\mu_n x)]}{kv_l a_n^d (\alpha^2 - \lambda_{n,l}^2)} \left[\frac{1}{\mu_n} x'^2 (c_n J_2(\mu_n x') + N_2(\mu_n x')) \right]_{R_0-a}^{R_0+a} \\ \frac{\partial j_\phi(x, P)}{\partial a_2} &= \frac{1}{x} + \alpha^2 \sum_{n=1}^M \sum_{l=0}^{M-1} \frac{2(-1)^l [c_n J_1(\mu_n x) + N_1(\mu_n x)]}{kv_l a_n^d (\alpha^2 - \lambda_{n,l}^2)} \left[\frac{1}{\mu_n} (c_n J_0(\mu_n x') + N_0(\mu_n x')) \right]_{R_0-a}^{R_0+a} \\ \frac{\partial j_\phi(x, P)}{\partial \alpha} &= \frac{2\alpha}{x} \psi(x, 0, P) + \alpha^3 \sum_{n=1}^M \sum_{l=0}^{M-1} \frac{4(-1)^l a_n^u}{kv_l a_n^d (\alpha^2 - \lambda_{n,l}^2)} [c_n J_1(\mu_n x) + N_1(\mu_n x)]\end{aligned}$$

With these, we have the partial derivatives (with respect to our parameters) of our objective function, and we have our objective function. Additionally, we have our data (emulated current density profile data, which we generated from the prescribed (a_1, a_2, α) values in Wang [30]). This is everything required by NLopt for performing an optimisation, and so the only thing left to do is optimise, and see our results!

4.2.2 Parameter Space

Initially, our efforts in fitting a set of parameters to a given input of current density profile data were without accomplishment, as the optimisation methods failed to converge (reaching max iteration counts before reaching an acceptable objective value). Given we have omnipotent knowledge as to what our $P = (a_1, a_2, \alpha)$ values should be, we can look at the parameter space around these values to see why our optimisation algorithm might be falling over. The results are telling, and can be seen in figures 4.1 to 4.3

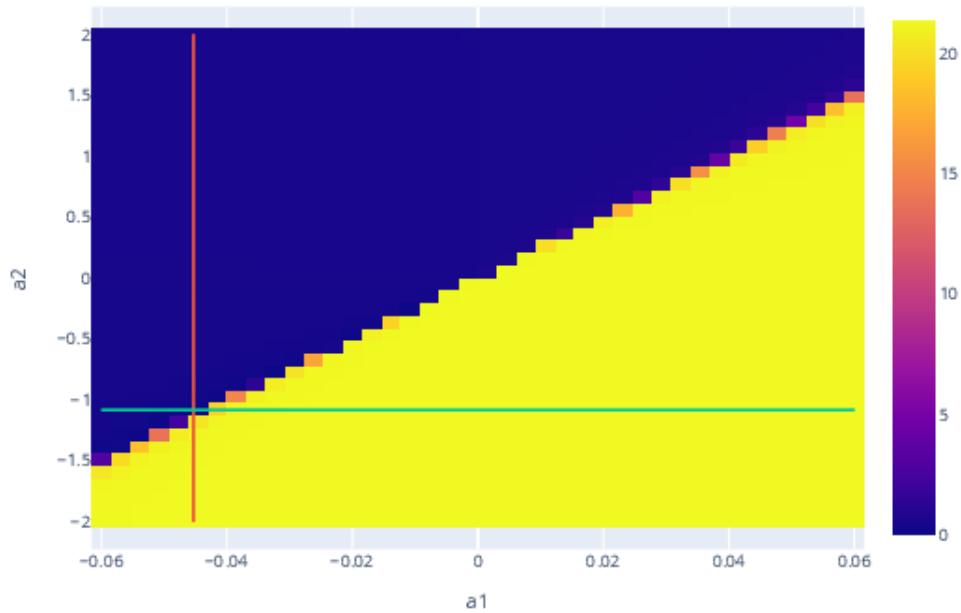


Figure 4.1: Objective value heatmap for fixed α , with varying a_1 and a_2

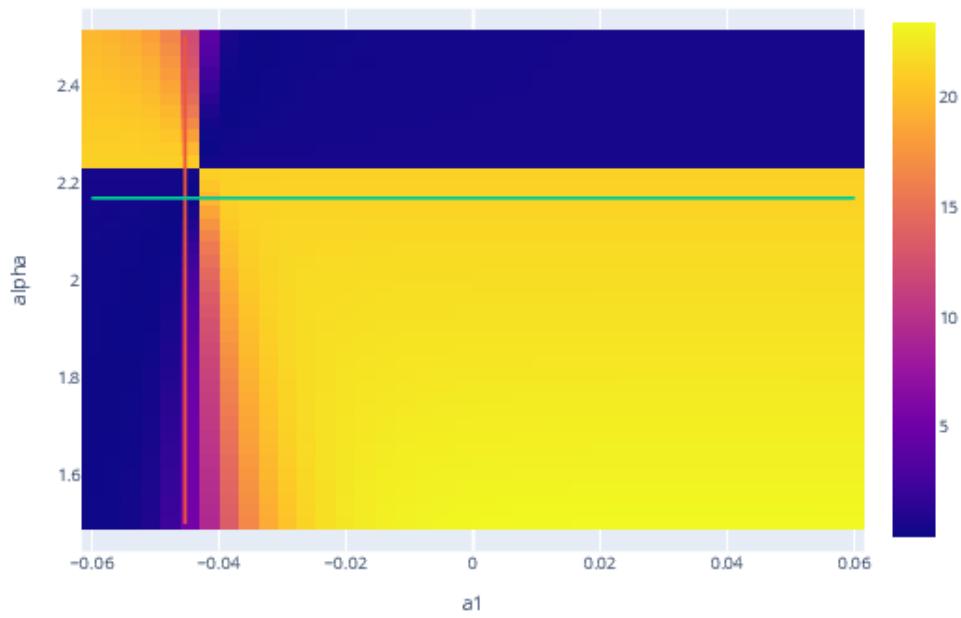


Figure 4.2: Objective value heatmap for fixed a_2 , with varying a_1 and α

The lines here point to the expected (a_1, a_2, α) values, and for each one we

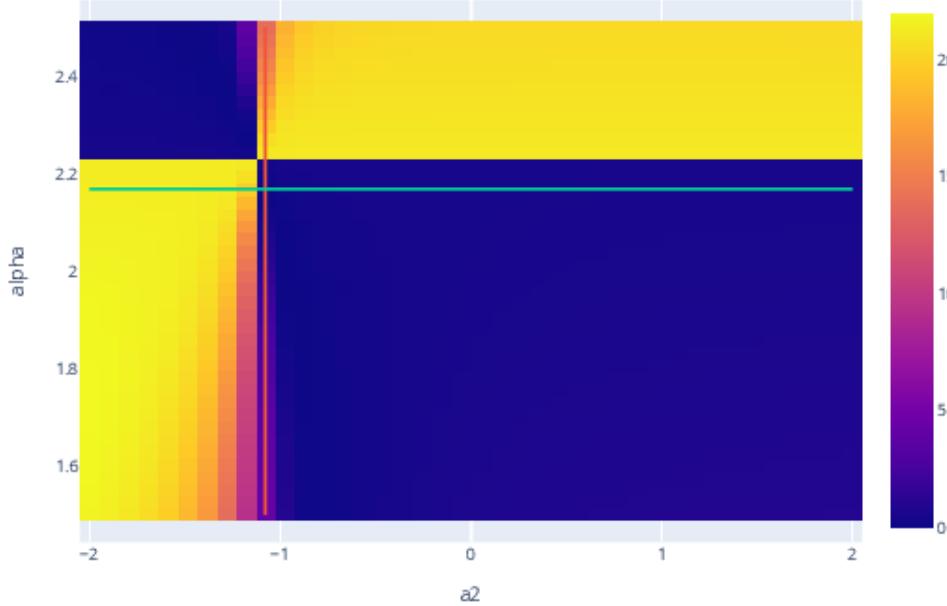


Figure 4.3: Objective value heatmap for fixed a_1 , with varying a_2 and α

fix the parameter not being varied to be its expected value. Nevertheless, as we can see, our parameter space is quite binary, in that the distinction between suitable parameters and not is quite sharp, but with plateau's on either side - a nightmare situation for optimisation algorithms. If we zoom in around our expected parameters we indeed see even more strongly the sensitivity of our objective value to fluctuations in input parameters - see figures A.1 to A.3 in Appendix A.

Some of this behaviour is the result of the $(\alpha^2 - \lambda_{n,l}^2)^{-1}$ term introduced by ψ in our objective function - it is this which motivates our usage of the MMA optimisation algorithm over more traditional methods like gradient descent, as we discussed in the previous section. However, as it stands - as an unconstrained optimisation problem - this is clearly still insufficient. One initial thought might be that our problem is simply underdetermined, and that by introducing more data we may improve the ability of our model to converge. However, as we see in figures A.4 to A.6, the inclusion of pressure density profile data in our fitting attempts does not vastly improve our parameter space.

As such, the inclusion of extra data does not seem promising in our efforts to find a suitable set of parameters. Instead, our next effort was to constraint our problem - given we know what to expect our $P = (a_1, a_2, \alpha)$ values to be, we can

constrain our parameter space by specifying a small offset for each. For example, if we know our parameter set is supposed to be $P' = (a'_1, a'_2, \alpha')$, then we can set

$$\Omega = [a'_1 - \varepsilon_{a_1}, a'_1 + \varepsilon_{a_1}] \times [a'_2 - \varepsilon_{a_2}, a'_2 + \varepsilon_{a_2}] \times [\alpha' - \varepsilon_\alpha, \alpha' + \varepsilon_\alpha]$$

For our initial parameter guess P_0 we can pick some random element of Ω . When we introduce these constraints and turn our problem into a constrained optimisation problem we see significantly improved results, and our model converges - see figure 4.4 as a recreation of figure 1 from Wang using simulated current density profile data.

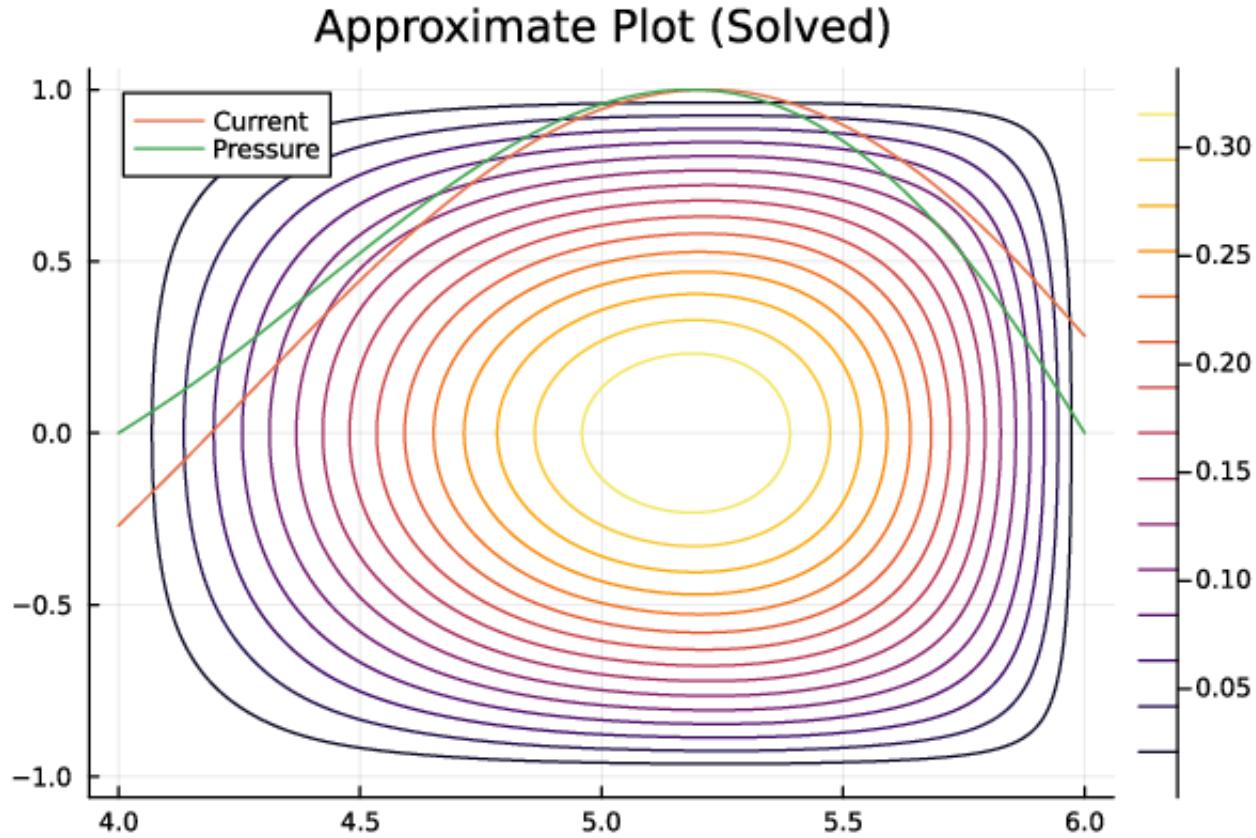


Figure 4.4: Equilibrium solution for the parameter set P that we determine via optimisation, starting with a set of current density profile data. This matches the expected figure 1 from Wang [30].

Now that we can, for a given set of data, accurately derive the associated equilibrium solution, our next goal is simulate a current inversion given said data.

```
Objective value: 0.012403462193520648
a1: 0.00935647192762522
a2: 0.22696301035377772
α : 4.9426985720035495
stopping reason: LOCALLY_SOLVED
.. : FTOL_REACHED
```

Figure 4.5: Text output of our code showing convergence of our model.

4.3 Simulated Current Reversals

We have the ability to solve for $P = (a_1, a_2, \alpha)$, and so now we will draw upon our work in chapter 3 in perturbing our equilibrium solution to simulate the effects of a time evolution, and in doing so, will observe how the system reacts.

4.3.1 Method of Reversal

Linearly vary current. Could talk about alternatives - i.e. if our time perturbation had not been linear but had been trigonometric (to fit with a larger scope AC simulation) then our current reversal method here could be different

Feed solved parameters into guess for next. Emphasise problem of not knowing the initial ones nonetheless

4.3.2 Results and Explanations

Have still frame shots of the result of that (3x3 tiles?)

Show the slow version, then the 'zoomed' in version

Highlight magnetic field topology breaking - suggests RE population generation

Comment on feasibility of the behaviour - e.g. pressure density profile follows magnetic field axis, but then there are things like current densities at the edge of the reactor

Chapter 5

Comparison with Experiment

5.1 ISTTOK

The ISTTOK project is a reactor based in Lisbon, Portugal, used for plasma science research by the University of Lisbon. It is this reactor which found no evidence for the existence of anti-parallel currents in the ramp down phase, contradicting the hypothesis put forward to explain the presence of a residual plasma during the quiescent current cycle, as we discussed in the introduction [18]. This hypothesis was initially supported by data from the CT-6B Tokamak by Huang et. al [13] however, and so no consensus exists on whether these currents exist, and what the cause of the runaway electrons is in cases of AC Tokamak ramp downs. In this chapter we attempt to use our model to describe the current density profile, pressure density profile, and poloidal magnetic field topology, given some data provided by ISTTOK.

5.1.1 Heavy Ion Beam Diagnostics

The data we have available to us is time slices of current density profile, pressure density profile, and v_{loop} data for the ramp down phase of a run of the ISTTOK reactor. Before we “plug and play” with this data though, it’s important to understand where it has come from.

One of the flagship features of the ISTTOK reactor is its heavy ion beam probe, which enables the measurement of the plasma temperature, electron density, and plasma potential (v_{loop}) data. Theoretically it is able to produce a one dimensional profile of each of these components - these can then be used to infer the value of other plasma properties, such as the current density profile [19].

We will stave off delving too deep into the physics here, and will instead seek to provide an intuition for the functioning of the beam. A heavy ion beam is a measurement tool that works by injecting heavy charged particles (ions) into a charged medium, at speed. The charged medium in our case is the plasma. For a Tokamak, a beam of positively charged ions (specifically, atoms with one extra electron) are injected into the plasma, known as the “primary beam”, and commonly consists of one of Cs^+ , Rb^+ . These ions are injected perpendicular to the toroidal magnetic field (that is to say, in line with the poloidal magnetic field). The idea behind this is that these positively charged atoms (I^+) will interact with the negatively charged electrons in our plasma in a way that we can measure, which will provide us insight into the properties of electron populations in our plasma. When I^+ interacts (collides) with an electron in the plasma, it is likely to become ionised further, becoming I^{2+} . The property in our plasma that we then exploit to measure this change is the Lorentz force, which, to recall, states

$$\vec{F} = q(\vec{E} + \vec{u} \times \vec{B})$$

where, crucially, q represents a particle’s charge as it moves through an magnetic field B with velocity \vec{u} , and an electric field E present. From this we can infer that the forces acting on a particle will increase as the charge increases. The effects of this increased ionisation of our heavy ions thus increases their Larmor radius (oscillatory movement of charged particles in a magnetic field) relative to their single-ionised counterparts, which changes their trajectory in the plasma. In our Tokamak we then position a charged particle detector on the opposite side of our heavy ion beam’s injection point, which we use to measure distributions of received charged particles. Form this we can infer information about the presence of negatively charged particles (predominantly electrons) in the $1D$ profile that our primary beam follows by measuring the quantity of charged particles detected at different positions on our detector. Ingenious! A graphic depicting this is given in figure 5.1.

While here we’ve provided what is hopefully an intuitive understanding for how a heavy ion beam is utilised for data collection in a Tokamak, we have skipped much of the nitty-gritty details for actual data derivation. Most importantly for us to note however, is that ISTTOK does not directly measure the current or pressure density profiles, but instead infers this data from poloidal magnetic field profile information afforded to us by the heavy ion beam. As one might expect, there is a large margin for error introduced in these measurements (and

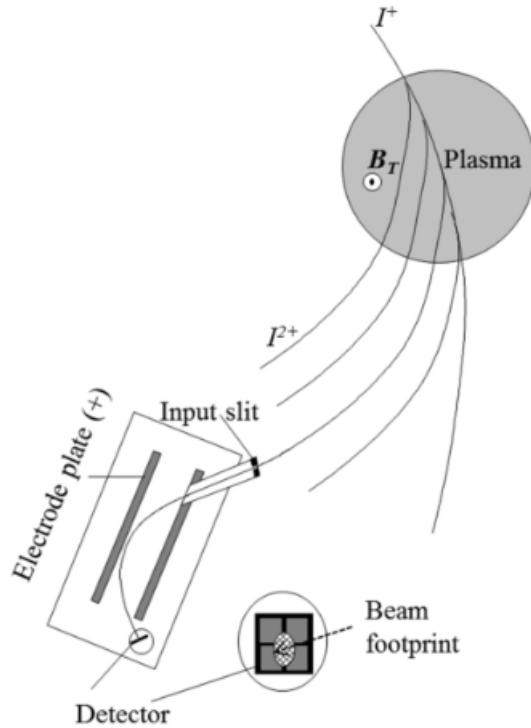


Figure 5.1: Simplified depiction of an ion beam's injected ions having their trajectories manipulated by the plasma they travel through, and how that is subsequently detected. Diagram taken from [19].

subsequent calculations) when doing so. Quoted unofficially, we can expect errors in current density profile data to be up to an order of $\pm 30\%$. While we don't have access to the explicit uncertainty in our data, we can keep this in mind when fitting our parameters (a_1, a_2, α) to the ISTTOK data, and justify some decisions we make further down.

5.1.2 Reactor Specification and Data

Reactor Specification

The ISTTOK reactor specification is given below:

Property	Value
Major Radius (a)	0.46m
Minor Radius (R_0)	0.085m
Plasma Current (I_p)	4 – 6kA

Table 5.1: Relevant ISTTOK reactor parameters [18]

Of particular emphasis, this configuration satisfies the requirement for a large

aspect ratio tokamak as required for our Grad-Shafranov-Helmholtz model, with $A = R/a = 5.4 \gg 1$ [30]. Note that aspect ratio in this paper refers to the ratio between major and minor radius, and not the aspect ratio of the cross poloidal cross section.

Available Data

The data we have is for a series of 20 time slices from 30.53ms – 31.93ms of a run of ISTTOK. Of interest, the average confinement time for a plasma in ISTTOK is approximately 0.4ms [18]. For each of these time slices there are 18 measurements at uniformly distributed discrete intervals across the cross section of the reactor; for $x \in [-8.5\text{cm}, 8.5\text{cm}]$ and $z = 0$. For this domain we have data for the current density profile, and the pressure density profile. From this data, we additionally have inferred I_p and V_{loop} data for each time slice.

In our simulations we will use the current density profile data to fit our parameters (a_1, a_2, α) . In the work that we did we did not initially seek to additionally use the pressure density profile data in fitting these parameters for two reasons:

1. We wish to have some data to compare our results to. From the (a_1, a_2, α) we get from our minimisation algorithm we can calculate a pressure density profile and compare it to the experimental data for accuracy
2. In our simulations we explored the case of having limited current density profile data and attempting to fit parameters to that, and found that we were able to determine parameters sufficiently accurate to represent the system.

While we will make comparisons with the pressure density profile data, another avenue we could have explored was to incorporate the pressure density profile into the data fitting. Initial testing with this did not improve convergence of the optimisation algorithm for those cases in which it struggled, however our testing was also not exhaustive in this manner, and the inclusion of this data combined with some changes in the approach to fitting our model to the data, could well result in a more accurate result. Ideas for further work around this are discussed in chapter 6.

A representative time slice is provided in figure 5.2

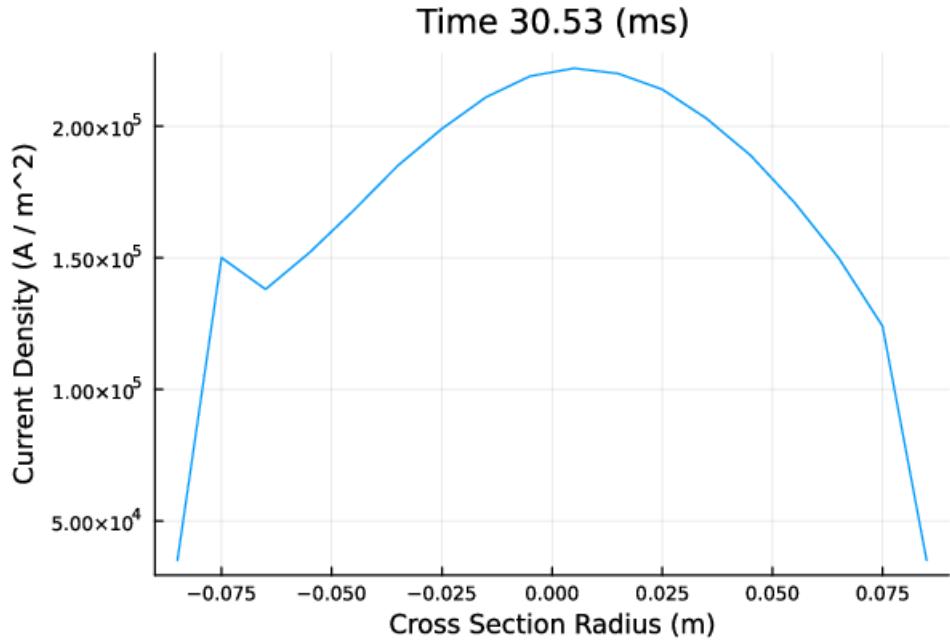


Figure 5.2: Current density profile for the beginning of a ramp down in ISTTOK. The cross section radius is centred with respect to the major radius.

The I_p time evolution is also provided for reference in figure 5.3. This showcases how the plasma's current changes during the ramp down phase.

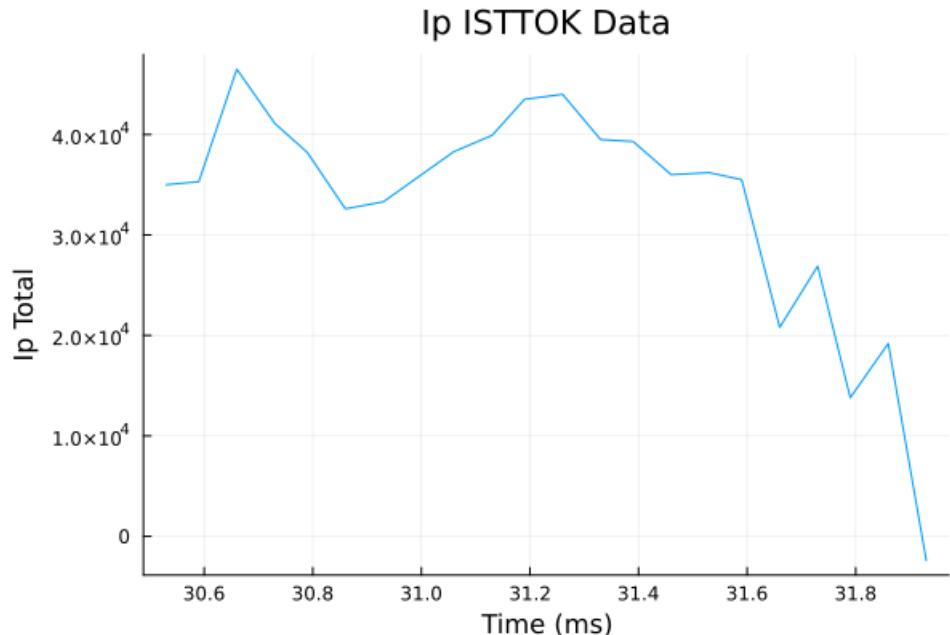


Figure 5.3: Plasma current (I_p) during the ramp down phase of a run of the ISTTOK reactor in AC operating mode.

5.2 Parameter Fitting

5.2.1 Results

Initial Attempts

We use the same process as we did when initially fitting the parameters (a_1, a_2, α) to data, though instead of contrived current inversion data we of course use the data we have from ISTTOK here. Our initial results showed a resistance of our model to fitting to the data, as can be seen by the poor fit for the current density profile in figure 5.4. All of our results are available in the source code under under the “experiments/isttok/current_solving_comparison/graphs/” directory [20]. Here we will highlight a couple representative examples, but will discuss the general trend of all our results.

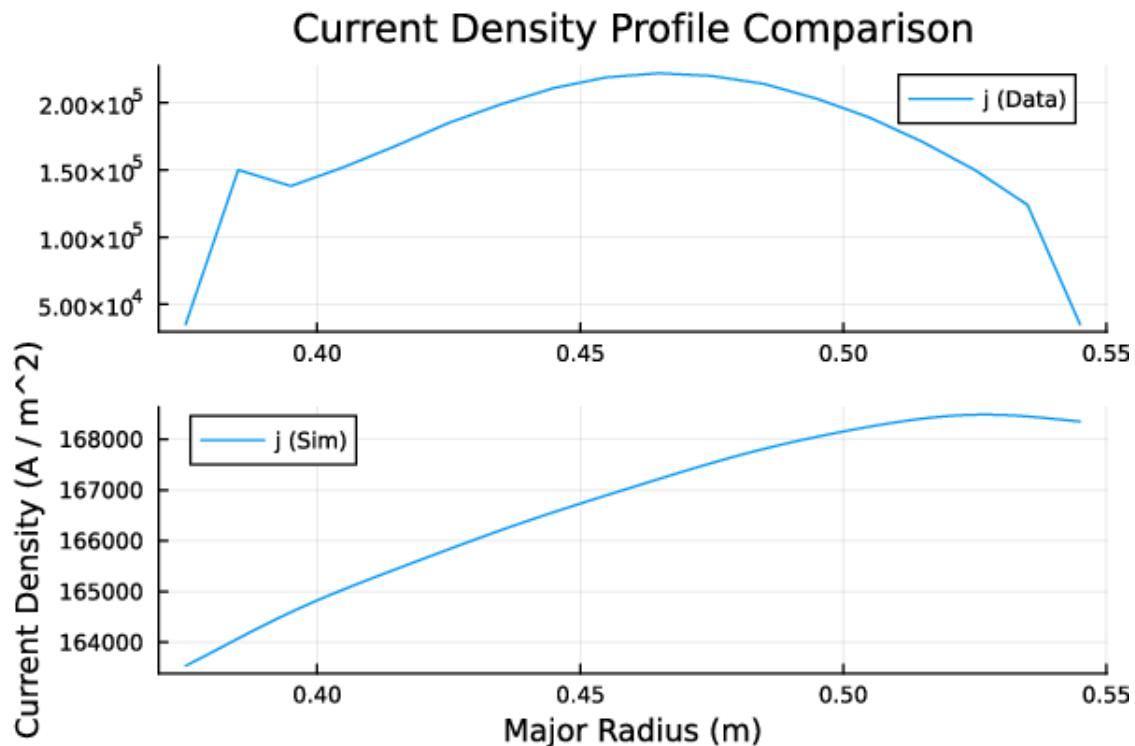


Figure 5.4: ISTTOK current density profile contrasted with our simulated current density profile after fitting against the former in order to derive (a_1, a_2, α) . This is the first time slice, i.e. at the start of the ramp down.

Purely by visual inspection we can see this is not a close fit, however this discrepancy becomes more pronounced as we progress through the remainder of our time slices. This effect can be seen in figure 5.5, where our simulation is even

more considerably differentiated from the data.

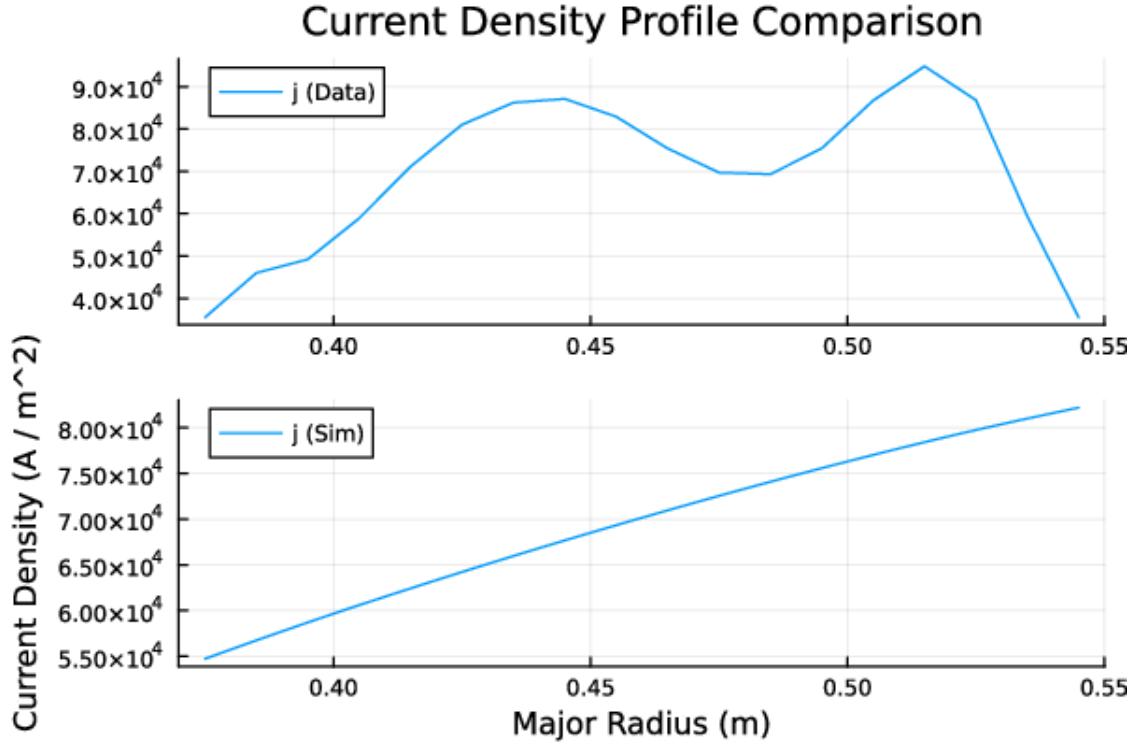


Figure 5.5: ISTTOK current density profile contrasted with our simulated current density profile after fitting against the former in order to derive (a_1, a_2, α) . This is the 15th time slice, i.e. approximately half way through the transition.

Here we begin to see a current hole develop at (what we for now presume, but will later observe) the position of the on-axis magnetic island. This is an important feature to be able to characterise in our simulations, however is clearly missed in our simulated current. In fact, it seems that the general features of our simulated current density profile remain largely unchanged from our initial results - we will comment on why this is so in a minute.

We can also (for completeness here) compare the last time slice, which is the first and only data entry we have showing the start of the ramp up phase, or in other words, the period immediately proceeding the conclusion of the ramp down phase we are investigating. This comparison is shown in figure 5.6.

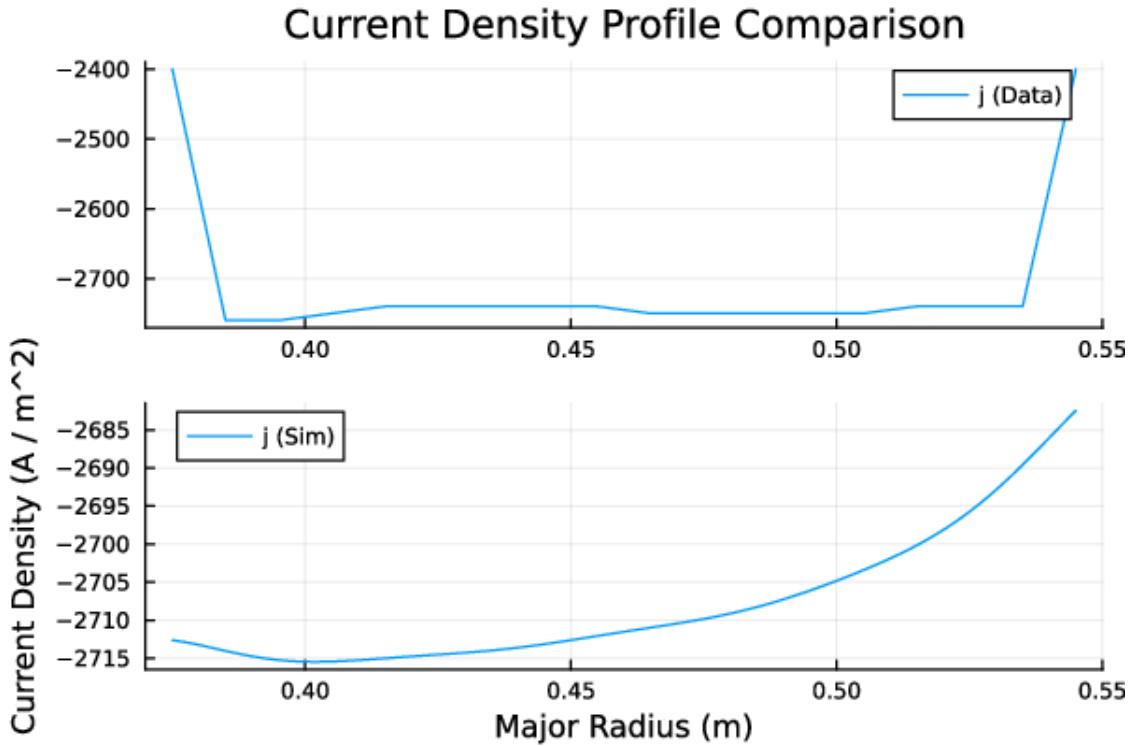


Figure 5.6: ISTTOK current density profile contrasted with our simulated current density profile after fitting against the former in order to derive (a_1, a_2, α) . This is the last time series data entry we have, and is after the ramp down phase has concluded / the start of the next cycle's ramp up phase.

Recall that, while the initial parameter guess is currently hand picked (with note on alternatives proposed in chapter 6), parameter guesses for subsequent iterations of our simulation are informed by the previously solved-for parameters. This means any initial error in choice of parameters will be detrimental to all following simulations.

We can see the effects of this more prominently by also observing the change in magnetic field topology as the current and pressure data vary. Or rather, we should say, the lack thereof. Figures 5.7 - 5.9 show the poloidal magnetic field topology for the representative time slices we've picked. Note again that results for all time slices can be found in the open sourced GitHub repository in [20].

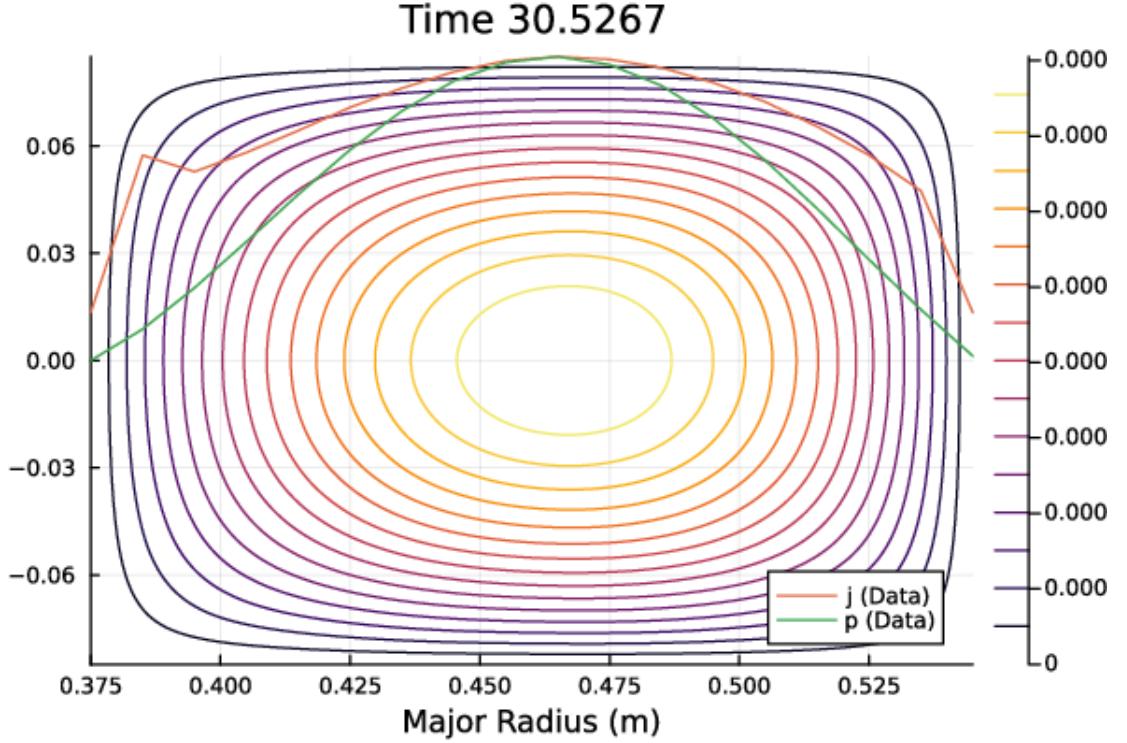


Figure 5.7: Poloidal magnetic field (simulated) for a given (a_1, a_2, α) parameter set which is derived from the current density profile data present in the graph. The pressure profile is similarly given. Both current and pressure density profiles are normalised.

We find there is minimal variation in magnetic field line topology between all the time slices, with exception of the last slice, which sees a drastic change. There is a slight drift tendency towards the external wall of the reactor (rightwards in our graphs), which might suggest some Shafranov shift of the plasma - though, given the misrepresentative nature of the simulated magnetic field lines generally, it is unlikely this is an actual result. We can further support the argument that this is not an accurate magnetic field topology as time progresses, as the pressure density profile in figure 5.8 suggests the formation of two distinct, or at least a partially split single, plasma, with two distinct current densities (i.e., that there are at least two magnetic islands present). However, the simulated magnetic field (which remains largely unchanged from the initial simulation iteration) suggests only the existence of a single magnetic island. We should note though, that figure 5.7 shows a magnetic field line topology which is consistent with where the data suggests the primary plasma density should be - this being despite a seemingly large discrepancy between our simulated current density and the provided data.

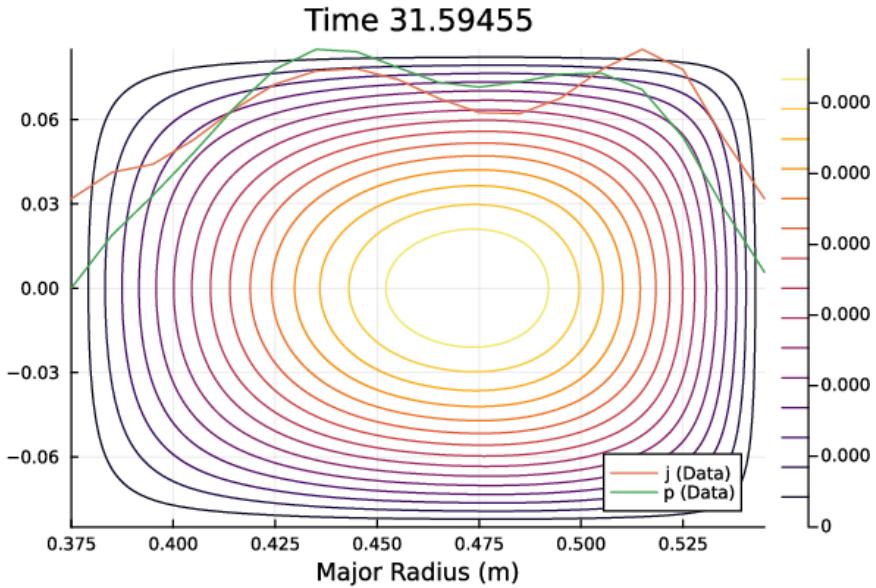


Figure 5.8: Poloidal magnetic field (simulated) for a given (a_1, a_2, α) parameter set which is derived from the current density profile data present in the graph. The pressure profile is similarly given. Both current and pressure density profiles are normalised.

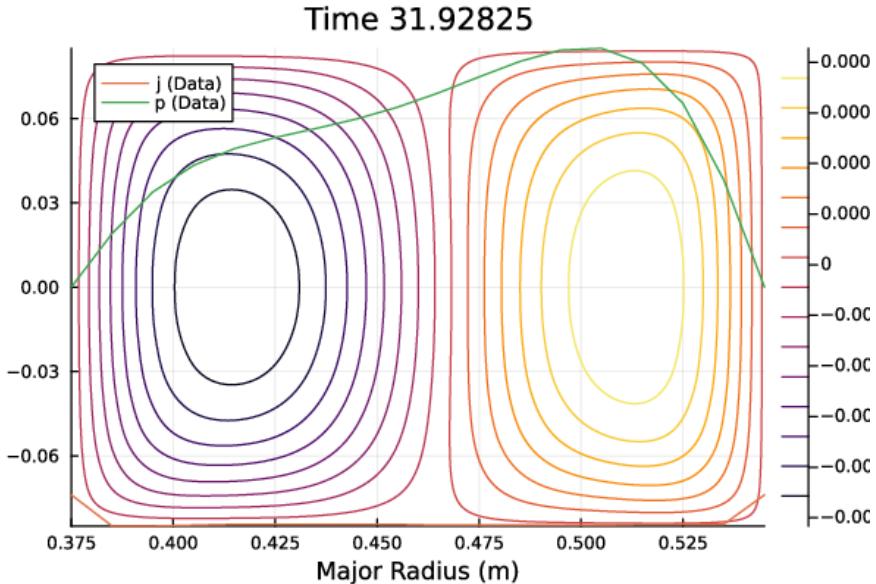


Figure 5.9: Poloidal magnetic field (simulated) for a given (a_1, a_2, α) parameter set which is derived from the current density profile data present in the graph. The pressure profile is similarly given. Both current and pressure density profiles are normalised.

Nevertheless, the result is clear - the simulation, in this form, is insufficient in modelling accurately the behaviour of the plasma, at least with respect to what

is expected from the data we have available.

Our initial efforts seem fleeting in their attempt to accurately model the system. However, that does not mean that our efforts should stop there. Whereas previously we politely requested any mathematicians avert their gaze at the crimes we were committing, it is only fair that we also at some point request any scientists to turn a blind eye to the crimes we will momentarily commit. This is that point. In the previous section we noted that there is a considerable uncertainty accompanying our data (see, for instance, figure 5.3, which intuitively we expect to be a smooth function of radial distance, yet is quite nonsmooth). We thus have some (limited) agency to manipulate our data to be more aligned with what we would expect.

Data Cleanup and Second Attempt

With the note that our data has considerably uncertainty in it, and considering the relatively few data entries we have, we might make some observations about the data we do have available, and the general physical properties we would expect it to exhibit with relation to the other data points we have available.

First, we'll note the (accusedly) extraneous data entries in the first time slice. Refer again to figure 5.2. There are two irregularities we can note:

- There is a sharp increase in current density at the -0.065m offset. Physically, unless there was anomalous behaviour inside the reactor, we would not expect this measurement, while the general behaviour of the rest of the profile relative to this entry is consistent with a plasma density being confined to the magnetic axis within the reactor. Thus, we can (to test) make an assumption that this entry is a result of imprecision in the data collection, and remove it from our fitting data.
- At the bounds of the reactor the current density decreases significantly faster than the general trend of the rest of the data. Physically this is consistent with what we would expect - that as we approach the edge of our reactor there is less plasma density and less current. However, we also know our analytic current density profile to be C^2 continuous, and such sharp jumps will affect our ability to fit to this data. Thus by removing this information we can potentially increase the performance of our data fitting.

We see the resulting current density profile in figure 5.10. It is clear that the

general behaviour of our current density profile is more consistent with that of the data now. We can observe that the magnitude of the densities is still, however, quite distinct. There are a couple considerations we could make here. First is that, this is evidence that our model is not fitting to the data as closely as we'd like yet, and so extra work can be done to attempt to improve this. The second would be that, while the scale of the current density is inconsistent with the data, it may be that we can still sufficiently make statements about the generation of runaway electron populations, and so it perhaps is (for our intended purposes) irrelevant whether the scale is precise or not. More important to us is the topology of the poloidal magnetic field, which in our model is informed more by the radial positioning of current density masses than the strength of those densities (as we highlighted in our simulations section). We can see in figure 5.11 that the magnetic field topology for the manually edited data is consistent with where the pressure and current density profiles suggest the primary magnetic island should be. Additionally, it appears this is more in agreement than the pre-processed data, as the positioning of the primary magnetic island is more accurate to what we expect than we observed in figure 5.7.

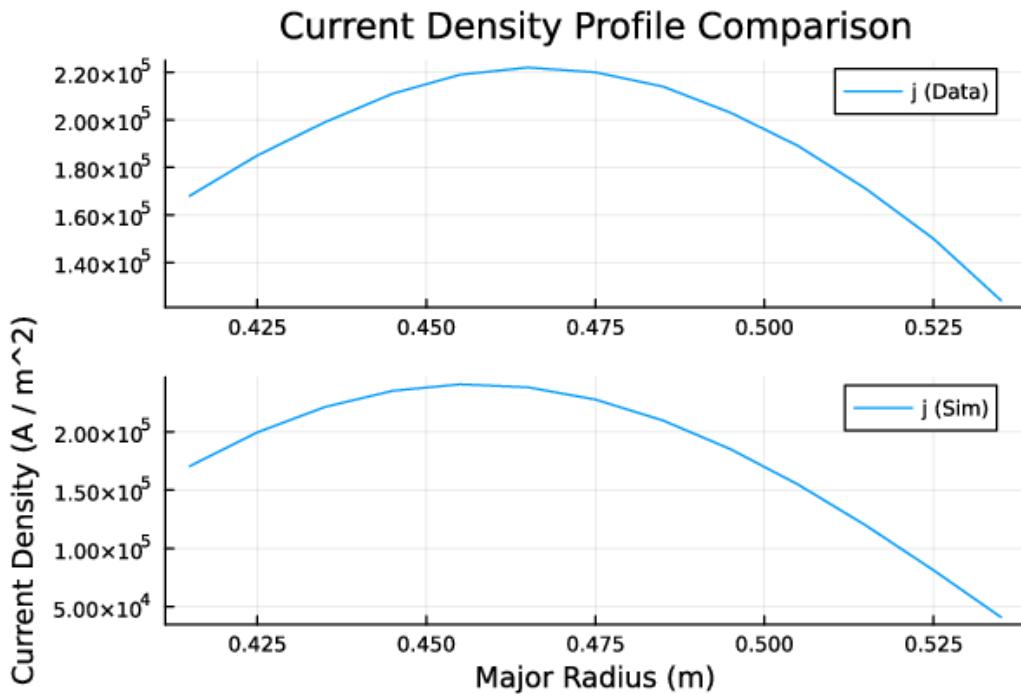


Figure 5.10: Comparison of simulated and data current density profiles, after we rid the data of potentially erroneous data entries.

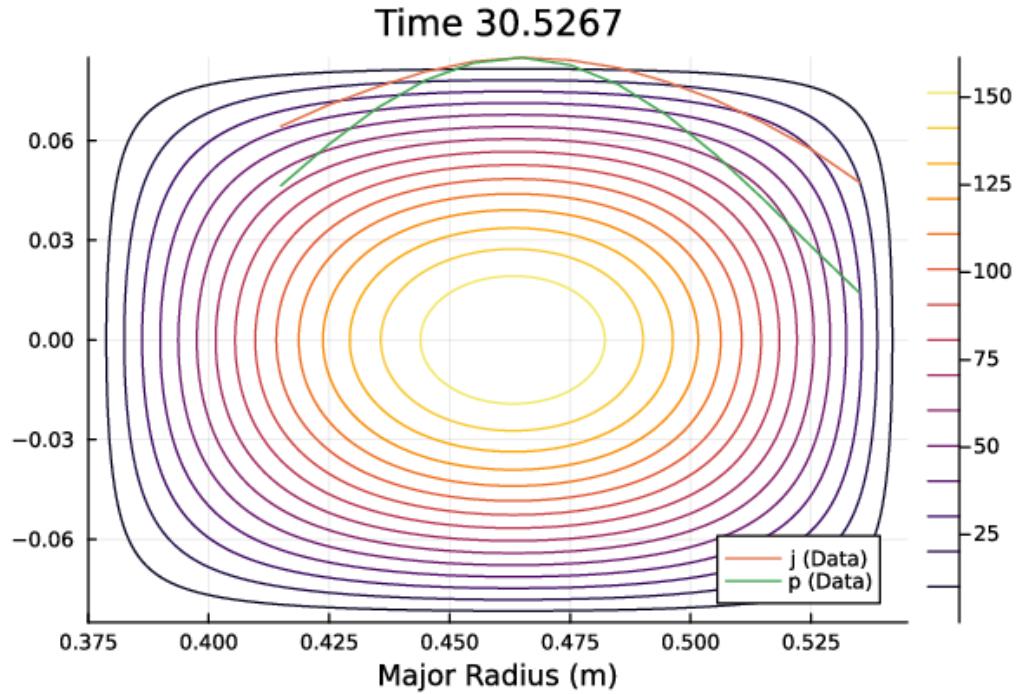


Figure 5.11: Simulated magnetic field topology for the stripped data. Current and pressure density profiles given are normalised, and provided by the data.

These (albeit preliminary) results would seem to suggest that, after some “cleaning” of the data, our model is able to produce results that are at least consistent with how we would expect our system to behave given the data.

Nevertheless we face another problem which, as of writing, remains unresolved. While we are able to clean our data for the first time slice as what we deemed “erroneous” data entries were easy to identify, this is not the case for the majority of the remaining data we have. This is partly a problem of labour, more problematic is the that of correctly determining which data points to classify as erroneous - if any can even be deemed so.

Alternative Data Coddling Methods and Steps Forward

Here we have simply removed potentially erroneous data points from our set, and continued on as if they never existed. However, especially as we don’t have access to a great many discrete measurements, this is not an ideal situation. There are alternative approaches to post-processing the data that could be used, and comparisons made to deem which is most suitable for our data fitting purposes. For example, it could be prudent to, after removing these erroneous data points, perform an interpolation using the remaining points. This would have the benefit

of providing a data set with data that was 1. more volumous, and 2. smoother.

We could also interpolate in another dimension - time. One potential explanation for the difficulty our model has in moving between time slices in the data is that the time variation between them is too large with respect to the time perturbation expansion we performed in chapter 3. As such, we could take two consecutive time slices of data and linearly interpolate them, introducing many more time slices in between. This would restrict our parameter space further, improving our optimisation, and, combined with the previous point about interpolating between data points, could see an improvement on the accuracy of our model.

Additionally, we face the issue that the simulation is very sensitive to the initial parameters chosen. This is directly related to the parameter space issue we identified in the chapter 4, which is that discontinuities introduced by the α parameter lead to a high false positive rate. For the above simulations we have achieved some results by hand picking an initial parameter guess via manual inspection of the parameter space - this, however, is obviously insufficient for more general simulations. A potential resolution for this is proposed in the next chapter, and is the subject of ongoing work.

Chapter 6

Blue Skies and Horizons

6.1 Conclusions

In this thesis we've built up an ability to reason about perturbations in time about equilibrium solutions to a variant of the Grad-Shafranov Equation. We did this with the intent of modelling the ISTTOK reactor's ramp down phase as a tool for identifying causal mechanisms for the generation of runaway electrons. Our model, using current density profile data provided by the ISTTOK project, is able to infer the topology of the poloidal magnetic field, with accompanying pressure density profile, though the accuracy of the model as it stands with respect to inputted data is in question. Our efforts in simulating the change in poloidal magnetic field topology as the current density profile varies nevertheless suggests there are mechanisms induced by the current inversion that can generate runaway electrons, providing one potential theoretical explanation for the observed spikes in runaway electron populations as observed by ISTTOK [18]. However, we can make no concrete statements in relation to the presence of two anti-parallel currents in the plasma.

There is still a lot of work that could be done to improve this model and compare it to literature in a more robust fashion. Some of the efforts that could be made to improve upon the work in this thesis are now presented.

6.2 Further Work

6.2.1 Simulated Electric Field via Fake Solenoids

In a meeting while presenting my findings to the Plasma Science group, I posed the question of extracting electric field information from the data we had available. There are many benefits to being able to describe the electric field for the confinement time of a plasma, however the most significant to our purpose is to positively identify the birth of runaway electrons. This information however is not readily available with the system we worked with.

David Pfefferle proposed a method of simulating the presence of an electric field instead. His idea was to introduce infinitesimal solenoids at the centre of magnetic islands, which would each contribute produce their own electric field. These would then interact, with the idea that the product would be an approximation to the expected electric field for a given state.

The physical justification for this is that the solenoid's magnetic field is emulating the magnetic field produced by current densities, which are themselves informed by magnetic islands. A toroidal current will produce an electric field, including a poloidal component, which will influence the behaviour of runaway electrons. Thus, this approach effectively emulates the presence of a poloidal electric field using position and strength information of current densities.

This approach could utilise work done by Nicholas Bohlsen in using topological data analysis to identify the presence of magnetic islands from poloidal magnetic flux data in his thesis.

6.2.2 Comparison with ISTTOK v_{loop} data

In the appendix of Wang's paper there are small extensions to their results if some slightly different assumptions are made about the model [30]. If these assumptions are made, we are presented with the ability to both derive information about the electric field, and calculate v_{loop} values. Both of these are of particular interest to us; if we are able to retrieve information about the electric field then we can use that to reason about the generation of runaway electrons directly by inspecting the strength of the electric field under certain conditions. Additionally, with the ability to calculate v_{loop} data, we would have another metric by which to test the accuracy of our simulation against experiment, as the ISTTOK data we have contained v_{loop} information, inferred from the raw data provided. This would more justify us in using pressure density profile data

6.2.3 Grid Based Initial Parameter Guessing

One of the identified problems with our model, at both the simulation stage and the data fitting, was that the initial parameter guess is just that; a guess. Due to the parameter space we are dealing with the system is very sensitive to this initial choice, and so it is crucial that a suitable initial guess is chosen for the optimisation algorithm to not fall into false positives. As calculating residuals is relatively cheap to perform, and we know our objective function has a smooth parameter space (with exception for $(\alpha^2 - \lambda_{n,l}^2)^{-1} = 0$ cases), we can utilise a grid-based search approach for identifying what could potentially be a good initial guess. If we describe some domain $\Sigma := \Sigma_1 \times \Sigma_2 \times \Sigma_3$ such that $(a_1, a_2, a) \in \Sigma$, then we can segment this cuboid into smaller, uniformly sized cuboids. We could then pick a “representative point” from these smaller cuboids (for simplicity, say the centre), and calculate the residual for that specific parameter set with respect to the provided data. Then we pick the parameter set that has the smallest residual, and use that parameter set as the initial guess for our optimisation algorithm.

Appendix A

Appendix

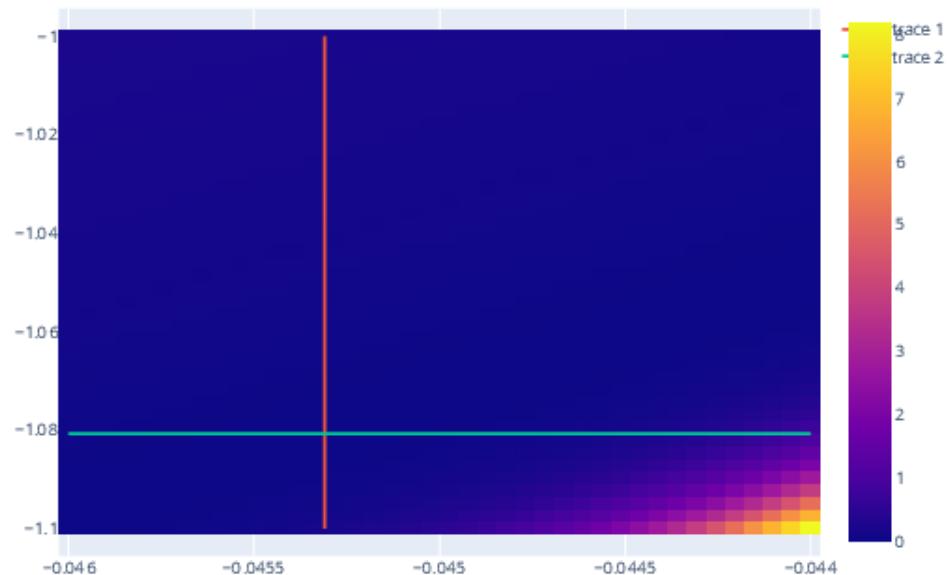


Figure A.1: Objective value heatmap for fixed α , with varying a_1 and a_2 , zoomed around expected values.

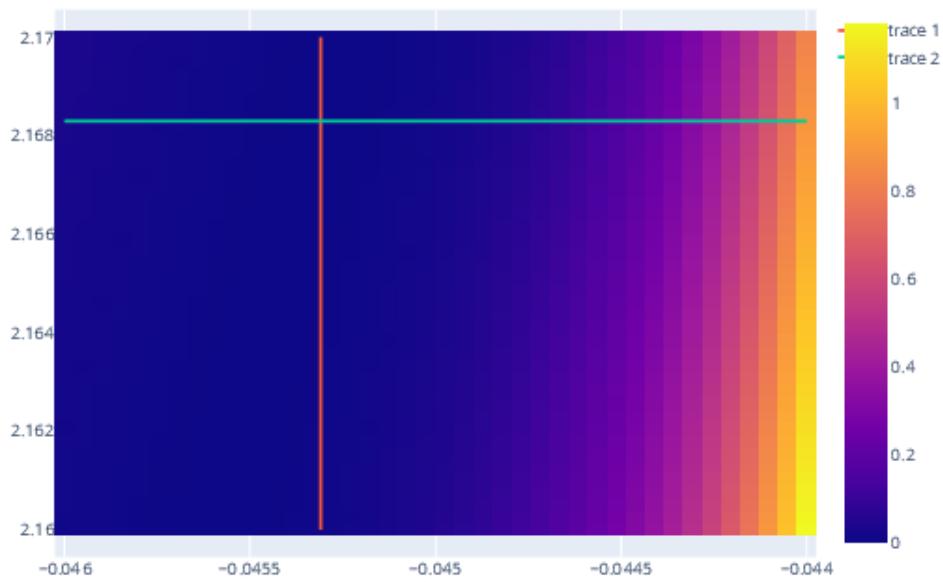


Figure A.2: Objective value heatmap for fixed a_2 , with varying a_1 and α , zoomed around expected values.

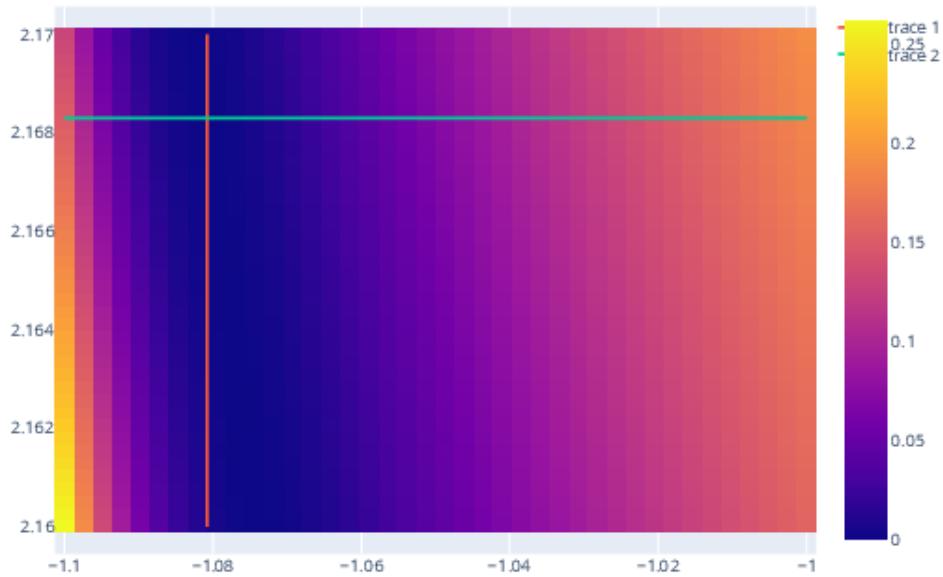


Figure A.3: Objective value heatmap for fixed a_1 , with varying a_2 and α , zoomed around expected values.

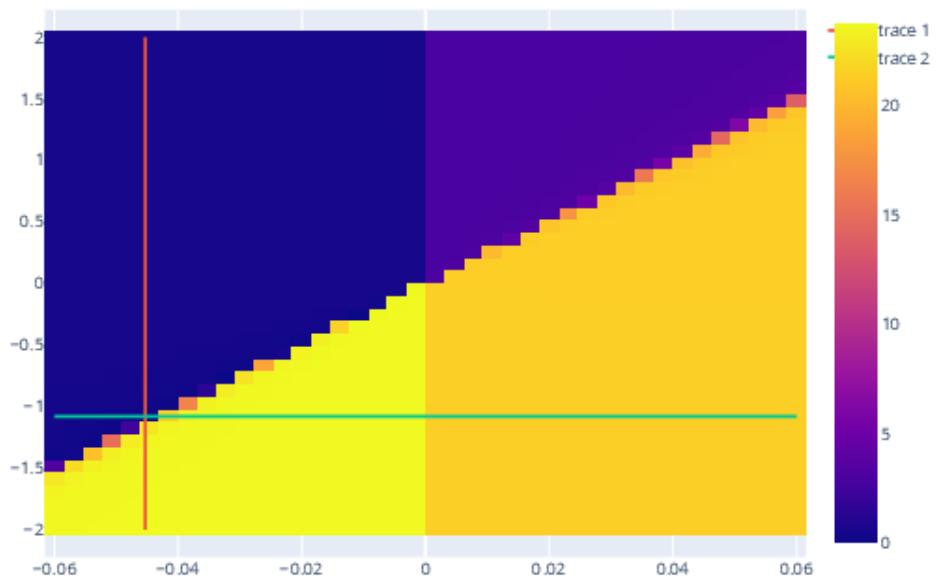


Figure A.4: Objective value heatmap for fixed α , with varying a_1 and a_2 , with pressure density profile data included.

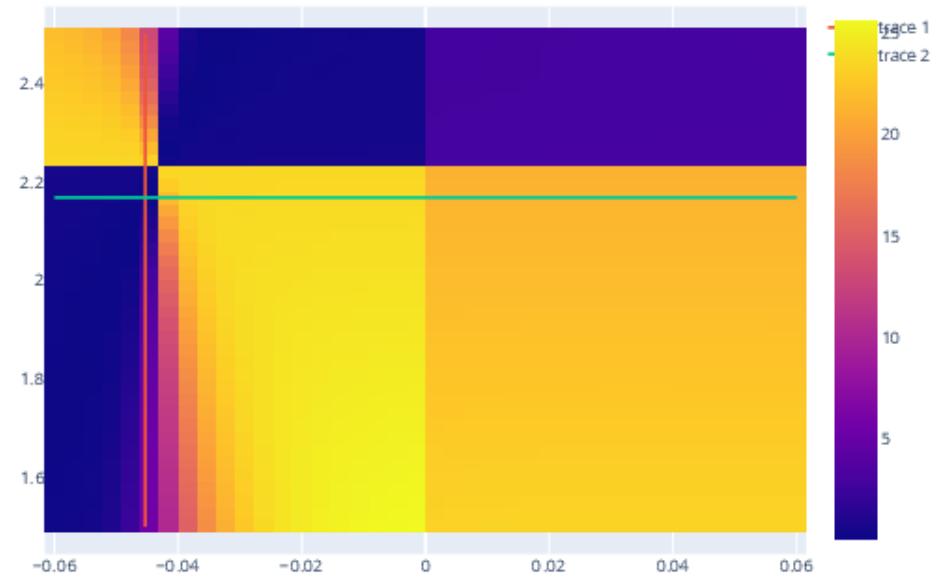


Figure A.5: Objective value heatmap for fixed a_2 , with varying a_1 and α , with pressure density profile data included.

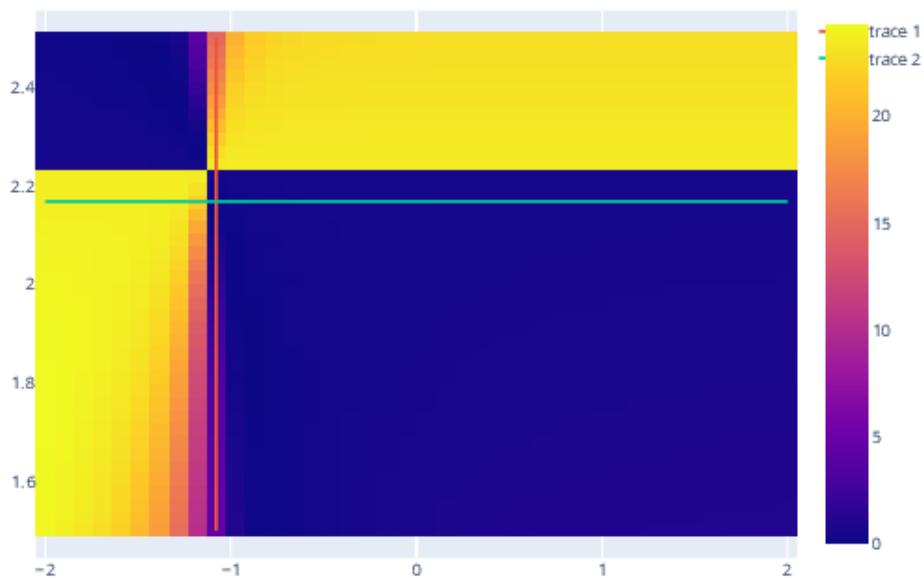


Figure A.6: Objective value heatmap for fixed a_1 , with varying a_2 and α , with pressure density profile data included.

Bibliography

- [1] H. ALFVÉN. Existence of electromagnetic-hydrodynamic waves. *Nature*, 150(3805):405–406, Oct 1942.
- [2] Allen H. Boozer. Runaway electrons and magnetic island confinement. *Physics of Plasmas*, 23(8), 8 2016.
- [3] Luca Carbone. Lcture 17-18: Least squares optimization. University Website at <https://vnav.mit.edu/material/17-18-NonLinearLeastSquares-notes.pdf>, 2021.
- [4] S. L. Chen, F. Villone, B. J. Xiao, L. Barbato, Z. P. Luo, L. Liu, S. Mastrotostefano, and Z. Xing. 3d passive stabilization of $n=0$ mhd modes in east tokamak. *Scientific Reports*, 6(1):32440, Sep 2016.
- [5] Pavel Cherenkov. Radiation of particles moving at a velocity exceeding that of light, and some of the possibilities for their use in experimental physics. Nobel Lecture, 1958.
- [6] C. Chiuderi and M. Velli. *Basics of Plasma Astrophysics*. Springer Milan, 2016.
- [7] William Cropper. *Great Physicists: The Life and Times of Leading Physicists from Galileo to Hawking*. Oxford University Press, 2004.
- [8] C. L. da Silva, R. G. Sonnenfeld, H. E. Edens, P. R. Krehbiel, M. G. Quick, and W. J. Koshak. The plasma nature of lightning channels and the resulting nonlinear resistance. *Journal of Geophysical Research: Atmospheres*, 124(16):9442–9463, 2019.
- [9] Dama Daniel. Derivation and applications of grad-shafranov equation in magnetohydrodynamics (mhd). *Journal of Research in Applied Mathematics*, 7(4), 2021.

- [10] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.
- [11] G. W. Hammett, S. C. Jardin, and B. C. Stratton. Non-existence of normal tokamak equilibria with negative central current. *Physics of Plasmas*, 10(10):4048–4052, 09 2003.
- [12] John Howard. Introduction to plasma physics. via ANU website at <https://people.physics.anu.edu.au/~jnh112/AIIM/c17/>, 2002.
- [13] Jianguo Huang, Xuanzong Yang, Shaobai Zheng, Chunhua Feng, Houxian Zhang, and Long Wang. The plasma current profile during current reversal in ac operation of the ct-6b tokamak. *Nuclear Fusion*, 40(12):2023, dec 2000.
- [14] S. R. Hudson, R. L. Dewar, G. Dennis, M. J. Hole, M. McGann, G. von Nessi, and S. Lazerson. Computation of multi-region relaxed magnetohydrodynamic equilibria. *Physics of Plasmas*, 19(11):112502, 11 2012.
- [15] ITER. Tokamak, 2023. <https://www.iter.org/mach/Tokamak>.
- [16] Steven G. Johnson. The NLopt nonlinear-optimization package. <https://github.com/stevengj/nlopt>, 2007.
- [17] Joseph Mahaffy. Ordinary differential equations - lecture notes - perturbation methods. University Website via <https://jmahaffy.sdsu.edu/courses/f19/math537/beamer/perturb.pdf>, 2019.
- [18] A Malaquias, M Hole, R B Henriques, I S Nedzelskiy, and V V Plyusnin. Experiments during ac current transition in isttok and the hypothesis of ballistic runaway electrons. *Plasma Physics and Controlled Fusion*, 64(9):095013, aug 2022.
- [19] A. Malaquias, I. S. Nedzelskiy, R. Henriques, and R. Sharma. The heavy-ion beam diagnostic of the isttok tokamak highlights and recent developments. *Sensors*, 22(11), 2022.
- [20] Tom Malcolm. Thesis source code. via GitHub at <https://github.com/Thomas-Malcolm/thesis-code>, 2023.
- [21] A. A. Martynov, S. Yu. Medvedev, and L. Villard. Tokamak equilibria with reversed current density. *Phys. Rev. Lett.*, 91:085004, Aug 2003.

- [22] Merriam-Webster. Plasma, 2023. <https://www.merriam-webster.com/dictionary/plasma>.
- [23] NASA. Fusion chemistry - a closer look, 2023. https://solarsystem.nasa.gov/genesismission/gm2/science/sunlight_solar-heat/fusion_chemistry.htm.
- [24] V V Plyusnin, L Jakubowski, J Zebrowski, P Duarte, K Malinowski, H Fernandes, C Silva, M Rabinski, and M J Sadowski. Characteristics of four-channel cherenkov-type detector for measurements of runaway electrons in the ISTTOK tokamak. *Rev Sci Instrum*, 81(10), October 2010.
- [25] Dalton D Schnack. *Lectures in Magnetohydrodynamics: With an Appendix on Extended MHD*. Lecture notes in physics. Springer, Berlin, Heidelberg, 2009.
- [26] V. D. Shafranov. Plasma Equilibrium in a Magnetic Field. *Reviews of Plasma Physics*, 2:103, January 1966.
- [27] Krister Svanberg. The method of moving asymptotes - a new method for structural optimization. *International Journal for Numerical Methods in Engineering*, 24, 1987.
- [28] Neil Trudinger. Fully nonlinear pdes in geometry. Lecture Slides, 2004.
- [29] Shaojie Wang. Theory of tokamak equilibria with central current density reversal. *Phys. Rev. Lett.*, 93:155007, Oct 2004.
- [30] Shaojie Wang and Jun Yu. An exact solution of the Grad-Shafranov-Helmhotz equation with central current density reversal. *Physics of Plasmas*, 12(6):062501, 05 2005.
- [31] J. Wesson and D.J. Campbell. *Tokamaks*. International series of monographs on physics. Clarendon Press, 2004.
- [32] Helen Wilson. Introduction to perturbation methods. University Website via <https://www.ucl.ac.uk/~ucahhwi/LTCC/section2-3-perturb-regular.pdf>, 2014.
- [33] J Zebrowski, L Jakubowski, M Rabinski, M J Sadowski, M J Jakubowski, R Kwiatkowski, K Malinowski, R Mirowski, J Mlynar, O Ficker, V Weinzettl, F Causa, COMPASS, and FTU Teams. Studies of runaway electrons

via cherenkov effect in tokamaks. *Journal of Physics: Conference Series*, 959(1):012002, jan 2018.