Nonlinear wave-particle resonance in deterministic and stochastic kinetic plasmas

Benjamin Jabulani Qondi Woods

Doctor of Philosophy

University of York
Physics

August 2019

Abstract

In kinetic plasma physics, BGK modes are ubiquitous solutions to the Vlasov equation, with particles travelling along orbits where the single particle energy is conserved. Approximate extensions of these exact solutions have been successfully used in the past to understand the formation and evolution of 'holes' and 'clumps', coherent structures on the particle distribution function which under certain conditions form in the nonlinear phase of the evolution of kinetic plasmas. In this thesis, analytical results are shown which consider perturbations and deformations to BGK orbits, allowing one to robustly construct more exotic orbits that allow for mode growth and frequency chirping. Computational results produced using the DARK code are presented, examining stochastic and deterministic populations in a 1D electrostatic plasma, and how they affect electrostatic waves exhibiting Landau resonance, based on Berk-Breizman models. A model is presented for parametric mode-mode destabilisation via holes and clumps interacting via the background distribution. Finally, work using the machine learning framework ERICSON is presented, analysing frequency spectrograms of magnetic perturbations in Alfvénic and sub-Alfvénic frequency ranges.

Contents

List of Figures

Preface

I started this project with a largely self taught knowledge of plasma physics, having read F. Chen's Introduction to Plasma Physics and Tokamaks by Wesson (of which I understood the first half, and a smattering of points respectively). In my third year of my undergraduate degree, I contacted S. Lebedev and A. Schekochihin who gave me valuable pointers on plasma physics, and allowed me to gain enough confidence to apply to study at doctorate level.

The beginning of the project was not without difficulty; the theory was quite different from my undergraduate research at the University of Exeter in experimental metamaterials and theoretical nanomaterials, and required me to look at problems from a largely numerical approach. However, I feel that my undergraduate background outside of plasma physics has allowed me to look at some of the problems engaged in this thesis from a more eclectic perspective.

This thesis is equipped with a glossary for the reader. In the digital version, many of the terms in the thesis are hyperlinked to the glossary, should the reader wish to quickly look up previously mentioned concepts.

Acknowledgments

I would like to thank first and foremost my supervisor Roddy Vann for his guidance over the course of this ever-evolving project. His support has proven invaluable as the work branched into unfamiliar lands.

Secondly, I am indebted to Vinícius Duarte, who has proven time and time again to be not only an outstanding physicist, but also a remarkable person. Thank you for everything.

Thirdly, I would like to thank Cristina Rea and Orso Meneghini, who have provided great discussions involving machine learning. The work both of these physicists do is in my opinion integral to the success of fusion in the future, and I hope to see their work for years to come.

Fourthly, I would like to formally acknowledge my sources of funding: EPSRC and EUROfusion.

Fifthly, I would like to thank everyone at the York Plasma Institute, and I would also like to thank other contacts and colleagues in the wider plasma physics community: M. Cole, A. Frank, E. Fredrickson, N. N. Gorelenkov, H. Hezaveh, M. Parsons, M. Podestà, C. Smiet, and many others.

Penultimately, I would like thank a few people who have proven instrumental in the path I have taken in life today:

- Graham Tavener, who inspired me to become a physicist
- Harriet Smith, who advised me not to study physics at university
- Sharon Strawbridge, who is the most inspiring academic I have had the pleasure of knowing
- Joe Dockrey, a fantastic scientist who motivated me to study for a doctorate
- Sergei Lebedev, who encouraged me to pursue plasma physics
- Joe Hanson, who introduced me to machine learning

Finally and most importantly, I would like to thank my parents Cynthia and Finlay. I am forever humbled by your strength of character, your tireless support, and the sacrifices and experiences you have faced. This work is, and always has been, for you.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

The work was primarily carried out at the York Plasma Institute, Department of Physics, University of York. Direct collaboration occurred at the Princeton Plasma Physics Laboratory (PPPL) with V. N. Duarte, E. Fredrickson, N. N. Gorelenkov, and M. Podestà.

The following publications were produced over the course of the thesis:

- B. J. Q. Woods, V. N. Duarte, N. N. Gorelenkov, A. P. De-Gol, and R. G. L. Vann. "Stochastic effects on phase-space holes and clumps". In: *Nucl. Fusion* 58.8 (2018), p. 082015.
- B. J. Q. Woods, V. N. Duarte, E. D. Fredrickson, N. N. Gorelenkov, M. Podestà, and R. G. L. Vann. "Machine learning characterisation of Alfvénic and sub-Alfvénic chirping and correlation with fast ion loss at NSTX". In: *IEEE Trans. Plasma Sci.* 48.1 (2020), pp. 71-81.
- B. J. Q. Woods. "Analytical solutions for nonlinear plasma waves with time-varying complex frequency". In: *Plasma Res. Express* 1.4 (2019), p. 045003.

Chapter 1

Introduction

1.1 Overview

This thesis began life as a study examining 1D electrostatic plasmas for the purpose of furthering understanding of tokamak plasmas. While some work was carried out to this end (see Chapter 5 and Chapter 6), this thesis is largely a mathematical work examining the origin of emergent phenomena in kinetic plasmas. These emergent phenomena are investigated analytically via generating functions and collective orbit effects (see Chapter 3), computationally through stochastic and deterministic means (see Chapter 4 and Chapter 5), and with machine learning (ML) techniques (see Chapter 6).

While this appears to make for quite a broad thesis, it is my belief that all of the approaches are merely manifestations of a few fundamental principles: composite mappings, nonlinear physics, and spectral analysis. Throughout this thesis, I reinforce these three concepts by treating almost all problems through the same rigorous mathematical notation.

However, this is not entirely an academic study, and the work was carried out for the express purpose of producing verifiable, real physics. To this end, this introductory chapter begins in Section 1.2 by covering motivation from a fundamental plasma physics perspective, with some additional motivation for tokamak plasmas; further motivation for the kinetic work and tokamak plasma work is included separately at the beginning of Chapter 2 to Chapter 6. Later in Section 1.3, I outline definitions, notation and shorthand which is used throughout the thesis. In Section 1.4, I derive the Boltzmann-Maxwell system from the background theory of kinetic plasmas. Finally, in Section 1.6, I discuss action-angle variables and how they are used with Noether's theorem[1] to analyse plasmas in a Hamiltonian framework.

1.2 Motivation

1.2.1 Nonlinear kinetic theory

The phenomenon of solitons is widely documented,[2–4] where nonlinear structures form due to spatial coupling between waves in a given system. In kinetic plasmas, nonlinear coupling in momentum space allows for structures to form on the particle distribution function (PDF) known as holes and clumps (H&Cs), referring to a local decrease or increase respectively on the spatially averaged PDF.[5] These structures propagate in a non-dispersive manner similar to solitons. However, as they move through momentum space they draw free energy from the system even when at constant amplitude. Understanding how these structures are formed may prove crucial for mitigating fast ion loss in tokamaks, as well as other types of kinetic instability (see Chapter 5).

The Vlasov equation[6] has been solved in previous work for the case where the particles follow orbits on which their energy is conserved.[7] These types of plasma wave, known as Bernstein-Green-Kruskal (BGK) modes, create phase space 'islands' where trapped particles resonate with the wave. Further work by Berk, Breizman and others[5,8,9] has explored the movement of H&Cs through phase space by assuming that the H&C evolve via a temporal continuum of BGK modes. While this work has proven very successful for predicting the frequency chirp associated with these structures, the work does not fully explain how these structures form, and assumes a constant slope on the PDF during structure formation. In addition, this theory is limited to systems near marginal stability; rapidly growing waves and rapidly chirping waves break the required adiabaticity conditions.

1.2.2 Nuclear fusion

Nuclear fusion is a process where two or more nuclei with binding energies per nucleon less than that of iron-56 combine to form a nucleus and daughter products which overall have less rest mass than the parent matter.

The process releases an amount of energy corresponding to the mass defect. The typical energy released per fusion event is on the order of MeV,[10] making it potentially a very large source of energy. However, overcoming the electromagnetic forces at play is tricky, requiring a great deal of finesse, and a large amount of energy. Unfortunately, due to this and other issues, nuclear fusion is not yet a viable source of commercial power. At the time of this thesis, many interesting experiments exist around the world examining a variety of devices for the purpose of eventually achieving commercial fusion, some of

which are currently under upgrade (MAST-U, NSTX-U), and some of which are currently operational (JET, DIII-D, W7-X).

Broadly speaking, the current primary roadblocks to fusion can be described as the following:

- **Suitable materials**

  Fusion neutrons are highly energetic (on the order of MeV) and can deal a large amount of damage to fusion devices. This includes ballistic damage as well as damage arising from nuclear reactions. Neutrons are ballistically devastating: the kinetic energy per unit cross sectional area of a single fusion neutron is on the order of 10^{17} $\mathrm{J\,m^{-2}}$, while for comparison the value for a bullet is about 8 orders of magnitude lower. In addition, neutron capture can lead to transmutation, rendering inert materials radioactive. As a result, finding suitable materials is a non-trivial endeavour.

- **Plasma confinement**

  Keeping plasma well confined is difficult: as highly ionised fluids, the constituent charged particles experience long range electromagnetic forces which make them susceptible to particle drifts (classical, neoclassical)[11] as well as nonlinear effects (wave-particle resonance) which can lead to confinement degradation.

- **Plasma stability**

  Emergent, transient phenomena in tokamaks and other fusion devices are highly problematic. Transients such as edge localised modes (ELMs), abrupt large events (ALEs), and runaway electrons can lead to large amounts of energy loss from the core, reducing plasma performance and causing large spikes in particle flux on the first wall of tokamaks.

This thesis focuses on the last two points: plasma confinement and stability. Obtaining a better understanding of the origin of transient phenomena such as nonlinear kinetic instabilities may allow one to mitigate them better in future, and potentially extinguish instabilities before they become problematic. While these transient phenomena occur on very rapid timescales which may be too hard to control by humans (< 10 ms), new work over the past few years using ML and neural networks is making strong strides towards real time feedback loops and system control.[12–14] With regards to confinement, understanding the mechanisms that drive wave-particle resonance better may allow one to control phase space islands and artificially move these using external influences.[15]

1.3 Definitions

1.3.1 Notation and shorthand

I will use the following shorthand here, and throughout the thesis:

- **Definition.** A quantity A is *defined* as B if written as:

$$A := B.$$

  Alternatively, $B =: A$ also has the *same* meaning.

- **Fourier series**. The Fourier coefficients under spatial decomposition of a quantity $Q(x)$ in a periodic system of length L are given by:

$$Q(x) =: \frac{1}{2}\sum_j Q_j \mathrm{e}^{\mathrm{i}k_j x}, \tag{1.1}$$

  noting that for real $Q(x)$, $Q_{-j} \equiv Q_j^*$. The coefficients are given by the following:

$$Q_j \equiv \frac{1}{L}\oint_0^L Q(x)\mathrm{e}^{-\mathrm{i}k_j x}\,\mathrm{d}x. \tag{1.2}$$

- **Laplace transforms**. The forward and backward Laplace transforms under temporal decomposition of a quantity $Q(t)$ are given respectively by:

$$\hat{L}[Q] \equiv \tilde{Q}(p) := \int_{\mathbb{R}} Q(t)\mathrm{e}^{-pt}\,\mathrm{d}t, \tag{1.3a}$$

$$\hat{L}^{-1}[Q] \equiv Q(t) \equiv \frac{1}{2\pi\mathrm{i}}\int_{\mathbb{R}+\mathrm{i}\sigma} \tilde{Q}(p)\mathrm{e}^{pt}\,\mathrm{d}p. \tag{1.3b}$$

- **Wave frame**. The component of quantity Q corresponding to the l^{th} wave frame (with phase velocity $u^{[l]}$) is denoted as $Q^{[l]}$.

- **Species**. The subscript $[l]$ is used to denote that a quantity $Q_{[l]}$ is of species l.

1.3.2 The δ-operator

Via Taylor expansion, one expects that if a functional $Q[\kappa]$ is differentiable:

$$Q[\kappa] = Q[\kappa_{(0)}] + \delta Q[\kappa_{(0)};\kappa] + \delta^2 Q[\kappa_{(0)};\kappa] + \ldots, \tag{1.4}$$

where $\kappa_{(0)}$ is a reference function. The δ-operator is defined here as returning the first variation of a functional with respect to a function κ and the reference function $\kappa_{(0)}$, such that:[16]

$$\delta Q[\kappa_{(0)};\kappa] := \int_{\Omega} \mathrm{d}\vec{\mu}\, \Delta\kappa \frac{\delta}{\delta\kappa_{(0)}(\vec{\mu})} Q[\kappa_{(0)}],$$

where $\Delta\kappa := \kappa - \kappa_{(0)}$, $\Omega \subset \mathbb{R}^D$ is a closed volume in which $\mathrm{d}\vec{\mu}$ is shorthand for the infinitesimal volume element, $\vec{\mu} \in \Omega$, and $\delta Q[\kappa]/\delta\kappa(\vec{\mu})$ is the functional derivative with respect to κ, given by:

$$\int_{\Omega} \mathrm{d}\vec{\mu}\, \xi(\vec{\mu}) \frac{\delta}{\delta\kappa(\vec{\mu})} Q[\kappa] := \lim_{\varepsilon\to 0} \frac{f[\kappa(\vec{\mu}) - \varepsilon\xi(\vec{\mu})] - f(\kappa)}{\varepsilon},$$

where for δQ, by definition the test function $\xi(\mu) := \Delta\kappa(\mu)/\varepsilon$. In general, higher order variations are therefore given by nonlocal functionals, arising from multiple iterations of the δ-operator:

$$\boxed{\delta^l Q[\kappa_{(0)};\kappa] := \frac{1}{l!} \int_{\bigcup_{m=1}^{l}\Omega_m} \left(\prod_{m=1}^{l} \mathrm{d}\vec{\mu}_{\{m\}}\, \Delta\kappa_{\{m\}} \frac{\delta}{\delta\kappa_{(0)}(\vec{\mu}_{\{m\}})} \right) Q[\kappa_{(0)}],} \tag{1.5}$$

where the subscript $\{m\}$ does not denote a component of a vector, but rather is shorthand for a distinct space, such that each $\mathrm{d}\vec{\mu}_{\{m\}}$ allows for multiple iterations of integration. In other words, the axioms of equivalence[17] hold for all spaces $\{\Omega_m\}$ with $\mu_{\{m\}} \in \Omega_m \,\forall\, \mu_{\{m\}}$ where $\vec{\mu}_{\{m\}} \mapsto \vec{\mu}_{\{n\}}$ is an identical transformation:

$$\boxed{\kappa(\vec{\mu}_{\{m\}}) \equiv \kappa(\vec{\mu}_{\{n\}}) \qquad \forall\, \vec{\mu}_{\{m\}} \equiv \vec{\mu}_{\{n\}}.} \tag{1.6}$$

The δ-operator therefore produces differentials which are *not infinitesimal*, and the gradient is not a true derivative. For completeness, the total change ΔQ is given by:

$$\Delta Q[\kappa_{(0)};\kappa] := Q[\kappa] - Q[\kappa_{(0)}] \equiv \sum_{l} \delta^l Q[\kappa_{(0)};\kappa],$$

whence the total infinitesimal is given by:

$$\mathrm{d}Q[\kappa_{(0)}] := \lim_{\kappa\to\kappa_{(0)}} \Delta Q[\kappa_{(0)};\kappa].$$

In this sense, $\Delta\kappa$ is simply the total change of κ under this representation, provided that one introduces the following shorthand:

$$Q_{(0)} := Q[\kappa] - \Delta Q[\kappa_{(0)};\kappa]. \tag{1.7}$$

To allow for an approximate calculus to exist, one requires in linearised models that:

$$\text{linearisation} : |\delta^2 Q|, |\delta Q|^2 \ll |\delta Q|. \tag{1.8}$$

This ordering means that one only keeps terms which are *linear* in δQ. In plasma systems, δQ is typically a time-dependent quantity. Therefore, for scenarios where δQ is not monotonic in time (i.e. oscillating electromagnetic fields), it is sometimes useful to instead consider ordering of the variation to the wave amplitude A, such that $|\delta^2 A| \ll |\delta A|$. For cases where δQ is affected by nonlinear processes such as frequency chirping, it sometimes becomes useful to look at temporal ordering instead, such that $(\delta t)^2 \ll (\delta t)$.

This final condition means that even for highly nonlinear plasma scenarios, if one uses very small timesteps, one could iteratively produce the nonlinear dynamics for the system by using a linearised eigensolver; this is an Eulerian decomposition of the problem in time. However, construction of this form of solver is not trivial.

1.4 Theoretical models

1.4.1 Classical action

The 4-position of the i^{th} particle in a plasma is given by:

$$X_i = \left(\; x_i(t_i), \quad y_i(t_i), \quad z_i(t_i), \quad t_i(t) \; \right),$$

where x_i, y_i, z_i denote the Cartesian 3-position of the particle, and t_i is the time coordinate for the particle. The set of 4-positions of all of the N_p particles in the system is given by $X = \{X_i\}_{i=1}^{N_p}$. The state of a plasma is given by the 4-position and 4-velocity of all of the N_p particles in the system, such that the state of the system is given by the set:

$$\mathcal{S}_C = \{X_i, \dot{X}_i\}_{i=1}^{N_p}.$$

One can define a functional J_c that maps $\mathcal{S}_C$ onto a real number:

$$J_C[\mathcal{S}_C] \; : \; \mathcal{S}_C \mapsto \mathbb{R}.$$

I refer to this functional as the *action*. One can formally represent the dynamics of the particle as being given by the trajectory which minimizes J_C. That is to say:

$$Y_i : \; \delta J_C[\mathcal{S}_C] = 0,$$

where $Y_i \in \mathbb{R}^3 \times \mathbb{R}[0,\infty]$ is the 4-trajectory for the particle. The action is given from classical mechanics with respect to the system Lagrangian L:

$$J_C[\mathcal{S}_C] := \int_{t_1}^{t_2} L(\mathcal{S}_C, t)\,\mathrm{d}t,$$

and as is given from elementary classical mechanics,[18] one arrives at the corresponding Euler-Lagrange equations:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial X_i}\right) - \left(\frac{\partial L}{\partial X_i}\right) = 0.$$

In plasma, there are few constraints to particle motion, and many degrees of freedom. Therefore, one expects that the Euler-Lagrange equations above are likely to be of similar complexity to simply solving the problem as a Newtonian many-body problem.

For the Newtonian problem, the forces here in the absence of external fields can be determined by the Lorentz force, Coulomb potential and the Biot-Savart law. However, solving this directly is incredibly laborious and largely unnecessary. However, for the case that there are external fields, one must produce self consistent equations which couple the particle motion and the evolution of the total electromagnetic field.

1.4.2 Field action

Still in the absence of external fields, it is possible for one to represent the self-consistent electromagnetic field via the Lorentz force and the Liénard-Weichert potentials. In this sense, one could construct a Lagrangian but it is simpler to move towards classical field theory. If one examines the 4-potential $\{A^\mu\}$:

$$\{A^\mu\} \,:\, (X, \dot{X}, X_0) \mapsto \mathbb{R}^3 \times \mathbb{R}[0,\infty],$$

where $X_0 = (\mathbf{x}, t)$ is the reference 4-position, and each element of the 4-potential is a scalar field. The 4-potential depends on the *relative* location between the observer at X_0 and all of the particles; accordingly $A^\mu = A^\mu(X, \dot{X}, X_0)$. Now, all the information is encoded in the positions of the particles, and the 4-potential. Therefore, our state is defined by the set:

$$\mathcal{S} = \{A^\mu(\mathcal{S}_C, X_0)\}_{\mu=1}^4 \cup \mathcal{S}_C.$$

Then, one can define a functional J that maps $\mathcal{S}$ onto a real number:

$$J_f[\mathcal{S}] \,:\, \mathcal{S} \mapsto \mathbb{R}.$$

I refer to this as the *field action*. This is given in the context of classical field theory by:[19]

$$J_f[\mathcal{S}] := \int_{\mathbb{R}^3} \int_{t_1}^{t_2} L_f(\mathcal{S}, t)\, \mathrm{d}t\, \mathrm{d}\mathbf{x},$$

where $X = X(\mathbf{x}, t)$, and where one integrates the Lagrangian density L_f over all space. Occasionally, one may find it appropriate to construct a model from as low-level a formalism as this; Berk *et al.* construct the base theory used for so-called Berk-Breizman models from a Lagrangian formalism.[8] Using a Lagrangian field theory or otherwise, one obtains the Lorentz-Maxwell dynamical system of equations:

$$\ddot{\mathbf{x}}^{(i)} = \frac{q^{(i)}}{m^{(i)}}(\mathbf{E}^K + \dot{\mathbf{x}}^{(i)} \times \mathbf{B}^K), \tag{1.9a}$$

$$\nabla \cdot \mathbf{E}^K = \frac{1}{\epsilon_0} \sum_{\text{particles}} \frac{q^{(i)}}{\int_{\mathbb{R}^3} \mathrm{d}\mathbf{x}}, \tag{1.9b}$$

$$\nabla \times \mathbf{E}^K = -\frac{\partial \mathbf{B}^K}{\partial t}, \tag{1.9c}$$

$$\nabla \cdot \mathbf{B}^K = 0, \tag{1.9d}$$

$$\nabla \times \mathbf{B}^K = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}^K}{\partial t} + \mu_0 \sum_{\text{particles}} q^{(i)} \dot{\mathbf{x}}^{(i)}, \tag{1.9e}$$

where $\mathbf{E}^K(X, \dot{X}, X_0)$ is the electric field, $\mathbf{B}^K(X, \dot{X}, X_0)$ is the magnetic flux density, $q^{(i)}$ is the charge of the i^{th} particle, $m^{(i)}$ is the mass of the i^{th} particle, ϵ_0 is the permittivity of free space and μ_0 is the permeability of free space. The equations form a system of $3(4 + N_p)$ differential equations to be solved. 4 vector equations (Maxwell's equations) are 1^{st} order partial differential equations (PDEs), while N_p vector equations (Lorentz force) are 2^{nd} order ordinary differential equations (ODEs). The entire system is nonlinear, and very challenging to solve computationally.

1.4.3 Klimontovich equation

Plasmas contain a very large amount of particles, and therefore it becomes favourable in many scenarios to use statistical physics. In 1967, Klimontovich showed that it is possible to represent the Lorentz-Maxwell system by replacing the particle positions and velocities with the particle phase space densities instead.[20] Each particle in the plasma occupies a

point in phase space, providing a coarse single PDF $f^K_{[l]}$ known as the Klimontovich density for a given species l:

$$f^K_{[l]}(\mathbf{x},\mathbf{v},t) := \sum_{m=1}^{N_{[l]}} \delta[\mathbf{x}-\mathbf{x}^{(m)}_{[l]}(t)]\,\delta[\mathbf{v}-\mathbf{v}^{(m)}_{[l]}(t)], \tag{1.10}$$

where one sums over $N_{[l]}$ particles for the species, $\mathbf{x}\in\mathbb{R}^3$ is the position, $\mathbf{v}\in\mathbb{R}^3$ is the velocity, $t\in\mathbb{R}[0,\infty]$ is the time, $\mathbf{x}^{(m)}_{[l]}(t)$ is the particle position, $\mathbf{v}^{(m)}_{[l]}(t)$ is the particle velocity, and $\delta[\cdot]$ is the Dirac delta function. As such, the the number density is given by:

$$n^K_{[l]} \equiv \int_{\mathbb{R}^3} f^K_{[l]}\,\mathrm{d}\mathbf{v} \equiv \sum_{m=1}^{N} \delta[\mathbf{x}-\mathbf{x}^{(m)}_{[l]}(t)].$$

Using the Klimontovich density allows one to combine the three 2^{nd} order ODEs given by Newton-Lorentz dynamics for each particle into a single 1^{st} order PDE for each species. Taking the total time derivative of f^K yields the Klimontovich equation:

$$\frac{\partial f^K_{[l]}}{\partial t} + \mathbf{v}\cdot\nabla f^K_{[l]} + \frac{q_{[l]}}{m_{[l]}}[\mathbf{E}^K+\mathbf{v}\times\mathbf{B}^K]\cdot\nabla_{\mathbf{v}} f^K_{[l]} = 0,$$

where one obtains an equation for each of the N_s species, $\mathbf{E}^K=\mathbf{E}^K(\mathbf{x},t)$, and $\mathbf{B}^K=\mathbf{B}^K(\mathbf{x},t)$. Therefore, the Klimontovich-Maxwell dynamical system of equations is as follows:

$$\ddot{\mathbf{x}}^{(i)}_{[l]} = \frac{q_{[l]}}{m_{[l]}}(\mathbf{E}^K + \dot{\mathbf{x}}^{(i)}_{[l]}\times\mathbf{B}^K), \tag{1.11a}$$

$$0 = \frac{\partial f^K_{[l]}}{\partial t} + \mathbf{v}\cdot\nabla f^K_{[l]} + \frac{q_{[l]}}{m_{[l]}}[\mathbf{E}^K+\mathbf{v}\times\mathbf{B}^K]\cdot\nabla_{\mathbf{v}} f^K_{[l]}, \tag{1.11b}$$

$$\nabla\cdot\mathbf{E}^K = \frac{1}{\epsilon_0}\sum_{\text{species}}\int_{\mathbb{R}^3} q_{[l]} f^K_{[l]}\,\mathrm{d}\mathbf{v}, \tag{1.11c}$$

$$\nabla\times\mathbf{E}^K = -\frac{\partial\mathbf{B}^K}{\partial t}, \tag{1.11d}$$

$$\nabla\cdot\mathbf{B}^K = 0, \tag{1.11e}$$

$$\nabla\times\mathbf{B}^K = \mu_0\epsilon_0\frac{\partial\mathbf{E}^K}{\partial t} + \mu_0\sum_{\text{species}}\int_{\mathbb{R}^3} q_{[l]}\mathbf{v} f^K_{[l]}\,\mathrm{d}\mathbf{v}. \tag{1.11f}$$

In practice the Klimontovich-Maxwell system is still difficult to solve; it is equivalent to the Lorentz-Maxwell system except for the parametrisation of the particle positions and velocities.

1.4.4 Boltzmann equation

Pre-dating the Klimontovich equation is the Boltzmann equation.[11] This equation requires one to examine a statistical ensemble of particles, and examines the full time derivative of the PDF, $f_{[l]}$:

$$\frac{\partial f_{[l]}}{\partial t} + \mathbf{v} \cdot \nabla f_{[l]} + \frac{\mathbf{F}_{\text{ext.},[l]}}{m} \cdot \nabla_{\mathbf{v}} f_{[l]} = \hat{C} f_{[l]}, \tag{1.12}$$

where $\mathbf{F}_{\text{ext.},[l]}$ is some external force, and $\hat{C}$ is an operator such that $\hat{C} f_{[l]}$ produces sources and sinks of the PDF. The Boltzmann equation is similar to the Klimontovich equation except for the fact that the Boltzmann equation uses a smooth function $f_{[l]}$ as opposed to $f^K_{[l]}$, and that the Boltzmann equation allows for sources and sinks. One can split the Klimontovich density (1.10) into an average $f_{[l]}$ and a fluctuating part $\delta f^K_{[l]}$:

$$f^K_{[l]}(\mathbf{x}, \mathbf{v}, t) \equiv f_{[l]}(\mathbf{x}, \mathbf{v}, t) + \delta f^K_{[l]}(\mathbf{x}, \mathbf{v}, t), \tag{1.13}$$

and similarly for the electric field and magnetic field:

$$\mathbf{E}^K_{[l]}(\mathbf{x}, t) \equiv \mathbf{E}_{[l]}(\mathbf{x}, t) + \delta \mathbf{E}^K_{[l]}(\mathbf{x}, t), \tag{1.14a}$$

$$\mathbf{B}^K_{[l]}(\mathbf{x}, t) \equiv \mathbf{B}_{[l]}(\mathbf{x}, t) + \delta \mathbf{B}^K_{[l]}(\mathbf{x}, t). \tag{1.14b}$$

Accordingly, one can identify that the sources and sinks in the Boltzmann equation may represented by the fluctuating terms in the Klimontovich equation. These small scale effects are important, and as such, defining 'small scale' is key. To use smooth distributions one must assume that *the dynamics of the system can be adequately described by ensemble averages.* To do this, one must divide the phase space into small grid cells which are not so small as to contain too few particles, but not too big as to allow for infinitesimal calculus to make poor approximations. Then:

$$f_{[l]}(\mathbf{x}, \mathbf{v}, t) \equiv \langle f^K_{[l]} \rangle := \frac{1}{\Delta\Gamma} \int_{\Delta\Gamma} f^K_{[l]} \, \mathrm{d}\Gamma, \tag{1.15}$$

where $\langle \cdot \rangle$ denotes ensemble averaging, and $\Delta\Gamma$ is a small closed 6D phase space volume defining a neighbourhood centred at $\mathbf{x}, \mathbf{v}$. Then, under ensemble averaging of the Klimontovich equation:

$$\underbrace{\frac{\partial f_{[l]}}{\partial t} + \mathbf{v} \cdot \nabla f_{[l]} + \frac{q_{[l]}}{m_{[l]}} [\mathbf{E} + \mathbf{v} \times \mathbf{B}] \cdot \nabla_{\mathbf{v}} f_{[l]}}_{\text{mean field}} = -\frac{q_{[l]}}{m_{[l]}} \overbrace{\langle [\delta \mathbf{E}^K + \mathbf{v} \times \delta \mathbf{B}^K] \cdot \nabla_{\mathbf{v}} \delta f^K_{[l]} \rangle}^{\text{fluctuations}}. \tag{1.16}$$

As such, small scale fluctuations in the electromagnetic field and Klimontovich density allow for the $\hat{C}$ operator to exist in the mean-field case, where $\{f_{[l]}\}$, $\mathbf{E}$ and $\mathbf{B}$ have been ensemble averaged. The external force in the Boltzmann equation by identification is therefore given as the mean-field Lorentz force, $q_{[l]}[\mathbf{E} + \mathbf{v} \times \mathbf{B}]$. The Boltzmann-Maxwell dynamical system of equations is therefore given in the rest of the thesis by:

$$\hat{C}f = \hat{V}f, \tag{1.17a}$$

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0} \sum_{\text{species}} q \int_{\mathbb{R}^3} f \, \mathrm{d}\mathbf{v}, \tag{1.17b}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{1.17c}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{1.17d}$$

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \sum_{\text{species}} q \int_{\mathbb{R}^3} \mathbf{v} f \, \mathrm{d}\mathbf{v}, \tag{1.17e}$$

where (1.17a) is the Boltzmann equation, (1.17b) to (1.17e) are Maxwell's equations, $\hat{V}$ is the 'Vlasov operator':

$$\hat{V} = \left[\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla + \frac{q}{m} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}} \right], \tag{1.18}$$

and other quantities are summarised as follows:

Quantity	**Definition**
q	particle charge for given species
m	particle mass for given species
ϵ_0	permittivity of free space
μ_0	permeability of free space
$\mathbf{x}$	position $\in \mathbb{R}^3$
$\mathbf{v}$	velocity $\in \mathbb{R}^3$
t	time $\in \mathbb{R}^{\geq 0}$
$f(\mathbf{x}, \mathbf{v}, t)$	PDF for given species
$\mathbf{E}(\mathbf{x}, \mathbf{v}, t)$	*mean-field* electric field
$\mathbf{B}(\mathbf{x}, \mathbf{v}, t)$	*mean-field* magnetic flux density
$\hat{C}$	phase space sources and sinks

where I use PDF to mean the ensemble averaged PDF for the remainder of the thesis. In Appendix C.1, I briefly outline how variants of the Boltzmann-Maxwell system can be derived from classical field theory. The $\hat{C}$ operator includes both collisions as well as other

phase space sources and sinks. These sources and sinks can arise from processes such as particle flux across the system boundary (for open systems), particle recombination, and particle excitation. For the scenarios examined in this thesis, particle number for each species can be considered to be conserved in all models. For this reason, $\hat{C}$ is simply the collision operator.

I have dropped the species superscript, noting that the collision operator will encapsulate interspecies collisions which may require reintroducing the superscript at a later point. I will herein refer to $\mathbf{E}$ and $\mathbf{B}$ as the *electric field* and *magnetic field* respectively. The $(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \nabla_{\mathbf{v}}$ leads to nonlinear coupling between the electromagnetic field and the PDF. As a result of this nonlinearity, it is difficult to solve systems with kinetic plasmas analytically.

From the Boltzmann equation, one can further reduce complexity by taking moments in velocity. Whichever moment of the Boltzmann equation is taken, the resultant equation features terms with the moment of order one greater. This leads to a problem with closure of the system. For this reason, so called 'fluid' models feature two key assumptions: all behaviour in velocity space is homogenised; the given closure for the system is valid. Different forms of closure lead to ideal magnetohydrodynamics (MHD), resistive MHD, and other models. This thesis is restricted to kinetic description of plasma, and therefore discussion of MHD is omitted.

1.5 Equilibrium distribution functions

There are infinitely many equilibrium PDFs that are solutions of the Boltzmann-Maxwell system, all of which form a subset of the general set of solutions. Here, I examine two equilibrium PDFs of physical significance which are later used in this thesis: the bump-on-tail (BOT) distribution, and the slowing-down distribution.

1.5.1 Bump-on-tail distribution

In this thesis, the 1D BOT distribution has the following form:

$$F_{\mathrm{BOT}}(\mathbf{x}, \mathbf{v}, t) = \frac{1}{\sqrt{2\pi}} \left\{ \eta \exp\left[-\frac{v^2}{2}\right] + \frac{1-\eta}{v_T} \exp\left[-\frac{1}{2}\left(\frac{v - v_B}{v_T}\right)^2\right] \right\}, \tag{1.19}$$

where v_T is the beam thermal velocity, v_B is the beam velocity, η is the fraction of particles in the bulk, velocity has been normalised to the bulk thermal velocity, and the number of particles has been normalised to 1. In the absence of electromagnetic fields and

wave-particle resonance, one expects that a Maxwellian represents the thermal equilibrium of the system. Therefore, a BOT distribution represents an equilibrium where two processes occur: particles of the given species are injected into the system with velocity v_B at the same rate that they thermalise with the bulk, and particles are removed from the bulk at that same rate.

However, this system is often kinetically unstable as the part of the PDF in the region $0 < v < v_B$ with positive slope allows for linear Landau instability.

1.5.2 Slowing-down distribution

In this thesis, the slowing-down distribution has the following form:[21,22]

$$F_{S,[i]}(v) = \frac{3}{4\pi \ln(1 + v_{\mathrm{c}}^3/v_{\mathrm{inj.}}^3)} \frac{n_{[i],(0)}}{v_{\mathrm{c}}^3 + v^3} H(v_{\mathrm{inj.}} - v), \tag{1.20}$$

where $F_{S,[i]}$ is the slowing-down distribution for the i^{th} fast ion species, H is the Heaviside step function, $\bar{n}_{[i],(0)}$ is the equilibrium number density of the i^{th} fast ion species, and v_{c} is the so-called critical speed:

$$v_{\mathrm{c}} \equiv v_{\mathrm{te}} \left(\frac{3\sqrt{\pi}}{4} \sum_i \frac{n_{[i],(0)} m_{\mathrm{e}}}{n_{\mathrm{e},(0)} m_{[i]}} Z_{[i]}^2 \right)^{1/3}, \tag{1.21}$$

where v_{te} is the electron thermal speed, $n_{\mathrm{e},(0)}$ is the equilibrium electron number density, $m_{[i]}$ is the mass of the i^{th} fast ion species, m_{e} is the electron rest mass, and $Z_{[i]}$ is the atomic number of the i^{th} fast ion species. In tokamaks, neutral beam injection (NBI) and radio frequency (RF) heating can lead the ions to exhibit a quasi-steady slowing-down distribution. Here, quasi-steady refers to the approximation that the distribution evolves smoothly as a function of $P_{\mathrm{inj.}}$; fundamentally, this assumes that the rate of increase of injected power is much less than the reciprocal of the Spitzer slowing down time.

1.6 Action-angle variables

Action-angle variables are a useful tool that allow one to deduce the form of the Hamiltonian in systems where the Hamiltonian is slowly evolving, commonly utilised in tokamak physics.[23,24] To illustrate this, I will examine the following canonical transformation of the Hamiltonian which is enabled by a type-2 generating function $Q_2(\mathbf{q}, \mathbf{J}, t)$:[18]

$$w_k = \frac{\partial Q_2}{\partial J_k}; \qquad p_k = \frac{\partial Q_2}{\partial q_k}; \qquad K(\mathbf{w}, \mathbf{J}, t) = H(\mathbf{q}, \mathbf{p}, t) + \frac{\partial Q_2}{\partial t},$$

where $\mathbf{J} = \{J_k\}$ are the action variables in the system, and $\mathbf{w} = \{w_k\}$ are the angle variables. The new Hamilton's equations are given by:

$$\frac{\mathrm{d}}{\mathrm{d}t} g(\mathbf{w}, \mathbf{J}, t) = \{g, K\} + \frac{\partial g}{\partial t},$$

where $\{g, K\}$ denotes the Poisson bracket in the new canonical phase space:

$$\{g, K\} := \frac{\partial g}{\partial w_k}\frac{\partial K}{\partial J_k} - \frac{\partial K}{\partial w_k}\frac{\partial g}{\partial J_k}.$$

I define the following new generalized momenta $\{J_k\}$:

$$J_k := \delta J_k(\mathbf{q}, t) + \oint_{H(q_k, p_k, t) = \text{const.}} p_k \, \mathrm{d}q_k,$$

where the integral is performed over a closed phase space orbit at time t (such that it is the *unperturbed orbit*), and the new generalized momenta are constants of motion. The last term on the right hand side represents an adiabatic invariant, while the first term on the right hand side represents deviation from adiabaticity. One can further demand that the new Hamiltonian is not explicitly a function of time. To achieve this, one must use a generating function such that:

$$\frac{\partial}{\partial t}\left(\frac{\partial Q_2}{\partial t} + H\right) = 0. \tag{1.22}$$

Then, from the new Hamilton's equations:

$$0 = -\frac{\partial K}{\partial w_k}; \qquad\qquad w_k = \int_0^t \frac{\partial}{\partial J_k} K(\mathbf{J}) \, \mathrm{d}t + \text{const.},$$

where one should note that K cannot be a function of $\mathbf{w}$. The canonical transformation is therefore such that the new Hamiltonian is solely a function of the constants of motion. Each one is representable by a quantity which resembles classical action, plus a perturbation. If one integrates over a full period $\tau_{(0)}(\mathbf{J})$ of the orbit at time t:

$$\Delta w_k = \frac{\partial}{\partial J_k} K(\mathbf{J}) \tau_{(0)}(\mathbf{J}, t).$$

But one can also represent the variation of the angle variable w_k by:

$$\Delta w_k = \delta w_k(\mathbf{J}, t) + \oint_{H(q_k, p_k, t) = \text{const.}} \frac{\partial w_k}{\partial q_k} \, \mathrm{d}q_k,$$

such that Δw_k is constructed from the unperturbed orbit contribution, plus the variation arising from the orbit changing as a function of time. Then:

$$\Delta w_k = 1 + \delta w_k,$$

and accordingly:

$$\frac{\partial}{\partial J_k} K(\mathbf{J}) = \frac{1 + \delta w_k(\mathbf{J}, t)}{\tau_{(0)}(\mathbf{J}, t)},$$

such that the left hand side is equal to the bounce frequency plus a contribution arising from temporal perturbations to the particle orbits. The angle variables are therefore given by:

$$w_k = \frac{1 + \delta w_k(\mathbf{J}, t)}{\tau_{(0)}(\mathbf{J}, t)} t + \text{const.}$$

Altogether, this is simply a representation of Noether's theorem[1]. The continuous symmetries here are that K is invariant under a temporal transformation or a translation in w_k. The former leads to conservation of energy, and the latter manifests with J_k as constants of motion. One finds that the Hamiltonian can be represented in the form:

$$H(\mathbf{q}, \mathbf{p}, t) = K(\mathbf{J}) + \delta H(\mathbf{q}, \mathbf{J}, t),$$

such that $\delta H := -\partial Q_2/\partial t$ represents a explicitly time-varying perturbation to the Hamiltonian. Later, in Section 3.2.1, I examine similar scenarios to this by considering how the period of an orbit changes as the kinetic energy of a particle changes in time. Here, I formally define adiabatic invariants as:

$$\boxed{\text{adiabatic} : \frac{\partial}{\partial t} \ln \left(J_k - \delta J_k \right) \ll \frac{\partial}{\partial t} \ln H,} \tag{1.23}$$

such that the logarithmic rate of change of $(J_k - \delta J_k)$ is small compared to the Hamiltonian. Under such an approximation, the orbits are approximately temporally static. In such a case, the Hamiltonian is approximately time-independent. Then, by virtue of (1.22), $\partial Q_2/\partial t$ must be time-invariant as well. The Hamiltonian then would take the approximate form:

$$H(\mathbf{q}, \mathbf{p}, t) \approx H_0(\mathbf{J}) + \text{const.}$$

allowing for H to be representable solely as a function of the constants of motion. For example, in tokamaks the equilbrium Hamiltonian (which can be identified as $K(\mathbf{J})$) is commonly taken to be the form:

$$H_0 = H_0(W, p_\varphi, \mu),$$

where W is the particle energy, p_φ is the toroidal angular momentum, and μ is the particle magnetic moment. The corresponding conjugate quantities for which continuous symmetries approximately exist are time, azimuthal angle, and the gyroangle.

1.7 Thesis outline

In this chapter, I have given a brief outline of the underlying theories one can use to represent kinetic plasmas.

In Chapter 2, I cover 'conservative orbit' theories, where the single particle energy is conserved on a particle orbit. I take a small detour via linearised kinetic plasma theory, obtaining a closed equation for the linear complex dispersion relation for general kinetic plasmas, with a later simplification to electrostatic plasmas. I close the chapter by defining theories which are safely approximated by conservative orbits: the approximations at hand are defined directly with respect to wave dynamics rather than adiabatic invariants.

In Chapter 3, I cover 'non-conservative orbit' theories, where the single particle energy is not conserved on a particle orbit. I discuss how such theories can allow for a fuller analytical description of frequency chirping, mode growth and wave-particle resonance than 'conservative orbit' theories. Furthermore, these theories allow for further explanation of the origin of phase space islands in a nonlinear plasma. I detail two approaches: a nonlinear basis expansion, and a generating function formulation. Both theories allow for solutions which extend beyond that which is given in the literature by 'near-conservative orbit' theories (BGK modes, H&C theory). Work from this chapter has been previously published.[25]

In Chapter 4, I detail the workings of the D-dimensional Augmented Resonance Kinetic solver (DARK) code written by myself over the course of 2015 - 2018 based on previous work by Arber, Vann and de Gol.[26,27] DARK is a sophisticated, new, modular code that was used to run the simulations shown in Chapter 5.

In Chapter 5, I cover computational simulations of 1D electrostatic kinetic plasmas. The chapter makes use of the Berk-Breizman augmentation commonly examined in the

literature,[todo2018introduction, 8,26,28–30] extending previous work in the literature by including stochastic electric fields and PDFs. I show that a stochastic subpopulation of particles can prevent phase space island formation and diminish frequency sweeping, and that the lifetime of H&Cs is stochastic, implying that the structures are not fully coherent and deterministic. I later show that one can predict behaviour surrounding hole-and-clump destabilisation to high accuracy, where a hole or clump destabilises nearby modes by locally deforming the PDF and increasing the Landau linear growth rate. These systems share heuristic similarities to Alfvén waves, and consequent predictions are further discussed in Chapter 7. This work was carried out with collaborators from Princeton Plasma Physics Laboratory (PPPL), and has been previously published.[31]

In Chapter 6, I cover ML methods employed to analyse datasets related to nuclear fusion. Here, I examine spectrograms of magnetic field perturbations, obtaining similar data to that observed in computational simulations in Chapter 5. I use the new ML framework Experimental Resonant Instability Correlation Studies on NSTX (ERICSON) to characterise the frequency behaviour of modes, allowing one to investigate subsequent correlations using traditional analysis. This work was carried out with collaborators from PPPL, and has been previously published.[32]

Finally in Chapter 7, I give more detailed conclusions drawn from the chapters, before going on to outline possible future work that could arise from the work.

Chapter 2

Conservative orbits and linear kinetic theory

2.1 Overview

Kinetic plasmas can be described by a Boltzmann equation[10] that describes the evolution of the particle distribution function (PDF), f, and Maxwell's equations which describe the evolution of the mean-field electromagnetic field. The force exerted on each of the particles in the system manifests as a term acting as $\mathbf{F}_{\text{ext.}} \cdot \nabla_{\mathbf{v}}$ on the PDF, where $\nabla_{\mathbf{v}}$ denotes a velocity space gradient, and $\mathbf{F}_{\text{ext.}}$ is typically taken to be the Lorentz force. The method of linearisation allows one to derive damping or inverse damping of electrostatic plasma waves in the system under the assumption that the nonlinear coupling is sufficiently small.[33] However, this coupling leads to wave-wave instabilities which affect the nonlinear stability of the system. This can lead to destabilisation[34,35] as well as stabilisation.[36,37] These nonlinear scenarios have been partially investigated analytically, however they are typically examined using computational means.[29,38,39]

I begin in Section 2.2 by discussing conservative orbits, and how a large amount of the existing theory of phase space structures is described using conservative orbit theories, or small variants thereof. In Section 2.2.1, I discuss previous work by Bernstein, Greene and Kruskal. In Section 2.2.2, I present my own work, rigorously proving certain properties of so-called Bernstein-Green-Kruskal (BGK) modes, and related phase space structures.

Then in Section 2.3, I explore a linearised theory of kinetic plasmas, developing from typical treatments in the literature[10,11,33] to produce extended solutions for collisional electrostatic plasmas (see (2.17) in Section 2.3.1, (2.23) in Section 2.3.5, and all of Section 2.3.6).

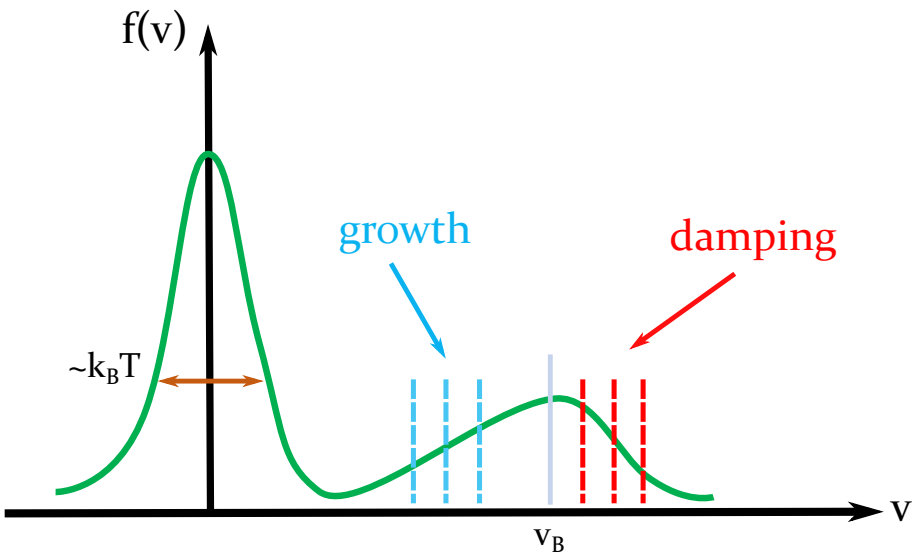


Figure 2.1: **Landau damping and inverse Landau damping**. Sketch of the spatially averaged PDF $f_0(v)$ for a 1D bump-on-tail (BOT) distribution with bulk temperature T, and a beam travelling at velocity v_B with respect to the bulk rest frame. Linear wave-particle resonance occurs at $v \approx \omega/k$, where ω/k is the phase velocity of the wave. Electrostatic waves with phase velocity coinciding with negative $\mathbf{v} \cdot \nabla_{\mathbf{v}} f_0$ give energy to the PDF and undergo damping (Landau damping). Conversely, electrostatic waves with phase velocity coinciding with positive $\mathbf{v} \cdot \nabla_{\mathbf{v}} f_0$ extract energy from the PDF and grow (inverse Landau damping).

2.2 Conservative orbits

Suppose that $f = f[\epsilon_{(0)}]$, such that $(x, v, t) \mapsto \epsilon_{(0)}$, and $\epsilon_{(0)} \mapsto \mathbb{R}$. Via Liouville's theorem[18], one expects that phase space is incompressible.[18] Therefore, particle continuity is required along contours of constant f. These contours exist in locations of constant $\epsilon_{(0)}$, and therefore these contours correspond to particle orbits. If $\epsilon_{(0)}$ has the form:

$$\epsilon_{(0)} = q\phi + \frac{1}{2}m(v-u)^2, \tag{2.1}$$

where $\phi(x,t)$ is the electric potential and u is a constant velocity, the orbit can be said to be *conservative* as the particle energy is conserved along the orbit.

A seminal 1957 work by Bernstein, Greene and Kruskal gives a family of exact nonlinear solutions to the 1D electrostatic Vlasov equation (BGK waves), where the particles take conservative orbits.[40] This theory successfully gives exact solutions for the PDF and electric potential provided that the wave is of constant amplitude and has a group velocity equal to the phase velocity. As only one frequency exists in the system, any wave-wave coupling

must be negligible in the system; the solution is nonlinear, but inherently assumes that waves only couple to affect the spatially averaged PDF.

This assumption is fairly useful for Alfvén waves in tokamak plasmas. Gap toroidal Alfvén eigenmodes (TAEs) exist in a very narrow bandwidth ($\Delta\omega \ll \omega$), and therefore wave-wave coupling typically excites modes in the Alfvén continuum or excite other harmonics of gap Alfvén wave. Continuum modes undergo continuum damping, which is relatively larger than other damping rates in the system ($|\gamma_{\text{cont.}}| > |\gamma_D|$). As a result, continuum modes can largely be neglected. Furthermore, the probability of exciting another harmonic of gap Alfvén wave is small as the bandwidth for the other harmonics is also small; the characteristic time for energy diffusion in k-space can therefore be thought to be very large with respect to other system timescales.

In Section 2.2.1, I recap the BGK theory, and pose some expected limitations of the model. In Section 2.2.2, I explicitly prove these aforementioned limitations.

2.2.1 BGK waves

BGK waves are plasma waves which are exact nonlinear solutions of the 1D electrostatic collisionless Vlasov-Maxwell system:

$$\frac{\partial f}{\partial t} = -v\frac{\partial f}{\partial x} + \frac{q}{m}\frac{\partial \phi}{\partial x}\frac{\partial f}{\partial v}, \tag{2.2a}$$

$$\nabla^2\phi = -\frac{1}{\epsilon_0}\sum_{\text{species}} q\int_{\mathbb{R}} f\,\mathrm{d}v. \tag{2.2b}$$

If $f = f[\epsilon_{(0)}]$, then $\epsilon_{(0)}$ must also satisfy the Vlasov equation:

$$\left[\frac{\partial}{\partial t} + v\frac{\partial}{\partial x} - \frac{q}{m}\frac{\partial \phi}{\partial x}\frac{\partial}{\partial v}\right]\epsilon_{(0)} = 0,$$

A solution of the above equation is given by $\epsilon_{(0)} = q\phi + \frac{1}{2}m(v-u)^2$. Upon insertion into the Vlasov equation, we find a differential constraint for $\phi(x,t)$.

$$\frac{\partial \phi}{\partial t} + u\frac{\partial \phi}{\partial x} = 0.$$

This advection equation requires the electric potential to take the following form:

$$\phi(x,t) = \frac{1}{2}\sum_j \left\{\phi_j \mathrm{e}^{\mathrm{i}k_j(x-ut)} + \text{c.c.}\right\}.$$

Accordingly, BGK theory finds that particles move along conservative orbits where the energy of *each and every particle* is conserved. This result is particularly remarkable,

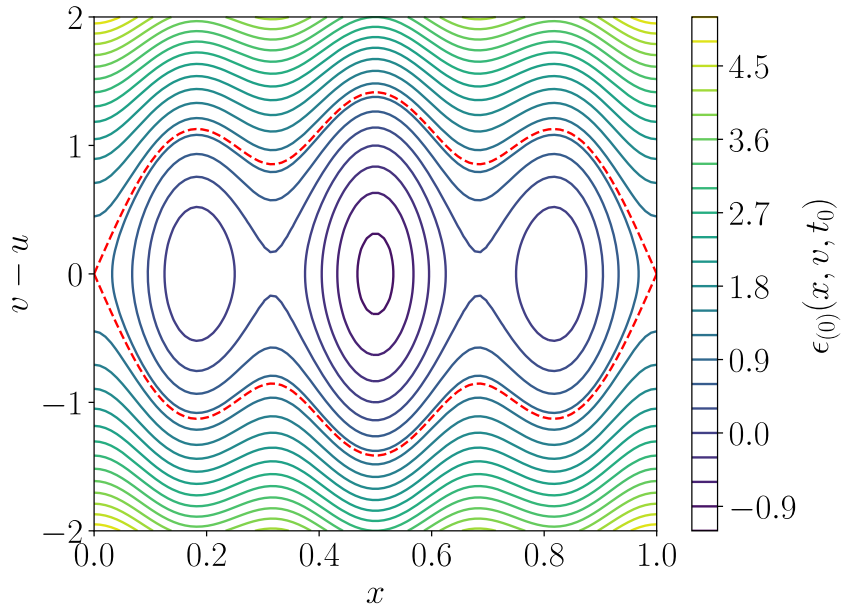


Figure 2.2: **Phase space contours for BGK modes.** Shown is a set of BGK islands with $q\phi \propto [\cos(kx - \omega t) + \cos(3kx - 3\omega t)]$, to illustrate BGK modes that can exist with complicated electric potentials. The first harmonic of the wavevector k corresponds to the fundamental wavenumber of the system ($k = 2\pi/L$). Note that the islands all are co-located at $v = u$, where the phase velocity $u \equiv \omega/k$. Energy is normalised to the maximum potential of the wave ($q\phi_{\text{max.}}$), and velocity is normalised with respect to $\sqrt{2q\phi_{\text{max.}}/m}$.

as one would only expect that the global energy of the system is conserved. Bernstein, Greene and Kruskal showed that the stronger case with the conservation of single particle energy was in fact an available solution. The resultant structure, as one might expect, is highly coherent. However, I will later show in Section 3.2.1 that deviation in the periods of single-particle orbits allows for phase mixing and wave-particle energy transfer.

BGK waves are highly ubiquitous in plasma physics, and a large number of phenomena can be explained using BGK theory or a small extension of it.[7,9,41]

However, the theory does fail under some circumstances:

1. *It cannot by itself recover solutions with time-varying frequency and growth rate.*

   BGK theory cannot give solutions with time-varying frequency and growth rate. However, provided that the change to the shape of the phase space island is small at a given sampling rate, one can approximate behaviour by assuming a temporal 'continuum of BGK modes' to exist in the system as shown by Berk *et al.*[5]

   This approach has led to a good deal of success, and is extensively used in hole and

clump (H&C) theory, allowing one to recover an approximate value for the chirping rate of modes during nonlinear saturation. However, it is only valid for adiabatically chirping modes.

2. *It cannot recover solutions where the topology of the phase space island changes.*

   In some scenarios, a phase space island may bifurcate, or sidelobes may appear.[42] These solutions cannot be explained by the BGK theory which only permits the existence of one island.

3. *It cannot recover stochastic solutions.*

   The BGK island is a highly coherent structure. As a result, any system with decoherence, such as that produced from stochastic effects and turbulence, will not be adequately described by the theory.

Wave packets feature constituent waves of differing phase velocity. The group velocity ($\nabla_k \omega$) is dependent on the linear complex dispersion relation for the whole wave, and the temporal envelope of the wave packet is dependent on the group velocity.

2.2.2 BGK modes as dispersionless waves

When considering a wave packet, one must reconsider BGK analysis. First, one can examine a superposition of energies:

Lemma 2.1. *For the energy given by:*

$$\epsilon_{(0)}^{[l]} = q\phi^{[l]}(x,t) + \frac{1}{2}m(v-u^{[l]})^2.$$

Any linear superposition over l of these energies only has one global minimum in velocity.

Proof. If one sums over N energies:

$$\sum_l \epsilon_{(0)}^{[l]} = \sum_l q\phi^{[l]} + \frac{1}{2}m\sum_l (v-u^{[l]})^2.$$

The rightmost sum is quadratic in v, and therefore should only produce one global minimum in velocity. To check, one seeks $(\partial/\partial v)\sum_l \epsilon_{(0)}^{[l]} = 0$:

$$0 = m\sum_l (v-u^{[l]}) = Nmv - m\sum_l u^{[l]}.$$

Therefore, one finds:

$$v_{\text{min.}} = \frac{\sum_l u^{[l]}}{N} \equiv \bar{u}.$$

□

This lemma allows one to examine two important cases, where $\epsilon^{[l]}_{(0)}$ is some form of energy in the l^{th} wave frame:

- **Single wave packet**: $\phi(x,t) = \phi^{[l]}(x,t)$

  The energy $\epsilon^{[l]}_{(0)}$ is as though particles are interacting with the full electric potential; $\epsilon^{[l]}_{(0)}$ is the particle energy in the l^{th} wave frame.

  This case only allows for a single phase velocity to exist. From substitution into the Vlasov equation:

  $$N\left[u^{[l]}\frac{\partial\phi}{\partial x} + \frac{\partial\phi}{\partial t}\right] = 0.$$

  This is only satisfied if ϕ consists of a single phase velocity. Therefore, $\bar{u} \equiv u^{[l]}\,\forall\, l$.

- **Multiple wave packets**: $\phi(x,t) = \sum_l \phi^{[l]}(x,t)$

  The energy $\epsilon^{[l]}_{(0)}$ is as though particles are interacting with only part of the electric potential; $\epsilon^{[l]}_{(0)}$ is not the particle energy in the l^{th} wave frame.

  This case would allow for multiple phase velocities to exist, but is expressly forbidden. From substitution into the Vlasov equation:

  $$\sum_l \left[u^{[l]}\frac{\partial\phi^{[l]}}{\partial x} + \frac{\partial\phi^{[l]}}{\partial t}\right] = \sum_l \sum_{l'\neq l} \frac{\partial\phi^{[l']}}{\partial x}(v - u^{[l]}).$$

  By comparing coefficients of v, the only allowed solution is where $\partial\phi^{[l']}/\partial x = 0\,\forall\, l' \neq l$. Therefore, this solution simply reduces to the 'full potential' solution, with only a single phase velocity.

At this point it becomes useful to prove the following:

Theorem 1. *Let $x \in X$, $v \in V$ with $X \subset \mathbb{R}^d$, and $V \subset \mathbb{R}^d$, such that $X \times V$ is a complete vector space. Let a phase space volume $\Gamma \subseteq X \times V$ be a closed volume, such that it includes its boundary $\partial\Gamma$. Let island separatrices be defined as the largest closed surfaces $\{\partial\Gamma_s\}$ such that $\forall\partial\Gamma_s : f(x,v,t_0) = const.$ at some time t_0. Then, it is impossible to represent a set of island separatrices existing on more than one finite subset of V if $f(x,v,t_0)$ is not at least cubic in v.*

Proof. For a function $f(x, v, t_0)$ with $x \in X$ and $v \in V$, in a finite subset of X there exists a finite subset of V where the value of f is constant. These subsets form either open or closed contours on the manifold.

If one supposes that x_0, t_0 is such that some points in V lie on an island separatrix (which in turn is a closed contour), then:

$$\forall v \in \{v_1, v_2, \dots\} : f(x_0, v, t_0) = \text{const.},$$

as shown in Figure 2.3. One can write this in the alternative form:

$$\forall v \in \{v_1, v_2, \dots\} : F(v' = v) = C(x = x_0, t = t_0),$$

where $F : v \mapsto \mathbb{R}$ and $C : (x, t) \mapsto \mathbb{R}$ such that we seek solutions for v where some function of v is equal to a value dependent on x_0 and t_0. F and C are surjective functions. The number of possible values that solves this equation is given by the number of roots.

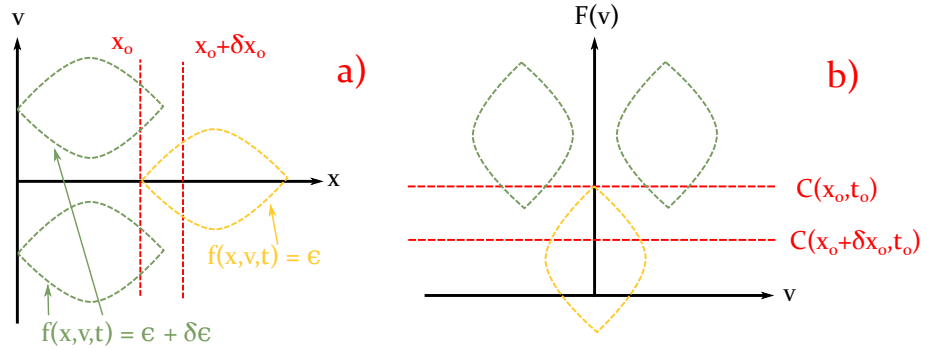


Figure 2.3: **Separatrices**. 1D sketches illustrating Theorem 1. **a:)** Sketches of separatrices of phase space islands with values of $f(x, v, t)$ given as ϵ for the gold island, and $\epsilon + \delta\epsilon$ for the green islands. The line $x = x_0$ intersects all three islands, and $x = x_0 + \delta x_0$ intersects only one island. **b:)** Only the gold island satisfies the equation $F(v) = C(x_0 + \delta x_0, t_0)$, while both the gold and green islands satisfy the equation $F(v) = C(x_0, t_0)$. As such, for separatrices to exist at the same point in x, the PDF must be locally at least cubic in v.

Under a variation from x_0 to $x_0 + \delta x_0$, one finds that $\{v_1, \dots\} \mapsto \{v_1 + \delta v_1, \dots\}$. Contours require a set of continuous points, and therefore one requires that there exists a root v_0 which bifurcates under the variation from x_0 to $x_0 + \delta x_0$, as shown in Figure 2.3. This variation simply changes the value of C, and so the topological requirements are that for a given island separatrix:

$$\exists\, v_0, \delta x_0 \; : \; v_1 = v_0 + \delta v_1; \quad v_2 = v_0 + \delta v_2.$$

Therefore, within the domain of C, $F(v)$ satisfies this in the vicinity of a local extremum. As such one can formally write that for $C \in [C_{\min.}, C_{\max.}]$:

$$F(v) \in [C_{\min.}, C_{\max.}] \, : \, \partial_v F = 0, \partial_v^2 F \neq 0.$$

Therefore, for islands to exist on more than one finite subset in V, there must be at least 2 extrema in v. Therefore, $f(x, v, t_0)$ must be at minimum cubic to represent a set of island separatrices existing on more than one finite subset of V. □

One can therefore extend Lemma 2.1 to show the following:

Lemma 2.2. *For the energy given by:*

$$\epsilon_{(0)}^{[l]} = q\phi^{[l]}(x,t) + \frac{1}{2}m(v - u^{[l]})^2,$$

where $\phi(x,t)$ is given by a sum over N phase velocities:

$$\phi(x,t) = \sum_l \phi^{[l]}(x,t).$$

A collisionless 1D electrostatic plasma does not permit a solution of the form $f[\epsilon_{(0)}]$ where:

$$\epsilon_{(0)} = \sum_l \epsilon_{(0)}^{[l]},$$

unless there is only a single wave packet in the system such that $\epsilon_{(0)} = \epsilon_{(0)}^{[l]}$. Such a solution only features a chain of phase space islands in x.

Proof. From Lemma 2.1, $\epsilon_{(0)}$ has only one minimum in velocity at $v = \bar{u}$, where $\bar{u} = \sum_l u^{[l]}/N$. Therefore, $\epsilon_{(0)}$ can be explicitly written in the form:

$$\epsilon_{(0)} = \sum_l q\phi^{[l]} + \frac{1}{2}m\sum_l (v - u^{[l]})^2 = \sum_l q\phi^{[l]} + c_1(v - \bar{u})^2 + c_2,$$

where c_1 and c_2 are constants. By inspection:

$$c_1 = \frac{1}{2}\sum_l m \quad ; \quad c_2 = \frac{1}{2}m\sum_l (\bar{u} - u^{[l]})^2.$$

Therefore:

$$\epsilon_{(0)} = \sum_l \left[q\phi^{[l]} + \frac{1}{2}m(v - \bar{u})^2 + \frac{1}{2}m(\bar{u} - u^{[l]})^2 \right].$$

In this form, one can explicitly see that $\epsilon_{(0)}$ is equal to:

$$\epsilon_{(0)} = q\phi' + \frac{1}{2}Nm(v-\bar{u})^2,$$

where ϕ' is a gauge shifted potential:

$$q\phi' \equiv \sum_l q\phi^{[l]} + \frac{1}{2}m\sum_l(\bar{u}-u^{[l]})^2.$$

If one substitutes $\epsilon_{(0)}$ into the Vlasov equation:

$$\frac{\partial\phi}{\partial t} + [v - N(v-\bar{u})]\frac{\partial\phi}{\partial x} = 0.$$

This is only satisfied for all v if $N = 1$; therefore there is only a single wave packet in the system with phase velocity $\bar{u}$. Via Theorem 1, as $\epsilon_{(0)}$ is quadratic in v, as contours of constant f are also contours of constant $\epsilon_{(0)}$, the solution only features a chain of phase space islands in x. □

If BGK modes can be superposed, then it may be possible to represent two phase space structures by superposing solutions. As such, it is useful to show the following:

Lemma 2.3. *A superposition of BGK islands of the form:*

$$f = f^{[1]}[\epsilon^{[1]}_{(0)}] + f^{[2]}[\epsilon^{[2]}_{(0)}],$$

where $\epsilon^{[1]}_{(0)} \neq \epsilon^{[2]}_{(0)}$*, and* $\epsilon^{[l]}_{(0)}$ *is given by:*

$$\epsilon^{[l]}_{(0)} = q\phi^{[l]}(x,t) + \frac{1}{2}m(v-u^{[l]})^2,$$

with $\phi^{[l]}$ *satisfying:*

$$u^{[l]}\frac{\partial\phi^{[l]}}{\partial x} + \frac{\partial\phi^{[l]}}{\partial t} = 0,$$

and $\phi(x,t) = \phi^{[1]} + \phi^{[2]}$ *only solves the Vlasov equation if the phase space islands exist on the same subset of* v*, such that* $u^{[1]} \equiv u^{[2]}$*.*

Proof. By substituting f into the Vlasov equation:

$$\left[\frac{\partial\epsilon^{[1]}_{(0)}}{\partial t} + v\frac{\partial\epsilon^{[1]}_{(0)}}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial\epsilon^{[1]}_{(0)}}{\partial v}\right]\frac{\mathrm{d}f^{[1]}}{\mathrm{d}\epsilon^{[1]}_{(0)}} = -\left[\frac{\partial\epsilon^{[2]}_{(0)}}{\partial t} + v\frac{\partial\epsilon^{[2]}_{(0)}}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial\epsilon^{[2]}_{(0)}}{\partial v}\right]\frac{\mathrm{d}f^{[2]}}{\mathrm{d}\epsilon^{[2]}_{(0)}}.$$

By substituting $\epsilon^{[l]}_{(0)}$:

$$\left(\frac{\partial\phi^{[1]}}{\partial t}+u^{[1]}\frac{\partial\phi^{[1]}}{\partial x}-\frac{\partial\phi^{[2]}}{\partial x}(v-u^{[1]})\right)\frac{\mathrm{d}f^{[1]}}{\mathrm{d}\epsilon^{[1]}_{(0)}}=-\left(\frac{\partial\phi^{[2]}}{\partial t}+u^{[2]}\frac{\partial\phi^{[2]}}{\partial x}-\frac{\partial\phi^{[1]}}{\partial x}(v-u^{[2]})\right)\frac{\mathrm{d}f^{[2]}}{\mathrm{d}\epsilon^{[2]}_{(0)}}.$$

One can evaluate this for $v\in\{u^{[1]},u^{[2]}\}$. First, at $u^{[1]}$:

$$\left(\frac{\partial\phi^{[1]}}{\partial t}+u^{[1]}\frac{\partial\phi^{[1]}}{\partial x}\right)\frac{\mathrm{d}f^{[1]}}{\mathrm{d}\epsilon^{[1]}_{(0)}}=-\left(\frac{\partial\phi^{[2]}}{\partial t}+u^{[1]}\frac{\partial\phi^{[2]}}{\partial x}\right)\frac{\mathrm{d}f^{[2]}}{\mathrm{d}\epsilon^{[2]}_{(0)}},$$

and then at $u^{[2]}$:

$$\left(\frac{\partial\phi^{[1]}}{\partial t}+u^{[2]}\frac{\partial\phi^{[1]}}{\partial x}\right)\frac{\mathrm{d}f^{[1]}}{\mathrm{d}\epsilon^{[1]}_{(0)}}=-\left(\frac{\partial\phi^{[2]}}{\partial t}+u^{[2]}\frac{\partial\phi^{[2]}}{\partial x}\right)\frac{\mathrm{d}f^{[2]}}{\mathrm{d}\epsilon^{[2]}_{(0)}}.$$

It is trivial to show that this system of equations only yields non-unique solutions in (x,t) if $u^{[1]}\equiv u^{[2]}$. □

Therefore, via Lemma 2.1, Lemma 2.2, and Lemma 2.3, one can state:

Theorem 2. *It is not possible to represent an exact solution via a superposition of distinct BGK islands with separate phase velocities, nor a BGK island that resonates with a dispersive wave; therefore, dispersive waves cannot exist in systems with only conservative orbits.*

When one considers multiple phase space islands as BGK modes, they are in fact solving the problem on discontinuous subsets of phase space, i.e.:

$$u=\begin{cases}u^{[1]} & \text{for } v\in V_1\\ u^{[2]} & \text{for } v\in V_2\\ \vdots & \vdots\end{cases}.$$

This is inherently unphysical; not only would one be assuming that the electric potential is velocity dependent (which is not true), one would also be assuming that its velocity derivatives are in fact singular at the edges of the subdomains in v. In Chapter 3, I show that one can achieve a similar effect by 'enveloping' the potential. This is achieved through the use of a basis expansion (see Section 3.4.3 for an example).

2.3 Linearised kinetic plasmas

In contrast to BGK theory, in this section I examine linear plasmas. The linearised theory of kinetic plasmas (as notably investigated by Landau[33]) allows one to recover the growth

rate and frequency of waves in the system under two caveats: that the perturbation to the system from an initial condition is small, and that the frequency and growth rate of each wave is approximately time invariant.

First, I extend the derivation for Landau damping to cover collisional and magnetized plasmas in Section 2.3.1. In Section 2.3.2 and Section 2.3.3, I discuss how one can recover the eigenstructure of the system and the linear complex dispersion relation for resonant kinetic waves. In Section 2.3.4 and Section 2.3.5, I simplify the equations showing that they reduce to familiar equations which I recap from the literature. Finally in Section 2.3.6, I discuss the existence of phase space structures via a new method, and motivate the use of BGK theory as discussed in Section 2.2.

2.3.1 Laplace transform method

I will take the Boltzmann-Maxwell system as given in (1.17), expand using the δ-operator, and then discards all terms of $\mathcal{O}(\delta f)^2$. The perturbation can be approximated as a linear functional:

$$f(x, v, t) \approx f_{(0)} + \delta f[\kappa_{(0)}; \kappa].$$

I define the function $\kappa_{(0)}$ such that:

$$\kappa_{(0)} \, : \, f_{(0)} = f_{(0)}(v),$$

that is to say, $f_{(0)}$ is considered here to be only a function of v. By approximating the perturbation as a linear functional, I am stating that:

$$\delta f[\kappa_{(0)}; \kappa] \approx f - f_{(0)}.$$

Typically in the literature, linearisation is analysed from the point of view of the full Δ-variation of the PDF. Here, I use linear functional theory to later allow for discussion of the limitations of the PDFs produced in Chapter 3 in finer detail. For cases where the PDF is analysed using affine parameters along orbits (as was the case in BGK theory), one must take note that the perturbations garnered from this theory ultimately discard any nonlocal contributions to Δf that may arise in nonlinear systems.

If one examines the part of the system with $O(\delta f)^0$, one yields a set of equations which defines the initial state of the system. Examining the part of the system with $O(\delta f)$):

$$\begin{aligned}\hat{C}(\delta f) = \frac{q}{m}\left\{[\delta\mathbf{E} + \mathbf{v}\times\delta\mathbf{B}]\cdot\nabla_{\mathbf{v}} f_{(0)} + \left[\mathbf{E}_{(0)} + \mathbf{v}\times\mathbf{B}_{(0)}\right]\cdot\nabla_{\mathbf{v}}\delta f\right\} \\ + \left[\frac{\partial}{\partial t} + \mathbf{v}\cdot\nabla\right]\delta f,\end{aligned} \tag{2.3a}$$

$$\nabla\cdot\delta\mathbf{E} = \frac{1}{\epsilon_0}\sum_{\text{species}} q\int_{\mathbb{R}^3}\delta f\,\mathrm{d}\mathbf{v}, \tag{2.3b}$$

$$\nabla\times\delta\mathbf{E} = -\frac{\partial}{\partial t}\delta\mathbf{B}, \tag{2.3c}$$

$$\nabla\cdot\delta\mathbf{B} = 0, \tag{2.3d}$$

$$\nabla\times\delta\mathbf{B} = \mu_0\epsilon_0\frac{\partial}{\partial t}\delta\mathbf{E} + \mu_0\sum_{\text{species}} q\int_{\mathbb{R}^3}\mathbf{v}\delta f\,\mathrm{d}\mathbf{v}. \tag{2.3e}$$

First, we will examine the Boltzmann equation, (2.3a). By Fourier decomposing the equation in x, and taking the Laplace transform in t:

$$\begin{aligned}\hat{L}[\hat{C}\delta f]_j = (p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v})\delta\tilde{f}_j - \delta f_j(t=0) \\ + \frac{q}{m}\left\{\left[\delta\tilde{\mathbf{E}}_j + \mathbf{v}\times\delta\tilde{\mathbf{B}}_j\right]\cdot\nabla_{\mathbf{v}} f_{(0)} + \left[\mathbf{E}_{(0)} + \mathbf{v}\times\mathbf{B}_{(0)}\right]\cdot\nabla_{\mathbf{v}}\delta\tilde{f}_j\right\},\end{aligned} \tag{2.4}$$

where $p_j \equiv \gamma_j - \mathrm{i}\omega_j$, with $(\gamma_j, \omega_j 0$ as the growth rate and frequency of a wave with wave vector $\mathbf{k}_j$. This is a vector differential equation in $\mathbf{v}$ for $\delta\tilde{f}_j$. At this point it becomes useful to separate $\hat{C}$ in Fourier space. Next, one can separate the operator into two parts: the contribution which is a differential operator (denoted with subscript (∇)), and the contribution which is non-differential (denoted with subscript (Δ)). As such, $\hat{C}$ is representable in the form:

$$[\hat{C}f]_j \equiv \hat{C}_{j,(\Delta)} f_j + \hat{C}_{j,(\nabla)} f_j, \tag{2.5}$$

with the operator $\hat{C}_{j,(\Delta)}$ defined as satisfying the following:

$$\hat{C}_{j,(\Delta)} f_j := P_j(\mathbf{v})(f - F)_j, \tag{2.6}$$

where $F(\mathbf{x}, \mathbf{v}, t)$ is a currently unspecified function. Then, one can recast (2.4) in the form:

$$\delta\tilde{f}_j = \frac{\delta f_j(t=0) - P_j F_j - \frac{q}{m}\left[\delta\tilde{\mathbf{E}}_j + \mathbf{v}\times\delta\tilde{\mathbf{B}}_j\right]\cdot\nabla_{\mathbf{v}} f_{(0)}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j} + \hat{O}_j\delta\tilde{f}_j, \tag{2.7}$$

where $\hat{O}_j$ is a scalar differential operator given by:

$$\hat{O}_j = \frac{\hat{C}_{j,(\nabla)} - \left[\mathbf{E}_{(0)} + \mathbf{v}\times\mathbf{B}_{(0)}\right]\cdot\nabla_{\mathbf{v}}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j}. \tag{2.8}$$

I declare the following rules for the operator:

- **Identity operator.** The identity operator is defined as:

$$\hat{1} =: (\hat{O}_j)^0, \tag{2.9}$$

  such that $\hat{1}Q = Q$.

- **Order of operation.** I use the following shorthand:

$$(\hat{O}_j)^{n+1}(\delta\tilde{f}_j) := \hat{O}_j((\hat{O}_j)^n(\delta\tilde{f}_j)) \qquad \forall\, n \geq 0. \tag{2.10}$$

It is fairly trivial to show that the operator satisfies the following:

$$\lim_{n\to\infty}\left|\int_{\mathbb{R}^3} v^\nu \left(\hat{O}_j\right)^n \delta\tilde{f}_j \,\mathrm{d}^3\mathbf{v}\right| = 0\,\forall\,\nu, \tag{2.11}$$

therefore, the L^1-norm of $\lim_{n\to\infty} v^\nu \left(\hat{O}_j\right)^n \delta\tilde{f}_j$ on $\mathbb{R}^3$ is identically zero for each and every ν. Therefore, by iteration it is possible to show that:

$$\delta\tilde{f}_j \equiv \hat{O}_{j,\infty}\left\{\frac{\delta f_j(t=0) - P_j F_j - \frac{q}{m}\left[\delta\tilde{\mathbf{E}}_j + \mathbf{v}\times\delta\tilde{\mathbf{B}}_j\right]\cdot\nabla_{\mathbf{v}} f_{(0)}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j}\right\}, \tag{2.12}$$

where $\hat{O}_{j,\infty}$ is defined as:

$$\hat{O}_{j,\infty} := \sum_{n=0}^{\infty}(\hat{O}_j)^n, \tag{2.13}$$

where $\hat{O}_{j,\infty}\delta f_j$ is not required to be a convergent series. In this form, it is clear that P_j serves to shift the location of the Landau resonance. Next, one can take the curl of the Faraday-Lenz law, and the curl of the Maxwell-Ampère law. Both of these combine all 4 Maxwell's equations, and allow one to obtain vector Laplacians in the **E** and **B** field. By using an identity for the vector Laplacian[43], one finds the equations:

$$\nabla^2\delta\mathbf{E} = \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\delta\mathbf{E} + \sum_{\text{species}} q\int_{\mathbb{R}^3}\left(\mu_0\mathbf{v}\frac{\partial}{\partial t} + \frac{\nabla}{\epsilon_0}\right)\delta f\,\mathrm{d}\mathbf{v}, \tag{2.14a}$$

$$\nabla^2\delta\mathbf{B} = \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\delta\mathbf{B} - \sum_{\text{species}} q\int_{\mathbb{R}^3}\mu_0(\nabla\times\mathbf{v}\delta f)\,\mathrm{d}\mathbf{v}. \tag{2.14b}$$

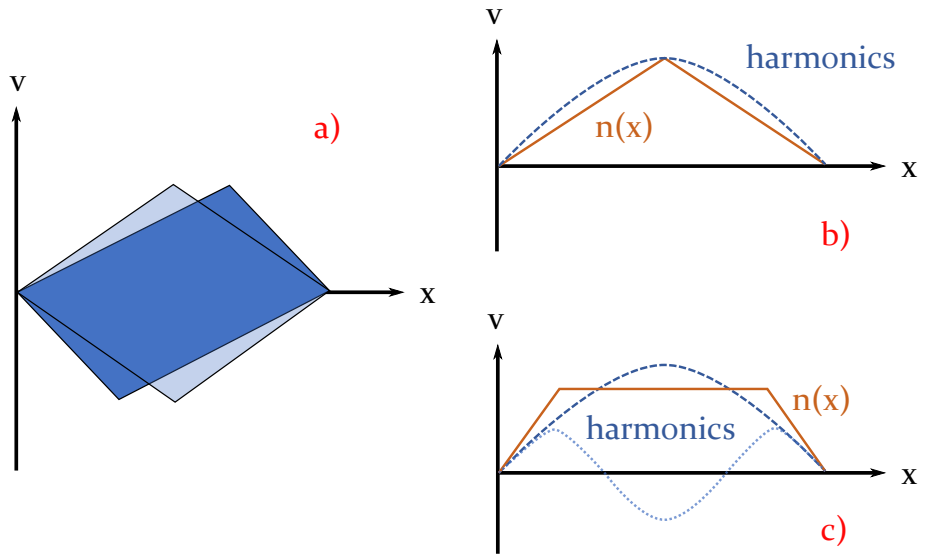


Figure 2.4: **Shear as nonlinear wave-wave coupling.** A 1D phase space structure undergoing shear strengthens higher order harmonics of the density perturbations in the system, corresponding to nonlinear coupling between particle density waves and electromagnetic waves via the Lorentz force. **a):** Sketching of a phase space structure with non-conservative orbits undergoing phase space shear. **b):** Sketch of the particle density prior to shear, and the fundamental harmonic under Fourier analysis. **c):** Sketch of the particle density after shear, showing that the amplitude of the third harmonic increases.

As expected from elementary electromagnetism, charge density gradients and current density fluctuations lead to an inhomogeneous wave equation for $\mathbf{E}$, while the vorticity of the particles leads to an inhomogeneous wave equation for $\mathbf{B}$.

Phase space structures can undergo shear due to nonlinear interactions in the system. As a structure undergoes shear, the Fourier decomposition of the structure changes. To illustrate this, as shown in Figure 2.4, a 1D box shaped structure in phase space shears into a parallelogram. The original Fourier decomposition features a primary wave, and a range of smaller harmonics. However, after shear, the Fourier decomposition now features a larger third harmonic. One can consider (2.14) as a set of independently driven simple harmonic motion (SHM) equations, one for each coupling of ω_j and $\mathbf{k}_j$. Therefore, if the Fourier decomposition for δf changes, the amplitude of the drive term given by charge and density currents changes for a given harmonic.

One can therefore consider the shear of phase space structures as nonlinearly modifying the stability of other waves in the system. Here, I only consider temporally static frequencies (viz. $\gamma_j(t) \approx \gamma_j(0)$, $\omega_j(t) \approx \omega_j(0)$).

In rectilinear coordinates, one can write $\delta\mathbf{E}_j$ and $\delta\mathbf{B}_j$ in the form:

$$\delta\tilde{\mathbf{E}}_j(p_j) = \delta\tilde{E}_j(p_j)\mathbf{e}(p_j), \tag{2.15a}$$

$$\delta\tilde{\mathbf{B}}_j(p_j) = \delta\tilde{B}_j(p_j)\mathbf{b}(p_j), \tag{2.15b}$$

where the vector direction of the field is dependent on p_j. Physically this means that every mode can propagate in its own direction. By performing spectral decomposition and Laplace transforming (2.14) one finds:

$$(p_j^2\mu_0\epsilon_0 + k_j^2)\delta\tilde{E}_j = -\sum_{\text{species}} q\mathbf{e}\cdot\int_{\mathbb{R}^3}\left(\mu_0\mathbf{v}p_j + \frac{\mathrm{i}\mathbf{k}}{\epsilon_0}\right)\delta\tilde{f}_j\,\mathrm{d}\mathbf{v}, \tag{2.16a}$$

$$(p_j^2\mu_0\epsilon_0 + k_j^2)\delta\tilde{B}_j = \sum_{\text{species}} q\mathbf{b}\cdot\int_{\mathbb{R}^3}\mu_0\mathrm{i}(\mathbf{k}\times\mathbf{v})\delta\tilde{f}_j\,\mathrm{d}\mathbf{v}. \tag{2.16b}$$

Therefore, one can construct a matrix equation of the form:

$$\boxed{[(p_j^2\mu_0\epsilon_0 + k_j^2)\underline{\mathbf{I}} - \underline{\mathbf{M}}]\begin{pmatrix}\delta\tilde{E}_j\\ \delta\tilde{B}_j\end{pmatrix} = \begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix},} \tag{2.17}$$

where $\underline{\mathbf{I}}$ is the identity matrix, β_1 and β_2 are given by:

$$\beta_1 = -\sum_{\text{species}} q\mathbf{e}\cdot\int_{\mathbb{R}^3}\left(\mu_0\mathbf{v}p_j + \frac{\mathrm{i}\mathbf{k}}{\epsilon_0}\right)\hat{O}_{j,\infty}\left[\frac{\delta f_j(t=0) - P_jF_j}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j}\right]\mathrm{d}\mathbf{v}, \tag{2.18a}$$

$$\beta_2 = \sum_{\text{species}} q\mathbf{b}\cdot\int_{\mathbb{R}^3}\mu_0\mathrm{i}(\mathbf{k}\times\mathbf{v})\hat{O}_{j,\infty}\left[\frac{\delta f_j(t=0) - P_jF_j}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j}\right]\mathrm{d}\mathbf{v}, \tag{2.18b}$$

and $\underline{\mathbf{M}}$ is given by the outer product:

$$\underline{\mathbf{M}} = \sum_{\text{species}}\frac{q^2}{m}\int_{\mathbb{R}^3}\mathrm{d}\mathbf{v}\begin{pmatrix}\mathbf{e}\cdot(\mu_0\mathbf{v}p_j + \frac{i\mathbf{k}}{\epsilon_0})\\ -\mathrm{i}\mu_0\mathbf{b}\cdot(\mathbf{k}\times\mathbf{v})\end{pmatrix}\hat{O}_{j,\infty}\begin{pmatrix}\mathbf{e}\cdot\nabla_{\mathbf{v}} & \mathbf{b}\cdot(\mathbf{v}\times\nabla_{\mathbf{v}})\end{pmatrix}f_{(0)}. \tag{2.19}$$

This form allows one to see the order in which different phenomena are mapped onto each other. First, the row vector considers the initial push of particles subject to the Lorentz force. Then, the operator $\hat{O}_{j,\infty}$ adds the effect of collisions and the initial electromagnetic field. Then, this is mapped onto the current density, charge density gradients, and vorticity. This then determines the evolution of the $\mathbf{E}$ and $\mathbf{B}$ field.

The velocity integrands have poles at:

$$\boxed{\gamma_j = \Im(\mathbf{v})\cdot\mathbf{k}_j + \Re(P_j),} \tag{2.20a}$$

$$\boxed{\omega_j = \Re(\mathbf{v})\cdot\mathbf{k}_j - \Im(P_j).} \tag{2.20b}$$

This resonance condition is known as the Landau resonance, where the particle velocity is matched with the plasma wave phase velocity ω_j/k_j. The location of the resonance affects the residues which contribute to the integral.

2.3.2 Eigenvalue representation

If one considers no initial perturbation to the system, then the β-vector is zero. This yields:

$$[(p_j^2\mu_0\epsilon_0 + k_j^2)\underline{\mathbf{I}} - \underline{\mathbf{M}}] \begin{pmatrix} \delta\tilde{E}_j \\ \delta\tilde{B}_j \end{pmatrix} = 0,$$

which is an eigenvalue problem. The eigenvalues are equal to the linear complex dispersion relation for the system, and are given by the secular determinant:

$$\det[(p_j^2\mu_0\epsilon_0 + k_j^2)\underline{\mathbf{I}} - \underline{\mathbf{M}}] = 0.$$

The corresponding eigenvectors define the allowed couplings between the **E**-field and **B**-field.

2.3.3 Pole representation

If there is an initial perturbation to the system, it becomes useful to consider how Landau's method reconstructs modes. Via the residue theorem[43], one finds that any closed contour integral in complex space is equal to the sum of the residues inside the contour. The forward Laplace transform of a quantity $Q(t)$ is analytic everywhere where $\Re(p) \geq \sigma$, where the real number σ is defined as:

$$\lim_{t\to\infty} Q(t) \propto \mathrm{e}^{\sigma t}.$$

As such, there are singularities in the forward Laplace transform at $\Re(p) < \sigma$. The backward Laplace transform typically uses the Bromwich contour, which is along constant $\Re(p)$, where $\Re(p) > \sigma$. One finds that as $\Re(p_j) \to 0$, the Bromwich contour reproduces the backward Fourier transform.

Suppose that one chose to allow the forward Laplace transform to be singular. Then, the poles in the complex plane would appear to the left of the Bromwich contour (see Figure 2.5). In Landau's 1946 paper,[33] he considered a deformation of the Bromwich contour by dragging the imaginary part to $-\infty$. Doing so means that one must form keyhole contours around the singularities to maintain the same value of the backward

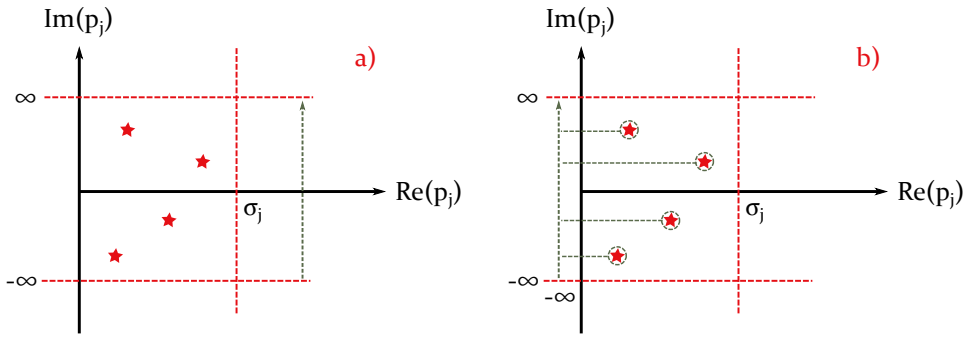


Figure 2.5: **The Landau contour.** For the Laplace transform of a quantity Q_j to be well defined, $\Re(p_j) \geq \sigma_j$ where $\lim_{t\to\infty} Q_j(t) \propto \mathrm{e}^{\sigma_j t}$. As such, the backward Laplace transform is typically taken over the Bromwich contour. **a):** Integration over the Bromwich contour (green dashed line) as utilised for the backward Fourier transform. This is a region where the Fourier transform is wholly analytic. **b:)** Integration over the Landau contour (green dashed line) as utilised for the backward Laplace transform. By dragging the integration contour to $-\infty$ in $\Re(p_j)$, as the value of the backward transform is independent of the value of $\Re(p_j)$ used along the contour, the value of the backward Laplace transform is equal to the sum of the residues at the singularities (see (2.21)).

transform, by virtue of Cauchy's theorem[43]. This contour is typically referred to as the *Landau contour*.

Inversion of this principle allows one to evaluate the backward transform equivalently by taking the residues of all the singular integrand values of the forward transform, and via Jordan's lemma[43] enclosing the contour at infinity.[43] Simple poles in the complex plane manifest as modes in the real space:

$$\int_{\Omega_{\mathrm{Lan.}}} \left\{ \prod_j \frac{\tilde{Q}(p)}{2\pi \mathrm{i}(p-p_j)} e^{pt} \right\} \mathrm{d}p = \sum_j \left(\tilde{Q}(p) e^{pt} \right)_{p=p_j} . \tag{2.21}$$

This elegant solution shows that if one examines $\{\delta\tilde{E}_j, \delta\tilde{B}_j\}$, any simple poles infact define the resonance condition for eigenmodes of the system. In matrix form, this manifests as scenarios where the matrix equation (2.19) cannot be inverted.

Remarkably for finding the linear complex dispersion relation, this gives exactly the same requirement given by the eigenvalue representation.

2.3.4 Light waves

In the case that the electromagnetic field is perpendicular to the perturbations to the current density, charge density gradients, and vorticity, the matrix $\underline{\mathbf{M}}$ is equal to the zero

operator.

Then, the linear complex dispersion relation is given by:

$$\det[[(p_j^2\mu_0\epsilon_0 + k_j^2)\underline{\mathbf{I}}] = 0.$$

This gives the eigenvalues as $p_j = \pm \mathrm{i}ck_j$ where c is the speed of light. Accordingly, these are light waves which simply do not couple to the plasma.

2.3.5 Electrostatic waves

For the electrostatic case there is no perturbation to the **B**-field. The matrix equation then yields an equation that is the separable sum of the Maxwell-Ampére law and Gauss' law, with each contribution independently equal to zero:

$$\overbrace{p_j\left(p_j\mu_0\epsilon_0\delta\tilde{E}_j + \sum_{\text{species}} q\mathbf{e}\cdot\int_{\mathbb{R}^3}\mu_0\mathbf{v}p_j\delta\tilde{f}_j\,\mathrm{d}\mathbf{v}\right)}^{\equiv 0} + \underbrace{k_j^2\delta\tilde{E}_j\sum_{\text{species}} q\mathbf{e}\cdot\int_{\mathbb{R}^3}\frac{i\mathbf{k}}{\epsilon_0}\delta\tilde{f}_j\,\mathrm{d}\mathbf{v}}_{\equiv 0} = 0.$$

As such, I will work directly with the contribution arising from the Maxwell-Ampére law for ease:

$$p_j\delta\tilde{\mathbf{E}}_j = -\sum_{\text{species}} q\int_{\mathbb{R}^3}\mathbf{v}\frac{\delta\tilde{f}_j}{\epsilon_0}\,\mathrm{d}\mathbf{v}. \tag{2.22}$$

Substituting (2.12), one finds that under no initial perturbation:

$$p_j\delta\tilde{\mathbf{E}}_j = -\sum_{\text{species}}\frac{q}{\epsilon_0}\int_{\mathbb{R}^3}\mathbf{v}\hat{O}_{j,\infty}\left\{\frac{\delta f_j(t=0) - P_jF_j - \frac{q}{m}\left[\delta\tilde{\mathbf{E}}_j\right]\cdot\nabla_{\mathbf{v}}f_{(0)}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j(\mathbf{v})}\right\}\mathrm{d}\mathbf{v}.$$

And so, the poles of $\delta\tilde{E}_j$ are given by:

$$\boxed{p_j = \sum_{\text{species}}\frac{q^2}{m\epsilon_0}\int_{\mathbb{R}^3}(\mathbf{e}\cdot\mathbf{v})\hat{O}_{j,\infty}\left\{\frac{\mathbf{e}\cdot\nabla_{\mathbf{v}}f_{(0)}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - P_j(\mathbf{v})}\right\}\mathrm{d}\mathbf{v}.} \tag{2.23}$$

When involving collisions, this integral equation is best solved numerically. It can be solved by using the shooting method and a finite term approximation for $\hat{O}_{j,\infty}$. One can then repeatedly iterate this solver with an increasing number of terms for $\hat{O}_{j,\infty}$, until a suitable level of convergence has been achieved.

Quasineutral plasma equilibrium with Krook collisions

If the plasma equilibrium is quasineutral, then the Lorentz force is initially zero. If one also examines only Krook collisions,[44] such that:

$$\hat{C}f = \nu_{\text{Kr.}}(f - F),$$

where $\nu_{\text{Kr.}}$ is the Krook collision frequency, then $\hat{O}_j = 0$, and therefore $\hat{O}_{j,\infty} = \hat{1}$ by virtue of (2.9). Therefore, one finds:

$$p_j = \sum_{\text{species}} \frac{q^2}{m\epsilon_0} \int_{\mathbb{R}^3} (\mathbf{e}\cdot\mathbf{v}) \left\{ \frac{\mathbf{e}\cdot\nabla_{\mathbf{v}} f_{(0)}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - \nu_{\text{Kr.}})} \right\} \mathrm{d}\mathbf{v}.$$

This can also be numerically solved by using the shooting method. In Appendix A.1, I show analytically that for waves travelling in rectilinear coordinates, the frequency of electrostatic waves in the system is given to zeroth order in $|\gamma_j - \nu_{\text{Kr.}}|$ by:

$$\omega_j(k_j, \gamma_j, \nu_{\text{Kr.}}) \approx \omega_{\text{pl.}}, \tag{2.24}$$

where $\omega_{\text{pl.}}$ is the plasma frequency:

$$\omega_{\text{pl.}}^2 := \frac{\bar{n}_{(0)} q^2}{m\epsilon_0}, \tag{2.25}$$

where $\bar{n}_{(0)}$ is the spatially averaged equilibrium number density. The growth rate is approximately given by:

$$\gamma_j(k_j, \omega_j, \nu_{\text{Kr.}}) \approx \frac{\nu_{\text{Kr.}}}{2} + \sum_{\text{species}} \frac{\pi\omega_j^3}{2k_{j,(p)}^2} \int_{\mathbb{R}^2} \left.\frac{\partial F}{\partial v_{(p)}}\right|_{v_{(p)}=\omega_j/k_{j,(p)}} \mathrm{d}A, \tag{2.26}$$

where $(\mathbf{e}\cdot\mathbf{v}) =: v_{(p)}$, and $\mathbf{e}$ is normal to $\mathrm{d}A$.

2.3.6 Dawson function representation

Suppose that one attempts to explicitly solve (2.4). If one takes the collisionless electrostatic case, using the definition $\mathbf{v}\cdot\mathbf{E}_{(0)} \equiv v_\parallel E_{(0)}$, one finds:

$$\frac{\partial}{\partial v_\parallel}\delta\tilde{f}_j + \frac{\mathrm{i}k_{j,\parallel}(v_\parallel - v_{\text{Lan.}})}{qE_{(0)}/m}\delta\tilde{f}_j = -\frac{\delta\tilde{\mathbf{E}}_j}{E_{(0)}}\cdot\nabla_{\mathbf{v}} f_{(0)}, \tag{2.27}$$

where the complex velocity $v_{\text{Lan.}}$ is given by:

$$v_{\text{Lan.}} = \frac{\mathrm{i}(p_j - \nu_{\text{Kr.}}) + k_{j,\parallel}v_\parallel - \mathbf{k}\cdot\mathbf{v}}{k_{j,\parallel}}. \tag{2.28}$$

This is a first order partial differential equation (PDE) soluble via integrating factor, yielding:

$$\delta\tilde{f}_j = \delta\tilde{f}_j\Big|_{v_\parallel = v_{\text{Lan.}}} - \mathrm{e}^{\left[-\frac{ik_{j,\parallel}}{2qE_{(0)}}(v_\parallel - v_{\text{Lan.}})^2\right]} \int_0^{v_\parallel - v_{\text{Lan.}}} \frac{\delta\tilde{\mathbf{E}}_j}{E_{(0)}} \cdot \nabla_{\mathbf{v}} f_{(0)} \mathrm{e}^{\left[\frac{ik_{j,\parallel}y^2}{2qE_{(0)}}\right]} \,\mathrm{d}y,$$

where $y = v_\parallel - v_{\text{Lan.}}$. If $\nabla_{\mathbf{v}} f_{(0)}$ is approximately constant, as is commonly approximated in the literature, this is representable in the form:

$$\delta\tilde{f}_j \approx \delta\tilde{f}_j\Big|_{v_\parallel = v_{\text{Lan.}}} - \frac{\delta\tilde{\mathbf{E}}_j}{E_{(0)}} \cdot \nabla_{\mathbf{v}} f_{(0)} \mathcal{D}\left[(v_\parallel - v_{\text{Lan.}})\sqrt{\frac{\mathrm{i}mk_{j,\parallel}}{2qE_{(0)}}}\right], \tag{2.29}$$

where $\mathcal{D}[x]$ is the Dawson function.[45] One expects the solution given in (2.29) to still obtain singularities which yield the eigenmodes of the system; however if one examines the case where $v_{\text{Lan.}} \in \mathbb{R}$, the Laplace transforms employed reduce to Fourier transforms, and therefore one expects no singular behaviour.

In this analytic region of $\delta\tilde{f}_j$, the fact that the solution can be represented in terms of the Dawson function is particularly interesting for a couple of reasons. Firstly, it is related to the plasma dispersion function $\zeta(y)$ (as defined by Fried and Conte[46]) as follows:

$$\mathcal{D}[x] = \frac{1}{2}\left[\mathrm{i}\sqrt{\pi}\exp(-x^2) - \zeta(x)\right].$$

This implies a strong link between δf and the linear complex dispersion relation, as one would intuitively expect for the non-analytic region. Secondly, the Dawson function is very closely related to the Hilbert transform of the Gaussian:

$$\mathcal{D}[x] = \frac{\sqrt{\pi}}{2}\left[\frac{1}{\pi}\int_{\mathbb{R}} \frac{e^{-y^2}}{x-y}\,\mathrm{d}y\right].$$

This definition is particularly interesting, as recent work by Heninger *et al.*[47] using the G-transform method (related in turn to the Hilbert transform) has allowed for linearised solutions for collisional plasmas with operators $\mathcal{O}(\partial_v^2)$. Indeed, the solution given in (2.29) is consistent with the G-transform theory if:

$$\frac{\partial f_{(0)}}{\partial v} \sim \exp\left[-\frac{m(v - v_{\text{Lan.}})^2}{2qE_{(0)}}\right],$$

where $\sigma \gg 1$. This form predicts two behaviours: that the PDF has a sigmoid shape near resonance, and that the resonance broadens with the strength of the electric field.

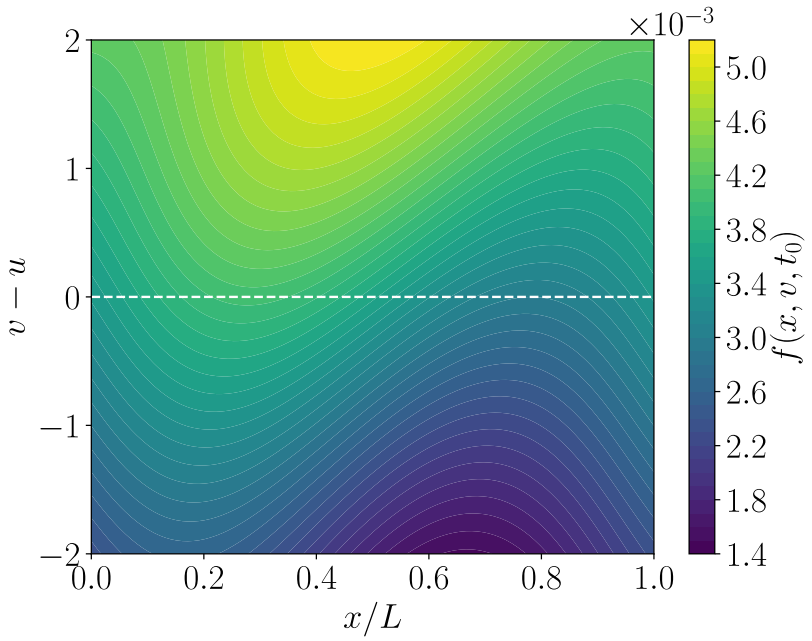


Figure 2.6: **Observation of phase space shear**. A 1D electrostatic plasma with a single ion species. The initial ion PDF is given by a 1D BOT distribution and normalised parameterisation $(\eta, v_T, v_B) = (0.95, 4, 10)$. The normalisation employed is the same as that used in Chapter 5. Phase space shear is observed with the direction of the shear (positive x) corresponding with the positive growth rate of a wave with $k = 2\pi/L$ with phase velocity $u \sim 6.66$ and a test value of $\gamma = 0.1$.

Suppose that one examines a test plasma with a single species of ions and single species of electrons. If the electrons are thermal, they cannot drive any Landau instabilities in the system. If the ions are thermal as well, the linear complex dispersion relation in a 1D electrostatic system is given by that which is expected in the literature for ion acoustic waves.[11] Accordingly, in a plasma with a small quantity of electron-ion collisions (such that for all j, the electron-ion collision frequency $\nu_{ei} \ll \gamma_j$, the linear growth rate of the j^{th} mode in the system), one can recover most of the dynamics by considering just the ion population; one can get a qualitative feel for the behaviour of the system by omitting the electrons and by shifting the frequencies in the system *post hoc* to account for the contribution to the wave speed from the electron density. If a population of suprathermal ions is added, such that the ions are now described by a 1D BOT distribution (1.19), one expects that the beam allows for inverse damping. Here, I will examine the following parameterisation:

$$(\eta, v_T, v_B) = (0.95, 4, 10),$$

defining a small beam population (5% of the particles) which is warmer and faster than the bulk ($v_B, v_T \gg 1$, the bulk thermal velocity). Using the normalisation given in Chapter 5, I briefly examine the linear growth of a single mode in a system defined such that the fundamental harmonic of the system is driven by the largest possible gradient on the bump at $v \sim 6.66$. One finds that for a system with length $L = 2\pi/0.15$, this is approximately satisfied as the frequency of the electrostatic wave in this ion only system is given by $\omega_j \approx 1$. By my choice of parameterisation, all higher harmonics resonate with the bulk ion population and are strongly Landau damped. Here, I will examine linear growth of this single mode in the system, as predicted by the Landau theory, and the corresponding phase space perturbations that can be approximated using this Dawson function analysis.

In Figure 2.6, I examine the point where the magnitude of the perturbation to the electric field $|\delta \mathbf{E}_j|$ is equal to 95% of the value at nonlinear saturation given by the literature[5] One can see even from this simple, quick model, the phase space is sheared as one would expect for a mode undergoing linear growth. Wave-particle resonance between the electrostatic wave and the PDF yields spatial perturbations in the PDF (δf) for finite wave amplitude. Then, the phase space structures shear via the advection term in the Boltzmann equation ($\mathbf{v} \cdot \nabla \delta f$).

2.3.7 Defining conservative theories

To conclude this chapter, I give a set of lemmas which can allow one to produce a conservative orbit theory of phase space island formation and migration in plasmas. While each one of the lemmas is not novel in itself, the collective use of them to describe 'conservative theories' is new:

Lemma 2.4. (Near conservative orbits) *In the linearised theory, structures grow on either side of the resonance, corresponding to wave-particle energy exchange. Locally (with respect to a structure located at $\mathbf{v} = \mathbf{u}_j$), particles approximately move along conservative orbits with energy $q\phi^{[j]} + \frac{1}{2}m|\mathbf{v} - \mathbf{u}_j|^2$, where $\phi^{[j]}$ is the potential energy of the interacting wave.*

Lemma 2.5. (Slowly chirping approximation) *In the limit that the chirping rate of the wave is much less than the bounce frequency of particles trapped in the resonant phase space islands ($\partial(\ln \omega_j)/\partial t \ll \omega_b$), one can assume a temporal continuum of linearised solutions, neglecting small perturbations to the particle orbits.*

Lemma 2.6. (Slowly growing approximation) *In the limit that the growth rate of the wave is much less than the bounce frequency of particles trapped in the resonant phase space islands ($\gamma_j \ll \omega_b$), one can assume a temporal continuum of linearised solutions, neglecting small perturbations to the particle orbits.*

Lemma 2.7. (Resonance broadening) *Under the slowly chirping approximation, the two perturbative structures near a given resonance move through phase space as the square root of the wave amplitude, corresponding to broadening of the resonance at $v = v_{Lan.}$.*

Lemma 2.4 justifies the use of BGK modes or at least 'BGK-like' modes for the case of distinct phase space structures, where each structure has all particles travelling along conservative orbits. It is largely supported by the literature,[5,8,9] and some of my simulations which are discussed in Chapter 5. However, these structures inherently assume that all particles interact with the wave. In Chapter 3, I show that this formally prevents the existence of two BGK islands in the same system, dispersive waves, and waves with finite growth rates.

However, one can consider each island separately as approximately given by a BGK island provided that they are spaced suitably far apart in phase space, and provided that Lemma 2.6 holds true. This lemma is similar to the adiabatic consideration used in the literature, however this lemma partially serves to illuminate the following point: the word 'adiabatic' here refers to analytical dynamics, not thermodynamics. That is to say, the consideration of the adiabatic evolution of H&Cs is with respect to an *adiabatic invariant* corresponding to the classical action of an unperturbed orbit in phase space, given by:

$$\oint_C \mathbf{v} \cdot \mathrm{d}\mathbf{x} \approx \text{const.},$$

where C is the unperturbed orbit. This can lead to some confusion as it is possible to have a plasma which is thermodynamically non-adiabatic (such as a tokamak plasma with radio frequency (RF) heating or neutral beam injection (NBI)) which is still heuristically described by phase space structures which evolve 'adiabatically'. The classical action is an adiabatic invariant if the corresponding angle variable evolves faster in time than the Hamiltonian of the system. Here, the rate of change of the angle variable here is the bounce frequency of particles in the orbit, and the rate of change of the Hamiltonian is proportional to the frequency chirping rate (as described in Section 1.6. As phase space structures move through phase space, the velocity of all of the resonant particles changes, causing the Hamiltonian to change. In Chapter 3, I show that the growth rate is intrinsically

linked to the chirping rate, and therefore one can consider adiabaticity through the use of appropriate ordering given by Lemma 2.5 and Lemma 2.6.

However, in reality, fewer and fewer particles interact with the wave the further they are from the resonance. This leads to deformation of the phase space contour, and deviation from conservative orbit. Lemma 2.7 agrees with the functional form of the chirping rate given by Berk *et al.*,[5] as the mode amplitude during nonlinear saturation can be shown to be proportional to the linear growth rate of the mode.

2.4 Brief conclusions

In this chapter, I showed a mixture of original work and background work relating to conservative orbit theories and linearised plasmas. While the theory of linearised plasmas is largely explored in the literature, the technique I utilised which allows one to recover collisions is novel. As such, everything in the chapter is novel except for the method of linearisation declared at the start of Section 2.3.1, and the Landau contour in Section 2.3.3 which have both been traditionally employed in the literature for some time.

The overall aim was to outline existing theory, with some small extensions: to highlight the limitations of BGK theory, and to derive the linear complex dispersion relation for electrostatic waves in collisional regimes.

Chapter 3

Non-conservative orbits and nonlinear kinetic theory

3.1 Overview

Work by Lilley *et al.* has shown computationally how holes and clumps (H&Cs) form via phase mixing of the wave within a phase space island.[41] Work by Wang *et al.* has recently examined frequency chirping for the late time evolution of plasma wave using a mixture of analytical theory and the CHIRP code.[48] However, the threshold for H&C formation and the formation of H&Cs on particle distribution functions (PDFs) with large curvature are both still relatively unexplored. As a result, understanding of H&C destabilisation is limited.

This chapter is comprised of original work. I begin in Section 3.2 by discussing 'non-conservative orbits', and how one can develop from conservative orbit theories such as Bernstein-Green-Kruskal (BGK) theory or extensions thereof.

In Section 3.2, I will compare results to conservative theories by showing that Lemma 2.4, Lemma 2.5, Lemma 2.6 and Lemma 2.7 are indeed all satisfied.

In Section 3.3, I discuss nonlinear kinetic theory. I show how one can represent the electromagnetic field in kinetic plasmas using a spectral decomposition where the waves have time-varying complex frequency. Then, I present a new representation of the Berk-Breizman energy sink[8] as an effective collision operator.

In Section 3.4 and Section 3.5, I show new results for solutions which are permitted for the orbit affine parameter ϵ along non-conservative orbit. In the former section, I prove which sets of basis expansions are permitted, and the corresponding non-holonomic constraints. In the latter section, I show that a perturbative theory allows one to generate

much of the desired behaviour using a generating function, ψ.

3.2 Non-conservative orbits

In Chapter 2, I examined particle orbits where the single particle energy is conserved (conservative orbits). Suppose that instead one considers orbits of the form:

$$\epsilon = \epsilon_0 + q\psi, \tag{3.1}$$

where ϵ is the orbit affine parameter such that $f = f[\epsilon]$, ϵ_0 is the following sum:

$$\epsilon_0 = \sum_l \left[q\phi^{[l]} + \frac{1}{2} m |\mathbf{v} - \mathbf{u}^{[l]}|^2 \right], \tag{3.2}$$

and ϕ is given by:

$$\phi(x,t) = \sum_l \phi^{[l]}(x,t). \tag{3.3}$$

Such orbits can be said to be *non-conservative* for $\psi \neq$ const., as the particle energy is not conserved along the orbit.

While all particles in real kinetic systems feel the electromagnetic field, only particles with velocity close to the phase velocity of electromagnetic waves in the system are able to exchange energy with the wave. This requires some form of velocity space 'shielding' akin to a Debye sheath.

I begin by examining the period of orbits in Section 3.2.1. By examining conservative orbits, one can infer how non-conservative orbits may behave.

The original BGK paper explicitly states that non-zero growth rates are not considered, however it is important to note that the solution will not work for finite growth rate. As will be shown in Section 3.2.3, the particle orbits deviate from conservative orbits when the wave amplitude changes. Furthermore, if the location of the phase space structures change, this also changes the shape of the particle orbits.

In addition, in Section 3.2.3, I discuss how waves interact directly with particle orbits, and how wave-orbit coupling allows for time-varying frequency and growth rate of waves.

In Section 3.2.4 and Section 3.2.5, I discuss how non-conservative orbits can be considered as particles living in an 'effective potential' which is velocity dependent. I then discuss how one can construct the PDF from a functional formulation, allowing one to represent the perturbation to the system using ϵ independent of the initial conditions.

3.2.1 Period of near-conservative orbits

As particles traverse phase space, they see an electric potential which varies in time and space. If one examines the Lorentz force for the i^{th} particle:

$$m\ddot{\mathbf{x}}^{(i)} = q[\mathbf{E} + \dot{\mathbf{x}}^{(i)} \times \mathbf{B}]_{\mathbf{x}=\mathbf{x}^{(i)}}.$$

One can calculate the work done by this force. As one might expect, the work done by the magnetic field is zero, giving:

$$\frac{1}{2}m|\dot{\mathbf{x}}^{(i)}|^2 = -q\phi_{\mathbf{x}=\mathbf{x}^{(i)}} + U(t), \tag{3.4}$$

where $U(t)$ is the single particle energy. Traditionally, one may consider orbits where $U(t)$ is conserved. Instead, I examine orbits with the following construction:

$$\text{orbits} := \left\{ \mathbf{x}(t) \in C \,\middle|\, U(t) = \frac{1}{2}m|\dot{\mathbf{x}}^{(i)}|^2 + q\phi_{\mathbf{x}=\mathbf{x}^{(i)}}; \quad U(t) - q\psi(t) = \text{const.} \right\}, \tag{3.5}$$

such that $\psi(t)$ is a function which enables orbits where $U(t)$ is allowed to vary. For conservative orbits, $\psi(t) = 0$, such that the constraint reduces to $U(t) = const.$ as is standard. I define the functional:

$$\epsilon[\psi, \phi; t] = \frac{1}{2}m|\dot{\mathbf{x}}^{(i)}|^2 + q\phi_{\mathbf{x}=\mathbf{x}^{(i)}} + q\psi[\mathbf{x}; t], \tag{3.6}$$

such that particles follow orbits where ϵ is constant. Note that ϵ is a functional which depends on ψ (encapsulating wave-particle energy exchange) and ϕ, but not the particle trajectory $\mathbf{x}(t)$. This is because the particle trajectory is determined by ϕ and ψ (viz. by virtue of (3.5), one uses the conditional trajectory $\mathbf{x}(t) \mid \phi, \psi$). One can find the time taken to complete a closed orbit by solving the differential equation given by (3.4):

$$\tau[t; \psi, \phi \mid \epsilon] = \oint_C \frac{\mathrm{d}|\mathbf{x}|}{\sqrt{\epsilon - q(\phi(\mathbf{x}, t) + \psi[\mathbf{x}; t])}},$$

where C is a given orbit, and τ is the period of orbit. The simplest family of orbits to consider are ones where $\psi(t)$ is completely in phase with $\phi(t)$. Here, I will consider a monochromatic wave:

$$\begin{pmatrix} \phi \\ \psi \end{pmatrix} = \begin{pmatrix} |\phi|(t) \\ |\psi|(t) \end{pmatrix} \cos\left(\mathbf{k} \cdot \mathbf{x}(t) - \int_0^t \omega \, \mathrm{d}t' + \theta \right).$$

Then, τ takes the form:

$$\tau[t;\psi,\phi\mid\epsilon]=\oint_C\frac{\mathrm{d}|\mathbf{x}|}{\sqrt{\epsilon-qA(t)\cos\left(\mathbf{k}\cdot\mathbf{x}(t)-\int_0^t\omega\,\mathrm{d}t'+\theta\right)}},$$

where $A(t)=|\phi|(t)+|\psi|(t)$. By defining $\chi=\mathbf{k}\cdot\mathbf{x}(t)-\int_0^t\omega\,\mathrm{d}t'$, one finds that as:

$$d|\mathbf{x}|\sim\frac{1}{k}(d\chi+\omega\,\mathrm{d}t).$$

Therefore, one finds that τ can split into a part in the wave frame, and a part arising from frequency sweeping:

$$\tau[t;\psi,\phi\mid\epsilon]\sim\oint_{\chi_0}^{\chi_1}\frac{\mathrm{d}\chi}{k\sqrt{\epsilon-qA(t)\cos(\chi+\theta)}}+\oint_t^{t+\tau}\frac{\omega\,\mathrm{d}t}{k\sqrt{\epsilon-qA(t)\cos(\chi+\theta)}}.$$

Inverting $\mathbf{x}(t)$ to obtain $t(\chi)$ is highly non-trivial, and requires knowledge of the distribution function. As such, I discard the second the term under the approximation that $\dot{\mathbf{x}}\gg\omega/k$. The validity of this approximation breaks down close to the resonance. One can Taylor expand the integrand about $A=A(t_0)$, yielding:

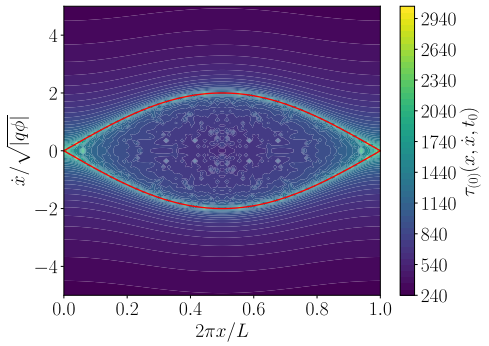
$$\tau[t;\psi,\phi\mid\epsilon]\approx\oint_{\chi_0}^{\chi_1}\frac{\mathrm{d}\chi}{k\sqrt{\epsilon-qA(t_0)\cos(\chi+\theta)}}\left[1+\frac{1}{2}\left(\frac{(A(t)-A(t_0))\cos(\chi+\theta)}{\epsilon-qA(t_0)\cos(\chi+\theta)}\right)+\dots\right].$$

τ is largely non-integrable except for the case where $A(t)$ is constant. This corresponds to the zeroth order contribution, which tells one the 'instantaneous' period of the orbit that the particle is on; if $U(t)$ was to then remain fixed for all time $t>t_0$, the particle would continue on a conservative orbit with period $\tau_{(0)}$. For passing orbits lying outside the separatrix, $\epsilon>|qA(t_0)|$, and:

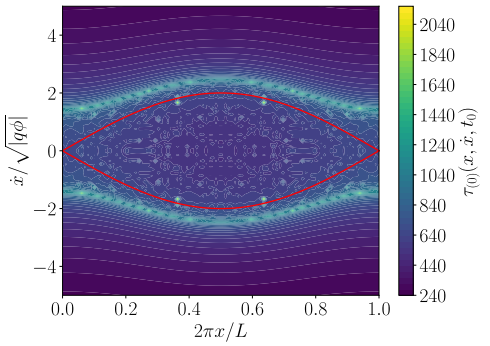
$$\begin{aligned}\tau_{(0)}[t;\psi,\phi\mid\epsilon,t=t_0]&:=\oint_0^{2\pi}\frac{\mathrm{d}\chi}{k\sqrt{\epsilon-qA(t_0)\cos(\chi)}}\\&\equiv\frac{2}{k}\left[\frac{1}{\sqrt{\epsilon-qA}}K\left(-\frac{2qA}{\epsilon-qA}\right)+\frac{1}{\sqrt{\epsilon+qA}}K\left(\frac{2qA}{\epsilon+qA}\right)\right]_{t=t_0},\end{aligned}$$

where $K(k)$ is the complete elliptic integral of the first kind.[45] Therefore, while the phase space structure is coherent (the entire structure moves with a single phase velocity), *the single-particle orbits are not coherent.* On the separatrix, $\tau\to\infty$. For trapped orbits lying inside the separatrix, $\chi_0>0$ and $\chi_1<2\pi$.

In Figure 3.1, I plot the period of near-conservative orbits as a function of x and $\dot{x}$ for particles in a single wavelength 1D potential. As one moves from an orbit outside the island towards the separatrix (decreasing U), the period of orbits increases. As one moves from



(a) **Case** $q\psi = 0$. The conservative orbit contours are exactly that which is expected from the BGK theory; each orbit corresponds to a different value of ϵ. This is reflected by the fact that the zeroth-order contribution to the period, $\tau_{(0)}$, tends to infinity along the separatrix.



(b) **Case** $q\psi > 0$. The contour where $\tau_{(0)} \to \infty$ occurs *outside* of the separatrix. Accordingly, particles which are outside of the separatrix but within the infinite $\tau_{(0)}$ contour may exhibit similar behaviour to trapped particles (which I refer to as 'quasi-passing particle' particles).

Figure 3.1: **Period of particle-orbits in a BGK island.** Two plots illustrating near-conservative orbits in a single wavelength 1D potential $\phi(x,t) = |\phi|\cos(kx - \omega t)$ with phase velocity $u = \omega/k$. Filled contours correspond to different values of $\tau_{(0)}$. The length of the 1D box is given by $L = 2\pi/k$.

an orbit on the separatrix to the centre of the island (decreasing U), the period of orbits decreases. But in addition, it is possible for the orbits of quasi-passing particles (particles which are not trapped, but interact somewhat with the wave) to decrease in period as one approaches the separatrix, provided that $q\psi > 0$.

As such, trapped particles which lose energy embark on orbits which spiral inwards, giving rise to phase mixing as observed by Lilley *et al.*[41] Contours of constant $\tau_{(0)}$ may be topologically challenging to represent analytically, as there will be finite discontinuities in the 6-dimensional gradient in the phase space. Conversely, passing particles which lose energy increase their period of orbit. I posit that as such, the period of the particle orbit may surpass the growth rate of a wave; in such a case, a passing particle becomes trapped.

3.2.2 Phase space shear

If the electric potential is not wholly periodic, then the orbit is deformed during the course of an orbit.

One can consider this to be a correction to the particle orbit, required when trans-

forming from the accelerating wave frame to the lab frame. As the amplitude of the wave changes, when the particle on a given orbit sees a different amplitude wave to when it started. This has the effect of elongating the path length of the orbit if the wave is growing.

3.2.3 Wave-orbit coupling

As the particle orbit changes to account for the finite growth rate, time-varying growth rate, and time-varying frequency, the generalised momentum for the system changes.

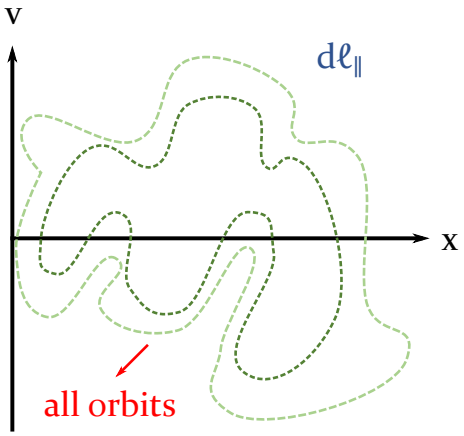


Figure 3.2: **Method of integration for wave-orbit coupling.** Integration is performed along an orbit (denoted as $d\ell_\parallel$), followed by integration across all available orbits.

In this subsection, I will derive the rate of energy transfer with the PDF as a function of orbit-based quantities. This will allow me to describe frequency chirping and wave growth as arising from properties of wave-orbit coupling, and to therefore discuss situations where the lemmas given in Section 2.3.7 may be violated. To start with, I will seek to perform a form of orbit integration:

$$\mathbf{P}_s(t) = \sum_{\mathbf{x}^{(i)} \in \text{orbit } s} m\dot{\mathbf{x}}^{(i)}, \tag{3.7}$$

where $\mathbf{P}_s$ is the contribution of every particle on orbit s to the momentum. By considering the Klimontovich density (1.10), one can write this as the following integral:

$$\mathbf{P}_s(t) = \oint_{\text{orbit } s} m\bar{\mathbf{v}}_s \, \mathrm{d}\ell_\parallel, \tag{3.8}$$

where $\mathrm{d}\ell_\parallel$ is an infinitessimal line element along the orbit, and $m\bar{\mathbf{v}}_s$ is given by:

$$m\bar{\mathbf{v}}_s := \sum_{\mathbf{x}^{(i)} \in \text{orbit } s} m\dot{\mathbf{x}}^{(i)} \delta[\mathbf{x} - \mathbf{x}^{(i)}] \delta[\mathbf{v} - \dot{\mathbf{x}}^{(i)}]. \tag{3.9}$$

However, as $a\,\delta(a-b) = b\,\delta(b-a)$, one can freely swap $\mathbf{x}$ and $\mathbf{x}^{(i)}$ under the sum, and $\mathbf{v}$ and $\dot{\mathbf{x}}^{(i)}$ under the sum.

One can then integrate over all orbits (see Figure 3.2) to give:

$$\mathbf{P}(t) = \int\limits_{\text{all orbits}} \oint\limits_{\text{orbit } s} m\bar{\mathbf{v}}_s |\mathcal{J}|\, \mathrm{d}\ell_{\parallel}\, \mathrm{d}^5\ell_{\perp}, \tag{3.10}$$

where $|\mathcal{J}|\, \mathrm{d}\ell_{\parallel}\, \mathrm{d}^5\ell_{\perp} = \mathrm{d}^3x^{(i)}\, \mathrm{d}^3\dot{x}^{(i)}$, such that $|\mathcal{J}|$ is the determinant of the Jacobian[43] of the transformation from $(\vec{\ell}_{\parallel}, \vec{\ell}_{\perp}) \to (\mathbf{x}^{(i)}, \dot{\mathbf{x}}^{(i)})$. The Jacobian[43] is required, as one expects that:

$$\mathbf{P}(t) := \iiint\limits_{\mathbb{R}^3} \iiint\limits_{\mathbb{R}^3} \left\{ \sum_{x^{(i)}} m\dot{\mathbf{x}}^{(i)} \delta[\mathbf{x} - \mathbf{x}^{(i)}] \delta[\mathbf{v} - \dot{\mathbf{x}}^{(i)}] \right\} \mathrm{d}^3x\, \mathrm{d}\mathbf{v}. \tag{3.11}$$

One finds that by choosing $\mathrm{d}\ell_{\parallel} = \mathrm{d}t$:

$$\mathrm{d}^3x^{(i)}\, \mathrm{d}^3\dot{x}^{(i)} = |\mathcal{J}|\, \mathrm{d}t\, \mathrm{d}^5\ell_{\perp}. \tag{3.12}$$

I shall now examine the form of the rate of change of momentum, without specifying the determinant of the Jacobian[43]. Therefore, (3.10) gives:

$$\mathbf{P}(t) = \mathbf{P}(0) + \int\limits_{\text{all orbits}} \oint\limits_{0}^{\tau_s(t,\dots)} m\bar{\mathbf{v}}_s |\mathcal{J}|\, \mathrm{d}\tau\, \mathrm{d}^5\ell_{\perp},$$

where $\tau_s(t, \dots)$ is the period of the s^{th} orbit. The rate of change of kinetic energy in the system for a given species is given by:

$$\dot{U} = \frac{1}{m} \mathbf{P} \cdot \dot{\mathbf{P}}. \tag{3.13}$$

If one takes the time derivative of the total momentum, via the Leibniz integral rule (see Appendix B.2):

$$\dot{\mathbf{P}}(t) = \int\limits_{\text{all orbits}} \left\{ \left[m\bar{\mathbf{v}}_s |\mathcal{J}| \frac{\partial \tau}{\partial t} \right]_0^{\tau_s(t,\dots)} + \oint\limits_0^{\tau_s(t,\dots)} \frac{\partial}{\partial \tau} \left[m\bar{\mathbf{v}}_s |\mathcal{J}| \right] \mathrm{d}\tau \right\} \mathrm{d}^5\ell_{\perp}\ .$$

By evaluating the time integral, one finds:

$$\boxed{\dot{U}(t) = \mathbf{P} \cdot \int\limits_{\text{all orbits}} \left[\bar{\mathbf{v}}_s |\mathcal{J}| \left(1 + \frac{\partial \tau}{\partial t} \right) \right]_{\tau=0}^{\tau=\tau_s(t,\dots)} \mathrm{d}^5\ell_{\perp}.} \tag{3.14}$$

For there to be no energy transfer between particles and the wave, the integrand must be zero. To enable this, the period of each orbit must be constant, the mean momentum along each orbit must be constant, and the determinant of the Jacobian[43] must be constant.

This allows one to identify three mechanisms which allow for energy transfer:

- **Perturbations to the period of each orbit.** If the period of an orbit is modified, typically one expects that the electric potential is changing. As one might expect, if there is a finite growth rate in the system, there is wave-particle resonance. However, the period (in $\mathbf{x}$) of an orbit can also oscillate if the wave is dispersive. The period of the oscillation is given by the beating frequency between the waves in the system; accordingly, waves with very similar frequency lead to oscillations in the period which occur over long time scales.

- **Perturbations to the mean momentum of each orbit.** If the orbit is translating through momentum space, this leads to wave-particle energy transfer. As such, frequency chirping waves transfer energy with translating phase space structures.

- **Perturbations to the Jacobian determinant.** This typically occurs when the shape of the phase space structure has been modified. Therefore, phase space structures undergoing shear transfer energy with the wave.

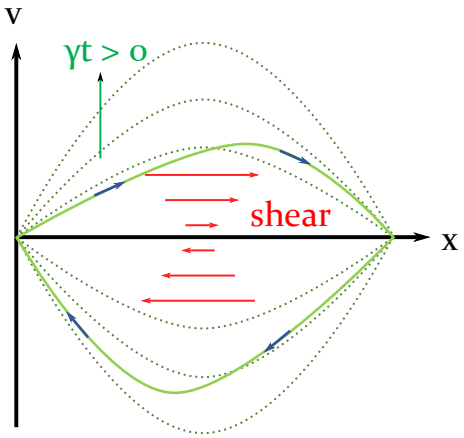


Figure 3.3: **Phase space shear induced by finite γ.** A passing or quasi-passing particle (see Section 3.2.1) interacting with a wave undergoing growth or decay can be thought of as moving across a continuum of conservative orbits. The resultant near-conservative orbit becomes elongated close to the X points of the orbit, resulting in phase space shear. The direction of the shear is related to the sign of the product $q\gamma$; as the sign of $q\gamma$ changes, the sign of the shear is also inverted.

As briefly discussed in Section 3.2.3, as a phase space structure undergoes shear, the

Fourier decomposition of the structure undergoes changes (see Figures 3.3 and 3.4). As the shape of phase space structures can be considered as given by contours of constant f, deformations to the determinant of the Jacobian[43] (and accordingly the shape of the orbit) allow for nonlinear modification of the stability of other waves in the system. Therefore, the determinant of the Jacobian[43] allows for wave-wave coupling.

This is an important point to note, as the phase space structure undergoes shear in a system with finite growth rate. Therefore, in all systems with finite growth rate there is inherently a finite amount of wave-wave coupling.

3.2.4 Effective potential

Intuitively, it is not possible to have a system with conservative orbits where there are finite growth rates.

$\phi + \psi$ can be thought of as acting together as an 'effective potential' manifesting from the collective effect of the charged particles on the electromagnetic field itself. Just as particles reorganise themselves to form a Debye sheath, one can consider a phase space island as a momentum space sphere which prevents all of the particles in the system from resonating with the wave. It is important to note that this is a **collective phenomenon**, and requires a suitably large number of particles; the fact that I consider the system as plasma already satisfies this condition.

3.2.5 Distribution function construction

In extension of the exposition of the δ-operator given in Section 1.3.2, I will now analyse how to reconstruct a 'first-order' PDF. One can then define κ_0 as the initial function used in the functional at $t = 0$:

Lemma 3.1. *If* $\exists \kappa_{(0)} : \partial f_{(0)}/\partial t = 0$*, then the PDF* $f[\kappa]$ *can be represented in the form:*

$$f[\kappa] = f(\mathbf{x}, \mathbf{v}, t = 0) + \Delta f[\kappa_{(0)}; \kappa].$$

Proof. If $\partial f_{(0)}/\partial t = 0$, then the functional form of $f_{(0)}$ is solely a function of x and v. One therefore has freedom of choice to define $\kappa_{(0)}$ such that:

$$\kappa_{(0)} \, : \, f_{(0)} = f(\mathbf{x}, \mathbf{v}, t = 0).$$

Therefore $f[\kappa]$ has been shown to have the aforementioned possible representation. □

This is particularly useful, as the perturbation to the PDF Δf can be represented using a completely independent functional; one could alternatively write:

$$f[\kappa] = f(\mathbf{x}, \mathbf{v}, t = 0) + g[h],$$

where g and h are arbitrary functions. It is also possible to construct a 'first-order' PDF by using Lemma 3.1. By splitting Δf into a linear and nonlinear contribution:

$$\Delta f = \int_{\Omega} \left\{ \mathrm{d}\vec{\mu}\, \Delta\kappa \frac{\delta}{\delta \kappa_{(0)}(\vec{\mu})} f_{(0)} \right\} + \mathcal{O}[\Delta\kappa^2]. \tag{3.15}$$

As such, one can approximate the PDF by discarding the terms which nonlinearly depend on $\Delta\kappa$. This generalises the technique used by Bernstein, Greene and Kruskal to obtain 'first-order' PDFs for BGK modes.[7]

3.3 Nonlinear kinetics

In this section, I build on work detailed in Section 3.2, allowing for an outline of the mathematical framework used for the remainder of this chapter. Nonlinear kinetic theory allows one to examine systems where nonlinear coupling occurs between the PDF f and the $\mathbf{E}$ and $\mathbf{B}$ field. As one expects from the convolution theorem, the product of two temporally/spatially varying quantities becomes the convolution of the two quantities in the Laplace/Fourier space.

The spatial and temporal dependence of f, $\mathbf{E}$ and $\mathbf{B}$ leads to wave-wave coupling. This wave-wave coupling means that energy can be transferred from one wave to another; accordingly the growth rates and frequencies for each wave in the system now become time-dependent.

In Section 3.3.1, I give exact allowed nonlinear solutions for the electric and magnetic fields under nonlinear theory. I discuss the asymptotic behaviour of the frequency and growth rate of modes, and give dynamic equations which allow for nonlinear wave-particle interactions.

In Section 3.3.2, I briefly touch on the expected energy balance for the system, and how sources and sinks in the Boltzmann equation affect the energy content of the system.

In Section 3.3.3, I derive a collision operator which yields the well known Berk-Breizman sink employed in the literature,[5] while maintain a consistent system of equations; in the literature, typically the Berk-Breizman sink is employed without a corresponding sink in the Boltzmann equation.

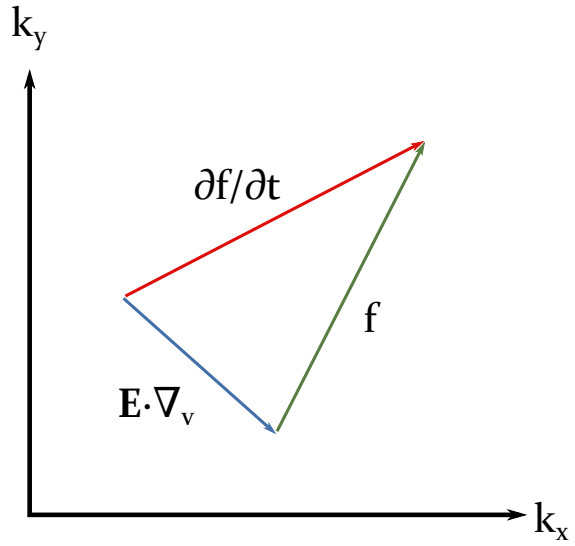


Figure 3.4: **Nonlinear wave-wave coupling.** Wave-wave coupling between the Lorentz force and the PDF directly affects the PDF, via (1.17a). Wave-wave coupling modifies the spatio-temporal structure of the waveform that defines the PDF, illustrated here spatially in 2D.

In Section 3.3.4, I discuss how the linear complex dispersion relation of waves is modified in models using the Berk-Breizman sink.

3.3.1 Complex frequency evolution

Similarly to (2.14), the evolution of the **E** and **B** fields are given by inhomogeneous wave equations, as expected from elementary electromagnetism:

$$\nabla^2\mathbf{E} = \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\mathbf{E} + \sum_{\text{species}} q\int_{\mathbb{R}^3}\left(\mu_0\mathbf{v}\frac{\partial}{\partial t} + \frac{\nabla}{\epsilon_0}\right) f\,\mathrm{d}\mathbf{v}, \tag{3.16a}$$

$$\nabla^2\mathbf{B} = \mu_0\epsilon_0\frac{\partial^2}{\partial t^2}\mathbf{B} - \sum_{\text{species}} q\int_{\mathbb{R}^3}\mu_0(\nabla\times\mathbf{v}f)\,\mathrm{d}\mathbf{v}. \tag{3.16b}$$

By coupling these to the Boltzmann equation (1.17a), one obtains a system of differential equations which features strong nonlinearities. One should note that $f = f(x, v, t)$, and therefore this is only part of a coupled system of equations. I shall solve (3.16) by representing the **E** and **B** field via a set of wavepackets $\{\phi_l\}$:

$$\begin{pmatrix}\mathbf{E}\\ \mathbf{B}\end{pmatrix}(x,t) := \sum_l \begin{pmatrix}\mathbf{E}^{[l]}\\ \mathbf{B}^{[l]}\end{pmatrix}(x,t). \tag{3.17}$$

Each wavepacket features waves which at a time t have a phase velocity $u^{[l]}$, such that

each wavepacket is dispersionless. Then, one can employ the following solution for each wavepacket:

$$\boxed{\begin{pmatrix} \mathbf{E}^{[l]} \\ \mathbf{B}^{[l]} \end{pmatrix}(x,t) := \frac{1}{2}\sum_j \left\{ \begin{pmatrix} \mathbf{E}_j^{[l]} \\ \mathbf{B}_j^{[l]} \end{pmatrix} \exp\left[\mathrm{i}\mathbf{k}_j \cdot \mathbf{x} + \int_0^t p_j^{[l]}(\tau)\,\mathrm{d}\tau \right] + \text{c.c.} \right\},} \tag{3.18}$$

where $\{\mathbf{k}_j\}$ is a set of wavevectors, $\{\phi_j^{[l]}\}$ is a set of wave amplitudes, and $\{\mathrm{i}p_j^{[l]}\}$ is a set of complex frequencies:

$$p_j^{[l]}(t) := \gamma_j^{[l]}(t) - \mathrm{i}\omega_j^{[l]}(t), \tag{3.19}$$

where $\{\gamma_j^{[l]}(t)\}$ is a set of nonlinear growth rates, and $\{\omega_j^{[l]}(t)\}$ is a set of constituent frequencies. As each wavepacket is dispersionless, the phase velocity of each wavepacket $\mathbf{u}^{[l]}$ is given by:

$$\omega_j^{[l]}(t) =: \mathbf{u}^{[l]}(t) \cdot \mathbf{k}_j^{[l]}. \tag{3.20}$$

Using this solution, one can define the following equation:

$$\boxed{(\Gamma_j^{[l]})^2 := \frac{\mathrm{d}p_j^{[l]}}{\mathrm{d}t} + (p_j^{[l]})^2,} \tag{3.21}$$

where $\{\Gamma_j^{[l]}\}$ is a set of constants, corresponding to rates of change that determine the dynamics of the system. One can show that by examining (3.16):

$$(\Gamma_j^{[l]})^2 = -\frac{1}{\mu_0\epsilon_0}\left\{ k_j^2 + \frac{\mathbf{E}_j^{[l]}}{|\mathbf{E}_j^{[l]}|^2} \cdot \sum_{\text{species}} q \int_{\mathbb{R}^3} \left(\mu_0 \mathbf{v} p_j^{[l]} + \frac{\mathrm{i}\mathbf{k}_j}{\epsilon_0} \right) f_j^{[l]}\,\mathrm{d}\mathbf{v} \right\}, \tag{3.22a}$$

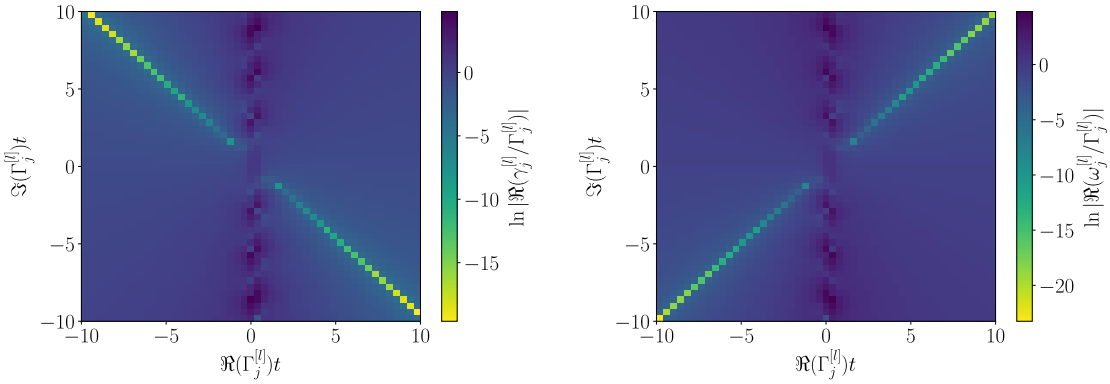
$$= -\frac{1}{\mu_0\epsilon_0}\left\{ k_j^2 - \frac{\mathbf{B}_j^{[l]}}{|\mathbf{B}_j^{[l]}|^2} \cdot \sum_{\text{species}} q \int_{\mathbb{R}^3} \mu_0(\mathrm{i}\mathbf{k}_j \times \mathbf{v} f_j^{[l]})\,\mathrm{d}\mathbf{v} \right\}. \tag{3.22b}$$

As there are multiple frequencies, it is instructive to note that the complex frequency of the l^{th} mode is made up of contributions from all of the branches.

Theorem 3. *There is no solution where* $\lim_{t\to\infty} \gamma_j^{[l]} \leq 0$ *except for the case where* $\gamma_j^{[l]}$ *is constant.*

Proof. Equation (3.21) yields:

$$p_j^{[l]}(t) = \Gamma_j^{[l]} \tanh\left[\varphi_j^{[l]} + \Gamma_j^{[l]} t \right], \tag{3.23}$$



(a) The growth rate is singular at $\Re(\Gamma_j^{[l]}) = 0$ for $t \neq 0$. The condition $\Re(\Gamma_j^{[l]}) = -\Im(\Gamma_j^{[l]})$ leads to $\gamma_j^{[l]} = 0$. The magnitude of the growth rate asymptotically approaches $\Re(\Gamma_j^{[l]})$.

(b) The frequency is singular at $\Re(\Gamma_j^{[l]}) = 0$ for $t \neq 0$. The condition $\Re(\Gamma_j^{[l]}) = \Im(\Gamma_j^{[l]})$ leads to $\omega_j^{[l]} = 0$. The magnitude of the frequency asymptotically approaches $\Re(\Gamma_j^{[l]})$.

Figure 3.5: **Complex frequency evolution.** Logarithmic plots of the angular frequency $\omega_j^{[l]}$ and growth rate $\gamma_j^{[l]}$ for waves in a single species system with $\omega(t=0) = \gamma(t=0) = 0$. The complex constant $\Gamma_j^{[l]}$ is related to the initial system conditions for a wave with wavevector $\mathbf{k}_j$. The x-axis features the real part of $\Gamma_j^{[l]}$, while the y-axis features the imaginary part of $\Gamma_j^{[l]}$. Evolution of the system can be traced by following a radial path from the centre of the plots. Where $|\Re(\Gamma_j^{[l]})t| \leq 1$, one observes undulating behaviour in the frequency and growth rate. The amplitude of the undulation decreases exponentially with $t/|\Gamma_j^{[l]}|$, the length of a radial trajectory from the temporal origin at $(0, 0)$.

where the complex hyperbolic angle $\varphi_j^{[l]} \equiv \operatorname{arctanh}(p_j^{[l]}(0)/\Gamma_{sl})$. I show plots of $p_j^{[l]}$ in Figure 3.5. If $\Gamma_j^{[l]}$ is equal to the initial complex frequency, then $p_j^{[l]}$ does not change value. For the case where $\Gamma_j^{[l]} = 0$, and $p_j^{[l]} \neq 0$, there is no stable solution. For all other cases:

$$\forall\, \Gamma_j^{[l]} \notin \{0, p_j^{[l]}\} \,:\, \lim_{t\to\infty} p_j^{[l]}(t) = |\Re(\Gamma_j^{[l]})| \pm \mathrm{i}\Im(\Gamma_j^{[l]}), \tag{3.24}$$

where $\Re(\Gamma_j^{[l]})$ denotes the real part of $\Gamma_j^{[l]}$, $\Im(\Gamma_j^{[l]})$ denotes the imaginary part of $\Gamma_j^{[l]}$, and $\pm$ corresponds to the sign of $\Re(\Gamma_j^{[l]})$. Therefore, no solution exists where $\lim_{t\to\infty} \gamma_j^{[l]} \leq 0$ except for the case where $\gamma_j^{[l]}$ is temporally static. □

By integrating equation (3.23), one can rewrite equation (3.18) as the following:

$$\mathbf{E}^{[l]}(\mathbf{x}, t) = \frac{1}{2} \sum_j \left\{ \mathbf{E}_j^{[l]} \mathrm{e}^{\mathrm{i}\mathbf{k}_j \cdot \mathbf{x}} \cosh\left[\varphi_j^{[l]} + \Gamma_j^{[l]} t\right] + \text{c.c.} \right\}. \tag{3.25}$$

One should note that for the case where $\varphi \to \infty$, $p_j^{[l]}(t) \to p_j^{[l]}(0)$ and therefore:

$$\lim_{\varphi_j^{[l]}\to\infty} \cosh\left[\varphi_j^{[l]} + \Gamma_j^{[l]} t\right] =: e^{p_j^{[l]}(0)t}.$$

I also impose the following physical limit:

$$\exists\,\{\Gamma_j^{[l]}\} : \lim_{t\to\infty} |\mathbf{E}| < \infty. \tag{3.26}$$

However, finding solutions that satisfy this limit is not trivial. One can decompose equation (3.21) into two real equations:

$$\frac{\mathrm{d}\gamma_j^{[l]}}{\mathrm{d}t} = -[(\gamma_j^{[l]})^2 - (\omega_j^{[l]})^2] + \Re[(\Gamma_j^{[l]})^2], \tag{3.27a}$$

$$\frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t} = -2\gamma_j^{[l]}\omega_j^{[l]} - \Im[(\Gamma_j^{[l]})^2]. \tag{3.27b}$$

These equations can be linearized from a time $t = 0$ where $\mathrm{d}\gamma_j^{[l]}/\,\mathrm{d}t = 0$. From equation (3.27a):

$$(\omega_j^{[l]})^2 \approx \omega_{\mathrm{pl.}}^2 + \text{const.},$$

where all terms $\mathcal{O}[(f-F)^2]$ have been discarded. As such, one can see that at $t = 0$, for $\gamma_j^{[l]} \ll \omega_{\mathrm{pl.}}$ the frequency is approximately equal to the plasma frequency.

3.3.2 Energy balance

By taking the second moment of the Boltzmann equation ((1.17a)), the total energy density T of the system is given by:

$$T(x,t) = \frac{\epsilon_0}{2}|\mathbf{E}|^2 + \frac{1}{2}\int_{\mathbb{R}^3} \sum_l f_{[l]} m_{[l]} |\mathbf{v}|^2 \,\mathrm{d}\mathbf{v}. \tag{3.28}$$

Use of equation (1.17e) allows one to find an energy conservation equation:

$$\left(\frac{\partial T}{\partial t}\right)_{\mathrm{C,T}} = \int_{\mathbb{R}^3} \left[\frac{1}{2} m |\mathbf{v}|^2 \frac{\partial f}{\partial t}\right] \mathrm{d}\mathbf{v} + \frac{1}{2}\frac{\partial}{\partial t}\left[\epsilon_0 |\mathbf{E}|^2 + \frac{1}{\mu_0}|\mathbf{B}|^2\right], \tag{3.29}$$

where the first term on the right hand side denotes the rate of change of energy density in the PDF, the second term denotes the rate of change of electromagnetic field energy density, and the third term denotes sources and sinks of energy density. $(\partial T/\partial t)_{\mathrm{C,T}}$ allows for sources and sinks of energy density:

$$\left(\frac{\partial T}{\partial t}\right)_{\mathrm{C,T}} \equiv \frac{1}{2} m \int_{-\infty}^{\infty} |\mathbf{v}|^2 (\hat{C}' f)\,\mathrm{d}\mathbf{v}, \tag{3.30}$$

where $\hat{C}' := \hat{C} - \mathbf{v} \cdot \nabla$ encapsulates collisions as well as spatial advection.

3.3.3 Berk-Breizman collision operator

Berk and Breizman[8] use an energy sink which dissipates plasma waves at a rate γ_D. This allows for a region of nonlinear stability for the system in the presence of instabilities. Here, I instead derive an equivalent formulation of the problem by using a collision operator.

One can represent this sink as the following:

$$\left(\frac{\partial T}{\partial t}\right)_{\text{coll.}} \equiv -\gamma_D \frac{1}{2}\left[\epsilon_0 |\mathbf{E}|^2 + \frac{1}{\mu_0}|\mathbf{B}|^2\right]. \tag{3.31}$$

Therefore, by using equations (3.29) to (3.31) one finds that for a vanishing integrand under v:

$$\frac{\partial}{\partial t}(\hat{C}' f) = -\gamma_D \left[(\hat{C}' f) - \frac{\partial f}{\partial t}\right]. \tag{3.32}$$

This is an inhomogeneous decay equation of the form $y(t)' = -ay + b(t)$. The solution, given by integrating factor, yields a collision operator that I will herein refer to as the Berk-Breizman operator:

$$\boxed{\hat{C}' f = \mathrm{e}^{-\gamma_D t}\left[\left.\frac{\partial f}{\partial t}\right|_{t=0} + \gamma_D \int_0^t \mathrm{e}^{\gamma_D \tau}\frac{\partial f}{\partial \tau}\,\mathrm{d}\tau\right].} \tag{3.33}$$

If one assumes the contribution from heat diffusion to be negligible in comparison to the contribution from collisions and other sources and sinks, then a Maclaurin expansion in γ_D yields:

$$\hat{C}' f \approx \left.\frac{\partial f}{\partial t}\right|_{t=0} \mathrm{e}^{-\gamma_D t} + \left[\gamma_D (f - F) + \mathcal{O}(\gamma_D^2)\right]. \tag{3.34}$$

As such, for $\gamma_D t << 1$, the first term in (3.33) dominates. At later times, a Krook-like collision operator safely approximates the behaviour to first order in γ_D.

3.3.4 Linear complex dispersion relation

The linear complex dispersion relation of waves can be analysed within a nonlinear framework. Here, I spectrally decompose the PDF and other spatially dependent quantities as described in Section 3.3.1, and analyse the nonlinear behaviour under pinning Berk-Breizman collisions in an equivalent linear framework (by approximating the collisions as Krook collisions).

Upon linearizing equations (1.17a) and (1.17e), the term proportional to $\partial F/\partial t$ produces a pole at $p_j = -\gamma_D$. Elsewhere, one obtains the following linear complex dispersion relation for a single mode, under the Krook-like approximation given by (3.34) to first order in γ_D:

$$p_j|_{t=0} \approx \sum_{\text{species}} \frac{q^2}{m\epsilon_0} \int_{\mathbb{R}^3} (\mathbf{e}\cdot\mathbf{v}) \left\{ \frac{\mathbf{e}\cdot\nabla_{\mathbf{v}} f_{(0)}}{p_j + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v} - \gamma_D} \right\} \mathrm{d}\mathbf{v}. \tag{3.35}$$

Overall, this approximately yields three branches:

$$\omega_j|_{t=0} \approx \{0, \pm\omega_{\text{pl.}}\} \quad ; \quad \gamma_j|_{t=0} \approx \left\{ -\gamma_D, \frac{\gamma_D}{2} + \sum_{\text{species}} \frac{\pi\omega_j^3}{2k_{j,(p)}^2} \int_{\mathbb{R}^2} \frac{\partial f}{\partial v_{(p)}}\bigg|_{v_{(p)}=\omega_j/k_{j,(p)}} \mathrm{d}A \right\},$$

where $\mathbf{E}$ is normal to $\mathrm{d}A$. Commonly in the literature, the 'modified' version of the Maxwell-Ampère law is used instead[26,29] to generate the same energy balance equation as equation (3.29):

$$\frac{\partial}{\partial t}\frac{\partial \phi_{\text{alt.}}}{\partial x} = \frac{q}{\epsilon_0} \int_{\mathbb{R}} v f \,\mathrm{d}v - \gamma_D \frac{\partial \phi_{\text{alt.}}}{\partial x}, \tag{3.36}$$

where $\phi_{\text{alt.}}$ is a modified electric potential. For these models to be consistent, Maxwell's equations must retain the same canonical form; previous work shows that transformation of the system Lagrangian that preserves the system Hamiltonian is indeed possible under certain conditions. However, for all counterexamples, these models do not form a fully consistent set of equations (see Appendix C.1). Using this alternative model, one obtains a very similar linear complex dispersion relation:

$$p_j|_{t=0} + \gamma_D \approx \sum_{\text{species}} \frac{q^2}{m\epsilon_0} \int_{\mathbb{R}^3} (\mathbf{e}\cdot\mathbf{v}) \left\{ \frac{\mathbf{e}\cdot\nabla_{\mathbf{v}} f_{(0)}}{p_j|_{t=0} + \mathrm{i}\mathbf{k}_j\cdot\mathbf{v}} \right\} \mathrm{d}\mathbf{v}. \tag{3.37}$$

This means that while the model is inconsistent (energy is typically not conserved), it produces similar dynamics to a model with Krook-like collisions; the primary difference is that the location of the Landau resonance is shifted from $p = \gamma_D - \mathrm{i}k_j v$ to $p = -\mathrm{i}k_j v$. Overall, this approximately yields the two branches:

$$\omega_j|_{t=0} \approx \pm\omega_{\text{pl.}} \quad ; \quad \gamma_j|_{t=0} \approx -\frac{\gamma_D}{2} + \sum_{\text{species}} \frac{\pi\omega_j^3}{2k_{j,(p)}^2} \int_{\mathbb{R}^2} \frac{\partial f}{\partial v_{(p)}}\bigg|_{v_{(p)}=\omega_j/k_{j,(p)}} \mathrm{d}A.$$

In the nonlinear phase of the systems evolution, it is a well known result that the original resonance undergoes broadening. Therefore, provided that γ_D/k_j is much smaller

than the resonance width, these models perform well except for energy conservation; as I have already utilised a small γ_D approximation to assume Krook-like dissipation, one is safe to make this assumption.

3.4 Basis decomposition formulation

In this section, I discuss how one can construct nonlinear solutions for the Vlasov-Maxwell system (1.17) by considering non-conservative orbits, and a basis decomposition of the orbit affine parameter $\epsilon(\mathbf{x}, \mathbf{v}, t)$ along particle orbits.

In Section 3.4.1, I show that as a proof of concept, one can easily generate van Kampen modes by using a basis decomposition.

Then, in Section 3.4.2, I show a general nonlinear basis decomposition which allows for wave-wave coupling, and discuss how it may not be possible to generate these solutions from an action-angle formalism.

In Section 3.4.3, I employ a basis of Gaussians, and give approximate solutions for the evolution of the growth rate and frequency of waves in the system provided that the corresponding phase space islands are far apart.

3.4.1 van Kampen modes

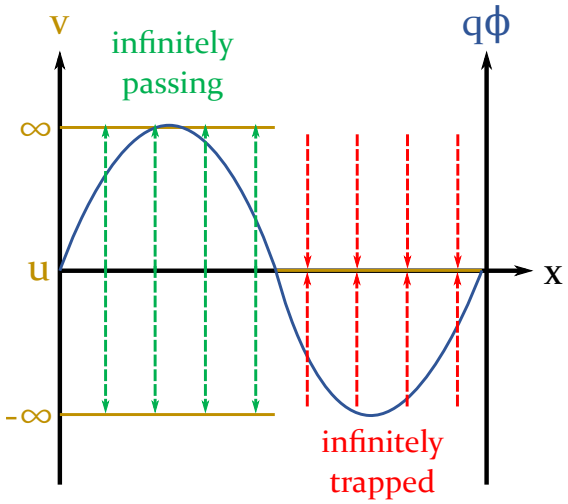


Figure 3.6: **van Kampen modes.** The sketch illustrates a single van Kampen mode with phase velocity 0. In the region where $q\phi > 0$, particles with $v = u$ are accelerated to $\pm\infty$ ('infinitely passing'), depending on the sign of u. In the region where $q\phi < 0$, particles with $v = u$ are infinitely trapped.

For van Kampen modes, particles with $v = u$ resonate, and particles elsewhere do not

resonate.

Lemma 3.2. *For $f = f[\epsilon]$, van Kampen modes are generated by the generating function given by:*

$$\psi(\mathbf{x}, \mathbf{v}, t) = \mathbf{v}_n \sum_l \phi^{[l]} \delta[\mathbf{v} - \mathbf{u}^{[l]}],$$

where $\mathbf{v}_n$ is a constant velocity that preserves dimensionality, and $\phi(\mathbf{x}, t)$ is given by:

$$\phi(\mathbf{x}, t) = \sum_l \phi^{[l]}(\mathbf{x}, t),$$

with each wavepacket $\phi^{[l]}$ containing a single constant phase velocity, $\mathbf{u}^{[l]}$.

Proof. van Kampen modes have fixed amplitude and fixed frequency, and are derived for collisionless plasmas. Therefore, $\hat{C} = 0$, $\gamma_j^{[l]} = 0 \, \forall \, j, l$ and $\omega_j^{[l]} = \text{const.} \, \forall \, j, l$. Then:

$$q\hat{V}\psi = 0.$$

If one takes the 0^{th} velocity moment of the above equation:

$$q\mathbf{v}_n \sum_l \left[\nabla \cdot (\mathbf{u}^{[l]} \phi^{[l]}) + \frac{\partial \phi^{[l]}}{\partial t} \right] = 0.$$

This is satisfied if each wavepacket $\phi^{[l]}$ contains a single phase velocity, $u^{[l]}$ which is not a function of $\mathbf{x}$. □

While it is mathematically allowed, the singularities makes ϵ no longer a physical quantity; it is impossible for a particle to have 'infinite energy'. Furthermore, $f[\epsilon]$ would also be singular at these points. One can however view the effective potential $\phi + \psi$ as accelerating a population of particles at $\mathbf{v} = \mathbf{u}^{[l]}$ to $|\mathbf{v}| = \pm\infty$ for $\phi^{[l]} > 0$, and infinitely trapping the particles at $\mathbf{v} = \mathbf{u}^{[l]}$ for $\phi^{[l]} < 0$ (see Figure 3.6). The particles which are accelerated to $|\mathbf{v}| = \pm\infty$ do not diminish the population as the PDF is singular at $\mathbf{v} = \mathbf{u}^{[l]}$, and the particles which are moved do not contribute to the electric potential.

The combined effective potential is dispersive, allowing for multiple resonances to exist.

3.4.2 Nonlinear basis decomposition

Suppose now one seeks general waves of varying complex frequency $p_{sl}(t) := \gamma_{sl}(t) - \mathrm{i}\omega_{sl}(t)$. I will now examine expansion using functions $g^{[l]}(\mathbf{x}, \mathbf{v} - \mathbf{u}^{[l]}(t), t)$ instead of Dirac delta functions, with the aim of examining some form of 'broadened' van Kampen modes. One finds that the derivation manifests as simply just a basis expansion of the problem.

I use a basis decomposition in $\mathbf{v}$ of ϵ:

$$\epsilon(\mathbf{x},\mathbf{v},t) := \sum_l c^{[l]}(\mathbf{x},t)g^{[l]}(\mathbf{x},\mathbf{v},t), \tag{3.38}$$

where the coefficients $c^{[l]}$ are position and time varying, and the basis functions in velocity $g^{[l]}$ are also position and time varying. This form of the basis decomposition is key, and corresponds to a different basis decomposition at every point in $\mathbf{x}$ and t. By allowing each $(\mathbf{x},t)$ to permit a different basis decomposition, one allows for nonlinear coupling which is typically not allowed using a variables separable method of the form $f = (XT)(\mathbf{x},t)V(\mathbf{v})$.

I desire that $\{g^{[l]}(\mathbf{x},\mathbf{v},t)\}$ are square-integrable functions on $v \in \mathbb{R}$, and to do so one can make the constraint that $\{g^{[l]}(\mathbf{x},\mathbf{v},t)\}$ vanish at infinity:

$$\lim_{v\to\pm\infty} g^{[l]}(\mathbf{x},\mathbf{v},t) = 0 \,\forall\, l. \tag{3.39}$$

Then, if one takes the 0^{th} moment of the Vlasov equation:

$$\int_{\mathbb{R}^3} \left[\nabla\cdot(\mathbf{v}\epsilon) + \frac{\partial\epsilon}{\partial t}\right] \mathrm{d}^3\mathbf{v} = 0.$$

Using the aforementioned basis expansion:

$$\sum_l \left[\nabla\cdot(c_l^{[l]}\mathbf{M}_1[g^{[l]}]) + \frac{\partial}{\partial t}(c_l^{[l]}M_0[g^{[l]}])\right] = 0, \tag{3.40}$$

where $M_0[f] := n(x,t)$ and $\mathbf{M}_1[f] \equiv \dfrac{1}{q}J(\mathbf{x},t)$ are the zeroth and first moments of f. Next, one can perform a co-moving Galilean transform such that $g^{[l]}(\mathbf{x},\mathbf{v},t) = g^{[l]}(\mathbf{x},\mathbf{v}-\mathbf{u}^{[l]}(t),t)$. Then:

$$\begin{aligned}\int_{\mathbb{R}^3} \mathbf{v}g^{[l]}\,\mathrm{d}^3\mathbf{v} &\approx \int_{\mathbb{R}^3} \mathbf{v}g^{[l]}\,\mathrm{d}^3(\mathbf{v}-\mathbf{u}^{[l]}) \\ &= \mathbf{u}^{[l]}(t)\int_{\mathbb{R}^3} g^{[l]}\,\mathrm{d}^3(\mathbf{v}-\mathbf{u}^{[l]}) + \int_{\mathbb{R}^3}(\mathbf{v}-\mathbf{u}^{[l]})g^{[l]}\,\mathrm{d}^3(\mathbf{v}-\mathbf{u}^{[l]}),\end{aligned}$$

where I have neglected a flux term, under the approximation that:

$$\int_{\mathbb{R}^3} \mathbf{v}g^{[l]}\,\mathrm{d}^3\mathbf{u}^{[l]} \equiv \int_0^t\int_{\mathbb{R}^2} \mathbf{v}g^{[l]}\frac{\mathrm{d}\mathbf{u}^{[l]}}{\mathrm{d}t}\,\mathrm{d}A(\mathbf{v})\,\mathrm{d}\tau \ll \int_{\mathbb{R}^3} \mathbf{v}g^{[l]}\,\mathrm{d}^3\mathbf{v},$$

where $\mathrm{d}\mathbf{A}$ is a velocity surface element perpendicular to $\mathbf{u}^{[l]}$. This is justified if the surface flow of $g^{[l]}$ perpendicular to the phase velocity $\mathbf{u}^{[l]}$ is negligible, or if the sweep rate

$\partial\omega_j^{[l]}/\partial t$ is small. If $g^{[l]}$ is symmetric about $\mathbf{v} = \mathbf{u}^{[l]}(t)$, then the second integral vanishes due to antisymmetry. Then:

$$\mathbf{M}_1[g^{[l]}] = \mathbf{u}^{[l]}(t) M_0[g^{[l]}]. \tag{3.41}$$

While it appears that this is a restriction on the solutions that are allowed, in fact any $\epsilon(\mathbf{x}, \mathbf{v}, t)$ can be represented via such a decomposition. The form of the basis functions has been restricted, but the function it represents has not (c.f. Fourier expansion with phase angles versus Fourier expansion with sines and cosines only).

Therefore from equation (3.40), one finds:

$$\sum_l \left[\mathbf{u}^{[l]} \cdot \nabla (c_l^{[l]} M_0[g^{[l]}]) + \frac{\partial}{\partial t} (c_l^{[l]} M_0[g^{[l]}]) \right] = 0. \tag{3.42}$$

If satisfied for each l independently, this takes the form of a continuity equation for each wave frame. It therefore becomes useful to exploit the conserved quantity, and via Noether's theorem[1] utilise the corresponding continuous symmetry. As such, one can define an integral mapping $(\mathbf{x}, t) \mapsto \chi$, where χ is a matrix of coordinates in instaneous co-moving frames:

$$\chi_j^{[l]} = \mathbf{k}_j \cdot \mathbf{x} - \int_0^t \omega_j^{[l]}(\tau)\, \mathrm{d}\tau, \tag{3.43}$$

representing co-moving Galilean transforms to the co-moving frame with time-dependent wave velocity $\mathbf{u}^{[l]}$ and growth rate $\gamma_j^{[l]}$. In this sense, χ contains all of the transformations possible with the available frequencies $\{\omega_j^{[l]}\}$ and available wavevectors $\{\mathbf{k}_j\}$ in the system.

The basis decomposition is therefore in such a form that the symmetry point in v of the even functions $\{g^{[l]}\}$ is the time-dependent wave speed of a wave with frequency $\omega_j^{[l]}$ and wavevector $\mathbf{k}_j$. Therefore, one can intuitively posit that such a decomposition yields phase space structures which move with the frequency of some form of wave in the system. If these waves are chosen to be electrostatic waves in the system, then one expects that the migration of phase space structures corresponds to nonlinear frequency chirping moving the wave-particle resonances.

Then, from equation (3.42):

$$\sum_{j,l} \left[\frac{\partial}{\partial t} \cdot (c^{[l]} M_0[g^{[l]}]) \right]_{\chi, v} = 0. \tag{3.44}$$

This forms a non-holonomic constraint on the system. This result is particularly interesting as it implies that it is not possible for one to set up a suitable Lagrangian for

the system with $c^{[l]}M_0[g^{[l]}]$ as generalised fields. It is therefore my belief that these solutions cannot be derived from an action-angle formalism. In such a formulation, Hamilton's principle assumes only holonomic constraints on the system. In contrast, the formalism yields an analog of d'Alembert's principle, allowing for more exotic scenarios involving non-conservative forces.

3.4.3 Gaussian expansion

Hermite functions form an orthonormal basis, and have recently been explored in the literature by J. M. Heninger *et al.* for G-transform based solutions of the linearized Vlasov-Poisson system.[47] Gaussian functions have a similar form to Hermite functions, but are easier to work with due to the lack of preceding Hermite polynomials in the function definition.

In this subsection, I will examine a one dimensional, collisionless system with no **B**-field. I will examine solutions of the form $f = f(x,v,t) + \Delta f[\epsilon]$, under the following Gaussian expansion:

$$\epsilon(x,v,t) = \sum_l W^{[l]}(x,t) \exp\left[-\left(\frac{v-u^{[l]}(t)}{v_N^{[l]}(\chi,v,t)}\right)^2\right], \tag{3.45}$$

where $v_N^{[l]}(\chi,v,t)$ is a normalising function with units of velocity, and $W^{[l]}(\chi,t)$ is an enveloping function with units of energy. By inserting ϵ into the Vlasov equation, $v_N^{[l]}$ is an approximate solution of the following differential equation under x,v,t coordinates:

$$\begin{aligned} -2W^{[l]}\frac{(v-u^{[l]})^2}{(v_N^{[l]})^3}\left[\frac{\partial}{\partial t} + v\frac{\partial}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial}{\partial v}\right] v_N^{[l]} \\ \approx 2(v-u^{[l]})\frac{W^{[l]}}{(v_N^{[l]})^2}\left[\frac{\mathrm{d}u^{[l]}}{\mathrm{d}t} + \frac{q}{m}\frac{\partial\phi}{\partial x}\right] + \left[\frac{\partial}{\partial t} + v\frac{\partial}{\partial x}\right] W^{[l]}. \end{aligned} \tag{3.46}$$

For the case of a single frequency wave in the system, (3.46) is exact. One must take the approximation that the overlap integral of two non-identical Gaussians is approximately zero. As such, for resonant structures which are close to each other in phase space, the validity of (3.46) breaks down.

It is possible to expand the left hand side of equation (3.46) using a Laurent series:

$$\begin{aligned}&\sum_{j=-\infty}^{\infty} L_j^{[l]}(x,t)(v-u^{[l]})^j \\ &:= -2W^{[l]}\frac{(v-u^{[l]})^2}{(v_N^{[l]})^3}\left[\frac{\partial}{\partial t}+v\frac{\partial}{\partial x}-\frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial}{\partial v}\right]v_N^{[l]},\end{aligned} \tag{3.47}$$

where $\{L_j^{[l]}\}$ are Laurent coefficients. Accordingly, one finds that for the analytic part of equation (3.46), in the limit that $v \to u^{[l]}$:

$$L_0^{[l]}(x,t) \approx \left[\frac{\partial}{\partial t}+u^{[l]}\frac{\partial}{\partial x}\right]W^{[l]}. \tag{3.48}$$

From this point onward, I will seek solutions where $L_0^{[l]}$ is the only term in the Laurent expansion.

Lemma 3.3. *If the Laurent series expansion of the left hand side of equation* (3.46) *only has one term,* $L_0^{[l]}(v-u^{[l]})^0$*, then, equation* (3.46) *is approximately solved with:*

$$(v_N^{[l]})^2 \approx -\frac{2W^{[l]}}{\partial_x W^{[l]}}\left[\frac{\mathrm{d}u^{[l]}}{\mathrm{d}t}+\frac{q}{m}\frac{\partial\phi}{\partial x}\right].$$

Proof. Using the expansion given by equation (3.47), equation (3.46) takes the form:

$$L_l^{(0)}(x,t) \approx 2(v-u^{[l]})\frac{W^{[l]}}{(v_N^{[l]})^2}\left[\frac{\mathrm{d}u^{[l]}}{\mathrm{d}t}+\frac{q}{m}\frac{\partial\phi}{\partial x}\right]+\left[\frac{\partial}{\partial t}+v\frac{\partial}{\partial x}\right]W^{[l]}.$$

By substituting equation (3.48) the lemma is proved if the trivial solution of $(v-u^{[l]}) \neq 0$ is ignored. □

Example solutions

Here, I will examine the possible solution:

$$W_l(x,t) = -q[\phi_l(x,t)+\tilde{W}_l(t)], \tag{3.49}$$

where $\phi = \sum_l \phi_l(x,t)$. If each ϕ_l has the form (from equation (3.18)):

$$\phi_l(x,t) \equiv |\phi_l^{[l]}|\exp\left[\int_0^t \gamma_l^{[l]}\,\mathrm{d}\tau\right]\cos\left[\chi_l^{[l]}+\theta^{[l]}\right], \tag{3.50}$$

where $\theta^{[l]}$ is some initial phase, then from Lemma 3.3:

$$(v_N^{[l]})^2 \approx \frac{2W_l \left[\dfrac{\mathrm{d}u^{[l]}}{\mathrm{d}t} + \dfrac{q}{m}\dfrac{\partial\phi}{\partial x}\right]}{k_l q \phi_l^{[l]} \exp\left[\displaystyle\int_0^t \gamma_l^{[l]}\,\mathrm{d}\tau\right] \sin\left[\chi_l^{[l]} + \theta^{[l]}\right]}. \tag{3.51}$$

I desire $\Delta\epsilon$ to be smooth, finite, and real-valued everywhere, which in turn requires $0 \leq v_{N,l} < \infty$. To enforce $|v_{N,j}| < \infty$ everywhere:

$$\left\{W_l \left[\frac{\mathrm{d}u_l}{\mathrm{d}t} + \frac{q}{m}\frac{\partial\phi}{\partial x}\right]\right\}_{\chi_l^{[l]} = -\theta^{[l]} + n\pi} = 0, \tag{3.52}$$

where n is an integer. This can be satisfied in two ways. Equation (3.52) can be satisfied if:

$$\left.\begin{aligned} \tilde{W}_l &= -|\phi_l^{[l]}| \exp\left[\int_0^t \gamma_l^{[l]}\,\mathrm{d}\tau\right] \\ 0 &= \frac{\mathrm{d}u^{[l]}}{\mathrm{d}t} + \frac{q}{m}\frac{\partial\phi}{\partial x}\bigg|_{\chi_l^{[l]} = -\theta^{[l]} + n\pi} \end{aligned}\right\} \quad \text{for } W_l \geq 0. \tag{3.53}$$

However, one cannot guarantee that this is true for each and every n. Therefore, we use the alternative solution:

$$\left.\begin{aligned} \tilde{W}_l &= +|\phi_l^{[l]}| \exp\left[\int_0^t \gamma_l^{[l]}\,\mathrm{d}\tau\right] \\ 0 &= \frac{\mathrm{d}u^{[l]}}{\mathrm{d}t} + \frac{q}{m}\frac{\partial\phi}{\partial x}\bigg|_{\chi_l^{[l]} = -\theta^{[l]}} \end{aligned}\right\} \text{for } W_l \leq 0, \tag{3.54}$$

such that $\tilde{W}_l$ is the amplitude of the wave with phase velocity $u^{[l]}$. Therefore, by examining the rate of change of $u^{[l]}$, one finds the equations:

$$\boxed{\frac{\mathrm{d}\omega_l^{[l]}}{\mathrm{d}t} \approx \sum_j \frac{q k_j k_l}{m} \tilde{W}_j(t) \sin\left[\left\{\int_0^t \left(\omega_j^{[j]} - \frac{k_j}{k_l}\omega_l^{[l]}\right)\mathrm{d}\tau - \left(\theta^{[j]} - \frac{k_j}{k_l}\theta^{[l]}\right)\right\}\right] \ \forall\, l.} \tag{3.55}$$

The system of equations (3.55) must be consistent with the solution for $\{p_{sl}(t)\}$ given by (3.19) determining the evolution of ϕ.

To enable positive $(v_N^{[l]})^2$ here is not trivial. By considering other terms in the Laurent series given by equation (3.47), one may be able to generate a form for $(v_N^{[l]})^2$ which permits more solutions.

Fixed frequency

For the case of a single wave in the system, equation (3.55) yields a fixed-frequency wave. This exact solution is consistent with equation (3.19) if the growth rate is also static:

$$\frac{\mathrm{d}p_l^{[l]}}{\mathrm{d}t} = 0.$$

Therefore, I find:

$$(v_N^{[l]})^2 \approx \frac{2q[\phi_l^{[l]}(x,t) + \tilde{W}_l(t)]}{m}.$$

This solution therefore gives to lowest order in $(v - u^{[l]})$:

$$\Delta\epsilon \approx \frac{1}{2}m(v - u^{[l]})^2 - q(\phi_l^{[l]} + \tilde{W}_l).$$

This gives an estimate for the separatrix width a factor of $\sqrt{2}$ higher, as the particle acts as though it is trapped in a potential well twice as large as the real potential.

In contrast, if one chose to use $W_l \leq 0$ instead, $(v_N^{[l]})^2$ would be negative, allowing for us to recover the BGK solution. In such a case, for the full Gaussian representation ϵ does not vanish as $v \to \infty$ (as is also the case for BGK modes), and moments of the ϵ are not real valued (as $v_N^{[l]}$ is imaginary).

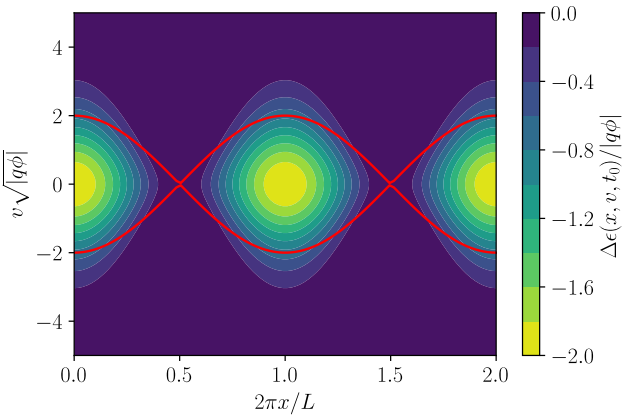


Figure 3.7: **BGK-like nonlinear mode**. Contours of constant $\epsilon(x, v, t)$ for a single electric potential wave with time-invariant frequency and constant amplitude undergoing a form of resonant interaction in a 1D kinetic system (derived in Section 3.4.3). Particles are approximately trapped within an island of width of $\sim \sqrt{4U/m}$ in velocity (a factor of $\sqrt{2}$ larger than that which is expected from BGK theory[7]) where $U = |q\phi_{\text{max.}}|$ is the electric potential energy, and m is the particle mass.

In Figure 3.7, I show the corresponding full solution for $\epsilon(x, v, t = 0)$ under this approximation, and how it deviates from BGK theory away from resonance. Here, I use normalised values of $|\phi_l^{[l]}| = q = m = 1$, and $k = 2\pi/L$.

Chirping frequency

For the case of two waves in the system with the same wavenumber:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\omega_+\\ \omega_-\end{pmatrix} \approx \frac{qk^2}{m}\sin\left[\left\{\int_0^t \tilde{\omega}\,\mathrm{d}\tau - \tilde{\theta}\right\}\right]\begin{pmatrix}\tilde{W}^{[-]}\\ -\tilde{W}^{[+]}\end{pmatrix}, \tag{3.56}$$

where $\omega_\pm := \omega_\pm^\pm$, $\tilde{\omega} := \omega_+ - \omega_-$, and $\tilde{\theta} \equiv \theta^{[+]} - \theta^{[-]}$. If one examines the case where:

$$\frac{\mathrm{d}\omega_+}{\mathrm{d}t} \approx -\frac{\mathrm{d}\omega_-}{\mathrm{d}t},$$

such that the frequency sweep is roughly symmetric, then $\tilde{W}^{[+]} \approx -\tilde{W}^{[-]}$. For either of the waves to be linearly stable, I require $\exists\, t_0 : \gamma_\pm(t = t_0) = 0$. However, the waves can still be nonlinearly unstable. To assess the nonlinear stability, one can combine equations (3.27a) and (3.27b):

$$\frac{\mathrm{d}\gamma_\pm}{\mathrm{d}t} = -\gamma_\pm^2 + \left(\frac{\dot{\omega} + \Im(\Gamma_\pm^2)}{2\gamma_\pm}\right)^2 + \Re(\Gamma_\pm^2), \tag{3.57}$$

where $\dot{\omega} = \mathrm{d}\tilde{\omega}/\,\mathrm{d}t$. For this to be nonlinearly stable, one requires that $\mathrm{d}\gamma_\pm/\,\mathrm{d}t < 0$. Therefore:

$$\lim_{t\to\infty}[\dot{\omega} + \Im(\Gamma_\pm^2)]^2 < \gamma_\pm^2[\gamma_\pm^2 - \Re(\Gamma_\pm^2)].$$

Additionally, for stability I require that $\lim_{t\to\infty}\dot{\omega} = 0$. Therefore, for nonlinear stability the following inequality must be satisfied:

$$\gamma_\pm^2 > \frac{1}{2}\left[\Re(\Gamma_\pm^2) + \sqrt{\Re(\Gamma_\pm^2)^2 + 4\Im(\Gamma_\pm^2)^2}\right], \tag{3.58}$$

where one can identify that $\Gamma_\pm^2 \equiv \lim_{t\to\infty} p_\pm^2$. In Figure 3.8, I sketch the regions defined by (3.58) with the corresponding sign of $\mathrm{d}\gamma_\pm/\,\mathrm{d}t$. Intersections between the line $\gamma_\pm$ and $\lim_{t\to\infty}\gamma_\pm$ show stable values for the growth rate, corresponding to nonlinearly stable (metastable limit with $\gamma_\pm < 0$) and nonlinearly unstable (stable limit with $\gamma_\pm > 0$) states.

Everywhere else above the line defined by the negative root of (3.58), the wave alternates between increasing and decreasing growth rates. As the choice of initial conditions

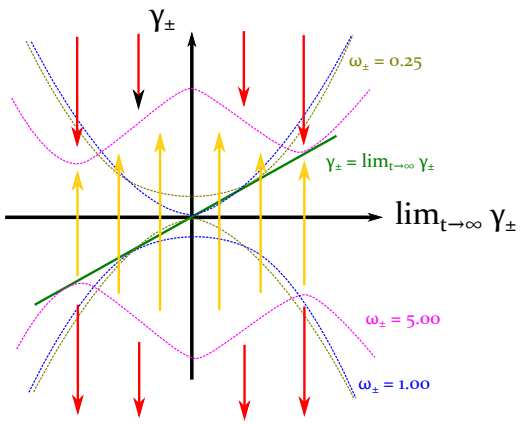


Figure 3.8: Sketches illustrating boundaries defined by (3.58), for $\omega_\pm \in \{0.25, 1, 5\}$. Yellow arrows and red arrows denote $\mathrm{d}\gamma_\pm/\,\mathrm{d}t > 0$ and $\mathrm{d}\gamma_\pm/\,\mathrm{d}t < 0$ respectively for $\omega_\pm = 5$. Points where the boundaries touch the line $\gamma_\pm = \lim_{t\to\infty}\gamma_\pm$ indicate stable or metastable values of $\gamma_\pm$ for $\gamma_\pm > 0$ and $\gamma_\pm < 0$ respectively. The stable value therefore corresponds to a nonlinearly unstable solution, while the metastable value corresponds to a nonlinearly stable solution.

influences $\Gamma_\pm$, some initial conditions will not have a limit for $\gamma_\pm$. This allows for either repeated chirping as observed in simulations and experiments,[29,49] or nonlinear instability (if $\int_{t_1}^{t_2}\gamma_\pm\,\mathrm{d}t > 0$ for $t_2 > t_1$). This may be a candidate for the rapid frequency chirping observed during abrupt large events or mode avalanching in tokamaks.[50–52]

Between $\gamma_\pm = 0$ and the line defined by the negative root of (3.58), the growth rate is negative, but decreasing in amplitude. Waves here are linearly stable, but may become nonlinearly unstable if they are able to cross the $\gamma_\pm = 0$ line with finite amplitude. Below the line defined by the negative root of (3.58), the growth rate is negative and always decreasing, corresponding to both linear and nonlinear stability.

As one expects $\lim_{t\to\infty}\omega_\pm = \Im(\Gamma_\pm)$, one finds that for the system to be nonlinearly stable, an appropriate temporal ordering is given by $\lim_{t\to\infty}|\dot{\omega}| \lessapprox 2\gamma\omega$. If the chirping rate is faster than this, the linearly stable state of the system can only exist for a finite amount of time. After this time, the system transitions to a nonlinearly unstable state.

From equation (3.56), one finds that the magnitude of the chirping rate is directly proportion to $\tilde{A}$. Therefore, if there is a large difference in the wave amplitudes, the system here is nonlinearly unstable.

It is worth noting that this solution is only approximately valid for structures which have a very small overlap integral in phase space. As such, one can consider these long-range, coupled structures in a approximate fashion by evolving a continuum of superposed,

fixed-frequency solutions.

Separatrix

I define the separatrix as the largest closed contour in phase space. Here, I examine only a single value of l active in the system, and examine $\Delta\epsilon(\chi, v, t) = c$, an unspecified constant. Points on such a contour only exist for:

$$0 \leq \frac{c}{W^{[l]}} < 1,$$

as $\exp(-z) \in (0, 1]$ for positive z. As such, if $c < W^{[l]}$, no point exists on the contour for the corresponding values of χ, t. Therefore, closed contours exist for $c > \min(W^{[l]})$, and the separatrix is given by $c = 0$. At this fixed value of $\Delta\epsilon(\chi, v, t)$:

$$v_{\text{sep.}} = u^{[l]} \pm \lim_{c\to 0} \sqrt{-\frac{2W^{[l]}}{m} \ln\left(\frac{W^{[l]}}{c}\right)}.$$

As a result, all the particles here are 'trapped'. However in reality, weakly bound particles would scatter out of the potential via neoclassical transport in a model that considers collisions.

In light of this, I use the following fit for an effective separatrix to compare to BGK theory:

$$v_{\text{sep.}} \approx u^{[l]} \pm \sqrt{-\frac{2W^{[l]}}{m}},$$

which is given when the structure is at 1/e height. In comparison to BGK theory, the value of $W^{[l]}$ has twice the amplitude of ϕ_l. Therefore, the width of the separatrix given by this theory is roughly a factor of $\sqrt{2}$ larger than that which is expected for a BGK island.

3.5 Generating function formulation

In contrast to the previous section which directly examines possible solutions for ϵ, in this section I detail solutions in terms of ψ. If one defines $f(x, v, t) = f(\epsilon)$, then, everywhere except on contours of constant ϵ:

$$\hat{C}\epsilon = \hat{V}\epsilon,$$

where ϵ is defined as in (3.1). If one substitutes ϵ into the Vlasov equation above:

$$\begin{aligned} q\hat{V}\psi &= \overbrace{\hat{C}\epsilon}^{\text{coll.}} + \overbrace{\sum_l \left[m(\mathbf{v}-\mathbf{u}^{[l]}) \cdot \left\{ \frac{\partial \mathbf{u}^{[l]}}{\partial t} + \mathbf{v}\cdot\nabla\mathbf{u}^{[l]} \right\} \right]}^{\text{frequency chirp}} \\ &\quad \underbrace{- q\sum_l \left[\frac{\partial \phi^{[l]}}{\partial t} + \mathbf{u}^{[l]}\cdot\nabla\phi^{[l]} \right]}_{\text{drive}} + \underbrace{\sum_l \left[\mathbf{u}^{[l]} \cdot (\mathbf{v} \times q\mathbf{B}) \right]}_{\text{gyration}}. \end{aligned} \tag{3.59}$$

In this sense, the generating function ψ provides a correction to ϵ. As contours of constant f (and therefore constant ϵ) denote particle orbits, systems with finite $\{\gamma_j^{[l]}\}$, finite $\{\partial_t \omega_j^{[l]}\}$, finite curvature of $\{\mathbf{k}_j\}$, or collisions require particles to take non-conservative orbits in phase space.

The 'frequency chirp' term and 'drive' term correspond to advection of ϵ_0. Advection of the kinetic energy of particles and the potential energy of particles perturbs the particle orbits (as expected from (3.14)).

Using the form of the electric potential given by (3.3), the electric potential is a solution of the coupled inhomogeneous advection equation:

$$\begin{aligned} \sum_l \left[\frac{\partial \phi^{[l]}}{\partial t} + \mathbf{u}^{[l]}\cdot\nabla\phi^{[l]} \right] &= \sum_{j,l} \Bigg\{ \left\{ \gamma_j^{[l]} - \mathbf{u}^{[l]} \cdot [\nabla(\mathbf{k}_j\cdot\mathbf{x}) - \mathbf{k}_j)] \right\} |\phi_j^{[l]}| \\ &\qquad \cdot \exp\left[\int_0^t \gamma_j^{[l]}\, \mathrm{d}\tau \right] \cos\left[\mathbf{k}_j\cdot\mathbf{x} - \int_0^t \omega_j^{[l]}\, \mathrm{d}\tau + \theta_j^{[l]} \right] \Bigg\}, \end{aligned} \tag{3.60}$$

where $\{\theta_j^{[l]}\}$ are phase angles. The term proportional to $\mathbf{u}^{[l]} \cdot [\nabla(\mathbf{k}_j\cdot\mathbf{x}) - \mathbf{k}_j]$ is a correction to the drive that is given by the curvature of the wavevector. If the wave is propagating in a rectilinear fashion, this term is identically zero.

In Section 3.5.1, I discuss an extension of the co-moving Galilean transform given in (3.43), yielding a velocity transformation matrix.

Then, in Section 3.5.2 and Section 3.5.3, I give functions which generate the 'frequency chirp' and 'growth' terms in (3.59).

Later, in Section 3.5.4, I give approximate solutions for the electrostatic Vlasov equation, under the limit that no wave-wave coupling occurs.

3.5.1 Velocity transformation matrix

Under the co-moving Galilean transform employed for χ (see (3.43)), one expects that the velocity is also transformed. Logically, one can also define a velocity transformation matrix

ν with the co-moving Galilean transform form:

$$\nu_j^{[l]} := \mathbf{k}_j \cdot \mathbf{v} - \omega_j^{[l]}. \tag{3.61}$$

Components of the matrix are zero at the point of Landau resonance, where $\omega_j^{[l]} = \mathbf{v}\cdot\mathbf{k}_j$. The derivatives transform covariantly under $(\mathbf{x}, \mathbf{v}, t) \mapsto (\chi, \nu, t)$. One must take special care as the direction of wave propagation can be curvilinear. Therefore:

$$\nabla = \sum_{j,l} \left\{\mathbf{k}_j - [\nabla(\mathbf{k}_j \cdot \mathbf{x}) - \mathbf{k}_j]\right\} \left.\frac{\partial}{\partial \chi_j^{[l]}}\right|_{\nu,t}, \tag{3.62a}$$

$$\nabla_{\mathbf{v}} = \sum_{j,l} \mathbf{k}_j \left.\frac{\partial}{\partial \nu_j^{[l]}}\right|_{\nu,t}, \tag{3.62b}$$

$$\left.\frac{\partial}{\partial t}\right|_{\mathbf{x},\mathbf{v}} = -\sum_{j,l} \left\{ \omega_j^{[l]} \left.\frac{\partial}{\partial \chi_j^{[l]}}\right|_{\nu,t} + \frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t} \left.\frac{\partial}{\partial \nu_j^{[l]}}\right|_{\chi,t} \right\} + \left.\frac{\partial}{\partial t}\right|_{\chi,\nu}. \tag{3.62c}$$

Under this transformation, one can split ψ into three parts: ψ_∂ encapsulating all of the curvilinear terms (given by $\sim \nabla(\mathbf{k}_j \cdot \mathbf{x}) - \mathbf{k}_j$), $\psi_{\mathbf{B}}$ encapsulating all of the magnetic field terms, and ψ_0 containing only the rectilinear terms and the electric field parts. Then, examining just the terms arising from $\hat{V}\psi_0$:

$$\begin{aligned} q\left[\left.\frac{\partial}{\partial t}\right|_{\chi,\nu} + \sum_{j,l}\left\{\nu_j^{[l]}\frac{\partial}{\partial \chi_j^{[l]}} - \frac{q}{m}\left[\frac{m}{q}\frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t} + |\mathbf{k}_j|^2 \sum_{l'}\frac{\partial\phi}{\partial\chi_j^{[l']}}\right]\frac{\partial}{\partial\nu_j^{[l]}}\right\}\right]\psi_0 \\ = [\hat{C}\epsilon]_0 + \sum_{j,l}\left[m\nu_j^{[l]}\frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t}\right] - q\sum_l \left.\frac{\partial\phi^{[l]}}{\partial t}\right|_{\chi,\nu}. \end{aligned} \tag{3.63}$$

This equation is very challenging to solve analytically. To enable solutions, I will solve parts of the equations separately.

3.5.2 Frequency chirp generating function

Here, I aim to solve the equation:

$$q\sum_{j,l}\left\{\nu_j^{[l]}\frac{\partial}{\partial\chi_j^{[l]}} - \frac{q}{m}|\mathbf{k}_j|^2\sum_{l'}\frac{\partial\phi}{\partial\chi_j^{[l']}}\frac{\partial}{\partial\nu_j^{[l]}}\right\}\psi_{\text{sw.}} = \sum_{j,l}\left[m\nu_j^{[l]}\frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t}\right]. \tag{3.64}$$

This allows one to generate the 'frequency chirp' term in equation (3.59). As the right hand side is independent of $\chi_j^{[l]}$, it is fairly straightforward to show that this has a solution

given by:

$$\psi_{\text{sw.}} = \sum_{j,l} \left[\frac{m}{q} \chi_j^{[l]} \frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t} \right]. \tag{3.65}$$

This solution is not periodic; I will address this later in Section 3.5.4.

3.5.3 Growth rate generating function

Here, I aim to solve the equation:

$$q \sum_{j,l} \left\{ \nu_j^{[l]} \frac{\partial}{\partial \chi_j^{[l]}} - \frac{q}{m} |\mathbf{k}_j|^2 \sum_{l'} \frac{\partial \phi}{\partial \chi_j^{[l']}} \frac{\partial}{\partial \nu_j^{[l]}} \right\} \psi_{\text{sw.}} = -q \sum_l \left. \frac{\partial \phi^{[l]}}{\partial t} \right|_{\chi,\nu}. \tag{3.66}$$

This allows one to generate the 'drive' term in equation (3.59). Here, I shall solve related equations with increasing complexity, until the full solution is found.

Single 1D wavepacket of constant frequency

To start with, one can attempt to solve a simplified version of (3.63):

$$\left[v \frac{\partial}{\partial x} - \frac{q}{m} \frac{\partial \phi}{\partial x} \frac{\partial}{\partial v} \right] \psi = -\gamma \phi, \tag{3.67}$$

where $\phi = \phi(x)$. This equation is equivalent to (3.63) when there is only a single 1D wavepacket of constant frequency.

It is worth noting that in reality, any wave with finite γ will have a finite $\partial\omega/\partial t$ by virtue of (3.21). Equation (3.67) has a solution (with derivation given in Appendix B.1.1) given by:

$$\psi = -\text{sign}(v) \gamma \sqrt{\frac{m}{2}} \int_{x_{(0)}}^{x} \frac{q\phi(x')}{\sqrt{\epsilon_{(0)}(x,v) - q\phi(x')}} \, \mathrm{d}x',$$

where $\epsilon_0 = q\phi + \frac{1}{2}mv^2$, and sign$(v)$ is the signum function. It is important to note that ψ is real and without singularities in the region:

$$q\phi(x) + \frac{1}{2}mv^2 > q|\phi|,$$

where $|\phi|$ is the amplitude of the wave. For values where $q\phi(x) + \frac{1}{2}mv^2 = q|\phi|$, ψ is singular. Furthermore, for $q\phi(x) + \frac{1}{2}mv^2 < q|\phi|$, ψ is imaginary. Therefore to enable real values of ϵ, one must find another solution for the regions where $q\phi(x) + \frac{1}{2}mv^2 \leq q|\phi|$.

In addition, the solution for ψ given above is not periodic. I will address this later in Section 3.5.4.

Single 3D wavepacket of constant frequency

Next, one can attempt to solve the following equation:

$$\left[\mathbf{v}\cdot\nabla - \frac{q}{m}(\nabla\phi)\cdot\nabla_{\mathbf{v}}\right]\psi = -\gamma\phi,$$

where $\phi = \phi(\mathbf{x})$. To do so, one requires an extension of the Leibniz integral rule to vector differential operators (see Appendix B.2). Using the vectorial form, one finds that the following integral is useful:

$$I_n := \frac{1}{v}\int_{\mathbf{x}_{(0)}}^{\mathbf{x}} \phi(\mathbf{x}')^n \mathbf{v}\cdot\mathrm{d}\mathbf{x}',$$

as it yields the following gradient:

$$\nabla I_n = \frac{\mathbf{v}}{|\mathbf{v}|}\phi(\mathbf{x})^n.$$

In Appendix B.1.2, I therefore show that the following solution for ψ is permitted:

$$\psi = -\gamma\sqrt{\frac{m}{2}}\int_{\mathbf{x}_{(0)}}^{\mathbf{x}} \frac{q\phi(\mathbf{x}')}{\sqrt{\epsilon_{(0)}(\mathbf{x},\mathbf{v}) - q\phi(\mathbf{x}')}}\frac{\mathbf{v}}{|\mathbf{v}|}\cdot\mathrm{d}\mathbf{x}', \tag{3.68}$$

where $\epsilon_0 = q\phi + \frac{1}{2}mv^2$. Similarly, this is only valid in the region where $q\phi(\mathbf{x}) + \frac{1}{2}mv^2 > q|\phi|$.

Multiple non-interacting wavepackets

The full equation is very difficult to solve. Instead, one can examine a simpler scenario by enforcing no wave-wave coupling here:

$$\left[\sum_{j,l}\left\{\nu_j^{[l]}\frac{\partial}{\partial\chi_j^{[l]}} - \frac{q}{m}|\mathbf{k}_j|^2\frac{\partial\tilde{\phi}_j^{[l]}}{\partial\chi_j^{[l]}}\frac{\partial}{\partial\nu_j^{[l]}}\right\}\right]\psi_\gamma \approx -\sum_{j,l}\gamma_j^{[l]}\tilde{\phi}_j^{[l]},$$

where $\tilde{\phi}_j^{[l]} = \frac{1}{2}\exp\left[\mathrm{i}\mathbf{k}_j\cdot\mathbf{x} - \int_0^t p_j^{[l]}\,\mathrm{d}\tau\right] + \text{c.c.}$, and all terms in the inner sum with $l \neq l'$ have been discarded. It is trivial to show that by extending (3.68) to an arbitrarily sized dimensional space and solving for each ψ_j independently, a solution is given by:

$$\boxed{\psi_\gamma = \psi_{\mathrm{ww}} - \sum_{j,l}\gamma_j^{[l]}\frac{\nu_j^{[l]}}{\sqrt{\sum_{j,l}(\nu_j^{[l]})^2}}\sqrt{\frac{m}{2}}\int_{\chi_{(0),j}^{[l]}}^{\chi_j^{[l]}}\frac{q\phi(\chi_j^{[l]\prime},\dots)}{\sqrt{\epsilon_{(0)}(\chi,\nu) - q\phi(\chi_j^{[l]\prime},\dots)}}\,\mathrm{d}\chi_j^{[l]\prime},} \tag{3.69}$$

such that one simply sums over all wavevectors and frequencies independently. ψ_{ww} encapsulates extra terms that would appear which are highly nonlinear, and feature wave-wave coupling. These interactions cannot be neglected for wave-wave interactions.

3.5.4 Collisionless system with no wave-wave coupling

For a collisionless system, one finds that:

$$\psi = \psi_{\gamma} + \psi_{\mathrm{sw.}} + \delta\psi, \tag{3.70}$$

where $\delta\psi$ is a highly nonlinear term that is given by all other contributions:

$$\delta\psi := \psi_{\partial} + \psi_{\mathbf{B}} + \psi_{\mathrm{ww}} + [\psi_0(\chi, \nu, t) - \psi_0(\chi, \nu, t = 0)]. \tag{3.71}$$

By operating on ψ, in a system with rectilinearly propagating waves one finds by inspection:

$$\hat{V}\delta\psi := \sum_{j,l}\left[\frac{m}{q}\chi_j^{[l]}\frac{\mathrm{d}^2\omega_j^{[l]}}{\mathrm{d}t^2}\right] - \sum_{j,l}\gamma_j^{[l]}\sqrt{\frac{m}{2}}\int_{\chi_{(0),j}^{[l]}}^{\chi_j^{[l]}}\frac{q\phi(\chi_j^{[l]\prime},\dots)}{2(\epsilon_{(0)}(\chi,\nu) - q\phi(\chi_j^{[l]\prime},\dots))^{3/2}}\hat{V}\left(\epsilon_{(0)}\frac{\nu_j^{[l]}}{\sqrt{\sum_{j,l}(\nu_j^{[l]})^2}}\right)\mathrm{d}\chi_j^{[l]\prime}.$$

As such, if the frequency chirp rate is static, there is a small $\mathbf{B}$-field and ϵ_0 is approximately static, then $\delta\psi$ can be considered to be mostly a small perturbation to the system; therefore:

$$\exists\,\mathbf{E},\mathbf{B}\ :\ \hat{V}\delta\psi \ll \hat{V}(\psi - \delta\psi). \tag{3.72}$$

However as expected from discussion at the end of Section 3.2.2, there is always a finite amount of wave-wave coupling in the system, as growth rate $\rightarrow$ modified orbit shape $\rightarrow$ shear $\rightarrow$ wave-wave coupling. Intrinsically, by discarding this term one assumes low shear of phase space islands, and negligible coupling between islands. One can assume that islands which are far apart do not interact with each other when considering gap toroidal Alfvén eigenmodes (TAEs) in tokamaks, however this approximation does not allow one to accurately examine the formation of H&Cs. For one to do so, one must retain all of the wave-wave coupling encapsulated by $\delta\psi$.

With regards to $(\psi - \delta\psi)$, in the limit of $\delta\psi \rightarrow 0$, one desires periodicity:

$$\lim_{\delta\psi\to 0}(\psi-\delta\psi)\ :\ \psi(\mathbf{x},\mathbf{v},t)\equiv\psi(\mathbf{x}+L_i\hat{\mathbf{x}}_i,\mathbf{v},t)\,\forall\, i,$$

where L_i is the box length in the x_i direction. To enforce this, one can demand that all terms proportional to $\chi_j^{[l]}$ must vanish. This requires:

$$\boxed{\sum_{j,l}\frac{m}{q}\frac{\mathrm{d}\omega_j^{[l]}}{\mathrm{d}t}\approx\sum_{j,l}\gamma_j^{[l]}\frac{\nu_j^{[l]}}{\sqrt{\sum_{j,l}(\nu_j^{[l]})^2}}\sqrt{\frac{m}{2}}\int_{\chi_{(0),j}^{[l]}}^{\chi_{(0),j}^{[l]}+2\pi}\frac{q\phi(\chi_j^{[l]\prime},\dots)}{\sqrt{\epsilon_0(\chi,\nu)-q\phi(\chi_j^{[l]\prime},\dots)}}\,\mathrm{d}\chi_j^{[l]\prime}\approx 0.}\tag{3.73}$$

This can be safely enforced for the case where the chirping rate is small, and where the wave amplitude is small. In such a scenario, one can approximate:

$$\psi\approx\psi_\gamma+\psi_{\text{sw.}}.$$

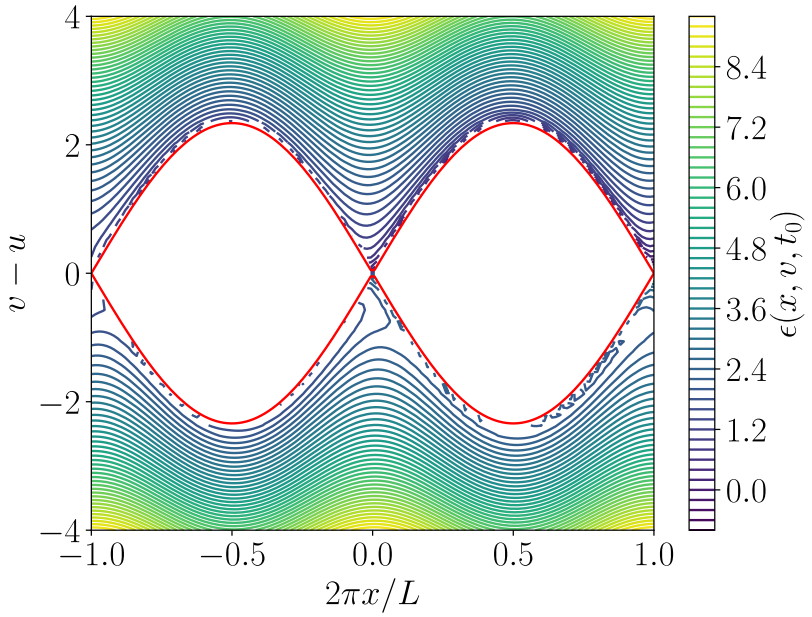


Figure 3.9: **Existence of quasi-passing particles**. Contours of constant $\epsilon(x,v,t)$ for a single 1D electrostatic wave with time-invariant frequency and growing amplitude undergoing a form of resonant interaction in a 1D kinetic system (derived in Section 3.5.4). Particles outside of the separatrix appear to exist in closed orbits; it is possible that these are quasi-passing particle particles as proposed in Section 3.5.4.

In Figure 3.9, I show an example case with a single electric potential wave with time-invariant frequency and growing amplitude. This plot was produced by numerically inte-

grating (3.68) using the trapezium rule, and plotting $\epsilon_{(0)}+q\psi$ for different values of (x, v, t). $x_{(0)}$ was taken to be 0. Close to the separatrix, some particles appear to be existing in closed orbits. It is worth noting that the condition as posed in (3.73) is violated here; as a result, there is a small amount of aperiodicity in the plots. In reality, this aperiodicity must be cancelled out by $\delta\psi$, as ψ must be periodic.

In addition, Figure 3.9 shows the existence of phase space shear. There is finite shear in the system arising from the finite growth rate in the system, as discussed in Chapter 2; here, I show that this effect occurs on an orbit level, as proposed in Section 3.2.2.

3.6 Brief conclusions

In this chapter, I showed completely original, novel work carried out by myself. In this chapter, I examined so-called 'non-conservative orbits' for a given plasma species, where the single particle energy is not conserved on a phase space orbit.

The overall aim was to consider the true orbit affine parameter, ϵ along particle orbits as the sum of a quantity representing the single particle energy, and a generating function ψ. As such, in the limit that the single particle energy is conserved on an orbit, ϵ naturally approaches the single particle energy, allowing the theory to directly extend BGK theory. This theory should allow one to recover analytic predictions for the form of phase space structures during events such as frequency bifurcation, mode growth, and frequency chirping.

Over the course of the chapter, I discussed a range of techniques, finishing with three solutions for the orbit affine parameter ϵ and the corresponding waves in the system given in (3.18). These allow one to describe frequency bifurcation and frequency chirping (shown in Section 3.4.3, as well as mode growth (shown in Section 3.5.4).

Chapter 4

DARK: D-dimensional Augmented Resonance Kinetic solver

4.1 Overview

In this chapter, I present D-dimensional Augmented Resonance Kinetic solver (DARK) a new, modular code based on previous work by Arber, Vann and De-Gol.[26,27] This code was written by myself in C99 during 2015 - 2018, and is a single framework which can incorporate fundamentally different models and approximations by using a Strang split[53,54] set of partial flows (simpler differential equations). Each partial flow is solved separately to yield the full solution across a timestep.

DARK makes use of core code which initialises fields such as the electromagnetic field and the particle distribution function (PDF), initialisation code which allows one to freely switch between different initial conditions, and modular code which allows one to freely switch between very different physical models. The modular code makes use of augmentation terms which are fields which are not found in the Boltzmann-Maxwell system, allowing one to explore more exotic models such as those explored in Chapter 5. The code is powerful and capable of running high dimensional simulations, however this thesis only makes use of a 1D electrostatic model, of which the workings are discussed in Section 4.4.

Spectral decomposition reduces N_x grid points in x to $(N_k + 1)$ equations. DARK natively works in $\mathbf{k} - \mathbf{v}$ space, but through the use of augmentation terms can be expanded within modules to explore different representations of phase space.

In Section 4.2, I will describe the iterative algorithm that DARK uses as default (Strang splitting), before describing in detail how memory is handled by DARK in Section 4.3. Then, in Section 4.4 I describe how one of DARK's modules simulates a family of models

of which one is explored in Chapter 5.

4.2 Strang splitting

Here, I briefly describe how Strang splitting works, from the viewpoint of dynamical systems. For a system whose n dynamical equations of motion are defined by a rank-2 tensor of differential operators $\hat{Z}^\mu_\nu$, acting in the phase space $(\mathbf{x}, \mathbf{v})$ on the rank-1 state tensor F^μ:

$$\frac{\partial}{\partial t} F^\mu = \hat{Z}^\mu_\nu F^\nu, \tag{4.1}$$

the flow operator tensor Φ^μ_ν yields the trajectory of the system:

$$F^\mu(\mathbf{x}, \mathbf{v}, t + \Delta t) = \Phi^\mu_\nu(\Delta t) \circ F^\nu(\mathbf{x}, \mathbf{v}, t), \tag{4.2}$$

where Δt is the timestep. Note that the contravariant and covariant indices are with respect to the vector space defined by the state tensor, *not* phase space. Next, one can setup an eigenvalue equation for $\hat{Z}^\mu_\nu$:

$$\underline{\underline{\zeta}} \cdot \underline{F} := \hat{Z}^\mu_\nu F^\nu,$$

such that the $n \times n$ matrix $\underline{\underline{\zeta}}$ contains the eigenvalues of the tensor decomposition of $\hat{Z}^\mu_\nu$ in a given Hilbert space, and $\underline{F}$ is the vector form of F^μ. Then, (4.1) solves directly to give:

$$F^\mu(\mathbf{x}, \mathbf{v}, t + \Delta t) = \exp[(\underline{\underline{\zeta}})\Delta t] \cdot \underline{F}(\mathbf{x}, \mathbf{v}, t), \tag{4.3}$$

where I have made use of the matrix exponential:

$$\exp[(\underline{\underline{\zeta}})\Delta t] \equiv \sum_{k=0}^{\infty} \frac{\Delta t^k}{k!} \underline{\underline{\zeta}}^k, \tag{4.4}$$

where $(\underline{\underline{\zeta}})^0$ is the $n \times n$ identity matrix, $(\underline{\underline{\zeta}})^1 := \underline{\underline{\zeta}}$, and $(\underline{\underline{\zeta}})^2 := \underline{\underline{\zeta}}\,\underline{\underline{\zeta}}$. As such, one can identify that $\exp[\underline{\underline{\zeta}}\Delta t]$ is simply an alternative representation of the flow operator.

The Baker-Campbell-Hausdorff formula[55] shows that for two matrices $\underline{\underline{\mathbf{A}}}$ and $\underline{\underline{\mathbf{B}}}$:

$$\exp[(\underline{\underline{\mathbf{A}}} + \underline{\underline{\mathbf{B}}})\Delta t] = \exp[\underline{\underline{\mathbf{A}}}\Delta t] \exp[\underline{\underline{\mathbf{B}}}\Delta t] + \exp\left[-\frac{1}{2}[\underline{\underline{\mathbf{A}}}, \underline{\underline{\mathbf{B}}}]\Delta t^2\right] + \mathcal{O}(\Delta t^3), \tag{4.5}$$

where $[\underline{\mathbf{A}}, \underline{\mathbf{B}}] := \underline{\mathbf{AB}} - \underline{\mathbf{BA}}$ is the commutator of matrices $\underline{\mathbf{A}}$ and $\underline{\mathbf{B}}$. Therefore, if $\hat{Z}^{\mu}_{\nu}$ is representable as a linear sum of m operators such that:

$$\hat{Z}^{\mu}_{\nu} = \sum_{j=1}^{m} (\hat{Z}_j)^{\mu}_{\nu}, \tag{4.6}$$

with $\underset{=j}{\zeta} \cdot \underline{F} := (\hat{Z}_j)^{\mu}_{\nu} F^{\nu}$, and with each operator yielding a corresponding flow operator:

$$(\hat{Z}_j)^{\mu}_{\nu} \mapsto (\Phi_j)^{\mu}_{\nu}(\Delta t)$$
$$\text{s.t.} \qquad F^{\mu}(\mathbf{x}, \mathbf{v}, t + \Delta t) = (\Phi_j)^{\mu}_{\nu}(\Delta t) \circ \underline{F}(\mathbf{x}, \mathbf{v}, t),$$

for $m = 2$, one expects that the exponential in (4.3) would yield under the Baker-Campbell-Hausdorff formula:

$$F^{\mu}(\mathbf{x}, \mathbf{v}, t + \Delta t) = \left\{ \exp[\underset{=0}{\zeta} \Delta t] \exp[\underset{=1}{\zeta} \Delta t] + \exp\left[\frac{1}{2}[\underset{=0}{\zeta}, \underset{=1}{\zeta}]\Delta t^2\right] + \mathcal{O}(\Delta t^3) \right\} \underline{F}(\mathbf{x}, \mathbf{v}, t).$$

If one evolves the system by half a timestep ($\Delta t/2$) iterating through a set of partial flows, and then evolves the system by half a timestep iterating *in reverse* through a set of partial flows, the term involving the commutator vanishes, as $[\underline{\mathbf{A}}, \underline{\mathbf{B}}] = -[\underline{\mathbf{B}}, \underline{\mathbf{A}}]$.

One finds that that this can be extended to an arbitrary number of partial flows, and therefore one can decompose the system of equations into a set of partial flows, allowing one to iterate via the symmetric Strang splitting method:[53,54]

$$\begin{aligned} F^{\mu}(\mathbf{x}, \mathbf{v}, t + \Delta t) = &\left[\prod_{j=1}^{m} (\Phi_j)^{\mu}_{\beta}(\Delta t/2) \circ \right] \\ &\left[\prod_{j=1}^{m} (\Phi_{m-j})^{\beta}_{\nu}(\Delta t/2) \circ \right] F^{\nu}(\mathbf{x}, \mathbf{v}, t) + \mathcal{O}(\Delta t^3), \end{aligned} \tag{4.7}$$

which is correct to $\mathcal{O}(\Delta t^2)$. To summarize the symmetric Strang splitting method in less mathematical terms:

1. Split the system of equations into a set of partial flows, and define an order for Strang iteration.
2. Solve each partial flow *independently* in that order for half a timestep.
3. Solve each partial flow *independently* in reverse order for half a timestep.
4. Repeat steps 2 and 3 until the simulation is complete.

When solving each partial flow, the only the element of the state tensor whose time derivative is being calculated is temporally evolved; every other element is fixed.

4.3 Memory handling

4.3.1 Parallelisation

DARK utilises a rich, flexible parallelisation setup. Parallelisation is natively performed in discrete **k**-space, and in **v**-space.

The data is split onto a process grid which is also temporally fixed and homogeneous across phase space. DARK is a D-dimensional code (D being the total number of position and velocity dimensions), and as such each process is designed to belong to $2 + D$ communicators: `MPI_COMM_WORLD`, `cart_comm`, `x_comms[]`, and `v_comms[]`.

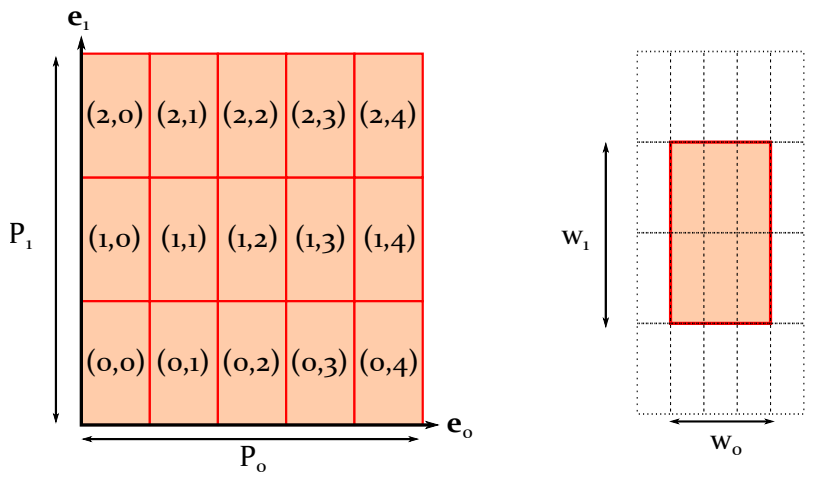


Figure 4.1: **Gridding in DARK.** A field on $\mathbf{e}_0 \times \mathbf{e}_1$ is defined over $P_0 w_0$ data points in the $\mathbf{e}_0$ direction, and $P_1 w_1$ points in the $\mathbf{e}_1$ direction. Employing a parallelisation scheme of (P_0, P_1) leaves $w_0 \times w_1$ points on each process.

`cart_comm` is a Cartesian communicator containing all of the processes on a $P_1 \times P_2 \times \ldots$ grid, where P_j is the parallelisation along the j^{th} dimension, as shown in Figure 4.1 Binary flags allow one to turn periodicity on/off for each dimension. For example in a 2D simulation, the array $\{1, 0, 0, 0\}$ would set the first position coordinate to have periodic boundaries, the second position coordinate to have aperiodic boundaries, and the velocity coordinates to have aperiodic boundaries. Each j^{th} dimension is illustrated here by its unit vector in phase space, $\mathbf{e}_j$.

The MPI group corresponding to `cart_comm` is dissected such that `x_comms[]` and `v_comms[]` are subsets of `cart_comm`. Unlike `cart_comm`, the processes which belong to the communicators `x_comms[]` and `v_comms[]` *are not the same on every process.* Rather, each communicator is defined such that is the dimensional index for position or velocity. For example:

- `x_comms[0]` : all processes sharing the same $x_1, \ldots$ and $v_0, v_1, \ldots$
- `v_comms[0]` : all processes sharing the same $x_0, x_1, \ldots$ and $v_1, \ldots$

This allows for easy collective communications for velocity/spatial reduction operations. Suppose that one had a complex quantity `quant` which was defined with respect to 4 variables (x_0, x_1, v_0, v_1), and one desired the sum over the $\mathbf{e}_3$ direction (v_1). Performing a collective communication across `MPI_COMM_WORLD` or equivalently `cart_comm` could prove to be computationally expensive. Instead, to sum over all of them one would could simply use:

```
MPI_Allreduce(MPI_IN_PLACE, &quant, 1, MPI_C_DOUBLE_COMPLEX,MPI_SUM,v_comms[1]);
```

Master processes are defined as ones where a subset of the coordinates are all 0. For example, in a 2D simulation, $(0, 0, c, d)$ are master processes for all c, d. These processes perform calculations that are done after collective communications, and are required for safe file I/O. The deposit is defined as the process at $(0, \ldots)$, and is typically used to write to `stdout` and `stderr`.

4.3.2 Data allocation

All of the position and velocity dependent quantities are stored in jagged arrays to enable non-square gridding. Each process holds $w_0 \times w_1 \times \ldots$ data points in a given array, where w_j are datasizes:

$$w_j := \frac{N_j}{P_j},$$

where N_j is the number of grid points in the $\mathbf{e}_j$ direction, and P_j is the number of processes in the $\mathbf{e}_j$ direction. On each process, the quantity `arr` can be described by coordinates:

`arr[i][j]...` : relative coordinate on the process given by $(i, j, \ldots)$.

The jagged arrays are defined dynamically via big-endian representation of the relative Cartesian coordinate on the process itself. For example, if the gridding is setup such that `arr` defined in x_0, x_1, v_0 has corresponding datasizes on a process given by $(w_0, w_1, w_2) = (2, 3, 4)$, all of (0,0,0) to (0,0,3) exists in one contiguous block of memory, while (0,1,0) to (0,1,3) exists in another contiguous block of memory. DARK uses pointer arrays to access the data in memory.

On input and output, the 'mixed-base' value given by the Cartesian communicator coordinates are converted to decimal for file naming. For example, for parallelisation setup such that x_0, x_1, v_0 are parallelised with $(P_0, P_1, P_2) = (2, 2, 2)$, the full grid contains $4 \times 6 \times 8$ points. The decimal value is given by:

$$(a, b, \cdots) \to a(P_1 w_1)(P_2 w_2)(\dots) + b(P_2 w_2)(\dots) + \dots$$

Therefore, for a process at $(0, 1, 1)$:

Comm.	Proc.	Full grid		Decimal
⋮	⋮	⋮		⋮
$(0,1,1)$	$(0,0,0)$	$(0,3,4)$	$\to$	28
$(0,1,1)$	$(0,0,1)$	$(0,3,5)$	$\to$	29
⋮	⋮	⋮		⋮
$(0,1,1)$	$(0,1,0)$	$(0,4,4)$	$\to$	36
$(0,1,1)$	$(0,1,1)$	$(0,4,5)$	$\to$	37
⋮	⋮	⋮		⋮
$(0,1,1)$	$(1,0,0)$	$(1,3,4)$	$\to$	76
$(0,1,1)$	$(1,0,1)$	$(1,3,4)$	$\to$	77
⋮	⋮	⋮		⋮

where the left hand side of the map is the relative Cartesian coordinate on the process, and the right hand side of the map is the decimal representation. Therefore, for each variable, there are $P_0 w_0 \times P_1 w_0 \times \cdots$ files. As an example, `arr_00036.csv` would contain data at $(0, 4, 4)$ on the *full* grid. As such, plotting software requires conversion from the decimal representation in the file name to the coordinate on the full grid.

During the simulation, data is written to temporary files in a local temporary folder, which can be defined as different for each process. This temporary directory can be supplied within a job script given to a queue managed by SGE or Slurm, which is then given as an argument upon executation of the DARK binary. These temporary files are held as binary data in `.DARKdat` format.

4.4 Dissipative 1D electrostatic models

4.4.1 Dynamical equations

Here, I introduce a family of 'Berk-Breizman' type models (strictly speaking these are dissipative 1D electrostatic models) of the form:

$$\frac{\partial f_{[l]}}{\partial t} = \hat{C}_{[l]} f_{[l]} - v\frac{\partial f_{[l]}}{\partial x} - \frac{q_{[l]}}{m_{[l]}} E \frac{\partial f_{[l]}}{\partial v}, \qquad \forall\, l \,\text{det.} \tag{4.8a}$$

$$f_{[l]} = f_{[l]}(x, v, t), \qquad \forall\, l \,\text{sto.} \tag{4.8b}$$

$$\frac{\partial E}{\partial t} = -\frac{1}{\epsilon_0}\int_{\mathbb{R}} v \sum_l (q_{[l]} f_{[l]})\ \mathrm{d}v - \alpha * E, \tag{4.8c}$$

where det. denotes deterministic populations. Later, in Chapter 5, I describe how such a model is constructed from physical considerations; here I simply cover how this is analysed computationally. One can construct the following state vector for these models:

$$(F^\mu) = (\{f_{[l]}\}, E).$$

The PDFs are given by the following:

$$f_{[l]}(x, v, t) = f_{[l],(0)}(x, v) + \Delta f_{[l]}(x, v, t),$$

such that $\{\Delta f_{[l]}\}$ denotes the perturbation to the PDFs. The species $\{l\}$ are separated into deterministic (det.) and stochastic (sto.) populations. Each deterministic population satisfies the Boltzmann equation, while each stochastic population is defined *ad hoc*. I prescribe the stochastic PDFs using the following form:

$$\mathcal{F}_v[f_{[l],j}](s, t) = \frac{1}{\sqrt{2\pi}}\epsilon_{f,j}\,\mathrm{sinc}\left(\frac{sL_v}{2}\right) + \mathcal{F}_v[\mathcal{N}_{f,j}],$$

where $\mathrm{sinc}(s) \equiv \sin(s)/s$ is the sinc function, $\mathcal{N}_f(x, v, t)$ is a noise term generated by a pseudorandom number generator. The electric field is split into three parts:

$$E := \delta E + E_{\mathcal{N}} + E_{\mathcal{S}},$$

where δE defines a perturbation, $E_{\mathcal{S}}$ is a seeded deterministic field, and $E_{\mathcal{N}}$ is a seeded stochastic field. As velocity space is infinite here, it serves one best to use Fourier trans-

forms than a discrete decomposition. I will use the following symmetric definitions:

$$\mathcal{F}_v[f_{[l],j}] := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f_{[l],j}(v,t) \mathrm{e}^{-\mathrm{i}sv} \, \mathrm{d}v, \tag{4.9a}$$

$$\mathcal{F}_v^{-1}[f_{[l],j}] := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathcal{F}_v[f_{[l],j}](s,t) \mathrm{e}^{\mathrm{i}sv} \, \mathrm{d}s. \tag{4.9b}$$

As is later discussed in further detail in Chapter 5, equation (4.8c) now takes the form:

$$\partial_t \delta E_j = - \int_{\mathbb{R}} W \sum_l q_{[l]} (\mathcal{F}_v \left[f_{[l],j} \right]) \, \mathrm{d}s - \frac{1}{2} \alpha_j \delta E_j + \mathcal{D}_j,$$

where $\mathcal{D}_j$ is given by:

$$\mathcal{D}_j \equiv - \left(\frac{\partial}{\partial t} + \frac{\alpha_j}{2} \right) (E_{\mathcal{S},j} + E_{\mathcal{N},j}),$$

and $W(s) \in \mathbb{I}$ is the complex formally divergent integral (which acts a removable singularity in the current):

$$W(s) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} v \mathrm{e}^{\mathrm{i}sv} \, \mathrm{d}v.$$

The terms $E_{\mathcal{S},j}$ and $E_{\mathcal{N},j}$ are defined *ad hoc*, as is shown later in Chapter 5. Linear collisions are implemented, such that the $\hat{C} f_{[l]}$ features no terms $\mathcal{O}(f_{[l]}^2)$:

$$\hat{C} f_{[l]} = \sum_m \nu_m \left(\frac{\partial}{\partial v} \right)^m (f_{[l]} - F_{[l],0}),$$

where $F_{[l],0}(v)$ is a spatially and temporally homogeneous function. I approximate that the Lorentz force is given by:

$$\left[E \frac{\partial f}{\partial v} \right]_j = \begin{cases} E_j \dfrac{\partial f_0}{\partial v} + E_0 \dfrac{\partial f_j}{\partial v} & \forall \, j \neq 0, \\ E_0 \dfrac{\partial f_0}{\partial v} + \dfrac{1}{4} \displaystyle\sum_j \left[E_j \frac{\partial f_j *}{\partial v} + \text{c.c.} \right] & j = 0. \end{cases}$$

In this sense, the discrete convolution that should arise under spectral decomposition of $E \partial_v f$ is simplified for $j \neq 0$. In this sense, waves only couple to their own wavelength and the spatial background, such that:

$$\exists \mathbf{k}_\alpha + \mathbf{k}_\beta = \mathbf{k}_\delta : \quad k_\delta \in 0 \cup \{k_\alpha\} \cup \{2k_\alpha\}.$$

Finally, I solve the entire problem in velocity Fourier space, and under spatial spectral decomposition. By using s-space, the code computes collisions and velocity advection at

the same time, giving a factor 2 increase in speed, and only requires backward transforms via Fastest Fourier Transform in the West (FFTW)[56] on PDF output, providing a potential $\mathcal{O}(N_v \ln N_v)$ decrease in computational time on each timestep. Listed together, all of the $(6N_k + 2)$ partial flows are as follows:

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[\Delta f_{[l],j}\right] = \mathcal{F}_v\left[\sum_l \nu_l \left(\frac{\partial}{\partial v}\right)^l f_{[l],j}\right], \qquad \forall\, l\,\text{det.}, j \neq 0 \tag{4.10a}$$

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[\Delta f_{[l],0}\right] = \mathcal{F}_v\left[\sum_l \nu_l \left(\frac{\partial}{\partial v}\right)^l (f_{[l],0} - F_{[l],0})\right], \qquad \forall\, l\,\text{det.} \tag{4.10b}$$

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[\Delta f_{[l],j}\right] = k_j \frac{\partial}{\partial s}\mathcal{F}_v\left[f_{[l]_j}\right], \qquad \forall\, l\,\text{det.}, j \neq 0 \tag{4.10c}$$

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[\Delta f_{[l],j}\right] = \frac{\mathrm{i}sq_{[l]}}{m_{[l]}}\left\{E_0\mathcal{F}_v\left[f_{[l],j}\right] + E_j\mathcal{F}_v\left[f_{[l],0}\right]\right\}, \qquad \forall\, l\,\text{det.}, j \neq 0 \tag{4.10d}$$

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[\Delta f_{[l],0}\right] = \frac{\mathrm{i}sq_{[l]}}{m_{[l]}}\left\{\begin{array}{l} E_0\mathcal{F}_v\left[f_{[l],0}\right] \\ +\frac{1}{4}\sum_j (E_j\mathcal{F}_v\left[f_{[l],j}\right]^* + \text{c.c.}) \end{array}\right\}, \qquad \forall\, l\,\text{det.} \tag{4.10e}$$

$$-\frac{\partial}{\partial t}\delta E_j = \int_{\mathbb{R}} W \sum_l (\mathcal{F}_v\left[f_{[l],j}\right])\,\mathrm{d}s - \frac{\alpha_j}{2}\delta E_j + \mathcal{D}_j, \qquad \forall\, j \neq 0 \tag{4.10f}$$

$$\mathcal{F}_v\left[f_{[l],j}\right] = \frac{1}{\sqrt{2\pi}}\epsilon_{f,j}\,\mathrm{sinc}\left(\frac{sL_v}{2}\right) + \mathcal{F}_v[\mathcal{N}_{f,j}], \qquad \forall\, l\,\text{sto.} \tag{4.10g}$$

$$\mathcal{D}_j = -\left(\frac{\partial}{\partial t} + \frac{\alpha_j}{2}\right)(D(1-\epsilon_{E,j})\mathrm{e}^{\mathrm{i}\omega_s t} + E_{\mathcal{N},j}). \qquad \forall\, j \neq 0 \tag{4.10h}$$

where sto. denotes stochastic populations. Here, I will show how DARK performs simulations for models of this type; equations (4.10g) and (4.10h) are given *ad hoc*, and therefore omitted from the following discussion.

4.4.2 Linear collisions

The collisional parts (4.10a) and (4.10b) are given by:

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[\Delta f_{[l],j}\right] = \mathcal{F}_v[\hat{C} f_{[l],j}],$$

with the collision operator taking the linear form:

$$\hat{C} f_{[l],j} = \begin{cases} \sum_m \nu_m \frac{\partial^m}{\partial v^m} f_{[l],j} & \forall\, j \neq 0, \\ \sum_m \nu_m \frac{\partial^m}{\partial v^m}(f_{[l],0} - F_{[l],0}(v)) & j = 0. \end{cases}$$

After forward velocity Fourier transforms one obtains:

$$\mathcal{F}_v\left[\hat{C} f_{[l],j}\right] = \begin{cases} g(s)\mathcal{F}_v[f_{[l],j}] & \forall\, j \neq 0, \\ g(s)\mathcal{F}_v[f_{[l],0} - F_{[l],0}] & j = 0, \end{cases}$$

where $g(s) \in \mathbb{C}$ is given by:

$$g(s) := \sum_m \nu_m (\mathrm{i}s)^m.$$

The fact that collisions spread energy gives a typical physical requirement that $\Re(g) < 0$; if this is not the case, the collisions cause the PDF to locally spike. Therefore, one obtains two trivial decoupled differential equations in s-space:

$$\begin{aligned} \frac{\partial}{\partial t}\mathcal{F}_v[f_{[l],j}] &= g(s)\mathcal{F}_v[f_{[l],j}] && \forall\, j \neq 0, \\ \frac{\partial}{\partial t}\mathcal{F}_v[f_{[l],0} - F_{[l],0}] &= g(s)\mathcal{F}_v[f_{[l],0} - F_{[l],0}] && j = 0, \end{aligned}$$

where one should note that $\partial\mathcal{F}_v\left[f_{[l],j}\right]/\partial t \equiv \partial\mathcal{F}_v\left[\Delta f_{[l],j}\right]/\partial t$. One therefore finds the following solutions in s-space:

$$\begin{aligned} \mathcal{F}_v[f_{[l],j}](s, t+\Delta t) &= \mathcal{F}_v[f_{[l],j}](s,t)\mathrm{e}^{g(s)\Delta t} && \forall j \neq 0, \\ \mathcal{F}_v[f_{[l],0}](s, t+\Delta t) &= \mathcal{F}_v[F_{[l],0}](s) + \left\{\mathcal{F}_v[f_{[l],0}](s,t) - \mathcal{F}_v[F_{[l],0}](s)\right\}\mathrm{e}^{g(s)\Delta t} && j = 0. \end{aligned}$$

By using the fact that $\mathcal{F}_v[\Delta f_{[l],j}] = \mathcal{F}_v[f_{[l],j}] - \mathcal{F}_v[f_{[l],(0),j}]$, one can evolve $\mathcal{F}_v[\Delta f_{[l],j}]$.

4.4.3 Spatial advection

The spatial advection term (4.10c) is given by:

$$\frac{\partial}{\partial t}\mathcal{F}_v\left[f_{[l],j}\right] = k_j \frac{\partial}{\partial s}\mathcal{F}_v\left[f_{[l]_j}\right] \qquad \forall\, l \,\text{det.}, j \neq 0.$$

This partial flow comes from the $v\partial f/\partial x$ term in the Boltzmann equation, and therefore represents spatial advection. On the right hand side, the Fourier transform in velocity yields $-\mathrm{i}\partial/\partial s$, while the spectral decomposition in space yields $\mathrm{i}k_j$. Altogether, one is now left with an advection in s-space, which is solved by DARK using the piecewise parabolic method[57] as used in previous work by Arber, Vann and de Gol.[26,27] This method is correct to $\mathcal{O}(\Delta t^2)$, making it suitable for this computational setup.

As this is an advection problem, it is computationally constrained by the Courant-Friedrich-Lewy (CFL) limit.[58] As the advection is performed in s-space, the CFL limit is not fully intuitive. Simply given:

$$\frac{k_j \Delta t}{\Delta s} < 1.$$

However the s-space discretization is simply given by the length in v; it is independent of the number of grid points:

$$\frac{k_j L_v \Delta t}{2\pi} < 1,$$

where L_v is the length of the velocity grid. This leads to a few counterposed requirements. In order to allow for one to adequately resolve features in s-space, one requires a large L_v to make Δs as small as possible. Naturally, one must also make Δt small to avoid violating the CFL limit. However, one also desires a suitable density of points in v-space to adequately sample the initial PDF. Counterintuitively, this requires one to sample quite a lot of points outside of the 'interesting region' of the initial PDF; one must add a lot of points quite far from the main part of the population to allow large L_v while preserving a high density of points in v-space. If one failed to include these points, features would not adequately be resolved in s-space, and the simulation would be computationally stable but inaccurate.

4.4.4 Lorentz force

The Lorentz force terms (4.10d) and (4.10e) are given by:

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{F}\left[\Delta f_{[l],j}\right] &= \frac{\mathrm{i}s q_{[l]}}{m_{[l]}}\left\{E_0 \mathcal{F}\left[f_{[l],j}\right] + E_j \mathcal{F}\left[f_{[l],0}\right]\right\} && \forall\, l \,\text{det.}, j \neq 0, \\
\frac{\partial}{\partial t}\mathcal{F}\left[\Delta f_{[l],0}\right] &= \frac{\mathrm{i}s q_{[l]}}{m_{[l]}}\left\{\begin{array}{l} E_0 \mathcal{F}\left[f_{[l],0}\right] \\ +\frac{1}{4}\sum_j (E_j \mathcal{F}\left[f_{[l],j}\right]^* + \text{c.c.}) \end{array}\right\} && \forall\, l \,\text{det.}
\end{aligned}$$

The velocity advection that would occur is now replaced by a simple product in s-space. Both of the equations above are first-order partial differential equations (PDEs) of the form $\dot{g}(t, \dots) = a g(t, \dots) + b$. This type of equation trivially solves to give:

$$g(t + \Delta t, \dots) = \begin{cases} g(t, \dots)\exp[a\Delta t] + \dfrac{b}{a}\left(\exp[a\Delta t] - 1\right) & \forall a \neq 0, \\ g(t, \dots) + b\Delta t & j = 0. \end{cases} \tag{4.11}$$

Simply replacing g, a and b with the relevant quantities yields the partial flows.

4.4.5 Dissipation

The evolution of the electric field perturbation (4.10f) is given by:

$$\frac{\partial}{\partial t}\delta E_j = -\int_{\mathbb{R}} W \sum_l (\mathcal{F}\left[f_{[l],j}\right])\,\mathrm{d}s - \frac{1}{2}\alpha_j \delta E_j + \mathcal{D}_j \qquad \forall\, j \neq 0.$$

Again, this equation is of the form $\dot{g}(t,\dots) = ag(t,\dots) + b$, and therefore simply replacing g, a and b in (4.11) with the relevant quantities yields the partial flow.

4.4.6 Energy flow

The total energy density T of the system is given by (3.28):

$$T(x,t) = \frac{\epsilon_0}{2}E^2 + \frac{1}{2}\int_{\mathbb{R}} \sum_l f_{[l]} m_{[l]} v^2\,\mathrm{d}v.$$

If one considers the rate of change of energy within the electromagnetic field:

$$\left.\frac{\mathrm{d}U}{\mathrm{d}t}\right|_{\mathrm{EM}} = \epsilon_0 \int_0^L E\,\frac{\partial E}{\partial t}\,\mathrm{d}x.$$

By substituting the modified Maxwell-Ampère law (4.8c):

$$\left.\frac{\mathrm{d}U}{\mathrm{d}t}\right|_{\mathrm{EM}} = \epsilon_0 \int_0^L E \left\{ \int_{\mathbb{R}} \left[v f_{[l]}\,\mathrm{d}v\right] - \alpha * E \right\} \mathrm{d}x.$$

By including the contribution from collisions and the distribution function:

$$\frac{\mathrm{d}U}{\mathrm{d}t} = \left.\frac{\mathrm{d}U}{\mathrm{d}t}\right|_{\mathrm{dist.}} + \left.\frac{\mathrm{d}U}{\mathrm{d}t}\right|_{\mathrm{EM}} + \left.\frac{\mathrm{d}U}{\mathrm{d}t}\right|_{\mathrm{coll.}},$$

where dist., EM, and coll. denote contributions from the PDF, electromagnetic field and collisions. The PDF contribution and the first term in the electromagnetic field contribution can be shown to cancel each other out, leaving:

$$\frac{\mathrm{d}U}{\mathrm{d}t} \approx \underbrace{-\frac{1}{4}\sum_j \alpha_j |E_j|^2}_{\text{dissipation}} + \overbrace{\frac{L}{2}\sum_{l,m}\int_{\mathbb{R}} \nu_m \frac{\partial^m}{\partial v^m}(f_{[l],0} - F_{[l],0})v^2\,\mathrm{d}v}^{\text{coll.}}.$$

Stochastic electric field

The stochastic contribution to the electric field can be seen as nothing more than a randomized source of electromagnetic energy in the system. By integrating over the typical period of pseudorandomization, $\tau_{\mathcal{N}}$ one is able to find noise period averaged quantities. Suppose that one investigates a simplified model where the noise term's contribution to the energy density is of the form:

$$\frac{\partial T}{\partial t}(x,t) \sim \epsilon_0 E \mathcal{N}.$$

One can expand $\mathcal{N}$ using a Taylor expansion:

$$\mathcal{N}(t') = \mathcal{N}(t) + \left.\frac{\partial \mathcal{N}}{\partial t}\right|_{t'=t} (t'-t) + \mathcal{O}(t'^2).$$

However, if one averages over the period, only the first term remains; there is no variation of the noise level on average. Therefore:

$$\Delta T(x, t+\Delta t) \sim \epsilon_0 \mathcal{N}(x)_{\text{avg}} \int_t^{t+\Delta t} E(x,t')\,\mathrm{d}t' + \mathcal{O}\left(\frac{\Delta t}{\tau_{\mathcal{N}}}\right)^2,$$

where $\Delta T(x, t+\Delta t)$ is the change in energy density from time t to $t+\Delta t$. If one examines the total energy instead:

$$\Delta U(t+\Delta t) \sim \epsilon_0 \mathcal{N}(x)_{\text{avg}} \int_0^L \int_t^{t+\Delta t} E(x,t')\,\mathrm{d}t'\,\mathrm{d}x + \mathcal{O}\left(\frac{\Delta t}{\tau_{\mathcal{N}}}\right)^2.$$

In this form one can see that to first order, the noise term performs work on the system based on the electric field at a given moment in time. As the energy density increases, the average electric field increases, causing the energy density to increase further in a feedback loop.

The energy density of the system increases like $\mathcal{N}_{\text{avg}}\sqrt{T}$, and therefore the growth is slower than exponential growth.

Even though this is only correct to first order, one can argue that if $\tau_{\mathcal{N}}$ is made large enough, then it is acceptable to treat the noise level as roughly constant for schemes that need to be correct to higher order in Δt. If the value of $\mathcal{N}_{\text{avg}}$ is zero, then to first order the noise term performs no net work on the system.

Dissipative sink

The dissipation term acts as a nonlinear sink for the electromagnetic energy density. To illustrate this, if one examines just the dissipative term's contribution:

$$\frac{\partial T}{\partial t} \sim -2\alpha T \rightarrow T(t+\Delta t) = T(t)e^{-2\alpha\Delta t}.$$

Therefore, if α is large enough, the sink will eventually quench any growth from noise. Accordingly, it is possible to define a system that can be seeded by noise, grow for a while, and then become dominated by dissipation; this allows for chirping behaviour.

4.5 Brief conclusions

In this chapter, I presented a new, modular code (DARK) which is capable of running sophisticated simulations for multi-dimensional kinetic systems with augmentation terms. In Section 4.2, I derived Strang splitting from tensor analysis of the dynamical equations of motion, showing that this method of iteration is correct to $\mathcal{O}(\Delta t^2)$. In Section 4.3, I covered the extensive parallelisation carried out by DARK, and how memory and data is handled by the software. In Section 4.4, I showed in depth how an extended family of 1D electrostatic Berk-Breizman like models can be simulated within DARK, solving a richer, more sophisticated system of equations than is typically examined in the literature. I showed how a stochastic population of particles and a drive field is modelled *ad hoc*, allowing one to solve the system of equations using a Strang split set of partial flows in $\mathbf{k} - \mathbf{s}$ space.

In Chapter 5, DARK is used to run simulations, investigating stochastic and deterministic modelling of holes and clumps.

Chapter 5

Stochastic and deterministic modelling of holes and clumps

5.1 Overview

5.1.1 Frequency modification

In linear systems, a monofrequency wave will have its frequency conserved provided that the energy of the wave is time-invariant. If the energy of the wave changes, this is typically due to an increase in wave amplitude, or slow changes to the system background. To illustrate this, if one examines the dispersion relation of a 1D electron electrostatic wave in the $k \to 0$ limit:

$$\lim_{k\to 0} \omega(k) = \sqrt{\frac{\bar{n}_\mathrm{e} q^2}{m_\mathrm{e}\epsilon_0}},$$

such that ω is equal to the plasma frequency (see Section 2.3.5), and $\bar{n}_\mathrm{e}$ is the spatially averaged electron density. Therefore, if $\bar{n}_\mathrm{e}$ changes, ω will change. This change typically occurs over much smaller timescales than the rate of evolution of the perturbations in the system; at the heart of the linearized theory, we assume that Δf evolves on a timescale faster than the background, and therefore we assume that the background is temporally static.

However, in nonlinear systems, it is possible for the frequency to change on timescales much quicker than the evolution of the background, allowing one to easily distinguish between nonlinear and linear modes in tokamak plasmas. While wave-wave coupling does exist in nonlinear systems, the frequency evolution is dominated by a much more collective behaviour given by the contribution of all active wavenumbers in the system. If one

examines the term in the Boltzmann equation that corresponds to the Lorentz force in an electrostatic plasma:

$$\frac{\partial f}{\partial t} \sim \frac{q}{m}\mathbf{E} \cdot \nabla_{\mathbf{v}} f.$$

If $|\partial(\ln \mathbf{E})/\partial t| \ll |\partial(\ln f)/\partial t|$, one can approximate that the $\mathbf{E}$ field is constant to lowest order, yielding an advection equation in velocity space. The rate of advection at each point is approximately determined by $\mathbf{E}(x, t = 0)$, leading to shock solutions for phase space structures as they advect under this term.

As an example involving real data, if one examines shot 139317 from NSTX at various times (see Figure 6.8, examined for other purposes in Chapter 6), we observe frequency streaks on spectrograms produced by analysing magnetic fluctuations in the system. The constituent frequencies of plasmas waves in the system can be shown to be undergoing both linear and nonlinear behaviour at this time. During 220 - 240 ms, some modes can be observed between 50 -80 kHz corresponding to toroidal Alfvén eigenmodes (TAEs) in the system whose rate of change of which the typical rate of change of frequency is quite small (on the order of 10 $\mathrm{GHz\,s^{-1}}$); this slow rate of change corresponds with the typical timescale that the background changes on, and therefore these modes are not undergoing kinetic resonance. During 100 - 120 ms, the frequency traces beteen 50- 80 kHz correspond to gap TAEs for which the typical rate of change of frequency is much higher (on the order of 100 $\mathrm{GHz\,s^{-1}}$). This fast rate of change implies very strong nonlinear behaviour.

5.1.2 Three-wave coupling

For a perturbation δf in the particle distribution function (PDF), nonlinear terms $\sim (\delta f)^2$ allow for interaction between waves of differing frequency. As shown in Section 2.3, the electric field is related to the PDF by:

$$(p_j^2 \mu_0 \epsilon_0 + k_j^2)\delta\tilde{\mathbf{E}}_j = -\sum_{\text{species}} q \int_{\mathbb{R}^3} \left(\mu_0 \mathbf{v} p_j + \frac{\mathbf{i}\mathbf{k}}{\epsilon_0}\right) \delta\tilde{f}_j \, \mathrm{d}\mathbf{v}.$$

Accordingly, the Lorentz force allows for terms to be generated with $\mathcal{O}(\delta f^2)$. One can easily illustrate how three-wave coupling occurs by using the convolution theorem:

$$(\delta f)^2 = \mathcal{L}^{-1}\left[\mathcal{L}(\delta f) * \mathcal{L}(\delta f)\right],$$

where $\mathcal{L}$ denotes the forward temporal Laplace transform, $\mathcal{L}^{-1}$ denotes the backward temporal Laplace transform. Therefore, if one were to examine a system with two discrete

frequencies ω_1 and ω_2, one would be able to generate terms with frequency 0, $2\omega_1$, $2\omega_2$, $\omega_1 + \omega_2$, and $|\omega_1 - \omega_2|$. This is three-wave coupling, and leads to **phase locking**; nonlinear interactions of this kind require the waves to be in phase with one another. Phase locking is therefore observable in tokamaks, as one can use bicoherence measurements and spectrogram analysis to determine whether two frequencies in the system are indeed locked. Furthermore, the locked nature of the waves means that frequency traces which have the same rate of change of frequency are strong contenders for phase locking. Phase locked waves exchange energy with one another, and due to this energy exchange can prove disruptive.

5.1.3 Frequency chirping

Fluctuations in the constituent frequencies of an oscillation can increase in value ('up-chirping') or decrease in value ('down-chirping'). It is well known that holes and clumps (H&Cs)[8,28,59] can form in the nonlinear phase of the evolution of an energetic particle driven mode (EPM), where resonant particles exchange energy with the corresponding plasma wave. For EPMs undergoing linear growth, as energy is exchanged between the PDF and the mode, the PDF becomes flattened in the vicinity of the Landau resonance, decreasing the linear growth rate. After a finite amount of time, the PDF is locally without gradient, and the mode reaches nonlinear saturation. In this phase, further energy exchange is allowed between the electromagnetic field and the PDF via H&Cs moving along the PDF; while the mode is linearly stable, the wave frequency changes as the resonant structures move, and the wave energy changes with the frequency.

5.1.4 Alfvénic chirping

In tokamaks, the Alfvén speed varies along a field line due to toroidal periodicity in the magnetic field. This phenomenon (analogous to Bragg reflection in periodic lattices) allows for gaps in the Alfvénic mode band structure. Due to defects in the magnetic field periodicity in real tokamaks, spatially localised modes can exist in the frequency gap between bands in the Alfvén continuum. These defects can include asymmetry produced by the triangularity and ellipticity of the magnetic field.[40]

TAEs are of particular interest to the fusion community. With frequencies on the order of 100 kHz,[40] they have the ability to become amplified via radio frequency (RF) heating and energetic particles. Gap TAEs can exist in the gap between TAEs and beta-induced Alfvén eigenmodes (BAEs), with frequency $< 50\,\text{kHz}$, close to geodesic acoustic mode

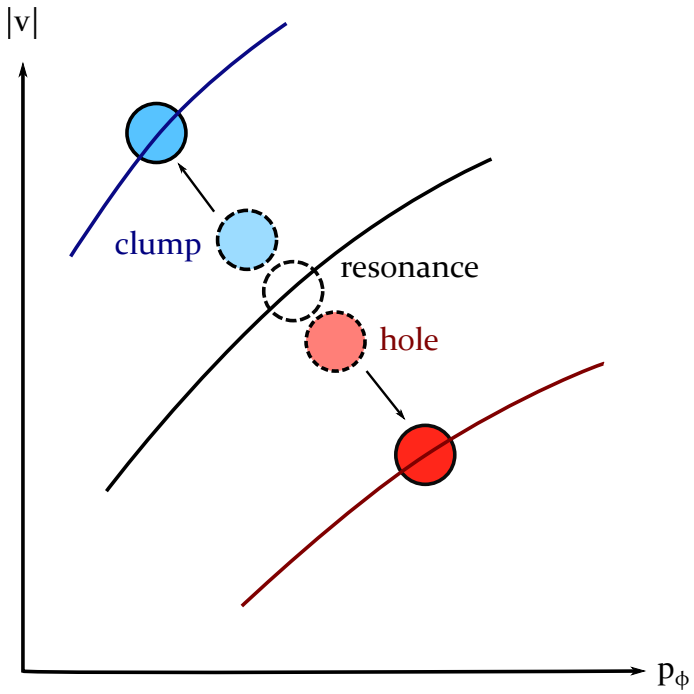


Figure 5.1: **Resonance bifurcation in a tokamak.** A sketch of a H&C on a PDF peaked near the origin for a chirping, resonant mode interacting with a tokamak plasma. The kinetic resonance $\omega+(m+l_p)\omega_\theta-n\omega_\varphi=0$ undergoes pitchfork bifurcation as the H&C move. Stochastic fluctuations modelled in this chapter heuristically map onto fluctuations a function of p_φ; these effect the formation and evolution of the H&C.

(GAM) frequency). Unlike continuum Alfvén eigenmodes, gap TAEs exist as coherent waveforms which are resilient to shear damping.[40] This allows gap TAEs to exist for longer timescales than continuum modes. As shown in Figure 5.1, a hole and clump emerge from a resonance (a contour in $(p_\phi, |\mathbf{v}|)$ space, where p_φ is the toroidal angular momentum, and $|\mathbf{v}|$ is the magnitude of the particle velocity). Holes move to higher p_φ and lower $|\mathbf{v}|$, while conversely clumps move to lower p_φ and higher $|\mathbf{v}|$.

Fluctuations in the magnetic field and particle density have been shown in some cases to undergo frequency chirping in tokamaks.[60–64] Energetic particles undergoing Landau resonance with plasma wave undergo radial diffusion, which can lead to large energetic particle losses in tokamaks.[60,65,66] It is understood that mode chirping is directly correlated with H&Cs; consequently, chirping modes can allow for greater radial diffusion. As a result, even in the case of continuum Alfvén eigenmodes, rapid mode chirping can also lead to a significant channel for fast ion loss; in such a case, the rate of energy loss via chirping is comparable to the sum of damping rates (e.g. collisional, radiative, continuum damping). Gap TAEs allow for an even greater particle loss than continuum Alfvén modes, due to

their resilience to shear dissipation.

These frequency chirping events in tokamaks are therefore directly correlated with fast ion losses in tokamaks, degrading plasma performance and thermal energy confinement in the core. Experimentally this has been observed through the use of fast ion loss detectors (FILD)[67] and by examining the neutron birth rate;[61] sharp peaks in FILD signal and sharp decreases in the neutral birth rate occur during frequency chirping events ('sierpes' mode chirping, Alfvénic chirping).[61,68] In some experimental scenarios, abrupt large events (ALEs) have been observed, where one frequency chirping event (typically down-chirping) is rapidly followed by other frequency chirping events.[50,51] In these scenarios, large fast ion losses are observed, even for modes which undergo continuum damping. Fredrickson *et al.* have previously proposed that mode-mode destabilisation may be a key mechanism for the onset of ALEs.[60]

In Section 5.2, I present the model used in this chapter, as briefly described in Chapter 4. The model allows for stochastic simulations of 1D electrostatic plasma with Berk-Breizman augmentation.

In Section 5.3 and Section 5.4, I present published work carried out in collaboration with Princeton Plasma Physics Laboratory (PPPL). We show that the lifetime of H&Cs is stochastic, implying that the extinction rate of H&Cs may be determined by a stochastic process, and examine how stochasticity in the PDF (mimicing turbulence) affects the coherence of H&Cs.

Finally, in Section 5.5 and Section 5.6 I present other work investigating H&C destabilisation via parametric modification of the background distribution function. I examine how this can lead to destabilisation that is distinctly different from that which would be expected from quasilinear diffusion.

5.1.5 Shorthand and notation

In this chapter, I will use the following shorthand:

$$\omega_j^{[l]} \to \omega_j,$$

that is to say, the superscript l and subscript j will be coalesced into one new subscript, j. The bracketed subscript $[l]$ now denotes species. For example:

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0} \sum_l q_{[l]} \int_{\mathbb{R}^3} f_{[l]} \, \mathrm{d}\mathbf{v}.$$

The PDF is given by:

$$f(x, v, t) = f_{(0)}(v) + \Delta f,$$

where $f_{(0)}$ has been defined such that:

$$f_{(0)}(v) := f(x, v, t = t_0).$$

5.2 Stochastic model

Here, we include stochasticity into 1D kinetic models, allowing us to examine in closer detail the resultant effects on the evolution of resonant modes.

5.2.1 Resonant damping

In a tokamak, plasma heating and particle confinement lead to local maxima in the ion PDF $f_{\text{ion}}(W, p_\varphi, \mu)$ where W is the particle kinetic energy, p_φ is the canonical toroidal angular momentum, and μ is the magnetic moment. The energy PDF (constant p_φ, μ) is peaked near energies associated with heating mechanisms such as neutral beam injection (NBI) and RF heating, and near the energy of ions produced as fusion products (such as fusion alphas) - this leads to damping of resonant modes.[40] The toroidal momentum PDF (constant W, μ) can be canonically mapped onto flux coordinate Ψ via:

$$p_\varphi = MRv_\varphi - Ze\Psi, \tag{5.1}$$

where M is the ion mass, R is major radius, v_φ is the ion toroidal velocity, and Ze is the ion charge. We expect that p_φ decreases as Ψ increases, as the particle density peaks near $r = 0$. Thus, the spatial gradient of particles leads to drive of resonant modes, in an analogous manner to inverse Landau damping.

TAEs resonate with a quasi-2D fast ion PDF where the linear stability is determined by competing $\mathrm{d}f/\,\mathrm{d}|v|$ and $\mathrm{d}f/\,\mathrm{d}p_\varphi$.[40] These correspond to kinetic resonance with the poloidal and toroidal transit frequencies ω_θ and ω_φ respectively,[40,69] given by $\omega + (m + l_p)\omega_\theta - n\omega_\varphi = 0$. Here, m and n are the poloidal and toroidal modenumbers, and $l_p \in \mathbb{Z} \neq 0$ correspond to poloidal harmonics of the drift velocity (see Figure 5.1).

In this chapter, we aim to model the same key instability physics by examining H&C formation on a 1D bump-on-tail (BOT) distribution. This allows us to model energetic particle drive via the positive slope of the PDF between the bulk and the beam, while modelling energetic losses as a damping term.[5,8,26,28,29] The evolution of the system is

determined by coupling the Boltzmann equation to Maxwell's equations; our model is given here by a multiple species generalisation of kinetic models used by Vann *et al.*[29] and De-Gol:[26]

$$\frac{\partial f_{[l]}}{\partial t} = \hat{C}_{[l]} f_{[l]} - v\frac{\partial f_{[l]}}{\partial x} - \frac{q_{[l]}}{m_{[l]}} E \frac{\partial f_{[l]}}{\partial v}, \qquad \forall l \tag{5.2a}$$

$$\frac{\partial E}{\partial t} = -\frac{1}{\epsilon_0}\int_{\mathbb{R}} v \sum_l (q_{[l]} f_{[l]})\ \mathrm{d}v - \alpha E, \tag{5.2b}$$

where $\{\hat{C}_{[l]}\}$ are collision operators, $f_{[l]}(x, v, t)$ is the energetic particle PDF for the l^{th} species, $E(x, t)$ is the electric field, and ϵ_0 is the permittivity of free space. Damping is effected in the system via $\alpha(x, t)$, and acts as a sink of electromagnetic field energy, emulating the effect of the Berk-Breizman collision operator (3.33). Formally, one can show that this augmentated model still preserves momentum and energy globally (see derivation in Appendix C.1 from classical field theory) for the case of a single active wave in the system. The corresponding 1D Landau resonance condition is:

$$\omega + (j + l_d)2\pi v/L_x = 0, \tag{5.3}$$

where j is the modenumber, $l_d \in \mathbb{Z} \neq 0$ correspond to harmonics of the 1D drift velocity, v is the particle velocity, and L_x is the length of the 1D box. For the electrons to be resonant, we require a suprathermal population of electrons. If this population does not exist, then the electrons can be treated as a fluid; no electron kinetic instabilities occur, and electrons only serve to weakly damp plasma waves in the system via Landau damping. Then, if the electrons are also treated as collisionless:

$$\frac{\partial f_{[l]}}{\partial t} = \hat{C}_{[l]} f_{[l]} - v\frac{\partial f_{[l]}}{\partial x} - \frac{q_{[l]}}{m_{[l]}} E \frac{\partial f_{[l]}}{\partial v}, \qquad \forall l \neq \text{el.} \tag{5.4a}$$

$$\frac{\partial n_{\text{el.}}}{\partial t} = -\frac{1}{q_{\text{el.}}}\frac{\partial J_{\text{el.}}}{\partial x}, \tag{5.4b}$$

$$\frac{\partial E}{\partial t} = -\frac{1}{\epsilon_0}\int_{\mathbb{R}} v \sum_{\text{ions}} (q_{[l]} f_{[l]})\ \mathrm{d}v - \frac{J_{\text{el.}}}{\epsilon_0} - \alpha E, \tag{5.4c}$$

where 'el.' denotes the electrons, $n_{\text{el.}}$ is the electron density, and $J_{\text{el.}}$ is the electron current density. If one examines the linear complex dispersion relation for this system, it is trivial to show that the electrons only serve to shift the frequency of the ions in the system. Therefore, we choose to use a model where we treat the electrons approximately as a neutralising background. This has the benefit of reducing the computational complexity

of the model, while still retaining the key physics, provided that the electrons are suitably cold (such that their contribution to the frequency of ion acoustic waves is negligible). We then assume a single species of fast ions, and apply the following normalisation:

$$\begin{aligned} \text{velocity} &= v_{\text{th}} \cdot v \quad ; \quad \text{time} = \frac{1}{\omega_{\text{pl.}}} \cdot t \\ \text{charge} &= q_{\text{ion}} \cdot q_{[l]} \quad ; \quad \text{mass} = M \cdot m_{[l]} \\ \text{dist. func.} &= \frac{N_{\text{ion}} \omega_{\text{pl.}}}{v_{\text{th}}^2} \cdot f_{[l]}, \end{aligned}$$

where v_{th} is a normalising velocity, $\omega_{\text{pl.}}$ is the ion plasma frequency, q_{ion} is the ion charge, M is the ion mass, and N_{ion} is the number of ions. The quasi-thermal quantity v_{th} normalises the energy of the system, but the equilibrium is not a thermal equilibrium. We choose the value to be the root-mean-squared velocity of ions after thermalization of the initial PDF, in the absence of electromagnetic interaction:

$$v_{\text{th}} \equiv \sqrt{\int v^2 f(x, v, t = 0)\, \mathrm{d}x\, \mathrm{d}v}. \tag{5.5}$$

This leads to the normalisation:

$$\text{position} = \left[v_{\text{th}} \omega_{\text{p}}^{-1} \right] \cdot x \quad ; \quad \text{electric field} = \left[N_{\text{ion}} \frac{m v_{\text{th}} \omega_{\text{p}}}{q} \right] \cdot E.$$

The model then takes the form:

$$\boxed{\begin{aligned} \frac{\partial f}{\partial t} &= \hat{C} f - v \frac{\partial f}{\partial x} - E \frac{\partial f}{\partial v}, && \text{(5.6a)} \\ \frac{\partial E}{\partial t} &= - \int_{\mathbb{R}} \sum_{\text{ions}} v f \, \mathrm{d}v - \alpha E, && \text{(5.6b)} \end{aligned}}$$

where we have dropped the species subscript, $f(x, v, t)$ is the ion PDF for the l^{th} ion species, $E(x, t)$ is the electric field.

5.2.2 Linear complex dispersion relation

The growth rate of an EPM is proportional to:[40]

$$\gamma \propto \left[\omega \frac{\partial f}{\partial W} + n \frac{\partial f}{\partial p_\varphi} \right]_{\omega + (m + l_p)\omega_\theta - n\omega_\varphi = 0}. \tag{5.7}$$

Here, we take the model used in (5.6a) and (5.6b) and show *a posteriori* that the growth rate has the same canonical form as (5.7). Assuming constant frequency and growth rate,

in the collisionless case we find the linear complex dispersion relation (see Section 2.3):

$$p_j(t=t_0) + \frac{\alpha_j}{2} = \int\limits_{\Omega_{\text{Lan.}}} \frac{v\dfrac{\partial}{\partial v} f_0(v, t=t_0)}{p_j(t=t_0) + \mathrm{i}k_j v}\,\mathrm{d}v, \tag{5.8}$$

where $\Omega_{\text{Lan.}}$ is the suitable Landau contour for the problem. One must take care to note that here we use $f_0(v, t=t_0)$, not $f_{(0)}$. This is because here (unlike in Section 2.3), we are performing a more general linearization at a time $t = t_0$, where t_0 may not be equal to 0. The complex frequency p_j is given by $p_j \equiv \gamma_j - i\omega_j$, where ω_j is the frequency of the j^{th} mode, and γ_j is the growth rate of the j^{th} mode. A first-order expansion in v for the integrand gives the growth rate as:

$$\gamma_j \sim -\frac{\alpha_j}{2} + \frac{\pi\omega_j}{2k_j^2}\left.\frac{\partial f_0}{\partial v}\right|_{v=\omega_j/k_j, t=t_0}. \tag{5.9}$$

By factorization, one finds that:

$$\gamma_j \propto -\omega_j \frac{\alpha_j}{\pi v_j^2} + \left.\frac{\partial f_0}{\partial v}\right|_{v=v_j, t=t_0}.$$

In this form, one can see that the term $\partial f/\partial W$ in a real tokamak (5.7) is modelled here by the analogous term $-\pi\alpha_j/v_j^2$. Therefore, our model captures the same damping as a tokamak under the following constraints:

1. The energy PDF has approximately constant gradient near resonance.

2. No resonance broadening occurs on the energy PDF.

The term $n\partial f/\partial p_\varphi$ in a real tokamak is modelled by the analogous term $\partial f_0/\partial v$. Therefore, our model captures the same drive mechanism as a tokamak under the following constraints:

1. The growth rate is dominated by the spatially averaged PDF f_0.

2. The gradient of f in the $v_\perp$ direction yields a negligible contribution to the growth rate.

We find that from simulations, collisions modify the linear growth rate $\gamma_j(\alpha_j, k_j)$, but for the case of sufficiently weak collisions, this correction is negligible.

5.2.3 Two species model for turbulence

Here, we model the plasma by using two separate PDFs of identical particles: a fast ion PDF f_{ion} with a deterministic phase space trajectory in the absence of electric field, and a turbulent PDF f_{tur} with a stochastic trajectory. By modelling the plasma using two separate PDFs, one is able to vary with ease the fraction of particles that are turbulent. Fluctuations in f_{tur} lead to fluctuations in the electric field via (5.6b). Fluctuations in the electric field interact with f_{ion} via (5.6a), heuristically mapping via αE onto the energy exchange associated with particle resonance along $\partial f/\partial W$. The model is therefore a two species plasma:

$$\frac{\partial f_{\text{ion}}}{\partial t} = \hat{C}_{\text{ion}} f_{\text{ion}} - v\frac{\partial f_{\text{ion}}}{\partial x} - E\frac{\partial f_{\text{ion}}}{\partial v}, \tag{5.10a}$$

$$\frac{\partial f_{\text{tur}}}{\partial t} = \hat{C}_{\text{tur}} f_{\text{tur}} - v\frac{\partial f_{\text{tur}}}{\partial x} - E\frac{\partial f_{\text{tur}}}{\partial v}, \tag{5.10b}$$

$$\frac{\partial E}{\partial t} = -\int_{\mathbb{R}} v(f_{\text{ion}} + f_{\text{tur}})\ \mathrm{d}v - \alpha E. \tag{5.10c}$$

Here, we evolve f_{ion}, but prescribe f_{tur} using pseudorandom noise. In lieu of an implemented model beyond Fokker-Planck theory, we instead define the trajectory of f_{tur} *ad hoc* to investigate the key resultant physics. For $\hat{C} = \mathcal{O}(\partial/\partial v)^3$, one can show that the resultant hyperjerk equation $\partial f/\partial t = \hat{C} f$ at fixed t can be represented in the form $\partial \mathbf{g}/\partial v = \mathbf{F}(\mathbf{g})$ where $\mathbf{g} \in \mathbb{R}^n$ where $n \geq 3$, and $\mathbf{F}(\mathbf{g})$ is a smooth function.[70] Via the Poincaré-Bendixson theorem,[71–73] this meets the minimum requirements for chaotic behaviour; a simple *reductio ad absurdum* shows that replacing n with a value less than 3 prevents the formation of chaotic solutions.

It is instructive to note that as two identical species, both the non-turbulent and turbulent fast ions experience the same collision operator. However, while for pitch angle scattering the individual particle trajectories are stochastic, the evolution of the PDF is deterministic in the absence of electric field; collective diffusion leads to a well defined trajectory for f_{ion}. In contrast, the turbulent population undergoes processes which lead to a stochastic evolution of f_{tur}.

Formally, we approximate that for f_{ion}, one can use $C = \mathcal{O}(\partial/\partial v)^2$. This is justified mathematically by arguing that the non-turbulent fast ion PDF exists on a subset of phase space $\mathbb{R}^2_{\text{ion}} \in \mathbb{R}^2$ where its evolution is non-chaotic in the absence of electric field. Therefore, we utilise a Fokker-Planck diffusive collision operator for f_{ion}, in alignment with kinetic

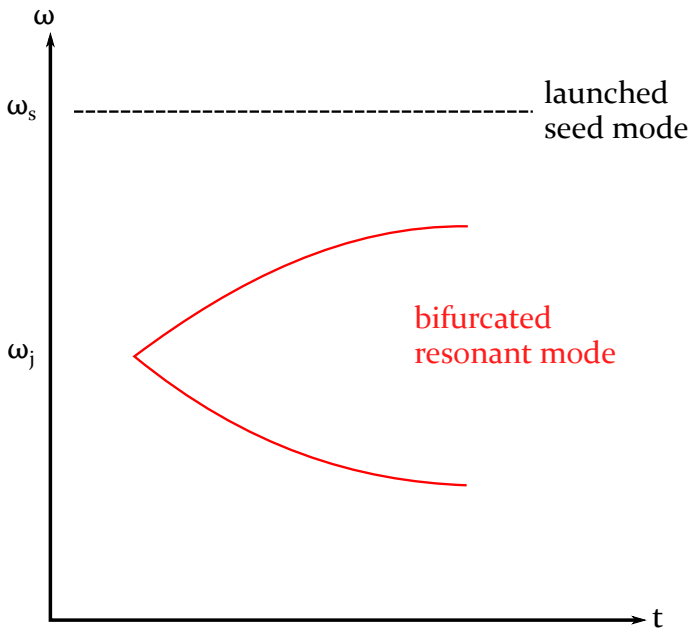


Figure 5.2: **Deterministic drive in reduced dimensionality simulations.** A resonant mode ($\omega_j(t=0) \sim \omega_{\text{pl.}}$) undergoing does so in a frequency bandwidth much lower than the frequency of the launched seed mode with mode drive generated via inverse Landau damping on the positive slope of $F_0(v)$ between 0 and v_B.

descriptions of mode chirping based on pitch angle scattering in the literature:[28,38,64,74]

$$\hat{C}_{\text{ion}} f_{\text{ion}} = \nu \frac{\partial^2}{\partial v^2}(f_{\text{ion}} - F_0), \tag{5.11}$$

where $f_{\text{ion}}(x, v, t=0) \equiv F_0(v)$ is the initially homogeneous ion PDF, and the effective collisional pitch angle scattering diffusion coefficient ν allows H&Cs to have a finite, non-zero lifetime. However, it is important to note that the operator relaxes to the initial PDF (F_0); energy transfer may occur during relaxation.

Similarly to the non-turbulent ions, we demand that f_{tur} exists on a non-intersecting subset of phase space $\mathbb{R}^2_{\text{tur}} \notin \mathbb{R}^2_{\text{ion}}$; that is to say it does not exist in non-chaotic regions. It is instructive to note that these spaces change as a function of time; the deterministic population f_{ion} and stochastic population f_{tur} exist on "transparent grids", such that there are no collisions between particles in f_{tur} and f_{ion}. We find that a second order collision operator is insufficient to generate f_{tur} as we require terms $\mathcal{O}(\partial/\partial v)^3$.

By defining f_{ion} and f_{tur} on separate subsets of phase space, we remove the ability to have interspecies collisions. Consequentially, we remove all anomalous thermal transport from f_{ion} that may occur via interactions with turbulent particles. As a result:

$$\frac{\partial f_{\text{ion}}}{\partial t} = \nu \frac{\partial^2}{\partial v^2}(f_{\text{ion}} - F_0) - v\frac{\partial f_{\text{ion}}}{\partial x} - E\frac{\partial f_{\text{ion}}}{\partial v}, \tag{5.12a}$$

$$f_{\text{tur}} = f_{\text{tur}}(x, v, t), \tag{5.12b}$$

$$\frac{\partial E}{\partial t} = -\int_{\mathbb{R}} v(f_{\text{ion}} + f_{\text{tur}})\ \mathrm{d}v - \alpha E, \tag{5.12c}$$

and f_{ion} and f_{tur} satisfy:

$$\begin{aligned} f_{\text{ion}} &= 0 \quad \forall \{x, v\} \notin \mathbb{R}^2_{\text{ion}}, \\ f_{\text{tur}} &= 0 \quad \forall \{x, v\} \notin \mathbb{R}^2_{\text{tur}}. \end{aligned} \tag{5.13}$$

5.2.4 Approximations

We make a few assumptions to simplify the model computationally. These rely on a Fourier series representation of the PDF and other quantities (as given in Section 1.3), and velocity Fourier transforms (as given in Section 4.4.1).

First, we replace $\alpha(x)E(x)$ with $\alpha(x) * E(x)$; via the convolution theorem, this allows for a piecewise product of α and E in k-space. We require this to be able to damp each mode separately. However, as it does not preserve the canonical form of the Hamiltonian (see C.1), conservation of energy is violated except for the trivial case where $\alpha(x)$ is a constant. Here, we examine this trivial case, and therefore $\forall j\ :\ \alpha_j = 2\alpha$.

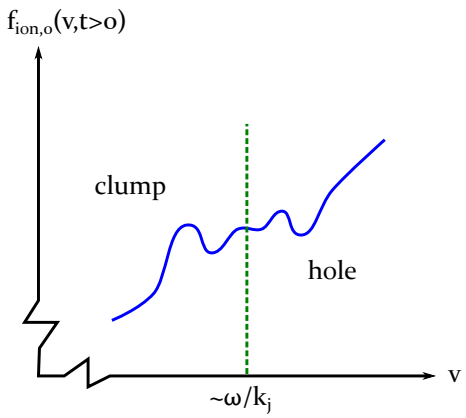
Secondly, we ignore three-wave coupling; this means that no modes exist except for harmonics of the initial set of modes. We allow this as three-wave coupling of gap TAEs will generate modes which exist in the Alfven continuum, which are quickly dissipated.[40]

Thirdly, we also ignore all harmonics of the initial set of modes except for the fundamental; this is justified by requiring physically that these harmonics are rapidly damped. If this assumption holds for the electromagnetic field, it also holds true for the PDF, as harmonics generated via $E\partial f/\partial v$ will also be rapidly dissipated.

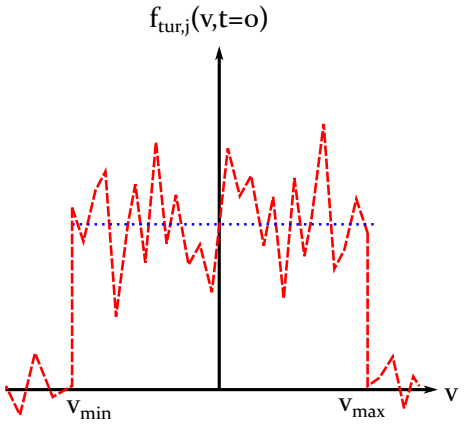
Finally, we force E_0 to be evanescent, and set it to 0 via boundary conditions, removing mean current from the Maxwell-Ampère law. The aforementioned approximations at the heart of the model lead to the following caveats:

$$\left|\alpha * E - \alpha E\right| \ll \left|\int_{\mathbb{R}} vf\ \mathrm{d}v\right|, \tag{5.14a}$$

$$\left|\alpha_0 E_0\right| \gg \left|\int_{\mathbb{R}} vf_0\ \mathrm{d}v\right|, \tag{5.14b}$$



(a) A sketch of a typical set of values for the fast ion PDF $f_{\text{ion},0}(v, t > 0)$: in the vicinity of $v \sim \omega/k_j$, a H&C can form during resonance.



(b) A sketch of a typical set of values for the turbulent PDF $f_{\text{tur},j}(v, t = 0)$. The blue dotted line denotes a background top-hat function whose first velocity moment corresponds to the turbulent number density. The red dashed line denotes noise which does not contribute to the turbulent number density or the turbulent particle flux.

Figure 5.3: **Sketches highlighting features of f_{tur} and f_{ion}.** Both f_{tur} and f_{ion} have a constant number of particles.

where (5.14a) justifies the discrepancy between using αE and $\alpha * E$, and (5.14b) justifies exclusion of the spatially averaged free current. Condition (5.14a) is remedied by using a warm beam with low population ($v_t \gg 1$, $(1 - \eta) \ll 1$) and by examining modes near marginal stability, while (5.14b) is never satisfied; however, we find that the free current is dominated by terms $\sim v f_j \, \mathrm{d}v$, and therefore we deem this discrepancy to be acceptable. In principle, (5.14b) is violated, but we find that the induced mean current is very small. However, it is important to note that there are known physical limitations of this model.

Overall, for a single mode simulation after spectral decomposition and velocity Fourier transforms, our partial flows examined in D-dimensional Augmented Resonance Kinetic

solver (DARK) are given by:

$$\frac{\partial}{\partial t}\mathcal{F}_v[f_{\text{ion},j}] = -\nu s^2 \mathcal{F}_v[f_{\text{ion},j}] + k_j \frac{\partial}{\partial s}\mathcal{F}_v[f_{\text{ion},j}] - \mathrm{i}sE_j\mathcal{F}_v[f_{\text{ion},0}], \tag{5.15a}$$

$$\frac{\partial}{\partial t}\mathcal{F}_v[f_{\text{ion},0}] = -\nu s^2(\mathcal{F}_v[f_{\text{ion},0}] - \mathcal{F}_v[F_0]) - \frac{\mathrm{i}s}{4}\left[E_j^*\mathcal{F}_v[f_{\text{ion},j}] + \text{c.c.}\right], \tag{5.15b}$$

$$\mathcal{F}_v[f_{\text{tur},j}] = \mathcal{F}_v[f_{\text{tur},j}](s,t), \tag{5.15c}$$

$$\frac{\partial E_j}{\partial t} = -\int_{\mathbb{R}} \left[W \cdot (\mathcal{F}_v[f_{\text{ion},j}] + \mathcal{F}_v[f_{\text{tur},j}])\right] \mathrm{d}s - \frac{1}{2}\alpha_j E_j, \tag{5.15d}$$

$$E_0 = 0, \tag{5.15e}$$

where $W(s) \in \mathbb{I}$ is the complex formally divergent integral (which acts a removable singularity in the current):

$$W(s) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} v\mathrm{e}^{\mathrm{i}sv} \, \mathrm{d}v.$$

5.2.5 Energy content

The total energy density T of the system is given by (3.28). By taking the partial derivative with respect to t and integrating over all space afterwards, one finds via the Leibniz integral rule:

$$\frac{\mathrm{d}U}{\mathrm{d}t} \equiv -\int_{\mathbb{R}} \alpha E^2 \, \mathrm{d}x + \frac{L}{2}\int_{\mathbb{R}} \nu \frac{\partial^2}{\partial v^2}(f_0 - F_0)v^2 \, \mathrm{d}v.$$

Therefore, energy is injected into the system via Fokker-Planck collisions, and lost via α_j. If we examine the contribution over time from the first term, in the absence of three-wave mixing one finds via (5.15c):

$$\Delta U(t) = \int_0^{t'}\int_0^L \alpha E(t)^2 \, \mathrm{d}x \, \mathrm{d}t \approx \frac{L}{4}\sum_j \int_0^t \alpha_j |E_j(t')|^2 \, \mathrm{d}t', \tag{5.16}$$

where $\Delta U(t)$ is herein refered to as the total energy loss (defined here as positive by convention).

5.2.6 Seed electric field

We define E using three parts: a perturbation δE, a deterministic part $E_{\mathcal{S}}$, and a stochastic part $E_{\mathcal{N}}$:

$$E := \delta E + E_{\mathcal{S}} + E_{\mathcal{N}}. \tag{5.17}$$

One can show that (5.15d) now takes the form:

$$\frac{\partial}{\partial t}\delta E_j = -\int_{\mathbb{R}} W \cdot (\tilde{f}_{\text{ion},j} + \tilde{f}_{\text{tur},j})\ \mathrm{d}s - \frac{1}{2}\alpha_j \delta E_j + \mathcal{D}_j,$$

where the drive term $\mathcal{D}_j$ is given by:

$$\mathcal{D}_j \equiv -\left(\frac{\partial}{\partial t} + \frac{\alpha_j}{2}\right)(E_{\mathcal{S},j} + E_{\mathcal{N},j}).$$

We desire that the seed terms $E_\mathcal{S}$ and $E_\mathcal{N}$ exchange no free energy with the system. *Ad hoc* definitions which are independent of δE and $\{f_l\}$ achieve this (see Appendix C.2):

$$E_{\mathcal{S},j} = D_j(1-\epsilon_{E,j})\mathrm{e}^{\mathrm{i}\omega_s t} \quad ; \quad E_{\mathcal{N},j} = E_{\mathcal{N},j}(t),$$

where ω_s is the seed mode frequency, and $\epsilon_{E,j}$ allows us to define stochastic seeding of the electric field with wavenumber k_j; $\epsilon_{E,j} = 1$ is fully stochastic seeding, $\epsilon_{E,j} = 0$ is non-stochastic seeding. D_j is the typical amplitude of $(E_{\mathcal{S},j} + E_{\mathcal{N},j})$.

For $E_{\mathcal{N},j}(t)$, we use Gaussian noise, with a mean value of 0, and a domain of $[-\infty, \infty]$. The standard deviation of the values $E_{\mathcal{N},j}(t)$ takes is $\sigma_{E,j}$, which we treat as the typical amplitude. Accordingly, $\sigma_{E,j} = \epsilon_{E,j} D_j$.

In order for the seed field $(E_\mathcal{N} + E_\mathcal{S})$ not to directly interact with the plasma (and attenuate), the constituent frequencies must be much greater than the plasma frequency. If we utilise pseudorandom noise, it oscillates on a timescale comparable to the timestep. For $E_\mathcal{S}$, we have to enforce $\omega_s \gg \omega_{\text{pl.}}$ (see Figure 5.2). Physically, this corresponds to an externally launched wave in the plasma.

5.2.7 Turbulent distribution

The fast ions are modelled using a 1D BOT distribution (see Figure 2.1):

$$F_0(v) = \frac{1}{\sqrt{2\pi}}\left[\eta \exp\left(-\frac{v^2}{2}\right) + \frac{1-\eta}{v_t}\exp\left(-\frac{(v-v_b)^2}{2v_t^2}\right)\right], \tag{5.18}$$

such that $(1-\eta)$ is the fraction of particles in the beam, v_t is the beam width, and v_b is the beam velocity. These act as a suitable analogue for the slowing-down distribution one typically expects to observe in tokamaks. We define the turbulent population as the sum of a top-hat function and a noise term; the noise term is defined to yield a particle population of zero (see Figure 5.3b). This allows us to parametrically modify the stochasticity of f_{tur}

via noise without changing the total number of particles. After Fourier transforms:

$$\boxed{\begin{aligned}\mathcal{F}_v[f_{\text{ion}}](x,s,t) &= \tilde{F}_0(s) + \mathcal{F}_v[\Delta f_{\text{ion}}](x,s,t), && \text{(5.19a)}\\ \mathcal{F}_v[f_{\text{tur},j}](s,t) &= \frac{1}{\sqrt{2\pi}}\epsilon_{f,j}\,\text{sinc}\left(\frac{sL_v}{2}\right) + \mathcal{F}_v[\mathcal{N}_{f,j}], && \text{(5.19b)}\\ \mathcal{F}_v[f_{\text{tur},0}](s,t) &= 0, && \text{(5.19c)}\end{aligned}}$$

where $\mathcal{F}_v[\mathcal{N}_f](x,s,t)$ is a noise term, and L_v is the length of the domain of $F_0(v)$. Accordingly, the fraction of non-turbulent particles is $1/[1+\epsilon_{f,j}]$. We assume that broadband noise in v-space will still be broadband noise in s-space, weighted accordingly via Parseval's theorem; accordingly we relate amplitudes as:

$$\frac{|\mathcal{F}_v[\mathcal{N}_{f,j}]|}{|\mathcal{N}_{f,j}|} \approx \frac{L_v}{\sqrt{2\pi}}.$$

For $\mathcal{N}_{f,j}(v,t)$, we use raised cosine noise, with a mean value of 0. The codomain is given by:

$$\mathcal{N}_{f,j}(v,t) \in \left[-\sigma_{f,j}\sqrt{\frac{3\pi^2}{\pi^2-6}}, \sigma_{f,j}\sqrt{\frac{3\pi^2}{\pi^2-6}}\right].$$

The typical amplitude is equal to the standard deviation $\sigma_{f,j}$, and accordingly to force positive f_{tur} everywhere:

$$\sigma_{f,j} \leq \sigma_{f,j}^{(\max)} \quad ; \quad \sigma_{f,j}^{(\max)} = \frac{\epsilon_{f,j}}{L_v}\sqrt{\frac{\pi^2-6}{3\pi^2}}. \tag{5.20}$$

We also require for conservation of energy and particle number (at constant ϵ_f) that the 0^{th} and 2^{nd} moment of f_{tur} vanish. To enable this, we enforce that $\mathcal{F}_v[\mathcal{N}_f](x,s,t)$ has a real part that is odd, and an imaginary part that is even. One finds that the net energy content of the turbulent PDF is given by:

$$U_{\text{tur}} = \frac{1}{24}\epsilon_{f,j}L_v^3. \tag{5.21}$$

5.3 Stochastic lifetime of hole and clump

Here, we explore how small-scale random fluctuations in the mode amplitude can lead to a stochastic lifetime of the H&C.

5.3.1 Simulation setup

We use the DARK code (see Chapter 4) to examine this model. Our global parameters were selected to be:

$$\begin{gathered} v_b = 10 \quad ; \quad v_t = 4 \quad ; \quad \eta = 0.95 \\ v \in [-28, 88] \quad ; \quad N_v = 8192. \end{gathered} \tag{5.22}$$

For this sub-model, we consider a system with noise only in the electric field; that is $\mathcal{N}_f = 0$. We consider no particles in the turbulent population f_{tur}, such that the simulations reduce to single species.

We examine a single resonant mode in the system with $k_j = 0.15$, which we define as the fundamental eigennumber of the system (the length of the 1D box $L_x = 2\pi/k_j$). This not only approximately coincides with where $\partial F_0/\partial v$ is maximized on the tail of the BOT distribution, but also all of the higher harmonics resonate with the bulk particles, where they undergo strong Landau damping. While some studies have shown that strongly Landau damped modes can be nonlinearly unstable,[35,75] here we assume that this is not possible due to small initial mode amplitude; this allows us to justify the lack of three-wave coupling in the model.

Noise in the system is provided by using a pseudorandom number generator (PRNG) with a given seed value; by using such a method of noise generation, the results are reproducible.

The electric field is only seeded by noise; $D_j = 10^{-7}$ was used, with $\epsilon_{E,j} = 0$. A set of 2500 simulations were employed, allowing for 50 varying values of $t_{\text{NL}} := (\gamma_{j,L})^2/[(\omega_{j,(0)})^2\nu]$, each tested for 50 different initial seeds of the PRNG. k_j and ν were fixed to 0.150 and 10^{-7} respectively, with 50 values of α_j on the interval $[0.06, 0.158]$. The low value of ν justifies a relatively large timestep of $\Delta t = 0.1$.

The timestep was picked to be $\Delta t = 0.1$. This is small enough to allow for a reasonable frequency analysis without becoming too computationally expensive. Using a window size of 2000 timesteps, we obtain a frequency resolution $\Delta\omega_{\text{res.}} = \pi/100$ provided that $\Delta\omega < \Delta\omega_{\text{res.}}$ across the timeframe of the bin. Furthermore, the Fourier spectrum of $E_{\mathcal{N}}$ should be dominated by structures in the region $\omega \gg \omega_{\text{pl.}}$.

It is also a requirement for the code to adhere to the Courant-Friedrich-Lewy (CFL)[58] limit for spatial advection in s-space via the piecewise parabolic method (PPM)[57] routine; the number of v-points N_v and the domain of v adheres to the CFL limit.

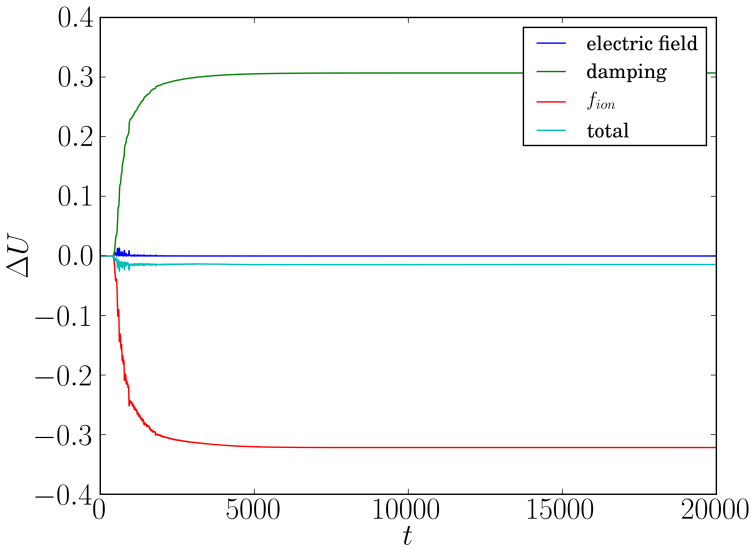


Figure 5.4: **Energy loss (ΔU) observed in a simulation from Section 5.3 with $\alpha_j = 0.6$.** Initially energy is approximately conserved, however the total energy in the system increases linearly in the nonlinear phase due to injective collisions.

5.3.2 Benchmarking

Energy is not conserved in these simulations; the Fokker-Planck collisions heat the fast ion PDF, always aiming to restore the energy content in the PDF to U_0:

$$U_0 \equiv \frac{1}{2}\int_{\mathbb{R}} F_0 v^2 \, \mathrm{d}v = \frac{1}{2}\left[\eta + (1-\eta)(v_b^2 + v_T^2)\right]. \tag{5.23}$$

For the purpose of benchmarking, we use a very simple model for symmetric mode flattening with a local population transfer $(f_0 - F_0) \sim -(v - v_0)\exp[-(v - v_0)^2]$ near resonance. This yields the corresponding energy flux from collisions in the weakly nonlinear regime:

$$\dot{U}_{\text{coll.}} \equiv \frac{1}{2}\int_{\mathbb{R}} \frac{\partial^2}{\partial v^2}(f_0 - F_0)v^2 \, \mathrm{d}v \sim -v_0\sqrt{\pi}. \tag{5.24}$$

We test the energy conservation by examining a simulation with the same parameters employed in Section 5.3, at $\alpha_j = 0.6$. As is shown in Figure 5.4, the total energy content in the system is roughly constant initially. As we approach the nonlinear phase, the energy lost from f_{ion} increases sharply. One expects this energy loss to be equal to that lost to the mode E_j, and damping α_j; the total energy injected into the system via ν at this time is very small.

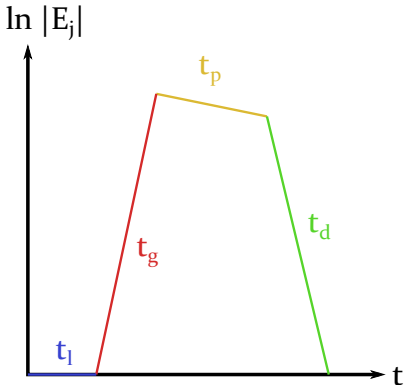


Figure 5.5: **Bursting mode near marginal stability**. Constituent regions t_l, t_g, t_p, and t_d are labelled, corresponding to the lag time, growth time, plateau time, and decay time respectively.

Once the PDF has suitably relaxed, the deficit in the energy content should asymptote to that lost via damping α_j, however we observe a small discrepancy in the energy content (roughly 0.45% of the total energy content). We believe that this discrepancy is due to approximations made regarding higher harmonics of $f_{\text{ion},j}$ and E_j, and deem this discrepancy to be negligible for the single burst simulations examined here.

5.3.3 Burst characterisation

To characterise the behaviour of a single bursting event, a set of simulations were used to produce data for the length of four temporal regions: lag, growth, plateau, and decay. Each of these regions are labelled in Figure 5.5 for a sketch burst.

For an overall burst time $t_b = t_g + t_p + t_d$, the constituent times can fluctuate (functional dependences determined from simulation). The theoretical maximum and minimum amplitudes were used to create a fit routine, allowing one to acquire from the mode amplitude $|E_j|$ the constituent times as a function of the parameters k_j, α_j, and ν. We find that these times are not deterministic, but instead are stochastic:

$$\begin{aligned} &\text{lag :} && t_l = t_{l,(0)} && + \delta t_l, \\ &\text{growth:} && t_g = t_{g,(0)}, && \\ &\text{plateau:} && t_p = t_{p,(0)} && + \delta t_p, \\ &\text{decay:} && t_d = t_{d,(0)} && + \delta t_d, \end{aligned} \tag{5.25}$$

where $\{\delta t_X\}$ denote stochastic terms. The fluctuation of each of the times has a well defined mean and standard deviation, and are analysed such that the mean is 0; here, we find that the lag, plateau, and decay times are highly stochastic. We find that at small

t_{NL} the mean plateau time $t_{p,(0)}$ is roughly constant, while the mean decay time $t_{d,(0)}$ is a linear function of t_{NL}, while the mean lag time $t_{l,(0)}$ is a nonlinear function of t_{NL}.

5.3.4 Linear phase

When a single mode bursts, the electric field grows linearly via inverse Landau damping, provided that the mode lies on the positive slope of the beam. For $\nu = E_{\mathcal{S}} = E_{\mathcal{N}} = 0$, the frequency and growth rate in the linear phase are determined by (5.8). The overall linear growth rate for the j^{th} mode is given by:

$$\gamma_j(k_j, \alpha_j) = \gamma_{j,L}(k_j, \alpha_j) - \frac{\alpha_j}{2}, \tag{5.26}$$

where the unperturbed linear growth rate $\gamma_{j,L}$ is equivalent to γ_j in the absence of dissipation.

The frequency $\omega_{j,(0)} = \omega_j(t = 0)$ is the initial eigenfrequency of the j^{th} mode. If one solves (5.15d) for negligible current, on average:

$$|\delta E_j|_{\min} \approx \frac{D_j}{\alpha_j/2}\left[1 - \mathrm{e}^{-\alpha_j \Delta t/2}\right] + \mathcal{O}\left(\mathrm{e}^{-(\alpha_j \Delta t)^2}\right). \tag{5.27}$$

The simulated noise is static over a timestep, leading to an error which manifests as the term $\sim \exp(-\alpha_j \Delta t/2)$.

One can interpret this physically as a finite bandwidth for the noise; we expect ITG (ion temperature gradient) turbulence spectral frequencies to typically be much lower than the plasma frequency,[63] however here we examine noise with a frequency spectrum that is typically above the plasma frequency, corresponding to high frequency turbulence. For both ITG and the synthetic turbulence we use, direct wave coupling between the turbulence and particles does not occur. As such, we believe that heuristically the results should be the same, as in both cases the Fourier spectrum of the turbulence is broadband, and not strongly peaked near the plasma frequency. The peak value for the electric field is the nonlinear saturation point, which can be approximated by the following value:[5]

$$|E_j|_{\max} \approx (\gamma_{j,L})^2. \tag{5.28}$$

Accordingly, as we expect exponential growth in the linear phase, the total time spent in the linear phase is given by:

$$(t_{g,(0)})_{\mathrm{theory}} \approx \frac{2}{\gamma_j} \ln\left[\frac{\alpha_j \gamma_{j,L}}{2D_j[1 - \exp(-\alpha_j \Delta t/2)]}\right]. \tag{5.29}$$

Simulations were found to strongly agree with this value; we find $t_{g,(0)} = (-91.0 \pm 8.1) + (1.15 \pm 0.01) \cdot (t_{g,(0)})_{\text{theory}}$. The quantity t_g does not appear to be stochastic; fluctuations in the value of t_g as a function of PRNG seed are typically about 2 or 3 orders of magnitude lower than the mean value $t_{g,(0)}$. One finds that the accuracy improves at low α_j; we find that this is in accordance with theory, as our value for $(t_{g,(0)})_{\text{theory}}$ assumes slow damping (small $\alpha_j \Delta t$).

At high α_j, the expected growth time grows logarithmically until $\gamma_{j,L}$ dominates:

$$\lim_{\alpha_j \to \infty} (t_{g,(0)})_{\text{theory}} = \frac{2}{\gamma_j} \left(\ln \left[\frac{\gamma_{j,L}}{2D_j} \right] + \ln \alpha_j \right).$$

Interestingly, one finds that even if the linear growth rate is non-zero, the seed electric field can prevent the mode from growing. This can be shown by setting the growth time to zero and solving for $\gamma_{j,L}$:

$$(\gamma_j)_{\min} \approx \frac{2D_j}{\alpha_j} - \frac{\alpha_j}{2}.$$

If the linear growth rate is below this minimum, the seed electric field quenches the mode before it has a chance to burst; in order to preserve the true meaning of the linear growth rate, one should impose a limit on D_j when using high α_j:

$$D_j \leq \frac{\alpha_j^2}{4}.$$

This hard limit on the noise level allows one to properly investigate simulations close to the linear stability boundary.

5.3.5 Nonlinear phase

Once the mode reaches the nonlinear saturation point, resonance broadening occurs, flattening the background PDF $f_{\text{ion},0}$ in the close vicinity of the resonant phase velocity $\omega_{j,(0)}/k_j$. If the mode is marginally unstable, to the extent that:[5]

$$\frac{\alpha_j}{2} > 0.2\gamma_{j,L}, \tag{5.30}$$

a phase space bifurcation in the form of a H&C form on $f_{\text{ion},0}$.

For diffusive Fokker-Planck collisions, the time spent in the plateau and decay regions has previously been shown to be a function of the timescale[5] t_{NL}. We take first order

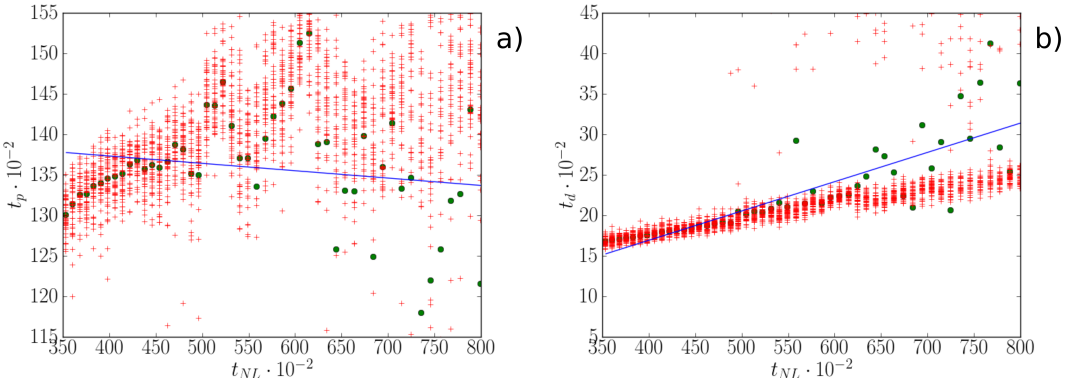


Figure 5.6: **Values for characteristic times associated with mode bursting.** Graphs showing calculated values for $t_p(t_{\mathrm{NL}})$ and $t_d(t_{\mathrm{NL}})$ from simulations in Section 5.3. Observed values ('+') and mean values ('•') from fit routines. Blue solid line ('-') denotes a linear best fit to mean values. **a**): Plateau time t_p: approximately constant as a function of t_{NL}, with a relative stochastic fluctuation $\sigma_p/t_{p,(0)} \sim 10^{-2}$. **b**): Decay time t_d: increases linearly as a function of t_{NL}, with a relative stochastic fluctuation $\sigma_d/t_{d,(0)} \sim 10^{-1}$.

Taylor expansions in t_{NL} as follows:

$$t_{X,(0)} = a_X + b_X t_{\mathrm{NL}} + \mathcal{O}(t_{\mathrm{NL}}^2) = \frac{1}{50}\sum_{\mathrm{seed}} t_X, \tag{5.31}$$

where we sum over 50 PRNG seeds. From the simulations, $t_{l,(0)}$ appears to be a nonlinear function of t_{NL} while $t_{d,(0)}$ appears to be linear.

The mean plateau time $t_{p,(0)}$ appears to be constant at low t_{NL}; however at high t_{NL}, large error in the linear fit reduces our ability to determine the mean time.

We find that $a_d = (2.56 \pm 2.01) \times 10^2$, $b_d = (3.62 \pm 0.35) \times 10^{-2}$, and $a_p = (1.41 \pm 0.04) \times 10^4$. One should note that the errors here are errors in the linear fit to mean values; they represent confidence in the functional dependence on t_{NL}, not the stochasticity. We find that $\mathcal{O}(b_p) = 10^{-3}$, allowing us to state $t_{p,(0)} \approx a_p$ for $t_{\mathrm{NL}} \ll 10^6$.

5.3.6 Burst stochasticity

We once again take first order Taylor expansions in t_{NL}, but now examining the standard deviations in $\{t_X\}$:

$$\sigma_X = c_X + d_X t_{\mathrm{NL}} + \mathcal{O}(t_{\mathrm{NL}}^2) = \frac{1}{50}\sum_{\mathrm{seed}} \delta t_X^2, \tag{5.32}$$

where we once again sum over 50 PRNG seeds. We find that generally, σ_X does not appear to be a function of t_{NL}. The lag time t_l is strongly stochastic, with $\sigma_l/t_{l,(0)} \sim 10^0$.

This is in accordance with theory, as at very low amplitude, $\partial|E_j|/\partial t$ is strongly dependent on the noise term, which is stochastic.

The plateau exists while the H&C have a static population of particles, and therefore, once the phase space structures dissipate, the mode drops significantly in amplitude. The time spent in this region, t_p, is stochastic; as is shown in Figure 5.6a, the relative fluctuation $\sigma_p/t_{p,(0)} \sim 10^{-2}$. We conclude that the point at which this occurs is stochastic, leading to a stochastic lifetime of the H&C.

The growth and decay times t_g and t_d are defined by the minimum and maximum mode amplitude. Therefore, any stochastic behaviour reflects fluctuation in the growth rate and decay rate of the mode. We find that $\sigma_g/t_{g,(0)}$ is negligible, implying that mode growth is not stochastic, as one might expect. However, we find that $\sigma_d/t_{d,(0)} \sim 10^{-1}$, implying a large fluctuation in the decay rate of the mode (see Figure 5.6b).

5.4 Stochastic suppression of hole and clump

The effects of random, small-scale phenomena have been previously examined in the literature:[5,28,59,66] mechanisms such as pitch-angle scattering can destroy H&Cs. However, the effect of microturbulence on the evolution of a bursting mode in its nonlinear phase is relatively unexplored. Recent work by Duarte *et al.*[64,74] proposes that enhanced stochasticity in resonant particle dynamics, in the form of fast ion microturbulence, can be a key mechanism for chirping suppression in several tokamak scenarios. The prediction stimulated dedicated experiments on DIII-D by Van Zeeland *et al.*[49] with negative triangularity, known for suppressing drift-like instabilities. The experiments have shown a clear correlation between chirping emergence and scenarios with very low turbulent activity.

Here, we consider a case with a wholly non-stochastic electric field ($\epsilon_E = 0$), and examined the effect of a stochastic PDF f_{tur} on mode chirping.

5.4.1 Simulation setup

We use the DARK code (see Chapter 4) to examine this model. Our global parameters were selected to be:

$$\begin{gathered} v_b = 10 \quad ; \quad v_t = 4 \quad ; \quad \eta = 0.95 \\ v \in [-28, 88] \quad ; \quad N_v = 8192. \end{gathered} \tag{5.33}$$

Here, we once again examine $k_j = 0.15$ (see Section 5.3.1 for justification). We use a seed electric field such that the deterministic part oscillates at $\omega_s = 2.0$.

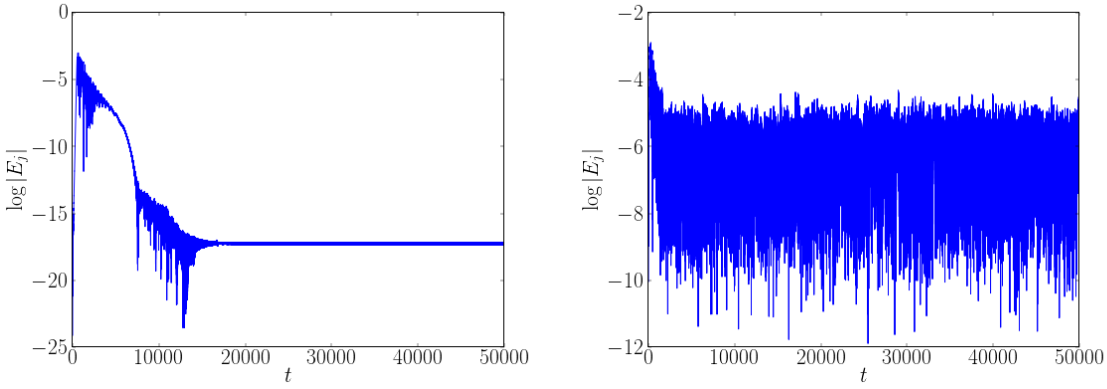


Figure 5.7: **Plots of $\ln|E_j|(t)$ from linearly unstable simulations in Section 5.4 with low effective collisionality.** The mode amplitude $|E_j|$ undergoes a single bursting event at $t \approx 0$, corresponding to mode chirping. Left plot (with deterministic field) shows $\nu = 10^{-6}$, $R_j = 0$; right plot (with stochastic field) shows $\nu = 10^{-5}$, $R_j = -2$.

One finds that in cases of high collisionality, we must enforce $\nu \lessapprox \Delta v^2/\Delta t$ to avoid numerical inaccuracies where collisions dissipate structures much faster than the timestep. We fix the linear growth rate to $\gamma_j = \alpha_j/2 = 0.0534$, to reduce the parameter space while still allowing H&C formation (see (5.30)). We fix D at 10^{-7} and $1/(1+\epsilon_{f,j}) = 98\%$ to investigate a small electric field drive and a small turbulent population. We define the relative stochasticity as $R_j \equiv \log_{10}(\sigma_{f,j}/\sigma_{f,j}^{(\max)})$.

In theory, f_{tur} can affect δE via (5.15d). As $(W \cdot f_{\text{tur}}) \in \mathbb{I}$, one can see that real stochastic noise will produce an imaginary stochastic term in (5.15d), which will lead to a scrambling of the phase of coherent structures with wavenumber k_j. Consequently, the coherence of H&Cs can be destroyed, allowing them to rapidly dissipate.

In contrast with $E_{\mathcal{N}}$, as f_{tur} exchanges no energy on pseudorandom time average with E (see Section 4.4.6), it can instantaneously create perturbations in E which cannot be induced by finite $E_{\mathcal{N}}$.

5.4.2 Stochastic and deterministic simulations

First, we show 3 simulations with varying collisionality $\log_{10}\nu = \{-2, -5, -6\}$, and no noise ($R_j = 0$). A timestep of $\Delta t = 0.01$ to allow us safe exploration of $\nu \sim 10^{-2}$. As is shown in Figures 5.7 to 5.9, high collisionality ($\nu = 10^{-2}$) suppresses mode chirping, medium collisionality ($\nu = 10^{-5}$) allows for repeated bursting, and low collisionality ($\nu = 10^{-6}$) allows for only a single event.

Next, we highlight the effect of $\sigma_{f,j}$ on the asymptotic behaviour of the mode in Figures

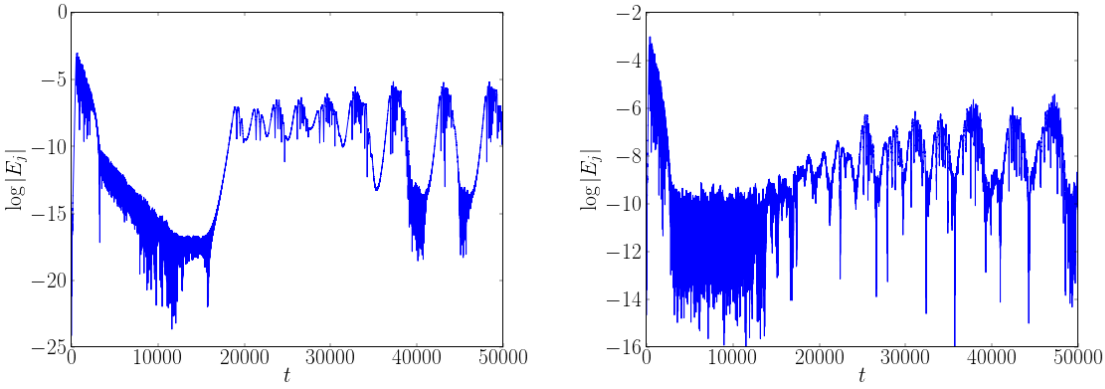


Figure 5.8: **Plots of $\ln|E_j|(t)$ from linearly unstable simulations in Section 5.4 with medium effective collisionality.** $|E_j|$ undergoes repeated bursting events ($t \approx \{0, 22000, 24500, \dots\}$), each corresponding to mode chirping. Left plot (with deterministic field) shows $\nu = 10^{-5}$, $R_j = 0$; right plot (with stochastic field) shows $\nu = 10^{-5}$, $R_j = -4$.

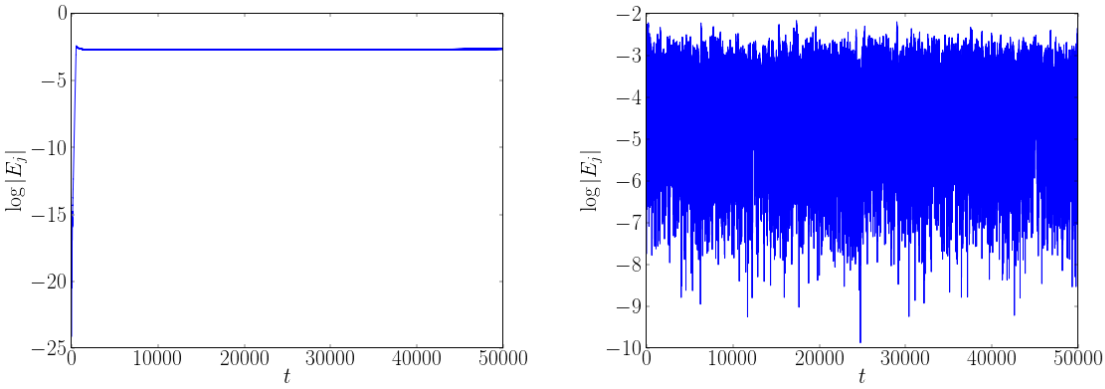


Figure 5.9: **Plots of $\ln|E_j|(t)$ from linearly unstable simulations in Section 5.4 with high effective collisionality.** High effective collisionality: the system does not undergo mode chirping. Left plot (with deterministic field) shows $\nu = 10^{-2}$, $R_j = 0$; right plot (with stochastic field) shows $\nu = 10^{-5}$, $R_j = -1$.

5.7 to 5.9, and Figures 5.10 to 5.12; we show results from 3 simulations with $\nu = 10^{-5}$ using a coarser timestep of $\Delta t = 0.1$, and $R_j = \{-1, -2, -4\}$. For low stochasticity ($R_j = -4$), the effective collisionality increases; we still observe repeated bursting, however the period between repeated bursts is characteristic of simulations with $\nu \sim 10^{-4}$. Here, f_{tur} affects the stability of f_{ion}, but does not make the fast ion population stable; as is shown in Figure 5.10, the initial H&C is undisturbed, and repeated bursts still occur.

For medium stochasticity ($R_j = -2$), we find that the repeated bursts are suppressed. We give two equivalent explanations: the electric field produced by f_{tur} approaches the

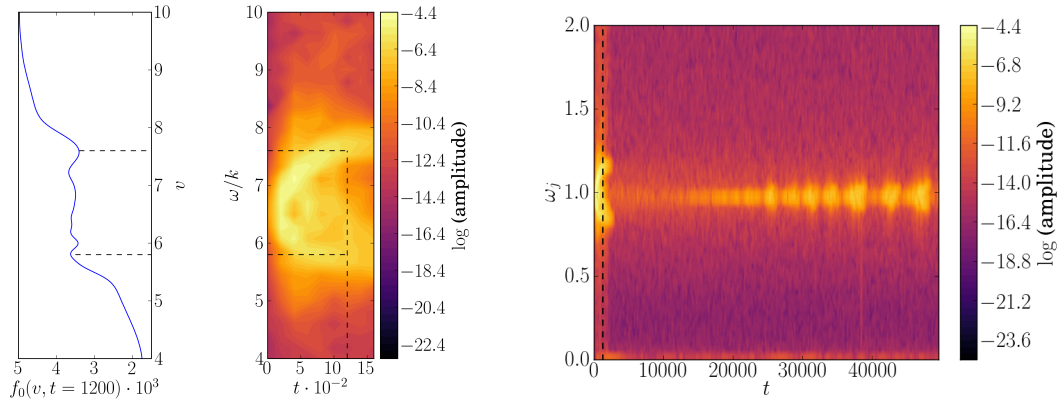


Figure 5.10: **Suppression of particle distribution functions from simulations in Section 5.4 with low relative stochasticity.** Black dotted lines in the bottom plots show the existence (or lack of) H&C at $t = 1200$. $R_j = -4$; noise in f_{tur} produces an electric field, but repeated bursting still occurs.

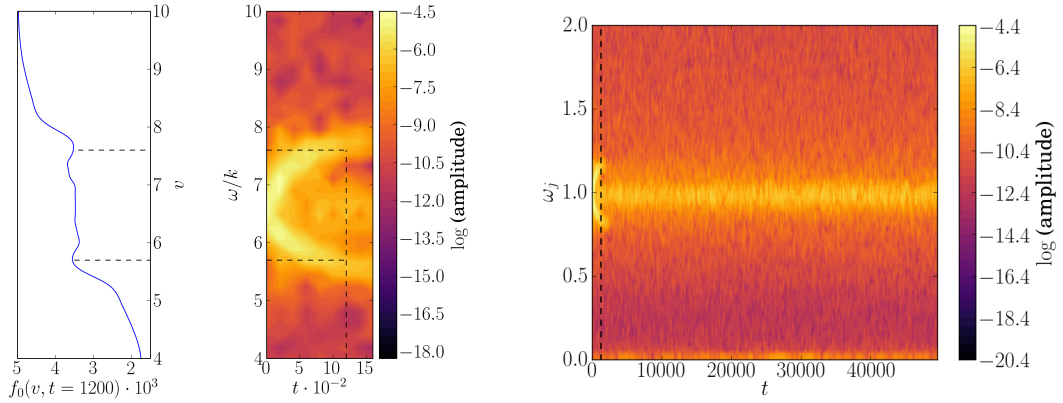


Figure 5.11: **Suppression of particle distribution functions from simulations in Section 5.4 with medium effective collisionality.** Black dotted lines in the bottom plots show the existence (or lack of) H&C at $t = 1200$. $R_j = -2$; noise in f_{tur} produces an electric field which prevents repeated bursts from occuring, but does not prevent the mode from initially chirping.

maximum amplitude of the repeated bursts, saturating them. As is shown in Figure 5.11, the initial H&C still exists, but repeated bursts do not occur.

As we increase to high stochasticity ($R_j = -1$), we find that the initial burst is suppressed. We find that this is when the electric field produced by f_{tur} has an amplitude close to the nonlinear saturation point; at this point, the mode does not resonate, even though it is unstable. Alternatively, f_{tur} prevents the mode from bursting by increasing the effective collisionality; as is shown in Figure 5.12, the electric field produced by f_{tur}

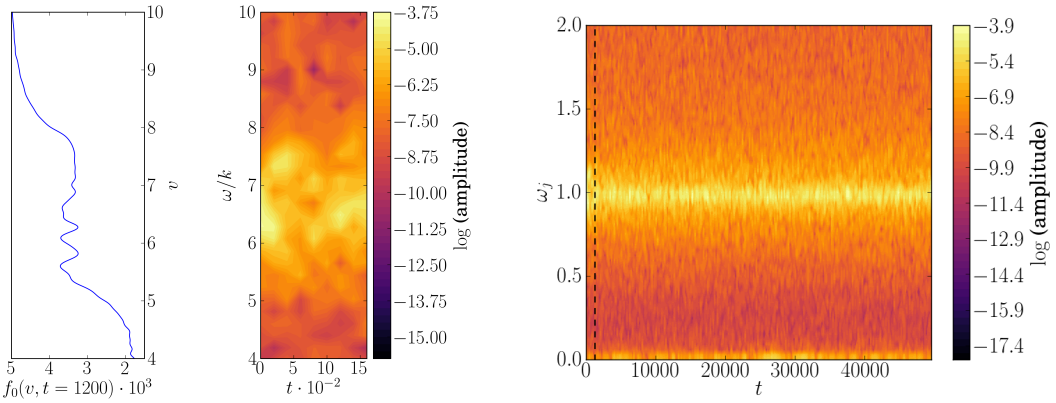


Figure 5.12: **Suppression of particle distribution functions from simulations in Section 5.4 with high effective collisionality.** Black dotted lines in the bottom plots show the existence (or lack of) H&C at $t = 1200$. $R_j = -1$; noise in f_{tur} produces an electric field which prevents a H&C from forming; the system is nonlinearly stable.

creates features on the spatially averaged PDF $f_0(v) = f_{\text{ion},0}(v, t = 1200)$ which affect the H&C.

5.5 Hole and clump destabilisation

In this section, I move away from the stochastic behaviour examined in Section 5.4, and instead examine wholly deterministic scenarios. H&Cs have been shown to form when a marginally unstable mode approaches the nonlinear phase of its evolution, leading to an adiabatic theory of H&C evolution.[5] Particles in the H&C regions resonate with an electrostatic field in Berk-Breizman models - under the adiabatic approximation, the resonating particles move through phase space, allowing for frequency chirping via nonlinear modification of the resonant velocity.[28,59] As they do so, they modify the local linear stability of resonant waves in the system by perturbing the local PDF.

Here, I investigate mode-mode destabilisation via phase space H&Cs in a 1D electrostatic system. By analysing the shape of the PDF, I later show that mode-mode destabilisation only occurs for modes that are sufficiently close together in k-space. I discuss the possibility for such a scenario to exist in tokamaks, and whether this is a suitable candidate for the origin of ALEs In this section, I do not include a stochastic population as such, here:

$$f := f_{\text{ion}}.$$

I herein refer to an initially chirping mode as the **master** mode (k_m), and a mode varying from linearly stable to linearly unstable as the **slave** mode ($k_{\text{sl.}}$). Depending on the location in parameter space, the slave mode and master mode may be able to destabilise each other.

5.5.1 Antisymmetric ansatz

In the construction of this theory of H&C destabilisation, I make a few justified assumptions.

As the primary ansatz, I assume that the H&C corresponding to the master mode form as antisymmetric structures of equal size and shape, that is to say:

$$\Delta f_0(v_m^+, t) = -\Delta f_0(v_m^-, t),$$

where $\Delta f_0 \equiv f_0(x, v, t = t_0) - f(x, v, t)$ is the spatially averaged perturbation to the PDF, and $v_m^{\pm}$ are the resonant velocities corresponding to the hole (+) and clump (-) prior to resonance broadening. The corresponding mode frequency of each branch is given by:

$$\omega_m^{\pm}(t) = \omega_{m,(0)} \pm \Delta\omega_m(t).$$

The initial frequency $\omega_{m,(0)}$ is the frequency of the mode prior to H&C formation, given by solving (5.8) for $t_0 = 0$. The frequency chirp $\Delta\omega_m(t)$ is predicted by Berk *et al.* for early times:[5]

$$\Delta\omega_m(t) = \frac{16}{3\sqrt{3}\pi^2}\gamma_{m,L}\sqrt{\alpha_j(t - t_m)}, \tag{5.34}$$

where t_m is the time at which the H&C form. The term $\gamma_{m,L}$ is equal to the growth rate of the mode in the absence of α_m damping, but is otherwise given by:

$$\gamma_{m,L} = \gamma_m + \frac{\alpha_m}{2}.$$

The ansatz is justified by observation of H&Cs in other simulations (see Section 5.4). However, this model is not valid for large $\partial^2 f_0/\partial v^2$. Particles resonating with the electric field do so by coupling to a BGK mode, and therefore one finds that individual particles have an energy ϵ_0 given by (after normalisation):

$$\epsilon_0 = \phi(x, t) + \frac{1}{2}(v - v_m)^2, \tag{5.35}$$

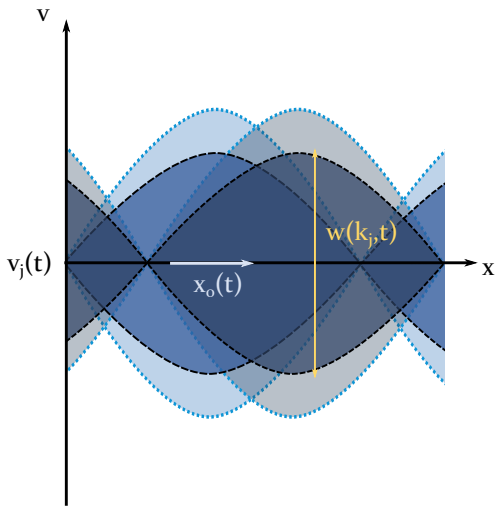


Figure 5.13: **A sketch of orbit affine parameter $\epsilon(x, v, t)$ at two separate times in a single mode system with an electric potential travelling with wave speed $v_j(t)$ relative to the observer frame.** Dotted lines denote contours of constant ϵ at an earlier time, while solid lines show contours of constant ϵ at a later time. The width λ of the phase space structure is dependent on the wavenumber of the Bernstein-Green-Kruskal (BGK) mode, and is taken to be the maximum width of the contour where $\epsilon = 0$ (the separatrix).

where $\phi(x, t)$ is the electric potential, and v_m is the phase velocity of the resonant mode. In this sense, I consider only conservative orbits (see Chapter 3 for other forms of the orbit affine parameter, ϵ). For the case where only the master mode has a non-zero amplitude, the electric field has a well-defined zero at a point $x_0(v_m, t) = x_0(v_m, 0) - \omega_j t/k_j$; via gauge freedom one can select $x_0(0) = 0$. One finds that particles with $\epsilon_0 < 0$ are bound within the BGK structure, while particles with $\epsilon_0 > 0$ lie outside of the corresponding separatrix in $x - v$ phase space. To first-order in the island width in velocity λ, bound particles resonate with the particle, while free-streaming particles do not (see Figure 5.13).

Here, we initially consider a single mode, and so by solving at the separatrix, one finds for $\lambda(k_m, t)$:

$$0 = |\phi_m(t)| + \frac{1}{2}(v - v_m)^2\,; \quad \therefore\ \lambda(k_m, t) = 2\sqrt{\frac{2|E_m|(t)}{k_m}}.$$

One should note that the electric field amplitude $|E_m| \equiv k_m|\phi_m|$ by virtue of:

$$\phi(x, t) = \int_{x_0}^{x} E(x', t)\,\mathrm{d}x',$$

where we have defined $\phi(x_0, t) = 0$. During H&C formation, I assume that the energy in the electric field is conserved, and therefore:

$$|E_m| = |E_m^{(+)}| + |E_m^{(-)}|,$$

where the indices $\pm$ refer to the mode amplitudes associated with the H&C branches $\omega_j^{\pm}$ respectively. This is assumption is justified as H&C formation occurs near the nonlinear saturation limit for a resonant mode, when $|E_j|$ approximately plateaus.

At $t = t_m$, the two branches have the same frequency ($\omega_{m,(0)}$). Next, we consider the H&C as two superposed BGK modes lying at the same point in phase space; as the mode undergoes chirping, we see the H&C move apart, resolving as two distinct structures. Overall, I consider the system to be degenerate in frequency for $t \leq t_m$, with both distinct mode branches manifesting with the same frequency and wavenumber.

As a result of the antisymmetry ansatz, I assume that the two structures have the same width, and therefore contain the same amount of electric field energy. This therefore assumes that $|E_m^{\pm}| \equiv |E_m|/2$, and so for the H&C:

$$\lambda(k_m, t) \approx 4\sqrt{\frac{|E_m^{\pm}|(t)}{k_m}}, \tag{5.36}$$

where λ corresponds to the H&C width in velocity.

5.5.2 Cosine perturbation model

With regards to the shape of the H&C, I demand Dirichlet boundary conditions that preserve the continuity of f_0 at the edges of the H&Cs:

$$\Delta f_0(v_m^{(+)} \pm \lambda/2, t) = 0 \quad ; \quad \Delta f_0(v_m^{(-)} \pm \lambda/2, t) = 0.$$

This allows for a well defined population of particles at each point in phase space. I also demand that the trough and peak of the H&C respectively are fixed in accordance with adiabaticity. Then, I note that the energy of trapped particles in BGK modes is symmetric with respect to the phase velocity v_m as given by (5.35):

$$\epsilon_0(\phi, v > v_m) = \epsilon_0(\phi, v < v_m).$$

As BGK modes are expressable solely as a function of energy $f[\epsilon_0]$,[7] I determine that the structure must also be symmetric about the resonance velocity, i.e.:

$$f(\epsilon(v > v_m)) \equiv f(\epsilon(v < v_m)). \tag{5.37}$$

Using these two conditions, we choose a cosine model for the H&C, due to the simple analytic form it generates for $\partial f_0/\partial v$. It is important to note that this shape does violate Neumann boundary conditions at the edge of the structures ($\partial f_0/\partial v$ is not continuous) however $\partial f_0/\partial v$ is continuous everywhere inside the structures. For $t_0 = 0$ (such that one examines the initial PDF):

$$\Delta f_0(v,t) \approx \left[F_0(v_{m,(0)}) - F_0(v_m)\right] \cdot \cos\left[\frac{\pi(v - v_m)}{\lambda}\right],$$

where $(v - v_m) \in [-\lambda/2, \lambda/2]$, and $\Delta f_0(v,t)$ is zero everywhere else. The phase velocity $v_m(t)$ chirps up or down depending on whether we examine the hole or clump respectively, such that $v_{m,(0)} \equiv v_m(t = t_0)$.

As a result, one finds that at the edge of the hole or clump, for $t > t_m$ the gradient is given by:

$$\frac{\partial}{\partial v}\Delta f_0\left(v_m \pm \frac{\lambda}{2}, t\right) = \frac{\pi}{\lambda}\left[F_0(v_{m,(0)}) - F_0(v_m)\right]. \tag{5.38}$$

The linear complex dispersion relation (5.8) considers the initial PDF as well as the perturbation to it at $t = t_0$. Therefore, the linear stability of a mode is parametrically affected by the mode amplitude of a chirping mode.

5.6 Two-mode destabilisation

In this section, I utilise computation techniques as well as the semi-analytical model from the previous chapter to examine two-mode destabilisation of H&Cs.

The perturbation Δf_0 modifies the local gradient of the spatially averaged PDF f_0, allowing a H&C to destabilise a mode that was previously linearly stable by instantaneously increasing the linear growth rate.

We can now illustrate the approximate form of the growth rate for the slave mode; using (5.9), (5.36) and (5.38), for a slave mode with a resonant velocity $v_{\text{sl.}}$ at the edge of a hole or clump being destabilised by a chirping mode with initial resonant velocity $v_{m,(0)}$:

$$\gamma_{\text{sl.}} \sim -\frac{\alpha_{\text{sl.}}}{2} + \frac{\pi\omega_{\text{sl.}}}{2k_{\text{sl.}}^2}\left.\frac{\partial F_0}{\partial v}\right|_{v=v_{\text{sl.}},t=t_0} + \frac{\pi^2\omega_{\text{sl.}}}{8k_{\text{sl.}}^2}\sqrt{\frac{k_m}{|E_m^{\pm}|}}[F_0(v_{m,(0)}) - F_0(v_{\text{sl.}})], \tag{5.39}$$

where t_0 is now the time taken for the chirping mode to reach the second mode. In reality the resonant velocity is modified by the PDF itself, as the linear complex dispersion relation is also dependent on the PDF itself. Therefore, in later analysis I substitute (5.38) into (5.8) and solve directly to give a more accurate result.

However, this approximate form allows one to predict two behaviours: $\gamma_{\text{sl.}}$ should decrease linearly with $\alpha_{\text{sl.}}$, and if the hole or clump dies before reaching the second mode, the final term is zero, leading to a sharp change in $\gamma_{\text{sl.}}$.

5.6.1 Simulation setup

I use the DARK code (see Chapter 4) to examine this model. The global parameters were selected to be:

$$\begin{gathered} v_b = 10 \quad ; \quad v_t = 4 \quad ; \quad \eta = 0.95 \\ v \in [-28, 88] \quad ; \quad N_v = 8192. \end{gathered} \tag{5.40}$$

For these simulations, I take $E_{\mathcal{N}} = 0$ and $\mathcal{N}_f = 0$, such that the simulations are fully deterministic. I use a timestep $\Delta t = 0.1$ (see Section 5.4.1 for justification).

A set of simulations were run for simulation time $40000\,\omega_{\text{pl.}}^{-1}$. The evolution of a hole was examined for a master mode fixed at $k_m = 0.200$, $\alpha_m = 0.034394$. This gives a linear growth rate $\gamma_m = \alpha_m/2$, which removes one dimension from the parameter space, and satisfies the inequality given by Berk *et al.*:[5]

$$0.4\gamma_{j,L} < \alpha_j.$$

One preliminary simulation was computed with only the master mode ($N_k = 1$) to allow for measurement of the lifetime of the hole, t_h. For $t > t_h$, the hole is rapidly dissipated, until the structure cannot affect the slave mode. From (5.34), the maximum frequency chirp is given when considering this lifetime:

$$\Delta\omega_m(t_h) = \frac{16}{3\sqrt{3}\pi^2}\alpha_m\sqrt{\alpha_m t_h}. \tag{5.41}$$

The maximum distance travelled in v-space is therefore given by $\Delta v_m(t_h) = \Delta\omega_m(t_h)/k_m$. Let us approximate for now that $\omega_m \sim \omega_{\text{sl.}} \sim 1$ (as $\omega_{\text{pl.}}$ is normalised to 1). Then, if we take a rough estimate with destabilisation occuring when $v_m(t_h) = v_s$:

$$\Delta k_{\text{est}} \sim \frac{\Delta\omega_m}{1 + \Delta\omega_m} k_m, \tag{5.42}$$

where $\Delta k := k_m - k_{\text{sl.}}$, and Δk_{est} denotes an estimate of the maximum distance in k-space where the mode-mode interaction still occurs.

I found that for $\nu = 10^{-7}$, the preliminary simulation had $t_h = 1.2 \times 10^4$, in turn giving an approximation of $\Delta k_{\text{est}} \sim 0.036$; this mode spacing corresponds to a maximum distance in v-space that is still on the positive slope of the beam ($v_h \sim 6.1$).

This allows one to construct a parameter space scan with the master mode and slave mode ($N_k = 2$), taking 100 values of $k_{\text{sl.}}$ on the interval $[0.100, 0.199]$ and 80 values of $\alpha_{\text{sl.}}$ on the interval $[0.1, 0.9]$ to form a set of 9000 simulations. This parameter space allows ones to examine initially linearly stable and unstable modes, and allows one to examine modes with $\Delta k < \Delta k_{\text{est}}$, and $\Delta k > \Delta k_{\text{est}}$; the simulations with $\Delta k \lessapprox \Delta k_{\text{est}}$ are expected to demonstrate mode-mode destabilisation.

A timestep of $\Delta t = 0.1$ was used to adhere to the CFL limit on s-space advection.[58] The length of the box is set to $L = 2\pi/0.001$ to enable high-fidelity scanning in $k_{\text{sl.}}$. A seed mode amplitude $D_j = 10^{-7}$ enables mode growth while giving a contribution to E lower than the typical nonlinear saturation limit ($\sim \text{e}^{-6}$).

5.6.2 Characterisation

For frequency analysis, a short-time Fourier transform (STFT) was used on the electric field components $\{E_j(t)\}$ to produce frequency spectra (see Section 6.4.1 for more information regarding the STFT procedure). I use a Hanning window, with a size of 8000 timesteps; this allows for a frequency resolution of $\Delta\omega = \pi/400$, as $\delta\omega_1 < \Delta\omega$ across the timeframe of the bin. I found that this window size suitably optimises the balance between frequency resolution and time resolution.

Here, I have recognised 6 key characterisations for discussion of the behaviour of each mode (see Figure 5.14): quiescence, fixed-frequency, up- and down-chirping, up-chirping, down-chirping, and repeated up- and down-chirping. Figure 5.15 contains four plots showing the different mode characterisations assigned to each simulation as a function of damping parameter α_j and mode spacing Δk.

Repeated and non-repeated characterisations for up-chirping and down-chirping are very similar; they are distinguished by examining whether a cluster of chirping events occurs after an initial cluster of events.

Each simulation was manually characterised by viewing $\omega(t)$ as given by the STFT routine, and $\ln|\delta E_j|(t)$ as outputted by the DARK code to produce a validation data set, containing 4450 simulations in the parameter subspace $\Delta k \in [0.0, 0.05]$, $\{\alpha_{\text{sl.}}\} \in [0.1, 0.89]$. Far from the characterisation boundaries, the simulations produce spectrograms which are 'typical' of a given characterisation; manual characterisation from these regions was used to construct a training data set for a machine learning (ML) algorithm.

As is later extended in Chapter 6 in the form of the ML framework Experimental Resonant Instability Correlation Studies on NSTX (ERICSON), the software `scikit-learn`

Characterisation	**E**-field	Freq. behaviour
● (quiescence)	$\frac{\partial}{\partial t}\ln\lvert\delta E_j\rvert = 0\,\forall\, t$	$\omega_j = 0$
■ (fixed-frequency)	$\lim_{t\to 0}\frac{\partial}{\partial t}(\ln\lvert\delta E_j\rvert) \approx \text{const.}$	$\omega_j \approx \omega_{j,(0)}\,\forall\, t$
★ (up- and down-chirping)	$\lim_{t\to 0}\frac{\partial}{\partial t}(\ln\lvert\delta E_j\rvert) \approx \text{const.}$	$\omega_j = \omega_{j,(0)} \pm \Delta\omega^{\pm}\,\forall\, t$
▲ (up-chirping)	$\lim_{t\to 0}\frac{\partial}{\partial t}(\ln\lvert\delta E_j\rvert) \approx \text{const.}$	$\omega_j = \omega_{j,(0)} + \Delta\omega^{+}\,\forall\, t$
▼ (down-chirping)	$\lim_{t\to 0}\frac{\partial}{\partial t}(\ln\lvert\delta E_j\rvert) \approx \text{const.}$	$\omega_j = \omega_{j,(0)} - \Delta\omega^{-}\,\forall\, t$
♦ (repeated up- and down-chirping)	$\lim_{t\to 0}\frac{\partial}{\partial t}(\ln\lvert\delta E_j\rvert) \approx \text{const.}$	$\omega_j = \omega_{j,(0)} \pm \Delta\omega^{\pm}\,\forall\, t$

Figure 5.14: **Table detailing the characterisations of the state evolution explored in Section 5.6.2.** The 6 characterisations examine the evolution of the electric field mode amplitude $|\delta E_j|$ and the mode frequency ω_j.

	Master mode	Slave mode
Training set accuracy	1.000	0.957
Test set accuracy	1.000	0.910
Validation set accuracy	0.978	0.606

was used to create a random forest classifier (RFC) which utilised fit data sets created from the manual characterisation data. The master mode training data set included 14.02% of the data (1248 simulations), while the slave mode training set included 4.14% (331 simulations). By using regions far from characterisation boundaries, the ML algorithm allows one to reliably determine the location of the characterisation boundaries without the inevitable cognitive bias from human classification. Furthermore, the relatively low ratio between fit data and full data allows one to provide confidence in the characterisation without overfitting.

As the frequency spectrogram is simply a human-readable representation of the data for characterisation, the RFC was given the original electric field data in the form of $\{\Re(\delta E_j(t)), \Im(\delta E_j(t))\}$ as training samples. Out of the fit data, 70% was used to construct the training set, and 10 estimators were used in the random forest. I obtained the following fit statistics from the RFC:

where the accuracy is the fraction of characterisations from the RFC that agree with the given data set. In simulations with 'quiescence' for a given mode, I find that the only contribution to the mode amplitude is the contribution given by the seed mode with amplitude D_j. As the seed mode does not exchange energy with the system, it should not affect the amplitude of δE. However, using a discrete time grid leads to a numerical flaw. As previously discussed in Section 5.3.4, for negligible current, one finds that an electric field is generated. In the absence of mode-mode interaction, one expects to always observe up-chirping and down-chirping for the master mode, and quiescence for the slave mode.

At low $\alpha_{\text{sl.}}$, the slave mode is initially linearly unstable ($\gamma_{\text{sl.}}(t = 0) > 0$). For low Δk simulations, the master's hole meets the slave's clump, creating a region on the PDF which is close to flat. For this phenomenon to occur, the separate H&C have to exist for long enough to meet, and therefore the modes have to be close enough in k-space - that is to say, at high Δk one simply observes uncorrelated chirping in both modes. The up-chirping branch from the master mode frequency spectrum is shown to rapidly vanish, coinciding with a vanishing branch in the slave mode frequency spectrum.

In these scenarios, one may expect to observe fixed-frequency behaviour in both of the branches; physically, this corresponds to particle transport from the clump into the hole, filling the population deficit as the structures merge in phase space. However, other H&Cs in the system saturate the frequency spectrograms and disturb the fixed-frequency regions (see Figure 5.16a).

As we increase $\alpha_{\text{sl.}}$ ($0.15 < \alpha_{\text{sl.}} < 0.2$) the slave mode is initially linearly stable ($\gamma_{\text{sl.}}(t = 0) < 0$). In these simulations, the slave mode has a possibility to become destabilised by a master hole; depending on Δk, one observes different characterisations (as shown in Figure 5.15). At sufficiently low Δk, the two modes are close enough to interact. The slave mode undergoes mode flattening, and the corresponding particle transport into the hole is small - one now observes H&C formation in the master mode at low Δk where previously one would have only observed clump formation. As one decreases in Δk, we begin to observe mode chirping in the form of down-chirping (corresponding to clumps).

For high $\alpha_{\text{sl.}}$ ($0.75 < \alpha_{\text{sl.}} < 0.9$) the slave mode tends to approximately undergo mode flattening if destabilised (see Figure 5.16b), with any frequency chirping being indistinguishable from the bandwidth in frequency of the resonance broadened mode.

As one might expect, I found that repeated chirping clusters in $(\Delta k, \alpha_m$ space for the master mode only occur when the slave mode produces clumps; in all other simulations, the master mode only exhibits one cluster of chirping events. Interestingly, we find that

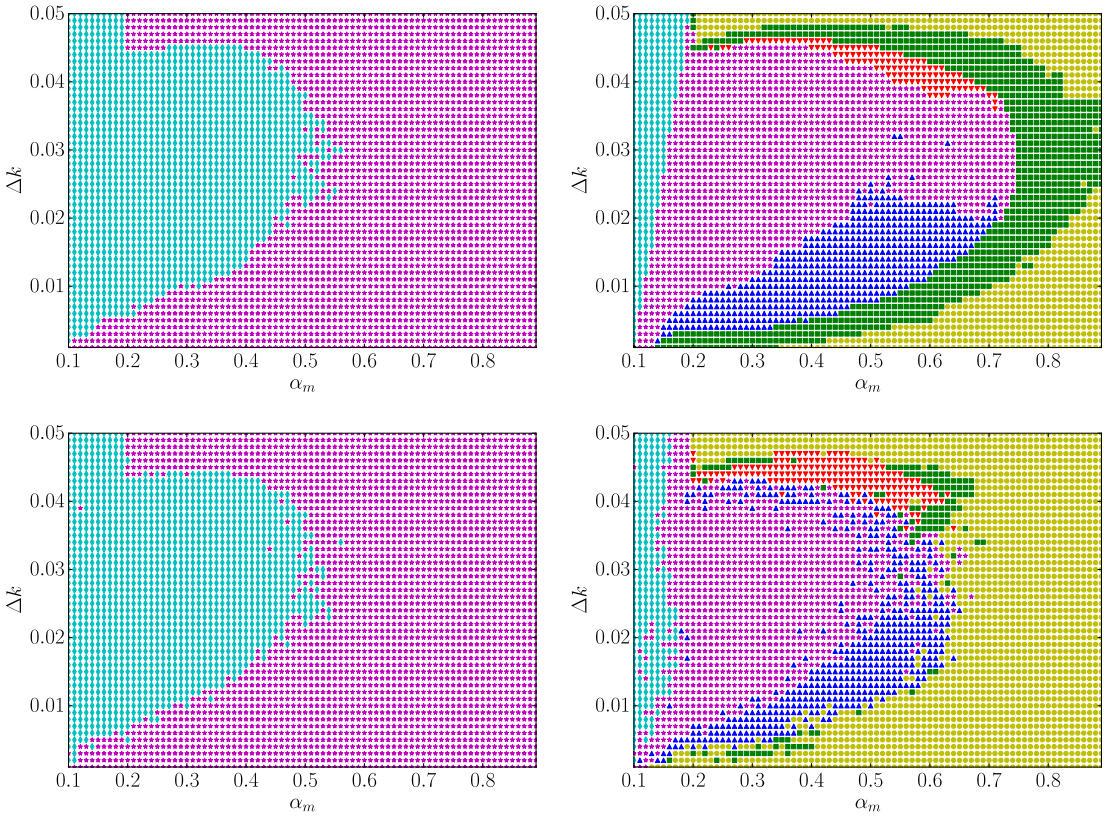
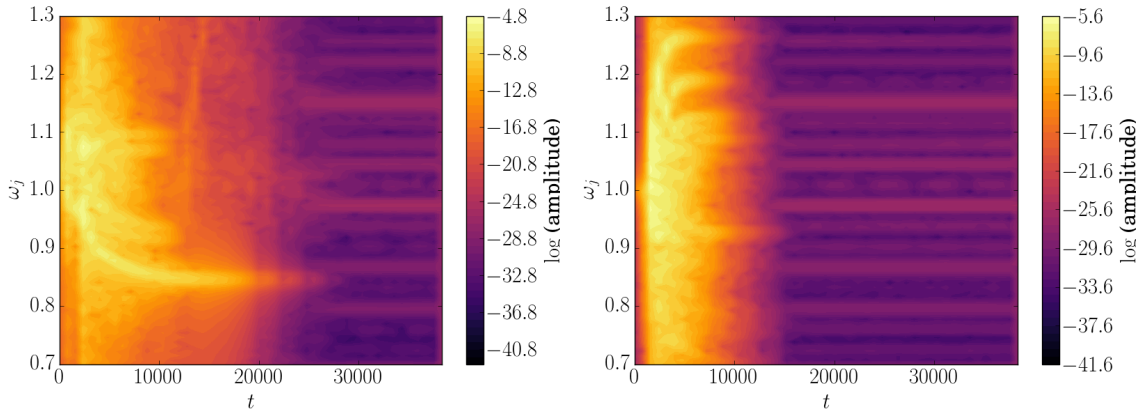


Figure 5.15: **Characterisation of the master mode (left plots) and slave mode (right plots) as function of $(\Delta k, \alpha_m)$ across the set of simulations.** The characterisation boundary (♦-★) for the master mode occurs at similar values to the union of three boundaries for the slave mode: (▼-■) ∪ (★-■) ∪ (★-▲). This phenomenon occurs because (♦) simulations require the production of slave clumps, which are missing for simulations where the outside of the boundary (▼-■) ∪ (★-■) ∪ (★-▲) for the slave mode.

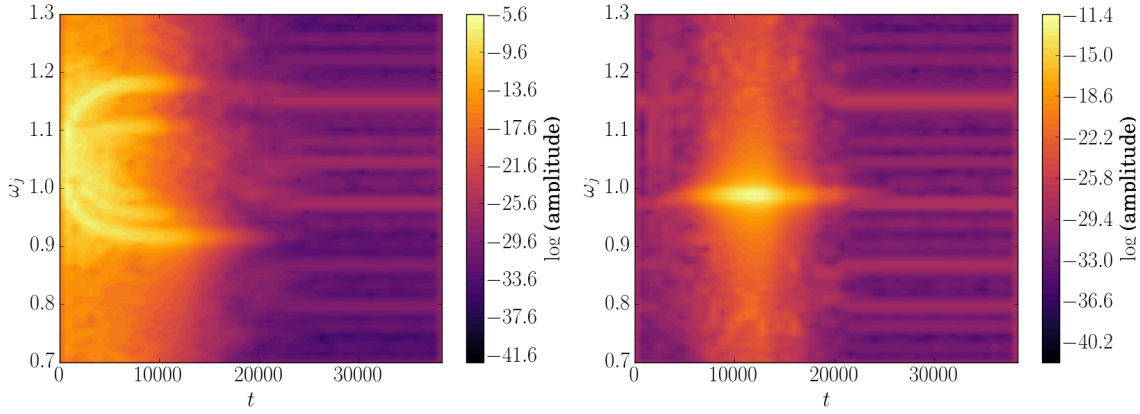
destabilised slave modes undergo preferential up-chirping or down-chirping depending on the mode spacing Δk. It is possible that due to asymmetry in $\partial^2 f/\partial v^2$ in proximity of the slave's resonance, the ansatz in Section 5.5.1 is violated (assumption of antisymmetry), leading to asymmetric chirping.

Unexpectedly, I found that highly stable slave modes (with high $\alpha_{\text{sl.}}$) are less likely to chirp when destabilised. While this might appear to be in contradiction with the inequality given by Berk *et al.*,[5] I argue that they simply give an upper bound on the equality. Including the upper bound, we find that if we consider the maximum $\alpha_{\text{sl.}}$ at which the RFC categorises the slave mode as chirping:

$$0.4\gamma_{s,L} < \alpha_{\text{sl.}} < 0.7. \tag{5.43}$$



(a) Master mode and slave mode characterised as ♦ (repeated chirping) and ★ up- and down-chirping) respectively. The slave mode is initially linearly unstable, undergoing chirping. The second chirping event in the master mode is due to destabilisation by a slave clump. The slight asymmetry is evidence of H&C annihilation; the master mode undergoes more down-chirping (fewer holes), while the slave mode undergoes more up-chirping (fewer clumps).



(b) Master mode and slave mode characterised as ★ (up- and down- chirping) and ■ (fixed-frequency) respectively. In this scenario, the slave mode is weakly destabilised by a master hole, allowing the mode to exhibit fixed frequency behaviour.

Figure 5.16: **Frequency spectrograms for the master mode, $\omega_m(t)$ (left plots) and slave mode, $\omega_{\text{sl.}}(t)$ (right plots) for $\Delta k = 0.045$.** Top plots are for $\alpha_{\text{sl.}} = 0.150$, bottom plots are for $\alpha_{\text{sl.}} = 0.750$.

5.6.3 Energy loss

Figure 5.17a allows one to easily analyse the total energy lost ΔU as a function of Δk and $\alpha_{\text{sl.}}$. The background colour shows a baseline energy loss dictated by bursting of the master mode in the absence of mode-mode interaction. Therefore, in the absence of destabilisation, there should be no extra energy lost; the increase in energy loss is a wholly nonlinear phenomenon. Added lines denote $\gamma_{\text{sl.}} = 0$ for the slave mode under certain considerations; the white line ignores destabilisation (and is therefore $\gamma_{\text{sl.}}(t = 0) = 0$),

while the cyan line allows for destabilisation by holes with a finite lifetime.

As one might expect, large energy losses occur for $\gamma_{\text{sl.}}(t = 0) > 0$ (to the left of the white dotted line). In these scenarios, both modes are unstable, losing a large amount of energy as expected from (5.16).

For $\gamma_{\text{sl.}}(t = 0) > 0$ (left of the white dotted line), two effects contribute to large energy loss. Firstly, the slave mode has $\alpha_{\text{sl.}} \gg \alpha_m$, and therefore yields a higher contribution to ΔU than the master mode. However, as the characterisation plots reveal this is not the only effect; close to the destabilisation boundary $\Delta k \sim \Delta k_{\text{est}}$, the slave mode emits a backpropagating clump that interacts with the master mode. The lifetime of the combined burst in master mode amplitude can be up to twice as long in some cases. The overall combined effect can lead to a substantially greater energy loss than is observed in the single mode case; as is shown in Figure 5.17a, the energy density loss $\Delta U/L$ can be up to two orders of magnitude greater, increasing from a value of ~ 0.02 to a value up to ~ 2.00. This result is somewhat counterintuitive, as one finds that injecting more stable modes in k-space (increasing the mode density) allows for a greater energy loss than occurs in the single mode case.

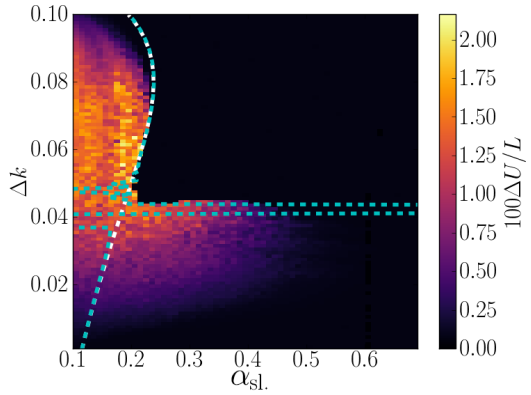
Furthermore, one can see that for some simulations with $\gamma_{\text{sl.}}(t = 0) < 0$ (right of the white dotted line in Figure 5.17a), there is an increase in $\Delta U/L$. This is due to mode-mode destabilisation; the slave mode is destabilised by the master mode, but as is shown in Figure 5.15, the master mode is destabilised in some scenarios by a clump corresponding to the slave mode. The combined effect leads to greater energy loss than one would observe in a single mode case.

5.6.4 Destabilisation boundary

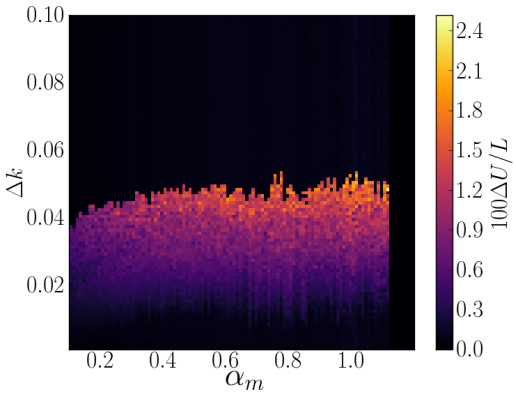
To calculate the maximum distance in k-space where mode-mode interaction still occurs, I build upon the analysis used in Section 5.6.2. First, we find $\omega_{m,(0)}$ (the initial mode frequency) exactly by using (5.8); the linear complex dispersion relation for the master mode can represented as two real integral equations with 2 unknowns: ω_m and γ_m.

Next, I use the cosine model given by (5.38), and therefore require the hole width λ. I assume that only a single master hole exists, with a mode amplitude approximately equal to the nonlinear saturation limit given by[5] $|E_m|^{(\text{NL})} \approx \alpha_m^2$.

Finally, I find $v_m(t_h)$ (the velocity of the hole after maximum chirp) using $\omega_{m,(0)}$ and (5.41). Then, I now substitute this and the hole width λ directly into (5.38), and in turn we substitute $\partial \Delta f_0/\partial t$ into (5.8) and solve directly. The resultant complex equation for



(a) Energy loss from the system via electric field damping $\alpha_m = 0.034394$ and $\alpha_{\text{sl.}} \in [0.1, 0.7]$. White dashed line denotes $\gamma_{\text{sl.}}(t = 0) = 0$ for the slave mode; for $\alpha_{\text{sl.}}$ less than these values, the mode is initially linearly stable. Cyan dashed line denotes the predicted $\gamma_{\text{sl.}}(t) = 0$ after the H&C have moved for a time t_h from the cosine model; the hole produced by the master mode k_m is able to destabilise the slave mode $k_{\text{sl.}}$.



(b) Energy loss from the system via electric field damping $\alpha_m \in [0.1, 1.2]$ and $\alpha_{\text{sl.}} = 0.3$. The threshold mode spacing at which destabilisation stops occurring (indicated by a sharp drop in ΔU with increasing Δk) initially increases weakly as a function of α_m; as we approach linear stability ($\alpha_m > 1.13$), the energy loss drops sharply again.

Figure 5.17: **Destabilisation plots from two-mode simulations in Section 5.6.3 and Section 5.6.4.** At suitably small mode spacing $\Delta k := k_m - k_{\text{sl.}}$, the slave mode $k_{\text{sl.}}$ can become destabilised.

the slave mode can represented as two real integral equations with 2 unknowns, $\omega_{\text{sl.}}$ and $k_{\text{sl.}}$. I have now fixed $\gamma_{\text{sl.}} = 0$, and use $\alpha_{\text{sl.}}$ as an input. Unfortunately, this equation is singular, owing to the pole $v = \omega_j/k_j$ on the real axis of the right hand integrand in (5.8), and so I instead seek a very small growth rate such that $\Re(p) = 0.001$. This small tolerance allows an integral solver to converge.

I therefore predict that the boundary exists for these simulations at $\Delta k_{\text{max}} \approx 0.043$ for $0.3 < \alpha_{\text{sl.}} < 0.7$, in very good agreement with the observed boundary between quiescence and mode growth for the slave mode (see Figure 5.15). The full boundary $\Delta k_{\text{max}}(\alpha_{\text{sl.}})$ as calculated here is shown in cyan in Figure 5.17a.

If one fixes α_s (mode damping for the slave mode) and instead varies α_m (mode damping for the master mode), one would expect from the form of (5.42) that Δk_{max} should be roughly constant at high α_m, and linear at low α_m. Figure 5.17b shows results from 11100 simulations with 111 values of $\alpha_m \in [0.1, 1.2]$, and 100 values of $k_{\text{sl.}} \in [0.1, 0.2]$. Indeed, I notice a similar trend; Figure 5.17b shows the energy ΔU as a function of Δk and α_m,

until the point where the master mode starts to approach linear stability ($\alpha_m > 1.13$), when the energy loss falls dramatically.

5.7 Brief conclusions

In this chapter, we investigate the role that stochasticity may play in 1D kinetic plasmas. In Section 5.3, we found that for the case of weak stochasticity, mode chirping is not wholly deterministic; the shape of the burst in mode amplitude can be determined on average analytically, but varies depending on the noise seed employed. We also found that stochasticity can suppress hole and clump formation when the distribution function is particularly noisy. Later in the chapter, I showed that holes can parametrically destabilise linearly stable modes in the system by causing strong, local deformation of the background distribution function.

Chapter 6

Identifying kinetic instabilities in tokamaks using ML

6.1 Overview

Tokamaks feature a high-dimensional parameter space (ion temperature, magnetic flux density, ion number density, etc.) with large operational domains. For certain plasma parameters, stability transition has been shown to suddenly occur over small regions of parameter space, as is observed with the L-H mode transition[76] and edge localised mode (ELM) crashes.[77] Accordingly, certain regions of parameter space have boundaries between different states of plasma stability. However, the transition that leads to abrupt large events (ALEs) is not fully understood. Previous work has shown correlation between fast ion beta and neutral beam injection (NBI) beam energy,[78] but a large area of parameter space still requires analysis.

A variety of machine learning (ML) algorithms can be used to analyse these data sets. In general, ML algorithms allow one to produce maps from the plasma parameter space to an output space which can take many different forms. The majority of work in nuclear fusion deals with two types of output spaces: characterisation spaces, and prediction spaces.

ML algorithms which map onto a set of characterisations allows one to understand the behaviour of the plasma in a more 'human' fashion. Typical scenarios involve image recognition and phase identification. Image recognition is an important part of physics, as humans commonly cast data in a visual format to enable easier interpretation, with examples including spectrogram analysis of plasma/field perturbations. Phase identification is also a crucial part of physics, where large changes in the system behaviour occur over small regions of parameter space (phase boundaries), with small changes occuring in given

regions (phases). By identifying spectrograms, one can gain further insight into the physics at play (see Section 6.4), and by mapping out different phases in parameter space, one can determine possible routes for tokamak operation that may optimise plasma performance.

ML algorithms which predict the time evolution of the plasma parameter space are also of great interest. Typical scenarios involve simulation emulation, and phase transition. Codes such as QualiKiZ[12] allow one to reproduce the results of simulations *without performing the simulations* (after training). This emulation leads to a vast increase in the computational speed of such a simulation, allowing one to reproduce the results of computationally expensive codes in a fraction of the time. ML for the purpose of phase transition has also been explored recently, primarily for disruption prediction.[13] By predicting phase transitions early, it is possible to create ML based 'smoke alarms' that may allow for real time avoidance of disruptions if considered as part of the systems control loop.

Sometimes overlap exists between some of these outcomes; generative topographic mapping (GTM) has recently been used to both map out disruptive and non-disruptive phases of the parameter space in a reduced space, and also to consequently predict the occurance of disruptions.[79]

In Section 6.2, I give a brief introduction to ML for the reader. In the remaining sections of this chapter, I detail work carried out in collaboration with Princeton Plasma Physics Laboratory (PPPL). In this work, we explored the possibility of using ML to enable rapid characterisation of frequency spectrograms, allowing us to produce correlations based on derived data sets with a high parameter space density.

6.2 Introduction to ML

For the benefit of the reader, the recent review paper by Spears *et al.*[80] provides a strong overview of accomplishments in the field, but also describes a variety of ML techniques at an overview level.

6.2.1 ML maps

Suppose we consider a map M_j from an input space X to an output space Y_j, as shown in Figure 6.1:

$$M_j : X \to Y_j. \tag{6.1}$$

This process is designed to somewhat mimic human cognitive processes. For example, a human capable of identifying Arabic numerals by eye $(0, \ldots, 9)$ is able to identify (in

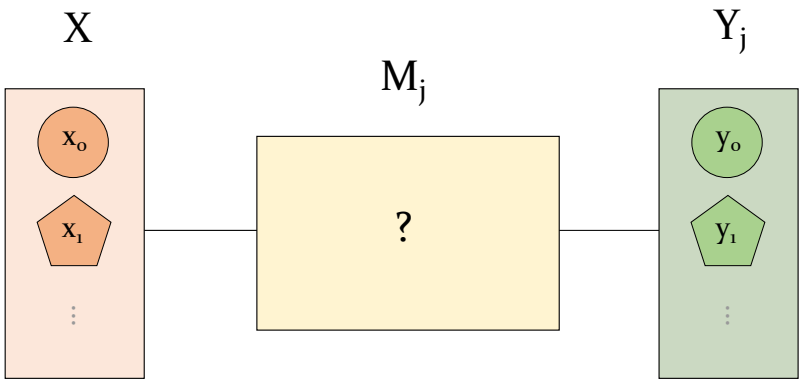


Figure 6.1: **ML maps.** An input space X is mapped onto an output space Y_j with the map $M_j \in M$, where $M = \{M_j\}$ is a set of allowed maps.

most scenarios) a number based on visual input. Interestingly, the human can never be said to be *100% accurate*. Nevertheless, most humans tend to be highly accurate at simple identification tasks such as this. The map M_j allows the human to cognitively analyse this problem, taking visual inputs and producing a classification.

One can therefore consider how to replicate this map with a ML algorithm. This can be achieved by considering the ML framework as a *set of allowed maps* ($M = \{M_j\}$), which after training takes the form of a specific map (M_T) by using some form of optimization. Just as a human may learn in a supervised or unsupervised manner, so too can a ML algorithm. This map is such that an overall 'figure-of-merit' is locally or globally maximized:

$$M_T := \text{best}(M, T), \tag{6.2}$$

where $\text{best}(M, T)$ is a function that returns the best map, and T is a set of training data (a null set for unsupervised methods). As a result, the map which is returned is a function of both the training data, and the available set of maps.

Delving further into the 'best map' function, one can represent this alternatively with respect to a loss function:

$$M_T = \arg\min_{M_j} L(M_j, T), \tag{6.3}$$

where $L(M_j, T) \in \mathbb{R}[0, \infty]$ is the loss function which is minimized for the best map, viz. this should be zero if $M_j = M_T$. One has freedom of choice for the functions $\text{best}(M, T)$ and $L(M_j, T)$; some training algorithms and optimization methods may work better than others. To conclude, one is left with four fundamental choices:

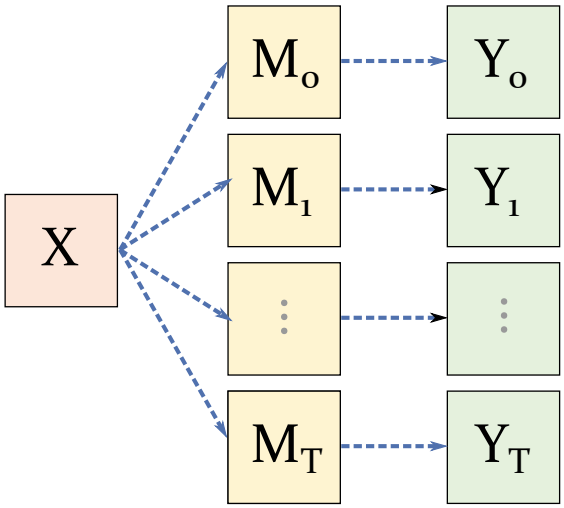


Figure 6.2: **Training.** With ML algorithms, one aims to produce a map $M_T : X \to Y_T$, where Y_T minimizes the loss function. Finding this map is called training the algorithm. By minimizing the loss function, the map M_T gives the 'closest' data set to the training data set. loss functions with multiple local minima may require global minimization schemes; for example, loss functions separable into a slowly varying part with one global minimum and a noisy/stochastic part may benefit from optimization methods such as Tabu search.[81]

- **The type of ML algorithm.** This defines the type of map that is used.
- **The optimization routine.** By choosing the loss function and the algorithm for minimizing the loss function, we may achieve differing degrees of success.
- **Hyperparameters.** These allow us to choose a specific map to use.
- **The training data.** The better quality the training data, the better the ML algorithm will perform.

6.2.2 Hyperparameters

The map M_j can alternatively be represented via a set of parameters P_j depending on the type of ML algorithm, $A = \{A_k\}$, represented as a space of generalised functions:

$$A_k : P(U_k) \to M, \tag{6.4}$$

where U_k is an element of the hyperparameter space U (i.e. number of nodes in a neural network), $P(U_k)$ is a map which produces parameter spaces determined by U_k (i.e. all possible weights in a neural network, given the number of nodes), and M_j is a map from input space X to output space Y_j. This is easiest illustrated through the use of an

example. One can examine ML algorithms which give the triangular numbers (one input x, one output y). To begin, one can consider the following simple algorithm, represented as a single summation operation:

$$A_k^{(sum)} \,:\, M = \left\{ \sum_{n=a}^{bx} n \,\middle|\, (a,b) \in P(U_k)\,;\, P(U_k) = \mathbb{R}^2 \right\}.$$

The two parameters (a, b) determine the value given. The set of available maps M is equivalently represented by the parameter space $P = \{P_j\}$. Note that M is not an element of a function space, as in general $Y_j \neq Y_i$ for some (i, j). One wishes to use a training algorithm which finds the best map M_T; this is achieved by using a training algorithm which tweaks the parameters to find the optimum set P_T. Training algorithms map from A_k onto a map M_j, as shown in Figure 6.2. To do this, they select parameters which are an element of the parameter space $P(U_k)$ which minimise the loss function. By inspection, if $a = b = 1$, then the map will correctly reproduce the triangular numbers (training set T) for $T = M_T(x)$.

However, the algorithm used to find the triangular numbers is quite specific. While this algorithm works in this example, one cannot use it for a wide variety of problems. In general, the more specific the algorithm is, the better it will perform. However on the whole, one does not fully understand how the outputs are generated for all ML problems *a priori*, and therefore it is advantageous on the whole to use more generally applicable maps.

A different algorithm which is much more general involves nesting operations of the form $x \mapsto ax + b$:

$$A_k^{(nest)} \,:\, M = \left\{ \prod_{k=1}^{N} \hat{O}_k x \,\middle|\, \hat{O}_k x = a_k x + b_k\,;\, (a_k, b_k, \cdots a_N, b_N) \in P(N)\,;\, P(N) = \mathbb{R}^{2\times N} \right\},$$

where the hyperparameter N determines how many times we nest operations. This is heuristically quite similar to an N-layer neural network with one node per layer. Again, one finds that this algorithm can produce the triangular numbers if $N = 2$, $a_1 = b_1$, $b_2 = 0$, and $a_1 b_1 = 0.5$. If the hyperparameter is too small ($N < 2$), the algorithm cannot reproduce the triangular numbers. However, if the hyperparameter is too large ($N \gg 2$), there are far too many local minima, and it is difficult to find parameters in $P(N)$ which produce the triangular numbers truly accurately.

The concept of using nested, general functions can be extended to ML algorithms on the whole. Hyperparameters are **typically selected by the trainer**, and are in general

not optimized by the routine. During training, for the given set of hyperparameters, the map is optimized.

6.2.3 Decision tree classifiers

Decision trees are perhaps the most intuitive ML algorithm, and are widely described in the literature.[82,83] Here, I will briefly cover how they work, what hyperparameters they take, and their flaws.

A decision tree takes a set of N_f inputs known as features. The features map onto a real number for output. Decision trees which perform discrete value output are known as **classifiers**, and trees that perform continous value output are known as **regressors**.

The most commonly used algorithm family for training both is the classification and regression tree (CART) algorithm. At the start of CART algorithms, the training data set exists on a root node. This data set is made up of a number of **samples**, each sample containing N_f features. At each node (starting at the root), the decision tree will examine different 'cuts' made in the input space, called **decisions**. Formally, this cut is an inequality (either $x_n < a$ or $x_n > b$). During training, the decision that is selected by CART is the one which minimizes a given loss function. After a decision has been made, the data on a node is split into two, and passed onto *leaf nodes*: data that satisfies the inequality is passed to one leaf, while data that violates the inequality is passed to another leaf. On each leaf, the process is repeated again, involving further and further splitting of the data. A connected path of leaf nodes is called a *branch*. A branch ends at a terminal node, with the termination condition dependent on a few conditions. For decision tree classifiers (DTCs), upon termination, the node is assigned a value for classification.

As an example, Figure 6.3 illustrates a possible tree that can be grown for an AND gate. The maximum depth of the tree (number of nodes in longest branch) is 2, and it is fairly trivial to obtain perfect accuracy with very little data. Typically, the termination condition for a decision tree is that the loss function reaches a given threshold, but other hyperparameters can also be employed. After training, the decision tree acts as a 'flowchart', such that a single new sample will be classified on input according to the tree.

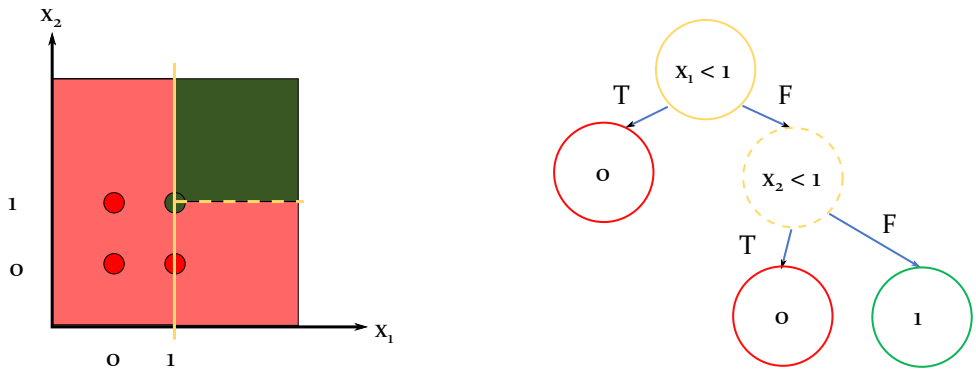


Figure 6.3: **Decision tree for an AND gate.** An AND gate can be correctly trained by using 4 data points in the training set T. A CART algorithm using Gini impurity will aim to minimize the Gini impurity in each leaf node on the tree. The first decision made will be to check whether $x_1 < 1$. If it is, then all of the training data points predict an output of 0. Therefore, the corresponding leaf node is a terminal node with a Gini impurity of 0 (all data is of the same classification, 0). If $x_1 \geq 1$, then the Gini impurity is 0.75 in the corresponding leaf node. By making a subsequent decision of $x_2 < 1$, the Gini impurity is minimized in the corresponding leaf node (all data is of the same classification, 0). All data with $x_2 \geq 1$ also minimizes the Gini impurity. The decision tree extrapolates beyond the given input space; it predicts an output of 1 for $x_1 \geq 1, x_2 \geq 1$ and an output of 0 elsewhere. As a result, generally it is best to restrain classification after training to being within the bounds of the training data. Otherwise, the accuracy of the algorithm outside of the neighbourhood of the training data cannot be adequately predicted, and may be quite low.

Gini impurity

One loss function that is very commonly used for DTCs uses the Gini impurity, I_G. This figure-of-merit is given on a single node by:

$$I_G := \sum_k \left[p_k (1 - p_k) \right], \tag{6.5}$$

such that it is the sum over all classifications of the product of the probability of a sample being of classification k, and the probability of it not being of classification k. The Gini impurity is minimized when all samples on a node are of the same classification, and varies between 0 and 1. By using CART and the Gini impurity, one can construct a loss function overall for a tree:

$$L(M_j, T) = \sum_{\text{nodes}} I_G(\{p_k\}),$$

where $p_k = p_k(M_j, T)$. The probabilities are determined by the training data T, and the decision on the node (which thereby creates the map M_j). This loss function is particularly useful, as one can minimize it by minimizing the Gini impurity in each leaf node, and then produce subsequent branches with minimized Gini impurity (allowing us to create a decision tree).

Decision contours

The decisions that the tree is trained to make leads to contours in the input space, splitting up different characterisations. As each decision is an inequality referring to a single feature, each cut made in the input space is made in Cartesian coordinates.

One can consider visually that every decision made (the parameters for the algorithm) lead to the decision contours observed. Therefore, by inverting the problem, it is trivially found that a non-Cartesian decision contour that exists in the *true map* between X and T will require many cuts in the input space. The number of cuts in the input space scales with the *tree depth* (maximum number of leaves on a branch), and therefore by increasing the branch depth, one expects higher accuracy in the decision tree algorithm.

If the tree depth is too small, not enough decisions are made, and the parameter space is poorly characterised (underfitting). However, if the tree depth is too large, too many decisions are made, and the parameter space can be poorly characterised (overfitting). This can occur if the algorithm hones in on anomalies in the input data set, or regions in the input space where characterisation is noisy/chaotic.

However, the number of cuts required to resolve a decision contour is maximized for diagonal lines and minimized for Cartesian-aligned lines. The former corresponds to highly collinear inputs, where as the later corresponds to uncorrelated inputs. One therefore expects that the accuracy of characterisation oscillates as $\sim \sin^2(2\theta)$, where θ is the angle between the line tangent to the contour, and a line aligned with a given Cartesian coordinate in the input space. This leads to an important problem with decision trees, where overfitting can occur in some regions of the input space, and underfitting in others.

Furthermore, the density of points in the training data plays a crucial role. If certain areas of the input space are uncharacterised, the predictions made by the algorithm will be poor in that section. If area of the input space are relatively densely populated in the training data, one may use a deeper tree, at detriment to the sparsely populated areas.

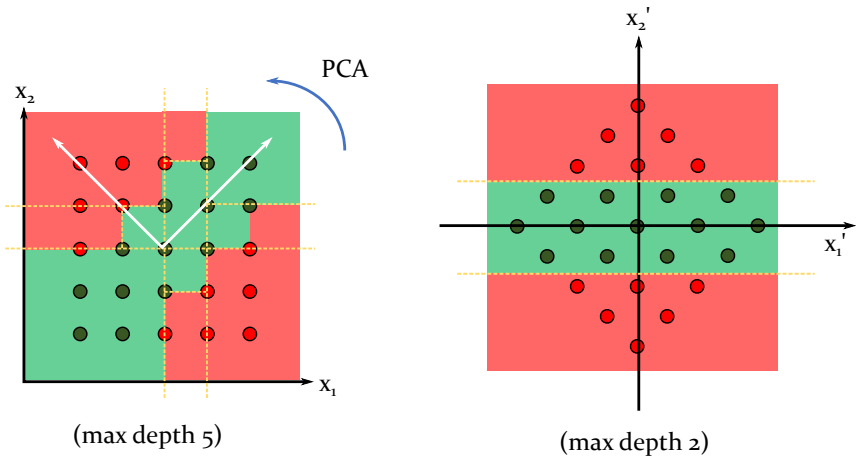


Figure 6.4: **Sketches to illustrate principal component analysis (PCA).** principal component analysis (PCA) performs a Cartesian translation and rotation of the input space onto a space where the unit vectors are given by the eigenvectors of the covariance matrix of the input data set. Mathematically, this produces a spatial representation of the data where the transformed inputs have the lowest possible amount of collinearity. An idealised case of PCA is shown in the figure, where a collinear 2D input space (nonlinearly correlated inputs (x_1, x_2)) is transformed onto a space with no collinearity (uncorrelated inputs (x_1', x_2')), with two corresponding classifications in the training data (red and green). Gold dotted lines denote the decisions made by a decision tree. For ease, the background colour immediately to the top right of a data point denotes the predicted classification. Datasets with high collinearity carry more information in their spatial representation, and therefore PCA can be thought of as a lossless form of information compression. This is advantageous for ML algorithms; the smaller the amount of collinearity, the simpler the algorithm that can be employed for the same degree of accuracy. In the figure it is shown that a decision tree with maximum depth of 5 is required to achieve 100% accuracy in the input space, and is likely to misclassify outside of the training data; the linear correlation is not preserved outside of the training data. After PCA, a decision tree would only require a maximum depth of 2 to achieve 100% accuracy in the input space, and is less likely to misclassify outside of the training data, as it would continue to exhibit the same linear correlation between x_1 and x_2 outside of the training data.

As such, decision trees perform best if:

- The density of samples in each terminal node is roughly equal across terminal nodes.
- The collinearity of the input space is small.

Methods such as PCA can improve the accuracy of trees by 'rotating' the input space

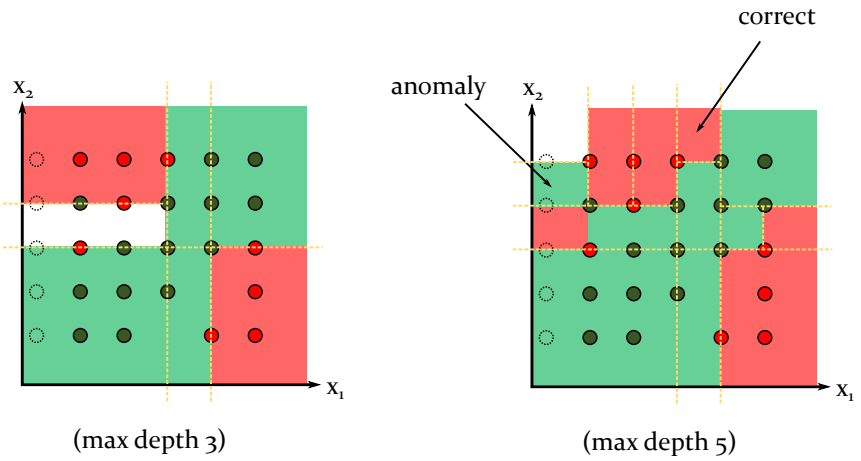


Figure 6.5: **Training data defects.** Here, I examine the same scenario as Figure 6.4, but include two types of data defects: anomalies/interstitials and vacancies. For ease, the background colour immediately to the top right of a data point denotes the predicted classification. Using PCA and a maximum tree depth of 2 would yield higher accuracy on the training data, but also high accuracy on similar data outside of the training data set. *Right plot*: Anomalies can be honed in on if the depth of a decision tree is deep enough. This means that while the data space requires a maximum tree depth of 5 for 100% accuracy, such a tree depth would cause the algorithm to erroneously learn from the defects that are present. While this would give 100% accuracy on the training data, the accuracy would only be 50% where we tested the algorithm with the vacant data points. *Left plot*: Reducing the tree depth prevents the algorithm from honing in on the defect, but yields some undefined classification (white area).

such that a set of correlated values become as close to linearly uncorrelated as possible (see Figure 6.4). The lack of correlation after PCA means that the Cartesian divisions made by decision trees are far more accurate, allowing for reduced overfitting.

6.2.4 Random forest classifiers

Random forest classifiers (RFCs) are a stochastic extension to the DTC first described by Tin Kam Ho and later developed and popularized by Breiman.[84] They allow one to smooth out inaccuracies in characterisation by adding a small amount of statistical noise, and then ensemble averaging.

Random trees work by selecting random subsets of the training data. Each random subset is used to train a tree, such that the random forest is an ensemble of decision trees.

Suppose that one uses a single decision tree and a set of training data to be characterised. If the tree characterises only 60% of the data correctly, it could be due to

imbalances in the data set, high collinearity from the input data, or a range of other reasons. By randomly selecting subsets of the data set and producing an ensemble of trees, we decrease the accuracy of each tree (the aforementioned statistical noise), however by averaging we are likely to produce a classifier of higher accuracy. Primarily, RFCs allow us to reduce noise that might occur from anomalies in the data set, or region of the input space where the characterisation is noisy/chaotic.

Due to the averaging process, RFCs give classifications as well as associated certainties for each classification; instead of a binary output (0 or 1) for a given class, a given class will have an output ranging from 0 to 1. A binary (two classification) RFC will simply pick the most likely class.

6.2.5 Confusion matrices

The confusion matrix C for a multi-class classifier is defined such that the element C_{ij} is the number of classifications made belong to class i which are actually of class j. One can construct some useful metrics from the confusion matrix:

Metrics	Definition
Number of classifications, N_C	$\sum_{ij} C_{ij}$
Accuracy	$\frac{\text{tr.}(C)}{N_C}$
Normalised skew	$\frac{1}{N_C}\sum_{ij}\text{sign}(i-j)C_{ij}$
R_K coefficient	$\frac{\sum_{klm}(C_{kk}C_{lm}-C_{kl}C_{mk})}{\sqrt{\sum_k(\sum_l C_{kl})\left(\sum_{l',k'=k} C_{k'l'}\right)}\sqrt{\sum_k(\sum_l C_{lk})\left(\sum_{l',k'=k} C_{l'k'}\right)}}$

where tr.(C) is the trace of C, and sign$(\cdot)$ is the signum function. If the classes are *ordered*, then the skew of the matrix represents the amount of bias in the system. Viz., for a system which is ordered such that class $i+1$ is higher up a heirarchy than class i, and we have enforced bias such that one is *always cautious* such that the algorithm only ever misclassifies by giving a classification that is too high in the heirarchy:

$$C_{ij} = \begin{cases} \delta_{ij} & \text{if accuracy} = 100\% \\ 0 & \text{if } i < j \end{cases},$$

such that the normalised skew of the matrix is 1, and the matrix is upper-triangular.

6.2.6 Cascading bias

Because of the aforementioned issues with density, DTCs and RFCs can perform badly if the data set is unbalanced. This can lead to a low degree of certainty from a RFC for each classification. One can correct somewhat for this by using a new hyperparameter in the form of bias. This is a design decision which was implemented in Experimental Resonant Instability Correlation Studies on NSTX (ERICSON) (see Section 6.4.2). Alternatively, one could employ PCA (as briefly shown in Figure 6.4), however this requires one to have an input space X for which any fields on X have a near constant gradient ∇X. For images, this is generally not the case, however as discussed at the end of Section 6.5.3, PCA may indeed be useful after suitable dimensionality reduction has been performed.

Additionally, as shown in Figure 6.5, errors in the output space prevent one from producing a characterisation method which is 100% accurate, and can potentially lead to overfitting. By applying bias to the certainty, it is possible for one to enable a binary RFC to perform at slightly higher accuracy, and allow for either more or less classifications to be made of a given classification. In essence, one 'relaxes' the decision contours produced by the RFC; these contours are not well defined for RFCs, but instead can be thought of as possessing some Gaussian blur, from the ensemble averaging.

It is advantageous for one to use another property of this hyperparameter in multi-class classification problems. A multi-class classifier may naïvely select the class which is most likely. Such a classifier may be extremely confident in one class which for now we call Q (ranging from 10 to 90% certainty), but not so confident in another two which one can call PD and D (ranging from 5 to 50% certainty). Then, the accuracy of the classifier is influenced by the certainty for class Q; only when the certainty for class Q drops is it possible for PD and D to be possible classifications. This classifier will perform poorly, as the accuracy of predicting class PD and D are by Bayesian statistics *also determined* by the accuracy of predictions of class Q. This is typically due to ambiguity between classifications which cannot be resolved, and the consequent overfitting/underfitting of certain areas of the input space. Conversely, an input space which involves no ambiguity in the classifications should feature a range of 0 to 100% certainty for each class.

Suppose that the resultant characterisations have a well defined hierarchy of importance, such as whether a plasma is quiescent (Q), pre-disruption (PD) or undergoing disruption (D). It is reasonable to assume that the input data for the pre-disruptive and disruptive phases may share some similarities. This can lead the aforementioned 'naïve' multi-class classifier to perform poorly. However, one desires that if there is ambiguity between two classes, the class with higher rank is used as the classification instead. It can then prove advantageous in some cases to perform separate binary classifications, and then construct a multi-class classifier with bias. By checking whether the sample is a given class or *not* a given class, one can apply bias to not only increase the accuracy, but also safely push the classifier to be 'cautious'. Each bias can be thought of as 'cascading', affecting the accuracy of not only one class, but also all other lower classes in the hierarchy. This bias can also serve to remove some of the effects that collinearity is having on multi-class classification.

It is still worth noting that using high dimensional, highly collinear data will still affect the *individual accuracy of each binary classifier*. However, this cascading bias does serve to improve the overall accuracy of multi-class classification in cases where the output space is somewhat collinear.

6.3 Fast ion instabilities

In this section, I take an aside from the mathematical discussion established in Section 6.2 to discuss fast ion instabilities in tokamaks. Later, in Section 6.4, I examine fast ion instabilities using techniques introduced in Section 6.2.

Fast ion instabilities could prove to be a serious limitation to the nominal ITER performance; these wave-particle instabilities transfer free energy between the particle distribution function (PDF) and plasma waves in the system, in certain cases leading to sudden degradation of plasma performance and energy confinement. ALEs are characterised by magnetic perturbations in the plasma undergoing very rapid frequency change ('chirping') across a broadband of frequencies, and are directly correlated with large energetic particle (EP) losses;[60,85] understanding the parametric dependencies on these losses is essential for good plasma performance. These events are sudden and highly distinguishable from the frequency behaviour at times when ALEs do not occur, and are typically observed in the upper part of the kink/tearing/fishbone (KTF) frequency band ($\sim 1\,\text{kHz}$ to $30\,\text{kHz}$) and the lower part of the toroidal Alfvén eigenmode (TAE) frequency band ($\sim 50\,\text{kHz}$ to $200\,\text{kHz}$).[52]

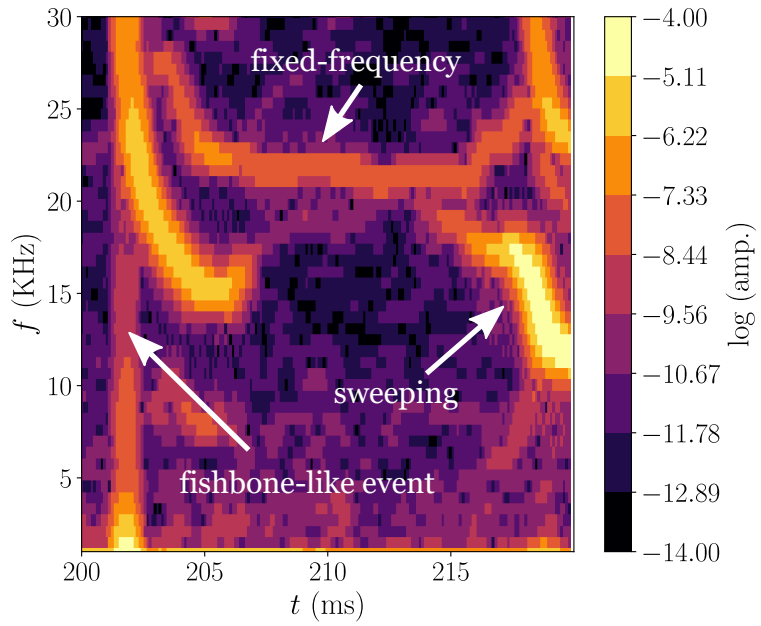


Figure 6.6: **Frequency spectrogram for magnetic fluctuations on NSTX in the** 1 kHz **to** 30 kHz **(kink/tearing/fishbone) band,** 200 ms **to** 220 ms **after the beginning of shot 139317.** 3 types of mode character are observed: fixed-frequency eigenmodes, sweeping eigenmodes, and a fishbone-like event.

Previous work has highlighted the relationship between microturbulence, stochastic effects and fast ion loss,[74] showing increased suppression of Alfvénic chirping as a function of microturbulence; theoretically microturbulence is treated as a stochastic term in the pitch angle scattering rate, while experimentally this can be heuristically inferred via the ion thermal conductivity χ_i. However, there are other correlations which may yet be undiscovered. Unfortunately, as one might expect in high-dimensional spaces, observing and predicting correlations becomes increasingly more difficult. With traditional computational analysis, there is no way to circumvent the amount of time required to perform human categorisation; a simple back-of-the-envelope calculation shows that even with an average characterisation time of 3 seconds per system state, the time taken to reach $\mathcal{O}(10^4)$ characterisations is over 8 hours. Previous work by Haskey *et al.* uses data mining techniques to to extract plasma fluctuations in time-series data, allowing one to identify different events using unsupervised classification.[86,87]

Magnetic perturbations in a very broadband range (0 to $\sim$ 500 MHz) are commonly measured on tokamaks such as NSTX by using Mirnov coils. Here, 5 characterisations of the frequency of magnetic perturbations are examined for KTF modes (of which 3 are shown

in Figure 6.6): quiescence (noise, or no frequency dependent behaviour), approximately fixed-frequency eigenmodes (herein referred to as simply fixed-frequency), sweeping eigenmodes (slow frequency variation due to time evolution of the plasma equilibrium), chirping (rapid frequency variation over a narrow frequency band due to wave-particle interaction), and fishbone-like (energetic particle mode with rapid frequency variation over a broad frequency band). For TAEs, we also examine 4 characterisations: quiescence, fixed-frequency, chirping, and ALEs (rapid frequency variation over a broad frequency band). Data from NSTX experiments in 2010 is utilised, revealing a rich set of correlations between different mode character and weighted averages of plasma parameters obtained from TRANSP.[88] Our results are in agreement with previous work by Fredrickson *et al.*[78] using human characterisation in a reduced parameter space, allowing us to heuristically confirm the validity of predictions made by ERICSON. We show new, strong correlations between moments of the spectrograms and mode character, as well as evidence of a TAE stability boundary along $v_\varphi \approx \frac{1}{4}(v_{\text{inj.}} - 3v_{\text{A}})$ where $v_{\text{inj.}} := \sqrt{2U_{\text{inj.}}/m}$ is the injection velocity, and v_{A} is the Alfvén speed.

Fast ions carry significantly more energy per mole than the thermal ions; on NSTX the NBI peak energy ($U_{\text{inj.}}$) lies at around 90 keV and the thermal peak lies at ~ 1 keV to 2 keV during typical tokamak operation.[89] If fast ions are lost from the plasma, the overall plasma performance is drastically reduced. Fast ions give a significant contribution to the plasma pressure and are essential to transfer heat to the thermal population. As described in Chapter 5, during frequency chirping events, nonlinear structures known as holes and clumps (H&Cs) can form on the ion PDF, existing respectively as a relative decrease and increase of the local PDF.

In the presence of background dissipation, kinetic instabilities that lead to H&C formation are triggered by gradients in the toroidal canonical angular momentum in a manner akin to inverse Landau damping; slowing-down distributions feature large gradients in the neighbourhood of $v_{\text{inj.}}$, and therefore wave-particle interaction is enabled for waves near $v_{\text{inj.}}$. As one attempts to increase the core temperature of the plasma, one creates sharper momentum gradients, leading to greater plasma instability. On NSTX, the NBI beam energy is super-Alfvénic ($v_{\text{inj.}} > v_{\text{A}}$), which allows for increased kinetic instability of Alfvén waves.

The formation of gap TAEs allow for relatively long lifetime waves; these waves are not dispersive in radius, and therefore allow for long range frequency chirping in the ~ 50 kHz to 200 kHz range;[40] the larger the frequency chirp, the further the momentum drift of the

resonant particles.

Furthermore, experiments and simulations in the literature[60,90–92] have shown that rapid long range frequency chirping across a wide range of frequency chirps (mode 'avalanching') is correlated with high amounts of fast ion loss. Other work has shown that mode-mode destabilisation may play a role in Alfvénic frequency chirping - activity in the KTF frequency range ($\sim 1\,\text{kHz}$ to $30\,\text{kHz}$) may instigate Alfvénic activity, and vice-versa.

Crucially, the conditions which trigger small scale frequency chirping and so-called mode avalanching are not fully understood. Intuitive knowledge related to momentum gradients provides insight, but factors such as the q-profile, energy confinement time, and the fast ion density gradient may also play a role.

6.4 ERICSON

ERICSON utilises four key parts: pre-processing NSTX data, mode characterisation, parameter space tracking from TRANSP data, and correlation studies. Here, we discuss the pre-processing, mode characterisation, and the parameter space tracking; later, we show the confusion matrices after training (see Figure 6.7), and results from some correlation studies.

Here, we describe the newly developed ML framework ERICSON, examining plasma wave frequency-chirping observed via Mirnov coils. ERICSON allows us to compare NSTX data with a wide range of parameters that would be largely unfeasible with human classification, allowing us to understand which parameters affect EP transport. After initial training of the algorithm, we find the time taken for a single process to reach $\mathcal{O}(10^4)$ characterisations is under 10 minutes, allowing for one to examine a much higher number of characterisations, and therefore a much richer set of correlations. While ML is not asymptotically convergent to perfect accuracy, it allows for broad, statistical recognition of the plasma stability boundaries that do exist. One expects any erroneous characterisations to still allow for asymptotically correct stability boundaries as one tends to an infinite number of characterisations.

6.4.1 Pre-processing (NSTX data)

The voltage drop was measured across Mirnov coils. Via the Faraday-Lenz law (1.17c):

$$V \equiv \oint_{\partial A} \mathbf{E} \cdot \mathrm{d}\ell = -\frac{\partial}{\partial t} \int_{A} \mathbf{B} \cdot \mathrm{d}\mathbf{a}, \tag{6.6}$$

where V is the voltage drop (positive by definition for a decrease in voltage), ∂A is the boundary of the area A of the coil, $\mathbf{E}$ is the electric field, $\mathbf{B}$ is the magnetic flux density, and $\hat{\mathbf{a}}$ is the unit vector normal to the surface.

We utilise the following asymmetric definitions for the discrete Fourier transform:

$$\mathcal{F}[B](\mathbf{r}, f[j]) = \sum_{l=0}^{N_t-1} B(\mathbf{r}, t[l]) \exp\left(i2\pi f[j]t[l]\right), \tag{6.7}$$

$$\mathcal{F}^{-1}[\tilde{B}](\mathbf{r}, t[j]) = \frac{1}{L_t} \sum_{l=0}^{N_t-1} \tilde{B}(\mathbf{r}, f[l]) \exp\left(-i2\pi f[l]t[j]\right), \tag{6.8}$$

where $\mathbf{r}$ is the position in real space, $f = \{j \in Z : 0 \leq j < N_t\}$ is a set of N_t frequencies, N_t is the number of time points in the dataset, and L_t is the temporal length of the dataset. Then, Fourier transforming the Faraday-Lenz law yields:

$$\tilde{V} \propto -\mathrm{i}2\pi f\tilde{B}. \tag{6.9}$$

One can immediately see that $\tilde{B}$ is singular at $f = 0$. Because discrete analysis is used, this singularity becomes broadened and can affect nearby points. For this reason, we employ a cut off frequency ($f_{\text{min.}} := 1\,\text{kHz}$) to avoid the singular value and nearby points saturating the data set. To minimise pick-up from the error-field correction switching power amplifiers (SPAs), we take the average time-domain signal from two Mirnov coils in close proximity of each other.

To enable analysis of time-dependent frequencies, we use a short-time Fourier transform (STFT) to track the frequency evolution of modes in the system. The frequency resolution Δf and maximum frequency $f_{\text{max.}}$ obtained via an STFT are given respectively by:

$$\Delta f = \frac{1}{L_t}; \qquad\qquad f_{\text{max.}} = \frac{1}{2\Delta t}, \tag{6.10}$$

where Δt is the time resolution of the data set, and the data set is now more specifically defined as the time points within the STFT window. The forward-difference STFT is calculated as follows:

$$\hat{V}(f, t) = \mathcal{F}\left[\left\{W(t')V(t') : t < t' < t + L_t\right\}\right], \tag{6.11}$$

where t is the time at which the STFT begins, $t'[j] = j\Delta t$ is a dummy time used for the transform, $W(t')$ is a window function, and $V(t)$ is now the two Mirnov coil average of the voltage drop. Many window functions are employed in the literature,[93] each producing

spurious sidebanding for a single frequency input. Here we employ a Hanning window due to its favourable decibel tapering for the erroneous signal produced, and its simple form:

$$W(t') = \sin^2\left(\frac{\pi t'}{N_t - 1}\right), \tag{6.12}$$

where $t'[j] = j\Delta t$. Using this method, one obtains the spectrogram signal:

$$\hat{B}(f,t) \propto \frac{\mathrm{i}}{2\pi f}\hat{V}(f,t). \tag{6.13}$$

Finally, one can separate the spectrogram into different subdomains of f, allowing for characterisation of the signal in different frequency bands. A human would typically examine spectrograms obtained from this pre-processing, and perform characterisation (see Figure 6.8 and Figure 6.9 for characterisations).

For ML training and analysis, 125 shots from the 2010 NSTX archives were selected due to the clear observation of KTF and TAE activity. The Mirnov coils produce a signal at sampling rate $f_V = 5\,\mathrm{MHz}$ (200 ns resolution).

The shots were split into 20 ms slices, allowing for a maximum of ~ 20 chirping events; typically chirping in these frequency bands occurs on a ~ 1 ms timescale.

The STFT time window L_t contained 2^{13} samples (1.6384 ms), allowing for a frequency resolution of $\sim 0.610\,\mathrm{kHz}$. By sliding the STFT using increments of 2^9 samples, we were able to obtain a time resolution τ on the spectrogram of 102.4 µs. It is instructive to note that by using a sliding increment smaller than the STFT window, to produce overlaid spectrograms that are plotted here (such as in Figure 6.8) one must add a 'lag' of 819.2 ns as the pixel is actually the pixel at the center of the STFT window.

The 2D data for each slice was then converted into a contiguous 1D array in memory, made by taking contiguous blocks in frequency of the original spectrogram. That is to say:

$$\begin{aligned} \mathrm{ad.}[\hat{B}(f_2,t)] >> \mathrm{sizeof}(\hat{B}) &= \mathrm{ad.}[\hat{B}(f_1, t+\tau)], \\ \forall f < f_2 : \mathrm{ad.}[\hat{B}(f,t)] >> \mathrm{sizeof}(\hat{B}) &= \mathrm{ad.}[\hat{B}(f+\Delta f, t)], \end{aligned}$$

where for a given spectrogram, f_1 is the minimum frequency, f_2 is the maximum frequency, $\mathrm{ad.}[\hat{B}(f,t)]$ is the address in memory of the spectrogram signal at (f,t), and $>>$ is the incremental bit-shift operator. A bit-shift by the datasize of the signal ($\mathrm{sizeof}(\hat{B})$) allows us to relate two neighbouring array elements.

$$C_{\mathrm{KTF}}^{(\mathrm{train})} : \begin{bmatrix} 140 & 0 & 0 & 0 & 0 \\ 1 & 54 & 0 & 0 & 0 \\ 0 & 0 & 12 & 0 & 0 \\ 0 & 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 0 & 18 \end{bmatrix} \qquad C_{\mathrm{TAE}}^{(\mathrm{train})} : \begin{bmatrix} 138 & 0 & 0 & 0 \\ 0 & 12 & 0 & 0 \\ 0 & 1 & 22 & 0 \\ 0 & 0 & 0 & 24 \end{bmatrix}$$

$$C_{\mathrm{KTF}}^{(\mathrm{test})} : \begin{bmatrix} 57 & 2 & 1 & 0 & 4 \\ 1 & 16 & 3 & 1 & 0 \\ 2 & 0 & 2 & 0 & 2 \\ 0 & 0 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 & 8 \end{bmatrix} \qquad C_{\mathrm{TAE}}^{(\mathrm{test})} : \begin{bmatrix} 59 & 0 & 0 & 1 \\ 1 & 2 & 1 & 2 \\ 0 & 2 & 7 & 1 \\ 0 & 0 & 3 & 6 \end{bmatrix}$$

Metric	Training set	Test set
Accuracy	99.2%	69.0%
R_K	0.996	0.824

Metric	Training set	Test set
Accuracy	98.9%	72.6%
R_K	0.995	0.871

Figure 6.7: **Confusion matrices and useful metrics from ERICSON trained to predict different types of classification for magnetic fluctuations on NSTX in the** 1 kHz **to** 30 kHz **(KTF) band and** 50 kHz **to** 200 kHz **(TAE) band.** From top-left to bottom-right, the leading diagonal of the confusion matrix for KTFs corresponds to correct classification for quiescence, fixed-frequency, frequency sweeping, chirping, and fishbone-like. From top-left to bottom-right, the leading diagonal of the confusion matrix for TAEs corresponds to correct classification for quiescence, fixed-frequency, chirping, and abrupt large events. The confusion matrices have high sparsity in the lower-diagonal part, owing to the preferential bias incorporated in ERICSON to increase prediction accuracy and enforce safer predictions.

In total, a database of 4773 slices was generated. It is worth noting that there is a finite probability of an exact overlap of the plasma parameters between two slices in the database, although this is quite unlikely.

6.4.2 ML algorithm and training

In this paper, we employ supervised ML to characterise the spectral data. We have used RFC, an ensemble variation of the well-established DTC. RFCs have been utilised and are well described in the literature[13,84] allowing for a simple, white-box approach of the problem, something we believe to be important for correlation studies.

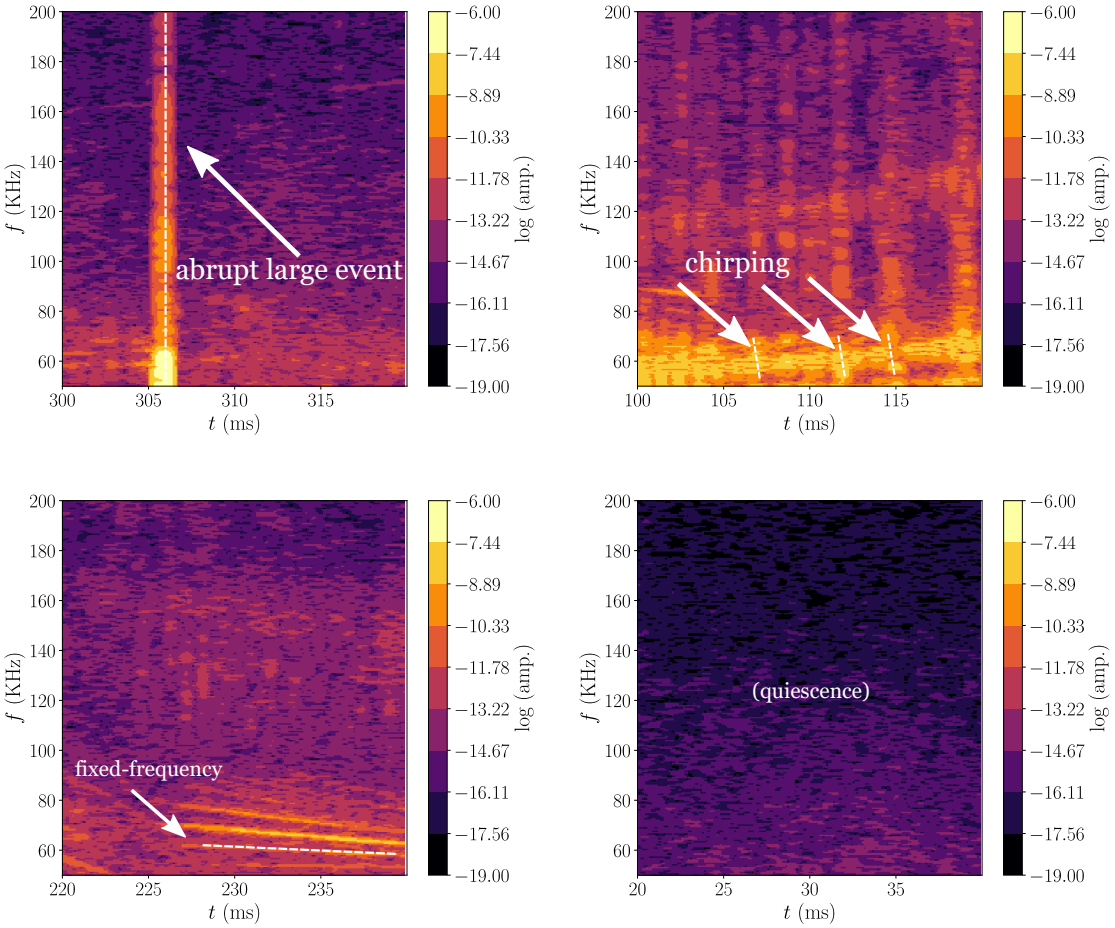


Figure 6.8: **Plots showing the 4 characterisations of TAE magnetic fluctuations obtained from shot 139317 on NSTX.** Top left: ALE, top right: chirping, bottom-left: fixed-frequency, bottom-right: quiescent.

DTCs act as flowchart-like algorithms which are optimised using a greedy algorithm. Therefore, while these classifiers are easy to operate, they sometimes only find local minima in the Gini impurity, leading to erroneous categorisation. Increasing the complexity of the tree (adding depth to the decision tree) can sometimes 'kick' the algorithm from a local minima into a global minima and improve accuracy, but too much depth leads to overfitting.

For this reason, we use RFCs which initialise an ensemble of decision trees with different random initial states. Then, the RFC returns the average probability that the data is of a given classification. This leads to a $\mathcal{O}(N_{\mathrm{DTC}})$ increase in accuracy for small N_{DTC}, where N_{DTC} is the number of trees.

Here, we examine a very high dimensional input space (each spectrogram pixel is a single dimension) with an extremely high amount of collinearity; mode character is observed by relating the spectrogram amplitude of local clusters of pixels. RFCs subdivide the

input space on a Cartesian grid, with hyperparameters determining the available number of subdivisions (tree depth) and the size of the minimum undivided volume (leaf size).

Unfortunately, highly collinear, high dimensional spaces are hard for RFCs to efficiently divide. Therefore we expect RFCs to yield a lower accuracy than other ML schemes such as convolutional neural networks (CNNs). However for correlation studies, we examine broad, collective behaviour in parameter space; as a result we do not require extremely high accuracies.

The RFCs here are implemented using the Python library `scikit-learn`;[94] it is feasible that one could perform similar analysis in `tensorflow`, or write an RFC from scratch.

a)

KTF character	Frequency traces
• (quiescent)	Noise
• (fixed-frequency)	Constant $\omega(t)$
• (sweeping)	Varying $\omega(t)$ ($>>$ 1 ms)
• (chirping)	Rapidly varying $\omega(t)$ ($\sim$ 1 ms)
• (fishbone-like)	Rapidly varying $\omega(t)$ (broadband, $\sim$ 1 ms)

b)

TAE character	Frequency traces
• (quiescent)	Noise
• (fixed-frequency)	Constant $\omega(t)$
• (chirping)	Rapidly varying $\omega(t)$ ($\sim$ 1 ms)
• (ALEs)	Rapidly varying $\omega(t)$ (broadband, $\sim$ 1 ms)

Figure 6.9: **Tables detailing the characterisations for each band.** The traces here refer to continuous streaks of slowly changing colour as one progresses in time on $\hat{B}(f,t)$ spectrograms. **a):** KTF band. **b):** TAE band.

Human classification was performed using 10 shots, producing 337 slices for the KTF band (1 kHz to 30 kHz) and 281 slices for the TAE band (50 kHz to 200 kHz). We utilised training slices which featured low ambiguity in the frequency behaviour; some of these slices included multiple characterisations during the same time slice. We desire that characterisation falls under the following hierarchy for the KTF band:

$$\text{fishbone-like} \rightarrow \text{chirping} \rightarrow \text{sweeping} \rightarrow \text{fixed-freq.} \rightarrow \text{quie.},$$

with leftmost as the most important feature. That is to say, we desire a multi-class

classifier such that any ambiguity leads to a more leftmost characterisation. For the TAE band, we desire the following hierarchy:

$$\text{ALEs} \rightarrow \text{chirping} \rightarrow \text{fixed-freq.} \rightarrow \text{quie.},$$

with leftmost as the most important feature. A set of RFCs were then trained to perform binary classification separately; each RFC yields a binary output for whether the system is exhibiting each of the types of character. We employ separate binary classification to allow us to tweak multi-class classification in a more direct fashion. The RFCs were trained using 75% of the learning set, and tested on the remaining 25%. The tree depth and number of trees were manually tweaked to optimise multi-class classification accuracy.

In Figure 6.10, we show a frequency spectrogram for magnetic fluctuations found during NSTX shot 139317, overlaid with the classification made by ERICSON. As can be seen, ERICSON categorizes fairly cautiously, and sometimes characterises chirping behaviour as ALEs. To classify, a naïve approach might simply yield the most likely classification for this; in pseudocode one can write this as:

```
for (i = 0; i < K; i ++) {
  prob[i] = RFC(i)
}
return max_index(prob)
```

where K is the number of binary classifiers used. This is typically the method employed by most multi-class classifiers. The classification accuracy from this method is typically $\min(\{p_i\})$ at best where p_i is the accuracy of the i^{th} RFC, making it a reasonable approach. However, due to the desired hierarchy, we impose a heirarchal method:

```
for (i = 0; i < K; i++) {
  if RFC(i) > RFC(j) + tol(i) for all j > i {
    return i
  }
}
```

Here, `tol(i)` denotes a tolerance factor. The tolerance factor for each characterisation is prescribed such that the confusion matrices for the training and test set meet two constraints: the confusion matrices are close to upper-triangular, and the confusion matrices are as close to diagonal as possible. The latter is the most important constraint, as a diagonal confusion matrix denotes perfect characterisation. By using the tolerance factor, we can force ERICSON to make better decisions by making it 'more cautious' (see Section 6.2.6); ERICSON would then preferentially classify TAE activity in a slice as chirping

when previously classified as quiescent, if the tolerance is set to be higher. For moderate tolerance levels, this reduces overall confusion between different characterisations, but naturally creates some off-diagonal elements in the above diagonal part of the confusion matrices. This is deemed to be an acceptable error as we have encoded a preference for caution; we would rather have ERICSON predict a more 'dangerous' behaviour if it does misclassify.

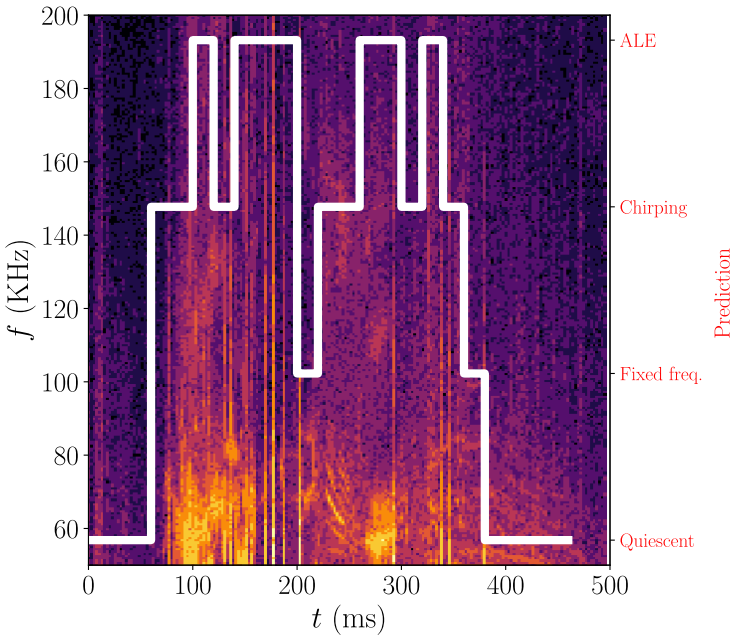


Figure 6.10: **Frequency spectrogram for magnetic fluctuations on NSTX in the** 50 kHz **to** 200 kHz **TAE band for the first 500ms of shot 139317, overlaid with the mode character classified by ERICSON.** 4 types of mode character are observed: quiescence, fixed-frequency eigenmodes, chirping modes, and ALEs

In general, many of the characterised slices will have multiple features. Unfortunately this means that while the 'most likely' method produces the best accuracy out of the multi-class algorithms we employed, it has a finite ceiling on the accuracy, owed to the fact that it will not adequately distinguish between two features which appear at the same time during a slice. While the ML multi-class classifier could be improved, we find the accuracy to still be acceptable. In kind, a human performing classification could quite easily misinterpret a slice; what may be recorded as slow frequency sweeping by one human could be recorded as fixed-frequency by another. For this reason, we include in Figure 6.7 the more conventional metric of accuracy, as well as the R_K coefficient[95] (see Section 6.2.5). This coefficient has an upper limit of 1, and a lower limit which is greater than or equal to -1 depending on

the dataset:

$$R_K = \frac{\sum_{klm}(C_{kk}C_{lm} - C_{kl}C_{mk})}{\sqrt{\sum_k(\sum_l C_{kl})\left(\sum_{l',k'=k} C_{k'l'}\right)}\sqrt{\sum_k(\sum_l C_{lk})\left(\sum_{l',k'=k} C_{l'k'}\right)}}. \tag{6.14}$$

The R_K coefficient is a generalisation of the Matthews correlation coefficient for multi-class classification. It is intrinsically related to the Pearson correlation coefficient, as it is directly expressed as the ratio of the covariance of two K-dimensional data sets corresponding to predicted classifications ($\underline{\mathbf{X}}$) and actual classifications ($\underline{\mathbf{Y}}$), and the product of the standard deviations of those two data sets:

$$R_K = \frac{\mathrm{cov}(\underline{\mathbf{X}}, \underline{\mathbf{Y}})}{\sqrt{\mathrm{cov}(\underline{\mathbf{X}}, \underline{\mathbf{X}}) \cdot \mathrm{cov}(\underline{\mathbf{Y}}, \underline{\mathbf{Y}})}}, \tag{6.15}$$

where $\mathrm{cov}(a, b)$ is the covariance of a and b. Because an unbalanced data set is used for training, we incur a reasonable amount of accuracy bias; the real data set is also unbalanced (for example, ALEs are much less common than periods of quiescence). While the R_K coefficient is not as intuitive as accuracy, this is a better measure of the quality of the ML classifier, as it weights the metric depending on this accuracy bias, and gives a metric whose value is resilient under the use of an unbalanced real data set.

6.4.3 Parameter space tracking (TRANSP data)

Each shot had a corresponding TRANSP simulation revealing the plasma parameters at each point in time. These simulations yield the classical or unperturbed predictions for the fast ion parameters. This temporal link allows for one to compare the multi-class classification given by the ML algorithm to plasma parameters.

As one performs more classification, one populates the parameter space with more points. One expects that for phase transitions occurring over small, finite regions of parameter space, ERICSON should reveal regions with largely the same classification. In Section 5.3, we showed that frequency chirps exist for stochastic lifetimes, with other work showing that stochastic transport mechanisms such as ion microturbulence can effect the likelihood of chirping and therefore the characterisations observed.[74]

For these reasons, we look at the overall behaviour of KTF and TAE magnetic activity in different regions of parameter space; one should not fully trust each individual classification, but rather the overall behaviour and the location of characterisation boundaries.

Each pixel shown in the figures in Section 6.5 is binned such that the value of the pixel is the most frequent characterisation in that bin.

6.5 Human correlation studies on output data from RFCs

6.5.1 Mode-weighted averaging

Chirping is a nonlinear phenomenon, requiring wave-particle nonlinearity. If one aims to correlate chirping with plasma parameters, some of these parameters may be spatially dependent; accordingly the mode structure plays an important role.

Here, we model the spatial distribution of physical quantities dependent on Alfvénic activity in the system via Bayesian inference, such that physical quantities are conditionally distributed *given the mode structure*. As such, by modelling the mode structure by using a normalized Gaussian (and therefore taking the prior to be Gaussian), quantities dependent on Alfvénic activity can be determined via a posterior distribution which is also Gaussian.

Accordingly, one can construct conditional expectations of plasma quantities, where integrals are weighted with a normalised Gaussian, given here by ρ:

$$\langle g \rangle \equiv \int_0^1 (g \cdot w)\, \mathrm{d}\rho, \tag{6.16}$$

where $\rho = \sqrt{\Psi/\Psi_0}$, Ψ_0 is the magnetic flux at the last closed flux surface, g is the quantity to be spatially averaged, and $\langle g \rangle$ is the mode-weighted average. Experimentally, the mode structure can be observed via reflectometry data, however this data cannot probe hollow density profiles. Unfortunately, for a sizeable number of chirping cases, beam deposition and other effects can lead to hollow density profiles, preventing inference of the mode structure via reflectometry.

In lieu of fully reliable values for every shot analysed, we make three assumptions. First, one expects Alfvén waves to have a fairly narrow mode structure, and as such we approximate the standard deviation to be ~ 0.1.

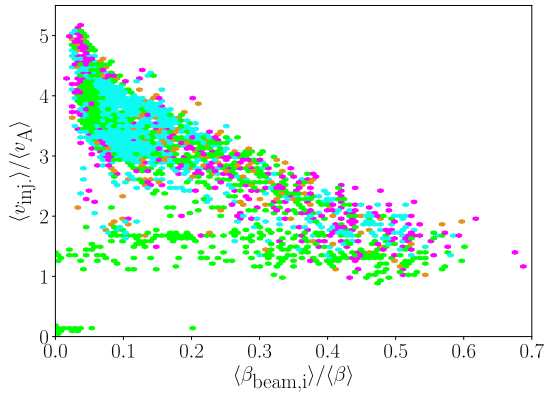
Second, we expect Alfvénic mode structure to peak at around $q_{\min}$,[40,65,96] which we approximate to be at $\rho = 0.5$ (see Levinton and Yuh[97]). Third, we extend this assumption also to the KTF modes in this analysis, however one would in reality expect the mode structure to be broader.[98] Therefore, we use the trial function for analysis:

$$w(\rho) \propto \exp\left[-\left(\frac{(\rho - \sqrt{0.5})^2}{2(0.1^2)}\right)\right], \tag{6.17}$$

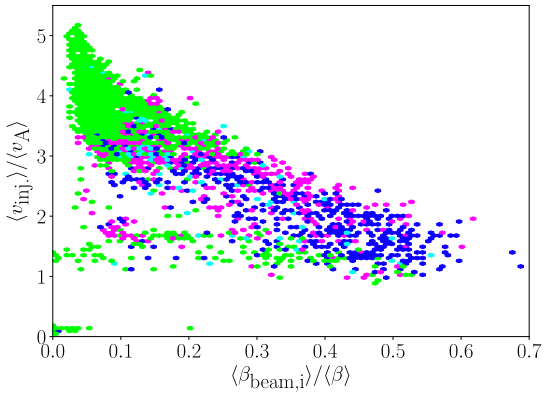
such that the mode structure is approximated by a Gaussian with peak at $\rho = 0.5$ and standard deviation 0.1.

6.5.2 Injection velocity

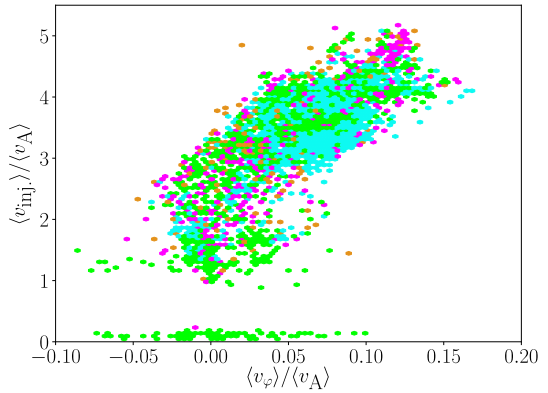
Figure 6.11 contains 4 plots showing differing mode character as a function of operational parameters at NSTX.



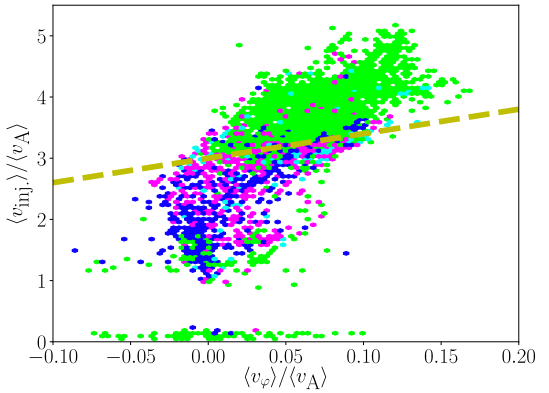
(a) Normalized injection velocity versus normalized beam ion beta for KTFs. Fixed-frequency modes largely confined to super-Alfvénic injection velocity ($v_{\text{inj.}} \gtrapprox 2v_{\text{A}}$) and low beam ion beta ($\beta_{\text{beam},i} \lessapprox 0.25\beta$).



(b) Normalized injection velocity versus normalized beam ion beta for TAEs. Quiescent behaviour largely confined to low beam ion beta ($\beta_{\text{beam},i} \lessapprox 0.2\beta$).



(c) Normalized injection velocity versus normalized plasma toroidal velocity for KTFs. Quiescent behaviour largely confined to sub-Alfvénic toroidal velocity ($v_{\varphi} \lessapprox 0.5v_{\text{A}}$). Fixed frequency behaviour dominates for $v_{\varphi} \gtrapprox 0.5v_{\text{A}}$.



(d) Normalized injection velocity versus normalized plasma toroidal velocity for TAEs. Quiescent behaviour largely confined to relatively low plasma rotation $\left(v_{\varphi} \lessapprox \frac{1}{4}(v_{\text{inj.}} - 3v_{\text{A}})\right)$.

Figure 6.11: **Plots showing differing mode character as a function of operational parameters at NSTX. Left plots:** KTF (1 kHz to 30 kHz modes); quiescent (green), fixed-frequency (cyan), sweeping (orange), chirping (blue), fishbone-like (magenta). **Right plots:** TAEs (50 kHz to 200 kHz modes); quiescent (green), fixed-frequency (cyan), chirping (blue), ALEs (magenta).

On the vertical axis for each plot is the normalized injection velocity $\langle v_{\text{inj.}}\rangle/\langle v_{\text{A}}\rangle$, where the Alfvén speed is given by:

$$\langle v_{\text{A}}\rangle(t) \approx \frac{\langle |B| \rangle}{\sqrt{2\mu_0 \langle n_e \rangle m_p}}, \tag{6.18}$$

where B is the magnetic flux density, n_e is the electron density, and m_p is the proton rest mass. This approximation assumes roughly 2 atomic mass units per electron, which is suitable for the NSTX plasma in these shots (typically featuring deuterium and carbon ions). The normalized beam ion beta is defined as $\langle \beta_{\text{beam},i}\rangle/\langle \beta \rangle$, where $\langle \beta_{\text{beam},i}\rangle$ is the beam ion beta, and $\langle \beta_{\text{beam}}\rangle$ is the total beta.

The left hand plots in Figure 6.11 show KTF mode character (1 kHz to 30 kHz modes); quiescent (green), fixed-frequency (cyan), sweeping (orange), chirping (blue), fishbone-like (magenta). The right hand plots in Figure 6.11 show TAE mode character (50 kHz to 200 kHz modes); quiescent (green), fixed-frequency (cyan), chirping (blue), ALEs (magenta).

For modes in the KTF band, fixed-frequency modes are largely confined to regions of parameter space where the injection velocity is super-Alfvénic, and the beam ion beta is relatively low. We found fixed-frequency mode behaviour for $v_{\text{inj.}} \gtrapprox 2v_{\text{A}}$, $\beta_{\text{beam},i} \lessapprox 0.25\beta$, $v_\varphi \gtrapprox 0.5v_{\text{A}}$. Quiescence is largely observed for sub-Alfvénic toroidal velocity, approximately given by $v_\varphi \lessapprox 0.5v_{\text{A}}$.

For modes in the TAE band, quiescent behaviour is largely confined to regions of parameter space where the beam ion beta is relatively low. We found quiescence for $\beta_{\text{beam},i} \lessapprox 0.2\beta$, $v_\varphi \lessapprox \frac{1}{4}(v_{\text{inj.}} - 3v_{\text{A}})$.

Plasma rotation shear can play a strong role in ideal MHD stabilisation; sub-Alfvénic plasma rotation serves to stabilise the plasma, leading to increased quiescence in the KTF frequency band. However, the TAEs are subject to kinetic instabilities for super-Alfvénic ($v_{\text{inj.}} > v_{\text{A}}$) near the Alfvén speed. Our results allow us to further posit that if the injection velocity is less than the Alfvén speed, one expects that a reversed toroidal plasma velocity leads to decreased TAE activity.

The beam ion beta plays a strong role in TAE destabilisation. Increased NBI power increases the kinetic drive for resonant modes; this leads to an increased likelihood for nonlinear wave-particle interaction and chirping.[99] As the gradient of the fast ion PDF determines the kinetic stability of nearby TAEs in momentum space, one expects theoretically that increased ion beam beta leads to increased TAE activity. The results from ERICSON show increased chirping and ALE activity at high beam ion beta in agreement

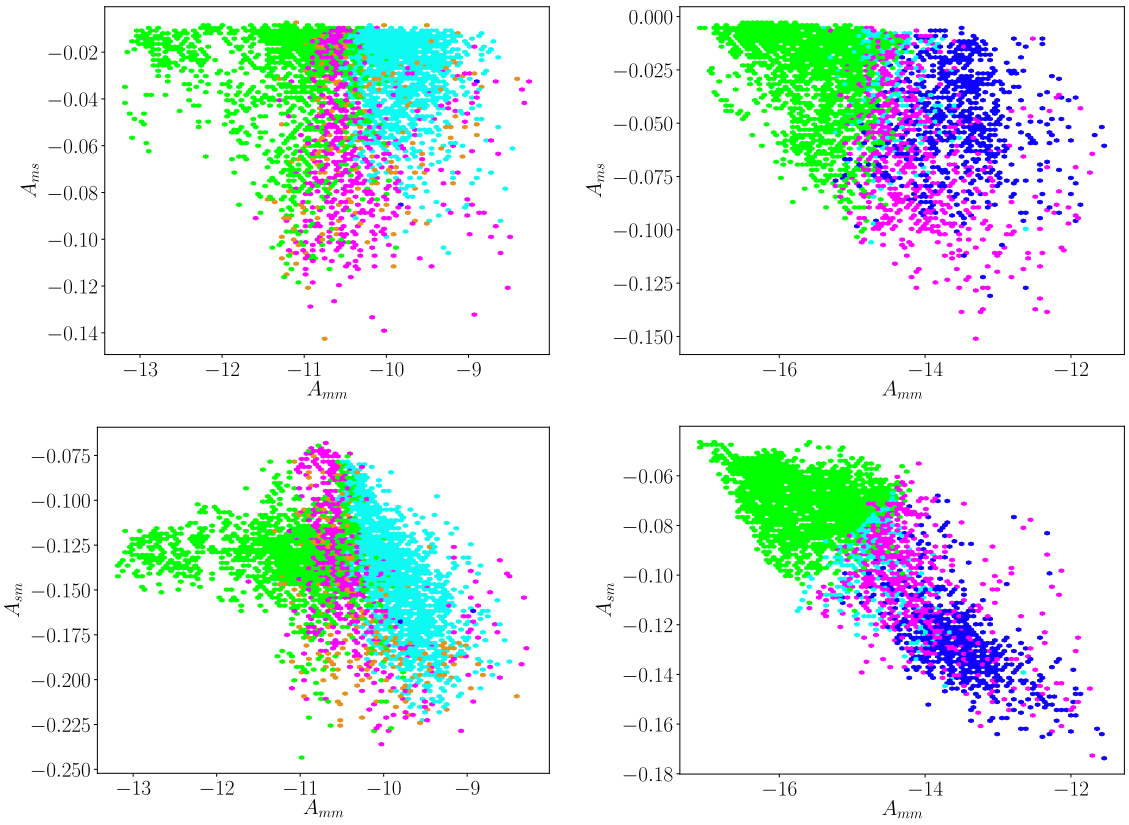


Figure 6.12: **Plots showing differing mode character as a function of spectrogram moments at NSTX. Left plots:** KTF (1 kHz to 30 kHz modes); quiescent (green), fixed-frequency (cyan), sweeping (orange), chirping (blue), fishbone-like (magenta). Spectrogram average (A_{mm}) is a good indicator of mode character; average frequency spread (A_{sm}) and temporal intermittency (A_{ms}) are poor indicators. **Right plots:** TAEs (50 kHz to 200 kHz modes); quiescent (green), fixed-frequency (cyan), chirping (blue), TAEs (magenta). Spectrogram average (A_{mm}) and average frequency spread (A_{sm}) are good indicators of mode character; temporal intermittency (A_{ms}) is a poor indicator.

with theory and observation.

6.5.3 Spectrogram moments

Figure 6.12 contains 4 plots showing differing mode character as a function of spectrogram moments at NSTX. The left hand plots show KTF mode character (1 kHz to 30 kHz modes); quiescent (green), fixed-frequency (cyan), sweeping (orange), chirping (blue), fishbone-like (magenta). The right hand plots show TAE mode character (50 kHz to 200 kHz modes); quiescent (green), fixed-frequency (cyan), chirping (blue), ALEs (magenta).

We define three simple metrics based on moments of the spectrogram signal given by

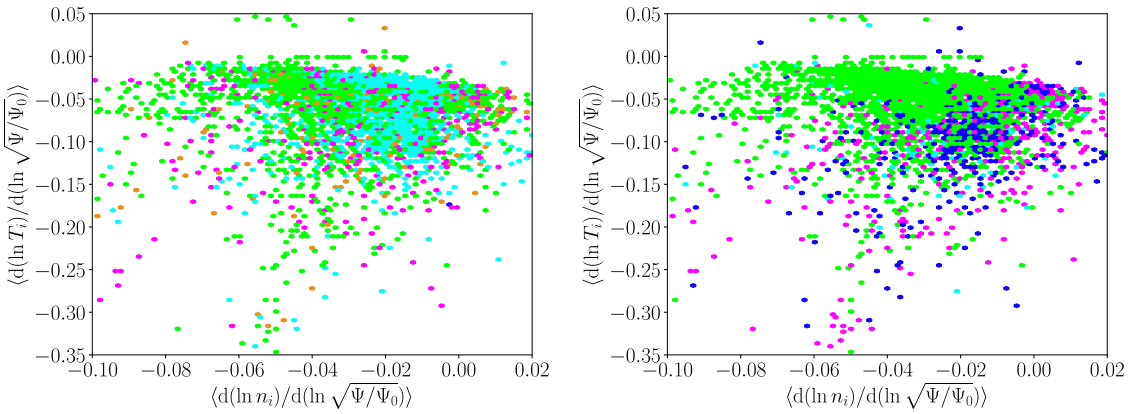


Figure 6.13: **Plots showing differing mode character as a function of normalized ion temperature gradient ($\mathbf{d}(\ln T_i)/\mathbf{d}(\ln \sqrt{\Psi/\Psi_0})$) and normalized ion density gradient ($\mathbf{d}(\ln n_i)/\mathbf{d}(\ln \sqrt{\Psi/\Psi_0})$) at NSTX. Left plot:** KTFs (1 kHz to 30 kHz modes); quiescent (green), fixed-frequency (cyan), sweeping (orange), chirping (blue), fishbone-like (magenta). Largely null result; no strong correlations observed. **Right plot:** TAEs (50 kHz to 200 kHz modes); quiescent (green), fixed-frequency (cyan), chirping (blue), ALEs (magenta). Chirping and ALEs occur at low $\eta = (\nabla T_i/\nabla n_i)$.

(6.13), taken both in frequency and time to yield a scalar. First, we define the spectrogram average A_{mm} as:

$$A_{mm} :\propto \sum_l \sum_j \hat{B}(f[j], t[l]). \tag{6.19}$$

A_{mm} is a measure of the amplitude of magnetic fluctuations in a given frequency band. The average frequency spread A_{sm} is defined as:

$$A_{sm} :\propto \sum_l \sqrt{\sum_j \left(\hat{B}(f[j], t[l]) - \mu_f(t[l]) \right)^2}, \tag{6.20}$$

where $\mu_f(t[l]) \propto \sum_j \hat{B}(f[j], t[l])$ is the frequency average of $\hat{B}$. A_{sm} is a measure of how broadband the signal is. The intermittency A_{ms} is defined as:

$$A_{ms} :\propto \sqrt{\sum_l \left(\left[\sum_j \hat{B}(f[j], t[l]) \right] - A_{mm} \right)^2}. \tag{6.21}$$

A_{ms} is a measure of how intermittent the signal is.

For modes in the KTF band, the spectrogram average heavily dominates over the other moments; the mode character is largely invariant as a function of the average frequency

spread and intermittency. For modes in the TAE band, the mode character is largely invariant as a function of the intermittency.

6.5.4 Ion η

Figure 6.13 contains 2 plots showing differing mode character as a function of ion temperature and pressure gradients at NSTX. The left hand plot shows KTF mode character (1 kHz to 30 kHz modes); quiescent (green), fixed-frequency (cyan), sweeping (orange), chirping (blue), fishbone-like (magenta). The right hand plot shows TAE mode character (50 kHz to 200 kHz modes); quiescent (green), fixed-frequency (cyan), chirping (blue), TAEs (magenta).

For modes in the KTF band, we observe a largely null result; no discernible correlation can be identified. However, for modes in the TAE band, chirping and ALEs occur at low $\eta = (\nabla T_i / \nabla n_i)$. This is consistent with observations in DIII-D,[74] reduced nonlinear kinetic simulations and nonlinear gyrokinetic simulations performed for NSTX experiments.[100] Higher η implies more drive for ion temperature gradient (ITG) modes and therefore more turbulent stochastisation of phase space. This leads to suppression of nonlinear structures, such as H&Cs, that sustain chirping. We note, however, that each turbulent mode has its own threshold in η, hence no global threshold in η can be identified as a well-defined transition for the chirping/fixed-frequency boundary.

6.6 Brief conclusions

In this chapter, I outlined how one can utilise ML algorithms (specifically DTCs and RFCs) to analyse fusion datasets for the purpose of experimentally identifying regions of parameter space with increased kinetic instability.

In collaboration with PPPL, I developed an ML framework to expedite the physics analysis of Alfvén waves at NSTX. We employed RFCs to study correlations between plasma parameters and the frequency response of Alfvén waves, which indicates the nature of fast ion losses.

We found correlations between the beam ion beta $\beta_{\text{beam},i}$ and the injection velocity $v_{\text{inj.}}$, as shown in Figure 6.11, which are in agreement with previous work by Fredrickson *et al.*,[78] forming a suitable benchmark test. Further correlations between the plasma toroidal velocity v_φ and the injection velocity $v_{\text{inj.}}$ were found, as also shown in Figure 6.11; we do not currently have physical explanation for this observation, but simply present the observed results.

Particularly interesting correlations were observed between moments of the magnetic fluctuation spectrograms, as shown in Figure 6.12. These correlations imply that it may be possible to expedite this analysis by simply using three moments of the spectrograms rather than the full spectrogram data; I will discuss this later in Chapter 7.

Chapter 7

Conclusions and summary

7.1 Conclusions

7.1.1 Conservative orbits and linear kinetic theory

Conservative orbits

In Section 2.2, I analysed so-called 'conservative orbits' where the single particle energy is conserved on a particle orbit. This preceded a discussion of Bernstein-Green-Kruskal (BGK) theory in Section 2.2.1, with some light motivation for its use in qualitatively modelling tokamak plasmas. Later in Chapter 5 and Chapter 6, I used reduced dimensionality models and BGK-like models for Alfvén waves, motivated by this discussion.

In Section 2.2.1, I qualitatively discussed the limitations of BGK theory. In later chapters, the plasmas described violated some of these limitations, and therefore the BGK theory does not adequately describe all of the dynamics. In Chapter 3, I examined models which relaxed these limitations, allowing one to better identify when BGK theory is a valid approximation; in the literature, BGK-like models are typically used with little consideration for these limitations.

In Section 2.2.2, I proved that BGK modes are dispersionless, and that it is not possible to represent a set of phase space structures via simple linear superpositions of BGK modes. Describing phase space structures as a superposition of BGK modes features widely in the literature; here, I showed that while this may be a suitable approximation, it is mathematically impossible for this to be an exact solution in 1D.

Linearized kinetic plasmas

In Section 2.3, I covered a linearised theory of kinetic plasmas. In Figure 2.4, I discussed that one can consider *evolving* shear in plasmas as nonlinear wave-wave coupling. This is important as while shear occurs in linearised systems, the degree of the shear is *constant*. Modification to the shear intrinsically requires nonlinear theory, as was later explored in Chapter 3.

Then, I derived the dispersion relation for Laplace decomposed 3D electromagnetic waves in collisional systems. Later in Section 2.3.5, I simplified the model to a 3D electrostatic waves with Krook collisions.

In Section 2.3.6, I directly solved for the perturbations to the particle distribution function (PDF). One finds that the solution has a similar functional form to the Dawson function. Interestingly, this foreshadows a key result which appears in Chapter 3: the use of a Gaussian basis decomposition in Section 3.4.3 with the same functional form of the normalising velocity in the exponential. In Figure 2.6, I further show that phase space shear occurs in this linearised model.

7.1.2 Non-conservative orbits and nonlinear kinetic theory

Non-conservative orbits

At the beginning of Section 3.2, I gave the functional form for the orbit affine parameter ϵ considered along these phase space orbits. Then, in Section 3.2.1, I examined the period of single particle orbits, and how to motivate the form of ψ from considering the work done under the Lorentz force for a single particle. Then, I analysed the periods of single particle orbits. Particles can be considered to evolve via continuum of conservative orbits if:

$$\left|\frac{\partial \ln(\delta U)}{\partial t}\right| \ll \omega_B,$$

such that the bounce frequency of particles is higher than the rate of change of energy. Accordingly, one finds that this theory is in accordance with the brief discussion in Section 2.3.7, and the adiabatic theory used in the literature.

I showed how by considering near-conservative orbits, one can define a subpopulation of particles which I refer to as quasi-passing particle. This expands discussion in the literature of passing and trapped particles;[101,102] the quasi-passing particle particles are in what is traditionally referred to as the passing region, but in fact interact strongly with the wave. As they are separated from the rest of the passing population by a singularity in the period

of orbit, they can be considered to act heuristically similar to trapped particles. Subfigure **a)** showed the existence of quasi-passing particle particles, while subfigure **b)** showed that these particles lie within an effective separatrix which is defined with respect to $q\psi$. This was illustrated further in Figure 3.1b, where the contour with $\tau \to \infty$ moved outside of the traditional separatrix for $q\psi > 0$.

Then, in Section 3.2.2, I briefly illustrated that a simple consideration of a single particle travelling in phase space along a continuum of near-conservative orbits and resonating with a wave undergoing growth ($\gamma t > 0$) leads to phase space shear. This showed *a posteriori* that the theory is in agreement with the collective behaviour expected for the shear of phase space structures on the PDFs described in Section 3.3.

In Section 3.2.3, I derived the rate of change of kinetic energy in the PDF by integrating along each particle orbit and then along all orbits. Using this method, I was able to show that perturbations to the kinetic energy in this orbit theory is determined by the perturbation of three orbit scale quantities: the period of orbits, the mean momentum of an orbit, and the shape of the orbit. This motivated later work in the chapter which treated ψ as a generating function which links wave-particle energy transfer and the formation and evolution of phase space islands.

In Section 3.2.4, I briefly discussed that $\phi + \psi$ acts as the effective electric potential for particles in this system. Indeed, as shown in Section 3.2.1, this allows for quasi-passing particle particles (due to modification of the separatrix). However, it also allows one to consider phase space structures as forming a momentum space 'sheath'. One expects that the characteristic width of this sheath in velocity is given by:

$$\lambda \sim 2\sqrt{\frac{2q|E_j|}{mk_j}}.$$

This corresponds with the width of the separatrix in BGK theory. However, *all* particles interact with the wave in the BGK theory. In contrast, for the theory established in Chapter 3, $\lim_{v\to\pm\infty} \delta f = 0$ such that particles far from the Landau resonance are 'shielded' from the resonance.

In Section 3.2.5, I showed that the PDF can be represented in a distributive form. While this was assumed *a priori* in Chapter 2, here I showed explicitly that this is possible. I therefore showed that one can construct a 'first-order' PDF, splitting the perturbation to the PDF into a linear and nonlinear functional. This generalises the technique used by Bernstein, Greene and Kruskal for BGK modes, allowing one to theoretically recover PDFs from the more exotic particle orbits explored with $\psi \neq 0$.

Nonlinear kinetics

In Section 3.3, I examined the Boltzmann-Maxwell system in a nonlinear framework.

In Section 3.3.1, I showed how the decomposition for the **E** and **B** field employed in (3.18) allows for self-consistent solutions of the Boltzmann-Maxwell system, with the time evolution of the complex frequencies ip_j determined by (3.19). In Figure 3.5, I illustrate that the complex frequencies asymptotically approach values determined by the initial conditions via $\Gamma_j^{[l]}$.

In Section 3.3.2, I briefly gave an energy conservation equation which is satisfied in the fully nonlinear consideration of the Boltzmann-Maxwell system. From this, in Section 3.3.3, I expanded the discussion in the literature of a Berk-Breizman sink of electromagnetic energy, deriving the collision operator which produces this energetic sink. This allows one to discuss the limitations of models with and without this collision operator. Models with the collision operator are full self-consistent, while models without this collision operator (as typically examined in the literature) are not self-consistent, and either involve discrepancies in the energy content of the system, or in the evolution of the PDF.

In Section 3.3.4, I derive the linear complex dispersion relation for electrostatic waves in models which use the Berk-Breizman sink explicitly via a collision operator (the self-consistent, fully correct model), and models which instead use an equation that resembles a modified version of the Maxwell-Ampère law. The models are heuristically similar except for the fact that the true model should instead produce two branches, with the $k_j = 0$ branch being damped by the Berk-Breizman collision operator, and the $k_j \neq 0$ being driven by the Berk-Breizman collision operator. Accordingly, for a single resonant mode in the system, the models are almost the same in the limit of $\gamma_D \to 0$. This motivates the use of the model used in Chapter 5.

Basis decomposition formulation

In Section 3.4, I examined basis decompositions of the orbit affine parameter ϵ, and the resultant modes that are allowed to exist in the system.

In Section 3.4.1, I examine a Dirac delta decomposition which recovers van Kampen modes. This decomposition allows for particles with velocity exactly equal to the phase velocity of a wave in the system to resonate, and permits now other wave-particle interaction. This is in stark contrast to the BGK theory which allows all particles to resonate with the wave; here, only an infinitesimal 'sheath' surrounds the resonance in the phase space.

In Section 3.4.2, I decomposed the orbit affine parameter ϵ using a novel nonlinear basis decomposition, where the decomposition is different at every point in $\mathbf{x}$ and t. By allowing each $(\mathbf{x}, t)$ to permit a different basis decomposition, I allowed for nonlinear coupling which is typically not allowed using a variables separable method of the form $f = XT(\mathbf{x}, t)V(\mathbf{v})$. I introduced a Galilean transformation matrix χ which contains co-moving coordinates in each possible wave frame. Remarkably, one finds that under the approximation that waves are slowly chirping or that the particle flux perpendicular to the wave is small, the evolution of the zeroth velocity moment of ϵ has a non-holonomic constraint when using this coordinate transformation, implying that it is not possible to derive the evolution of this form of solution via Lagrangian theory or related formalisms such as action-angle perturbation theory.

Generating function formulation

In Section 3.5, I analyse ϵ using an alternative method, treating ψ as a generating function which generates collisions, frequency sweeping, drive, and particle gyration.

In Section 3.5.1, I introduced a velocity Galilean transformation matrix ν, and show that by using χ and ν, one can significantly decompose the differential equation which governs the evolution of ψ; under this decomposition, one can represent ψ via a sum of separate contributions.

In Section 3.5.2, I derive the contribution that generates the frequency sweep, and in Section 3.5.3, I derive the contribution outside the separatrix that generates mode growth. I utilise both of these in Section 3.5.4, where I approximate the form of ψ in a collisionless system with no wave-wave coupling. From this form, I illustrate the approximate form of phase space contours under this formulation, and note that appearance of quasi-trapped particles near the separatrix, in agreement with earlier theory in Section 3.2.1.

7.1.3 Stochastic and deterministic modelling of holes and clumps

From Section 5.3, we conclude that the lifetime of a hole and clump (H&C) is stochastic, and that the decay rate of the mode is also stochastic. We hope that the accuracy of analysis for t_p and t_d improves with further work; in reality, the plateau region has a slight negative slope. It is our belief that an upgraded model with a negative gradient for t_p would yield greater accuracy on the nonlinear dependence of t_p on t_{NL}.

In Section 5.4, we found that increasing the stochasticity in the system is initially equivalent to increasing the effective diffusive collisionality. This is as one might expect

from theory; stochasticity in the turbulent PDF or electric field affects the damping term αE, which is analogous to the energetic particle drive $\mathrm{d}f/\mathrm{d}p_\varphi$. Accordingly, simulations with increased stochasticity produce similar overall results to theory based on stochastic perturbations to momentum scattering via microturbulence-induced radial diffusion.[103] However, an important nuance appears when considering repeated bursting; low amplitude repeated bursts are saturated by the electric field produced by f_{tur}, leading to an asymptotic behaviour characteristic of a decrease in the effective collisionality.

As a result, we posit that in a given regime, an increase in micro-turbulence leads to an anomalous decrease in the effective collisionality; it is implied that in this regime, microturbulence reduces the ability for the PDF to reconstitute via pitch angle scattering.

In Section 5.5, I found that a marginally unstable mode can serve to destabilise a marginally stable mode, serving to sharply increase energetic loss, leading to losses in some cases that are up to two orders of magnitude greater. The mode-mode interaction can be explained using semi-analytic theory detailing H&C destabilisation of the marginally stable (slave) mode; time-dependent modification of the PDF leads to a modified linear growth rate for the slave mode. I demonstrated that a simple cosine model for the hole adequately allows for a prediction of the maximum mode spacing at which destabilisation occurs. However, one still requires a single simulation to observe the lifetime of the H&C.

I have shown that marginally stable modes can lose much more energy than marginally unstable modes. The consequences are initially counterintuitive: by implication, an increased ratio of marginally stable to marginally unstable modes in a tokamak plasma can serve to increase overall energy loss rather than decrease it, provided that the mode spacing is sufficiently small. This means that in large tokamaks, increasing mode stability of gap toroidal Alfvén eigenmodes (TAEs) might actually serve to increase overall energy loss, as opposed to the expected reduction that would be observed in small tokamaks where destabilisation may be less prevalent.

However, at very small mode separation, the energy loss for two modes rapidly decreases. This is likely to be due to both increased mode overlap (slave mode flattening, no behaviour) and preferential up-chirping. This is a particularly interesting result, as it implies that tightly packed modes close to the quasilinear limit have a preferred direction of chirping. I posit that this may explain preferential down-chirping as observed in mode avalanching in tokamaks: a down-chirping mode close to the poloidal axis $\Psi = 0$ could destabilise a mode which preferentially down-chirps when destabilised, triggering other down-chirps in a cascading fashion, leading to large fast ion losses.

7.1.4 Identifying kinetic instabilities in tokamaks using machine learning (ML)

In conclusion, in collaboration with PPPL I have developed an ML framework to expedite the physics analysis of Alfvén waves at NSTX. We employed random forest classifiers (RFCs) to study correlations between plasma parameters and the frequency response of Alfvén waves, which indicates the nature of fast ion losses. Our results are remarkably good, considering the high dimensionality of the input space. Furthermore, for both the TAE and KTF bands, very strong correlations are found between mode character and moments of the spectrograms of magnetic fluctuations in the plasma found in Section 6.5.3. Accordingly, one could use these moments as features instead; one can expect that reduction to this three dimensional space yields higher accuracy at quicker computational speeds.

7.2 Summary and future work

In summary, this thesis has covered a range of topics, covering analytical theory, computational physics, and machine learning driven data analysis. I have presented a breadth of work, but each of the three 'parts' to this thesis have blended together smoothly to make for some interesting research.

7.2.1 Analytical work

The work carried out in Chapter 3 could easily be developed upon. To start, a logical next step would be to reconstruct first-order distribution functions from values of ϵ, rather than to simply examine particle orbits. Much more work is yet to be done regarding solving the equations in Section 3.5: aiming to recover the asymmetric terms in $\delta\psi$ to enable symmetry in ψ would be very interesting, as would be solutions that allow one to peer into the separatrix. The solutions found for ψ in the generating function sadly did not allow one to look inside the separatrix. It would be interesting to know whether this has physical consequence (it is not immediately clear whether the lack of closed orbits inside the separatrix are problematic for the theory or not).

7.2.2 Computational work

Expanding the theory of burst stochasticisation in Chapter 5 to include the time between bursts could lead to predictions of the burst frequency between Alfvén mode chirping events in tokamaks, which would allow for a greater understanding of fast ion loss. Additionally,

further study building on previous work may allow for empirical values that allow for prediction of the lifetimes of holes and clumps without simulations. This work enabled myself to examine hole and clump destabilisation later in the chapter (Section 5.5), and further work building on this could lead to a richer understand how to create and destroy holes and clumps in an easier fashion.

As was shown in Section 5.5, hole-and-clump destabilisation preferentially creates holes and clumps depending on how far apart the modes are, and depending on how fast the master mode chirps. It is my opinion that this could almost definitely be utilised to create (at least in theory or in simulations at first) a shock front of holes or clumps. It is again of my opinion that this mechanism may be a contender for the origin of abrupt large events and mode avalanching in tokamaks. From simulation carried out outside of this thesis but in immediate extension appeared to show that holes pass almost unperturbed through holes, and clumps pass almost unperturbed through clumps, similar to solitons. However, further work could be carried out in DARK to investigate whether hole and clump annihilation can be easily enabled in simulations by choosing the correct parameters. This could then open up the avenue for potential work into 'anti-avalanches' where many holes are launched towards an avalanche of clumps, to attempt to mitigate fast ion loss.

The language that is starting to be employed here (and has been employed in discussions around this work) seems to push towards treating holes and clumps as macroparticles. As such, another avenue of further work would involve creating macro-theories of holes and clumps. Simulations based on simulating a small collection of holes and clumps may offer the benefit of simpler simulations while retaining much of the key physics. However, more research into hole and clump interaction would need to be carried out first.

7.2.3 Machine learning work

The work in this chapter could be expanded to examine automatic dataset development (such as archival searches for shots with given mode behaviour), real-time feedback control, and real-time modelling. Recent work using ML has investigated real-time capable modelling of NSTX-U using neural networks;[14] by also employing classification data, it may be possible to further enhance the predictive capabilities of such a framework.

Further work building on this analysis could aim to examine correlations in the full parameter space. Transitions such as the L-H mode transition have been partially explained in reduced parameter spaces (such as the 2D parameter space with magnetic shear $\hat{s}$ and normalised pressure gradient α).[104] This transition is shown by projecting the full

parameter space onto a plane, similar to the plots we show; in reality, this transition occurs over more exotic surfaces in parameter space which are topologically challenging to analyse. Furthermore, it is entirely feasible that a coordinate transform in the parameter space would change the topology of stability boundaries - one cannot predict *a priori* what is the most sensible representation of the parameter space for a given stability condition.

One particularly interesting avenue for future work involves investigating the strong correlations between spectrogram moments observed in Figure 6.12. As discussed in principal component analysis (PCA) would be highly advantages for this derived dataset: a workflow involving reclassification based on the spectrogram moments instead is currently underway along with some small projects I hope to carry out in my spare time using ERICSON and DARK.

Appendices

Appendix A

Conservative orbits and linear kinetic theory

A.1 Electrostatic plasma with Krook-type collisions

If there is no initial electric or magnetic field, equation (2.23) simplifies to give for an electrostatic plasma with Krook collisions:

$$p_j - \sum_{\text{species}} \frac{q^2}{m\epsilon_0} \int_{\mathbb{R}^3} \frac{v_{(p)}}{p_j + \mathrm{i}\mathbf{k}_j \cdot \mathbf{v} - \nu_{\text{Kr.}}} \frac{\partial F}{\partial v_{(p)}} \, \mathrm{d}^3 v = 0,$$

where $P(v) = \nu_{\text{Kr.}}$, and I have stated that $v_{(p)}$ is the component of the velocity parallel to the **E**-field:

$$(\mathbf{e} \cdot \mathbf{v}) =: v_{(p)}.$$

To avoid confusion with the parallel and perpendicular directions considered in tokamaks, I choose not to use the parallel symbol $\|$. If one examines the integrand, the volume element is given by:

$$\mathrm{d}^3 v = |\mathcal{J}| \, \mathrm{d}v_{(p)} \, \mathrm{d}A,$$

where $|\mathcal{J}|$ is the determinant of the Jacobian[43] of the mapping from Cartesian coordinates to the **E**-field aligned coordinates. $v_{(p)}$ is the component of $\mathbf{v}$ *normal* to the surface area differential $\mathrm{d}A$. To generalise the integral for the complex plane, one must define a complex velocity $v + \mathrm{i}w$, and extend F for complex arguments.

Expressing the contour integral explicitly in terms of its pole with respect to the $v_{(p)}$ coordinate:

$$\int\limits_{\mathbb{R}^3} \frac{v_{(p)}}{p_j + \mathrm{i}\mathbf{k}_j \cdot \mathbf{v} - \nu_{\mathrm{Kr.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}^3 v = -\frac{\mathrm{i}}{k_{j,(p)}} \int\limits_{\mathbb{R}^2} \int\limits_{\Omega} \frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\mathrm{Lan.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} \,\mathrm{d}A,$$

where Ω is an undefined contour, and the complex velocity $v_{\mathrm{Lan.}}$ is given by:

$$v_{\mathrm{Lan.}} = \frac{\mathrm{i}(p_j - \nu_{\mathrm{Kr.}}) + k_{j,(p)} v_{(p)} - \mathbf{k} \cdot \mathbf{v}}{k_{j,(p)}}.$$

For $\Im(v_{\mathrm{Lan.}}) > 0$, the forward Laplace transform used to produce $\tilde{f}_j$ and associated quantities is not expected to be singular. Therefore, the integration contour is simply given by integration along the $v_{(p)}$ axis. However, as $\Im(v_{\mathrm{Lan.}}) \to 0$, one must keep the pole above the contour integrated along. If $\Im(v_{\mathrm{Lan.}}) < 0$, one must encircle the pole, and therefore via Cauchy's theorem[43] one must consider the residue of the pole. If $\Im(v_{\mathrm{Lan.}}) = 0$, one must add a semicircular part to the contour; this yields a value equal to half that obtained for $\Im(v_{\mathrm{Lan.}}) < 0$.

Therefore, in conclusion:

$$\int\limits_{\Omega} \frac{v_{(p)}|\mathcal{J}|}{p_j + \mathrm{i}\mathbf{k}_j \cdot \mathbf{v} - \nu_{\mathrm{Kr.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} = \begin{cases} \displaystyle\int\limits_{\mathbb{R}} \frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\mathrm{Lan.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} & \text{for } \Im(v_{\mathrm{Lan.}}) > 0, \\ \mathrm{i}\pi\mathcal{R} + \mathcal{P} \displaystyle\int\limits_{\mathbb{R}} \frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\mathrm{Lan.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} & \text{for } \Im(v_{\mathrm{Lan.}}) = 0, \\ 2\mathrm{i}\pi\mathcal{R} + \displaystyle\int\limits_{\mathbb{R}} \frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\mathrm{Lan.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} & \text{for } \Im(v_{\mathrm{Lan.}}) < 0, \end{cases}$$

where $\mathcal{R}$ is the residue in $v_{(p)}$ of the integral at the simple pole $v_{(p)} - v_{\mathrm{Lan.}} = 0$, and where the Cauchy principal value of an integral which is singular at $Q(v = v_0)$ is given by (note the difference between this and Lebesgue integration):[43]

$$\mathcal{P} \int\limits_a^b Q(v) \,\mathrm{d}v := \lim_{\epsilon \to 0^+} \left\{ \int\limits_a^{v_0 - \epsilon} Q(v) \,\mathrm{d}v + \int\limits_{v_0 + \epsilon}^b Q(v) \,\mathrm{d}v \right\}.$$

It is possible to estimate the linear stability boundary (where $\Im(v_{\mathrm{Lan.}}) = 0$) by using $\lim_{\Im(v_{\mathrm{Lan.}}) = 0^-}$; this is an adequate quick way to estimate the value by setting $\Im(p_j) \approx \nu_{\mathrm{Kr.}}$. However, one should note that there is a finite discontinuity as $\Im(p_j) \to \nu_{\mathrm{Kr.}}$, with one only travelling π radians around the singularity for the $\Im(p_j) = \nu_{\mathrm{Kr.}}$ solution. In short, iterating for smaller $(\Im(p_j) - \nu_{\mathrm{Kr.}})$ *does not produce a convergent answer.*

By using integration by parts, one finds:

$$\int\limits_{\mathbb{R}} \frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\mathrm{Lan.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} = -\int\limits_{\mathbb{R}} F \frac{\partial}{\partial v_{(p)}} \left[\frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\mathrm{Lan.}}} \right] \mathrm{d}v_{(p)},$$

with use of the product rule yielding:

$$\int\limits_{\mathbb{R}} \frac{v_{(p)}|\mathcal{J}|}{v_{(p)} - v_{\text{Lan.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}v_{(p)} = -\left\{ \int\limits_{\mathbb{R}} \frac{F(W + v_{(p)}\partial_{v_{(p)}}|\mathcal{J}|)}{v_{(p)} - v_{\text{Lan.}}} \,\mathrm{d}v_{(p)} - \int\limits_{\mathbb{R}} \frac{F v_{(p)}|\mathcal{J}|}{(v_{(p)} - v_{\text{Lan.}})^2} \,\mathrm{d}v_{(p)} \right\}.$$

Now, I shall expand the reciprocals in the integrands on the right hand side; ignoring high order terms assumes a fundamentally 'cold plasma', where F is dominated by values at low $|\mathbf{v}|$. Each term in the expansion is proportional to increasing moments of the particle distribution function (PDF). Under a full expansion, if $|\mathcal{J}|$ is equal to unity:

$$\begin{aligned}
&\sum_{\text{species}} \frac{q^2}{m\epsilon_0} \int\limits_{\mathbb{R}^3} \frac{v_{(p)}}{p_j + \mathrm{i}\mathbf{k}_j \cdot \mathbf{v} - \nu_{\text{Kr.}}} \frac{\partial F}{\partial v_{(p)}} \,\mathrm{d}^3 v \\
&\quad = \frac{\mathrm{i}}{k_{j,(p)}} \sum_{\text{species}} \left\{ \frac{\bar{n}|_{t=0} q^2}{m\epsilon_0 v_{\text{Lan.}}} + \frac{q(\mathbf{e}\cdot\mathbf{k}_j)\mathbf{e}\cdot(\bar{\mathbf{J}}|_{t=0} - \bar{\mathbf{J}}|_{t=0})}{m\epsilon_0 v_{\text{Lan.}}^2} + \cdots \right\} \\
&\quad = \frac{\mathrm{i}}{k_{j,(p)} v_{\text{Lan.}}} \sum_{\text{species}} (\omega_{\text{pl.}}^2) + \ldots
\end{aligned}$$

The residue $\mathcal{R}$ is given by:

$$\mathcal{R} = \int\limits_{\mathbb{R}^2} v_{\text{Lan.}} \left. \frac{\partial F}{\partial v_{(p)}} \right|_{v_{(p)} = v_{\text{Lan.}}} |\mathcal{J}| \,\mathrm{d}A.$$

One should note that the residue can be safely Taylor expanded around $\Im(v_{\text{Lan.}}) = 0$ to use the solely real PDF:

$$\mathcal{R} = \int\limits_{\mathbb{R}^2} \left\{ v_{\text{Lan.}} \left. \frac{\partial F}{\partial v_{(p)}} \right|_{v_{(p)} = \Re(v_{\text{Lan.}})} + \mathcal{O}[\Im(v_{\text{Lan.}})]^2 \right\} |\mathcal{J}| \,\mathrm{d}A.$$

Therefore, to first order in $\Im(v_{\text{Lan.}})$ (such that $|\gamma_j - \nu_{\text{Kr.}}|$ is small) and for a suitably cold plasma (such that higher order moments of the PDF have a negligible contribution), the linear complex dispersion relation is given by:

$$p_j + \frac{\mathrm{i}}{k_{j,(p)} v_{\text{Lan.}}} \sum_{\text{species}} (\omega_{\text{pl.}}^2) \approx \frac{\pi}{k_{j,(p)}} \sum_{\text{species}} \left\{ \frac{\omega_{\text{pl.}}^2}{\bar{n}|_{t=0}} \int\limits_{\mathbb{R}^2} v_{\text{Lan.}} \left. \frac{\partial F}{\partial v_{(p)}} \right|_{v_{(p)} = \Re(v_{\text{Lan.}})} |\mathcal{J}| \,\mathrm{d}A \right\},$$

where I have used $\pi\mathrm{i}$ times the residue, as the residue was expanded in the vicinity of $\Im(\gamma_j - \nu_{\text{Kr.}}) \sim 0$. In rectilinear coordinates, we find that after multiplying through by $(p_j - \nu_{\text{Kr.}})$:

$$p_j(p_j - \nu_{\text{Kr.}}) + \sum_{\text{species}} (\omega_{\text{pl.}}^2) \approx \mathrm{i}(p_j - \nu_{\text{Kr.}})^2 \rho_j,$$

where I have defined a part ρ_j that represents the contribution from wave-particle resonance:

$$\rho_j := \frac{\pi}{k_{j,(p)}^2} \sum_{\text{species}} \left\{ \frac{\omega_{\text{pl.}}^2}{\bar{n}|_{t=0}} \int_{\mathbb{R}^2} \frac{\partial F}{\partial v_{(p)}}\bigg|_{v_{(p)}=\omega_j/k_{j,(p)}} \mathrm{d}A \right\}.$$

For $\gamma_j, \nu_{\text{Kr.}} << \omega_j$, one finds that:

$$\begin{aligned} p_j(p_j - \nu_{\text{Kr.}}) &\approx -\omega_j^2 - 2\omega_j(\gamma_j - \nu_{\text{Kr.}}/2)\mathrm{i}, \\ (p_j - \nu_{\text{Kr.}})^2 &\approx -\omega_j^2 - 2\omega_j(\gamma_j - \nu_{\text{Kr.}})\mathrm{i}. \end{aligned}$$

By examining the real part of the former equation, one finds:

$$\omega_j^2 + 2(\gamma_j - \nu_{\text{Kr.}})\rho_j\omega_j - \sum_{\text{species}} (\omega_{\text{pl.}}^2) \approx 0.$$

This yields the following dispersion:

$$\omega_j \approx -(\gamma_j - \nu_{\text{Kr.}})\rho_j \pm \sqrt{(\gamma_j - \nu_{\text{Kr.}})\rho_j + \frac{\sum_{\text{species}} (\omega_{\text{pl.}}^2)}{2(\gamma_j - \nu_{\text{Kr.}})}}.$$

One can see that in the limit of $(\gamma_j - \nu_{\text{Kr.}})\rho_j \to 0$, the frequency of the wave is equal to the plasma frequency, as one might expect. However, finite $(\gamma_j - \nu_{\text{Kr.}})\rho_j$ enables a band gap between the upper and lower branches.

By examining the imaginary part of the latter equation, one finds:

$$\gamma_j \approx \frac{1}{2}(\omega_j\rho_j + \nu_{\text{Kr.}}).$$

For the case of a single species plasma, and with $\nu_{\text{Kr.}} = 0$, one obtains an expression very similar to collisionless Landau damping:

$$\gamma_j \approx \frac{\nu_{\text{Kr.}}}{2} + \frac{\pi\omega_j^3}{2k_{j,(p)}^2\bar{n}|_{t=0}} \int_{\mathbb{R}^2} \frac{\partial F}{\partial v_{(p)}}\bigg|_{v_{(p)}=\omega_j/k_{j,(p)}} \mathrm{d}A.$$

Appendix B

Non-conservative orbits and nonlinear kinetic theory

B.1 Growth rate generating function

B.1.1 Infinite sum solution in 1D

Here I shall solve the following equation given in Section 3.5.3:

$$\left[v\frac{\partial}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial}{\partial v}\right]\psi = -\gamma\phi.$$

Lemma B.1. *The following equation:*

$$\left[v\frac{\partial}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial}{\partial v}\right]\psi = -\gamma\phi,$$

permits a solution:

$$\psi = \psi'(x,v) - \frac{\gamma}{v}\sum_{n=0}^{\infty}\left(\frac{q}{\frac{1}{2}mv^2}\right)^n \frac{(-1)^n(2n!)}{2^{2n}n!}C_n,$$

where $\psi'(x,v)$ is a generating function, C_n are integral polynomials:

$$C_n := \sum_{l=0}^{n}\frac{(-1)^l}{l(n-l)!}\phi^{n-l}I_{l+1},$$

and $I_n(x)$ is defined as:

$$I_n(x) := \int_{x_{(0)}}^{x}\phi^n(x')\,\mathrm{d}x'. \tag{B.1}$$

Proof. Seek a term which when operated on with $v\partial/\partial x$ returns the right hand side, $-\gamma\phi$. Then, operate on this with the left hand side operator and see what extra term is generated. By using $\gamma I_1/v$, one finds:

$$\left[v\frac{\partial}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial}{\partial v}\right]\left(\psi + \gamma\frac{I_1}{v}\right) = \gamma\frac{q}{m}\frac{\partial\phi}{\partial x}\frac{I_1}{v^2},$$

where the term on the right hand side is generated by the Lorentz force term (v-derivative). Iteratively, one can find the series by repeating this technique:

$$\left[v\frac{\partial}{\partial x} - \frac{q}{m}\frac{\partial\phi}{\partial x}\frac{\partial}{\partial v}\right]\left(\psi + \gamma\left[\frac{I_1}{v} - \frac{q}{m}\frac{\phi I_1 - I_2}{v^3}\right]\right) = \gamma\frac{q^2}{m^2}\frac{\partial\phi}{\partial x}\frac{3(\phi I_1 - I_2)}{v^4}.$$

At this point the following identity becomes useful:

$$(a+1)\phi^a I_b\frac{\partial\phi}{\partial x} = \frac{\partial}{\partial x}[\phi^{a+1}I_b - I_{a+b+1}].$$

One finds that on each iteration one produces a term which is q/mv^2 times an integral polynomial. Therefore, after an infinite number of iterations, one finds:

$$\left[v\partial_x - \frac{q}{m}\partial_x\phi\partial_v\right]\psi' \equiv \lim_{n\to\infty}\gamma(2n-1)\left(\frac{q}{\frac{1}{2}mv^2}\right)^n\frac{(-1)^n(2n!)}{2^{2n}n!}C_n, \tag{B.2}$$

where ψ' is given by:

$$\psi' = \psi + \frac{\gamma}{v}\sum_{n=0}^{\infty}\left(\frac{q}{\frac{1}{2}mv^2}\right)^n\frac{(-1)^n(2n!)}{2^{2n}n!}C_n.$$

One finds that via l'Hopital's rule, the right hand side of (B.2) vanishes everywhere except at $v = 0$. Therefore:

$$\psi = \psi' - \frac{\gamma}{v}\sum_{n=0}^{\infty}\left(\frac{q}{\frac{1}{2}mv^2}\right)^n\frac{(-1)^n(2n!)}{2^{2n}n!}C_n.$$

□

Next, I shall show that this has a closed form solution as a single integral function:

Lemma B.2. *The following function:*

$$\psi = \psi' - \frac{\gamma}{v}\sum_{n=0}^{\infty}\left(\frac{q}{\frac{1}{2}mv^2}\right)^n\frac{(-1)^n(2n!)}{2^{2n}n!}C_n,$$

is representable in the form:

$$\psi = \psi' - \operatorname{sign}(v)\gamma\sqrt{\frac{m}{2}}\int_{x_{(0)}}^{x}\frac{q\phi(x')}{\sqrt{\epsilon_{(0)}(x,v) - q\phi(x')}}\,\mathrm{d}x',$$

where $\epsilon_{(0)}$ is defined as:

$$\epsilon_{(0)}(x,v) = q\phi(x) + \frac{1}{2}mv^2.$$

Proof. First, to switch the order of the summation:

$$\begin{array}{ccccccc} n=0 & ; & l\in[0,0] & \rightarrow & l=0 & : & n\in[0,\infty] \\ n=1 & ; & l\in[0,1] & \rightarrow & l=1 & : & n\in[1,\infty] \\ n=2 & ; & l\in[0,2] & \rightarrow & l=2 & : & n\in[2,\infty] \\ \vdots & & \vdots & & \vdots & & \vdots \end{array}$$

Therefore one finds that switching the order of summation requires:

$$\sum_{n=0}^{\infty}\sum_{l=0}^{n} \rightarrow \sum_{l=0}^{\infty}\sum_{n=l}^{\infty}.$$

With this information, one can evaluate the n sum first:

$$\psi = \psi' - \frac{\gamma q}{v}\sum_{l=0}^{\infty}\sum_{n=l}^{\infty}\left(\frac{q}{\frac{1}{2}mv^2}\right)^n \frac{(-1)^n(2n!)}{2^{2n}n!}\frac{(-1)^l}{l(n-l)!}\phi^{n-l}I_{l+1}.$$

This allows one to find a function of ϕ which is *independent* of integration. Noting that:

$$\begin{aligned} \frac{\phi^{n-l}}{(n-l)!} &= \frac{(n-l+1)(n-l+2)\dots(n)}{n!}\phi^{n-l} \\ &= \frac{\phi^{n-l}}{n!}\prod_{r=1}^{l}(n-l+r) \\ &= \frac{1}{n!}\left(\frac{\mathrm{d}}{\mathrm{d}\phi}\right)^l \phi^n, \end{aligned}$$

it is possible to show that:

$$\begin{aligned} \psi &= \psi' - \frac{\gamma q}{v}\sum_{l=0}^{\infty}\frac{(-1)^l}{l!}I_{l+1}\left(\frac{\mathrm{d}}{\mathrm{d}\phi}\right)^l\sum_{n=l}^{\infty}R^n\frac{(-1)^n(2n!)}{2^{2n}(n!)^2} \\ &= -\frac{\gamma q}{v}\sum_{l=0}^{\infty}\frac{(-1)^l}{l!}I_{l+1}\left(\frac{q}{\frac{1}{2}mv^2}\frac{\mathrm{d}}{\mathrm{d}R}\right)^l\sum_{n=0}^{\infty}R^n\frac{(-1)^n(2n!)}{2^{2n}(n!)^2}, \end{aligned}$$

where I have used the shorthand:

$$R := \frac{q\phi}{\frac{1}{2}mv^2}.$$

This quantity has physical significance; it is the ratio of the potential energy and kinetic energy of the particle. The sum over n is nothing more than a fractional binomial expansion:

$$\sum_{n=0}^{\infty} R^n \frac{(-1)^n (2n!)}{2^{2n}(n!)^2} = (1+R)^{-0.5},$$

and therefore one now has:

$$\psi = \psi' - \frac{\gamma q}{v} \sum_{l=0}^{\infty} \frac{(-1)^l}{l!} I_{l+1} \left(\frac{q}{\frac{1}{2}mv^2} \frac{\mathrm{d}}{\mathrm{d}R} \right)^l (1+R)^{0.5}.$$

By induction, one finds:

$$\begin{aligned} \left(\frac{\mathrm{d}}{\mathrm{d}R} \right)^l (1+R)^{0.5} &= \left(\frac{\mathrm{d}}{\mathrm{d}R} \right)^{l-1} (1+R)^{-1.5} \left(\frac{-1}{2} \right) \\ &= \left(\frac{\mathrm{d}}{\mathrm{d}R} \right)^{l-2} (1+R)^{-2.5} \left(\frac{-1}{2} \right) \left(\frac{-3}{2} \right) \\ &= (1+R)^{-l-0.5} \frac{(-1)^l (2l)!}{2^{2l} l!}, \end{aligned}$$

and therefore:

$$\psi = \psi' - \frac{\gamma q}{v} \sum_{l=0}^{\infty} I_{l+1} \left(\frac{q}{\frac{1}{2}mv^2} \right)^l (1+R)^{-l-0.5} \frac{(2l)!}{2^{2l}(l!)^2}.$$

It is possible to combine terms by noting the following:

$$\frac{(1+R)^{-l-0.5}}{(\frac{1}{2}mv^2)^l} = \sqrt{\frac{mv^2}{2\epsilon_{(0)}}} (\epsilon_{(0)})^{-l}.$$

Therefore, ψ now takes the form:

$$\begin{aligned} \psi &= \psi' - \mathrm{sign}(v)\gamma q \sqrt{\frac{m}{2\epsilon_{(0)}}} \sum_{l=0}^{\infty} \left(\frac{q}{\epsilon_{(0)}} \right)^l I_{l+1} \frac{(2l)!}{2^{2l} l!} \\ &= \psi' - \mathrm{sign}(v)\gamma q \sqrt{\frac{m}{2\epsilon_{(0)}}} \int_{x_{(0)}}^{x} \left\{ \phi(x') \sum_{l=0}^{\infty} \left(\frac{q\phi(x')}{\epsilon_{(0)}} \right)^l \frac{(2l)!}{2^{2l} l!} \right\} \mathrm{d}x', \end{aligned}$$

where I note that $v = \mathrm{sign}(v)\sqrt{v^2}$, with $\mathrm{sign}(v)$ as the signum function. Finally one finds that:

$$\sum_{l=0}^{\infty} \left(\frac{q\phi(x')}{\epsilon_{(0)}} \right) \frac{(2l!)}{2^{2l}(l!)^2} = \left(1 - \frac{q\phi(x')}{\epsilon_{(0)}} \right)^{-0.5}.$$

And therefore one arrives at the closed form solution:

$$\psi = \psi' - \mathrm{sign}(v)\gamma \sqrt{\frac{m}{2}} \int_{x_{(0)}}^{x} \frac{q\phi(x')}{\sqrt{\epsilon_{(0)}(x,v) - q\phi(x')}} \mathrm{d}x'$$

□

By substitution, one therefore finds the following condition:

$$\psi' = \text{const.}$$

Via gauge freedom one can choose ψ' to be zero. Therefore:

$$\psi = -\text{sign}(v)\gamma\sqrt{\frac{m}{2}}\int_{x_{(0)}}^{x}\frac{q\phi(x')}{\sqrt{\epsilon_{(0)}(x,v)-q\phi(x')}}\,\mathrm{d}x'.$$

B.1.2 Infinite sum solution in 3D

Here I shall solve the following equation given in Section 3.5.3:

$$\left[\mathbf{v}\cdot\nabla-\frac{q}{m}(\nabla\phi)\cdot\nabla_{\mathbf{v}}\right]\psi=-\gamma\phi.$$

Lemma B.3. *The following equation:*

$$\left[\mathbf{v}\cdot\nabla-\frac{q}{m}\nabla\phi)\cdot\nabla_{\mathbf{v}}\right]\psi=-\gamma\phi,$$

permits a solution:

$$\psi=-\frac{\gamma}{|\mathbf{v}|}\sqrt{\frac{m}{2}}\int_{\mathbf{x}_{(0)}}^{\mathbf{x}}\frac{q\phi(\mathbf{x}')}{\sqrt{\epsilon_{(0)}(\mathbf{x},\mathbf{v})-q\phi(\mathbf{x}')}}\mathbf{v}\cdot\mathrm{d}\mathbf{x}',$$

where $\epsilon_{(0)}$ is defined as:

$$\epsilon_{(0)}(\mathbf{x},\mathbf{v})=q\phi(\mathbf{x})+\frac{1}{2}m|\mathbf{v}|^2,$$

and where C_n are integral polynomials:

$$C_n\equiv\sum_{l=0}^{n}\frac{(-1)^l}{l(n-l)!}\phi^{n-l}I_{l+1},$$

with $I_n(x)$ is defined as:

$$I_n(x):=\frac{1}{v}\int_{\mathbf{x}_{(0)}}^{\mathbf{x}}\phi^n(x')\mathbf{v}\cdot\mathrm{d}\mathbf{x}'. \tag{B.3}$$

Proof. Seek a term which when operated on with $\mathbf{v}\cdot\nabla$ returns the right hand side, $-\gamma\phi$, in a similar manner to Appendix B.1.1. Then, operate on this with the left hand side operator and see what extra term is generated. By using $\gamma I_1/|\mathbf{v}|$, one finds:

$$\left[\mathbf{v}\cdot\nabla-\frac{q}{m}(\nabla\phi)\cdot\nabla_{\mathbf{v}}\right]\left(\psi+\gamma\frac{I_1}{|\mathbf{v}|}\right)=\gamma\frac{q}{m}\frac{\partial\phi}{\partial x}\frac{I_1}{|\mathbf{v}|^2},$$

where the term on the right hand side is generated by the Lorentz force term ($\nabla_{\mathbf{v}}$). Iteratively, one can find the series by repeating this technique:

$$\left[\mathbf{v}\cdot\nabla-\frac{q}{m}(\nabla\phi)\cdot\nabla_{\mathbf{v}}\right]\left(\psi+\gamma\left[\frac{I_1}{|\mathbf{v}|}-\frac{q}{m}\frac{\phi I_1-I_2}{|\mathbf{v}|^3}\right]\right)=\gamma\frac{q^2}{m^2}\frac{\partial\phi}{\partial x}\frac{3(\phi I_1-I_2)}{|\mathbf{v}|^4}.$$

At this point the following identity becomes useful:

$$(a+1)\phi^a I_b\nabla\phi=\nabla[\phi^{a+1}I_b-I_{a+b+1}].$$

One therefore ends up with the same closed form solution as in Appendix B.1.1, but with different integral functions $\{I_n\}$, and with $|\mathbf{v}|$ used instead:

$$\psi=\psi'-\frac{\gamma}{|\mathbf{v}|}\sum_{n=0}^{\infty}\left(\frac{q}{\frac{1}{2}m|\mathbf{v}|^2}\right)^n\frac{(-1)^n(2n!)}{2^{2n}n!}C_n.$$

The use of sign(v) to fulfill symmetry requirements has now been incorporated in higher dimensionality by using a more complete form of I_n. Therefore, by virtue of Lemma B.2, one finds:

$$\psi=\psi'-\frac{\gamma}{|\mathbf{v}|}\sqrt{\frac{m}{2}}\int_{\mathbf{x}_{(0)}}^{\mathbf{x}}\frac{q\phi(\mathbf{x}')}{\sqrt{\epsilon_{(0)}(\mathbf{x},\mathbf{v})-q\phi(\mathbf{x}')}}\mathbf{v}\cdot\mathrm{d}\mathbf{x}'.$$

Again similarly to Appendix B.1.1, ψ' can be shown to be a constant, and therefore:

$$\psi=-\frac{\gamma}{|\mathbf{v}|}\sqrt{\frac{m}{2}}\int_{\mathbf{x}_{(0)}}^{\mathbf{x}}\frac{q\phi(\mathbf{x}')}{\sqrt{\epsilon_{(0)}(\mathbf{x},\mathbf{v})-q\phi(\mathbf{x}')}}\mathbf{v}\cdot\mathrm{d}\mathbf{x}'.$$

□

B.2 Leibniz path integral rule

One can extend the Leibniz integral rule to path integrals. Suppose that one starts with the following integral:

$$I:=\int_{\mathbf{a}(\mathbf{r})}^{\mathbf{b}(\mathbf{r})}\mathbf{f}(\mathbf{r},\mathbf{t})\cdot\mathrm{d}\mathbf{t}\equiv\sum_j\int_{a_j(\mathbf{r})}^{b_j(\mathbf{r})}f_j(\mathbf{r},\mathbf{t})\,\mathrm{d}t_j.$$

This integral is a path integral between two points on the D-dimensional manifold $\mathbf{R}^D$ where the position on the manifold is given by $\mathbf{r} = r_i \mathbf{e}^i$. Suppose that one acts on I with a vector differential operator $\hat{\mathbf{O}}$:

$$\hat{\mathbf{O}} I = \sum_j \left[\hat{\mathbf{O}}(b_j(\mathbf{r})) f_j(\mathbf{r}, \mathbf{b}) - \hat{\mathbf{O}}(a_j(\mathbf{r})) f_j(\mathbf{r}, \mathbf{a}) + \int_{a_j(\mathbf{r})}^{b_j(\mathbf{r})} \hat{\mathbf{O}} f_j(\mathbf{r}, \mathbf{t}) \, \mathrm{d}t_j \right],$$

which can be alternatively represented by using dyads:

$$\hat{\mathbf{O}} I = \hat{\mathbf{O}}\mathbf{b}(\mathbf{r}) \cdot \mathbf{f}(\mathbf{r}, \mathbf{b}) - \hat{\mathbf{O}}\mathbf{a}(\mathbf{r}) \cdot \mathbf{f}(\mathbf{r}, \mathbf{a}) + \int_{\mathbf{a}(\mathbf{r})}^{\mathbf{b}(\mathbf{r})} \hat{\mathbf{O}}\mathbf{f}(\mathbf{r}, \mathbf{t}) \cdot \mathrm{d}\mathbf{t}.$$

Appendix C

Stochastic and deterministic modelling of holes and clumps

C.1 Resonant damping

The classical Lagrangian and Hamiltonian densities of the electromagnetic field are given by [19]:

$$\mathcal{L}_{(0)} = -\frac{F^{\alpha\beta}F_{\alpha\beta}}{4\mu_0} - A_\alpha J^\alpha \quad ; \quad \mathcal{H}_{(0)} = \Pi^{\beta\alpha}\partial_\beta A_\alpha - \mathcal{L}_{(0)}, \tag{C.1}$$

where $F^{\alpha\beta}$ is the electromagnetic force tensor, μ_0 is the permeability of free space, J^α is the 4-current. A_α and $\Pi^{\beta\alpha}$ are the 4-potential and conjugate Π-tensor:

$$A_\alpha = (\phi/c, \mathbf{A}) \quad ; \quad \Pi^{\beta\alpha} = \frac{\partial\mathcal{L}}{\partial(\partial_\beta A_\alpha)}, \tag{C.2}$$

where ϕ is the electric scalar potential, and $\mathbf{A}$ is the magnetic vector potential.

C.1.1 Augmentation tensor, $G^{\beta\alpha}$

One can define $\mathcal{L}(A_\alpha, \partial_\beta A_\alpha) =: \mathcal{L}_{(0)} + \delta\mathcal{L}$. Then, $\mathcal{H}(A_\alpha, \Pi^{\beta\alpha})$ is given by the appropriate Legendre transformation:

$$\begin{aligned}\mathcal{H} &= \Pi^{\beta\alpha}\partial_\beta A_\alpha - \mathcal{L} \\ &= \underbrace{\left[\frac{\partial\mathcal{L}_0}{\partial(\partial_\beta A_\alpha)}\partial_\beta A_\alpha - \mathcal{L}_0\right]}_{\mathcal{H}_{(0)}} + \underbrace{\left[\frac{\partial(\delta\mathcal{L})}{\partial(\partial_\beta A_\alpha)}\partial_\beta A_\alpha - \delta\mathcal{L}\right]}_{\delta\mathcal{H}}.\end{aligned}$$

One seeks the perturbation $\delta\mathcal{H} = 0$ so as to preserve the canonical form of the Hamiltonian (bearing some similarities to setting up Hamilton-Jacobi mechanics). Therefore:

$$\delta\mathcal{L} = \frac{\partial(\delta\mathcal{L})}{\partial(\partial_\beta A_\alpha)}\partial_\beta A_\alpha.$$

This trivial partial differential equation solves to give:

$$\delta\mathcal{L} = G^{\beta\alpha}(A_\mu) \cdot \partial_\beta A_\alpha,$$

where the augmentation tensor $G^{\beta\alpha}(A_\mu)$ preserves the Lorentz invariance of the Lagrangian density, but is only a function of the 4-potential.

C.1.2 Canonical form

From $\mathcal{L} = \mathcal{L}_{(0)} + \delta\mathcal{L}$, the augmented Maxwell's equations for the system are given by the generalized Euler-Lagrange equations:

$$\partial_\beta \left[\frac{\partial\mathcal{L}}{\partial(\partial_\beta A_\alpha)}\right] - \frac{\partial\mathcal{L}}{\partial A_\alpha} = 0. \tag{C.3}$$

By examining a 1D Cartesian space with no **B**-field, one seeks an augmentation tensor that satisfies:

$$-\alpha E = \frac{1}{\epsilon_0}\left[A_\mu \frac{\partial G^{0\mu}}{\partial A_x} + \partial_x A_\mu \frac{\partial G^{1\mu}}{\partial A_x} - (G^{01} + G^{11})\right]. \tag{C.4}$$

Doing so allows the augmentation to the Gauss-Ampere law *a posteriori* to manifest as a Berk-Breizman sink of energy via a global dissipation channel [5, 29]. This is non-trivially satisfied, but the simplest case is when $\partial G^{\nu\mu}/\partial A = 0$, and:

$$G^{01} + G^{11} = -\alpha\epsilon_0(\partial_x\phi + \dot{A}). \tag{C.5}$$

One can also show that further constraints on the augmentation tensor allow Gauss' law to retain the exact same form as before. Therefore, this family of augmentation tensors produce the modified Maxwell-Ampere law and preserve the canonical form of the energy density:

$$\partial_t E = -\frac{J}{\epsilon_0} - \alpha E \quad ; \quad U = \frac{\epsilon_0}{2}\int_{\mathbb{R}} E^2 \, \mathrm{d}x. \tag{C.6}$$

The Lorentz force on charged particles due to the electromagnetic field is given by the particle Lagrangian [19]. If we examine a single particle:

$$L_p = \left[\frac{1}{2}mu_\mu u^\mu + qu^\mu A_\mu\right], \tag{C.7}$$

where m and q are the particle mass and charge, and u^μ is the 4-velocity. One can show that if one constrains the definition for the 4-potential to be invariant under the augmentation, then:

$$m\frac{\mathrm{d}v_\alpha}{\mathrm{d}t} = -q\mu_0\Pi_{\alpha\beta}\frac{\mathrm{d}x^\beta}{\mathrm{d}t}. \tag{C.8}$$

The augmentation does perturb $\Pi_{\alpha\beta}$, however the force does no work; it is in fact a fictious force, and therefore can be omitted. Finally, I omit the spatially averaged current to avoid a build up of loop voltage; one can show that if one take the spatially averaged part of the Maxwell-Ampere law:

$$-\int_{\mathbb{R}} v(f_{\mathrm{ion},0} + f_{\mathrm{tur},0})\ \mathrm{d}v = \partial_t E_0 + \alpha_0 E_0.$$

We require that the mean current is very small, and is dominated by exponential decay. In such a case, one finds that the spatially averaged electic field $E_0(t)$ must be temporally evanescent:

$$E_0(t) \approx E_0(t=0)\mathrm{e}^{-\alpha_0 t}.$$

This in turn allows one to remove the spatially averaged electric field by setting $E_0(t = 0) = 0$ as a boundary condition.

C.2 Seed electric field

If one chooses the following form for the seed contribution:

$$\mathcal{S}_j := -\frac{1}{2}\sum_s \left[A_{js}\mathrm{e}^{-\mathrm{i}\omega_s t} + \mathrm{c.c.}\right],$$

where $\{A_{js}\} \in \mathbb{C}$, one finds the following partial differential equation:

$$\partial_t E_{\mathcal{S},j} + \frac{1}{2}\alpha_j E_{\mathcal{S},j} = \frac{1}{2}\sum_s \left[A_{js}\mathrm{e}^{-\mathrm{i}\omega_s t} + \mathrm{c.c.}\right].$$

Under a forward Laplace transformations:

$$\tilde{E}_{\mathcal{S},j} = \frac{1}{p + \alpha_j/2} \left\{ E_{\mathcal{S},j}(t=0) + \sum_s \left[\frac{A_{js}}{p + \mathrm{i}\omega_s} + \frac{A_{js}^*}{p + \mathrm{i}\omega_s} \right] \right\}.$$

Formally, for convergence of the forward transform:

$$\exists \sigma < \mathrm{Re}(p) : \lim_{t\to\infty} |f(t)| = \mathrm{e}^{\sigma t}.$$

Therefore, if the traditional Bromwich contour is shifted to examine a line integral along $\Re(p) < \sigma$, one can examine singularities in $\tilde{E}_{\mathcal{S},j}$; residues of these singularities allow us to recover the solution for the electric field via the residue theorem. By examining $\Re(p) \to -\infty$, one findsthat the only remaining contribution to the backward transform is given by the residues, as the rest of the integral becomes exponentially small:

$$E_{\mathcal{S},j} = 2\pi\mathrm{i} \cdot \sum_r \mathrm{Res}(\tilde{E}_{\mathcal{S},j}\mathrm{e}^{pt}, p_r),$$

where p_r are the locations of the singularities for $\Re(p) < \sigma$. When calculated, this yields:

$$E_{\mathcal{S}}(x,t) = E_{\mathcal{S}}^{(\mathrm{ev})}(x)\mathrm{e}^{-\alpha_j t/2} + \frac{1}{2}\sum_s \left[E_{\mathcal{S},js}(x)\mathrm{e}^{-\mathrm{i}\omega_s t} + \mathrm{c.c.} \right],$$

where $E_{\mathcal{S}}^{(\mathrm{ev})}$ refers to non-propagating evanescent modes given by the simple pole at $p = -\alpha_j/2$. One can deny these modes from existing by setting $E_{\mathcal{S}}^{(\mathrm{ev})} = 0$ as a boundary condition. $E_{\mathcal{S},js}(x)$ is given by:

$$E_{\mathcal{S},js} = \frac{A_{js}}{\alpha_j/2 - \mathrm{i}\omega_s}.$$

If one selects $A_s \in \mathbb{R}$, we find that this reduces to the form:

$$E_{\mathcal{S}}(x,t) = \sum_s \frac{A_s}{\alpha_j^2/4 + \omega_s^2} \cos(\omega_s t).$$

If one solves the modified Maxwell-Ampere law when there is very little change to the particle distribution function (PDF) (negligible instability drive) and no noise, one finds the solution:

$$\delta E(t) = -\epsilon_E E_{\mathcal{S}}(x,t).$$

As can be seen, for these modes there is no net electric field overall; the seed mode and the perturbation are counterpropagating.

One can then consider the addition of a noise term $E_{\mathcal{N}}$. By similar analysis, one can show that the noise is representable as a distinct set of frequencies, $\{\omega_n\}$. However, in the limit that $\{\omega_n\}$ form a continuum, one can represent the noise term contribution in the form:

$$\lim_{\Delta\omega_n\to 0}\frac{1}{2}\sum_n\left[A_n\mathrm{e}^{-\mathrm{i}\omega_n t}+\text{c.c.}\right]\equiv-\mathcal{N}_E,$$

where $\mathcal{N}_E$ is a pseudorandom noise term that seeds instabilities. One can once again ignore evanescent effects via boundary conditions, leaving only the propagating contribution. Again, one finds from a similar analysis that the noise term leads to no net perturbation of E, in accordance with the conservation of energy.

However, one can show that by solving (5.15d) for negligible current:

$$\delta E_j(t+\Delta t)\approx\frac{\mathcal{D}_j}{\alpha_j/2}\left[1-\exp\left(-\frac{1}{2}\alpha_j\Delta t\right)\right]+\mathcal{O}(\Delta t^2).$$

The numerical flaw associated with timestep size is what gives the initial drive; clearly, the limit as $\Delta t\to 0$ yields $\delta E_j=0$.

Bibliography

[1] E. Noether. “Invariante Variationsprobleme”. In: *Machr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math.–Phys. Klasse* (1918), pp. 235–257.

[2] D. J. Korteweg and G. de Vries. “XLI. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves”. In: *London, Edinburgh, Dublin Philos. Mag. J. Sci.* 39.240 (1895), pp. 422–443. ISSN: 1941-5982. DOI: 10.1080/14786449508620739.

[3] N. J. Zabusky and M. D. Kruskal. “Interaction of "Solitons" in a Collisionless Plasma and the Recurrence of Initial States”. In: *Phys. Rev. Lett.* 15.6 (1965), pp. 240–243. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.15.240.

[4] P. Emplit et al. “Picosecond steps and dark pulses through nonlinear single mode fibers”. In: *Opt. Commun.* 62.6 (1987), pp. 374–379. ISSN: 0030-4018. DOI: 10.1016/0030-4018(87)90003-4.

[5] H. L. Berk, B. N. Breizman, and N. V. Petviashvili. “Spontaneous hole-clump pair creation in weakly unstable plasmas”. In: *Phys. Lett. A* 234.3 (1997), pp. 213–218. ISSN: 0375-9601. DOI: 10.1016/S0375-9601(97)00523-9.

[6] A. A. Vlasov. “О вибрационных свойствах электронного газа”. In: *ЖЭТФ* 8.3 (1938), pp. 291–318.

[7] I. B. Bernstein, J. M. Greene, and M. D. Kruskal. “Exact Nonlinear Plasma Oscillations”. In: *Phys. Rev.* 108.3 (1957), pp. 546–550. ISSN: 0031-899X. DOI: 10.1103/PhysRev.108.546.

[8] H. L. Berk, B. N. Breizman, and M. Pekker. “Numerical simulation of bump-on-tail instability with source and sink”. In: *Phys. Plasmas* 2.8 (1995), pp. 3007–3016. ISSN: 1070-664X. DOI: 10.1063/1.871198.

[9] H. L. Berk, B. N. Breizman, and M. Pekker. "Nonlinear Dynamics of a Driven Mode near Marginal Stability". In: *Phys. Rev. Lett.* 76.8 (1996), pp. 1256–1259. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.76.1256.

[10] J. Wesson. *Tokamaks*. Oxford University Press, 2011.

[11] F. F Chen. *Introduction to Plasma Physics and Controlled Fusion*. 2nd. Vol. 1. Plenum Press, 1984.

[12] J. Citrin et al. "Tractable flux-driven temperature, density, and rotation profile evolution with the quasilinear gyrokinetic transport model QuaLiKiz". In: *Plasma Phys. Control. Fusion* 59.12 (2017), p. 124005. DOI: 10.1088/1361-6587/aa8aeb.

[13] C. Rea and R. S. Granetz. "Exploratory Machine Learning Studies for Disruption Prediction Using Large Databases on DIII-D". In: *Fusion Sci. Technol.* 74.1-2 (2018), pp. 89–100. DOI: 10.1080/15361055.2017.1407206.

[14] M. D. Boyer, S. Kaye, and K. Erickson. "Real-time capable modeling of neutral beam injection on NSTX-U using neural networks". In: *Nucl. Fusion* 59.5 (2019), p. 056008. DOI: 10.1088/1741-4326/ab0762.

[15] L. Friedland, P. Khain, and A. G. Shagalov. "Autoresonant Phase-Space Holes in Plasmas". In: *Phys. Rev. Lett.* 96.22 (2006), p. 225001. DOI: 10.1103/PhysRevLett.96.225001.

[16] R. M. Dreizler and J. da Providência. *Density Functional Methods in Physics*. 1st. Springer-Verlag, 2013.

[17] E. Kreyszig. *Differential Geometry*. Dover Publications, 1991, p. 352. ISBN: 0486667219.

[18] H. Goldstein, C. P. Poole, and J. L. Safko. *Classical Mechanics*. 3rd. 1950, p. 646. ISBN: 8131758915.

[19] M. Burgess. *Classical Covariant Fields*. 1st. Cambridge University Press, 2002.

[20] Y. L. Klimontovich and D. ter Haar. *The statistical theory of non-equilibrium processes in a plasma*. Pergamon Press, 1967. ISBN: 9781483214627.

[21] J. G. Cordey. "Effects of particle trapping on the slowing-down of fast ions in a toroidal plasma". In: *Nucl. Fusion* 16.3 (1976), p. 499. DOI: 10.1088/0029-5515/16/3/014.

[22] J. D. Gaffey. "Energetic ion distribution resulting from neutral beam injection in tokamaks". In: *J. Plasma Phys.* 16.2 (1976), pp. 149–169. DOI: 10.1017/S0022377800020134.

[23] R. B. White. *The Theory of Toroidally Confined Plasmas*. IMPERIAL COLLEGE PRESS, 2014. ISBN: 978-1-78326-363-9. DOI: 10.1142/p916.

[24] V. N. Duarte. "Quasilinear and nonlinear dynamics of energetic-ion-driven Alfvènic eigenmodes". PhD thesis. Universidade de São Paulo, 2017.

[25] B. J. Q. Woods. "Analytical solutions for nonlinear plasma waves with time-varying complex frequency". In: *Plasma Res. Express* 1.4 (2019), p. 045003. ISSN: 2516-1067. DOI: 10.1088/2516-1067/ab5052. arXiv: 1905.03104. URL: https://doi.org/10.1088/2516-1067/ab5052.

[26] A. J. De-Gol. "Nonlinear wave-particle phenomena in a Berk-Breizman Vlasov-Maxwell system". PhD thesis. University of York, 2010.

[27] T. D. Arber and R. G. L. Vann. "A Critical Comparison of Eulerian-Grid-Based Vlasov Solvers". In: *J. Comput. Phys.* 180.1 (2002), pp. 339–357. DOI: 10.1006/JCPH.2002.7098.

[28] B. N. Breizman et al. "Critical nonlinear phenomena for kinetic instabilities near threshold". In: *Phys. Plasmas* 4.5 (1997), pp. 1559–1568. DOI: 10.1063/1.872286.

[29] R. G. L. Vann et al. "Fully nonlinear phenomenology of the Berk–Breizman augmentation of the Vlasov–Maxwell system". In: *Phys. Plasmas* 10.3 (2003), pp. 623–630. ISSN: 1070-664X. DOI: 10.1063/1.1539854.

[30] M. Lesur, Y. Idomura, and X. Garbet. "Fully nonlinear features of the energetic beam-driven instability". In: *Phys. Plasmas* 16.9 (2009), p. 092305. DOI: 10.1063/1.3234249.

[31] B. J. Q. Woods et al. "Stochastic effects on phase-space holes and clumps in kinetic systems near marginal stability". In: *Nucl. Fusion* 58.8 (2018), p. 082015. ISSN: 17414326. DOI: 10.1088/1741-4326/aaa9fd.

[32] B. J. Q. Woods et al. "Machine Learning Characterization of Alfvénic and Sub-Alfvénic Chirping and Correlation with Fast-Ion Loss at NSTX". In: *IEEE Trans. Plasma Sci.* 48.1 (2020), pp. 71–81. ISSN: 19399375. DOI: 10.1109/TPS.2019.2960206. arXiv: 1903.05213. URL: https://doi.org/10.1109/TPS.2019.2960206.

[33] L. D. Landau. "On the vibrations of the electronic plasma". In: *Zh. Eksp. Teor. Fiz.* 10.1 (1946), p. 25.

[34] M. N. Rosenbluth and S. V. Putvinski. "Theory for avalanche of runaway electrons in tokamaks". In: *Nucl. Fusion* 37.10 (1997), pp. 1355–1362. ISSN: 0029-5515. DOI: 10.1088/0029-5515/37/10/I03.

[35] M. Lesur et al. "Nonlinear excitation of subcritical fast ion-driven modes". In: *Nucl. Fusion* 56.5 (2016), p. 056009. ISSN: 0029-5515. DOI: 10.1088/0029-5515/56/5/056009.

[36] Z. Lin et al. "Turbulent transport reduction by zonal flows: massively parallel simulations". In: *Science* 281.5384 (1998), pp. 1835–7. ISSN: 1095-9203. DOI: 10.1126/SCIENCE.281.5384.1835.

[37] J. Anderson and B. Hnat. "Statistical analysis of Hasegawa-Wakatani turbulence". In: *Phys. Plasmas* 24.6 (2017), p. 062301. ISSN: 1070-664X. DOI: 10.1063/1.4984985.

[38] J. Lang, G.-Y. Fu, and Y. Chen. "Nonlinear simulation of toroidal Alfvén eigenmode with source and sink". In: *Phys. Plasmas* 17.4 (2010), p. 042309. ISSN: 1070-664X. DOI: 10.1063/1.3394702.

[39] N. N. Gorelenkov et al. "Resonance broadened quasi-linear (RBQ) model for fast ion distribution relaxation due to Alfvénic eigenmodes". In: *Nucl. Fusion* 58.8 (2018), p. 082016. ISSN: 0029-5515. DOI: 10.1088/1741-4326/aac72b.

[40] W. W. Heidbrink. "Basic physics of Alfvén instabilities driven by energetic particles in toroidally confined plasmas". In: *Phys. Plasmas* 15.5 (2008), p. 055501. ISSN: 1070-664X. DOI: 10.1063/1.2838239.

[41] M. K. Lilley, B. N. Breizman, and S. E. Sharapov. "Effect of dynamical friction on nonlinear energetic particle modes". In: *Phys. Plasmas* 17.9 (2010), p. 092305. ISSN: 1070-664X. DOI: 10.1063/1.3486535.

[42] M. K. Lilley and R. M. Nyqvist. "Formation of Phase Space Holes and Clumps". In: *Phys. Rev. Lett.* 112.15 (2014), p. 155002. DOI: 10.1103/PhysRevLett.112.155002.

[43] G. B. Arfken, H.-J. Weber, and F. E. Harris. *Mathematical methods for physicists*. Academic, 2012. ISBN: 0123846544.

[44] P. L. Bhatnagar, E. P. Gross, and M. Krook. "A Model for Collision Processes in Gases. I. Small Amplitude Processes in Charged and Neutral One-Component Systems". In: *Phys. Rev.* 94.3 (1954), pp. 511–525. DOI: 10.1103/PhysRev.94.511.

[45] M. Abramowitz and I. A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables*. 1965.

[46] B. D. Fried and S. D. Conte. *The Plasma Dispersion Function: The Hilbert Transform of the Gaussian*. Academic Press, New York, 1961.

[47] J. M. Heninger and P. J. Morrison. "An integral transform technique for kinetic systems with collisions". In: *Phys. Plasmas* 25.8 (2018), p. 082118. ISSN: 1070-664X. DOI: 10.1063/1.5046194.

[48] G. Wang et al. "Frequency chirping in the Alfvén continuum". In: *Nucl. Fusion* 58.8 (2018), p. 082014. ISSN: 0029-5515. DOI: 10.1088/1741-4326/aab502.

[49] M. A. Van Zeeland et al. "Alfvén eigenmodes and fast ion transport in negative triangularity DIII-D plasmas". In: *Nucl. Fusion* 59.8 (2019), p. 086028. DOI: 10.1088/1741-4326/ab2488.

[50] S. E. Sharapov et al. "Alfvén wave cascades in a tokamak". In: *Phys. Plasmas* 9.5 (2002), pp. 2027–2036. ISSN: 1070-664X. DOI: 10.1063/1.1448346.

[51] M. P. Gryaznevich and S. E. Sharapov. "Perturbative and non-perturbative modes in START and MAST". In: *Nucl. Fusion* 46.10 (2006), S942–S950. ISSN: 0029-5515. DOI: 10.1088/0029-5515/46/10/S11. URL: http://stacks.iop.org/0029-5515/46/i=10/a=S11?key=crossref.33bda8e4535b45517b40e08662c3b345.

[52] M. Podestà et al. "Experimental studies on fast-ion transport by Alfvén wave avalanches on the National Spherical Torus Experiment". In: *Phys. Plasmas* 16.5 (2009), p. 056104. ISSN: 1070-664X. DOI: 10.1063/1.3080724.

[53] G. Strang. "On the Construction and Comparison of Difference Schemes". In: *SIAM J. Numer. Anal.* 5.3 (1968), pp. 506–517. DOI: 10.1137/0705041.

[54] C. Z. Cheng and G. Knorr. "The integration of the vlasov equation in configuration space". In: *J. Comput. Phys.* 22.3 (1976), pp. 330–351. DOI: 10.1016/0021-9991(76)90053-X.

[55] F. Schur. "Neue Begründung der Theorie der endlichen Transformationsgruppen". In: *Math. Ann.* 35.1-2 (1889), pp. 161–197. DOI: 10.1007/BF01443876.

[56] M. Frigo and S. G. Johnson. "The Design and Implementation of FFTW3". In: *Proc. IEEE* 93.2 (2005), pp. 216–231. DOI: 10.1109/JPROC.2004.840301.

[57] P. Colella and P. R. Woodward. "The Piecewise Parabolic Method (PPM) for gas-dynamical simulations". In: *J. Comput. Phys.* 54.1 (1984), pp. 174–201. DOI: 10.1016/0021-9991(84)90143-8.

[58] R. Courant, K. Friedrichs, and H. Lewy. "Über die partiellen Differenzengleichungen der mathematischen Physik". In: *Math. Ann.* 100.1 (1928), pp. 32–74. DOI: 10.1007/BF01448839.

[59] H. Wang, Y. Todo, and C. C. Kim. "Hole-Clump Pair Creation in the Evolution of Energetic-Particle-Driven Geodesic Acoustic Modes". In: *Phys. Rev. Lett.* 110.15 (2013), p. 155006. ISSN: 0031-9007. DOI: 10.1103/PhysRevLett.110.155006.

[60] E. D. Fredrickson et al. "Fast ion loss in a 'sea-of-TAE'". In: *Nucl. Fusion* 46.10 (2006), S926–S932. ISSN: 0029-5515. DOI: 10.1088/0029-5515/46/10/S09.

[61] E. D. Fredrickson et al. "Collective fast ion instability-induced losses in National Spherical Tokamak Experiment". In: *Phys. Plasmas* 13.5 (2006), p. 056109. DOI: 10.1063/1.2178788.

[62] R. Nazikian et al. "Intense Geodesic Acousticlike Modes Driven by Suprathermal Ions in a Tokamak Plasma". In: *Phys. Rev. Lett.* 101.18 (2008), p. 185001. DOI: 10.1103/PhysRevLett.101.185001.

[63] J. Citrin et al. "Comparison between measured and predicted turbulence frequency spectra in ITG and TEM regimes". In: *Plasma Phys. Control. Fusion* 59.6 (2017), p. 064010. DOI: 10.1088/1361-6587/aa6d1d.

[64] V. N. Duarte et al. "Theory and observation of the onset of nonlinear structures due to eigenmode destabilization by fast ions in tokamaks". In: *Phys. Plasmas* 24.12 (2017), p. 122508. DOI: 10.1063/1.5007811.

[65] S. D. Pinches et al. "The role of energetic particles in fusion plasmas". In: *Plasma Phys. Control. Fusion* 46.12B (2004), B187–B200. ISSN: 0741-3335. DOI: 10.1088/0741-3335/46/12B/017.

[66] N. N. Gorelenkov, S. D. Pinches, and K. Toi. "Energetic particle physics in fusion research in preparation for burning plasma experiments". In: *Nucl. Fusion* 54.12 (2014), p. 125001. ISSN: 0029-5515. DOI: 10.1088/0029-5515/54/12/125001.

[67] M. García-Muñoz et al. "NTM induced fast ion losses in ASDEX Upgrade". In: *Nucl. Fusion* 47.7 (2007), pp. L10–L15. DOI: 10.1088/0029-5515/47/7/L03.

[68] M. García-Muñoz et al. "Fast-Ion Losses due to High-Frequency MHD Perturbations in the ASDEX Upgrade Tokamak". In: *Phys. Rev. Lett.* 100.5 (2008), p. 055005. DOI: 10.1103/PhysRevLett.100.055005.

[69] B. N. Breizman. "Nonlinear travelling waves in energetic particle phase space". In: *Nucl. Fusion* 50.8 (2010), p. 084014. DOI: 10.1088/0029-5515/50/8/084014.

[70] S. J. Linz. "Nonlinear dynamical models and jerky motion". In: *Am. J. Phys.* 65.6 (1997), pp. 523–526. DOI: 10.1119/1.18594.

[71] H. Poincaré. "Sur les courbes definies par les equations differentielles". In: *J. Math. Pures Appl.* 4 (1885), pp. 167–244.

[72] I. Bendixson. "Sur les courbes définies par des équations différentielles". In: *Acta Math.* 24.0 (1901), pp. 1–88. DOI: 10.1007/BF02403068.

[73] K. E. Chlouverakis and J. C. Sprott. "Chaotic hyperjerk systems". In: *Chaos, Solitons & Fractals* 28.3 (2006), pp. 739–746. DOI: 10.1016/J.CHAOS.2005.08.019.

[74] V. N. Duarte et al. "Prediction of nonlinear evolution character of energetic-particle-driven instabilities". In: *Nucl. Fusion* 57.5 (2017), p. 054001. DOI: 10.1088/1741-4326/aa6232.

[75] M. Lesur and Y. Idomura. "Nonlinear categorization of the energetic-beam-driven instability with drag and diffusion". In: *Nucl. Fusion* 52.9 (2012), p. 094004. DOI: 10.1088/0029-5515/52/9/094004.

[76] S.-I. Itoh and K. Itoh. "Model of L to H-Mode Transition in Tokamak". In: *Phys. Rev. Lett.* 60.22 (1988), pp. 2276–2279. DOI: 10.1103/PhysRevLett.60.2276.

[77] J. W. Connor. "Edge-localized modes - physics and theory". In: *Plasma Phys. Control. Fusion* 40.5 (1998), pp. 531–542. DOI: 10.1088/0741-3335/40/5/002.

[78] E. D. Fredrickson et al. "Parametric dependence of fast-ion transport events on the National Spherical Torus Experiment". In: *Nucl. Fusion* 54.9 (2014), p. 093007. DOI: 10.1088/0029-5515/54/9/093007.

[79] A. Pau et al. "A Machine Learning approach based on Generative topographic mapping for disruption prevention and avoidance at JET". In: *Nucl. Fusion* (2019). DOI: 10.1088/1741-4326/ab2ea9.

[80] B. K. Spears et al. "Deep learning: A guide for practitioners in the physical sciences". In: *Phys. Plasmas* 25.8 (2018), p. 080901. DOI: 10.1063/1.5020791.

[81] F. Glover. "Future paths for integer programming and links to artificial intelligence". In: *Comput. Oper. Res.* 13.5 (1986), pp. 533–549. DOI: 10.1016/0305-0548(86)90048-1.

[82] L. Breiman et al. *Classification And Regression Trees*. 1st. Routledge, 1984. ISBN: 9781315139470. DOI: 10.1201/9781315139470.

[83] J. R. Quinlan. "Induction of decision trees". In: *Mach. Learn.* 1.1 (1986), pp. 81–106. DOI: 10.1007/BF00116251.

[84] L. Breiman. "Random Forests". In: *Mach. Learn.* 45.1 (2001), pp. 5–32. DOI: 10.1023/A:1010933404324.

[85] A. Bierwage et al. "Simulations tackle abrupt massive migrations of energetic beam ions in a tokamak plasma". In: *Nat. Commun.* 9.1 (2018), p. 3282. DOI: 10.1038/s41467-018-05779-0.

[86] S. R. Haskey, B. D. Blackwell, and D. G. Pretty. "Clustering of periodic multi-channel timeseries data with application to plasma fluctuations". In: *Comput. Phys. Commun.* 185.6 (2014), pp. 1669–1680. ISSN: 00104655. DOI: 10.1016/j.cpc.2014.03.008.

[87] S. R. Haskey et al. "Experiment-theory comparison for low frequency BAE modes in the strongly shaped H-1NF stellarator". In: *Plasma Phys. Control. Fusion* 57.9 (2015). ISSN: 13616587. DOI: 10.1088/0741-3335/57/9/095011.

[88] J. Breslau et al. *TRANSP*. 2018. DOI: 10.11578/dc.20180627.4.

[89] E. J. Synakowski et al. "The national spherical torus experiment (NSTX) research programme and progress towards high beta, long pulse operating scenarios". In: *Nucl. Fusion* 43.12 (2003), p. 1653. DOI: 10.1088/0029-5515/43/12/011.

[90] D. S. Darrow et al. "Stochastic orbit loss of neutral beam ions from NSTX due to toroidal Alfvén eigenmode avalanches". In: *Nucl. Fusion* 53.1 (2013), p. 013009. DOI: 10.1088/0029-5515/53/1/013009.

[91] G. J. Kramer et al. "Improving fast-ion confinement in high-performance discharges by suppressing Alfvén eigenmodes". In: *Nucl. Fusion* 57.5 (2017), p. 056024. DOI: 10.1088/1741-4326/aa6456.

[92] C. S. Collins et al. "Phase-space dependent critical gradient behavior of fast-ion transport due to Alfvén eigenmodes". In: *Nucl. Fusion* 57.8 (2017), p. 086005. DOI: 10.1088/1741-4326/aa720c.

[93] F.J. Harris. "On the use of windows for harmonic analysis with the discrete Fourier transform". In: *Proc. IEEE* 66.1 (1978), pp. 51–83. ISSN: 0018-9219. DOI: 10.1109/PROC.1978.10837. URL: http://ieeexplore.ieee.org/document/1455106/.

[94] F. Pedregosa et al. "Scikit-learn: Machine Learning in {P}ython". In: *J. Mach. Learn. Res.* 12 (2011), pp. 2825–2830.

[95] J. Gorodkin. "Comparing two K-category assignments by a K-category correlation coefficient". In: *Comput. Biol. Chem.* 28.5-6 (2004), pp. 367–374. DOI: 10.1016/J.COMPBIOLCHEM.2004.09.006.

[96] H. L. Berk et al. "Theoretical Interpretation of Alfvén Cascades in Tokamaks with Nonmonotonic q Profiles". In: *Phys. Rev. Lett.* 87.18 (2001), p. 185002. DOI: 10.1103/PhysRevLett.87.185002.

[97] F. M. Levinton and H. Yuh. "The motional Stark effect diagnostic on NSTX". In: *Rev. Sci. Instrum.* Vol. 79. 10. 2008. DOI: 10.1063/1.2968699.

[98] A. Ödblom et al. "Nonlinear magnetohydrodynamical effects in precessional fishbone oscillations". In: *Phys. Plasmas* 9.1 (2002), p. 155. ISSN: 1070664X. DOI: 10.1063/1.1421373.

[99] W. W. Heidbrink. "Beam-driven chirping instability in DIII-D". In: *Plasma Phys. Control. Fusion* 37.9 (1995), pp. 937–949. DOI: 10.1088/0741-3335/37/9/002.

[100] V. N. Duarte et al. "Study of the likelihood of Alfvénic mode bifurcation in NSTX and predictions for ITER baseline scenarios". In: *Nucl. Fusion* 58.8 (2018), p. 082013. DOI: 10.1088/1741-4326/aab37c.

[101] H. Hezaveh et al. "Impact of energetic particle orbits on long range frequency chirping of BGK modes". In: *Nucl. Fusion* 57.12 (2017). ISSN: 17414326. DOI: 10.1088/1741-4326/aa80a9.

[102] A. V. Dudkovskaia et al. "Island Stability in Phase Space". In: *J. Phys. Conf. Ser.* Vol. 1125. 1. Institute of Physics Publishing, 2018. DOI: 10.1088/1742-6596/1125/1/012009.

[103] J. Lang and G.-Y. Fu. "Nonlinear simulation of toroidal Alfvén eigenmode with microturbulence-induced radial diffusion". In: *Phys. Plasmas* 18.5 (2011), p. 055902. DOI: 10.1063/1.3574503.

[104] F. Wagner. "A quarter-century of H-mode studies". In: *Plasma Phys. Control. Fusion* 49.12B (2007), B1–B33. DOI: 10.1088/0741-3335/49/12B/S01.

Glossary

Symbols

δ-operator An operator that returns the variation of a functional. See Section 1.3.2.

A

action-angle variables A set of phase-space variables (generalised momentum and generalised coordinates respectively), for which the momenta resemble the classical action. See Section 1.6.

adiabaticity Of or pertaining to a slow or zero change of the following integral:

$$\oint_{H=H(\mathbf{x},\mathbf{v},t)} \mathbf{v}\cdot\mathrm{d}\mathbf{x},$$

such that one integrates over a contour of constant energy at time t. See Section 2.3.7.

Alfvén waves[11] A family of magnetohydrodynamic wave. Dispersion relations are defined with respect to the *Alfvén* speed:[11]

$$v_A = \frac{B}{\sqrt{\sum_i \mu_0 m_i n_i}}.$$

B

Berk-Breizman sink A sink of electromagnetic energy in plasmas popularised by Berk and Breizman[5] and later widely used in the literature. See Section 3.3.3

Bernstein-Green-Kruskal Of or pertaining to structures, waves or theory surrounding the 1958 paper by Bernstein, Greene and Kruskal. See Section 2.2.1.

Boltzmann equation Kinetic equation which governs the evolution of a particle distribution function (PDF), f, for the l^{th} species. See Section 1.4.4.

Boltzmann-Maxwell Of or pertaining to the coupled system of equations featuring the Boltzmann equation and Maxwell's equations. See (1.17).

bounce frequency Also sometimes called the transit frequency. This is the time it takes for an angle variable to make a full rotation.

Bromwich contour The contour used for the inverse Laplace transform $\tilde{f}(p) \to f(t)$. Given by $\Omega_{\text{Br.}} : p \in [-\infty + \mathrm{i}\sigma, \infty + \mathrm{i}\sigma]$. See Landau's method.

bump-on-tail distribution A simple particle distribution function featuring suprathermal particles. Typically used in reduced dimensionality models. See Section 1.5.1.

bursting A plasma wave is 'bursting' if it is undergoing one or more cycles in the mode amplitude including the following stages: linear growth, nonlinear saturation, relaxation. For a sketch of a burst, see Figure 5.5.

C

canonical transformation A phase space transformation which preserves the same form of Hamilton's equations.

Cauchy's theorem[43] The line integral on a closed contour C in the complex plane of a function that is holomorphic in a subset of the complex plane A is equal to zero:

$$\oint_C f(z)\,\mathrm{d}z = 0,$$

where f is holomophic in A, and C is a closed contour in or on the boundary of A.

co-moving Galilean transform Two tensors describe Galilean transforms in this thesis: spatial transform (χ, see Section 3.4.2) and velocity transform (ν, see Section 3.5.1).

confusion matrix The confusion matrix C for a multi-class classifier is defined such that the element C_{ij} is the number of classifications made belong to class i which are actually of class j. See Section 6.2.5.

conservative orbit A conservative orbit is defined in this thesis as a single particle orbit upon which the particle energy is conserved. See Chapter 2.

D

decision tree classifier A decision tree setup to perform classification. It takes an input space (features) and maps those features onto an output space (classification). In this thesis, I examine multi-class classification. See Chapter 6.

Dirichlet boundary condition Also known as a 'first-type' boundary condition. A condition that a solution to a function which satisfies a differential equation must be equal to a given value at a given point in its codomain.

E

energetic particle For a given plasma species, energetic particles are a subpopulation with mean velocity much greater than the bulk thermal velocity, who have not yet thermalised with the bulk.

ensemble average Averaging process to remove fluctuations from a quantity, leaving only the average/mean-field contribution. See Section 1.4.4.

F

frequency sweeping A slow change in frequency for a wave (as opposed to frequency chirping). Typically caused by changes to the background. See Section 5.1.1.

frequency chirping A fast change in frequency for a wave (as opposed to frequency sweeping). Typically caused by nonlinear kinetic resonance. See Section 5.1.1.

G

Gini impurity A nonlinear function which determines how well sorted the data is in a decision tree. Value given by:

$$I_G := \sum_k \left[p_k (1 - p_k) \right].$$

See Section 6.2 for more details.

H

hole and clump During kinetic resonance between a plasma species and an electromagnetic wave, two nonlinear structures can form on the distribution function near the resonance. These manifest as a relative decrease or increase on the distribution function (hole and clump). See Chapter 5.

I

input space This can essentially can be thought of as the codomain of a function or map. Example: $\mathbb{R}$ forms the input space for the function $f(x) = x^2$. See Section 6.2.

J

Jacobian[43] A matrix containing all of a vector function's first-order partial derivatives; commonly used for area and volume transformations. For a mapping $\{X_j\} \to \{Y_i\}$, using matrix indices:

$$\mathcal{J}_{ij} = \frac{\partial Y_i}{\partial X_j}.$$

Jordan's lemma[43] If a function $f(z)$ is of the form $f(x) = e^{iaz}g(z)$, on a semicircular contour $C_R = Re^{i\theta}|\theta \in [0, \pi]$, if $\lim_{R\to\infty} g(Re^{i\theta}) = 0$:

$$\lim_{R\to\infty} \int_{C_R} f(z)\,\mathrm{d}z = 0.$$

K

Klimontovich density A coarse distribution function constructed from Dirac delta functions. See Section 1.4.3.

L

Landau's method A method of obtaining the complex linear dispersion relation. See Section 2.3.3.

leaf node A node on a branch of a decision tree. At each leaf, the tree branches. See Section 6.2.

Leibniz integral rule Also known as differentiation under the integral sign. See Appendix B.2.

linear complex dispersion relation The linear complex dispersion relation. See Section 2.3.1.

Liouville's theorem[18] The phase space distribution is constant along trajectories of the system. Alternatively, phase space acts as an incompressible fluid.

Lorentz force The force exerted on a single charged particle by an electromagnetic field; depending on the context, this may be the full electromagnetic field, or the mean-field.

loss function A loss function L that maps an event nto a real number to give some notion of 'loss' associated with that event. In the context of machine learning, loss functions map the output space onto a real number which gives an overall 'figure-of-merit' for

how well the machine thinks it is performing. A machine learning algorithm then try to decrease the loss function until optimization is achieved.

N

neighbourhood On a manifold, all points which are within a distance of d away from $\mathbf{x}$ are said to be in the neighbourhood of $\mathbf{x}$.

Neumann boundary condition Also known as a 'second-type' boundary condition. A condition that the first derivative of a function which satisfies a differential equation must be equal to a given value at a given point in its codomain.

Noether's theorem[1] For each and every continuously differentiable symmetry there is a corresponding conservation law.

non-conservative orbit A non-conservative orbit is defined in this thesis as a single particle orbit upon which the particle energy is *not* conserved. See Chapter 3.

O

orbit affine parameter A quantity ϵ that is conserved along particle orbits. See Section 3.2

output space This can essentially can be thought of as the domain of a function or map. Example: $\mathbb{R}[0, \infty]$ forms the output space for the function $f(x) = x^2$. See Section 6.2.

Q

quasi-passing particle A passing particle which appears to exhibit some properties of trapped particles: restricted orbit; wave-particle resonance. See Section 3.2.1

R

random forest classifier A random ensemble of decision trees. See Section 6.2 and .

residue theorem[43] The line integral on a closed contour C in the complex plane of a function that is meromorphic in a subset of the complex plane A is equal to the product of $2\pi\mathrm{i}$ and the sum of the residues in the region bounded by C:

$$\oint_C f(z)\,\mathrm{d}z = 2\pi\mathrm{i}\sum_j \mathrm{Res}(f, p_j),$$

where f is meromorphic in A, C is a closed contour in or on the boundary of A, $\{p_j\}$ are all of the poles of f in the region bounded by C, and $\text{Res}(f, p_j)$ gives the residue of f at p_j.

S

separatrix The largest closed phase-space contour/boundary. See Theorem 1

slowing-down distribution A particle distribution function featuring suprathermal ions which are steadily injected into a system. Typically used with regards to tokamaks. See Section 1.5.2.

T

training data The set of data used to train and/or test a machine learning algorithm. See Section 6.2.

U

unperturbed orbit The orbit a particle travels along before some perturbation is added to the system. Generally superceded: see conservative orbits.

V

van Kampen mode Solutions to the Vlasov equation characterised by singular behaviour.

Vlasov equation Collisionless variant of the Boltzmann equation. See Boltzmann equation.

List of Abbreviations

A

ALE abrupt large event

B

BAE beta-induced Alfvén eigenmode

C

CART classification and regression tree

CFL Courant-Friedrich-Lewy

CNN convolutional neural network

D

DARK D-dimensional Augmented Resonance Kinetic solver

E

ELM edge localised mode

EPM energetic particle driven mode

ERICSON Experimental Resonant Instability Correlation Studies on NSTX

F

FFTW Fastest Fourier Transform in the West

FILD fast ion loss detectors

G

GAM geodesic acoustic mode

GTM generative topographic mapping

K

KTF kink/tearing/fishbone

M

MHD magnetohydrodynamics

ML machine learning

N

NBI neutral beam injection

O

ODE ordinary differential equation

P

PCA principal component analysis

PDE partial differential equation

PDF particle distribution function

PPPL Princeton Plasma Physics Laboratory

PRNG pseudorandom number generator

R

RF radio frequency

S

SHM simple harmonic motion

STFT short-time Fourier transform

T

TAE toroidal Alfvén eigenmode

Other Equations and Identities

C

complete elliptic integral of the first kind[45]

$$K(k) = \int_0^{\pi/2} \frac{\mathrm{d}y}{\sqrt{1 - k^2 \sin y}} \tag{E.1}$$

convolution theorem

$$\mathcal{F}[fg] = \mathcal{F}[f] * \mathcal{F}[g] \tag{E.2}$$

D

Dawson function[45]

$$\mathcal{D}[x] = \mathrm{e}^{-x^2} \int_0^x \mathrm{e}^{y^2} \,\mathrm{d}y \tag{E.3}$$

F

Fourier transform (velocity)

$$\mathcal{F}_v[f(v)] := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(v) \mathrm{e}^{-\mathrm{i}sv} \,\mathrm{d}v$$
$$\mathcal{F}_v^{-1}[g(s)] := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} g(s) \mathrm{e}^{\mathrm{i}sv} \,\mathrm{d}s$$

Fourier series (spatial)

$$Q(x) =: \frac{1}{2} \sum_j Q_j \mathrm{e}^{\mathrm{i}k_j x}$$
$$Q_j \equiv \frac{1}{L} \oint_0^L Q(x) \mathrm{e}^{-\mathrm{i}k_j x} \,\mathrm{d}x$$

L

Laplace transform

$$\hat{L}[Q] \equiv \tilde{Q}(p) := \int_{\mathbb{R}} Q(t)\mathrm{e}^{-pt}\,\mathrm{d}t$$
$$\hat{L}^{-1}[Q] \equiv Q(t) \equiv \frac{1}{2\pi\mathrm{i}} \int_{\mathbb{R}+\mathrm{i}\sigma} \tilde{Q}(p)\mathrm{e}^{pt}\,\mathrm{d}p$$

Laurent series

$$f(v) = \sum_{j=-\infty}^{\infty} L_j(v-u)^j \tag{E.4}$$

V

vector Laplacian[43]

$$\nabla^2\mathbf{Q} = \nabla(\nabla\cdot\mathbf{Q}) - \nabla\times(\nabla\times\mathbf{Q}) \tag{E.5}$$