The Dissertation Committee for Kun Huang
certifies that this is the approved version of the following dissertation:

A Numerical and Analytical Study of Kinetic Models for Particle-Wave Interaction in Plasmas

Committee:

Irene M. Gamba, Supervisor

Boris Breizman

Clint Dawson

Thomas J.R. Hughes

Philip J. Morrison

Robert Moser

A Numerical and Analytical Study of Kinetic Models for Particle-Wave Interaction in Plasmas

by

Kun Huang

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2023

Dedication

Dedicated to my parents, Xiaolan Liu and Xiaobing Huang, and my sister Yiming.

Acknowledgments

I would like to thank the multitudes of people who have helped me in completing this research.

First of all, I would like to express my sincere gratitude to my advisor, Professor Irene M. Gamba. I would not have completed this work without her dedicated support and guidance. Her passion and insight in research brought me the confidence to tackle problems that no one has ever attempted.

I would also like to thank Professor Boris Breizman for introducing the fascinating problems in plasma physics. I also thank Professor Chi-Wang Shu, from whom I learned a lot about DG methods. I owe special thanks to Michael Abdelmalik, who provided precious instructions on finite element methods and programming techniques. Thanks should also be given to my fellows students in the Applied Mathematics Group, for their support and encouragement during the preparation of my dissertation.

Finally, I would like to thank my parents and my sister for their unconditional support and inspiration. I thank them as well as my grandparents for their understanding and caring words across the Pacific Ocean, which brought me infinite power in my darkest moments.

Abstract

A Numerical and Analytical Study of Kinetic Models for Particle-Wave Interaction in Plasmas

Kun Huang, PhD
The University of Texas at Austin, 2023

SUPERVISOR: Irene M. Gamba

This dissertation presents a study of particle-wave interaction in plasmas. It focuses on a kinetic model called quasilinear theory, which is a reduction of Vlasov-Maxwell (or Vlasov-Poisson) system in the weak turbulence regime. The quantized waves in plasmas, known as plasmons, are absorbed or emitted by charged particles. Meanwhile, the particles change their states due to such emission/absorption process, therefore resulting in a nonlinear kinetic system for the *pdf* (probability density function) of particles and plasmons. The research presented here unfolds in two main topics: structure-preserving numerical solvers, and solvability of the kinetic model.

On the first topic, we are interested in numerical simulation of non-uniform magnetized plasmas, which involves two processes: particle-wave interaction and wave propagation (plasmon advection).

For particle-wave interaction in homogeneous magnetized plasmas, we propose a finite element scheme that preserves all the conservation laws. Firstly, an unconditionally conservative weak form is constructed. By "unconditional" we mean that conservation is independent of the transition probabilities. Then we design a discretization that preserves such unconditional conservation property, and discuss the conditions for positivity and stability. We present numerical examples with a "bump

on tail" initial configuration, showing that the particle-wave interaction results in a strong anisotropic diffusion of the particles. We generalize the strategy to obtain a conservative DG (discontinuous Galerkin) scheme.

The evolution of plasmon *pdf* is governed by a Liouville equation with additional reaction term caused by particle-wave interaction, where the dominant Poisson bracket term necessitates trajectorial average. Hence, we propose a Galerkin approach for trajectorial average in dynamical systems. The weak form of averaged equation is derived, and the concept of trajectory bundle is introduced. To compute and store the trajectory bundles, we propose a novel algorithm, named connection-proportion algorithm, which transforms a continuous topological problem into a discrete graph theory problem.

The conservative DG scheme, combined with our trajectorial average method, renders a structure-preserving solver for particle-wave interaction in non-uniform magnetized plasmas. We demonstrate that discrete weak form with/without average differs only in the choice of test/trial spaces. The complexity of each procedure is analyzed. Finally, a numerical example for a non-uniform magnetized plasma in an infinitely long symmetric cylinder is presented. It is verified that the connection-proportion algorithm allows to distinguish different trajectory bundles, and the proposed DG scheme rigorously preserves all the conservation laws.

On the second topic, the existence of global weak solution to quasilinear theory for electrostatic plasmas is proved. In the one-dimensional case, both the particle *pdf* and the plasmon *pdf* can be expressed with the same auxiliary function. The auxiliary function itself, is the solution of a porous medium equation with nonlinear source terms, defined on an unbounded domain. The solvability is then proved in two steps: Firstly, the equation on finite cut-off domain with Dirichlet's boundary condition is solved. Next, the solution, extended by zero outside the cut-off domain, turns out to be a solution to the same equation on the unbounded domain.

Table of Contents

List of Figures 11
Chapter 1: Introduction 12
Chapter 2: Preliminaries 17
2.1 Generalized Hamiltonian dynamics 17
2.1.1 Hamiltonian system 17
2.1.2 Nambu system 18
2.2 Waves in plasmas 19
2.2.1 Waves in unmagnetized plasmas 20
2.2.2 Waves in magnetized plasmas 20
2.2.3 Plasmon 21
2.3 Quasilinear theory from a probablistic perspective 22
2.3.1 Master equation for unmagnetized plasmas 23
2.3.2 Master equation for magnetized plasmas 24
2.3.3 Singular transition rates and properties 26
2.4 Kinetic model for electron runaway in tokamaks 28
2.4.1 Geometry 28
2.4.2 Kinetic model 29
Chapter 3: Conservative schemes for particle-wave interaction in homogeneous magnetized plasmas 31
3.1 Introduction 31
3.2 Electron energy and plasmon energy 33
3.3 Equations reformulated 39
3.3.1 Conditionally conservative weak form 39
3.3.2 Numerical integration on the resonance manifold 43
3.3.3 Unconditionally conservative weak form 45
3.4 The conservative discretization 49
3.4.1 The finite element discretization 49
3.4.2 The conservative semi-discrete form 52
3.5 Sparse interaction tensors 55
3.6 Stability and positivity 58
3.6.1 Stability of the semi-discrete form 58
3.6.2 Time discretization 59

3.6.3 Positivity-preserving technique for the wave SED 61
3.7 Numerical results 66
3.7.1 Problem Setting 66
3.7.2 Temporal evolution 67
3.7.3 Verification of conservation 67
3.7.4 Comparison of different dispersion relations 69
3.8 Generalization of the strategy: a conservative LDG scheme 70
3.8.1 Bilinear form of local DG method 71
3.8.2 Conservative DG discretization 73
3.8.3 Time discretization 75
3.8.4 Implementation and numerical results 76
Chapter 4: A Galerkin approach for trajectorial average in maximally superintegrable systems 77
4.1 Introduction 77
4.2 Weak form and its discretization 79
4.3 The connection-proportion algorithm 82
4.4 Properties 89
Chapter 5: A structure-preserving solver for particle-wave interaction in non-uniform magnetized plasmas 90
5.1 Electron-plasmon kinetic system 90
5.2 Averaged system and properties 91
5.3 Partition of the plasmon phase space 93
5.3.1 Interpolated Hamiltonian 93
5.3.2 Cut-off domains 94
5.4 DG formulation 96
5.4.1 Cut-off domains 96
5.4.2 Discrete test spaces 97
5.4.3 Unconditionally conservative semi-discrete form 97
5.5 Sparse interaction tensors 98
5.6 Complexity analysis 100
5.7 Numerical results 102

Chapter 6: Existence of global weak solutions to quasilinear theory for electrostatic plasmas 108
6.1 Introduction 108
6.2 From kinetic system to porous medium equation 109
6.3 Lifted extension of approximate solutions 115
6.4 Regularity estimates on approximate solutions 119
6.5 Convergence results 126
6.6 Existence of global weak solution 134
6.7 Appendix 138
Chapter 7: Summary 141
Appendix A: Generalized inverse inequality 143
Appendix B: Quasilinear theory: a weak turbulence model 144
B.1 Weak turbulence model in fluid dynamics 144
B.2 Weak turbulence model in plasma physics 145
Works Cited 146
Vita 151

List of Figures

1.1 Hierarchy of mathematical models 12
1.2 Magnetized plasmas 14

2.1 A trajectory in a 3-dimensional Nambu system is the intersection of two level sets. 19
2.2 Bump-on-tail velocity distribution excites waves. 22
2.3 Feynman diagrams. 23
2.4 Magnetic field configurations 29
2.5 Local frame 30

3.1 wave modes. 38
3.2 Temporal evolution of the electron *pdf* and wave *sed*. 68
3.3 The characteristics of directional differential operator $\mathcal{L}$ given $(k_\parallel, k_\perp) = (0.2, 0.3)$. 69
3.4 Comparison between ω_{exact} and ω_{approx} 70

4.1 A nice box has no saddle point on the boundary of its inverse image. 81
4.2 Examples of complicated topologies 83
4.3 The minimal triangle covers replicate the topological relation between the corresponding trajectory bundles. 87

5.1 Trajectory bundles in the plasmon phase space. 95
5.2 Complete and incomplete trajectory bundles in a cut-off domain. . . . 95
5.3 Trajectory bundles and their minimal triangle covers 104
5.4 Spatial distribution at $t = 3 \times 10^6 \frac{1}{2\pi\omega_0}$ 105
5.5 Temporal evolution at $r = \frac{31R_{max}}{40}$. 106
5.6 Conservation laws 107

6.1 The lifted extension 136

Chapter 1: Introduction

Plasma, as the most abundant form of ordinary matter in the universe, has been studied because of its intriguing properties and vast range of applications.

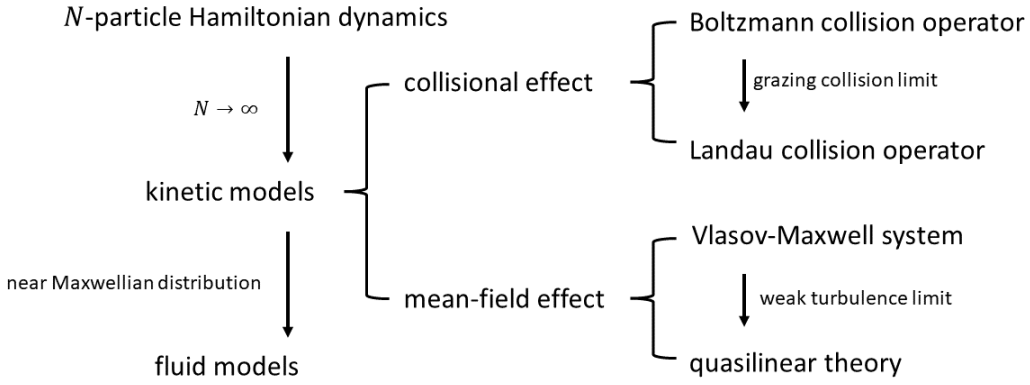


Figure 1.1: Hierarchy of mathematical models

As shown in Figure 1.1, there is a hierarchy of mathematical models for a plasma. The most detailed description, N-particle Hamiltonian dynamics, tracks the position and momentum of every particle at all times. The kinetic description is concerned about the particle probability density function in phase space. The fluid model describes the plasma based on macroscopic quantities, such as density and mean velocity. Although the fluid model is the easiest one to solve, it has a very limited scope of application because the velocity distributions are assumed to be close to the Maxwellian distribution, which is not true when collision is not dominant. In other words, a lot of information is lost after taking the moments of a distribution function. In this dissertation, we focus on the kinetic description.

Apart from the external electromagnetic field which might be generated by coils or other objects outside the plasma, the motion of a particle is influenced by other particles in within the plasma. Such influence can be categorized into two effects in kinetic model: collisional and mean-field(collective) effect. The former one can be regarded as particle-particle interaction, while the latter one can be regarded as particle-wave interaction. In general, the collisional effect is more significant in low-temperature plasma, while in high-temperature plasma, the mean-field effect dominates.

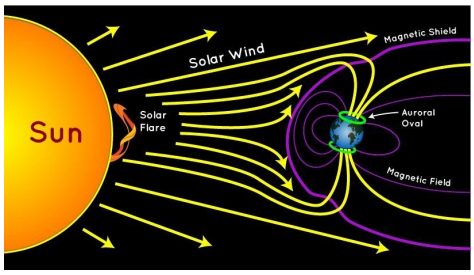
Binary collisions due to Coulomb potential is described with the Fokker-Planck-Landau operator(Landau, 1936), which is the grazing limit of a Boltzmann collision operator.

The Vlasov-Maxwell system and the Vlasov-Poisson system are widely used to describe the mean-field effect. Although a lot of work has been done in numerical methods for these systems(Heath et al., 2012; Cheng et al., 2014), in practice a reduced model is often preferred when the problem is in high dimension and some loss of details is justified from physics consideration. For example, in weak turbulence regime, the Vlasov-Maxwell(Poisson) system is asymptotic to the quasilinear theory. The quasilinear theory for unmagnetized plasmas was proposed by Vedenov et al. (1961) and Drummond and Pines (1962). It was later generalized by Shapiro and Shevchenko (1962) to model the magnetized plasma. The same idea has been used extensively in the following years, for example in the work of Kennel and Engelmann (1966), Lerche (1968), and Kaufman (1971), etc. The validity of such a model reduction was studied numerically by Besse et al. (2011), and analytically by Bardos and Besse (2021) for Vlasov-Poisson system. The quasilinear theory describes particle-wave interaction by treating waves as quasi-particles, named plasmons. A particle changes its momentum by absorbing or emitting plasmons, instead of "feeling" the Lorentz force. The absorption/emission process is governed by a resonant transition probability, which means particles and plasmons do not interact unless their momenta

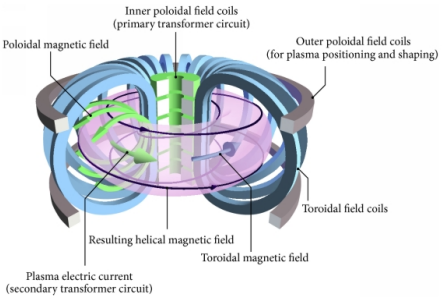
satisfy certain resonance condition. Due to such singular transition probability, the well-posedness of quasilinear theory remains an open problem.

In application aspect, we are most interested in magnetized plasmas. A magnetized plasma is one in which the ambient magnetic field $\mathbf{B}$ is strong enough to alter fluid behavior(Hazeltine and Waelbroeck, 2018). The most significant consequence of a strong background magnetic field is that electrons will gyrate around magnetic field lines with high frequency and small radius due to the strong Lorentz force. Therefore one can expect that magnetized plasmas will display totally different characteristics from unmagnetized ones.

Imagine a single electron in a magnetized plasma. It is continuously pushed by the background electromagnetic field, sometimes collides with other electrons or ions. Meanwhile, it "surfs" the waves in plasma, or in other words, emits/absorbs plasmons. Note that the plasmons are moving as well. Their motion is governed by Hamiltonian dynamics, thus can be described with Liouville equation. All these effects and behaviors intertwine with each other, rendering interesting phenomena.



(a) Solar wind(sol)



(b) Tokamak(tok)

Figure 1.2: Magnetized plasmas

People have been studying the magnetized plasma behavior in various fields, for instance astrophysics (Figure 1.2a), magnetic confinement fusion (Figure 1.2b), etc. Among which, simulating runaway electrons in tokamaks(Breizman et al., 2019) is the main motivation of my work. Runaway electrons are a group of extremely fast

electrons generated inside the tokamak, the release of which will damage the wall of the nuclear fusion reactors like ITER(International Thermonuclear Experimental Reactor). Thus it is important to have an answer for the questions like how are they generated and how to mitigate them.

Electron runaway occurs when the background electric field in a tokamak is strong enough to overcome the friction force caused by collision. In such situation, electrons are accelerated to high velocity, where relativity effect must be taken into consideration. Meanwhile, since this small group of electrons are far away from equilibrium, fluid model based on near Meaxwellian distribution assumption is no longer sufficient to describe their behavior, thus kinetic theory becomes necessary.

The study on collisional and mean-field transport of electrons in plasmas poses interesting problems in both analytical and numerical aspects.

On the analytical side, the following questions are of great concern: In what sense is the Vlasov-Maxwell system asymptotic to the quasilinear theory? Is the kinetic system given by quasilinear theory solvable? If it is solvable, what is the condition for uniqueness of solution? Does the solution change a lot when the initial value is perturbed? Those questions are not only out of mathematicians' interest, but also the foundation of rigorous error analysis for numerical schemes.

On the numerical side, in general people pursue two goals at the same time: accuracy and efficiency, and try to maintain a perfect balance between them. That is extremely challenging in solving kinetic models due to the following facts. First of all, the unknown probability density functions reside in high dimensional phase spaces, therefore it requires huge amount of resource and time to store and compute. Secondly, the kinetic equations are usually nonlinear and in integro-differential form, causing extra troubles in discretization. Thirdly, the singular transition probabilities as a result of resonance often requires carefully designed numerical integration approaches. Moreover, the multi-scale feature of kinetic models necessitates appropriate averaging methods.

However, besides accuracy and efficiency, an important factor has not gained enough attention: the delicate structure in kinetic equations originating from the underlying basic physics principles. For example, the conservation law for N-particle Hamiltonian system is inherited by kinetic models, embedded in the specific form of transition probabilities and divergence-free advection fields. Numerical schemes fail to capture physics if those structures are ruined due to discretization.

The rest of this dissertation is organized as follows.

Chapter 2 provides basic background knowledge about kinetic model of plasmas by giving a brief introduction on three topics: how waves propagate in plasmas, how waves interact with particles in plasmas and how particles interact with particles in plasmas.

Chapter 3 is devoted to the spatially homogeneous particle-wave kinetic system. In this chapter we introduce the concept of unconditional conservation. And based on that, a continuous Galerkin scheme is formulated, which preserves all the conservation laws rigorously. The strategy will then be generalized, rendering a conservative discontinuous Galerkin scheme.

Chapter 4 concerns trajectorial average. The weak form of the averaged equation with special trial/test spaces will be presented. An algorithm based on graph theory will be proposed in order to discretize that equation.

Chapter 5 tackles a problem closely related to real-world applications: particle-wave interaction in non-uniform magnetized plasmas. Combining the numerical methods introduced in Chapter 3 and Chapter 4, we formulate a structure-preserving solver.

A proof of the existence of global weak solution to quasilinear theory for one-dimensional electrostatic plasmas will be given in Chapter 6.

Finally, in Chapter 7, we summarize this dissertation and discuss future research plans.

Chapter 2: Preliminaries

2.1 Generalized Hamiltonian dynamics

2.1.1 Hamiltonian system

Consider a Hamiltonian dynamical system with canonical coordinates $x_i \in \mathbb{R}_x^n$ and conjugate momenta $k_i \in \mathbb{R}_k^n$. Given Hamiltonian $H(\mathbf{k}, \mathbf{x})$, its evolution is governed by Hamilton's equations,

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} = \dfrac{\partial H}{\partial \mathbf{k}}, \\ \dfrac{d\mathbf{k}}{dt} = -\dfrac{\partial H}{\partial \mathbf{x}}. \end{cases} \tag{2.1}$$

For an ensemble of independent identical Hamiltonian systems, the probability density function $\rho(\mathbf{x}, \mathbf{k}, t)$ must satisfy Liouville equation:

$$\partial_t \rho + \{\rho, H\} = 0, \tag{2.2}$$

where the Poisson bracket takes the form

$$\{f, g\} \coloneqq \sum_i \left(\frac{\partial f}{\partial x_i} \frac{\partial g}{\partial k_i} - \frac{\partial f}{\partial k_i} \frac{\partial g}{\partial x_i} \right).$$

An invariant of the Hamiltonian system refers to a function $I : \mathbb{R}_k^n \times \mathbb{R}_x^n \to \mathbb{R}$ such that

$$\{I, H\} = 0.$$

The value of an invariant remains constant along any trajectory in phase space $\mathbb{R}_k^n \times \mathbb{R}_x^n$:

$$I\left(\mathbf{k}(t), \mathbf{x}(t)\right) \equiv I\left(\mathbf{k}(0), \mathbf{x}(0)\right),$$

which is equivalent to say any trajectory lies in a certain level set of invariant I. Obviously, any Hamiltonian system has at least one invariant: the Hamiltonian per se. And there can be at most $2n - 1$ independent invariants in total, because it only

needs $2n-1$ hypersurfaces to determine a 1-dimensional curve in $2n$-dimensional phase space. A Hamiltonian system with $2n-1$ invariants is called maximally superintegrable.

2.1.2 Nambu system

Note that the dimensionality of phase space in a Hamiltonian system is always even, but in fact, that is not necessary. The concept of maximally superintegrable Hamiltonian system was generalized by Nambu (1973), who no longer makes a distinction between momentum and position. The generalized Hamilton's equation reads

$$\frac{dq_j}{dt} = \sum_{i_1 i_2 \cdots i_{n-1} \in \pi(\cdots, j-1, j+1, \cdots)} \epsilon^{i_1 i_2 \cdots i_{n-1}} \left(\frac{\partial H_1}{\partial q_{i_1}} \cdots \frac{\partial H_{n-1}}{\partial q_{i_{n-1}}} \right), \tag{2.3}$$

where $\epsilon^{i_1 i_2 \cdots i_n}$ is the Levi-Civita tensor and π represents the collection of all permutations. The generalized Poisson bracket, i.e. Nambu bracket, is defined as

$$\{\rho, H_1, H_2, \cdots, H_{n-1}\} \coloneqq \sum_{i_1 i_2 \cdots i_n \in \pi(1,2,\cdots,n)} \epsilon^{i_1 i_2 \cdots i_n} \left(\frac{\partial H_1}{\partial q_{i_2}} \cdots \frac{\partial H_{n-1}}{\partial q_{i_n}} \right) \frac{\partial \rho}{\partial q_{i_1}}, \tag{2.4}$$

The invariants $\{H_j(\mathbf{q}), j = 1, \cdots, n-1\}$ generate an incompressible flow in n-dimensional phase space:

$$\partial_t \rho(q_1, \cdots, q_n, t) = \{\rho, H_1, H_2, \cdots, H_{n-1}\}. \tag{2.5}$$

Remark 2.1. The Nambu bracket can be even further generalized when the Levi-Civita tensor is replaced by other fully antisymmetric forms (Bialynicki-Birula and Morrison, 1991).

In what follows we provide two examples of Nambu bracket:

For $n = 2$,

$$\{\rho, H_1\} = \frac{\partial H_1}{\partial q_2} \frac{\partial \rho}{\partial q_1} - \frac{\partial H_1}{\partial q_1} \frac{\partial \rho}{\partial q_2},$$

note that letting $q_1 = k$ and $q_2 = x$, we recover the Poisson bracket.

For $n = 3$,

$$\{\rho, H_1, H_2\} = \nabla \rho \cdot (\nabla H_1 \times \nabla H_2).$$

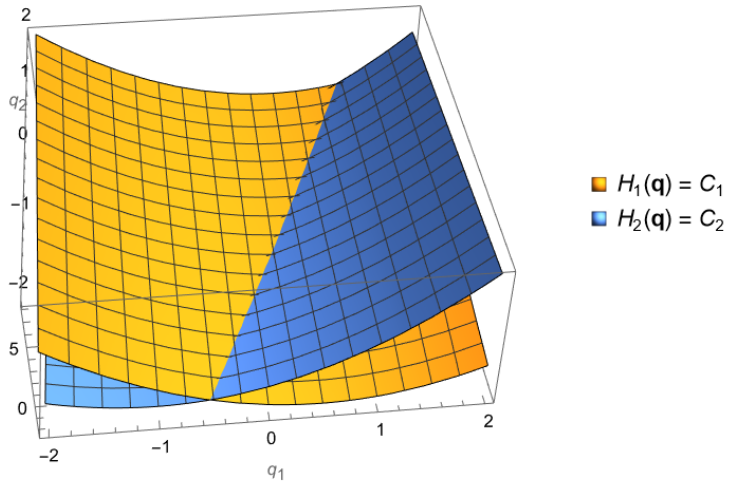


Figure 2.1: A trajectory in a 3-dimensional Nambu system is the intersection of two level sets.

Although it is still a divergence-free advection operator acting on probability density function ρ, the Nambu bracket cannot be equivalent to any Poisson bracket because the dimensionality of phase space is odd. As shown in Figure 2.1, the level sets of two invariants, $\{\mathbf{q} \in \mathbb{R}^3 : H_1(\mathbf{q}) = c_1\}$ and $\{\mathbf{q} \in \mathbb{R}^3 : H_2(\mathbf{q}) = c_2\}$, determine a trajectory in 3-dimensional phase space.

The following property of Nambu bracket will be useful in Chapter 4:

Theorem 2.1. *If a function* $g(\mathbf{q}) = \xi\left(H_1(\mathbf{q}), \cdots, H_{n-1}(\mathbf{q})\right)$*, then it must be an invariant of the Nambu system, i.e.*

$$\{g, H_1, H_2, \cdots, H_{n-1}\} = 0.$$

2.2 Waves in plasmas

The dispersion relation of waves in plasmas can be derived by analyzing the Vlasov-Maxwell(Poisson) system with small perturbations. In what follows, only the

results will be presented. For details, we refer the readers to (Stix, 1992; Thorne and Blandford, 2017; Hazeltine and Waelbroeck, 2018).

2.2.1 Waves in unmagnetized plasmas

If the thermal motion of electrons in an unmagnetized plasma is negligible, then the charge density oscillates at the electron plasma frequency:

$$\omega(\mathbf{k}) = \omega_{pe} \coloneqq \sqrt{\frac{n_e e^2}{m_e \epsilon_0}}. \tag{2.6}$$

For homogeneous warm plasmas, the linearized Vlasov-Poisson system renders the Bohm-Gross dispersion relation,

$$\omega(\mathbf{k}) = \left(1 + 3\lambda_D^2 \mathbf{k}^2\right)^{1/2} \omega_{pe},$$

where $\lambda_D \coloneqq \sqrt{\frac{\epsilon_0 k_B T_e}{n_e e^2}}$ is the Debye length, a parameter proportional to the electron thermal speed. In particular, for cold plasmas where $T_e = 0$, we recover Equation(2.6).

2.2.2 Waves in magnetized plasmas

Suppose a plasma is embedded in background magnetic field $\mathbf{B} = B_0 \mathbf{e}_z$, then electrons gyrate around magnetic field lines at the following gyrofrequency:

$$\omega_{ce} \coloneqq \frac{eB_0}{m_e}. \tag{2.7}$$

For homogeneous cold magnetized plasmas, the Vlasov-Maxwell system, after linearization, yields the following equation for $\widehat{\delta \mathbf{E}}$, the Fourier spectrum of a small electric field perturbation:

$$\left(\mathbf{N} \otimes \mathbf{N} - (\mathbf{N} \cdot \mathbf{N})\mathcal{I} + \varepsilon(\omega)\right) \widehat{\delta \mathbf{E}} = 0, \tag{2.8}$$

where $\mathbf{N} = \frac{c\mathbf{k}}{\omega}$ is the refractive index, and $\mathcal{I}$ is the identity map. Define $\mathbf{b} = \mathbf{B}/\left|\mathbf{B}\right|$, the dielectric tensor $\varepsilon(\omega)$ takes the form

$$\varepsilon_{\alpha\beta}(\omega) \equiv \varepsilon \delta_{\alpha\beta} + ige_{\alpha\beta\gamma} b_\gamma + (\eta - \varepsilon) b_\alpha b_\beta,$$

where

$$\begin{aligned}\varepsilon = \varepsilon^H &\equiv 1 - \frac{\omega_{pe}^2}{\omega^2 - \omega_{ce}^2},\\ g = g^H &\equiv -\frac{\omega_{ce}}{\omega}\frac{\omega_{pe}^2}{\omega^2 - \omega_{ce}^2},\\ \eta = \eta^H &\equiv 1 - \frac{\omega_{pe}^2}{\omega^2}.\end{aligned} \tag{2.9}$$

Equation(2.8) gives nontrivial solution only if

$$\det\mathcal{M}(\omega, \mathbf{k}) \coloneqq \det\left(\mathbf{N}\otimes\mathbf{N} - (\mathbf{N}\cdot\mathbf{N})\mathcal{I} + \varepsilon(\omega)\right) = 0. \tag{2.10}$$

The above equation renders the graph of implicit function $\omega(\mathbf{k})$, which is known as the dispersion relation.

Remark 2.2. For non-uniform plasmas, the dielectric tensor ε containing information on the medium property varies in space, hence the wave frequency will also depend on spatial coordinates, i.e. $\omega = \omega(\mathbf{k}, \mathbf{x})$.

2.2.3 Plasmon

Just like photons are quantizations of electromagnetic oscillations and phonons are quantizations of mechanical vibrations, a plasmon is a quantum of plasma oscillation. Each plasmon in state $\mathbf{k}$ carries momentum $\hbar\mathbf{k}$ and energy $\hbar\omega(\mathbf{k})$. By WKB(Wentzel–Kramers–Brillouin) approximation, the motion of plasmons is governed by Hamiltonian dynamics:

$$\begin{cases}\dfrac{d\mathbf{x}}{dt} = \nabla_k\omega(\mathbf{k}, \mathbf{x})\\ \dfrac{d\mathbf{k}}{dt} = -\nabla_x\omega(\mathbf{k}, \mathbf{x}).\end{cases}$$

Consequently, the evolution of plasmon probability density function $N(\mathbf{k}, \mathbf{x}, t)$ can be described with the following Liouville equation:

$$\partial_t N + \{N, \omega(\mathbf{k}, \mathbf{x})\} = 0. \tag{2.11}$$

2.3 Quasilinear theory from a probablistic perspective

In the last section we omitted an important fact that the wave frequency is actually complex, containing a real part and also an imaginary part. The real part accounts for advection in the phase space, while the imaginary part accounts for damping/growing. The damping effect was discovered by Landau (1946), now known as Landau damping. For some configurations such as bump-on-tail, as shown in Figure 2.2, the waves can also grow. The damping/growing of waves in a plasma implies a process of energy transfer. The study of quasilinear theory is motivated by a simple but fundamental principle: if waves are affected by the particles, there must exist a back-reaction of the waves on particles.

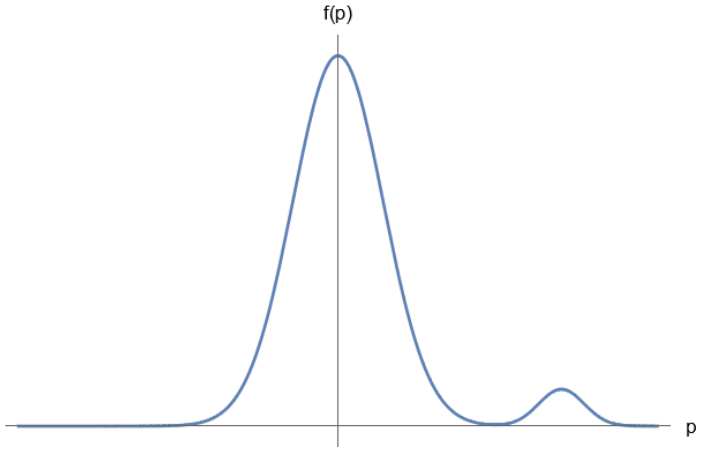
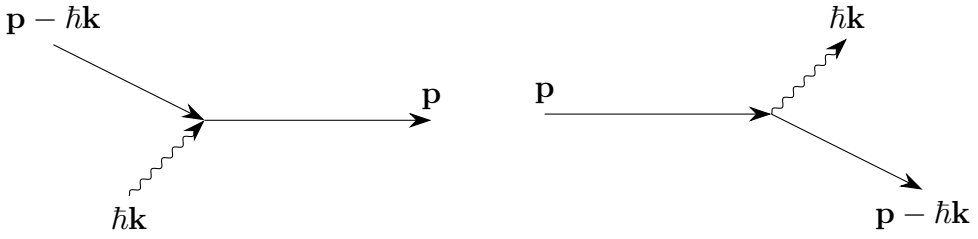


Figure 2.2: Bump-on-tail velocity distribution excites waves.

The electron-plasmon kinetic system must be derived from Vlasov-Maxwell(Poisson) system in weak turbulence limit, but it can also be interpreted from a probablistic perspective using quantum mechanical language. In this section we follow the probablistic approach as that will be sufficient for understanding the rest of this dissertation. For detailed derivation, see Vedenov et al. (1961); Shapiro and Shevchenko (1962); Stix (1992); Thorne and Blandford (2017). We will also present a brief introduction in Appendix B.

2.3.1 Master equation for unmagnetized plasmas



(a) absorption (b) emission

Figure 2.3: Feynman diagrams.

Consider a particle in a plasma, as shown in Figure 2.3, as it encounters a plasmon, by some probability it will absorb it and gain energy, or induced by the oscillating electromagnetic field, it emits another identical plasmon and loses energy. The transition probability is only relevant to the state of the particle and the plasmon. Therefore the emission/absoprtion processs, as a continuous-time Markov process, can be described with a master equation.

Remark 2.3. Spontaneous emission is usually negligible, therefore not discussed in this dissertation. For more discussion on this issue, see Thorne and Blandford (2017).

Suppose the corresponding probability of electrons in state $\mathbf{p} + \hbar\mathbf{k}$ emitting a plasmon in state $\mathbf{k}$ is $Q(\mathbf{p} + \hbar\mathbf{k}, \mathbf{k}) f(\mathbf{p} + \hbar\mathbf{k}) N(\mathbf{k})$, then for electrons in state $\mathbf{p} - \hbar\mathbf{k}$ absorbing a plasmon in state $\mathbf{k}$, the probability should be $Q(\mathbf{p}, \mathbf{k}) f(\mathbf{p} - \hbar\mathbf{k}) N(\mathbf{k})$. Consequently, the master equations read:

$$\begin{aligned} \frac{\partial f(\mathbf{p}, t)}{\partial t} &= \int Q(\mathbf{p} + \hbar\mathbf{k}, \mathbf{k}) \left[f(\mathbf{p} + \hbar\mathbf{k}, t) N(\mathbf{k}, t) - f(\mathbf{p}, t) N(\mathbf{k}, t) \right] d\mathbf{k}, \\ &\quad - \int Q(\mathbf{p}, \mathbf{k}) \left[f(\mathbf{p}, t) N(\mathbf{k}, t) - f(\mathbf{p} - \hbar\mathbf{k}, t) N(\mathbf{k}, t) \right] d\mathbf{k} \\ \frac{\partial N(\mathbf{k}, t)}{\partial t} &= \int Q(\mathbf{p}, \mathbf{k}) \left[f(\mathbf{p}, t) N(\mathbf{k}, t) - f(\mathbf{p} - \hbar\mathbf{k}, t) N(\mathbf{k}, t) \right] d\mathbf{p}. \end{aligned} \tag{2.12}$$

Observe that Equation(2.12) is in finite difference form and $|\hbar\mathbf{k}| \ll |\mathbf{p}|$, hence the

first order terms in Taylor expansion yield

$$\begin{aligned}\partial_t f =& \nabla_p \cdot \left\{ \left[\int Q(\mathbf{p},\mathbf{k}) N \left(\hbar\mathbf{k} \otimes \hbar\mathbf{k}\right) d\mathbf{k} \right] \cdot \nabla_p f \right\}, \\ \partial_t N =& \left[\int Q(\mathbf{p},\mathbf{k}) \left(\hbar\mathbf{k} \cdot \nabla_p f\right) d\mathbf{p} \right] N.\end{aligned} \tag{2.13}$$

In Equation(2.13), the particle probability density function $f(\mathbf{p},t)$ satisfies a diffusion equation with diffusion coefficients depending on plasmon probability density function $N(\mathbf{k},t)$:

$$D(\mathbf{p},t) = \int Q(\mathbf{p},\mathbf{k}) N \left(\hbar\mathbf{k} \otimes \hbar\mathbf{k}\right) d\mathbf{k}. \tag{2.14}$$

Meanwhile, the plasmon probability density function $N(\mathbf{k},t)$ satisfies a reaction equation with growth/damping rate depending on particle probability density function $f(\mathbf{p},t)$:

$$\Gamma(\mathbf{k},t) = \int Q(\mathbf{p},\mathbf{k}) \left(\hbar\mathbf{k} \cdot \nabla_p f\right) d\mathbf{p}. \tag{2.15}$$

2.3.2 Master equation for magnetized plasmas

Given magnetic field $\mathbf{B}$, any vector $\mathbf{u}$ can be decomposed in the following sense:

$$\begin{aligned} u_\parallel &\coloneqq \frac{\mathbf{B}}{|\mathbf{B}|} \cdot \mathbf{u}, \\ u_\perp &\coloneqq \left| \left(\mathcal{I} - \frac{\mathbf{B}}{|\mathbf{B}|} \otimes \frac{\mathbf{B}}{|\mathbf{B}|} \right) \cdot \mathbf{u} \right|. \end{aligned}$$

For homogeneous plasmas embedded in uniform magnetic field $\mathbf{B}$, there are master equations for electron-plasmon interaction analogously. But note that the electrons are gyrating around magnetic field lines, hence it only needs two coordinates instead of three to describe the state of an electron: momentum parallel to magnetic field $p_\parallel$ and gyro-motion quantum number n. The relation between quantum number n and classical perpendicular momentum $p_\perp$ is

$$n\hbar\omega_{ce} = \frac{p_\perp^2}{2m_e}.$$

Suppose the corresponding probability of electrons in state $\left(p_\parallel + \hbar k_\parallel, n + l\right)$ emitting a plasmon in state $\mathbf{k}$, and jumping into state $\left(p_\parallel, n\right)$ is $Q(p_\parallel + \hbar k_\parallel, n + l, l, \mathbf{k}) f(p_\parallel + \hbar k_\parallel, n + l) N(\mathbf{k})$, then for electrons in state $\left(p_\parallel - \hbar k_\parallel, n - l\right)$ absorbing a plasmon in state $\mathbf{k}$ and jumping into state $\left(p_\parallel, n\right)$, the probability should be $Q(p_\parallel, n, l, \mathbf{k}) f(p_\parallel - \hbar k_\parallel, n - l) N(\mathbf{k})$.

Consequently the master equations take the form

$$\begin{aligned}
\frac{\partial f(p_\parallel, n, t)}{\partial t} =& \sum_l \int \left\{ Q(p_\parallel + \hbar k_\parallel, n + l, l, \mathbf{k}) \left[f(p_\parallel + \hbar k_\parallel, n + l, t) N(\mathbf{k}, t) - f(p_\parallel, n, t) N(\mathbf{k}, t) \right] \right\} d\mathbf{k} \\
& - \sum_l \int Q(p_\parallel, n, l, \mathbf{k}) \left[f(p_\parallel, n, t) N(\mathbf{k}, t) - f(p_\parallel - \hbar k_\parallel, n - l, t) N(\mathbf{k}, t) \right] d\mathbf{k}, \\
\frac{\partial N(\mathbf{k}, t)}{\partial t} =& \sum_n \sum_l \int Q(p_\parallel, n, l, \mathbf{k}) \left[f(p_\parallel, n, t) N(\mathbf{k}, t) - f(p_\parallel - \hbar k_\parallel, n - l, t) N(\mathbf{k}, t) \right] dp_\parallel.
\end{aligned}$$

Since $\left|\hbar k_\parallel\right| \ll \left|p_\parallel\right|$ and $|l| \ll |n|$, the first order terms in Taylor expansion yield

$$\begin{aligned}
\partial_t f =& \sum_l \int \frac{\partial}{\partial p_\parallel} \left[\hbar k_\parallel Q(p_\parallel, n, l, \mathbf{k}) \left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l \frac{\partial f}{\partial n} \right) \right] N d\mathbf{k} \\
& + \sum_l \int \frac{\partial}{\partial n} \left[l Q(p_\parallel, n, l, \mathbf{k}) \left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l \frac{\partial f}{\partial n} \right) \right] N d\mathbf{k}, . \qquad (2.16) \\
\partial_t N =& \sum_n \sum_l \int Q(p_\parallel, n, l, \mathbf{k}) N \left[\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l \frac{\partial f}{\partial n} \right] dp_\parallel
\end{aligned}$$

Recall that $n\hbar\omega_{ce} = \frac{p_\perp^2}{2m_e}$, hence for any probability distribution g we have the following relation between quantum states and the classical momentum,

$$\sum_n \int g(p_\parallel, n) dp_\parallel = \int g(p_\parallel, \frac{p_\perp^2}{2\hbar\omega_{ce} m_e}) 2\pi p_\perp dp_\perp dp_\parallel.$$

Replace $\frac{1}{\hbar\omega_{ce}}\frac{\partial}{\partial n}$ with $\frac{m_e}{p_\perp}\frac{\partial}{\partial p_\perp}$ in Equation(2.16) to obtain

$$\begin{aligned}\partial_t f =& \sum_l \int \frac{\partial}{\partial p_\parallel}\left[\hbar k_\parallel Q(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}) N\left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l\frac{m_e\hbar\omega_{ce}}{p_\perp}\frac{\partial f}{\partial p_\perp}\right)\right] d\mathbf{k} \\ &+ \sum_l \int l\frac{m_e\hbar\omega_{ce}}{p_\perp}\frac{\partial}{\partial p_\perp}\left[Q(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}) N\left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l\frac{m_e\hbar\omega_{ce}}{p_\perp}\frac{\partial f}{\partial p_\perp}\right)\right] d\mathbf{k}, \\ \partial_t N =& \left[\sum_l \int Q(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k})\left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l\frac{m_e\hbar\omega_{ce}}{p_\perp}\frac{\partial f}{\partial p_\perp}\right) d\mathbf{p}\right] N.\end{aligned} \tag{2.17}$$

Same as in unmagnetized plasmas, electron probability density function f satisfies a diffusion equation with coefficients:

$$\begin{aligned} D_{\parallel,\parallel}(\mathbf{p}, t) &= \sum_l \int_k N(\mathbf{k}, t)\left(\hbar k_\parallel\right)^2 Q\left(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}\right), \\ D_{\parallel,\perp}(\mathbf{p}, t) &= \sum_l \int_k N(\mathbf{k}, t)\left(\hbar k_\parallel\right)\left(l\frac{m_e\hbar\omega_{ce}}{p_\perp}\right) Q\left(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}\right), \\ D_{\perp,\perp}(\mathbf{p}, t) &= \sum_l \int_k N(\mathbf{k}, t)\left(l\frac{m_e\hbar\omega_{ce}}{p_\perp}\right)^2 Q\left(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}\right). \end{aligned} \tag{2.18}$$

Meanwhile plasmon probability density function N satisfies a reaction equation with growth/damping rate:

$$\Gamma(\mathbf{k}, t) = \sum_l \int_p Q\left(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}\right)\left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l\frac{m_e\hbar\omega_{ce}}{p_\perp}\frac{\partial f}{\partial p_\perp}\right). \tag{2.19}$$

2.3.3 Singular transition rates and properties

The quantum mechanical approach is just a way of interpretation, not derivation. In the previous sections we mentioned the transition rates Q but did not give their explicit forms. The explicit forms can only be obtained by analyzing the original Vlasov-Maxwell(Poisson) system. The derivations are too long and not necessary, therefore only only the results are listed below.

Define Lorentz factor

$$\gamma(\mathbf{p}) \coloneqq \sqrt{1 + p^2/m_e^2c^2}.$$

The kinetic energy $\mathcal{E}$ of an electron with momentum $\mathbf{p}$ is

$$\mathcal{E}(\mathbf{p}) \coloneqq \begin{cases} \dfrac{p^2}{2m_e}, & \text{non-relativistic} \\ \gamma(\mathbf{p})m_e c^2 = \sqrt{m_e^2 c^4 + p^2 c^2}, & \text{relativistic} \end{cases}$$

Denote electron velocity as $\mathbf{v}(\mathbf{p}) \coloneqq \nabla_p \mathcal{E}(\mathbf{p})$, then the transition rates take the following form:

$$Q(\mathbf{p}, \mathbf{k}) = \frac{\pi e^2}{\epsilon_0} \frac{\omega}{\hbar |\mathbf{k}|^2} \delta\left(\omega - \mathbf{k} \cdot \mathbf{v}\right), \qquad \text{Vedenov et al. (1961)}$$

$$Q(p_\parallel, \frac{p_\perp^2}{2m_e \hbar \omega_{ce}}, l, \mathbf{k}) = \frac{\left(p/\gamma m_e\right)^2}{\hbar\omega} U_l(\mathbf{p}, \mathbf{k}) \delta(\omega - k_\parallel v_\parallel - l\omega_{ce}/\gamma), \quad \text{Shapiro and Shevchenko (1962)}$$

where the coefficients take the form(Breizman et al., 2019)

$$U_l(\mathbf{p}, \mathbf{k}) = 8\pi^2 e^2 \frac{\left\{ \frac{l\omega_{ce}}{k_\perp p} J_l + E_3 \cos\theta J_l + i E_2 \sin\theta J_l' \right\}^2}{(1 - E_2^2) \frac{1}{\omega} \frac{\partial}{\partial \omega}(\omega^2 \varepsilon) + 2i E_2 \frac{1}{\omega} \frac{\partial}{\partial \omega}(\omega^2 g) + E_3^2 \frac{1}{\omega} \frac{\partial}{\partial \omega}(\omega^2 \eta)}. \tag{2.20}$$

In the above formula, the dielectric tensor components ε, g, η are given in Equation(2.9). The wave polarization vector components E_j are defined as follows:

$$\begin{aligned} E_1(\mathbf{k}) &= 1, \\ E_2(\mathbf{k}) &= i \frac{g}{\varepsilon - \mathbf{N}^2}, \\ E_3(\mathbf{k}) &= -\frac{N_\parallel N_\perp}{\eta - N_\perp^2}, \end{aligned} \tag{2.21}$$

where vector $\mathbf{N} = \frac{\mathbf{k}}{\omega}$ is the refractive index. In addition, the argument of Bessel functions J_l is $k_\perp p_\perp / m_e \omega_{ce}$,

Note that both transition rates include a Dirac delta term, which suggests the resonance nature of electron-plasmon interaction: electrons and plasmons do not interact unless their momenta satisfy certain condition.

Define the resonance indicator functions as follows

$$\begin{aligned} s\left(\mathbf{p}, \mathbf{k}\right) &\coloneqq \omega(\mathbf{k}) - \mathbf{k} \cdot \mathbf{v}, && \text{(unmagnetized plasmas)} \\ s_l\left(\mathbf{p}, \mathbf{k}\right) &\coloneqq \omega(\mathbf{k}) - k_\parallel v_\parallel - l\omega_{ce}/\gamma, && \text{(magnetized plasmas)} \end{aligned}$$

then the resonance condition take the form $s(\mathbf{p},\mathbf{k})=0$, which determines a resonance manifold embedded in the joint momentum space $\mathbb{R}^3_p\times\mathbb{R}^3_k$:

$$\begin{aligned}\mathcal{S}&=\left\{(\mathbf{p},\mathbf{k})\in\mathbb{R}^3_p\times\mathbb{R}^3_k : s(\mathbf{p},\mathbf{k})=0\right\}, \quad \text{(unmagnetized plasmas)}\\ \mathcal{S}_l&=\left\{(\mathbf{p},\mathbf{k})\in\mathbb{R}^3_p\times\mathbb{R}^3_k : s_l(\mathbf{p},\mathbf{k})=0\right\}, \quad \text{(magnetized plasmas)}\end{aligned} \tag{2.22}$$

Now we can say that electrons with momentum $\mathbf{p}$ emit or absorb plasmons with wave vector $\mathbf{k}$ only when $(\mathbf{p},\mathbf{k})$ belongs to a resonance manifold.

As can be verified easily, quasilinear theory for unmagnetized plasmas conserves total mass, total momentum and total energy as follows

$$\begin{aligned}\partial_t\mathcal{M}_{tot}&=\partial_t\left((f,1)_p+(N,0)_k\right)=0,\\ \partial_t\mathcal{P}_{i,tot}&=\partial_t\left((f,p_i)_p+(N,\hbar k_i)_k\right)=0,\ i=1,2,3,\\ \partial_t\mathcal{E}_{tot}&=\partial_t((f,\mathcal{E})_p+(N,\hbar\omega)_k)=0.\end{aligned}$$

For magnetized plasmas, the second line above has to be replaced with

$$\partial_t\mathcal{P}_{\parallel,tot}=\partial_t\left((f,p_\parallel)_p+(N,\hbar k_\parallel)_k\right)=0.$$

Just like binary collision process between particles, the emission/absorption process is also entropy-dissipative:

$$\partial_t(f,\log f)_p\le 0.$$

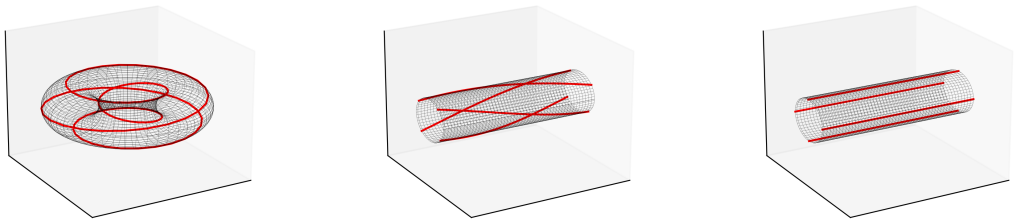
2.4 Kinetic model for electron runaway in tokamaks

2.4.1 Geometry

Tokamak, as the most promising candidate for a practical fusion reactor, is known for its complex geometry. The complexity and cost for kinetic model in realistic geometry is beyond capability. As a consequence, simplification and compromise is necessary.

As shown in Figure(2.4a), the red line represents the magnetic field line in a tokamak. The magnetic field has both poloidal and toroidal components. A simplified

configuration, the infinitely long symmetric cylinder, is shown in Figure(2.4b). By symmetric we mean the magnetic field and the plasma density do not depend on azimuthal angle ϕ, nor axial coordinate z. It can be regarded as a limit case of Figure(2.4a), when torus radius goes to infinity. In the rest of this dissertation, we will constrain our scope on a further simplified configuration, shown in Figure(2.4c), assuming axial magnetic field.



(a) Tokamak (b) Cylinder (c) Cylinder($B_\phi = 0$)

Figure 2.4: Magnetic field configurations

It is natural to use cylindrical coordinates inside a cylinder. As shown in Figure(2.5). At position (r, ϕ, z), any vector $\mathbf{u}$ can be decomposed in two ways.

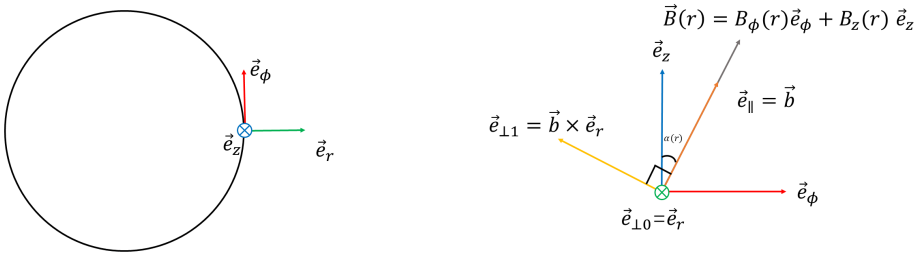
1. $\mathbf{u} = u_r\mathbf{e}_r + u_\phi\mathbf{e}_\phi + u_z\mathbf{e}_z$.

2. $\mathbf{u} = u_\parallel\mathbf{e}_\parallel + u_{\perp 0}\mathbf{e}_{\perp 0} + u_{\perp 1}\mathbf{e}_{\perp 1}$.

Wave vector $\mathbf{k}$ is usually decomposed in the first way, while electron momentum $\mathbf{p}$ is decomposed in the second way. In particular, when $B_\phi = 0$, we have $\mathbf{e}_\parallel = \mathbf{e}_z$. Consequently $u_\parallel = u_z$ and $u_\perp = \sqrt{u_r^2 + u_\phi^2}$.

2.4.2 Kinetic model

Combining Equation(2.11) and (2.17), we obtain a kinetic model for non-uniform plasma embedded in background magnetic field $\mathbf{B} = B_0(\mathbf{x})\mathbf{e}_z$. The electron



(a) Cylinder Cross Section (b) Magnetic field

Figure 2.5: Local frame

probability density function $f(\mathbf{p}, \mathbf{x}, t)$ and the plasmon probability density function $N(\mathbf{k}, \mathbf{x}, t)$ satisfy

$$\begin{aligned} \partial_t f + v_z \partial_z f &= \nabla_p \cdot (D[N] \cdot \nabla_p f)\,, \\ \partial_t N + \{N, \omega\} &= \Gamma[f] N, \end{aligned} \tag{2.23}$$

where $D[N]$ and $\Gamma[f]$ are given in Equation(2.18) and (2.19).

Electron runaway is determined by multiple factors competing with each other. According to Breizman et al. (2019), the complete kinetic model for electron runaway takes the following form:

$$\begin{aligned} \partial_t f + v_z \partial_z f &= \nabla_p \cdot (D[N] \cdot \nabla_p f) + e E_z \frac{\partial f}{\partial p_z} + \mathcal{C} f, \\ \partial_t N + \{N, \omega\} &= \Gamma[f] N, \end{aligned} \tag{2.24}$$

where E_z is the parallel component of external electric field, and operator $\mathcal{C}$ represents collision.

Chapter 3: Conservative schemes for particle-wave interaction in homogeneous magnetized plasmas

3.1 Introduction

In this chapter, we consider particle-wave interaction in homogeneous magnetized plasmas(Huang et al., 2023a)[1]. The word "homogeneous" here refers to the case where the plasma, as a medium for wave propagation, is homogeneous when inspected on a scale much greater than the characteristic wave length. Since the quasilinear theory studies the spectrum of waves(density of plasmons) and the averaged particle distribution function, it does not require a small time step to characterize the high wave frequency numerically. However, the numerical computation of the particle-wave resonance system is still challenging, due to the resonance condition described with the Dirac delta function, the complicated dispersion relation, high dimension, nonlinearity, and conservation laws consisting of integrals in two different spaces. Therefore, although the theory has been widely used in physics(Pokol et al., 2008; Liu et al., 2018; Jeong et al., 2020), there is no preceding work focusing on the numerical method for quasilinear theory in magnetized plasmas.

Despite being a paradigm approach in the analysis and discretization of other kinetic equations, the weak formulation of the quasilinear model has not gained enough attention, partly because the equation for particles was usually written in a nonlinear diffusion form, and the equation for waves was treated as independent first-order ODEs with parameters. There are infinitely many equivalent forms to the same equation because of the resonance condition. Among all the equivalent forms, some are superior to the others, the reason is as follows.

[1]Huang, K., Abdelmalik, M., Breizman, B. and Gamba, I.M., 2023. A conservative Galerkin solver for the quasilinear diffusion model in magnetized plasmas. Journal of Computational Physics, 488, p.112220. The dissertator's contribution includes proposing and implementing the scheme, analyzing the data and writing the article.

The quasilinear theory inherits the conservation laws from the original Vlasov-Maxwell system. However, generally the conservation is conditional, which means the gain and loss parts only offset each other on the resonance manifold. When the resonance manifold is broadened or approximated, conservation laws are no longer guaranteed. In this chapter, we propose a novel integro-differential form and the corresponding unconditionally conservative weak form.

It is desired that the discrete weak form will preserve the unconditional conservation property above, unfortunately, naive standard finite element discretizations turn out to fail. We located the cause of discretization errors by analyzing the weak form, and managed to construct a perfect discretization by replacing some quantities with their projection in the discrete finite element spaces.

Apart from that, for numerical integration on resonance manifold, we adopt the marching simplex algorithm(Doi and Koide, 1991; Min and Gibou, 2007), which enables us to deal with arbitrary wave modes.

This chapter is organized as follows. Section 3.2 focuses on the energy of a single electron and a single plasmon, which play the key role in calculation of transition rates. Section 3.3 introduces the unconditionally conservative weak form. The conservative semi-discrete system, as the main result of this chapter, will be presented in Section 3.4. In Section 3.5, we derive the nonlinear ODE system associated with our conservative semi-discrete form, and the relation between two interaction tensors is proved. Stability and positivity will be discussed in Section 3.6. The numerical results are presented in Section 3.7. Finally in Section 3.8, we adopt the same strategy to construct a conservative discontinuous Galerkin scheme.

3.2 Electron energy and plasmon energy

As mentioned in Chapter 2, the equations for quasilinear particle-wave interaction share the following structure,

$$\begin{aligned} \partial_t f(\mathbf{p},t) &= \nabla_p \cdot (D[N](\mathbf{p},t) \cdot \nabla_p f(\mathbf{p},t)), \\ \partial_t N(\mathbf{k},t) &= \Gamma[f](\mathbf{k},t) N(\mathbf{k},t). \end{aligned} \tag{3.1}$$

Both relations, $D : L^1(\mathbb{R}^3_k) \to (L^\infty(\mathbb{R}^3_p))^{3\times 3}$ and $\Gamma : H^1(\mathbb{R}^3_p) \to L^\infty(\mathbb{R}^3_k)$, as given in Equation(2.14), (2.15), (2.18) and (2.19), are determined by transition probabilities of the stochastic emission/absorption process. The transition probabilities per se, depend solely on pre-interaction and post-interaction kinetic variables: particle momentum $\mathbf{p}$, particle energy $\mathcal{E}(\mathbf{p})$, plasmon momentum $\hbar\mathbf{k}$ and plasmon energy $\hbar\omega(\mathbf{k})$. Hence the particle energy relation $\mathcal{E}(\mathbf{p})$ and plasmon dispersion relation $\omega(\mathbf{k})$ must be specified before numerical simulation.

The kinetic energy $\mathcal{E}$ of an electron with momentum $\mathbf{p}$ takes the following form:

$$\mathcal{E}(\mathbf{p}) \coloneqq \begin{cases} \dfrac{p^2}{2m_e}, & \text{non-relativistic} \\ \sqrt{m_e^2 c^4 + p^2 c^2}, & \text{relativistic} \end{cases}$$

As for plasmons, the wave dispersion relation $\omega(\mathbf{k})$ depends on the medium, i.e. the plasma itself, which is evolving. Since the computational cost for an accurate dispersion relation $\omega := \omega(\mathbf{k})$ can be quite high, there is, in practice, a tendency to use low-order approximations based on appropriate assumptions, for example, the cold plasma assumption. Nevertheless, even for a cold magnetized plasma, there can be multiple wave modes, i.e. multiple "species" of plasmons, each having a distinct dispersion relation $\omega(\mathbf{k})$.

Recall Equation(2.10) which renders the graph of implicit function $\omega(\mathbf{k})$ for cold magnetized plasmas:

$$\det \mathcal{M}(\omega,\mathbf{k}) \coloneqq \det\left(\mathbf{N}\otimes\mathbf{N} - (\mathbf{N}\cdot\mathbf{N})\mathcal{I} + \varepsilon(\omega)\right) = 0.$$

According to textbooks(Thorne and Blandford, 2017; Stix, 1992), the function $\omega(\mathbf{k})$ implicitly given by the above equation is multi-valued, with four branches. But we would like to know, for a specific branch, whether $\omega(\mathbf{k})$ is well-defined for any $\mathbf{k} \in \mathbb{R}^3_k$. Such issue is important for numerical implementation, but textbooks never elaborate on that, therefore we provide an answer here.

Define $k_\parallel \coloneqq \mathbf{k} \cdot \mathbf{B}/|\mathbf{B}|$, $k_\perp \coloneqq \sqrt{\mathbf{k}^2 - k_\parallel^2}$, $k \coloneqq |\mathbf{k}|$, and $\xi \coloneqq k_\parallel/k$. We find that

$$Det[M(\omega, \mathbf{k})] = 0 \Leftrightarrow F(\omega/\omega_{pe}, kc/\omega_{pe}, \xi; \omega_{ce}/\omega_{pe}) = 0.$$

Scaling ω and ω_{ce} with ω_{pe}, scaling k with ω_{pe}/c, we take the numerator of $F(\omega, k, \xi; \omega_{ce})$ and get the following algebraic equation:

$$\begin{aligned} &k^4 \left(\omega_{ce}^2 \left(\xi^2 - \omega^2\right) + \omega^2 \left(\omega^2 - 1\right)\right) - k^2\omega^2 \left(\omega_{ce}^2 \left(\xi^2 - 2\omega^2 + 1\right)\right) \\ &+\omega^2 \left(\omega^2 - 1\right) \left(-\left(\omega_{ce}^2 + 2\right)\omega^2 + \omega^4 + 1\right) = 0. \end{aligned} \tag{3.2}$$

Let $\chi = k^2$, $\varsigma = \omega^2$, we obtain a second order algebraic equation for k:

$$A(\varsigma)\chi^2 + B(\varsigma)\chi + C(\varsigma) = 0, \tag{3.3}$$

with

$$\begin{aligned} A(\varsigma) &= \varsigma^2 - (1 + \omega_{ce}^2)\varsigma + \omega_{ce}^2\xi^2, \\ B(\varsigma) &= -\varsigma \left(\omega_{ce}^2 \left(\xi^2 - 2\varsigma + 1\right) + 2\left(\varsigma - 1\right)^2\right), \\ C(\varsigma) &= \varsigma \left(\varsigma - 1\right) \left(\varsigma^2 - \left(\omega_{ce}^2 + 2\right)\varsigma + 1\right). \end{aligned} \tag{3.4}$$

Consider the discriminant of quadratic equation(3.3), for $\xi \in (0, 1)$ and $\varsigma \in (0, +\infty)$,

$$(B(\varsigma))^2 - 4A(\varsigma)C(\varsigma) = \omega_{ce}^2\varsigma(\omega_{ce}^2\varsigma(1 - \xi^2)^2 + 4\xi^2(\varsigma - 1)^2) > 0, \tag{3.5}$$

thus there are always two real roots for each $\varsigma \in (0, +\infty)$. Define these two implicit functions as

$$\begin{aligned} \chi_1(\varsigma) &\coloneqq \frac{-B + \sqrt{B^2 - 4AC}}{2A}, \\ \chi_2(\varsigma) &\coloneqq \frac{-B - \sqrt{B^2 - 4AC}}{2A}. \end{aligned} \tag{3.6}$$

Meanwhile, equation(3.2) can also be written as a fourth order algebraic equation for ς.

$$\mathcal{A}(\chi)\varsigma^4 + \mathcal{B}(\chi)\varsigma^3 + \mathcal{C}(\chi)\varsigma^2 + \mathcal{D}(\chi)\varsigma + \mathcal{E}(\chi) = 0. \tag{3.7}$$

Fundamental theorem of algebra states that equation(3.7) has four complex roots, however, what really matters is the number of positive roots, for $\omega = \sqrt{\varsigma}$ has to be a positive real number.

Proposition 3.1. *$\forall\ (k_\parallel, k_\perp) \in (0,+\infty)\times(0,+\infty)$, $\exists\ 0 < \omega_1 < \omega_2 < \omega_3 < \omega_4 < \infty$ s.t.*

$$Det[M(\omega_i, k_\parallel, k_\perp)] = 0,\ i = 1,2,3,4$$

i.e. the equation admits exactly 4 positive single-value implicit functions $\omega_i(k_\parallel, k_\perp)$, $i = 1,2,3,4$, on domain $(0,+\infty)\times(0,+\infty)$, moreover, $\omega_i(k,\xi) \coloneqq \omega_i(k_\parallel(k,\xi), k_\perp(k,\xi))$ satisfy that $\frac{\partial}{\partial k}\omega_i(k,\xi) \geq 0,\ \forall\ \xi \in (0,1)$.

Proof. Note that $(k_\parallel, k_\perp) \in (0,+\infty)\times(0,+\infty) \Leftrightarrow (k,\xi) \in (0,+\infty)\times(0,1)$, thus it suffices to show that the equation

$$A(\varsigma)\chi^2 + B(\varsigma)\chi + C(\varsigma) = 0$$

admits 4 distinct roots $0 < \varsigma_1 < \varsigma_2 < \varsigma_3 < \varsigma_4 < \infty$ for any given $\chi > 0$, and $\frac{d}{d\chi}\varsigma_i(\chi) \geq 0$.

Consider $\chi_1(\varsigma) = \frac{-B+\sqrt{B^2-4AC}}{2A}$ from (3.6) first:

Recall from (3.4) that

$$A(\varsigma) = \varsigma^2 - (1+\omega_{ce}^2)\varsigma + \omega_{ce}^2\xi^2.$$

It can be verified by substituting ς with $0, 1, \omega_{ce}^2, \omega_{ce}^2 + 1$ that

$$A(\varsigma) = (\varsigma - \varsigma_L)(\varsigma - \varsigma_R),$$

and

$$0 < \varsigma_L < \min\{1, \omega_{ce}^2\} \leq \max\{1, \omega_{ce}^2\} < \varsigma_R < \omega_{ce}^2 + 1. \tag{3.8}$$

Meanwhile, from (3.4),

$$B(\varsigma) = -\varsigma\left(\omega_{ce}^2\left(\xi^2 - 2\varsigma + 1\right) + 2\left(\varsigma - 1\right)^2\right) = -\varsigma(A(\varsigma) + (\varsigma - \omega_{ce}^2 - 2)(\varsigma - 1)),$$

thus by the inequalities in (3.8)

$$B(\varsigma_L) = -\varsigma_L(\varsigma_L - \omega_{ce}^2 - 2)(\varsigma_L - 1) < 0,$$
$$B(\varsigma_R) = -\varsigma_R(\varsigma_R - \omega_{ce}^2 - 2)(\varsigma_R - 1) > 0.$$

Now we claim that although $A(\varsigma)$ has two positive roots ς_L and ς_R, $\chi_1(\varsigma) = \frac{-B+\sqrt{B^2-4AC}}{2A}$ is only singular at $\varsigma = \varsigma_L$, while continuous at $\varsigma = \varsigma_R$. The reason is as follows.

Notice that Equation(3.6) can also be written as

$$\chi_1(\varsigma) = \frac{-B + \sqrt{B^2 - 4AC}}{2A} = \frac{2C}{-B - \sqrt{B^2 - 4AC}}. \tag{3.9}$$

Since $B(\varsigma_L) < 0$, use the first expression at $\varsigma = \varsigma_L$, then

$$\lim_{\varsigma \to \varsigma_L} \chi_1(\varsigma) = \frac{|B(\varsigma_L)|}{(\varsigma_L - \varsigma_R)} \cdot \frac{1}{\varsigma - \varsigma_L},$$

therefore

$$\lim_{\varsigma \to \varsigma_L^-} \chi_1(\varsigma) = +\infty, \tag{3.10}$$

and

$$\lim_{\varsigma \to \varsigma_L^+} \chi_1(\varsigma) = -\infty. \tag{3.11}$$

At $\varsigma = \varsigma_R$, use Equation(3.9) to obtain that

$$\lim_{\varsigma \to \varsigma_R} \chi_1(\varsigma) = \frac{C(\varsigma_R)}{-B(\varsigma_R)}.$$

Additionally, it can be verified that

$$\lim_{\varsigma \to 0^+} \chi_1(\varsigma) = 0, \tag{3.12}$$

and

$$\lim_{\varsigma \to +\infty} \chi_1(\varsigma) = +\infty. \tag{3.13}$$

Conbining (3.12) and (3.10), since $\chi_1(\varsigma)$ is continuous on $(0, \varsigma_L)$, by intermediate value theorem, for any given $\kappa > 0$, there must be a $\varsigma_{11}(\kappa) \in (0, \varsigma_L)$ such that $\chi_1(\varsigma_{11}(\kappa)) = \kappa$. Conbining (3.13) and (3.11), since $\chi_1(\varsigma)$ is continuous on $(\varsigma_L, +\infty)$, by intermediate value theorem, for any given $\kappa > 0$, there must be a $\varsigma_{12}(\kappa) \in (\varsigma_L, +\infty)$ such that $\chi_1(\varsigma_{12}(\kappa)) = \kappa$.

Similar conclusion can be proved for $\chi_2(\varsigma)$: For any given $\kappa > 0$, there must be a $\varsigma_{21}(\kappa) \in (0, \varsigma_R)$ such that $\chi_2(\varsigma_{21}(\kappa)) = \kappa$. Also, for any given $\kappa > 0$, there must be a $\varsigma_{22}(\kappa) \in (\varsigma_R, +\infty)$ such that $\chi_2(\varsigma_{22}(\kappa)) = \kappa$.

Now we have 4 solutions $\varsigma_{11}, \varsigma_{12}, \varsigma_{21}, \varsigma_{22}$, sort them in an increasing order, and we refer them as $0 < \varsigma_1 < \varsigma_2 < \varsigma_3 < \varsigma_4 < \infty$. For example, as shown in Figure(3.1a), the blue curves represent $\varsigma_{11}, \varsigma_{12}$, and the yellow curves represent $\varsigma_{21}, \varsigma_{22}$.

Next, we prove that $\varsigma_i(\chi)$'s are distinct and monotonic increasing in χ by contradiction:

Suppose $\exists\ \kappa > 0$ such that $\varsigma_i(\kappa) = \varsigma_j(\kappa) = w > 0$, then $\chi_1(w) = \chi_2(w)$, i.e.

$$B(w)^2 - 4A(w)C(w) = 0,$$

contradictory to (3.5), thus four solutions ς_i are distinct.

Recall that $A(\varsigma)\chi^2 + B(\varsigma)\chi + C(\varsigma) = 0$ can also be rewritten as

$$\mathcal{A}(\chi)\varsigma^4 + \mathcal{B}(\chi)\varsigma^3 + \mathcal{C}(\chi)\varsigma^2 + \mathcal{D}(\chi)\varsigma + \mathcal{E}(\chi) = 0,$$

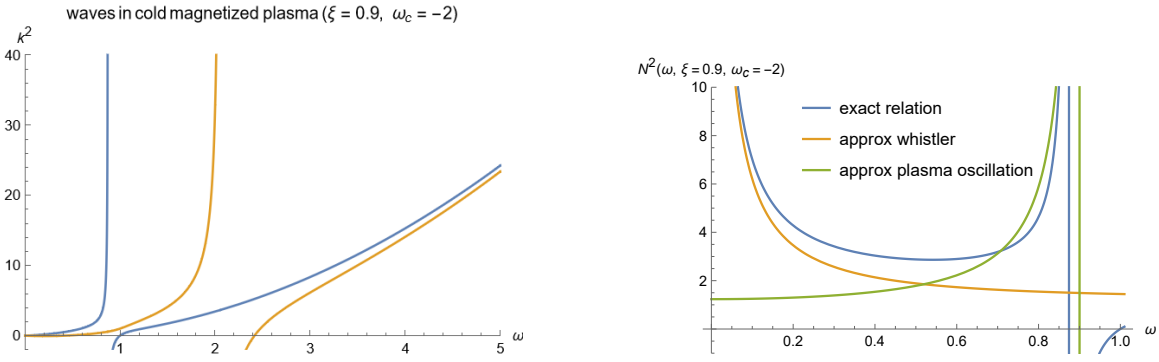
thus there can only be exactly 4 solutions ς_i.

Suppose $\exists\ w \in (0, \varsigma_L)$ such that $\chi_1(w) = \kappa > 0$ and $\frac{d}{d\varsigma}\chi_1(w) < 0$, then through intermediate value theorem for continuous function $\chi_1(\varsigma)$ on $(0, \varsigma_L)$, there will be at least 3 distinct solutions in $(0, \varsigma_L)$ for equation $\mathcal{A}(\kappa)\varsigma^4 + \mathcal{B}(\kappa)\varsigma^3 + \mathcal{C}(\kappa)\varsigma^2 + \mathcal{D}(\kappa)\varsigma +$

$\mathcal{E}(\kappa) = 0$, in addition with 3 solutions on other branches of the graph, there will be 6 solutions in total, contradictory to the fundamental theorem of algebra.

As a consequence, for any $w \in (0, \varsigma_L)$ such that $\chi_1(w) = \kappa > 0$, $\frac{d}{d\varsigma}\chi_1(w) \geq 0$, equivalently, $\frac{d}{d\chi}\varsigma_{11}(\chi) \geq 0$. Similar conclusions can be proved for $\varsigma_{12}(\chi)$, $\varsigma_{21}(\chi)$ and $\varsigma_{22}(\chi)$. □

We plot $\chi_1(\omega^2)$ and $\chi_2(\omega^2)$ as functions of ω given $\xi = 0.9, \omega_{ce} = -2$ in Figure(3.1a). The figure shows exactly what we have proved, given a point (k, ξ) (or equivalently $(k_\parallel, k_\perp)$), there are 4 values of ω which satisfy Equation(2.10), i.e. 4 formulas for $\omega(k_\parallel, k_\perp)$.



(a) waves (b) N^2

Figure 3.1: wave modes.

Remark 3.1. In the paper of Aleynikov and Breizman (2015), the authors gave two approximated dispersion relations as Equation (8) and (9). They are nearly accurate only in two limit conditions $k \to 0$ and $k \to \infty$. To illustrate, we plot $N^2 = \frac{k^2}{\omega^2}$ as a function of ω, see Figure(3.1b).(But for $|\omega_{ce}| < \omega_{pe}$, the plasma oscillation is not asymptotic to any branch.)

Remark 3.2. In our numerical experiment, we use the dispersion relation of whistler waves in a cold magnetized plasma. Nevertheless, our numerical method is compatible with any dispersion relation, and can be used to simulate multiple wave modes at the same time.

3.3 Equations reformulated

In principle, with $\omega(\mathbf{k})$ given, the formulas (2.18) and (2.19) combined with Equation(3.1) are sufficient to perform a trivial numerical simulation: treat Equation(3.1) as a normal diffusion equation and a normal reaction equation with time-varying coefficients. The challenging part, however, is to preserve conservation, especially when there are integrals containing the Dirac delta function. Numerical integrals are always performed by quadrature rules, however, the quadrature points usually do not reside exactly on the resonance manifold. In this chapter, we propose an unconditionally conservative approach by employing a novel equivalent form of the original equation. By "unconditional" we mean that the scheme is conservative no matter how the resonance manifold is discretized or broadened.

3.3.1 Conditionally conservative weak form

Recall the kinetic system(2.17) for magnetized plasmas:

$$\begin{aligned}
\partial_t f =& \sum_l \int \frac{\partial}{\partial p_\parallel} \left[\hbar k_\parallel Q(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}) N \left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l \frac{m_e\hbar\omega_{ce}}{p_\perp} \frac{\partial f}{\partial p_\perp} \right) \right] d\mathbf{k} \\
&+ \sum_l \int l \frac{m_e\hbar\omega_{ce}}{p_\perp} \frac{\partial}{\partial p_\perp} \left[Q(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}) N \left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l \frac{m_e\hbar\omega_{ce}}{p_\perp} \frac{\partial f}{\partial p_\perp} \right) \right] d\mathbf{k}, \\
\partial_t N =& \left[\sum_l \int Q(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}) \left(\hbar k_\parallel \frac{\partial f}{\partial p_\parallel} + l \frac{m_e\hbar\omega_{ce}}{p_\perp} \frac{\partial f}{\partial p_\perp} \right) d\mathbf{p} \right] N.
\end{aligned}$$

To rephrase it into a new form, let us first introduce two important concepts, the emission/absorption kernel and the directional differential operator, along with some necessary notations.

Emission/absorption kernel and directional differential operator

Analogous to the definition of collisional kernels in Boltzmann equations and Fokker-Planck-Landau equations, we define the *emission/absorption kernel* which characterizes the probability for a particle with momentum $\mathbf{p}$ to absorb or emit a

plasmon with wave vector $\mathbf{k}$,

$$\mathcal{G}_l(\mathbf{p},\mathbf{k}) \coloneqq \frac{\hbar\omega}{\left(p/\gamma m_e\right)^2} Q\left(p_\parallel, \frac{p_\perp^2}{2m_e\hbar\omega_{ce}}, l, \mathbf{k}\right) = U_l(\mathbf{p},\mathbf{k})\delta\left(\omega - k_\parallel v_\parallel - l\omega_{ce}/\gamma\right). \quad (3.14)$$

As we have mentioned above, interaction happens only if the resonance condition is satisfied, so the emission/absorption kernel contains a Dirac delta function. The coefficients $U_l(\mathbf{p},\mathbf{k})$ are given in Equation(2.20). They take finite non-negative values for any coordinates $(\mathbf{p},\mathbf{k}) \in \mathbb{R}_p^3 \times \mathbb{R}_k^3$.

Interaction with a plasmon results in diffusion of particle *pdf* $f(\mathbf{p},t)$ along a particular direction $\beta(\mathbf{p},\mathbf{k})$, thus we define the directional differential operator

$$\mathcal{R}_l g \coloneqq \frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel} + l\frac{\omega_{ce}/\gamma}{\omega}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}, \quad (3.15)$$

Further, define the L^2 inner product in particle momentum space $(u(\mathbf{p}), v(\mathbf{p}))_p \coloneqq \int_{\mathbb{R}_p^3} uvd^3p$, and the L^2 inner product in wave spectral space $(U(\mathbf{k}), V(\mathbf{k}))_k = \int_{\mathbb{R}_k^3} UVd^3k$.

Denote the adjoint operator of $\mathcal{R}_l$ by $\mathcal{R}_l^*$, then by definition, we have

$$(\mathcal{R}_l^* u, v)_p = (u, \mathcal{R}_l v)_p.$$

Bilinear integro-differential operators

Now all the ingredients are prepared, we claim that the diffusion term and reaction term can be rewritten as bilinear integro-differential operators.

Theorem 3.2. *The particle-wave interaction system in Equation(2.17) is equivalent to*

$$\begin{aligned} \partial_t f = \mathbb{B}(N,f) &= -\sum_l \int_{\mathbb{R}_k^3} \mathcal{R}_l^*\left(\mathcal{G}_l N\hbar\omega\mathcal{R}_l f\right) d\mathbf{k}, \\ \partial_t N = \mathbb{H}(f,N) &= \sum_l \int_{\mathbb{R}_p^3} \mathcal{G}_l N\left(\mathcal{R}_l f\right)\left(\mathcal{R}_l \mathcal{E}\right) d\mathbf{p}. \end{aligned} \quad (3.16)$$

Remark 3.3. Both the particle diffusion operator $\mathbb{B}$ and the wave reaction operator $\mathbb{H}$ mix particle momentum $\mathbf{p}$ and plasmon wave vector $\mathbf{k}$ through the absorption/emission kernel $\mathcal{R}_l(\mathbf{p}, \mathbf{k})$.

Remark 3.4. One might have noticed that $\mathcal{R}_l\mathcal{E}(\mathbf{p}) = v(\mathbf{p})$. The reason we write $\mathcal{R}_l\mathcal{E}(\mathbf{p})$ rather than v is to induce our conservative semi-discrete form and to save preprocessing time. The details will be addressed later.

Conditionally conservative weak form

Now test Equation(3.16) with $\varphi(\mathbf{p})$ and $\eta(\mathbf{k})$ to obtain its weak form:

$$\begin{aligned}
&\int_{\mathbb{R}^3_p} \varphi \partial_t f d\mathbf{p} + \int_{\mathbb{R}^3_k} \eta \partial_t N d\mathbf{k} \\
=&\int_{\mathbb{R}^3_k} d\mathbf{k} \int_{\mathbb{R}^3_p} d\mathbf{p} \left[\eta \left(\mathcal{R}_l\mathcal{E}\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l\right] - \int_{\mathbb{R}^3_k} d\mathbf{p} \int_{\mathbb{R}^3_p} d\mathbf{k} \left[\hbar\omega \left(\mathcal{R}_l\varphi\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l\right].
\end{aligned}$$

Note that the order of integration here is different for the terms on the right-hand side. In what follows, assume that

$$\int_{\mathbb{R}^3_k \times \mathbb{R}^3_p} d\mathbf{k} d\mathbf{p} |\hbar\omega \left(\mathcal{R}_l\varphi\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l|,$$

and

$$\int_{\mathbb{R}^3_k \times \mathbb{R}^3_p} d\mathbf{k} d\mathbf{p} |\eta \left(\mathcal{R}_l\mathcal{E}\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l|,$$

are finite, therefore by Fubini's theorem, the order of integration does not matter:

$$\begin{aligned}
\int_{\mathbb{R}^3_p} \left(\int_{\mathbb{R}^3_k} \hbar\omega \left(\mathcal{R}_l\varphi\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l d\mathbf{k} \right) d\mathbf{p} &= \int_{\mathbb{R}^3_k} \left(\int_{\mathbb{R}^3_p} \hbar\omega \left(\mathcal{R}_l\varphi\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l d\mathbf{p} \right) d\mathbf{k} \\
&= \iint_{\mathbb{R}^3_p \times \mathbb{R}^3_k} d\mathbf{k} d\mathbf{p} \left[\hbar\omega \left(\mathcal{R}_l\varphi\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l\right], \\
\int_{\mathbb{R}^3_k} \left(\int_{\mathbb{R}^3_p} \eta \left(\mathcal{R}_l\mathcal{E}\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l d\mathbf{p} \right) d\mathbf{k} &= \int_{\mathbb{R}^3_p} \left(\int_{\mathbb{R}^3_k} \eta \left(\mathcal{R}_l\mathcal{E}\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l d\mathbf{k} \right) d\mathbf{p} \\
&= \iint_{\mathbb{R}^3_p \times \mathbb{R}^3_k} d\mathbf{k} d\mathbf{p} \left[\eta \left(\mathcal{R}_l\mathcal{E}\right) \left(\mathcal{R}_l f\right) N\mathcal{G}_l\right].
\end{aligned}$$

It follows that

$$\int_p \varphi \partial_t f + \int_k \eta \partial_t N = \sum_l \iint_{pk} \mathcal{G}_l N \mathcal{R}_l f \left(\eta \mathcal{R}_l \mathcal{E} - \hbar\omega \mathcal{R}_l \varphi \right). \tag{3.17}$$

Then the conservation laws can be easily verified as we substitute the test functions with conserved quantities. For mass conservation, we have

$$\begin{aligned}
\partial_t \mathcal{M}_{tot} &= \partial_t \left((f,1)_p + (N,0)_k \right) \\
&= (\partial_t f, 1)_p + (\partial_t N, 0)_k \\
&= \sum_l \iint_{pk} \mathcal{G}_l N \mathcal{R}_l f \left(0 \cdot \mathcal{R}_l \mathcal{E} - \hbar\omega \cdot \mathcal{R}_l 1 \right) \\
&= 0.
\end{aligned}$$

For momentum conservation along magnetic field line, we have

$$\begin{aligned}
\partial_t \mathcal{P}_{\parallel,tot} &= \left(\partial_t f, p_\parallel \right)_p + \left(\partial_t N, \hbar k_\parallel \right)_k \\
&= \sum_l \iint_{pk} \mathcal{G}_l N \mathcal{R}_l f \left(\hbar k_\parallel \cdot \mathcal{R}_l \mathcal{E} - \hbar\omega \cdot \mathcal{R}_l p_\parallel \right) \\
&= 0.
\end{aligned}$$

And analogously for energy conservation,

$$\begin{aligned}
\partial_t \mathcal{E}_{tot} &= (\partial_t f, \mathcal{E})_p + (\partial_t N, \hbar\omega)_k \\
&= \sum_l \iint_{pk} \mathcal{G}_l N \mathcal{R}_l f \left(\hbar\omega \cdot \mathcal{R}_l \mathcal{E} - \hbar\omega \cdot \mathcal{R}_l \mathcal{E} \right) \\
&= 0.
\end{aligned}$$

Nevertheless, if one observe in detail, the total mass and energy are conserved unconditionally, because the identities $0 \cdot \mathcal{R}_l \mathcal{E} - \hbar\omega \cdot \mathcal{R}_l 1 = 0$ and $\hbar\omega \cdot \mathcal{R}_l \mathcal{E} - \hbar\omega \cdot \mathcal{R}_l \mathcal{E}$ are true for any point $(\mathbf{p}, \mathbf{k})$ in the joint momentum space $\mathbb{R}^3_p \times \mathbb{R}^3_k$. While the total momentum is conserved conditionally, because

$$\hbar k_\parallel \cdot \mathcal{R}_l \mathcal{E} - \hbar\omega \cdot \mathcal{R}_l p_\parallel = \hbar k_\parallel v \left(\frac{k_\parallel v_\parallel}{\omega} + l \frac{\omega_{ce}/\gamma}{\omega} - 1 \right),$$

while $\frac{k_\parallel v_\parallel}{\omega} + l\frac{\omega_{ce}/\gamma}{\omega} - 1 = 0$ is not true unless the states $(\mathbf{p}, \mathbf{k})$ reside on the resonance manifold

$$\mathcal{S}_l = \left\{ (\mathbf{p}, \mathbf{k}) \in \mathbb{R}^3_p \times \mathbb{R}^3_k : s_l(\mathbf{p}, \mathbf{k}) \coloneqq \omega(\mathbf{k}) - k_\parallel v_\parallel - l\omega_{ce}/\gamma = 0 \right\}.$$

Therefore we refer to Equation(3.17) as the conditionally conservative weak form.

3.3.2 Numerical integration on the resonance manifold

From the analytical point of view, the distinction between conditional conservation and unconditional conservation does not matter, since the emission/absorption kernel

$$\mathcal{G}_l(\mathbf{p},\mathbf{k}) = U_l(\mathbf{p},\mathbf{k})\delta\left(\omega - k_\parallel v_\parallel - l\omega_{ce}/\gamma\right)$$

is always zero outside the resonance manifold. However, for the purpose of either modeling or numerical implementation, the absorption/emission kernel is usually replaced with its approximation $\mathcal{G}_l^{\varepsilon}(\mathbf{p},\mathbf{k})$, which implies distortion of the resonance manifold.

In what follows we present two examples for such approximation that enables numerical integration on resonance manifold.

Approximation to the identity

In this approach, the emission/absorption kernel is approximated with

$$\mathcal{G}_l^{ati}(\mathbf{p},\mathbf{k};\varepsilon) = U_l(\mathbf{p},\mathbf{k})\frac{1}{\varepsilon}\psi\left(\frac{s_l(\mathbf{p},\mathbf{k})}{\varepsilon}\right), \tag{3.18}$$

where the compactly supported and positive function $\psi(z)$ has unit mass, i.e.

$$\int_{\mathbb{R}} \psi(z)dz = 1.$$

Recall the definition of resonance manifold in Equation(2.22), it is a hypersurface implicitly determined by the resonance condition. The above approximation is equivalent to a broadening of the resonance manifold, as the approximated hypersurface has finite "widt" proportional to ε.

Now the singular Dirac delta has been "mollified", the integrals in the weak form (3.17) can be obtained with Gaussian quadrature over the joint momentum space $\mathbb{R}^3_p \times \mathbb{R}^3_k$.

Marching cube/simplex algorithm

As opposed to the method introduced above, which aims at modifying the Dirac delta, in what follows we pursue an alternative resonance condition.

By the coarea formula from geometric measure theory, the following identity holds:

$$\int_{\mathbb{R}^n} g(\mathbf{x})\,\delta(s(\mathbf{x}))\,d\mathbf{x} = \int_{\mathcal{S}} \frac{g(\mathbf{x})}{|\nabla s|} d\sigma(\mathbf{x}),$$

where $\mathcal{S} \coloneqq \{\mathbf{x} \in \mathbb{R}^n : s(\mathbf{x}) = 0\}$ is an $(n-1)$-dimensional hypersurface and σ represents its measure. This identity, however, cannot be applied to our problem straightforwardly, since $\mathcal{S}$ is not flat in our problem.

Note that the level set of a linear function must be a $(n-1)$-dimensional hyperplane. Therefore, the level set of a piecewise linear function must be a disjoint union of $(n-1)$-dimensional simplexes. That observation inspires the famous marching cube/simplex algorithm in computational graphics(Doi and Koide, 1991), to discretize a surface.

Suppose that inside an n-dimensional simplex V, the linear interpolation of a function $s(\mathbf{x})$ is $Ls(\mathbf{x}) = a_0 + \sum_{i=1}^n a_i x_i$, then by Min and Gibou (2007),

$$\int_V g(\mathbf{x})\,\delta(Ls(\mathbf{x}))\,d\mathbf{x} = \int_{\mathcal{S}} \frac{g(\mathbf{x})}{|\nabla Ls|} d\sigma(\mathbf{x}) = \int_{\mathcal{S}} \frac{g(\mathbf{x})}{\sqrt{\sum_{i=1}^n a_i^2}} d\sigma(\mathbf{x}),$$

where $\mathcal{S}$ is a $(n-1)$-dimensional polytope, which can be decomposed into several $(n-1)$-dimensional simplexes. Such integral can easily be calculated with Gaussian quadrature rule.

Now let us go back to our specific problem. The marching simplex algorithm implies the following approximation of the emission/absorption kernel:

$$\mathcal{G}_l^{\mathrm{msa}}(\mathbf{p}, \mathbf{k}; \varepsilon) = U_l(\mathbf{p}, \mathbf{k})\delta(L_\varepsilon s_l(\mathbf{p}, \mathbf{k})), \tag{3.19}$$

where $L_\varepsilon s_l$ represents the piecewise linear interpolation of s_l, and ε indicates the mesh size.

Remark 3.5. The piecewise linear interpolation L_ε is not necessarily for every dimension. For instance, it can act only on $\mathbb{R}^3_k$, taking the form

$$L_\varepsilon s_l\left(\mathbf{p},\mathbf{k}\right)=a_0\left(\mathbf{p}\right)+\sum_{i=1}^{3}a_i\left(\mathbf{p}\right)k_i$$

inside each simplex mesh. In that case, the quadrature rule is in hybrid-type:

$$\begin{aligned}\int_{pk}g\left(\mathbf{p},\mathbf{k}\right)\delta\left(L_\varepsilon s_l\left(\mathbf{p},\mathbf{k}\right)\right)&=\int_{\mathbb{R}^3_p}\left(\int_{\mathbb{R}^3_k}g\left(\mathbf{p},\mathbf{k}\right)\delta\left(L_\varepsilon s_l\left(\mathbf{p},\mathbf{k}\right)\right)d\mathbf{k}\right)d\mathbf{p}\\&\approx\int_{\mathbb{R}^3_p}^{\text{Gaussian}}\left(\int_{\mathbb{S}}^{\text{MSA}}g\left(\mathbf{p},\mathbf{k}\right)\delta\left(L_\varepsilon s_l\left(\mathbf{p},\mathbf{k}\right)\right)d\sigma(\mathbf{k})\right)d\mathbf{p},\end{aligned}$$

where "MSA" represents marching simplex algorithm.

3.3.3 Unconditionally conservative weak form

Once the emission/absorption kernel is replaced, even the continuous system(3.17) no longer conserves momentum, needless to say about any discrete form.

However, observe that, due to resonance, there are infinitely many equivalent forms for the same equation(3.17), for example,

$$\delta\left(s_l(\mathbf{p},\mathbf{k})\right)\mathcal{R}_l g=\delta\left(\omega-k_\parallel v_\parallel-l\omega_{ce}/\gamma\right)\left(\frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel}+l\frac{\omega_{ce}/\gamma}{\omega}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}\right)\tag{3.20}$$

is always equal to

$$\delta\left(\omega-k_\parallel v_\parallel-l\omega_{ce}/\gamma\right)\left(\left(\frac{k_\parallel v_\parallel+l\omega_{ce}/\gamma}{\omega}\right)^{\alpha}\frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel}+l\frac{\omega_{ce}/\gamma}{\omega}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}\right)$$

for any constant $\alpha>0$.

We claim that among all the equivalent forms, there exists a special form which unconditionally preserves momentum. And it can be derived as follows.

Note that

$$l=\frac{\gamma}{\omega_{ce}}\left(\omega-k_\parallel v_\parallel\right)$$

on the resonance manifold $\mathcal{S}_l$. Substitute the above equation into the formula(3.20), it follows that

$$\begin{aligned}\delta\left(s_l(\mathbf{p},\mathbf{k})\right)\mathcal{R}_l g &= \delta\left(\omega - k_\parallel v_\parallel - l\omega_{ce}/\gamma\right)\left(\frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel} + \frac{\gamma}{\omega_{ce}}\left(\omega - k_\parallel v_\parallel\right)\frac{\omega_{ce}/\gamma}{\omega}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}\right)\\ &= \delta\left(\omega - k_\parallel v_\parallel - l\omega_{ce}/\gamma\right)\left(\frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel} + \left(1 - \frac{k_\parallel v_\parallel}{\omega}\right)\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}\right).\end{aligned}$$

Now we define a new directional differential operator

$$\mathcal{L}g \coloneqq \frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel} + \left(1 - \frac{k_\parallel v_\parallel}{\omega}\right)\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp},$$

as opposed to $\mathcal{R}_l$ in Equation(3.15), the new operator $\mathcal{L}$ does not depend on l. In addition, define a new emission/absorption kernel as follows:

$$\mathcal{B}(\mathbf{p},\mathbf{k}) \coloneqq \sum_l \mathcal{G}_l(\mathbf{p},\mathbf{k}) = \sum_l U_l\left(\mathbf{p},\mathbf{k}\right)\delta\left(\omega - k_\parallel v_\parallel - l\omega_{ce}/\gamma\right).$$

Note that $\mathcal{B}$ is also independent of l.

It can be verified that the weak form(3.17) is equivalent to

$$\int_p \varphi\partial_t f + \int_k \eta\partial_t N = \iint_{pk} \mathcal{B}N\mathcal{L}f\left(\eta\mathcal{L}\mathcal{E} - \hbar\omega\mathcal{L}\varphi\right). \tag{3.21}$$

In the following theorem, we prove the superiority of the proposed form, i.e. the unconditional conservation property of Equation(3.21).

Theorem 3.3 (unconditional conservation)**.** *If $f(\mathbf{p},t)$ and $N(\mathbf{k},t)$ solve Equation(3.21) with emission/absorption kernel being replaced by $\mathcal{B}_\varepsilon$, then for any $\mathcal{B}_\varepsilon$ we have the following conservation laws*

$$\begin{aligned}\partial_t\mathcal{M}_{tot} &= \partial_t\left(\left(f,1\right)_p + \left(N,0\right)_k\right) = 0,\\ \partial_t\mathcal{P}_{\parallel,tot} &= \partial_t\left(\left(f,p_\parallel\right)_p + \left(N,\hbar k_\parallel\right)_k\right) = 0,\\ \partial_t\mathcal{E}_{tot} &= \partial_t\left(\left(f,\mathcal{E}\right)_p + \left(N,\hbar\omega\right)_k\right) = 0.\end{aligned}$$

Proof. It is sufficient to prove the second row, i.e. unconditional momentum conservation, as the other two can be easily verified. Indeed, the following identity

$$\hbar k_\parallel \mathcal{L}\mathcal{E} - \hbar\omega \mathcal{L} p_\parallel = \hbar k_\parallel v - \hbar\omega \frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel} = 0$$

is true for any $(\mathbf{p}, \mathbf{k}) \in \mathbb{R}^3_p \times \mathbb{R}^3_k$. Hence for any approximated emission/absorption kernel $\mathcal{B}_\varepsilon$ we have

$$\int_p p_\parallel \partial_t f + \int_k \hbar k_\parallel \partial_t N = \iint_{pk} \mathcal{B}_\varepsilon N \mathcal{L} f \left(\hbar k_\parallel \mathcal{L}\mathcal{E} - \hbar\omega \mathcal{L} p_\parallel\right) = 0.$$

□

Remark 3.6. If one considers Equation(2.13), the quasilinear theory for non-relativistic unmagnetized plasmas, it is not hard to find out that its weak form

$$(\partial_t f, \varphi)_p + (\partial_t N, \eta) = \iint_{pk} Q(\mathbf{p}, \mathbf{k}) \frac{\hbar k^2}{\omega} N \left(\widehat{\mathbf{k}} \cdot \nabla_p f\right) \left(\eta \widehat{\mathbf{k}} \cdot \nabla_p \mathcal{E} - \hbar\omega \widehat{\mathbf{k}} \cdot \nabla_p \varphi\right) \quad (3.22)$$

also conserves total momentum in each direction conditionally.

One might wonder whether it is possible to derive an unconditionally conservative form, analogous to that in the magnetized plasma case. It turns out that it does exist, but only for equations with cylindrical symmetry. In fact, for $f(p_\parallel, p_\perp, \alpha)$ independent of α, if the emission/absorption kernel is defined as

$$\mathcal{B}^{\text{unmag}}(\mathbf{p}, \mathbf{k}) \coloneqq \frac{1}{2\pi} \int_0^{2\pi} \frac{\pi e^2}{\epsilon_0} \frac{\omega^2}{k^2 v^2} \delta(\omega - k_\parallel v_\parallel - k_\perp v_\perp \cos\alpha) d\alpha,$$

then Equation(3.22) can be rewritten in the following form

$$\int_p \varphi \partial_t f + \int_k \eta \partial_t N = \iint_{pk} \mathcal{B}^{\text{unmag}} N \mathcal{L} f \left(\eta \mathcal{L}\mathcal{E} - \hbar\omega \mathcal{L}\varphi\right),$$

which is exactly the same as Equation(3.21) for magnetized plasmas. Moreover, they even share the same directional differential operator:

$$\mathcal{L} g \coloneqq \frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel} + \left(1 - \frac{k_\parallel v_\parallel}{\omega}\right)\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp},$$

although the $\mathbf{v}(\mathbf{p})$ here is defined as

$$\mathbf{v}(\mathbf{p}) \coloneqq \nabla_p \left(\frac{p^2}{2m_e} \right) = \frac{\mathbf{p}}{m_e},$$

instead of

$$\mathbf{v}(\mathbf{p}) \coloneqq \nabla_p \left(\sqrt{m_e^2 c^4 + p^2 c^2} \right) = \frac{\mathbf{p}}{\gamma(\mathbf{p}) m_e}.$$

Remark 3.7. Since the conservation laws solely depend on the fact that $\eta \mathcal{L}\mathcal{E} - \hbar\omega \mathcal{L}\varphi = 0$ for any $(\mathbf{p}, \mathbf{k}) \in \mathbb{R}^3_p \times \mathbb{R}^3_k$, the unconditional conservative form and the scheme we are going to propose can be generalized for time-dependent dispersion relation $\omega(\mathbf{k}; t)$ with no extra effort. An important example is the self-consistent dispersion relation $\omega = \omega(\mathbf{k}; f(t))$. The only obstacle is the extra computational cost of updating the interaction tensors in each step. As will be shown in Section 3.5, that calculation can be expensive.

To simplify the notation, we can define trilinear forms B and H as follows:

$$B(f, N, \varphi) \coloneqq \iint_{\mathbb{R}^3_p \times \mathbb{R}^3_k} d\mathbf{k} d\mathbf{p} \left\{ \hbar\omega \mathcal{L}\varphi \mathcal{L} f N \mathcal{B} \right\},$$
$$H(N, f, \eta) \coloneqq \iint_{\mathbb{R}^3_p \times \mathbb{R}^3_k} d\mathbf{k} d\mathbf{p} \left\{ \eta \mathcal{L}\mathcal{E}\mathcal{L} f N \mathcal{B} \right\}.$$

As a result, the unconditionally conservative weak form(3.21) can be written as,

$$\begin{cases} (\partial_t f, \varphi)_p = -B(f, N, \varphi), \\ (\partial_t N, \eta)_k = H(N, f, \eta). \end{cases}$$

With this form it is convenient to verify the dissipation of entropy. Recall the definition of emission/absorption kernel $\mathcal{B}$,

$$\mathcal{B}(\mathbf{p}, \mathbf{k}) = \sum_{l=-\infty}^{+\infty} U_l(\mathbf{p}; \mathbf{k}) \delta(s_l(\mathbf{p}; \mathbf{k})).$$

Test the equation for particle *pdf* with $\varphi = \log f$, since U_l and N are non-negative, the right-hand side will be non-positive:

$$(\partial_t f, \log f)_p = -\iint d\mathbf{k} d\mathbf{p} \frac{1}{f} (\mathcal{L} f)^2 N \hbar\omega \mathcal{B} \leq 0.$$

The H-theorem for the particle *pdf* follows from the above inequality, as

$$\partial_t \left(f, \log f\right)_p = \left(\partial_t f, \log f\right)_p + \left(\partial_t f, 1\right)_p \leq 0.$$

3.4 The conservative discretization

In this section, we pursue a semi-discrete problem that consistently approximates the original system, and at the same time preserves discrete conservation laws. In the following subsections, we will first introduce our finite element discretization, the necessary projection operators, and then elaborate on the conservation technique.

Before continuing, note that is also possible to use wave spectral energy density $W(\mathbf{k}, t) \coloneqq N(\mathbf{k}, t)\hbar\omega(\mathbf{k})$ instead of $N(\mathbf{k}, t)$ as the unknown. The weak form for particle probability density function f and wave spectral energy density W reads:

$$\int_p \varphi \partial_t f + \int_k \eta \partial_t W = \iint_{pk} \mathcal{B} W \mathcal{L} f \left(\eta \mathcal{L}\mathcal{E} - \mathcal{L}\varphi\right). \tag{3.23}$$

The above form is also unconditionally conservative.

Equation(3.21) for plasmon probability density N and (3.23) for wave spectral energy density W are equally suitable for conservative discretization. In the rest of this section, we show the discretization of Equation(3.23), while the conservative discretization of Equation (3.21) will be given Section 3.8 with DG schemes. Nevertheless, the readers should be aware that all of the four types of combinations are feasible.

3.4.1 The finite element discretization

Cut-off domains and boundary conditions

Analogous to existing work on kinetic equations, for example, Zhang and Gamba (2017, 2018), we assume that given any $0 < \epsilon_p \ll 1$ and $0 < \epsilon_k \ll 1$, there exists finite cylindrical domains $\Omega_p^L \subsetneq \mathbb{R}_p^3$ and $\Omega_k^L \subsetneq \mathbb{R}_k^3$ such that for any $t \geq 0$,

$$\left| 1 - \frac{\int_{\Omega_p^L} f(\mathbf{p}, t) d^3p}{\int_{\mathbb{R}_p^3} f(\mathbf{p}, t) d^3p} \right| \leq \epsilon_p,$$

and

$$\left| 1 - \frac{\int_{\Omega_k^L} W(\mathbf{k}, t) d^3k}{\int_{\mathbb{R}_k^3} W(\mathbf{k}, t) d^3k} \right| \leq \epsilon_k.$$

The particle momentum cut-off domain Ω_p^L is supposed to be adaptive, while in our numerical experiments it turns out that, as a result of anisotropic diffusion, there is no need to extend it.

Then it is reasonable to solve the equations in cut-off domains Ω_p^L and Ω_k^L. For the wave *sed* $W(\mathbf{k})$, there is no need for a boundary condition since there is no flux in wave vector space. For the particle *pdf*, we have the following choices, and when the domain Ω_p^L is large enough, they are actually equivalent.

On the boundary $\partial\Omega_p^L$ of cut-off domain Ω_p^L, $|f|$ and $|\nabla_{\mathbf{p}} f|$ are nearly zero, two types of boundary conditions can be applied,

1. The zero-value boundary condition

$$f = 0, \ \forall \mathbf{p} \in \partial\Omega_p^L.$$

2. The zero-flux boundary condition

$$(D[W]\nabla_{\mathbf{p}} f) \cdot \mathbf{n} = 0, \ \forall \mathbf{p} \in \partial\Omega_p^L.$$

Suppose we test the diffusion equation with $\varphi_h \in V_h$. With Neumann's boundary condition, i.e. in the zero-flux case, the semi-discrete weak form reads:

$$(\frac{\partial f_h}{\partial t}, \varphi_h) + (D[W_h]\nabla_{\mathbf{p}} f_h, \nabla_{\mathbf{p}} \varphi_h) = 0.$$

For Dirichlet's boundary conditions given by to zero-value on the discretized boundary, i.e. $f_{\text{ini}} \mid_{\partial\Omega_p^L} \equiv 0$, Nitsche's method (Nitsche, 1971) applies, hence the weak the semi-discrete form reads

$$(\frac{\partial f_h}{\partial t}, \varphi_h) + (D[W_h]\nabla_{\mathbf{p}} f_h, \nabla_{\mathbf{p}} \varphi_h) - \langle (D[W_h]\nabla_{\mathbf{p}} f_h) \cdot \mathbf{n}_p, \varphi_h \rangle_{\partial\Omega_p^L} + \langle (D[W_h]\nabla_{\mathbf{p}} \varphi_h) \cdot \mathbf{n}_p, f_h \rangle_{\partial\Omega_p^L} = 0.$$

The only difference between them is the boundary integral, which can be below machine epsilon for large enough Ω_p^L, because D and φ_h are finite, while $|f_h|$ and $|\nabla_{\mathbf{p}} f_h|$ goes to zero as we enlarge the domain. Stability can be proved for both formulations, in the rest of the article, for simplicity, we will use the zero-flux boundary condition.

Discrete test/trial function spaces

Since we have assumed cylindrical symmetry, the 3P-3K problem actually becomes 2P-2K.

$$\mathbf{p} = (p_1, p_2, p_3) \in \Omega_p^L \Leftrightarrow (p_\parallel, p_\perp) \in \tilde{\Omega}_p^L \subset \mathbb{R} \times \mathbb{R}^+$$
$$\mathbf{k} = (k_1, k_2, k_3) \in \Omega_k^L \Leftrightarrow (k_\parallel, k_\perp) \in \tilde{\Omega}_k^L \subset \mathbb{R} \times \mathbb{R}^+$$

Let $\mathcal{T}_h^p = \{R_p\}$, $\mathcal{T}_h^k = \{R_k\}$ be rectangular partitions of $\tilde{\Omega}_p^L$ and $\tilde{\Omega}_k^L$ respectively. We define the meshsize for momentum space as $h_p = \max_{R_p \in \mathcal{T}_h^p} \mathrm{diam}(R_p)$ and the meshsize for wave vector space as $h_k = \max_{R_k \in \mathcal{T}_h^k} \mathrm{diam}(R_k)$.

The test space for particle *pdf* consists of continuous piecewise polynomials with degree α_1,

$$\mathcal{G}_h^{\alpha_1} = \{f(p_\parallel, p_\perp) \in C^0(\Omega_p) : f|_{R_p} \in Q^{\alpha_1}(R_p), \forall\ R_p \in \mathcal{T}_h^p\}. \tag{3.24}$$

The test space for wave *sed* consists of discontinuous piecewise polynomials with degree α_2,

$$\mathcal{W}_h^{\alpha_2} = \{W(k_\parallel, k_\perp) : W|_{R_k} \in Q^{\alpha_2}(R_k), \forall\ R_k \in \mathcal{T}_h^k\}. \tag{3.25}$$

To ensure positivity of W_h, it is required that $\alpha_2 = 0$ or $\alpha_2 = 1$, the reason will be addressed later.

As will be shown in the next section, one of the key points to conservation is replacing $v_\parallel$, $v_\perp$ and $k_\parallel/\omega$ with $\frac{\partial \mathcal{E}_h(\mathbf{p})}{\partial p_\parallel}$, $\frac{\partial \mathcal{E}_h(\mathbf{p})}{\partial p_\perp}$ and $N_{\parallel,h}$, where $\mathcal{E}_h = \Pi_{p,h}\mathcal{E}(\mathbf{p})$ is the discrete particle kinetic energy, and $N_{\parallel,h} = \Pi_{k,h} N_\parallel$ is the discrete refraction index.

The projection operators can be arbitrarily chosen as long as they satisfy the following conditions:

1. The projection $\Pi_{p,h}$ into test space $\mathcal{G}_h^{\alpha_1}$ must satisfy that

$$\lim_{h\to 0} \|\Pi_{p,h} g(\mathbf{p}) - g(\mathbf{p})\|_{L^2(\Omega_p^L)} = 0, \ \forall g \in L^2(\Omega_p^L),$$

and

$$\lim_{h\to 0} \|\Pi_{p,h} \mathcal{E}(\mathbf{p}) - \mathcal{E}(\mathbf{p})\|_{H^1(\Omega_p^L)} = 0.$$

2. The projection $\Pi_{k,h}$ into test space $\mathcal{W}_h^{\alpha_2}$ must satisfy that

$$\lim_{h\to 0} \|\Pi_{k,h} \xi(\mathbf{k}) - \xi(\mathbf{k})\|_{L^2(\Omega_k^L)} = 0, \ \forall \xi \in L^2(\Omega_k^L). \tag{3.26}$$

There is no need to specify particular projections until we implement them in the numerical examples, our method works with any of them.

3.4.2 The conservative semi-discrete form

Adopting the zero-flux boundary condition, testing the system on the cut-off domain with $\varphi_h \in \mathcal{G}_h^{\alpha_1}$ and $\eta_h \in \mathcal{W}_h^{\alpha_2}$, we write the following semi-discrete weak form,

$$\begin{aligned}
(\frac{\partial f_h}{\partial t}, \varphi_h)_p &= -B_L^u(f_h, W_h, \varphi_h) \coloneqq -\iint_{\Omega_k^L \times \Omega_p^L} d\mathbf{k}d\mathbf{p}\{\mathcal{L}\varphi_h \mathcal{L} f_h W_h \mathcal{B}\}, \\
(\frac{\partial W_h}{\partial t}, \eta_h)_k &= H_L^u(W_h, f_h, \eta_h) \coloneqq \iint_{\Omega_k^L \times \Omega_p^L} d\mathbf{k}d\mathbf{p}\{\eta_h \mathcal{L} E \mathcal{L} f_h W_h \mathcal{B}\},
\end{aligned}$$

where the subscript L means integral on cut-off domain, the superscript u means unconservative. We will first analyze the source of conservation errors and then present our conservative semi-discrete trilinear forms B_L and H_L.

Source of conservation errors

Suppose different quadrature rules R_1 and R_2 are used for different equations,

$$(\frac{\partial f_h}{\partial t}, \varphi_h)_p = R_1 \left[- \int_{\Omega_p^L} d\mathbf{p} \int_{\Omega_k^L} d\mathbf{k} \{ \mathcal{L}\varphi_h \mathcal{L} f_h W_h \mathcal{B} \} \right],$$
$$(\frac{\partial W_h}{\partial t}, \eta_h)_k = R_2 \left[\int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p} \{ \eta_h \mathcal{L} E \mathcal{L} f_h W_h \mathcal{B} \} \right].$$

The error of conservation laws can be decomposed into three terms,

$$\begin{aligned}
&\frac{\partial}{\partial t}((f_h, \Pi_{p,h}\varphi)_p + (W_h, \Pi_{k,h}\eta)_k) \\
=&R_1 \left[- \int_{\Omega_p^L} d\mathbf{p} \int_{\Omega_k^L} d\mathbf{k} \left\{ (\mathcal{L}\Pi_{p,h}\varphi) \mathcal{L} f_h W_h \mathcal{B} \right\} \right] + R_2 \left[\int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p} \left\{ (\Pi_{k,h}\eta \mathcal{L} E) \mathcal{L} f_h W_h \mathcal{B} \right\} \right] \\
=&A_1 + A_2 + A_3,
\end{aligned}$$

where

$$A_1 = (R_1 - I) \left[\int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p} \left\{ (-\mathcal{L}\Pi_{p,h}\varphi) \mathcal{L} f_h W_h \mathcal{B} \right\} \right],$$
$$A_2 = (I - R_2) \left[\int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p} \left\{ (-\mathcal{L}\Pi_{p,h}\varphi) \mathcal{L} f_h W_h \mathcal{B} \right\} \right],$$
$$A_3 = R_2 \left[\int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p} \left\{ (\Pi_{k,h}\eta \mathcal{L} E - \mathcal{L}\Pi_{p,h}\varphi) \mathcal{L} f_h W_h \mathcal{B} \right\} \right].$$

The error terms A_1 and A_2 are caused by inconsistent numerical integration on the resonance manifold. Suppose that $R_1 - I$ is of the same order as $O(h^a)$, and quadrature rule R_2 has error $O(h^b)$, then the sum will be roughly $O(h^{\min\{a,b\}})$. The last error term A_3 is a result of projection error, whose order depends on the degree of test spaces, α_1 and α_2.

Note that A_1 and A_2 cancel out when we use the same quadrature rules, i.e. $R_1 = R_2$. In what follows, we will introduce a conservative semi-discrete form such that A_3 disappears.

Conservative semi-discrete form

Recall the definition of directional differential operator $\mathcal{L}$,

$$\mathcal{L}g = \frac{k_\parallel v_\parallel}{\omega}\frac{p}{p_\parallel}\frac{\partial g}{\partial p_\parallel} + (1 - \frac{k_\parallel v_\parallel}{\omega})\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp} = N_\parallel \frac{\partial \mathcal{E}}{\partial p_\perp}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\parallel} + (1 - N_\parallel \frac{\partial \mathcal{E}}{\partial p_\parallel})\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}.$$

We propose a discretized operator $\mathcal{L}_h$ defined as follows,

$$\mathcal{L}_h g \coloneqq N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\perp}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\parallel} + (1 - N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\parallel})\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp}, \tag{3.27}$$

where the discretized kinetic energy is defined as $\mathcal{E}_h = \Pi_{p,h}\mathcal{E}(\mathbf{p})$, and the discretized wave refractive index is defined as $N_{\parallel,h} = \Pi_{k,h}N_\parallel = \Pi_{k,h}\frac{k_\parallel}{\omega(\mathbf{k})}$.

The main result of this chapter is stated in the following theorem.

Theorem 3.4. *If $f_h(\mathbf{p},t)$ and $W_h(\mathbf{k},t)$ are solutions of the following semi-discrete weak form,*

$$\begin{aligned}
(\frac{\partial f_h}{\partial t}, \varphi_h)_p &= -B_L(f_h, W_h, \varphi_h) \coloneqq -\iint_{\Omega_k^L\times\Omega_p^L} d\mathbf{k}d\mathbf{p}\{\mathcal{L}_h\varphi_h\mathcal{L}_h f_h W_h \mathcal{B}_h\},\\
(\frac{\partial W_h}{\partial t}, \eta_h)_k &= H_L(W_h, f_h, \eta_h) \coloneqq \iint_{\Omega_k^L\times\Omega_p^L} d\mathbf{k}d\mathbf{p}\{\eta_h\mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h W_h \mathcal{B}_h\},
\end{aligned} \tag{3.28}$$

then the following discrete conservation laws hold,

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{M}_{tot,h} &= \frac{\partial}{\partial t}((f_h, \Pi_{p,h}1)_p + (W_h, 0)_k) = 0,\\
\frac{\partial}{\partial t}\mathcal{P}^\parallel_{tot,h} &= \frac{\partial}{\partial t}\left((f_h, \Pi_{p,h}p_\parallel)_p + (W_h, \Pi_{k,h}N_\parallel)_k\right) = 0,\\
\frac{\partial}{\partial t}\mathcal{E}_{tot,h} &= \frac{\partial}{\partial t}((f_h, \Pi_{p,h}\mathcal{E}(\mathbf{p}))_p + (W_h, \Pi_{k,h}1)_k) = 0.
\end{aligned}$$

Proof. Substitute the discrete conservation pairs $\{\Pi_{p,h}1, 0\}$, $\{\Pi_{p,h}p_\parallel, \Pi_{k,h}N_\parallel\}$ and $\{\Pi_{p,h}\mathcal{E}(\mathbf{p}), \Pi_{k,h}1\}$ into semi-discrete form (3.28) and use the definition of $\mathcal{L}_h$. □

Corollary 3.5. *If in addition to the assumptions of Theorem(3.4), the projections are L^2 orthogonal projections, i.e.*

$$(u - \Pi_{p,h}u, v)_p = 0, \forall v \in \mathcal{G}_h^{\alpha_1}$$

and

$$(U - \Pi_{k,h}U, V)_k = 0, \forall V \in \mathcal{W}_h^{\alpha_2},$$

then the exact conservation laws are preserved, i.e.

$$\begin{aligned}
\frac{\partial}{\partial t}\mathcal{M}_{tot,h} &= \frac{\partial}{\partial t}((f_h, 1)_p + (W_h, 0)_k) = 0, \\
\frac{\partial}{\partial t}\mathcal{P}^{\|}_{tot,h} &= \frac{\partial}{\partial t}\left((f_h, p_{\|})_p + (W_h, N_{\|})_k\right) = 0, \\
\frac{\partial}{\partial t}\mathcal{E}_{tot,h} &= \frac{\partial}{\partial t}((f_h, \mathcal{E}(\mathbf{p}))_p + (W_h, 1)_k) = 0.
\end{aligned}$$

Proof. Use the fact that $f_h \in \mathcal{G}_h^{\alpha_1}$ and $W_h \in \mathcal{W}_h^{\alpha_2}$. □

Remark 3.8. Same as stated in Theorem(3.3), our semi-discrete weak form is also unconditionally conservative, i.e. the conservation does not depend on a particular discrete emission/absorption kernel $\mathcal{B}_h$.

3.5 Sparse interaction tensors

Suppose that the test spaces are spanned by basis functions, i.e. $\mathcal{G}_h^{\alpha_1} = \text{span}\{\varphi_i\}$ and $\mathcal{W}_h^{\alpha_2} = \text{span}\{\eta_j\}$. Then we can express the discrete particle *pdf* f_h and wave *sed* W_h as a linear combination of basis functions.

$$\begin{aligned}
f_h(\mathbf{p}, t) &= \sum_{i=1}^{N_f} a_i(t)\varphi_i(\mathbf{p}), \\
W_h(\mathbf{k}, t) &= \sum_{j=1}^{N_w} w_j(t)\eta_j(\mathbf{k}).
\end{aligned}$$

By definition, $\mathcal{E}_h = \Pi_{p,h}\mathcal{E} \in \mathcal{G}_h^{\alpha_1}$, therefore it is also a linear combination of basis functions, $\mathcal{E}_h = \sum_{q=1}^{N_f} \mathcal{E}_q\varphi_q$.

Substitute the above expressions into Equation(3.28), then the semi-discrete system becomes a first-order finite dimension ODE system:

$$\sum_{i=1}^{N_f} \frac{\partial a_i}{\partial t} \int \varphi_i \varphi_m d^3p = -\sum_{n=1}^{N_f} \sum_{k=1}^{N_w} a_n w_k \int_{\Omega_p^L} d\mathbf{p} \int_{\Omega_k^L} d\mathbf{k} \{\mathcal{L}_h \varphi_m \mathcal{L}_h \varphi_n \eta_k \mathcal{B}(\mathbf{p}; \mathbf{k})\},$$
$$\sum_{j=1}^{N_w} \frac{\partial w_j}{\partial t} \int \eta_j \eta_q d^3k = \sum_{k=1}^{N_w} \sum_{n=1}^{N_f} w_k a_n \int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p} \{\eta_q \mathcal{L}_h \mathcal{E}_h \mathcal{L}_h \varphi_n \eta_k \mathcal{B}(\mathbf{p}; \mathbf{k})\}.$$

Denote the mass matrix for particle *pdf* as $A_{im} = (\varphi_i, \varphi_m)_p$, and denote the mass matrix for wave *sed* as $G_{jq} = (\eta_j, \eta_q)_k$.

Analogously, define the interaction tensors B and H corresponding to the trilinear forms.

$$B_{nkm} = B(\varphi_n, \eta_k, \varphi_m),$$
$$H_{knq} = H(\eta_k, \varphi_n, \eta_q).$$

As a result, we obtain the nonlinear ODE system corresponding to semi-discrete weak form(3.28):

$$\begin{aligned} \frac{\partial a_i}{\partial t} A_{im} &= -a_n w_k B_{nkm}, \\ \frac{\partial w_j}{\partial t} G_{jq} &= w_k a_n H_{knq}. \end{aligned} \tag{3.29}$$

The interaction tensors B and H are both sparse tensors for two reasons: compactly supported basis and the resonant feature of trilinear forms. Taking particle interaction tensor B as an example, $B_{nkm} = 0$ when

1. φ_m and φ_n are not in neighboring elements.

2. φ_m and η_k do not "resonate", i.e. $\text{supp}(\varphi_m) \times \text{supp}(\eta_k)$ does not intersect with the resonant manifold.

Suppose in each dimension we have $O(n)$ meshes, then the shape of particle interaction tensor B is roughly $O(n^2) \times O(n^2) \times O(n^2)$, while the number of nonzero

elements will be only $O(n^3)$, i.e. the sparsity of tensor B is about $1-\frac{1}{O(n^3)}$. A similar analysis can also be applied to the wave interaction tensor H.

We observed that the trilinear forms B and H defined in Equation(3.3.3) have similar structures. Therefore one might wonder if there is any relation between the interaction tensors B and H. It turns out that when $\alpha_2 = 0$, i.e piecewise constant basis functions are used for wave *sed* W_h, we can infer any nonzero element of wave interaction tensor H from particle interaction tensor B. In practice, the interaction tensors are precomputed and saved for later use. Taking advantage of this relation, we can save half the time of preprocessing. The derivation is as follows.

When $\alpha_2 = 0$, $\mathcal{W}_h^{\alpha_2} = \text{span}\{\eta_j\}$ are piecewise constant functions, we have

$$\eta_i(\mathbf{k})\eta_j(\mathbf{k}) = \delta_{ij}\eta_i(\mathbf{k}).$$

Then the mass matrix for wave *sed* is diagonal,

$$G_{jq} = (\eta_j, \eta_q)_k = \int_{R_k^j} \delta_{jq} d^3k = \text{diag}(\mu(R_k^j)),$$

where $\mu(R_k^j) = \int_{R_k^j} 1 d^3k$ is the measure of j-th element in Ω_k^L.

Moreover, note that if we define a 4-th order tensor

$$\tilde{H}_{mknq} \coloneqq \int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\varphi_m\mathcal{L}_h\varphi_n\eta_k\eta_q\mathcal{B}(\mathbf{p};\mathbf{k})\}.$$

Recall the expansion $\mathcal{E}_h = \sum_{q=1}^{N_f} \mathcal{E}_q\varphi_q$, and substitute it into the definition of wave interaction tensor H, we obtain the relation between H_{knq} and $\tilde{H}_{mknq}$,

$$H_{knq} = H(\eta_k, \varphi_n, \eta_q) = \sum_{m=1}^{N_f} \mathcal{E}_m \tilde{H}_{mknq}.$$

It can be observed that the form of $\tilde{H}_{mknq}$ is almost identical to the definition of particle tensor B_{mnk}, except for the extra η_q. Replace $\eta_k(\mathbf{k})\eta_q(\mathbf{k})$ with $\delta_{kq}\eta_k(\mathbf{k})$, we obtain the relation between $\tilde{H}_{mknq}$ and B_{mnk},

$$\tilde{H}_{mknq} = \delta_{kq} \int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\varphi_m\mathcal{L}_h\varphi_n\eta_k\mathcal{B}(\mathbf{p};\mathbf{k})\} = \delta_{kq}B_{mnk}.$$

Therefore B_{mnk} and $\mathcal{E}_m$ is all we need to calculate H_{knq},

$$H_{knq} = \sum_{m=1}^{N_f} \mathcal{E}_m \tilde{H}_{mknq} = \sum_{m=1}^{N_f} \mathcal{E}_m \delta_{kq} B_{mnk} = \delta_{kq} \sum_{m=1}^{N_f} \mathcal{E}_m B_{mnk} = \begin{cases} 0, & k \neq q \\ \sum_{m=1}^{N_f} \mathcal{E}_m B_{mnk} & k = q \end{cases} \tag{3.30}$$

3.6 Stability and positivity

In this section, we investigate the stability of the fully discretized nonlinear system. With semi-implicit time discretization, there is no constraint on time step size from the CFL condition. However, the stability will rely on the positivity of W_h, which results in a condition for the time step size, relevant to the gradient of particle *pdf* f_h. The condition will not cause any trouble for implementation, because we can always adapt the step size a posteriori.

3.6.1 Stability of the semi-discrete form

Consider the equation for particle *pdf* only, it has the form of a diffusion equation, thus its stability relies on the fact that the diffusion coefficient is positive semi-definite, which further relies on the positivity of wave *sed* W_h.

Lemma 3.6 (L^2 stability of $f_h(\mathbf{p})$ and L^1 bound of $W_h(\mathbf{k})$)**.** *Suppose* $f_h(\mathbf{p}, t)$ *and* $W_h(\mathbf{k}, t)$ *are the solution of equation(3.28) with the following initial condition:*

$$f_h(\mathbf{p}, 0) = f_h^0(\mathbf{p}),$$
$$W_h(\mathbf{k}, 0) = W_h^0(\mathbf{k}).$$

If W_h *always takes non-negative values, i.e.* $W_h(\mathbf{k}, t) \geq 0, \forall\ \mathbf{k} \in \Omega_k^L, \forall t \geq 0$, *then* f_h *has* L^2 *stability*

$$\|f_h\|_{L^2(\Omega_p^L)} \leq \|f_h^0\|_{L^2(\Omega_p^L)},$$

and W_h *has bounded* L^1 *norm.*

$$\|W_h\|_{L^1(\Omega_k^L)} \leq \mathcal{E}_{tot,h}^0 + \|f_h^0\|_{L^2(\Omega_p^L)} \cdot \|\Pi_{p,h}\mathcal{E}\|_{L^2(\Omega_p^L)}.$$

Proof. Since f_h belongs to the test space $\mathcal{G}_h^{\alpha_1}$, we test the equation for particles with f_h, we obtain that

$$(\frac{\partial f_h}{\partial t}, f_h)_p = -\int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p}\{(\mathcal{L}_h f_h)^2 W_h \mathcal{B}\}.$$

The right hand side is non-positive as long as W_h always take non-negative values, therefore the L^2 norm of f_h always decreases,

$$\frac{1}{2}\frac{\partial}{\partial t}\|f_h\|^2_{L^2(\Omega_p^L)} \leq 0 \Rightarrow \|f_h\|_{L^2(\Omega_p^L)} \leq \|f_h^0\|_{L^2(\Omega_p^L)}.$$

Now consider W_h, by definition,

$$\|W_h\|_{L^1(\Omega_k^L)} = (W_h, \mathrm{sgn}(W_h))_k = (W_h, 1)_k.$$

Recall the energy conservation property in Theorem 3.4:

$$\|W_h\|_{L^1(\Omega_k^L)} + (f_h, \Pi_{p,h}\mathcal{E})_p = \mathcal{E}^0_{tot,h}.$$

Use Holder's inequality

$$\begin{aligned}\|W_h\|_{L^1(\Omega_k^L)} &= \mathcal{E}^0_{tot,h} - (f_h, \Pi_{p,h}\mathcal{E})_p \\ &\leq \mathcal{E}^0_{tot,h} + |(f_h, \Pi_{p,h}\mathcal{E})_p| \\ &\leq \mathcal{E}^0_{tot,h} + \|f_h\|_{L^2(\Omega_p^L)} \cdot \|\Pi_{p,h}\mathcal{E}\|_{L^2(\Omega_p^L)}.\end{aligned}$$

By the L^2-stability of f_h, we obtain the upper bound of W_h's L^1 norm,

$$\|W_h\|_{L^1(\Omega_k^L)} \leq \mathcal{E}^0_{tot,h} + \|f_h\|_{L^2(\Omega_p^L)} \cdot \|E\|_{L^2(\Omega_p^L)} \leq \mathcal{E}^0_{tot,h} + \|f_h^0\|_{L^2(\Omega_p^L)} \cdot \|\Pi_{p,h}\mathcal{E}\|_{L^2(\Omega_p^L)}.$$

□

3.6.2 Time discretization

Recall our conservative semi-discrete weak form,

$$\begin{aligned}(\frac{\partial f_h}{\partial t}, \varphi_h)_p &= -B_L(f_h, W_h, \varphi_h) \coloneqq -\iint_{\Omega_k^L \times \Omega_p^L} d\mathbf{k}d\mathbf{p}\{\mathcal{L}_h\varphi_h \mathcal{L}_h f_h W_h \mathcal{B}_h\}, \\ (\frac{\partial W_h}{\partial t}, \eta_h)_k &= H_L(W_h, f_h, \eta_h) \coloneqq \iint_{\Omega_k^L \times \Omega_p^L} d\mathbf{k}d\mathbf{p}\{\eta_h \mathcal{L}_h \mathcal{E}_h \mathcal{L}_h f_h W_h \mathcal{B}_h\},\end{aligned}$$

The time step size of the explicit scheme for diffusion equations is restricted by the CFL condition. Two reasons urge us to avoid explicit schemes,

1. The CFL bound of step size may be too restrictive, and we might lose efficiency.

2. The upper bound depends on the eigenvalues of time-varying diffusion coefficients. However, in the proposed scheme, we never calculate the diffusion coefficient explicitly, instead, we compute the interaction tensor B associated with the trilinear form B.

On the other hand, due to nonlinearity, a fully implicit scheme requires fixed-point iteration involving both particle *pdf* f_h and wave *sed* W_h, which can be time-consuming. Therefore, the objective is to find a scheme that is only implicit for f_h, and at the same time preserves discrete conservation laws.

We propose the following semi-implicit scheme,

$$\begin{aligned} (\frac{f_h^{s+1}-f_h^s}{\Delta t},\varphi_h)_p + B_L(f_h^{s+1},W_h^s,\varphi_h) = 0,\\ (\frac{W_h^{s+1}-W_h^s}{\Delta t},\eta_h)_k - H_L(W_h^s,f_h^{s+1},\eta_h) = 0. \end{aligned} \tag{3.31}$$

The scheme is implicit for particle *pdf* f_h if we focus on the first line, meanwhile it is explicit for wave *sed* W_h, considering the second line. For implementation, we solve the first row and then substitute the next step particle *pdf* f_h^{s+1} into the second row. It can be easily verified that the discrete conservation laws still hold, i.e. we have

$$(f_h^{s+1},\varphi_{c,h})_p + (W_h^{s+1},\eta_{c,h})_k = (f_h^s,\varphi_{c,h})_p + (W_h^s,\eta_{c,h})_k.$$

The following theorem is the fully discrete version of Lemma 3.6, giving the unconditional L^2-stability of f_h^s when W_h^s is non-negative.

Theorem 3.7. *Suppose* $f_h^s(\mathbf{p})$ *and* $W_h^s(\mathbf{k})$ *are the solution of Equation(3.31).*

If W_h^s always takes non-negative values, i.e. $W_h^s(\mathbf{k}) \geq 0, \forall\ \mathbf{k} \in \Omega_k^L, \forall s \geq 0$, then f_h^s has L^2 stability

$$\|f_h^s\|_{L^2(\Omega_p^L)} \leq \|f_h^0\|_{L^2(\Omega_p^L)}.$$

and W_h^s has bounded L^1 norm,

$$\|W_h^s\|_{L^1(\Omega_k^L)} \leq \mathcal{E}_{tot,h}^0 + \|f_h^0\|_{L^2(\Omega_p^L)} \cdot \|\Pi_{p,h}\mathcal{E}\|_{L^2(\Omega_p^L)}.$$

Proof. Given that $W_h^s(\mathbf{k}) \geq 0, \forall \mathbf{k} \in \Omega_k^L$, we have $B_L(f_h^{s+1}, W_h^s, f_h^{s+1}) \geq 0$. Therefore, f_h^s has unconditional L^2-stability.

Since the scheme(3.31) preserves energy conservation, the L^1 bound of W_h can be proved in the same approach as we have done in Lemma 3.6. □

Note that the stability depends on our assumption that W_h^s is non-negative. Therefore, in what follows, we will discuss the positivity-preserving technique of W_h^s.

3.6.3 Positivity-preserving technique for the wave SED

To ensure positivity of wave *sed* W_h^s, we draw the strategy from Zhang and Shu (2010):

1. Use a small enough time step to ensure positive cell-average of a temporary wave *sed* $W_h^{s+1,*}$, given that we have pointwise positivity of last step wave *sed* W_h^s.

2. Apply a slope limiter on $W_h^{s+1,*}$ which preserves cell-average at the same time, then we obtain a pointwise positive W_h^{s+1} as our solution of the next step wave *sed*. (Obviously, if we use piecewise constant basis functions, this step is not necessary).

Firstly we will derive the constraint on time step size. After that, we explain why the slope limiter will not break discrete conservation laws.

Suppose $\eta_{j,0}$ is the characteristic function of the j-th element $R_k^j \subset \Omega_k^L$, i.e. $\eta_{j,0} = 1_{\mathbf{k}\in R_k^j}$, which belongs to the test space $\mathcal{W}_h^{\alpha_2}$. According to the time discretization in Equation(3.31),

$$(\frac{W_h^{s+1,*} - W_h^s}{\Delta t}, \eta_{j,0})_k = \int_{\Omega_k^L} \frac{W_h^{s+1,*} - W_h^s}{\Delta t}\eta_{j,0} d\mathbf{k} = \int_{\Omega_k^L} d\mathbf{k} \int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h^{s+1} W_h^s \eta_{j,0}\mathcal{B}_h\},$$

which is equivalent to

$$\int_{R_k^j} W_h^{s+1,*} d^3k = \int_{R_k^j} W_h^s (1 + \Delta t \int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h^{s+1}\mathcal{B}_h\}) d\mathbf{k}.$$

To ensure positive cell-average, i.e. $\int_{R_k^j} W_h^{s+1,*} d\mathbf{k} \geq 0$, we require that there exists a constant $\epsilon > 0$ such that

$$1 + \Delta t \int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h^{s+1}\mathcal{B}_h\} \geq \epsilon, \ \forall \mathbf{k} \in R_k^j. \tag{3.32}$$

As long as the time step size Δt satisfy condition(3.32), we have $\int_{R_k^j} W_h^{s+1,*} d\mathbf{k} \geq \epsilon \int_{R_k^j} W_h^s d\mathbf{k} \geq 0$.

The following theorem guarantees that our bound for Δt will not shrink over time.

Theorem 3.8. *For any $\epsilon > 0$, given a regular enough discrete emission/absorption kernel $\mathcal{B}_h$, there exists a constant Δt_M determined by ϵ, f^0, Ω_p^L, Ω_k^L and h, such that any $\Delta t < \Delta t_M$ satisfies condition (3.32).*

Proof. By Hölder's inequality, the "growth rate" is bounded as follows,

$$\begin{aligned}
\left| \int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h^{s+1}\mathcal{B}_h\} \right| \leq & \int_{\Omega_p^L} d\mathbf{p} \left| \mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h^{s+1}\mathcal{B}_h \right| \\
\leq & \left\| \frac{p_\perp}{p}\mathcal{L}_h f_h^{s+1} \right\|_{L^\infty(\Omega_p^L)} \cdot \left\| \frac{p}{p_\perp}\mathcal{B}_h\mathcal{L}_h\mathcal{E}_h \right\|_{L^1(\Omega_p^L)} \\
= & 2\pi \left\| \frac{p_\perp}{p}\mathcal{L}_h f_h^{s+1} \right\|_{L^\infty(\Omega_p^L)} \cdot \left(\int p\mathcal{B}_h\mathcal{L}_h\mathcal{E}_h dp_\perp dp_\parallel \right).
\end{aligned} \tag{3.33}$$

Firstly, consider the L^∞-norm factor from inequality (3.33). Recall the definition of $\mathcal{L}_h$,

$$\frac{p_\perp}{p}\mathcal{L}_h f_h^{s+1} \coloneqq N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\perp}\frac{\partial f_h^{s+1}}{\partial p_\parallel} + (1 - N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\parallel})\frac{\partial f_h^{s+1}}{\partial p_\perp}.$$

Both of the coefficients $N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\perp}$ and $(1 - N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\parallel})$ are bounded by some constant C dependent on Ω_p^L and Ω_k^L, hence it follows that,

$$\begin{aligned}
\left\|\frac{p_\perp}{p}\mathcal{L}_h f_h^{s+1}\right\|_{L^\infty(\Omega_p^L)} &= \left\|N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\perp}\frac{\partial f_h^{s+1}}{\partial p_\parallel} + (1 - N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\parallel})\frac{\partial f_h^{s+1}}{\partial p_\perp}\right\|_{L^\infty(\Omega_p^L)} \\
&\leq C(\Omega_p^L, \Omega_k^L)\left\|\frac{\partial f_h^{s+1}}{\partial p_\parallel}\right\|_{L^\infty(\Omega_p^L)} + C(\Omega_p^L, \Omega_k^L)\left\|\frac{\partial f_h^{s+1}}{\partial p_\perp}\right\|_{L^\infty(\Omega_p^L)} \\
&\leq C_1(\Omega_p^L, \Omega_k^L)\cdot\left\|\nabla_p f_h^{s+1}\right\|_{L^\infty(\Omega_p^L)}.
\end{aligned} \tag{3.34}$$

We claim that $\left\|\nabla_p f_h^{s+1}\right\|_{L^\infty(\Omega_p^L)}$ is bounded uniformly in time. Indeed, since the domain Ω_p^L is finite, all L^r norms are equivalent, therefore,

$$\left\|\nabla_p f_h^{s+1}\right\|_{L^\infty(\Omega_p^L)} \leq C_2(\Omega_p^L)\left\|\nabla_p f_h^{s+1}\right\|_{L^2(\Omega_p^L)}.$$

Moreover, the inverse inequality for finite element spaces,

$$\left\|\nabla_p f_h^{s+1}\right\|_{L^2(\Omega_p^L)} \leq \frac{C_3}{h_p}\left\|f_h^{s+1}\right\|_{L^2(\Omega_p^L)}, \tag{3.35}$$

and the L^2 stability estimate from Theorem 3.7

$$\left\|f_h^{s+1}\right\|_{L^2(\Omega_p^L)} \leq \left\|f_h^0\right\|_{L^2(\Omega_p^L)},$$

leads to the following estimate for $\nabla_p f_h^{s+1}$,

$$\left\|\nabla_p f_h^{s+1}\right\|_{L^\infty(\Omega_p^L)} \leq C_2(\Omega_p^L)\frac{C_3}{h_p}\left\|f_h^0\right\|_{L^2(\Omega_p^L)}. \tag{3.36}$$

Therefore, the L^∞-norm factor from inequality (3.33) is bounded as follows,

$$\left\|\frac{p_\perp}{p}\mathcal{L}_h f_h^{s+1}\right\|_{L^\infty(\Omega_p^L)} \leq C_1(\Omega_p^L, \Omega_k^L)C_2(\Omega_p^L)\frac{C_3}{h_p}\left\|f_h^0\right\|_{L^2(\Omega_p^L)}.$$

Next, consider the L^1-norm factor from inequality (3.33), and write it as follows,

$$\int p\mathcal{B}_h\mathcal{L}_h\mathcal{E}_h dp_\perp dp_\parallel = \int \left[pU_l(\mathbf{p};\mathbf{k})\delta_h(\omega(\mathbf{k}) - k_\parallel v_\parallel - l\omega_{ce}/\gamma(\mathbf{p}))\frac{p}{p_\perp}\frac{\partial \mathcal{E}_h}{\partial p_\perp}\right] dp_\perp dp_\parallel,$$

where δ_h represents an approximation of Dirac delta, see Equation (3.18) and 3.19.

We discuss the following two cases,

- When $l \neq 0$, since

$$\lim_{p_\perp \to 0^+} U_l(\mathbf{p};\mathbf{k})\frac{p^2}{p_\perp} = \lim_{p_\perp \to 0^+} 8\pi^2 e^2 \frac{(iE_2 J_l')^2 \left(\frac{p_\perp}{p}\right)^2}{(1-E_2^2)\frac{1}{\omega}\frac{\partial}{\partial\omega}(\omega^2\varepsilon) + 2iE_2\frac{1}{\omega}\frac{\partial}{\partial\omega}(\omega^2 g) + E_3^2\frac{1}{\omega}\frac{\partial}{\partial\omega}(\omega^2\eta)}\frac{p^2}{p_\perp} = 0,$$

the integral is bounded as follows,

$$\begin{aligned}\int p\mathcal{B}_h\mathcal{L}_h\mathcal{E}_h dp_\perp dp_\parallel &= \int \left[\left(U_l(\mathbf{p};\mathbf{k})\frac{p^2}{p_\perp}\right)\frac{\partial \mathcal{E}_h}{\partial p_\perp}\delta_h(\omega(\mathbf{k}) - k_\parallel v_\parallel - l\omega_{ce}/\gamma(\mathbf{p}))\right] dp_\perp dp_\parallel \\ &\leq \sup_{\Omega_p^L}\left(U_l(\mathbf{p};\mathbf{k})\frac{p^2}{p_\perp}\right)\frac{\partial \mathcal{E}_h}{\partial p_\perp}\int \left[\delta_h(\omega(\mathbf{k}) - k_\parallel v_\parallel - l\omega_{ce}/\gamma(\mathbf{p}))\right] dp_\perp dp_\parallel \\ &\leq C_4(\Omega_p^L, \Omega_k^L).\end{aligned} \tag{3.37}$$

- When $l = 0$, the above trick does not work, because $\lim_{p_\perp \to 0^+} U_0(\mathbf{p};\mathbf{k}) > 0$. For this special case, as an alternative to the original operator

$$\mathcal{L}_h g \coloneqq N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\perp}\frac{p}{p_\perp}\frac{\partial g}{\partial p_\parallel} + (1 - N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\parallel})\frac{p}{p_\perp}\frac{\partial g}{\partial p_\perp},$$

we adopt a new discrete operator,

$$\mathcal{L}_h^0 g \coloneqq N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\perp}\frac{p}{\mathcal{E}_h\frac{\partial \mathcal{E}_h}{\partial p_\perp}}\frac{\partial g}{\partial p_\parallel} + (1 - N_{\parallel,h}\frac{\partial \mathcal{E}_h}{\partial p_\parallel})\frac{p}{\mathcal{E}_h\frac{\partial \mathcal{E}_h}{\partial p_\perp}}\frac{\partial g}{\partial p_\perp}.$$

The operator is still a consistent discretization since $E = \sqrt{1 + p_\parallel^2 + p_\perp^2}$. It can also be easily verified that the L^∞ bound in inequality (3.34) is still true with this new operator $\mathcal{L}_h^0$.

In addition, the L^1-norm factor becomes,

$$\int p\mathcal{B}_h\mathcal{L}_h^0\mathcal{E}_h dp_\perp dp_\parallel = \int \left[pU_l(\mathbf{p};\mathbf{k})\delta_h(\omega(\mathbf{k}) - k_\parallel v_\parallel - l\omega_{ce}/\gamma(\mathbf{p}))\frac{p}{\mathcal{E}_h}\right] dp_\perp dp_\parallel.$$

Therefore inequality (3.37) still holds.

Combine inequalities (3.34), (3.36), and (3.37) to obtain

$$\left|\int_{\Omega_p^L} d\mathbf{p}\{\mathcal{L}_h\mathcal{E}_h\mathcal{L}_h f_h^{s+1}\mathcal{B}_h\}\right| \leq 2\pi\frac{C_1\cdot C_2\cdot C_3\cdot C_4}{h_p}\left\|f_h^0\right\|_{L^2(\Omega_p^L)},$$

which enables us to define the uniform-in-time upper bound,

$$\Delta t_M := \frac{(1-\epsilon)h_p}{2\pi\cdot C_1\cdot C_2\cdot C_3\cdot C_4\left\|f_h^0\right\|_{L^2(\Omega_p^L)}}.$$

It can be easily verified that any $\Delta t < \Delta t_M$ satisfies condition (3.32). □

The condition does not need to be calculated explicitly, because we can adapt time step size a posteriori in the code: monitor the cell averages, if any cell average of the temporary solution $W_h^{s+1,*}$ is non-positive, replace Δt with $0.5\Delta t$ and calculate $W_h^{s+1,*}$ again.

Now let us discuss the effect of slope limiters on conservation laws. If $\alpha_2 = 0$, there is no need for any slope limiter. If $\alpha_2 = 1$, we apply the slope limiter θ and obtain $W_h^{s+1} = \theta(W_h^{s+1,*})$. According to Zhang and Shu (2010), the cell average is preserved, i.e. $\int_{R_k^j} W_h^{s+1,*}d\mathbf{k} = \int_{R_k^j} W_h^{s+1}d\mathbf{k}$. In other words, $\left(W_h^{s+1,*},\eta\right)_k = \left(W_h^{s+1},\eta\right)_k$, for any piecewise constant test function, i.e. $\forall\eta\in\mathcal{W}_h^0$. Therefore, to preserve discrete conservation laws, in the definition of the discrete directional differential operator $\mathcal{L}_h$, we need to pick a projection $\Pi_{k,h}$ such that $\Pi_{k,h}U$ belongs to $\mathcal{W}_h^0\subset\mathcal{W}_h^{\alpha_2}$ for any function U.

3.7 Numerical results

3.7.1 Problem Setting

Although the emission/absorption kernel contains a summation from $l = -\infty$ to $l = +\infty$, it is not practical to perform that numerically. In practice, we keep the dominant part of those terms. In the following example, we will only consider one term with $l = 1$, associated with the anomalous Doppler resonance. We used the dispersion relation $\omega(\mathbf{k})$ of the whistler wave in cold magnetized plasma, with electron gyro-frequency $\omega_{ce} = -2\omega_{pe}$.

Set the cut-off computational domain as follows,

$$\Omega_p^L = \{(p_\parallel, p_\perp) : p_\parallel \in (-5mc, 25mc),\ p_\perp \in (0, 15mc)\},$$
$$\Omega_k^L = \{(k_\parallel, k_\perp) : k_\parallel \in (0.05\frac{\omega_{pe}}{c}, 0.65\frac{\omega_{pe}}{c}),\ k_\perp \in (0, 0.6\frac{\omega_{pe}}{c})\}.$$

Take piecewise linear quadrilateral basis $\mathcal{G}_h^1 = \{f(p_\parallel, p_\perp) \in C^0(\Omega_p^L) : f|_{R_p} \in Q^1(R_p), \forall\ R_p \in \mathcal{T}_h^p\}$ and piecewise constant basis $\mathcal{W}_h^0 = \{W(k_\parallel, k_\perp) : W|_{R_k} \in Q^0(R_k), \forall\ R_k \in \mathcal{T}_h^k\}$ as our test spaces. Choose the L^2 orthogonal projections $\Pi_{p,h}$ and $\Pi_{k,h}$ as stated in Corollary 3.5.

The numerical experiment is performed with 75×75 elements in Ω_p^L, and 40×40 elements in Ω_k^L. The initial time step size is set as $\Delta t = 1.0 \times 10^4 \frac{1}{2\pi\omega_{pe}}$.

The integration on resonance manifold is performed with Gauss-Legendre quadrature on Ω_p^L and the marching simplex method on Ω_k^L.

Consider the following initial conditions, which is the so-called 'bump on tail instability' configuration.

$$\begin{cases} f(p_\parallel, p_\perp)|_{t=0} = \left[10^{-5}\frac{1}{\sqrt{\pi}}\exp\left(-\left(\frac{p_\parallel}{mc} - 20\right)^2 - \left(\frac{p_\perp}{mc}\right)^2\right)\right]\frac{n_0}{m^3c^3}, \\ W(k_\parallel, k_\perp)|_{t=0} = 10^{-5}\frac{n_0 mc^2}{(\omega_{pe}/c)^3}. \end{cases}$$

Remark 3.9. The bump on tail configuration actually refers to the sum of a bulk and a bump, i.e. $f(\mathbf{p}, t) = f_c(\mathbf{p}) + f_b(\mathbf{p}, t)$, where the cold bulk $f_c(\mathbf{p}) \approx \frac{n_0}{m^3c^3}\delta(\mathbf{p})$, and the

bump $f_b(\mathbf{p}, t)$ is a peak with a much smaller population, centered far from the origin. However, as shown in the following equation,

$$\frac{\partial(f - f_c)}{\partial t} = \mathbb{B}(W, f) = \mathbb{B}(W, f - f_c) + \mathbb{B}(W, f_c) = \mathbb{B}(W, f - f_c),$$
$$\frac{\partial W}{\partial t} = \mathbb{H}(f, W) = \mathbb{H}(f - f_c, W) + \mathbb{H}(f_c, W) = \mathbb{H}(f - f_c, W),$$

we do not have to really compute the contribution from f_c.

3.7.2 Temporal evolution

In analogy to the analysis done by Kennel and Engelmann (1966), for a given wave vector $\mathbf{k}$, the characteristics associated with directional differential operator $\mathcal{L}$ is

$$z(p_\parallel, p_\perp) \coloneqq \frac{\omega}{k_\parallel} p_\parallel - \mathcal{E}(p_\parallel, p_\perp) = \frac{\omega}{k_\parallel} p_\parallel - \sqrt{m^2c^4 + p_\parallel^2 c^2 + p_\perp^2 c^2} = \text{const}, \tag{3.38}$$

which is the isoenergy contour in the reference frame moving at the wave's phase velocity.

When the wave *sed* $W(\mathbf{k}, t)$ is concentrated around the given $\mathbf{k}$, the contours as illustrated in Figure(3.3) indicates the principal diffusion direction. For the specific problem setting, $\frac{\omega}{k_\parallel}$ is small, hence the contour lines are almost concentric circles.

In Figure(3.2) we show the evolution of electron *pdf* $f(p_\parallel, p_\perp, t)$ and wave *sed* $W(k_\parallel, k_\perp, t)$. It can be observed that the bump on tail results in the excitation of the approximate waves in a narrow region of spectral space Ω_k^L, and as predicted by Equation(3.38), the whistler waves in turn cause anisotropic diffusion of electron *pdf* almost along the contour lines in Figure(3.3).

3.7.3 Verification of conservation

To verify the discrete conservation property of the proposed scheme, we define the relative error for conserved quantity as follows,

$$e_{rel}\left(\mathcal{Q}_{tot,h}\right) \coloneqq \frac{\|\mathcal{Q}_{tot,h} - \mathcal{Q}_{tot,h}^0\|_{L^\infty(0,T_{max})}}{\mathcal{Q}_{tot,h}^0},$$

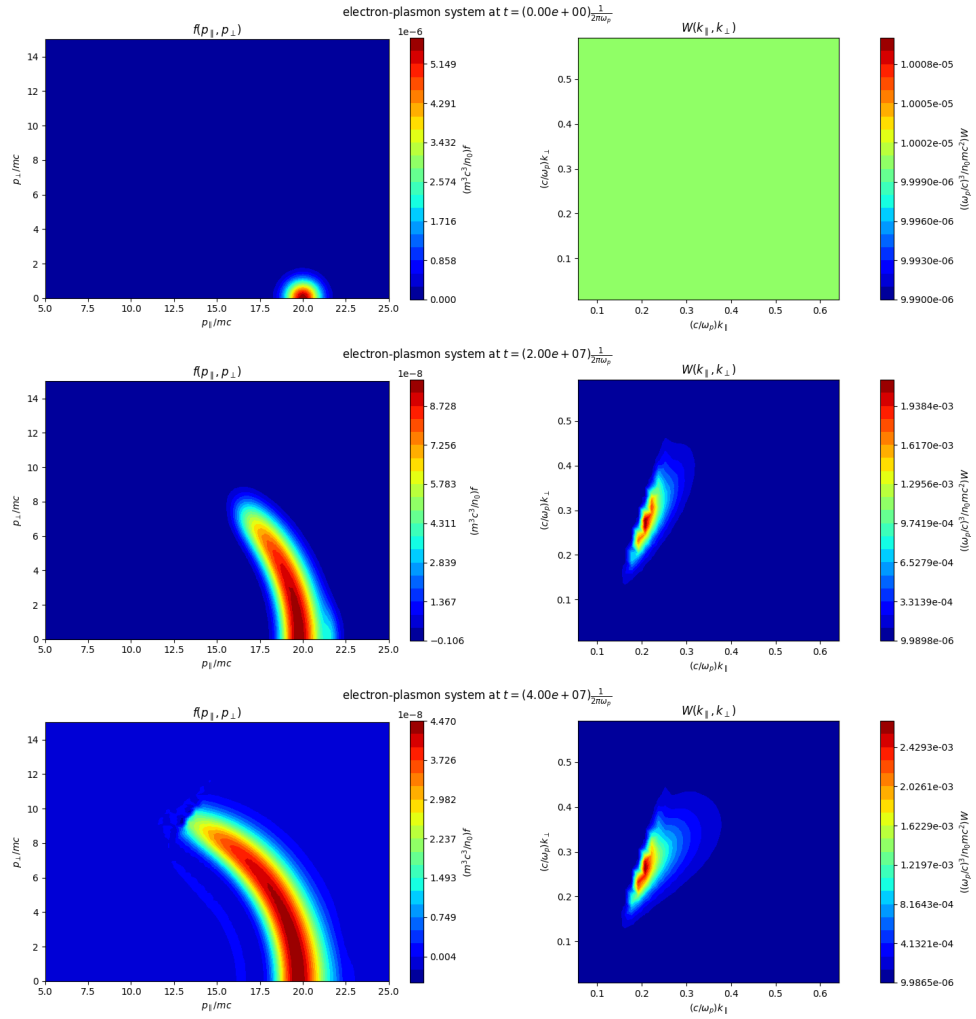


Figure 3.2: Temporal evolution of the electron *pdf* and wave *sed*.

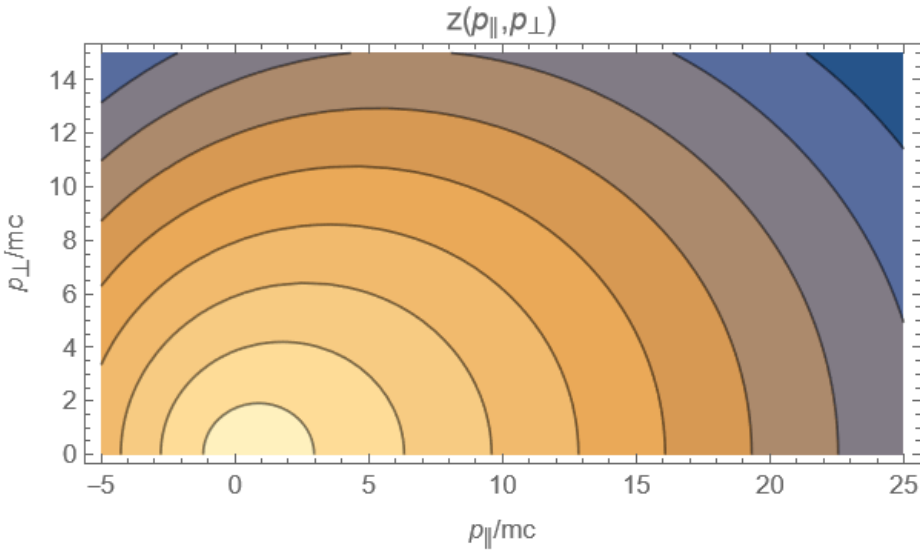


Figure 3.3: The characteristics of directional differential operator $\mathcal{L}$ given $(k_\parallel, k_\perp) = (0.2, 0.3)$.

where $\mathcal{Q}$ is the conserved quantities defined in Theorem 3.3.

Then with $T_{max} = 8.0 \times 10^7 \frac{1}{2\pi\omega_{pe}}$, we have

$$e_{rel}(\mathcal{M}_{tot,h}) = 4.96 \times 10^{-14},$$
$$e_{rel}(\mathcal{P}_{\parallel,tot,h}) = 4.58 \times 10^{-14},$$
$$e_{rel}(\mathcal{E}_{tot,h}) = 4.69 \times 10^{-14}.$$

For the evolution of the electron-plasmon system momentum and energy, see solid lines in Figure(3.4).

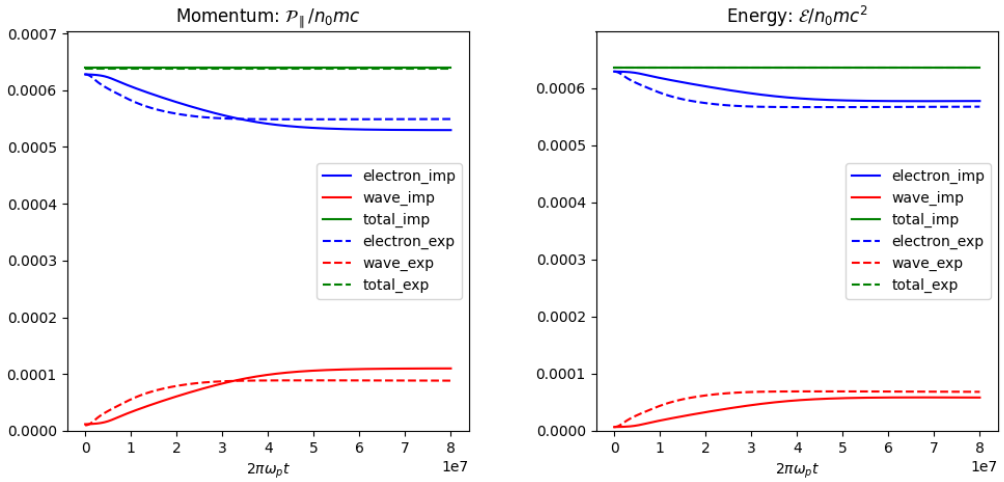
3.7.4 Comparison of different dispersion relations

The above results were obtained with the exact whistler wave dispersion relation $\omega_{\text{imp}}(\mathbf{k})$ for cold magnetized plasma, given implicitly by Equation(2.10). One might wonder what if we replace it with a simpler explicit approximate relation, for instance,

$$\omega_{\text{exp}}(\mathbf{k}) = |\omega_{ce}| \frac{|k_\parallel| k c^2}{\omega_p^2 \sqrt{1 + k^2 c^2 \omega_{ce}^2 / \omega_{pe}^4}},$$

which is asymptotic to the implicit relation $\omega_{\text{imp}}(\mathbf{k})$ when k is small, i.e.

$$\lim_{k\to 0}\left(\omega_{\text{exp}}(\mathbf{k})-\omega_{\text{imp}}(\mathbf{k})\right)=0.$$



(a) Momentum evolution. (b) Energy evolution.

Figure 3.4: Comparison between ω_{exact} and ω_{approx}

As shown in Figure(3.4), for both cases, energy and momentum are transferred from particles to waves. Meanwhile, we do observe a different transfer rate for the approximate whistler dispersion relation when compared to the exact implicit dispersion relation derived from Equation(2.10).

3.8 Generalization of the strategy: a conservative LDG scheme

For the electron *pdf*, induced emission and absorption process yields a pure diffusion equation, given by the first line of Equation(3.1):

$$\partial_t f(\mathbf{p},t)=\nabla_p\cdot\left(D[N](\mathbf{p},t)\cdot\nabla_p f(\mathbf{p},t)\right).$$

However, as shown in Equation(2.24), electron runaway is determined by multiple factors competing with each other. Both collision and external electric field causes

electron advection in momentum space $\mathbb{R}^3_p$. As is well-known, standard finite element scheme is not applicable for advection equations due to numerical instability. Therefore, the other terms have to be solved with stabilized methods, for example, upwind FD, DG and SUPG. If one uses different method for different terms, then accuracy will be affected due to operator splitting. Moreover, the projection back and forth between different function spaces takes more time. Among those methods, SUPG does not require projection onto different trial function spaces. However, its stabilized formulation interferes with conservation of total momentum.

In what follows, a conservative scheme based on the bilinear form of local DG method will be proposed. Since it is in the same framework as the conservative finite element scheme, this section will only focus on issues that are specific to DG.

3.8.1 Bilinear form of local DG method

Consider the following diffusion equation in bounded domain Ω,

$$\partial_t f = \nabla \cdot (D \cdot \nabla f), \tag{3.39}$$

with Neumann's boundary condition on $\partial\Omega$.

$$D \cdot \nabla f = \mathbf{0}.$$

Suppose the domain Ω is partitioned into elements $\{R\}$. The standard local discontinuous Galerkin method uses the following weak form,

$$\begin{aligned}
&\sum_R (\partial_t f, \varphi)_R = -\sum_R (Z, \nabla\varphi)_R + \sum_R \langle \widehat{Z}, \varphi^- \mathbf{n}^- \rangle_{\partial R/\Gamma}, \\
&\sum_R \left(Z, \widetilde{V}\right)_R = \sum_R \left(D \cdot \widetilde{Z}, \widetilde{V}\right)_R, \\
&\sum_R \left(\widetilde{Z}, V\right)_R = -\sum_R (f, \nabla \cdot V)_R + \sum_R \langle \widehat{f}, V^- \cdot \mathbf{n}^- \rangle_{\partial R}.
\end{aligned} \tag{3.40}$$

Choosing an appropriate reference vector $\mathbf{u}$, we use alternating flux as follows:

$$\widehat{Z} = \begin{cases} Z^-, & \mathbf{n}^- \cdot \mathbf{u} > 0 \\ Z^+, & \text{otherwise} \end{cases} \tag{3.41}$$

and

$$\widehat{f} = \begin{cases} f^-, & \mathbf{n}^- \cdot \mathbf{u} < 0 \text{ or } \partial R \subset \partial\Omega \\ f^+, & \text{otherwise} \end{cases} \tag{3.42}$$

Definition 3.1 (discrete gradient operator(Di Pietro and Ern, 2011))**.** We define $(\nabla\varphi)_u$ as the function in $\mathcal{V}$ such that,

$$-\sum_R \left((\nabla\varphi)_u, V\right)_R = -\sum_R (\nabla\varphi, V)_R + \sum_R \langle\varphi^-\mathbf{n}^-, \widehat{V}\rangle_{\partial R/\Gamma}, \forall V \in \mathcal{V}. \tag{3.43}$$

Analogously, define $(\nabla\varphi)_d$ as the function in $\mathcal{V}$ such that

$$\sum_R \left((\nabla\varphi)_d, V\right)_R = -\sum_R (\varphi, \nabla\cdot V)_R + \sum_R \langle\widehat{\varphi}, V^-\cdot\mathbf{n}^-\rangle_{\partial R}, \forall V \in \mathcal{V}. \tag{3.44}$$

Proposition 3.9. *If the flux terms are as defined in Equation(3.41) and Equation(3.42), then*

$$(\nabla\varphi)_u = (\nabla\varphi)_d. \tag{3.45}$$

Proof. Integrate by parts on the right hand side of Equation(3.44), we have

$$-\sum_R (\varphi, \nabla\cdot V)_R + \sum_R \langle\widehat{\varphi}, V^-\cdot\mathbf{n}^-\rangle_{\partial R} = \sum_R (\nabla\varphi, V)_R + \sum_R \langle\widehat{\varphi}-\varphi^-, V^-\cdot\mathbf{n}^-\rangle_{\partial R/\Gamma}.$$

Summing Equation(3.43) and Equation(3.44) to obtain,

$$\sum_R \left((\nabla\varphi)_d - (\nabla\varphi)_u, V\right)_R = \sum_R \langle\widehat{\varphi}-\varphi^-, V^-\cdot\mathbf{n}^-\rangle_{\partial R/\Gamma} + \sum_R \langle\varphi^-\mathbf{n}^-, \widehat{V}\rangle_{\partial R/\Gamma}. \tag{3.46}$$

Since

$$\sum_R \langle\widehat{\varphi}, V^-\cdot\mathbf{n}^-\rangle_{\partial R/\Gamma} + \sum_R \langle\varphi^-\mathbf{n}^-, \widehat{V}\rangle_{\partial R/\Gamma} = \sum_R \langle\varphi^-, V^-\cdot\mathbf{n}^-\rangle_{\partial R/\Gamma},$$

the right hand side of Equation(3.46) must be zero.

□

As have been discussed in Arnold et al. (2002), the LDG weak form can also be written in bilinear form.

Theorem 3.10. *The weak form (3.40) is equivalent to*

$$(\partial_t f, \varphi)_\Omega + ((\nabla f)_d \,, D \cdot (\nabla \varphi)_d)_\Omega = 0.$$

Proof. By Definition 3.1, the weak form (3.40) is equivalent to

$$\begin{aligned} \sum_R (\partial_t f, \varphi)_R &= -\sum_R ((\nabla \varphi)_u \,, Z)_R \,, \\ \sum_R \left(Z, \widetilde{V}\right)_R &= \sum_R \left(D \cdot \widetilde{Z}, \widetilde{V}\right)_R \,, \\ \widetilde{Z} &= (\nabla f)_d \,. \end{aligned}$$

Therefore, the corresponding bilinear form can be derived as follows,

$$\begin{aligned} \sum_R (\partial_t f, \varphi)_R &= -\sum_R ((\nabla \varphi)_u \,, Z)_R \\ &= -\sum_R \left(D \cdot \widetilde{Z}, (\nabla \varphi)_u\right)_R \\ &= -\sum_R ((\nabla f)_d \,, D \cdot (\nabla \varphi)_u)_R \\ &= -\sum_R ((\nabla f)_d \,, D \cdot (\nabla \varphi)_d)_R \,, \end{aligned}$$

where Proposition 3.9 is used in the last row. □

3.8.2 Conservative DG discretization

The cut-off domains and boundary conditions have been discussed in Section 3.4, thus that is not repeated here.

As opposed to the standard finite element method, the new test space for electron *pdf* consists of discontinuous piecewise polynomials as follows:

$$\mathcal{G}_h^{\alpha_1} = \{f(p_\parallel, p_\perp) : f|_{R_p} \in Q^{\alpha_1}(R_p), \forall\ R_p \in \mathcal{T}_h^p\},\ \alpha_1 = 0, 1, \cdots .$$

Analogous to Section 3.4, the projection $\Pi_{p,h}$ onto test space $\mathcal{G}_h$ must satisfy that

$$\lim_{h \to 0} \|\Pi_{p,h} g(\mathbf{p}) - g(\mathbf{p})\|_{L^2(\Omega_p^L)} = 0,\ \forall g \in L^2(\Omega_p^L),$$

and

$$\lim_{h\to 0} \left\| \nabla_{p,h} \left(\Pi_{p,h} \mathcal{E}(\mathbf{p}) \right) - \nabla_p \mathcal{E}(\mathbf{p}) \right\|_{L^2(\Omega_p^L)} = 0,$$

where the discrete gradient operator $\nabla_{p,h}$ is defined in Definition 3.1.

Combining the above projection operator and a projection $\Pi_{k,h}$ onto test space $\mathcal{W}_h$ satisfying Equation(3.26), we propose the following discretized operator:

$$\mathcal{L}_h g_h \coloneqq \frac{1}{\omega} \left[k_{\parallel,h} \left(\partial_{\perp,h} \mathcal{E}_h \right) \cdot \frac{p}{p_\perp} \left(\partial_{\parallel,h} g_h \right) + \left(\omega_h - k_{\parallel,h} \left(\partial_{\parallel,h} \mathcal{E}_h \right) \right) \frac{p}{p_\perp} \left(\partial_{\perp,h} g_h \right) \right], \quad (3.47)$$

where

$$\begin{aligned} k_{\parallel,h} &\coloneqq \Pi_{k,h} k_\parallel, \\ \omega_h &\coloneqq \Pi_{k,h} \omega, \\ \mathcal{E}_h &\coloneqq \Pi_{p,h} \mathcal{E}, \\ \partial_{\parallel,h} g_h &\coloneqq \frac{\mathbf{B}}{|\mathbf{B}|} \cdot \nabla_{p,h} g_h, \\ \partial_{\perp,h} g_h &\coloneqq \left| \left(I - \frac{\mathbf{B}}{|\mathbf{B}|} \otimes \frac{\mathbf{B}}{|\mathbf{B}|} \right) \cdot \nabla_{p,h} g_h \right|. \end{aligned}$$

The operator $\mathcal{L}_h$ also guarantees unconditional conservation, as proved below. This theorem is the DG version of Theorem 3.4.

Theorem 3.11. *If $f_h(t) \in \mathcal{G}_h$ and $N_h(t) \in \mathcal{W}_h$ satisfies that*

$$\int_p \varphi_h \partial_t f_h + \int_k \eta_h \partial_t N_h = \iint_{pk} N_h \mathcal{B}_\varepsilon \left(\eta_h \mathcal{L}_h \mathcal{E}_h - \hbar \omega_h \mathcal{L}_h \varphi_h \right) \mathcal{L}_h f_h, \quad (3.48)$$

for any $\varphi_h(t) \in \mathcal{G}_h$ and $\eta_h(t) \in \mathcal{W}_h$, with $\mathcal{L}_h$ defined by Equation(3.47), then

$$\begin{aligned} \partial_t \mathcal{M}_{tot}^h &\coloneqq \partial_t \left((f_h, 1)_p + (N_h, 0)_k \right) = 0, \\ \partial_t \mathcal{P}_{\parallel,tot}^h &\coloneqq \partial_t \left((f_h, \Pi_{p,h} p_\parallel)_p + (N_h, \hbar \Pi_{k,h} k_\parallel)_k \right) = \hat{0}, \\ \partial_t \mathcal{E}_{tot}^h &\coloneqq \partial_t \left((f_h, \Pi_{p,h} \mathcal{E})_p + (N_h, \hbar \Pi_{k,h} \omega)_k \right) = 0. \end{aligned}$$

Remark 3.10. By $\hat{0}$ we mean zero or a value below machine epsilon when the cut-off domain is large enough, depending on the specific projection operator $\Pi_{p,h}$, the details will be given in the following proof.

Proof. Mass conservation and energy conservation can be easily verified. For momentum conservation, note that

$$\begin{aligned}
&\left(\hbar\Pi_{k,h}k_{\|}\right)\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\left(\Pi_{p,h}p_{\|}\right)\\
=&\left(\hbar\Pi_{k,h}k_{\|}\right)\frac{\omega_h}{\omega}\frac{p}{p_\perp}\left(\partial_{\perp,h}\mathcal{E}_h\right)-\frac{\hbar\omega_h}{\omega}k_{\|,h}\left(\partial_{\perp,h}\mathcal{E}_h\right)\frac{p}{p_\perp}\left(\partial_{\|,h}p_{\|,h}\right)\\
=&\hbar\frac{\omega_h}{\omega}k_{\|,h}\left(\partial_{\perp,h}\mathcal{E}_h\right)\frac{p}{p_\perp}\left[1-\left(\partial_{\|,h}p_{\|,h}\right)\right].
\end{aligned}$$

When $\alpha_1\geq 1$ and $\Pi_{p,h}p_{\|}=p_{\|}$, the error term $1-\left(\partial_{\|,h}p_{\|,h}\right)$ is equal to zero on any element.

When $\alpha_1=0$, i.e. with piecewise constant basis function, we have $1-\left(\partial_{\|,h}p_{\|,h}\right)=0$, except on some elements near boundary $\partial\Omega_p^L$. That can be observed from the explicit formula of discrete gradient operator in Equation(5.9). Since $\nabla_{p,h}f_h$ on the boundary can be arbitrarily small for large enough cut-off domain, we conclude that

$$\iint_{pk}N_h\mathcal{B}_\varepsilon\hbar\frac{\omega_h}{\omega}k_{\|,h}\left(\partial_{\perp,h}\mathcal{E}_h\right)\frac{p}{p_\perp}\left[1-\left(\partial_{\|,h}p_{\|,h}\right)\right]\mathcal{L}_hf_h=\hat{0}.$$

□

3.8.3 Time discretization

There are four types of first order conservative time discretizations,

$$\begin{aligned}
&\int_p\varphi_h\frac{f_h^{s+1}-f_h^s}{\Delta t}+\int_k\eta_h\frac{N_h^{s+1}-N_h^s}{\Delta t}\\
=&\begin{cases}
\iint_{pk}N_h^s\mathcal{B}_\varepsilon\left(\eta_h\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\varphi_h\right)\mathcal{L}_hf_h^s, & \text{fully-explicit}\\
\iint_{pk}N_h^s\mathcal{B}_\varepsilon\left(\eta_h\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\varphi_h\right)\mathcal{L}_hf_h^{s+1}, & \text{semi-implicit}\\
\iint_{pk}N_h^{s+1}\mathcal{B}_\varepsilon\left(\eta_h\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\varphi_h\right)\mathcal{L}_hf_h^s, & \text{semi-implicit}\\
\iint_{pk}N_h^{s+1}\mathcal{B}_\varepsilon\left(\eta_h\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\varphi_h\right)\mathcal{L}_hf_h^{s+1}, & \text{fully-implicit}
\end{cases}
\end{aligned}$$

The first and third row are explicit for particle *pdf* $f_h(t)$, therefore the time stepsize Δt has to satisfy the CFL condition for diffusion equations in order to preserve the L^2-stability. The second and the fourth row are implicit for particle *pdf* $f_h(t)$, therefore

unconditionally L^2-stable. Nevertheless, all of them relies on the positivity of plasmon *pdf* N_h.

In Theorem 3.8 it has been proved that N_h^s remains positive as long as the time stepsize Δt is small enough. The only difference here is that $\nabla_p f_h$ has been replaced with the discrete gradient $\nabla_{p,h} f_h$. To complete the proof, one need a bound similar to Equation(3.35) in the proof of Theorem 3.8:

$$\left\|\nabla_{p,h} f_h^{s+1}\right\|_{L^2(\Omega_p^L)} \leq \frac{C_3}{h_p} \left\|f_h^{s+1}\right\|_{L^2(\Omega_p^L)}.$$

Obviously, one can no longer resort to the inverse inequality, since $\nabla_{p,h} \neq \nabla_p$. But it is indeed provable, the detailed proof is provided in Appendix A.

3.8.4 Implementation and numerical results

Details on implementation and the numerical results of the DG method will be provided in Chapter 5, as a part of a larger problem.

Chapter 4: A Galerkin approach for trajectorial average in maximally superintegrable systems

4.1 Introduction

The method proposed in this chapter was initially designed for solving the second line of Equation(2.23):

$$\partial_t N + \{N, \omega\} = \Gamma[f]N.$$

But it turns out to be applicable for more general problems. Therefore we discuss the general problems first and elaborate on specific ones in the next chapter.

Recall the Liouville equation(2.5) for Nambu systems:

$$\partial_t \rho(q_1, \cdots, q_n, t) = \{\rho, H_1, H_2, \cdots, H_{n-1}\}.$$

In what follows we denote the Nambu bracket as advection operator $\mathcal{T}$. Suppose a small perturbation is applied to the Liouville equation, for example, a reaction term as follows,

$$\partial_t \rho + \frac{1}{\varepsilon}\mathcal{T}\rho = \gamma\rho. \tag{4.1}$$

Note that we use a dimensionless parameter $\varepsilon \ll 1$ to indicate the scale.

In practice, the time scale of interest is often much greater than τ_{adv}, the characteristic time scale of advection. Therefore, averaging over the flow is necessary. For Equation(4.1) where operator $\mathcal{T}$ represents a divergence-free advection field, there is already a standard tool for multiscale analysis: the Hilbert expansion. Assume that ρ has the following Hilbert expansion,

$$\rho = \rho_0 + \varepsilon\rho_1 + \varepsilon^2\rho_2 + \cdots. \tag{4.2}$$

Substitute Equation(4.2) into Equation(4.1) and collect the terms according to the power of ε. The ε^{-1} term being zero yields that,

$$\mathcal{T}\rho_0 = 0. \tag{4.3}$$

And the ε^0 term being zero yields that,

$$\partial_t \rho_0 + \mathcal{T}\rho_1 = \gamma\rho_0. \tag{4.4}$$

Suppose $\mathcal{P}$ is the orthogonal projection operator onto $\ker\mathcal{T}$. Apply it on both sides of the above equation to obtain

$$\partial_t \rho_0 = \mathcal{P}\left(\gamma\rho_0\right). \tag{4.5}$$

The first term on the left hand side of Equation(4.4) remains the same because $\rho_0 \in \ker\mathcal{T}$ as shown in Equation(4.3). The second term on the left hand side vanishes because operator $\mathcal{T}$ is skew-symmetric. The right hand side $\mathcal{P}\left(\gamma\rho_0\right)$ is called the average of function $\gamma\rho_0$ over the flow, as it can be verified that on any advection field line, $\mathcal{P}g$ is always equal to the average of function g along this line. For details, we refer the readers to Bostan (2010).

The above technique has been successful in tackling problems with simple advection field lines. For example, the gyro-average of Vlasov equations, where the average operator is simply integration along circles. For more complicated problems, there is the ray-tracing method(Aleynikov and Breizman, 2015), where sample trajectories are calculated with the Hamilton's equation(2.1), and numerical averaging is performed on these trajectories. However, such approach is expensive, not conservative, and not compatible with finite element or finite difference solvers for kinetic equations.

In this chapter we propose a new method from a completely different perspective: instead of approximating the operator $\mathcal{P}$, we focus on the discretization of function spaces.

The rest of this chapter is organized as follows. Section 4.2 concerns the weak form of averaged equation and its discretization. In Section 4.3 we introduce the connection-proportion algorithm to construct discrete test/trial spaces. Finally in Section 4.4 the property of the proposed Galerkin approach is discussed. Numerical examples will be provided in Chapter 5, as a part of a larger problem.

4.2 Weak form and its discretization

Existing methods are all about solving the strong form of Equation(4.5) straightforwardly. Indeed, it does have a corresponding weak form, although the test/trial space is special.

Define function space $\mathcal{U} \coloneqq L^2(\mathbb{R}^n) \cap \ker\mathcal{T}$ and the orthogonal projection operator $\mathcal{P} : L^2(\mathbb{R}^n) \to \mathcal{U}$. Test Equation(4.5) with $\eta \in \mathcal{U}$, it becomes

$$(\partial_t \rho_0, \eta) = (\mathcal{P}(\gamma\rho_0), \eta).$$

Note that the orthogonal projection operator $\mathcal{P}$ is symmetric, therefore

$$(\mathcal{P}(\gamma\rho_0), \eta) = (\gamma\rho_0, \mathcal{P}\eta) = (\gamma\rho_0, \eta),$$

where the second identity comes from the fact that $\eta \in \mathcal{U}$.

To conclude, the corresponding weak formulation of Equation(4.5) is as follows: find a $\rho_0 \in \mathcal{U}$ such that for any $\eta \in \mathcal{U}$,

$$(\partial_t \rho_0, \eta) = (\gamma\rho_0, \eta). \tag{4.6}$$

Note that the trajectorial averaging operator $\mathcal{P}$ does not even appear in the weak form, hence there is no need to approximate it with some discrete operator $\mathcal{P}_h$. Instead, we follow the Galerkin approach and seek for a series of subspaces $\mathcal{U}_h$ dense in the test/trial space $\mathcal{U}$.

Constructing a finite dimensional space $\mathcal{U}_h$ means constructing all the basis functions that span it. In order to do that, we first introduce the concept of trajectory bundles.

Consider a particle moving in phase space, denote the trajectory generated by initial state $\mathbf{q}_0$ as $S(\mathbf{q}_0)$:

$$S(\mathbf{q}_0) \coloneqq \{\mathbf{q} \in \mathbb{R}^n : \mathbf{q} = \mathbf{q}(t; \mathbf{q}_0) \text{ for some } t \in \mathbb{R}\},$$

where $\mathbf{q}(t;\mathbf{q}_0)$ represents the solution of generalized Hamilton's equation(2.3) with initial state $\mathbf{q}_0$. By the property of Nambu system introduced in Section 2.1, the generalized Hamiltonians must be invariant along the trajectory:

$$H_j(\mathbf{q}) = H_j(\mathbf{q}_0), \forall \mathbf{q} \in S(\mathbf{q}_0), j = 1, \cdots, n-1.$$

Define the trajectory bundle generated by connected open set $\mathcal{A}$ as $S(\mathcal{A}) = \cup_{\mathbf{q}_0 \in \mathcal{A}} S(\mathbf{q}_0)$. The set $S(\mathcal{A})$ is the union of all the trajectories generated by initial states inside $\mathcal{A}$, therefore they share the same range of Hamiltonians,

$$H_j(S(\mathcal{A})) = H_j(\mathcal{A}), j = 1, \cdots, n-1.$$

We present the above definition just to give the readers some physics intuition. The following approach is more convenient for numerical implementation. Without loss of generality, we consider a finite domain $\Omega_b \subset \mathbb{R}^n$ such that

$$H_1(\mathbf{q}) = \text{Const}, \ \forall \mathbf{q} \in \partial\Omega_b.$$

Such condition ensures no flux across the boundary.

Definition 4.1 (Trajectory Bundles)**.** Given a nice box in the $\mathbb{R}^{n-1}$ space:

$$I = \left(H_1^l, H_1^r\right) \times \left(H_2^l, H_2^r\right) \times \cdots \times \left(H_{n-1}^l, H_{n-1}^r\right) \subset \mathbb{R}^{n-1},$$

there exists a family $\mathcal{F}$ of subsets S of Ω_b with finite cardinality such that

1. The union of these subsets is the inverse image of the box I,

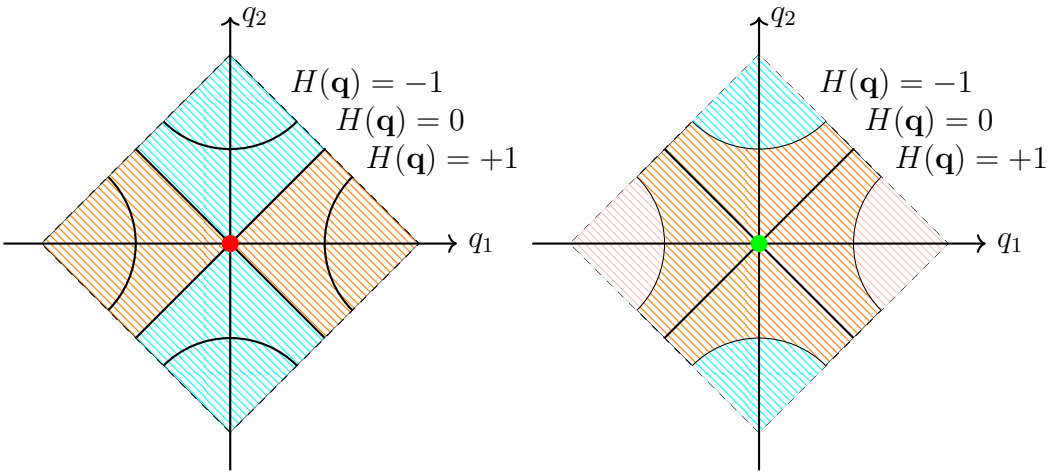
$$\cup_{S \in \mathcal{F}} S = \overrightarrow{H}^{-1}(I) \coloneqq \left\{\mathbf{q} \in \Omega_b : (H_1(\mathbf{q}), \cdots, H_{n-1}(\mathbf{q})) \in I\right\}. \tag{4.7}$$

2. Any subset $S \in \mathcal{F}$ is connected.

3. If $S_1 \in \mathcal{F}$, $S_2 \in \mathcal{F}$ and $S_1 \neq S_2$, then $\overline{S_1} \cap \overline{S_2} = \emptyset$.

Then each subset $S \in \mathcal{F}$ is called a trajectory bundle. And $\mathcal{F}$ is called the family of trajectory bundles generated by the interval (H_a, H_b).

Remark 4.1. By "nice box" we mean a box I such that gradient vectors $\{\nabla H_j(\mathbf{q})\}$ are well defined and linearly independent for any $\mathbf{q} \in \partial\overrightarrow{H}^{-1}(I)$ in Equation(4.7). This condition enables us to avoid the discussion of singular scenarios such as saddle points, shown in Figure 4.1



(a) Bad box (b) Nice box

Figure 4.1: A nice box has no saddle point on the boundary of its inverse image.

The basis function induced by a trajectory bundle S is defined as

$$1_S(\mathbf{q}) \coloneqq \begin{cases} 1, & \mathbf{q} \in S \\ 0, & \text{otherwise} \end{cases}$$

By Theorem 2.1, we have

$$\{1_S, H_1, \cdots, H_{n-1}\} = 0.$$

In other words, $1_S \in \mathcal{U} \coloneqq \ker\mathcal{T} \cap L^2(\Omega_b)$.

Given a rectangular segmentation $\{I_\alpha\}$ of the space $\mathbb{R}^{n-1}$, each box I_α generates a family of trajectory bundles $\mathcal{F}_\alpha$. Denote the basis functions induced by all these trajectory bundles as η_i, $i = 1, \cdots, M$. They have the following properties:

1. Partition of unity,

$$\cup_{i=1}^{M} \eta_i(\mathbf{q}) = 1, \ \forall \mathbf{q} \in \Omega_b. \tag{4.8}$$

2. Orthogonality

$$(\eta_i, \eta_j) = \delta_{ij} (\eta_i, \eta_i),$$

where δ_{ij} is the Kronecker delta.

It can be verified that $\mathcal{U}_h \coloneqq \operatorname{span}\{\eta_i\}$ is dense in $\mathcal{U} \coloneqq L^2(\Omega_b) \cap \ker\mathcal{T}$. Therefore, the discrete weak form can be stated as follows: find a $\rho_h(\cdot, t) \in \mathcal{U}_h$ such that for any $\eta_h \in \mathcal{U}_h$,

$$(\partial_t \rho_h, \eta_h) = (\gamma \rho_h, \eta_h). \tag{4.9}$$

Suppose that $\rho_h(\mathbf{q}, t) = \sum_i \alpha_i(t)\eta_i(\mathbf{q})$. Define two matrices: $M_{ij} = (\eta_i, \eta_j)$ and $\Gamma_{ij} = (\eta_i, \gamma\eta_j)$. Then $\{\alpha_i(t)\}$ satisfy the following ODE:

$$M_{ij}\partial_t \alpha_i = \Gamma_{ij}\alpha_i. \tag{4.10}$$

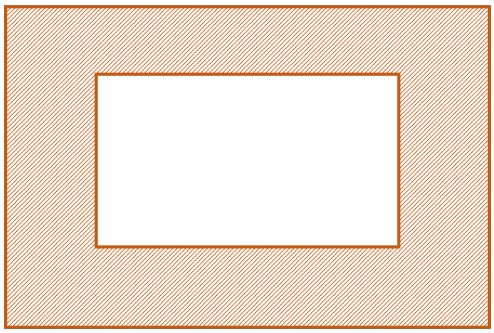
Now it remains to construct all the basis functions and evaluate those two matrices numerically, which is going to be discussed in the next section.

4.3 The connection-proportion algorithm

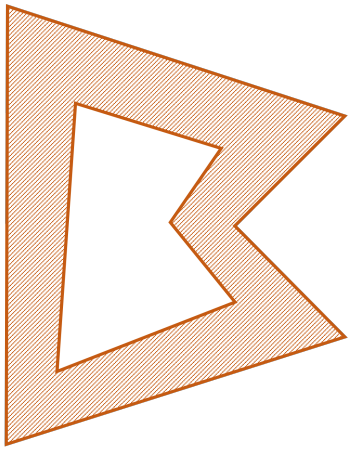
In what follows, we assume that the domain $\Omega_b \subset \mathbb{R}^n$ is partitioned into simplex meshes $\{V_i\}$, and the generalized Hamiltonians $\{H_j(\mathbf{q})\}$ are continuous and piecewise linear on the mesh. Under such assumption, the trajectory bundles are areas between polytopes: the level sets of generalized Hamiltonians. The challenge here is twofold:

1. How to distinguish disconnected trajectory bundles in the same family?

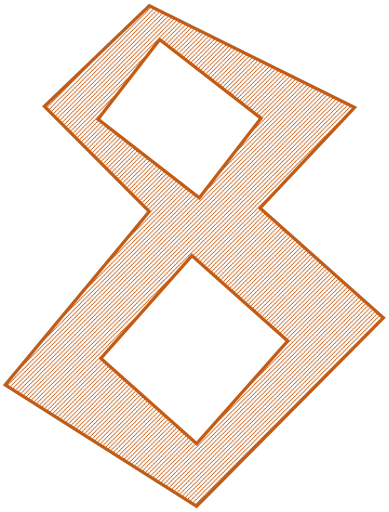
2. How to store all the necessary information associated to the trajectory bundles?



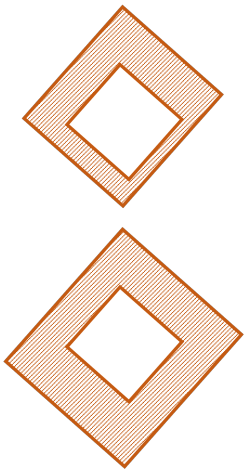
(a) Region between convex polygons



(b) Region between non-convex polygons



(c) Complicated topology



(d) Disconnected regions.

Figure 4.2: Storing the nodes of polygons does not solve the problem. For example, it is hard to tell the inside/outside of a non-convex polygon, needless to say about more complicated topologies.

Taking the two-dimensional problem as an example: as shown in Figure 4.2, storing all the nodes of polygons does not solve the problem.

We propose the connection-proportion algorithm to construct the family $\mathcal{F}$ of trajectory bundles generated by a box I. The key is to turn a continuous topological problem into a discrete graph theory problem.

To begin with, let us introduce some preliminary concepts.

Definition 4.2 (minimal simplex cover)**.** A simplex cover $\mathcal{R}$ of trajectory bundle S is a set of simplex meshes such that

$$S \subset \cup_{V_i \in \mathcal{R}} \overline{V_i}.$$

Now define

$$\mathcal{R}_{min}(S) \coloneqq \{V_i : V_i \cap S \neq \emptyset\}.$$

It can be easily verified that for any simplex cover $\mathcal{R}$ of S we have $\mathcal{R}_{min}(S) \subset \mathcal{R}$. Hence we call $\mathcal{R}_{min}(S)$ the minimal simplex cover of S.

As shown in the following proposition, the minimal simplex cover replicates the topological relation between the corresponding trajectory bundles.

Proposition 4.1. *If $S_i, S_j \in \mathcal{F}$ and $S_i \neq S_j$, then*

$$\mathcal{R}_{min}(S_i) \cap \mathcal{R}_{min}(S_j) = \emptyset.$$

Corollary 4.2. *If a point $\mathbf{q}$ belongs to $\cup_{V_i \in \mathcal{R}_{min}(S)} \overline{V_i}$ and $(H_1(\mathbf{q}), H_2(\mathbf{q}), \cdots, H_{n-1}(\mathbf{q})) \in I$, then*

$$\mathbf{q} \in S.$$

Corollary 4.2 actually points out the data structure to be used in storing the necessary information of a basis function. As shown in Algorithm 1, once the minimal simplex cover $\mathcal{R}_{min}(S)$ is known, it only takes two steps to evaluate the basis function 1_S. Note that the evaluation process has nothing to do with the nodes of polygons.

Algorithm 1 Evaluation of basis function 1_S.

input: coordinates $\mathbf{q}$, minimal simplex cover $\mathcal{R}_{min}(S)$, box $I(S)$
calculate $\overrightarrow{H}(\mathbf{q})$
find the simplex V s.t. $\mathbf{q} \in V$
if $\overrightarrow{H}(\mathbf{q}) \in I(S)$ **then**
  if $V \in \mathcal{R}_{min}(S)$ **then**
    output 1
  else
    output 0
  end if
  output 0
end if

Next, we propose the algorithm to construct the minimal simplex covers. As mentioned before, the goal is to obtain a graph theory problem. Therefore we introduce the following connection matrix.

Definition 4.3 (connection matrix)**.** For a given box $I \subset \mathbb{R}^{n-1}$, the connection matrix associated with I is defined as

$$A_{ij}(I) = \begin{cases} 1, & \overline{V_i} \cap \overline{V_j} \cap \overrightarrow{H}^{-1}(I) \neq \emptyset \\ 0, & \text{otherwise,} \end{cases} \tag{4.11}$$

where the inverse image $\overrightarrow{H}^{-1}(I)$ is defined in Equation(4.7).

Remark 4.2. The connection matrix is easy to evaluate because

$$\overline{V_i} \cap \overline{V_j} \cap \overrightarrow{H}^{-1}(I) \neq \emptyset \Leftrightarrow \overrightarrow{H}\left(\overline{V_i} \cap \overline{V_j}\right) \cap I \neq \emptyset.$$

The set $\overrightarrow{H}\left(\overline{V_i} \cap \overline{V_j}\right) \subset \mathbb{R}^{n-1}$ can be easily obtained due to the following fact: Linear functions only attain their extrema on the boundary of convex sets, in this case, the vertices.

The connection matrix $A_{ij}(I)$ renders a finite family of connected components $\{\mathcal{R}_\nu\}$. The following theorem shows that each of them corresponds to a trajectory bundle.

Theorem 4.3. *Suppose that $\mathcal{F}(I)$ is the family of trajectory bundles generated by box I and $\{\mathcal{R}_\nu\}$ is the set of connected components determined by the connection matrix $A_{ij}(I)$ as defined in Equation(4.11), then*

$$\{\mathcal{R}_\nu\} = \{\mathcal{R}_{min}(S) : S \in \mathcal{F}(I)\}.$$

Proof. It takes two steps to prove the theorem.

1. Prove that $\{\mathcal{R}_\nu\} \subset \{\mathcal{R}_{min}(S) : S \in \mathcal{F}(I)\}$.

   By definition, any connected component $\mathcal{R}_\nu$ contains at least two elements, say V_m and V_n.

   Since $H^{-1}(I)$ is open, $\overline{V_m} \cap \overline{V_n} \cap H^{-1}(I) \neq \emptyset$ implies that

   $$V_m \cap H^{-1}(I) \neq \emptyset.$$

   Considering that

   $$V_m \cap H^{-1}(I) = V_m \cap (\cup_{\mathcal{F}} S) = \cup_{\mathcal{F}} (V_m \cap S),$$

   there must exist a trajectory bundle $S_0 \in \mathcal{F}$ such that

   $$V_m \cap S_0 \neq \emptyset.$$

   It remains to show that $\mathcal{R}_\nu = \mathcal{R}_{min}(S_0)$.

   (a) Prove that $\mathcal{R}_\nu \subset \mathcal{R}_{min}(S_0)$.

   Recall that $V_m \cap S_0 \neq \emptyset$, then for any $V_l \in \mathcal{R}_\nu$, there must be a continuous path inside $H^{-1}(I)$ from $\mathbf{q}_1 \in V_l \cap H^{-1}(I)$ to $\mathbf{q}_2 \in V_m \cap H^{-1}(I)$. Since $\mathbf{q}_2 \in S_0$ implies that $\mathbf{q}_1 \in S_0$, it follows that $V_l \cap S_0 \neq \emptyset$. Therefore $\mathcal{R}_\nu \subset \mathcal{R}_{min}(S_0)$.

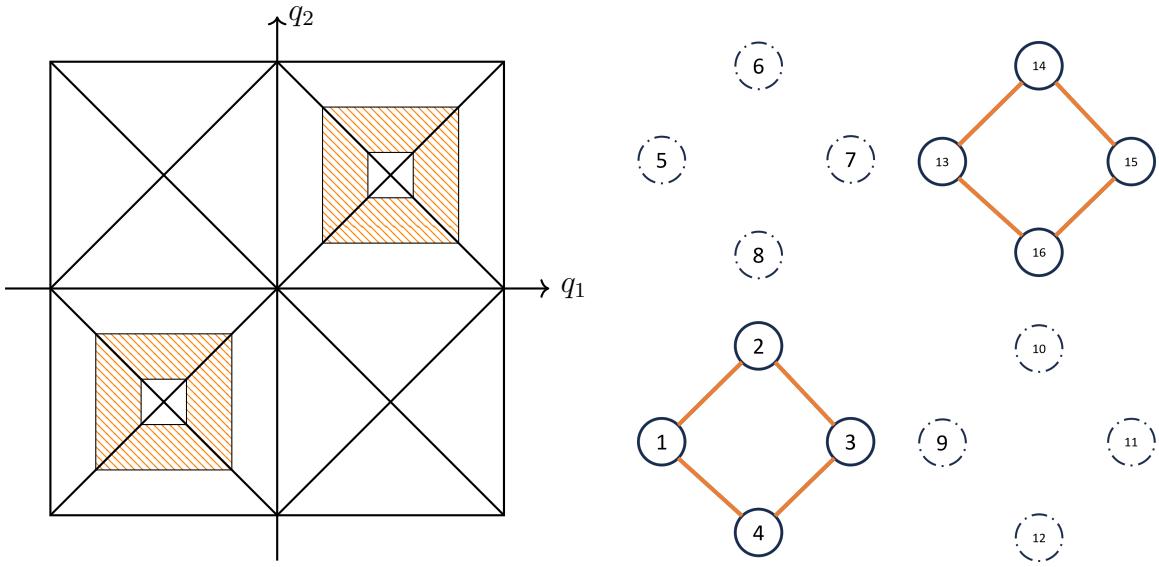


Figure 4.3: The minimal triangle covers replicate the topological relation between the corresponding trajectory bundles.

(b) Prove that $\mathcal{R}_\nu \supset \mathcal{R}_{min}(S_0)$.

Consider a simplex $V_l \in \mathcal{R}_{min}(S_0)$, by definition $V_l \cap S_0 \neq \emptyset$. Suppose that $V_l \notin \mathcal{R}_\nu$, then $\mathbf{q}_1 \in V_l \cap S_0$ is not connected to any point $\mathbf{q}_2 \in S_0 \cap \cup_{i \in \mathcal{R}_\nu} \overline{V_i}$, hence S_0 is not a connected set, which is contradictory to its definition. Therefore, V_l must belong to $\mathcal{R}_\nu$.

2. Prove that $\{\mathcal{R}_\nu\} \supset \{\mathcal{R}_{min}(S) : S \in \mathcal{F}(I)\}$.

By definition $\cup_\nu \mathcal{R}_\nu = \cup_{\mathcal{F}} \mathcal{R}_{min}(S)$, hence for any S_0 there exists an $\mathcal{R}_\nu$ such that $\mathcal{R}_\nu \cap \mathcal{R}_{min}(S_0) \neq \emptyset$. In the same way as above, it can be proved that $\mathcal{R}_\nu = \mathcal{R}_{min}(S_0)$.

□

In principle, once the minimal simplex covers are constructed, any integral can be performed by quadrature. However, the mass matrix $M_{ij} = (\eta_i, \eta_j)$ plays the key role here, so we expect to know the precise values. That is possible thanks to the fact

that

$$(\eta_i, \eta_j) = \delta_{ij} (\eta_i, \eta_i) = \delta_{ij} (\eta_i, 1) = \delta_{ij} \cdot \mu (S_i),$$

where $S_i = \operatorname{supp}(\eta_i)$ is a trajectory bundle and $\mu(S_i)$ represents its measure.

Theorem 4.4. *For a n-dimensional simplex V with vertices $\mathbf{q}_i,\ i = 1, \cdots, n+1$, the proportion $\frac{\mu(V \cap \overrightarrow{H}^{-1}(I))}{\mu(V)}$ only depends on the box $I = \left(H_1^l, H_1^r\right) \times \left(H_2^l, H_2^r\right) \times \cdots \times \left(H_{n-1}^l, H_{n-1}^r\right)$ and the node values $\overrightarrow{H}(\mathbf{q}_i) = (H_1(\mathbf{q}_i), H_2(\mathbf{q}_i), \cdots, H_{n-1}(\mathbf{q}_i)),\ i = 1, \cdots, n+1$.*

Proof. For any simplex V there exists an affine transformation φ such that $\varphi(V)$ is a standard simplex. The ratio

$$\frac{\mu\left(\varphi(V) \cap \varphi\left(\overrightarrow{H}^{-1}(I)\right)\right)}{\mu(\varphi(V))}$$

only depends on the box I and node values $\overrightarrow{H}(\mathbf{q}_i) = (H_1(\mathbf{q}_i), H_2(\mathbf{q}_i), \cdots, H_{n-1}(\mathbf{q}_i)),\ i = 1, \cdots, n+1$.

By the property of affine transformations,

$$\frac{\mu(V \cap \overrightarrow{H}^{-1}(I))}{\mu(V)} = \frac{\mu\left(\varphi(V) \cap \varphi\left(\overrightarrow{H}^{-1}(I)\right)\right)}{\mu(\varphi(V))}.$$

Therefore the original proportion $\frac{\mu(V \cap \overrightarrow{H}^{-1}(I))}{\mu(V)}$ only depends on the box I and the node values $\overrightarrow{H}(\mathbf{q}_i),\ i = 1, \cdots, n+1$. □

Define the proportion vector as

$$r_m := \frac{\mu(V_m \cap H^{-1}(I))}{\mu(V_m)}.$$

Then for any given trajectory bundle $S \in \mathcal{F}(I)$, its measure can be calculated with the following formula:

$$\mu(S) = \sum_{V_m \in \mathcal{R}_{min}(S)} r_m \cdot \mu(V_m).$$

The whole process introduced above includes the evaluation of connection matrix and the proportion vector, therefore we name the proposed algorithm as the connection-proportion algorithm.

Remark 4.3. All divergence-free advection equation can be averaged, but only Nambu systems allow connection-proportion algorihm.

4.4 Properties

The Galerkin method for trajectorial average perserves two important properties of the unaveraged equation(4.1):

1. The growth rate of invariants. The solution ρ to the unaveraged equation(4.1) does not necessarily belong to $\ker\mathcal{T}$, however, if tested with some invariant $\eta \in \ker\mathcal{T}$, it follows that
$$(\partial_t \rho, \eta) = (\gamma\rho, \eta)\,.$$
The solution ρ_h of discrete averaged weak equation(4.9) preserves this identity. Indeed, denote the interpolation operator as $\Pi_h : \mathcal{U} \to \mathcal{U}_h$, the discrete weak form renders
$$(\partial_t \rho_h, \Pi_h\eta) = (\gamma\rho_h, \Pi_h\eta)\,.$$
In particular, thanks to the partition of unity shown in Equation(4.8), we have $\Pi_h 1 = 1$, hence the mass growth rate is exact.

2. Positivity. The solution ρ to the unaveraged equation(4.1) remains postive if $\rho(\mathbf{q}, 0) \geq 0$. The solution ρ_h of discrete averaged weak equation(4.9) preserves positivity as well, because the mass matrix M_{ij} and the reaction matrix Γ_{ij} in Equation(4.10) are both diagonal matrices.

As will be seen in the next chapter, the above properties are the key of a structure-preserving solver for Equation(2.23).

Chapter 5: A structure-preserving solver for particle-wave interaction in non-uniform magnetized plasmas

5.1 Electron-plasmon kinetic system

Consider an infinitely long cylinder $\Omega_x = (0, R) \times (0, 2\pi) \times \mathbb{R}_z$. At any given location $(r, \phi, z) \in \Omega_x$, there is a set of local orthonormal basis $(\mathbf{e}_z, \mathbf{e}_r, \mathbf{e}_\phi)$.

Suppose a plasma is confined in the cylinder, embedded in a magnetic field along the z-axis, $\mathbf{B}(\mathbf{x}) = B(r)\mathbf{e}_z$. For simplicity, we focus on the z, ϕ-symmetric case, which means any function $g(\mathbf{x})$ depends only on r.

For highly magnetized plasma, the electron probability density function $f(\mathbf{p}, \mathbf{x}, t) = f(p_\parallel, p_\perp, r, t)$, where momentum vector $\mathbf{p}$ is decomposed into a parallel component $p_\parallel$ and a perpendicular component $p_\perp$ relative to the magnetic field line.

The momentum $\mathbf{k}$ of a plasmon (the wave vector of a wave packet) can also be decomposed as $\mathbf{k} = k_z\mathbf{e}_z + k_r\mathbf{e}_\mathbf{r} + k_\phi\mathbf{e}_\phi \in \mathbb{R}^3_k$. Define angular momentum $q_\phi = k_\phi \cdot r$, then (k_r, q_ϕ, k_z) and (r, ϕ, z) are a set of canonical coordinates. Given z, ϕ-symmetry, the plasmon probability density function $N(\mathbf{k}, \mathbf{x}, t) = N(k_r, q_\phi, k_z, r, t)$.

The following system governs the evolution of particle *pdf* $f(\mathbf{p}, \mathbf{x}, t)$ and plasmon *pdf* $N(\mathbf{k}, \mathbf{x}, t)$,

$$\begin{cases} \partial_t f + v_z \partial_z f & = \nabla_p \cdot (D[N] \cdot \nabla_p f), \\ \partial_t N + \{\omega, N\} & = \Gamma[f] N. \end{cases} \tag{5.1}$$

The advection term $v_z\partial_z f$ comes from the gyro-averaged Vlasov equation, it vanishes since f does not depend on z. The Poisson bracket

$$\{\omega, N\} = \frac{\partial N}{\partial r}\frac{\partial \omega}{\partial k_r} - \frac{\partial \omega}{\partial r}\frac{\partial N}{\partial k_r}$$

is a result of WKB approximation. Denote it as $\mathcal{T}N$, then obviously the operator $\mathcal{T}$ is skew-symmetric. The quasilinear diffusion term and the reaction term account for particle-plasmon interaction. Test the equations with $\varphi(\mathbf{p}, \mathbf{x})$ and $\eta(\mathbf{k}, \mathbf{x})$ to obtain the following weak form,

$$\iint_{px} \varphi \partial_t f + \iint_{kx} \eta \partial_t N = \iint_{kx} N \mathcal{T} \eta + \iiint_{pkx} \mathcal{B} N \mathcal{L} f \left(\eta \mathcal{L} \mathcal{E} - \hbar \omega \mathcal{L} \varphi \right).$$

Due to nonlinearity, high-dimensionality, multi-scale, and resonant transition rates, the problem is so hard that there is no numerical solver for it yet. In this chapter, a structure-preserving solver will be presented(Huang et al., 2023b)[1], putting together the scheme for the homogeneous system in Chapter 3 and the algorithm for trajectorial average in Chapter 4.

This chapter is organized as follows. Section 5.2 introduces the weak form of the averaged kinetic system. Section 5.3 concerns the trajectory bundle partition of the plasmon phase space. In Section 5.4 we discuss the LDG discretization. Section 5.5 focuses on the details of implementation. The complexity analysis will be provided in Section 5.6. Numerical results are presented in Section 3.7.

5.2 Averaged system and properties

As can be observed in Equation(5.1), there are three different time scales associated with the model, τ_{diff}, τ_{adv} and τ_{reac}. According to Kiramov and Breizman (2021); Aleynikov and Breizman (2015), the advection process is much faster than the other two: $\tau_{\text{adv}} \sim \varepsilon \tau_{\text{diff}} \sim \varepsilon \tau_{\text{reac}}$, where $\varepsilon \ll 1$. In practice, particle-wave interaction is of interest, rather than plasmon advection. Hence it is reasonable to eliminate the Poisson bracket term through trajectorial average.

[1]Kun Huang, Irene M Gamba, and Chi-Wang Shu. Structure-preserving solvers for particle-wave interaction in non-uniform magnetized plasmas. in preparation, 2023. The dissertator's contribution includes proposing and implementing the scheme, analyzing the data and writing the article.

Assume that $N(\mathbf{k}, \mathbf{x}, t)$ has the following Hilbert expansion:

$$N = N_0 + \varepsilon N_1 + \varepsilon^2 N_2 + \cdots .$$

Then the diffusion equation for particle *pdf* in (5.1) can be approximated with

$$\partial_t f = \nabla_p \cdot (D[N_0] \cdot \nabla_p f).$$

Meanwhile the advection-reaction equation for plasmon *pdf* renders

$$\partial_t N_0 = \mathcal{P}\left(\Gamma[f] N_0\right).$$

Define test/trial spaces as follows

$$\begin{aligned} \mathcal{G} &\coloneqq H^1(\mathbb{R}^3_p) \otimes L^2(\Omega_x), \\ \mathcal{W} &\coloneqq \ker\mathcal{T} \cap L^2(\mathbb{R}^3_k \times \Omega_x). \end{aligned} \tag{5.2}$$

The averaged system, tested with $\varphi(\mathbf{p}, \mathbf{x}) \in \mathcal{G}$ and $\eta(\mathbf{k}, \mathbf{x}) \in \mathcal{W}$, yields

$$\iint_{px} \varphi \partial_t f + \iint_{kx} \eta \partial_t N = \iiint_{pkx} \mathcal{B} N \mathcal{L} f \left(\eta \mathcal{L}\mathcal{E} - \hbar\omega \mathcal{L}\varphi\right), \tag{5.3}$$

where we have used N in place of N_0 since the higher order terms in the Hilbert expansion (4.2) will be ignored in the rest of this chapter.

Remark 5.1. Note that Equation(5.3) has the same structure as Equation(3.21) in Chapter 3 for homogeneous plasmas. They only differ in the choice of test/trial spaces. Therefore the techniques used in Chapter 3 can be translated here without extra effort.

In Chapter 3, it has been proved that our choice of directional differential operator $\mathcal{L}$ guarantees unconditional conservation. In the following theorem, we show that the unconditional conservation property is preserved even after trajectorial average.

Theorem 5.1 (unconditional conservation)**.** *If $f(\mathbf{p}, \mathbf{x}, t)$ and $N(\mathbf{k}, \mathbf{x}, t)$ solve Equation(5.3) with emission/absorption kernel replaced by $\mathcal{B}_\varepsilon$, then for any $\mathcal{B}_\varepsilon$ we have the following conservation laws,*

$$\begin{aligned} \partial_t \mathcal{M}_{tot} &= \partial_t \left((f, 1)_{px} + (N, 0)_{kx}\right) = 0, \\ \partial_t \mathcal{P}_{z,tot} &= \partial_t \left((f, p_z)_{px} + (N, \hbar k_z)_{kx}\right) = 0, \\ \partial_t \mathcal{E}_{tot} &= \partial_t \left((f, \mathcal{E})_{px} + (N, \hbar\omega)_{kx}\right) = 0. \end{aligned}$$

Proof. We consider the conservation laws one by one. Mass conservation is trivial. The energy conservation law can also be verified easily, since $\omega \in \ker\mathcal{T}$.

For momentum conservation, note that $k_z \in \ker\mathcal{T}$, hence

$$\partial_t \mathcal{P}_{z,tot} = \iiint_{pkx} (\hbar k_z \mathcal{L}\mathcal{E} - \hbar\omega \mathcal{L} p_z)\, \mathcal{L} f N \mathcal{B}_\varepsilon.$$

Recall the definition

$$\mathcal{L} g \coloneqq \frac{k_\parallel}{\omega} \frac{\partial \mathcal{E}}{\partial p_\perp} \frac{p}{p_\perp} \frac{\partial g}{\partial p_\parallel} + \left(1 - \frac{k_\parallel}{\omega} \frac{\partial \mathcal{E}}{\partial p_\parallel}\right) \frac{p}{p_\perp} \frac{\partial g}{\partial p_\perp}.$$

Since $\mathbf{B} \parallel \mathbf{e}_z$, $k_z = k_\parallel$ and $p_z = p_\parallel$, therefore

$$k_z \mathcal{L}\mathcal{E} - \omega \mathcal{L} p_z = k_z \frac{p}{p_\perp} \frac{\partial \mathcal{E}}{\partial p_\perp} - k_\parallel \frac{\partial \mathcal{E}}{\partial p_\perp} \frac{p}{p_\perp} = 0.$$

It follows that $\partial_t \mathcal{P}_{z,tot} = 0$ regardless of $\mathcal{B}_\varepsilon$. □

5.3 Partition of the plasmon phase space

To discretize the test/trial space $\mathcal{W} \coloneqq \ker\mathcal{T} \cap L^2(\mathbb{R}^3_k \times \Omega_x)$, the only possible approach is the connection-proportion algorithm proposed in Chapter 4. However, that algorithm is designed only for Nambu systems. The Poisson bracket $\{\omega, N\}$, indeed, is not a Nambu bracket defined in Equation(2.4). Fortunately, thanks to the symmetry along ϕ and z direction, for fixed q_ϕ and k_z, it can be regarded as a Nambu bracket in 2-dimensional phase space.

5.3.1 Interpolated Hamiltonian

Recall the formulation of Poisson bracket

$$\mathcal{T} N \coloneqq \{\omega, N\} = \frac{\partial N}{\partial r} \frac{\partial \omega}{\partial k_r} - \frac{\partial \omega}{\partial r} \frac{\partial N}{\partial k_r}.$$

The operator $\mathcal{T}$ determines a flow on $\mathbb{R}^2$. However, the Hamiltonian ω for plasmons is a function of four variables, $\omega = \omega(k_r, q_\phi, k_z, r)$, which means we have to

construct the corresponding trajectory bundles for infinitely many (q_ϕ, k_z). To avoid that, for implementation, we propose the following interpolation of the dispersion relation.

To begin with, partition the space $\Omega_a = \mathbb{R}_{q_\phi} \times \mathbb{R}_{k_z}$ into rectangular meshes $\{R\}$, and partition the domain $\Omega_b = \mathbb{R}_{k_r} \times (0, R_{max})$ into triangular meshes $\{T\}$.

Next, define the following space of piecewise polynomials.

$$\mathcal{H} = \left\{ H(k_r, r, q_\phi, k_z) \in C^0(\Omega_b) : H|_{R\times T} \in Q^0(R) \times P^1(T), \forall\ R \subset \Omega_a, T \subset \Omega_b \right\}.$$

Project the function $\omega = \omega(k_r, q_\phi, k_z, r)$ onto $\mathcal{H}$ to obtain the interpolated Hamiltonian $\overline{\omega} = \overline{\omega}(k_r, q_\phi, k_z, r)$. Obviously, for any given (q_ϕ, k_z), $\overline{\omega}$ is a continuous piecewise linear function on Ω_b, while for any given (k_r, r), $\overline{\omega}$ is a piecewise constant function on Ω_a.

The purpose of such interpolation is as follows.

Theorem 5.2. *If $\omega \in \mathcal{H}$, and $S \subset \Omega_b$ is a trajectory bundle for $\omega(\cdot, \cdot, q_\phi, k_z)$ with $(q_\phi, k_z) \in R \subset \Omega_a$, then*

$$\{\omega, \mathbf{1}_{S\times R}\} = 0.$$

In other words, $\widetilde{S} \coloneqq S \times R$ is a trajectory bundle w.r.t. the operator $\mathfrak{T}N \coloneqq \{\omega, N\}$.

Theorem 5.2 guarantees that it is sufficient to construct only a finite number of trajectory bundles, as long as there is a representative (q_ϕ, k_z) sample for each rectangular mesh $R \subset \Omega_a \coloneqq \mathbb{R}_{q_\phi} \times \mathbb{R}_{k_z}$, see Figure 5.1.

5.3.2 Cut-off domains

Another concern is the unboundedness of domain $\Omega_{kx} = \Omega_b \times \Omega_a = \mathbb{R}_{k_r} \times (0, R_{max}) \times \mathbb{R}_{q_\phi} \times \mathbb{R}_{k_z}$. As introduced in Chapter 4, the connection-proportion algorithm is designed for bounded domains.

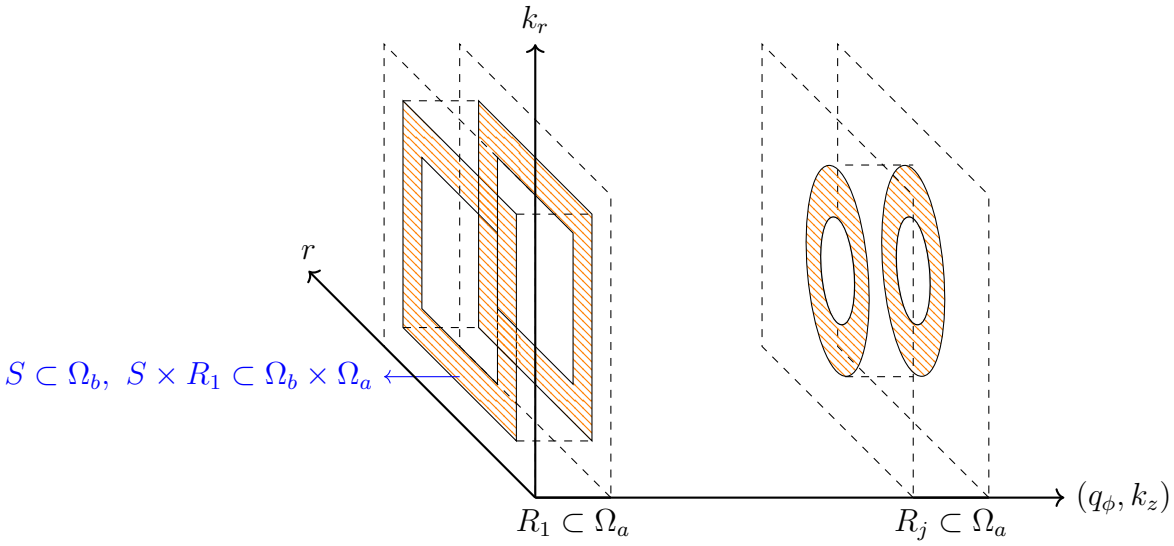


Figure 5.1: Trajectory bundles in the plasmon phase space.

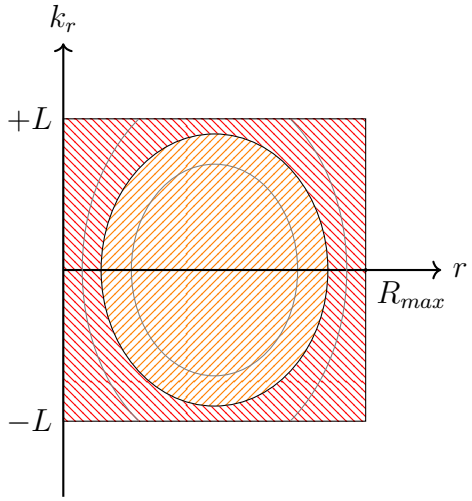


Figure 5.2: Complete and incomplete trajectory bundles in a cut-off domain.

To resolve that, we consider the following cut-off domain:

$$\Omega_b^L \times \Omega_a^L = [(-L, L) \times (0, R_{max})] \times \left[\left(q_\phi^l, q_\phi^r\right) \times \left(k_z^l, k_z^r\right)\right],$$

then the connection-proportion algorithm is feasible. Denote the trajectory bundles generated by the algorithm as $\{\widetilde{S}^L\}$.

If a trajectory bundle $\widetilde{S}^L$ does not intersect with the k_r boundary $e^L = \{-L, L\} \times (0, R_{max}) \times \Omega_a^L$, then it is also a complete trajectory bundle for the original domain, otherwise it is just part of an original trajectory bundle. As shown in Figure 5.2, the red region contains incomplete trajectory bundles.

There are two choices for implementation:

1. "Peel off" the incomplete trajectory bundles and define the cut-off domain as $\Omega_{kx}^L = \cup_{\widetilde{S}^L \text{ is complete}} \widetilde{S}^L$.
2. Keep all the incomplete trajectory bundles and define cut-off domain as $\Omega_{kx}^L = \cup \widetilde{S}^L = \Omega_b^L \times \Omega_a^L$, which is actually a box.

For both choices we have $\lim_{L\to\infty} \Omega_{kx}^L = \Omega_{kx}$.

5.4 DG formulation

As can be observed, Equation(5.3) shares the same form as Equation(3.21). The only difference lies in the test/trial spaces. So this section focuses mainly on the discretization of function spaces.

5.4.1 Cut-off domains

Analogous to the discussion in Chapter 3, it is assumed that given any $0 < \epsilon_p \ll 1$ and $0 < \epsilon_k \ll 1$, there exists bounded domains $\Omega_{px}^L = \Omega_p^L \times \Omega_x \subsetneq \mathbb{R}_p^3 \times \Omega_x$ and $\Omega_{kx}^L \subsetneq \mathbb{R}_k^3 \times \Omega_x$ such that for any $t \geq 0$,

$$\left| 1 - \frac{\int_{\Omega_{px}^L} f(\mathbf{p}, \mathbf{x}, t)}{\int_{\mathbb{R}_p^3 \times \Omega_x} f(\mathbf{p}, \mathbf{x}, t)} \right| \leq \epsilon_p,$$

and

$$\left| 1 - \frac{\int_{\Omega_{kx}^L} N(\mathbf{k}, \mathbf{x}, t)}{\int_{\mathbb{R}_k^3 \times \Omega_x} N(\mathbf{k}, \mathbf{x}, t)} \right| \leq \epsilon_k.$$

The cut-off domain Ω_{px}^L for particles is supposed to be adaptive to ensure that $|f|$ and $|\nabla_{\mathbf{p}} f|$ are nearly zero on the boundary, while in our numerical experiments it turns out that, as a result of anisotropic diffusion, there is no need to extend it. The cut-off domain Ω_{kx}^L for plasmons consists of trajectory bundles, as have been discussed in the last section.

5.4.2 Discrete test spaces

Note that the particle *pdf* is symmetric, i.e. $f(\mathbf{p}, \mathbf{x}, t) = f(p_\parallel, p_\perp, r, t)$, we have

$$\mathbf{p} = (p_1, p_2, p_3) \in \Omega_p^L \Leftrightarrow (p_\parallel, p_\perp) \in \tilde{\Omega}_p^L \subset \mathbb{R} \times \mathbb{R}^+.$$

Let $\{R_p\}$ be the rectangular partition of $\tilde{\Omega}_p^L$, let $\{R_x\}$ be the partition of interval $(0, R_{max})$, and let $\{\widetilde{S}\}$ be the trajectory bundle partition of Ω_{kx}^L.

The test space for particle *pdf* consists of discontinuous piecewise polynomials,

$$\mathcal{G}_h \coloneqq \{f(p_\parallel, p_\perp, r) : f|_{R_p \times R_x} \in Q^{\alpha_1}(R_p) \times Q^{\alpha_2}(R_x), \forall\ R_p, R_x\}. \tag{5.4}$$

The test space for plasmon *pdf* consists of indicator functions of trajectory bundles,

$$\mathcal{W}_h \coloneqq \left\{ N(\mathbf{k}, \mathbf{x}) : N = \sum_{\widetilde{S}} n_{\widetilde{S}} \mathbf{1}_{\widetilde{S}} \right\}. \tag{5.5}$$

5.4.3 Unconditionally conservative semi-discrete form

Define the projection operator $\Pi_{kx,h}$ onto $\mathcal{W}_h$ as follows:

$$\Pi_{kx,h} G(\mathbf{k}, \mathbf{x}) = \sum_{\widetilde{S}} \left(\sup_{(\mathbf{k},\mathbf{x}) \in \widetilde{S}} G(\mathbf{k}, \mathbf{x}) \right) \cdot 1_{\widetilde{S}}. \tag{5.6}$$

Discretize ω and $k_\parallel$ as $\omega_h \coloneqq \Pi_{kx,h}\omega$ and $k_{\parallel,h} \coloneqq \Pi_{kx,h} k_\parallel$.

The discrete directional differential operator takes the form as follows:

$$\mathcal{L}_h g_h \coloneqq \frac{1}{\omega}\left[k_{\parallel,h}\left(\partial_{\perp,h}\mathcal{E}_h\right)\cdot\frac{p}{p_\perp}\left(\partial_{\parallel,h}g_h\right)+\left(\omega_h-k_{\parallel,h}\left(\partial_{\parallel,h}\mathcal{E}_h\right)\right)\frac{p}{p_\perp}\left(\partial_{\perp,h}g_h\right)\right].$$

Using the operator $\mathcal{L}_h$ above, the semi-discrete DG form preserves unconditional conservation laws, as shown in the following theorem.

Theorem 5.3. *If $f_h(\cdot,t)\in\mathcal{G}_h$ and $N_h(\cdot,t)\in\mathcal{W}_h$ solves the following equation*

$$\iint_{px}\varphi_h\partial_t f_h+\iint_{kx}\eta_h\partial_t N_h=\iiint_{pkx}\left(\eta_h\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\varphi_h\right)\mathcal{L}_h f_h N_h\mathcal{B}_\varepsilon,$$

for any $\phi_h\in\mathcal{G}_h$ and $\eta_h\in\mathcal{W}_h$. Then the conservation laws hold regardless of emission/absorption kernel $\mathcal{B}_\varepsilon$.

5.5 Sparse interaction tensors

Recall the semi-discrete weak form

$$\iint_{px}\varphi_h\partial_t f_h+\iint_{kx}\eta_h\partial_t N_h=\iiint_{pkx}\left(\eta_h\mathcal{L}_h\mathcal{E}_h-\hbar\omega_h\mathcal{L}_h\varphi_h\right)\mathcal{L}_h f_h N_h\mathcal{B}_\varepsilon,$$

Suppose that the test space $\mathcal{G}_h$ consists of basis functions $\{\varphi_i\}$ and the test space $\mathcal{W}_h$ consists of basis functions $\{\eta_j\}$, and let

$$f_h(\mathbf{p},\mathbf{x},t)=\sum a_i(t)\varphi_i(\mathbf{p},\mathbf{x}),$$
$$N_h(\mathbf{k},\mathbf{x},t)=\sum N_j(t)\eta_j(\mathbf{k},\mathbf{x}).$$

The weak form is then equivalent to the following system,

$$\begin{aligned}\sum_i\partial_t a_i\iint_{px}\varphi_i\varphi_m&=-\sum_n\sum_q\sum_s a_n N_q\hbar\omega_s\iiint_{pkx}\eta_q\eta_s\left(\mathcal{L}_h\varphi_m\right)\left(\mathcal{L}_h\varphi_n\right)\mathcal{B}_\varepsilon,\\ \sum_j\partial_t N_j\iint_{kx}\eta_j\eta_s&=\sum_n\sum_q\sum_m a_n N_q\mathcal{E}_m\iiint_{pkx}\eta_q\eta_s\left(\mathcal{L}_h\varphi_m\right)\left(\mathcal{L}_h\varphi_n\right)\mathcal{B}_\varepsilon.\end{aligned}\tag{5.7}$$

Denote the mass matrix for particle *pdf* as $M_{im}=\iint_{px}\varphi_i\varphi_m$, and denote the mass matrix for plasmon *pdf* as $G_{js}=\iint_{kx}\eta_j\eta_s$. Define the sparse interaction tensor

$$B_{qmn}=\iiint_{pkx}\eta_q\left(\mathcal{L}_h\phi_m\right)\left(\mathcal{L}_h\phi_n\right)\mathcal{B}_\varepsilon.$$

Following the same procedure as in Chapter 3, Equation(5.7) can be written as follows,

$$\begin{aligned}
\sum_i M_{im}\partial_t a_i &= -\sum_n \sum_q \sum_s a_n N_q \omega_s \delta_{qs} B_{qmn} = -\sum_n \sum_q a_n N_q \hbar \omega_q B_{qmn}, \\
\sum_j G_{js}\partial_t N_j &= \sum_n \sum_q \sum_m a_n N_q \mathcal{E}_m \delta_{qs} B_{qmn} = N_s \sum_n \sum_m a_n \mathcal{E}_m B_{smn}.
\end{aligned} \tag{5.8}$$

Recall the definition of test space for particle *pdf*,

$$\mathcal{G}_h \coloneqq \{f(p_\parallel, p_\perp, r) : f|_{R_p \times R_x} \in Q^{\alpha_1}(R_p) \times Q^{\alpha_2}(R_x), \forall\ R_p, R_x\}.$$

Here we present the explicit formulation of interaction tensors for the simplest case: when $\alpha_1 = \alpha_2 = 0$, i.e. the basis functions are piecewise constant functions on the rectangular meshes.

It can be verified that the discrete gradient operators read,

$$\begin{aligned}
\left(\partial_{\parallel,h} g\right)_{i,j} &= \frac{g_{i+1,j} - g_{i,j}}{\left(\Delta p_\parallel\right)^i}, \\
\left(\partial_{\perp,h} g\right)_{i,j} &= \frac{p_\perp^{i,j+1/2}}{p_\perp^{i+1/2,j}} \frac{g_{i,j+1} - g_{i,j}}{\left(\Delta p_\perp\right)^j}.
\end{aligned} \tag{5.9}$$

By i, j we mean the ith mesh in $p_\parallel$ axis and the jth mesh in $p_\perp$ axis. For elements on the upper-right boundary, we define

$$\begin{aligned}
g_{I+1,j} &= g_{I,j}, \\
g_{i,J+1} &= g_{i,J}.
\end{aligned}$$

Suppose that

$$\begin{aligned}
\varphi_m(p_\parallel, p_\perp, r) &= \lambda_\xi(r)\tilde{\varphi}_\mu(p_\parallel, p_\perp), \\
\varphi_n(p_\parallel, p_\perp, r) &= \lambda_\zeta(r)\tilde{\varphi}_\nu(p_\parallel, p_\perp).
\end{aligned}$$

Then by definition

$$\begin{aligned}
B_{qmn} &= \iiint_{pkx} \eta_q \left(\mathcal{L}_h\varphi_m\right)\left(\mathcal{L}_h\varphi_n\right)\mathcal{B}_\varepsilon \\
&= \sum_{R_{ij}} \int_{R_{ij}} \left(\iint_{kx} \eta_q\lambda_\xi\lambda_\zeta \left(\mathcal{L}_h\tilde{\varphi}_\mu\right)\left(\mathcal{L}_h\tilde{\varphi}_\nu\right)\mathcal{B}_\varepsilon \right) \\
&\approx \sum_{R_{ij}} \left(\Delta h_\parallel\right)^i \left(\Delta h_\perp\right)^j 2\pi p_\perp^j \iint_{kx} \left[\eta_q\lambda_\xi\lambda_\zeta \left(\mathcal{L}_h\tilde{\varphi}_\mu\right)\left(\mathcal{L}_h\tilde{\varphi}_\nu\right)\mathcal{B}_\varepsilon\right]\Big|_{p_\parallel^i, p_\perp^j} \\
&= \sum_{R_{ij}} \delta_{\xi\zeta} \left(\Delta h_\parallel\right)^i \left(\Delta h_\perp\right)^j 2\pi p_\perp^j \left[\left(\nabla_h\tilde{\varphi}_\mu\right)_{ij} \cdot D^{ij}[\eta_q,\lambda_\xi] \cdot \left(\nabla_h\tilde{\varphi}_\nu\right)_{ij}\right],
\end{aligned}$$

where the piecewise constant diffusion coefficient is defined as follows

$$D^{ij}[\eta_q,\lambda_\xi] = \left(\iint_{kx} \eta_q\lambda_\xi \left(\beta_h^{ij}\otimes\beta_h^{ij}\right)\mathcal{B}_\varepsilon \right)\Big|_{p_\parallel^i, p_\perp^j},$$

with

$$\beta_h^{ij} = \left[\ k_{\parallel,h}\left(\partial_{\perp,h}E_h\right)^{ij}\frac{p^{ij}}{p_\perp^j},\quad \left(\omega_h - k_{\parallel,h}\left(\partial_{\parallel,h}E_h\right)^{ij}\right)\frac{p^{ij}}{p_\perp^j}\ \right].$$

The advantage of storing $D^{ij}[\eta_q,\lambda_\xi]$ instead of the tensor B_{qmn} is that, in this way, the CFL condition can be explicitly calculated.

5.6 Complexity analysis

Suppose that the r-domain $(0, R_{max})$ is partitioned into $n_r \sim \mathcal{O}(n)$ grids, the cut-off p-domain $(P_\parallel^l, P_\parallel^r)\times(0, P_\perp^m)$ is partitioned into $n_p \sim \mathcal{O}(n^2)$ grids.

In addition, suppose that $n_{z\phi} \sim \mathcal{O}(n^2)$ is the number of grids in cut-off (k_z, q_ϕ)-domain $(K_z^l, K_z^r)\times(0, Q_\phi^m)$, and $n_\Delta \sim \mathcal{O}(n^2)$ is the number of triangular meshes for interpolation of the Hamiltonian $H(r, k_r)$. For each discrete Hamiltonian, we stratify it into $n_s \sim \mathcal{O}(n)$ layers, then $n_b \sim \mathcal{O}(n^3)$ trajectory bundles are generated in total.

data

The degree of freedom for discrete particle *pdf* f_h is $n_r \times n_p \sim \mathcal{O}(n^3)$, and for discrete plasmon *pdf* N_h it is $n_b \sim \mathcal{O}(n^3)$.

solving

The linear map from plasmon *pdf* to element-wise diffusion coefficient is a $[n_r \times n_p] \times [n_b]$ sparse matrix with approximately $n_r \times n_p \times n_{z\phi} \sim \mathcal{O}(n^5)$ non-zero elements.

In each time step, the most expensive part is to obtain element-wise diffusion coefficients by matrix-vector product, taking $\mathcal{O}(n^5)$ flops. Note that since the diffusion process happens only in momentum space, the solving procedure is naturally parallelizable along the r-axis.

preparing

For each grid in (k_z, q_ϕ)-domain, we interpolate the Hamiltonian $H(r, k_r; k_z^c, q_\phi^c)$ with piecewise linear basis on triangular meshes, which requires evaluating ω on n_Δ nodes. For each interval (H_a, H_b), constructing the connection matrix as defined in (4.11) takes n_Δ flops, distinguishing all its connected components takes $\mathcal{O}(n)$ flops. Based on the above analysis, the time complexity for constructing all the trajectory bundles is $\mathcal{O}(n_{z\phi} \times n_s \times n_\Delta) \sim \mathcal{O}(n^5)$. However, since the procedure is independent along k_z-axis and q_ϕ-axis, the time complexity can be reduced in practice through parallel computing.

For each of these trajectory bundles we have to store its minimal triangle cover, thus the space complexity is $\mathcal{O}(n^4)$.

Constructing the diffusion coefficients takes $[n_r \times n_p] \times [n_{z\phi} \times n_\Delta] \sim \mathcal{O}(n^6)$ flops, the procedure is also naturally parallelizable since all the evaluations are independent.

5.7 Numerical results

Problem setting

Analogous to Chapter 3, only anomalous Doppler resonance with $l = 1$ is considered. For dispersion relation $\omega(\mathbf{k}, \mathbf{x})$ we take the lowest branch given by Equation(2.10).

Set the cut-off computational domains as follows:

$$\Omega_{px}^{L} = \left\{ \left(r, p_{\parallel}, p_{\perp}\right) : r \in (0, R_{max}) \,,\ p_{\parallel} \in (10 m_e c, 25 m_e c) \,,\ p_{\perp} \in (0, 15 m_e c) \right\},$$

$$\Omega_{kx}^{L} = \left\{ (r, k_r, q_\phi, k_z) : r \in (0, R_{max}) \,,\ k_r \in \left(-\frac{\omega_0}{c}, \frac{\omega_0}{c}\right) \,,\ q_\phi \in \left(0, 0.5\frac{\omega_0 R_{max}}{c}\right) \,,\ k_z \in \left(0, \frac{\omega_0}{c}\right) \right\}.$$

The domain Ω_{px}^{L} is partitioned into rectangular meshes, with 75 meshes in $p_{\parallel}$ axis, 75 meshes in $p_{\perp}$ axis, and 20 meshes in r axis.

The domain Ω_{kx}^{L} is partitioned as follows. First of all, partition the ϕz-domain into 20×40 rectangular meshes. In each ϕz-mesh interpolate $\omega(\cdot, \cdot, q_\phi, k_z)$ on $20 \times 20 \times 2$ triangular meshes, as illustrated in Figure 5.3. For each Hamiltonian $\omega(\cdot, \cdot, q_\phi, k_z)$, we construct trajectory bundles generated by 10 boxes with the connection-proportion algorithm.

Choose piecewise constant functions as the test/trial spaces:

$$\mathcal{G}_h = \left\{ \varphi\left(r, p_{\parallel}, p_{\perp}\right) : \varphi|_{R_p \times R_x} \in Q^0(R_p) \times Q^0(R_x), \forall\ R_p, R_x \right\},$$

$$\mathcal{W}_h = \left\{ \eta\left(r, k_r, q_\phi, k_z\right) : \eta = \sum_{\widetilde{S}} n_{\widetilde{S}} \mathbf{1}_{\widetilde{S}} \right\}.$$

The projection operators $\Pi_{kx,h}$ is already given in Equation(5.6). Meanwhile, we define projection $\Pi_{p,h}$ as

$$\Pi_{p,h} G(p_{\parallel}, p_{\perp}) \coloneqq G\left(p_{\parallel}^{i-1/2}, p_{\perp}^{j-1/2}\right) \,,\ \forall (p_{\parallel}, p_{\perp}) \in R_p^{i,j}.$$

Consider a magnetized plasma with non-uniform electron density:

$$n_e(r) = n_0 \left[1 - \left(\frac{r}{R_{max}}\right)^2\right],$$

embedded in a uniform external magnetic field

$$\mathbf{B}(r) = B_0 \mathbf{e}_z.$$

Analogous to Chapter 3, we only compute the "bump" part of electron *pdf* $f(\mathbf{p}, \mathbf{x}, t)$, which takes the following initial configuration:

$$f(p_\parallel, p_\perp, r)\big|_{t=0} = \left[10^{-5} \frac{1}{\sqrt{\pi}} \exp\left(-\left(\frac{p_\parallel}{m_e c} - 20\right)^2 - \left(\frac{p_\perp}{m_e c}\right)^2\right)\right] \frac{n_0}{m_e^3 c^3}.$$

Meanwhile, the plasmon *pdf* $N(\mathbf{k}, \mathbf{x}, t)$ is initialized as follows:

$$N\left(r, k_r, q_\phi, k_z\right)\big|_{t=0} = 10^{-5} \frac{n_0 m c^2}{(\omega_0/c)^3} \frac{1}{\hbar \omega_0}.$$

Trajectory bundles

As shown in Figure 5.3, the triangular partition of the (r, k_r)-domain is done by dividing every rectangular mesh into two. The connection-proportion algorithm successfully distinguishes different trajectory bundles in the inverse image of the same box(in this case, interval).

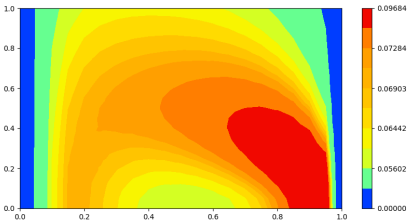
Probability density functions

Figure 5.5 shows the evolution of electron *pdf* $f(p_\parallel, p_\perp, r, t)$ at $r = \frac{31 R_{max}}{40}$. Similar to the result of homogeneous plasmas, the particles are scattered due to interaction with plasmons.

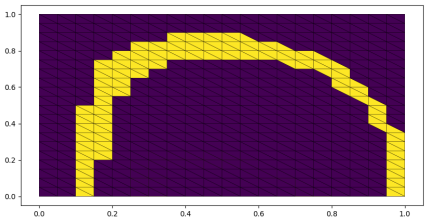
In Figure 5.4 we present the electron *pdf* at the same time point $t = 3 \times 10^6 \frac{1}{2\pi\omega_0}$ in different positions.

Conservation verification

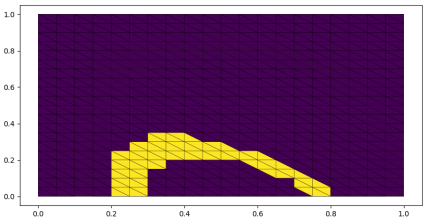
For the evolution of the electron-plasmon system momentum and energy, see Figure(5.6).



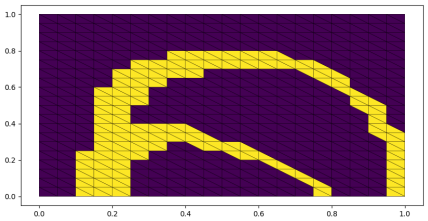
(a) $\omega(r, k_r; q_\phi, k_z)/\omega_0$ at $(q_\phi, k_z) = (\frac{9\omega_0}{80c}, \frac{9\omega_0 R_{max}}{80c})$



(b) The minimal triangle cover of S_{11}

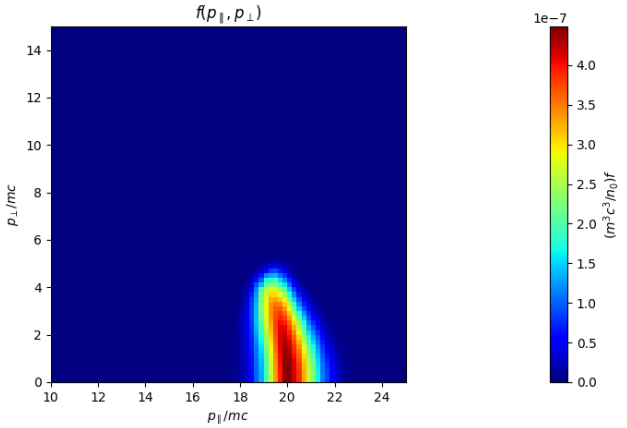


(c) The minimal triangle cover of S_{12}

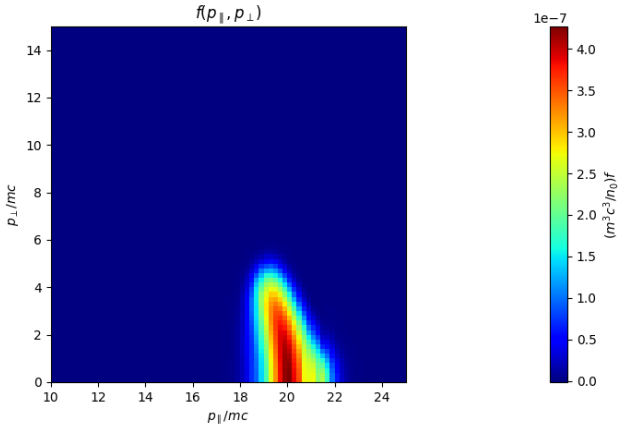


(d) The minimal triangle cover of S_{13}

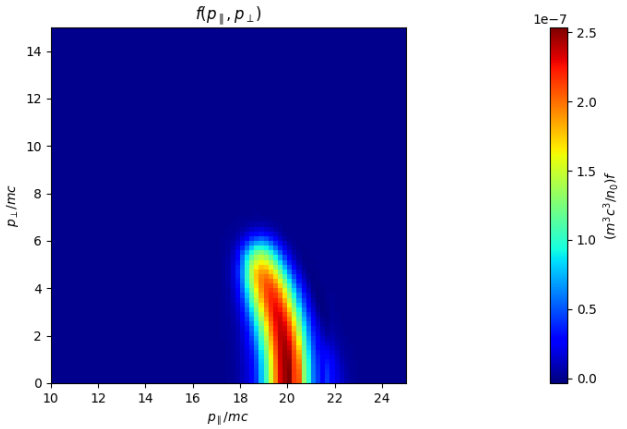
Figure 5.3: Trajectory bundles and their minimal triangle covers. The x-axis represents r/R_{max}, and the y-axis represents $k_r c/\omega_0$. Since the Hamiltonian is symmetric for $\pm k_r$, we only plot half of the domain. Note that S_{11} and S_{12} are two trajectory bundles generated by the same Hamiltonian range interval. And S_{13} is not a single strip because it contains a saddle point.



(a) $r = \frac{11R_{max}}{40}$

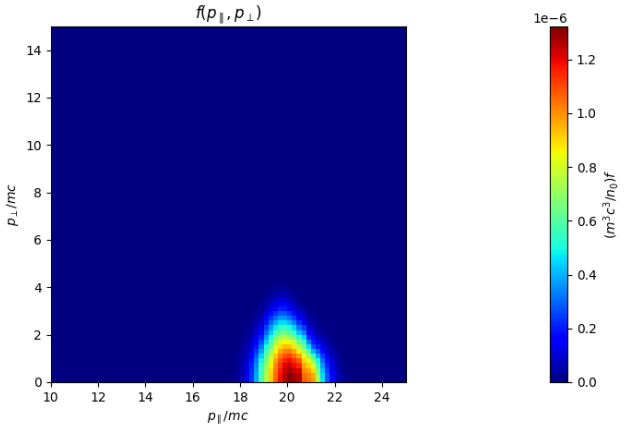


(b) $r = \frac{21R_{max}}{40}$

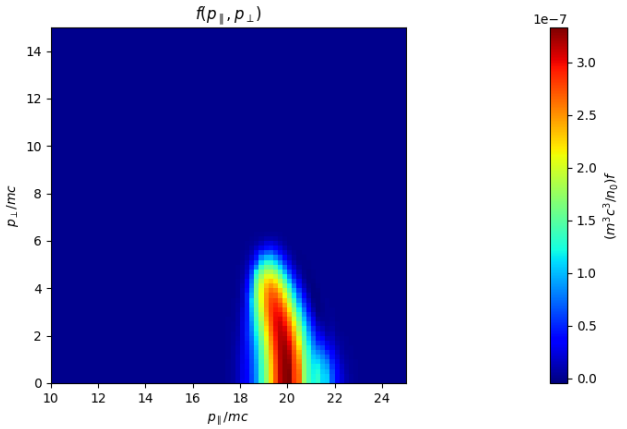


(c) $r = \frac{31R_{max}}{40}$

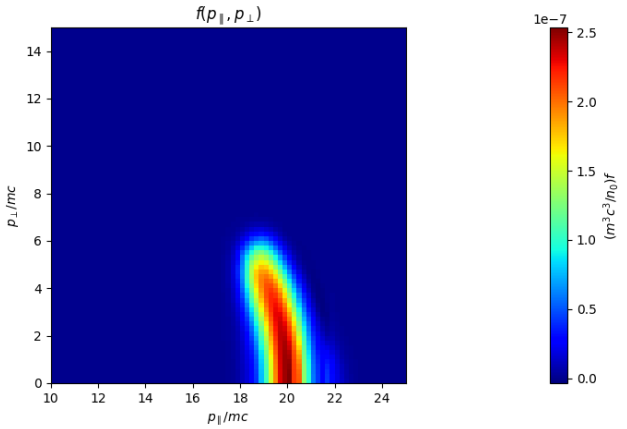
Figure 5.4: Spatial distribution at $t = 3 \times 10^6 \frac{1}{2\pi\omega_0}$



(a) $t = 1.0 \times 10^6 \frac{1}{2\pi\omega_0}$



(b) $t = 2.0 \times 10^6 \frac{1}{2\pi\omega_0}$

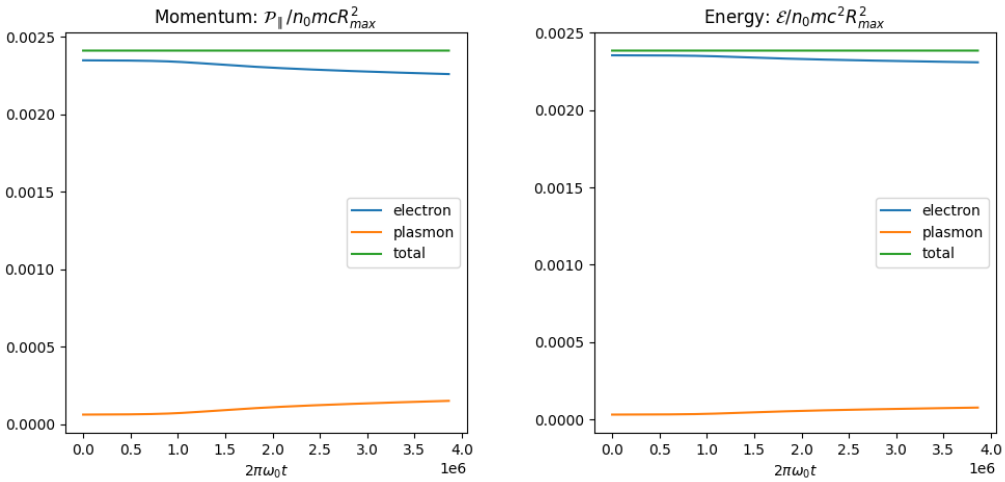


(c) $t = 3.0 \times 10^6 \frac{1}{2\pi\omega_0}$

Figure 5.5: Temporal evolution at $r = \frac{31R_{max}}{40}$.

With $T_{max} = 3.86 \times 10^6 \frac{1}{2\pi\omega_0}$, we have the following relative errors:

$$e_{rel}(\mathcal{M}_{tot,h}) = 4.6 \times 10^{-16},$$
$$e_{rel}(\mathcal{P}_{\parallel,tot,h}) = 1.8 \times 10^{-14},$$
$$e_{rel}(\mathcal{E}_{tot,h}) = 5.5 \times 10^{-16}.$$



(a) Momentum evolution.

(b) Energy evolution.

Figure 5.6: Conservation laws

Chapter 6: Existence of global weak solutions to quasilinear theory for electrostatic plasmas

6.1 Introduction

The quasilinear theory is a model reduction of the Vlasov-Maxwell(or Vlasov-Poisson) system in weak turbulence regime. Since first proposed by Vedenov et al. (1961) and Drummond and Pines (1962), it has found extensive use in plasma physics to describe the interaction between particles and waves(plasmons). Despite that, not much work from analysis point of view can be found. A recent paper by Bardos and Besse (2021) discussed the diffusion limit of the Vlasov-Poisson system, and showed that they can retreive formally the diffusion matrix of the quasilinear theory when the typical autocorrelation time of particles goes to infinity. The well-posedness of quasilinear systems remains an open problem. This chapter is devoted to proving the existence of weak solutions to the one dimensional problem(Huang and Gamba, 2023)[1].

Our strategy can be summarized as follows. Firstly, we generalize the trick of Ivanov and Rudakov (1967) to show that, in one dimensional case, the system is equivalent to a porous medium equation with nonlinear source terms. This allows us to use existing techniques from that field. In particular, the proposed proof is inspired by the book of Vázquez (2007). The basic idea is to analyze a series of approximate problems with parameter n and try to establish regularity estimates uniform in n. Then it is possible to pass the limit to infinity.

In order to pass the limit, the strong convergence of the gradient term is necessary. And that is the most challenging part. Similar problems has been tackled

[1]Huang, K. and Gamba, I.M., 2023. Existence of global weak solutions to quasilinear theory for electrostatic plasmas. arXiv preprint arXiv:2304.12430. The dissertator's contribution includes proposing the proof and writing the article.

by Abdellaoui et al. (2015), where the nonlinear source contains a gradient term to some power. A significant difference is that they require positive source, while in our problem, the gradient square term has a minus sign. Nevertheless, their proof for a.e. convergence can still be adopted if we can find an alternative way to prove the prerequisites.

The chapter is organized as follows. In Section 6.2, we derive the equivalent porous medium equation and state the main result. In Section 6.3, we resort to the book of Ladyzhenskaya et al. (1968) for the maximum principle and the well-posedness of the strictly coercive approximate problems with parameter n. Section 6.4 aims at proving the regularity estimates uniform in parameter n, which paves the way to convergence results. In Section 6.5, the a.e. convergence of gradient term is proved, using the technique from Abdellaoui et al. (2015). The proof of main theorem is contained in Section 6.6.

6.2 From kinetic system to porous medium equation

The quasilinear system for particle-wave interaction contains two equations, as there are two unknown functions: the particle probability density function $f(\mathbf{p},t)$: $\mathbb{R}^d_p \times \mathbb{R}^+ \to [0,+\infty)$ and the wave spectral energy density $W(\mathbf{k},t)$: $\mathbb{R}^d_k \times \mathbb{R}^+ \to [0,+\infty)$. In general they are in the following form,

$$\begin{cases} \partial_t f(\mathbf{p},t) = \nabla_p \cdot \left(\left[\displaystyle\int_{\mathbb{R}^d_k} W(\mathbf{k},t)\Phi(\mathbf{p},\mathbf{k})d\mathbf{k} \right] \cdot \nabla_p f(\mathbf{p},t) \right), & \text{(Vedenov et al., 1961)} \\ \partial_t W(\mathbf{k},t) = \left[\displaystyle\int_{\mathbb{R}^d_p} (\nabla_p f(\mathbf{p},t)) \cdot \Phi(\mathbf{p},\mathbf{k}) \cdot (\nabla_p E(\mathbf{p}))\, d\mathbf{p} \right] W(\mathbf{k},t), & \text{(Landau, 1946)} \end{cases} \tag{6.1}$$

In electrostatic case,

$$\Phi(\mathbf{p},\mathbf{k}) = \frac{\pi e^2}{\epsilon_0} \left(\hat{\mathbf{k}} \otimes \hat{\mathbf{k}} \right) \delta(\omega(\mathbf{k};f) - \nabla_p E(\mathbf{p}) \cdot \mathbf{k}), \tag{6.2}$$

where $E(\mathbf{p}) = \frac{\mathbf{p}^2}{2m_e}$ is the kinetic energy of a single particle, and $\omega(\mathbf{k})$ is the dispersion relation of plasma waves, which depends on the medium, i.e. the particle *pdf*.

Remark 6.1. For magnetized plasma, we have

$$\Phi(\mathbf{p}, \mathbf{k}) = (\beta \otimes \beta)\, \mathcal{B}(\mathbf{p}, \mathbf{k}),$$

see Huang et al. (2023a).

In this chapter we pursue a weak solution in the following sense,

$$\begin{cases} -\int_0^T (f, \partial_t \zeta)_p \, dt - (f_0, \zeta_0)_p = \int_0^T \langle \nabla_p \cdot \Phi, fW \nabla_p \zeta \rangle_{pk} dt + \int_0^T \langle \Phi, fW \nabla_p \nabla_p \zeta \rangle_{pk} dt, \\ -\int_0^T (W, \partial_t \xi)_k \, dt - (W_0, \xi_0)_k = -\int_0^T \langle \nabla_p \cdot \Phi, fW \xi \nabla_p E \rangle_{pk} dt - \int_0^T \langle \Phi, fW \xi \nabla_p \nabla_p E \rangle_{pk} dt. \end{cases} \tag{6.3}$$

For scenarios of interest in physics, the Bohm-Gross dispersion relation is widely used to model warm plasmas,

$$\omega(\mathbf{k}; f) = \left(1 + 3\lambda_D^2(f)\mathbf{k}^2\right)^{1/2} \omega_{pe},$$

where the Debye length λ_D is proportional to the thermal speed, i.e. the variance of particle *pdf*, and the constant ω_{pe} is the plasma frequency.

Assuming that the diffusion process rarely affects the particles' temperature, it is reasonable to write the following approximation,

$$\omega(\mathbf{k}) = \left(1 + 3\lambda_D^2(f_0)\mathbf{k}^2\right)^{1/2} \omega_{pe}.$$

The most important feature of the system lies in the fact that particles and waves interact through the absorption/emission kernel, which contains a Dirac delta. In physics this is called "resonance", since particles with certain momentum only interact with waves with some particular wave vectors. The one dimensional plasma is special, because in this case each momentum $\mathbf{p}$ corresponds to only one wave vector $\mathbf{k}$, and vice versa. This is the reason that the following trick is feasible.

Normalize all the quantities with time unit $\frac{1}{\omega_{pe}}$, mass unit m_e and length unit $\frac{c}{\omega_{pe}}$, The equation in one dimensional case reads,

$$\begin{cases} \partial_t f = \partial_p \left(\left[\int_{\mathbb{R}_k} W \delta \left(\left(1+\lambda^2 k^2\right)^{1/2} - pk \right) dk \right] \partial_p f \right), \\ \partial_t W = \left[\int_{\mathbb{R}_p} \left(\partial_p f\right) p \delta \left(\left(1+\lambda^2 k^2\right)^{1/2} - pk \right) dp \right] W, \end{cases} \tag{6.4}$$

where $\lambda = \frac{\sqrt{3}\omega_{pe}\lambda_D}{c}$.

We define the distributions in (6.3) as follows,

$$\begin{aligned} \langle \delta \left(\left(1+\lambda^2 k^2\right)^{1/2} - pk \right), y(p)z(k) \rangle_{pk} &\coloneqq \int_{(-\infty,-\lambda)\cup(\lambda,+\infty)} \frac{|p|}{p^2-\lambda^2} y(p) z \left(\frac{\operatorname{sgn}(p)}{\sqrt{p^2-\lambda^2}} \right) dp, \\ \langle \partial_p \left(\delta \left(\left(1+\lambda^2 k^2\right)^{1/2} - pk \right) \right), y(p)z(k) \rangle_{pk} &\coloneqq \int_{(-\infty,-\lambda)\cup(\lambda,+\infty)} y(p) \partial_p \left(\frac{|p|}{p^2-\lambda^2} z(\frac{\operatorname{sgn}(p)}{\sqrt{p^2-\lambda^2}}) \right) dp. \end{aligned} \tag{6.5}$$

Note that the resonance condition

$$\left(1+\lambda^2 k^2\right)^{1/2} - pk = 0,$$

is equivalent to

$$p = \operatorname{sgn}(k) \left(k^{-2} + \lambda^2\right)^{1/2},$$

and

$$k = \operatorname{sgn}(p) \left(p^2 - \lambda^2\right)^{-1/2}.$$

As a generalization of the trick in Ivanov and Rudakov (1967), we introduce the following auxiliary function $u(p) : (-\infty, -\lambda) \cup (\lambda, \infty) \to \mathbb{R}^+ \cup \{0\}$,

$$u(p,t) \coloneqq \left(p^2-\lambda^2\right)^{-3/2} W \left(\operatorname{sgn}(p) \left(p^2-\lambda^2\right)^{-1/2}, t \right). \tag{6.6}$$

Without loss of generality, consider the positive half domain, i.e. $p \in (\lambda, +\infty)$. The kinetic system (6.4) is formally equivalent to

$$\begin{cases} \partial_t f = \partial_p \left(p \left(p^2 - \lambda^2 \right)^{1/2} u \partial_p f \right), \\ \partial_t u = p \left(p^2 - \lambda^2 \right)^{1/2} \left(\partial_p f \right) u. \end{cases} \tag{6.7}$$

Denote $u(p,0)$ as $\varphi_0(p)$, note that

$$\partial_t f = \partial_p(\partial_t u) \Rightarrow f = \partial_p(u - \varphi_0) + f_0.$$

Substitute the above identity into the second row of (6.7), it follows that

$$\partial_t u = p \left(p^2 - \lambda^2 \right)^{1/2} \partial_p \left(\partial_p u - \partial_p \varphi_0 + f_0 \right) u(p,t) = \gamma u \partial_p^2 u + g_0 u, \tag{6.8}$$

where $g_0(p) \coloneqq p \left(p^2 - \lambda^2 \right)^{1/2} \partial_p \left(f_0 - \partial_p \varphi_0 \right)$ and $\gamma(p) \coloneqq p\sqrt{p^2 - \lambda^2}$.

The above equation can also be written in the following divergence form,

$$\partial_t u = \partial_p \left(\frac{\gamma}{2} \partial_p u^2 \right) - \gamma \left(\partial_p u \right)^2 - \frac{\partial_p \gamma}{2} \partial_p u^2 + g_0 u. \tag{6.9}$$

Note that once we have a solution u to the above equation, both particle *pdf* f and wave *sed* W are formally determined. Inspired by the above formal derivation, we present a rigorous statement in the following theorem.

Theorem 6.1. *Given initial condition,*

$$\begin{cases} f(p,0) = f_0(p), & p \in \mathbb{R}_p, \\ W(k,0) = W_0(k), & k \in \mathbb{R}_k, \end{cases}$$

where $f_0 \in C^\infty(\mathbb{R}_p)$ *and* $W_0 \in C^\infty(\mathbb{R}_k)$.

Define the domain for auxiliary function as follows,

$$\begin{aligned} \Omega^* &\coloneqq (\lambda, +\infty), \\ Q_T^* &\coloneqq \Omega^* \times (0,T). \end{aligned} \tag{6.10}$$

In addition, define $\varphi_0 : \Omega^* \to [0, +\infty)$ *and* $g_0 : \Omega^* \to \mathbb{R}$,

$$\begin{aligned} \varphi_0(p) &\coloneqq \left(p^2 - \lambda^2 \right)^{-\frac{3}{2}} W_0 \left(\frac{1}{\sqrt{p^2 - \lambda^2}} \right), \\ g_0(p) &\coloneqq p\sqrt{p^2 - \lambda^2} \partial_p \left(f_0(p) - \partial_p \varphi_0(p) \right). \end{aligned}$$

Suppose there is a solution $u \in L^2(0,T;W_0^{1,2}(\Omega^*))$ *such that, for any* $\eta \in \mathcal{V} \coloneqq \{\eta \in C^1(Q_T^*) \cap L^\infty(Q_T^*) : \eta(\cdot,T) = 0\}$*, the following identity holds,*

$$-(u,\partial_t\eta)_{Q_T^*} + \left(\frac{\gamma}{2}\partial_p u^2, \partial_p\eta\right)_{Q_T^*} = \left(-\gamma(\partial_p u)^2,\eta\right)_{Q_T^*} + \left(-\left(\frac{\partial_p\gamma}{2}\right)\partial_p u^2,\eta\right)_{Q_T^*} + (g_0 u,\eta)_{Q_T^*} + (\varphi_0,\eta(p,0))_{\Omega^*}.$$

Define

$$W(k,t) \coloneqq \begin{cases} \frac{1}{k^3}u\left(\frac{\left(1+\lambda^2k^2\right)^{1/2}}{k},t\right), & k \in (0,+\infty) \\ 0, & \textit{otherwise} \end{cases}$$
$$f(p,t) \coloneqq f_0(p) + \begin{cases} \partial_p u(p,t) - \partial_p\varphi_0(p), & p \in (\lambda,+\infty) \\ 0, & \textit{otherwise} \end{cases} \tag{6.11}$$

If there exists a positive constant ε *such that* $supp\,(W(\cdot,t)) \subset (\varepsilon,\frac{1}{\varepsilon})$ *for any* $t \in (0,T)$*, then* $f(p,t)$ *and* $W(k,t)$ *satisfy Equation (6.3) for any*

$$\zeta \in \left\{\zeta(p,t) \in C^\infty\left(\mathbb{R}_p \times [0,T]\right) : \zeta(\cdot,t) \in C_c^\infty\left(\mathbb{R}_p\right)\ \forall t \in [0,T] \textit{ and } \zeta(\cdot,T) = 0\right\},$$

and

$$\xi \in \left\{\xi(k,t) \in C^\infty\left(\mathbb{R}_k \times [0,T]\right) : \xi(\cdot,t) \in C_c^\infty\left(\mathbb{R}_k\right)\ \forall t \in [0,T] \textit{ and } \xi(\cdot,T) = 0\right\}.$$

Proof. By definition of $f(p,t)$ in (6.11),

$$\begin{aligned}
&-\int_0^T (f,\partial_t\zeta)_p\,dt - (f_0,\zeta_0)_p \\
&= -\int_0^T\int_\lambda^{+\infty} \left[\left(\partial_p u - \partial_p\varphi_0\right)\partial_t\zeta\right] dpdt \\
&= \int_0^T\int_\lambda^{+\infty} \left[\left(u-\varphi_0\right)\partial_t\partial_p\zeta\right] dpdt - \int_0^T \left[\left(u-\varphi_0\right)\partial_p\zeta\right] dt\Big|_\lambda^{+\infty} \\
&= \int_0^T\int_\lambda^{+\infty} \left[\left(u-\varphi_0\right)\partial_t\partial_p\zeta\right] dpdt. \\
&= (u,\partial_t\left(\partial_p\zeta\right))_{Q_T^*} + (\varphi_0,\partial_p\zeta_0)_{\Omega^*} \\
&= \left(\frac{\gamma}{2}\partial_p u^2,\partial_p^2\zeta\right)_{Q_T^*} + \left(\gamma(\partial_p u)^2,\partial_p\zeta\right)_{Q_T^*} + \left(\left(\frac{\partial_p\gamma}{2}\right)\partial_p u^2,\partial_p\zeta\right)_{Q_T^*} - (g_0 u,\partial_p\zeta)_{Q_T^*}.
\end{aligned}$$

Meanwhile, by definition of the distributions in (6.5),

$$\begin{aligned}&\int_0^T \langle \nabla_p \cdot \Phi, fW\nabla_p\zeta\rangle_{pk}dt + \int_0^T \langle \Phi, fW\nabla_p\nabla_p\zeta\rangle_{pk}dt\\ =&\int_0^T \int_{+\lambda}^{+\infty} \left[f\partial_p\left(\frac{p}{p^2-\lambda^2}W(\frac{1}{\sqrt{p^2-\lambda^2}},t)\partial_p\zeta\right)\right]dpdt.\end{aligned}$$

Note that $W(\frac{1}{\sqrt{p^2-\lambda^2}},t) = (p^2-\lambda^2)^{\frac{3}{2}}\,u(p,t)$ on $(\lambda,+\infty)$, hence

$$\begin{aligned}&\int_0^T \langle \nabla_p \cdot \Phi, fW\nabla_p\zeta\rangle_{pk}dt + \int_0^T \langle \Phi, fW\nabla_p\nabla_p\zeta\rangle_{pk}dt\\ =&\int_0^T \int_{+\lambda}^{+\infty} \left[f\partial_p\left(p\sqrt{p^2-\lambda^2}u(p,t)\partial_p\zeta\right)\right]dpdt\\ =&\int_0^T \int_{+\lambda}^{+\infty} \left[f\partial_p\left(\gamma(p)u(p,t)\partial_p\zeta\right)\right]dpdt.\end{aligned}$$

Then it can be verified that since $\partial_p\zeta(p,t)\in\mathcal{V}$,

$$-\int_0^T (f,\partial_t\zeta)_p\,dt - (f_0,\zeta_0)_p = \int_0^T \langle \nabla_p \cdot \Phi, fW\nabla_p\zeta\rangle_{pk}dt + \int_0^T \langle \Phi, fW\nabla_p\nabla_p\zeta\rangle_{pk}dt.$$

Next, consider the equation for $W(k,t)$. By definition of $W(k,t)$ in (6.11),

$$\begin{aligned}\int_0^T\int_{-\infty}^{+\infty}\left[W\partial_t\xi\right]dtdk &= \int_0^T\int_0^{+\infty}\left[\frac{1}{k^3}u\left(\frac{\sqrt{1+\lambda^2k^2}}{k},t\right)\partial_t\xi(k,t)\right]dkdt\\ &= \int_0^T\int_\lambda^{+\infty}\left[u\left(p,t\right)\partial_t\left(p\xi\left(\frac{1}{\sqrt{p^2-\lambda^2}},t\right)\right)\right]dpdt.\end{aligned}$$

Meanwhile, by definition of the distributions in (6.5),

$$\begin{aligned}&-\int_0^T \langle \nabla_p \cdot \Phi, fW\xi\nabla_pE\rangle_{pk}dt - \int_0^T \langle \Phi, fW\xi\nabla_p\nabla_pE\rangle_{pk}dt\\ =&-\int_0^T \int_\lambda^{+\infty}\left[f(p)\partial_p\left(p\frac{p}{p^2-\lambda^2}W\left(\frac{1}{\sqrt{p^2-\lambda^2}},t\right)\xi\left(\frac{1}{\sqrt{p^2-\lambda^2}},t\right)\right)\right]dpdt\\ =&-\int_0^T \int_\lambda^{+\infty}\left[f(p)\partial_p\left(\gamma u\left(p\xi\left(\frac{1}{\sqrt{p^2-\lambda^2}},t\right)\right)\right)\right]dpdt.\end{aligned}$$

Analogous to the procedure for $f(p,t)$, since $p\cdot\xi\left(\frac{1}{\sqrt{p^2-\lambda^2}},t\right)\in\mathcal{V}$, it can be verified that

$$-\int_0^T (W,\partial_t\xi)_k\,dt-(W_0,\xi_0)_k=-\int_0^T\langle\nabla_p\cdot\Phi,fW\xi\nabla_pE\rangle_{pk}dt-\int_0^T\langle\Phi,fW\xi\nabla_p\nabla_pE\rangle_{pk}dt.$$

□

Remark 6.2. It takes no extra effort to extend the above conclusion to the negative half domain $p\in(-\infty,-\lambda)$ by concatenating solutions. In that case, instead of requiring $\operatorname{supp}(W(\cdot,t))\subset(\varepsilon,\frac{1}{\varepsilon})$, we need $\operatorname{supp}(W(\cdot,t))\subset(-\frac{1}{\varepsilon},-\varepsilon)\cup(\varepsilon,\frac{1}{\varepsilon})$.

Now it remains to show the existence of u. The main result of the chapter can be stated as follows.

Note that in the rest of the chapter, to keep it consistent with existing literature in mathematics, we will use x instead of p.

Theorem 6.2. *If $g_0(x)\in C^\infty(\Omega^*)\cap L^\infty(\Omega^*)$, $\varphi_0(x)\in C_0^\infty(\Omega^*)$, then there exists a non-negative weak solution to Equation(6.9), $u(x,t)\in L^q(0,T;W_0^{1,q}(\Omega^*))$ with $q\in[1,+\infty)$ such that, for any $\eta(x,t)\in\mathcal{V}:=\{\eta\in C^1(Q_T^*)\cap L^\infty(Q_T^*):\eta(\cdot,T)=0\}$, the following identity holds,*

$$\begin{aligned}-(u,\partial_t\eta)_{Q_T^*}+\left(\frac{\gamma}{2}\partial_x u^2,\partial_x\eta\right)_{Q_T^*}&=\left(-\gamma(\partial_x u)^2,\eta\right)_{Q_T^*}+\left(-\left(\frac{\partial_x\gamma}{2}\right)\partial_x u^2,\eta\right)_{Q_T^*}\\&\quad+(g_0u,\eta)_{Q_T^*}+(\varphi_0,\eta(x,0))_{\Omega^*}.\end{aligned}\tag{6.12}$$

6.3 Lifted extension of approximate solutions

The equation (6.9) is difficult to tackle due to its degeneracy. Therefore we consider a series of approximate problems on a cut-off domain first. They are arbitrarily close to the original problem, but each one of them is strictly coercive, which ensures the existence of classical solutions. Furthermore, these approximate solutions have enough regularity, allowing us to test with various functions and to obtain bounds that are uniform in the parameter n.

Given initial condition $\varphi_0 \in C_0^\infty(\Omega^*)$, we choose a cut-off domain $\Omega = (a, b)$ where

$$\begin{aligned} a &= \frac{1}{2}\inf\left(\text{supp}\varphi_0\right) > \lambda, \\ b &= 2\sup\left(\text{supp}\varphi_0\right) < +\infty, \end{aligned} \tag{6.13}$$

apparently $\varphi_0 \in C_0^\infty(\Omega)$.

Consider the following approximate problem $(\mathcal{S}_n)$ of Equation(6.9) on the cut-off domain $\Omega = (a, b)$,

$$\begin{cases} \partial_t u_n = \gamma \mathcal{P}_n(u_n)\partial_x^2 u_n + g_0 u_n, & \text{in } \Omega \\ u_n(x,t) = 0, & \text{on } \partial\Omega \\ u_n(x,0) = \varphi_0(x), & \text{in } \Omega \end{cases} \tag{$\mathcal{S}_n$}$$

where $\mathcal{P}_n \in C^\infty(\mathbb{R})$ is a family of functions with the following properties,

1. $\mathcal{P}_n(y) \geq \frac{1}{2n}, \forall y \in \mathbb{R}$.
2. $\mathcal{P}_n(y) = y + \frac{1}{n}, \forall y \in \mathbb{R}^+$.
3. $\mathcal{P}_n'(y) \geq 0$.

The following maximum principle can be found in Theorem 2.1, Chapter I of Ladyzhenskaya et al. (1968). Note that the bounds are independent of the parameter n.

Theorem 6.3. *(maximum principle) If u_n is a classical solution to the approximate problem $(\mathcal{S}_n)$, then $u_n(x,t)$ satisfies the maximum principle on $\overline{Q_T}$:*

$$\begin{cases} u_n(x,t) \geq 0 \\ u_n(x,t) \leq max(\varphi_0)\exp(\max(|g_0|)t) \end{cases}$$

For the existence of classical solution to the approximate problems $(\mathcal{S}_n)$, we refer the readers to Theorem 6.1, Chapter V of Ladyzhenskaya et al. (1968).

Theorem 6.4 (**classical solution**). *For any $T > 0$, the problem ($\mathcal{S}_n$) admits an unique classical solution $u_n(x,t)$ on $\overline{Q_T}$. Moreover, $u_n(x,t)$ belongs to Hölder space $\mathcal{H}^{2+\beta,(2+\beta)/2}(\overline{Q_T}) \cap \mathcal{H}^{3+\beta,(3+\beta)/2}(Q_T)$.*

Remark 6.3. The Hölder space $\mathcal{H}^{2+\beta,(2+\beta)/2}(\overline{Q_T})$ consists of all the functions that are β-Hölder continuous in $\overline{Q_T}$, together with all derivatives of the form $D_t^r D_x^s$ for $2r + s < 2$. Same for $\mathcal{H}^{3+\beta,(3+\beta)/2}(Q_T)$.

Proof. To begin with, write the equation in divergence form.

$$\begin{aligned}\mathcal{L}_n u \equiv& \partial_t u_n - \gamma \mathcal{P}_n(u_n)\partial_x^2 u_n - g_0 u_n \\ =& \partial_t u_n - \partial_x\left(\gamma \mathcal{P}_n(u_n)\partial_x u_n\right) + \gamma \mathcal{P}_n'(u_n)\left(\partial_x u_n\right)^2 + \left(\partial_x \gamma\right)\mathcal{P}_n(u_n)\partial_x u_n - g_0 u_n.\end{aligned}$$

Define

$$\varphi(x,t) \coloneqq \varphi_0(x) + \left[\gamma \mathcal{P}_n'(\varphi_0)\partial_x^2 \varphi_0 + g_0 \varphi_0\right] t.$$

Since $\varphi_0 \in C_0^\infty(\Omega)$, the initial-boundary condition can be written as

$$u_n|_{\Gamma_T} = \varphi|_{\Gamma_T}$$

where $\Gamma_T = \partial\Omega \times [0,T] \cup \{(x,t) : x \in \Omega, t = 0\}$.

In accordance with the notation of Ladyzenskaja, define

$$\begin{aligned} a_1(x,t,u,p) &\coloneqq \gamma(x)\mathcal{P}_n(u)p, \\ a(x,t,u,p) &\coloneqq \gamma(x)\mathcal{P}_n'(u)p^2 + \partial_x\gamma(x)\mathcal{P}_n(u)p - g_0(x)u, \\ A(x,t,u,p) &\coloneqq -g_0(x)u. \end{aligned}$$

Note that the p variable here is just a notation from Ladyzenskaja representing $\partial_x u$, not particle momentum. The conditions in Ladyzenskaja's theorem can be easily verified. See Section 6.7.

□

By Theorem 6.3, u_n is always non-negative, therefore by the second property of function $\mathcal{P}_n$, the problem $(\mathcal{S}_n)$ in divergence form is as follows,

$$\begin{cases} \partial_t u_n = \partial_x \left(\gamma(u_n + \frac{1}{n}) \partial_x u_n \right) - \gamma \left(\partial_x u_n \right)^2 - \dfrac{\partial_x \gamma}{2} \partial_x (u_n + \frac{1}{n})^2 + g_0 u_n, & \text{in } \Omega \\ u_n(x) = 0, & \text{on } \partial\Omega \\ u_n(x,0) = \varphi_0(x), & \text{in } \Omega \end{cases} \tag{6.14}$$

Define the following continuous lifted extension of u_n on $\Omega^* = (\lambda, +\infty)$:

$$\tilde{u}_n(x,t) \coloneqq \begin{cases} \frac{1}{n}, & x \in (\lambda, a) \\ u_n + \frac{1}{n}, & x \in \Omega = (a,b) \\ \frac{1}{n} h(x), & x \in (b, +\infty) \end{cases} \tag{6.15}$$

with the tail $h(x)$ s.t. $h(b) = 1$ decaying fast enough. In particular,

$$h(x) = e^{-(x-b)} \tag{6.16}$$

is a suitable choice, in which case the derivative of lifted extension $\tilde{u}_n$ is:

$$\partial_x \tilde{u}_n(x,t) = \begin{cases} 0, & x \in (\lambda, a) \\ \partial_x u_n, & x \in \Omega = (a,b) \\ -\frac{1}{n} h(x), & x \in (b, +\infty) \end{cases} \tag{6.17}$$

Note that $\tilde{u}_n$ is strictly positive, and it solves the following problem on the cut-off domain $\Omega = (a,b)$,

$$\begin{cases} \partial_t \tilde{u}_n = \partial_x \left(\gamma \tilde{u}_n \partial_x \tilde{u}_n \right) - \gamma \left(\partial_x \tilde{u}_n \right)^2 - \dfrac{\partial_x \gamma}{2} \partial_x \tilde{u}_n^2 + g_0 \left(\tilde{u}_n - \dfrac{1}{n} \right), & \text{in } \Omega \\ \tilde{u}_n(x,t) = \dfrac{1}{n}, & \text{on } \partial\Omega \\ \tilde{u}_n(x,0) = \varphi_0(x) + \dfrac{1}{n}, & \text{in } \Omega \end{cases} \tag{6.18}$$

Obviously, $\tilde{u}_n$ and $\partial_x \tilde{u}_n$ converge to zero uniformly on $\Omega^* - \Omega$. In the next section, we are going to prove the existence of a limit function u on $\Omega = (a,b)$ through regularity estimates uniform in n.

6.4 Regularity estimates on approximate solutions

In order to prove a.e. convergence of the gradient term in the next section, several regularity estimates are of necessity. In the work of Abdellaoui et al. (2015), the authors used existing results for an elliptic-parabolic problem with measure data. However for the problem we are dealing with, the nonlinear source term is not "measure data", thus it calls for a different approach.

To begin with, we introduce the following estimate on the spatial derivative of lifted extension $\tilde{u}_n$.

According to Ladyzhenskaya et al. (1968), our assumption on the smoothness of data yields enough smoothness of solutions, even the third order derivative is Hölder continuous. And that allows us to study a parabolic equation for the first derivative $\partial_x u_n$, which renders the following regularity estimates.

Proposition 6.5 (**Uniform $W^{1,q}(\Omega^*)$ estimate of the lifted extension**). *If $\tilde{u}_n$ are the lifted extension of classical solutions u_n as defined in (6.15), then their gradient in space are uniformly bounded in $L^q(\Omega^*)$,*

$$\sup_{t\in[0,T]} \|\partial_x \tilde{u}_n\|_{L^q(\Omega^*)} \leq C_1(\varphi_0, g_0, T, q),\ \forall q \in [1, +\infty).$$

Consequently,

$$\|\tilde{u}_n\|_{L^\infty(0,T;W^{1,q}(\Omega^*))} \leq C_2(\varphi_0, g_0, T, q),\ \forall q \in [1, +\infty). \tag{6.19}$$

Proof. To begin with, consider the non-divergence form in $(\mathcal{S}_n)$,

$$\partial_t u_n - \gamma \left(u_n + \frac{1}{n}\right) \partial_x^2 u_n - g_0 u_n = 0, \tag{6.20}$$

Note that $\partial_x \tilde{u}_n = \partial_x u_n$ in cut-off domain $\Omega = (a, b)$ as given in Equation(6.17).

Let $z_n = \partial_x u_n$, then by Theorem 6.4, $z_n \in \mathcal{H}^{1+\beta,(1+\beta)/2}(\overline{Q_T}) \cap \mathcal{H}^{2+\beta,(2+\beta)/2}(Q_T)$. Taking first derivative on both sides of Equation(6.20), every term is still continuous. Therefore z_n satisfies the following linear parabolic equation, if we regard u_n as data.

$$\partial_t z_n - \partial_x \left(\gamma \left(u_n + \frac{1}{n}\right) \partial_x z_n\right) - g_0 z_n - u_n \partial_x g_0 = 0.$$

In addition, from Equation(6.20) and the boundary condition for u_n, it can be derived that

$$\partial_x z_n = \partial_x^2 u_n = \frac{\partial_t u_n - g_0 u_n}{\gamma\left(u_n + \frac{1}{n}\right)} = 0 \text{ on } \partial\Omega.$$

To summarize, z_n satisfies the following equations,

$$\begin{cases} \partial_t z_n - \partial_x \left(\gamma\left(u_n + \dfrac{1}{n}\right)\partial_x z_n\right) - g_0 z_n - u_n \partial_x g_0 = 0, & \text{in } \Omega \\ \partial_x z_n = 0, & \text{on } \partial\Omega \\ z_n(x,0) = \partial_x \varphi_0(x) \end{cases}$$

Test the equation with z_n^{2l+1} and perform integration by parts to obtain

$$\left(\partial_t z_n, z_n^{2l+1}\right)_\Omega + \left(\gamma\left(u_n + \frac{1}{n}\right)\partial_x z_n, (2l+1) z_n^{2l}\partial_x z_n\right)_\Omega = \left(g_0 z_n, z_n^{2l+1}\right)_\Omega + \left(u_n \partial_x g_0, z_n^{2l+1}\right)_\Omega. \tag{6.21}$$

Since the second term on the left hand side is non-negative, the following inequality holds,

$$\frac{1}{2l+2}\frac{d}{dt}\left(\int_\Omega z_n^{2l+2}\right) \le \int_\Omega g_0 z_n^{2l+2} + \int_\Omega u_n \partial_x g_0 z_n^{2l+1}.$$

Again, use the maximum principle for u_n in Theorem 6.3, and by the assumption on the data g_0,

$$\frac{d}{dt}\left(\int_\Omega z_n^{2l+2}\right) \le C_1(\varphi_0, g_0, T)\int_\Omega z_n^{2l+2} + C_2(\varphi_0, g_0, T)\int_\Omega |z_n|^{2l+1}.$$

Let $I = \int_\Omega z_n^{2l+2}$, by Hölder's inequality, the above is equivalent to

$$\frac{d}{dt} I \le C_1 I + C_3 I^{\frac{2l+1}{2l+2}}.$$

Apply Young's inequality on the second term to obtain

$$\frac{d}{dt}(I + C_5) \le C_4 (I + C_5).$$

Therefore by Grönwall's lemma we have

$$I \le (I_0 + C_5)\exp(C_4 t) - C_5 \le (I_0 + C_5)\exp(C_4 T) - C_5.$$

Consequently,

$$\sup_{t\in[0,T]} \|\partial_x u_n\|_{L^{2l}(\Omega)} \leq C_6(\varphi_0, g_0, T, l),\ l = 1, 2, \cdots$$

Note that the cut-off domain $\Omega = (a, b)$ defined in Equation(6.13) is bounded, therefore

$$\sup_{t\in[0,T]} \|\partial_x u_n\|_{L^q(\Omega)} \leq C_6(\varphi_0, g_0, T, q),\ \forall q \in [1, +\infty).$$

By definition we have,

$$\begin{aligned}
\|\partial_x \tilde{u}_n(\cdot, t)\|^q_{L^q(\Omega^*)} &= \int_{\Omega^*} |\partial_x \tilde{u}_n(\cdot, t)|^q dx \\
&= \int_0^a 0 \cdot dx + \int_\Omega |\partial_x u_n(\cdot, t)|^q dx + \int_b^\infty \left|\frac{1}{n}\partial_x h\right|^q dx \\
&= \|\partial_x u_n(\cdot, t)\|^q_{L^q(\Omega)} + \int_b^\infty \left|\frac{1}{n}\partial_x h\right|^q dx,
\end{aligned}$$

hence

$$\sup_{t\in[0,T]} \|\partial_x \tilde{u}_n\|_{L^q(\Omega^*)} \leq C_7(\varphi_0, g_0, T, q),\ \forall q \in [1, +\infty).$$

In addition, it can be easily verified that

$$\sup_{t\in[0,T]} \|\tilde{u}_n\|_{L^q(\Omega^*)} \leq C_8(\varphi_0, g_0, T, q),\ \forall q \in [1, +\infty).$$

The above two bounds lead to the following uniform estimate,

$$\|\tilde{u}_n\|_{L^\infty(0,T;W^{1,q}(\Omega^*))} \leq C_9(\varphi_0, g_0, T, q),\ \forall q \in [1, +\infty).$$

□

Corollary 6.6. *If $\tilde{u}_n$ is the lifted extension of classical solution u_n as defined in (6.15), then the following inequality holds,*

$$\|\partial_x \tilde{u}_n^2\|_{L^2(Q_T^*)} \leq C(\varphi_0, g_0, T),$$

where the constant C is independent of n.

Proof. The conclusion is a straight forward consequence of the maximum principle in Theorem 6.3 combined with Proposition 6.5. □

Proposition 6.7 (**Uniform $W^{1,2}(Q_T^*)$ estimate of $\tilde{u}_n^2$**). *The sequence $\tilde{u}_n^2$ is uniformly bounded in $W^{1,2}(Q_T^*)$.*

Proof. To begin with, we estimate the $L^2(Q_T^*)$ norm of the temporal derivative.

Take Equation(6.18) and test it with $\partial_t \tilde{u}_n^2$. Integrate by parts in x, and the trace integral vanishes because $\partial_t \tilde{u}_n^2 = 0$ on $\partial\Omega$, hence

$$\left(\partial_t \tilde{u}_n, \partial_t \tilde{u}_n^2\right)_{Q_T} = -\left(\frac{\gamma}{2}\partial_x \tilde{u}_n^2, \partial_t \partial_x \tilde{u}_n^2\right)_{Q_T} + \left(-\gamma\left(\partial_x \tilde{u}_n\right)^2 - \frac{\partial_x \gamma}{2}\partial_x \tilde{u}_n^2 + g_0\left(\tilde{u}_n - \frac{1}{n}\right), \partial_t \tilde{u}_n^2\right)_{Q_T}.$$

Applying the fundamental theorem of calculus with respect to $t \in (0,T)$ on the first term of the right hand side,

$$\begin{aligned}\left(\partial_t \tilde{u}_n, \partial_t \tilde{u}_n^2\right)_{Q_T} = &-\left(\frac{\gamma}{4}, \left(\partial_x \tilde{u}_n^2(T)\right)^2\right)_{\Omega} + \left(\frac{\gamma}{4}, \left(\partial_x \tilde{u}_n^2(0)\right)^2\right)_{\Omega} \\ &+ \left(-\gamma\left(\partial_x \tilde{u}_n\right)^2 - \frac{\partial_x \gamma}{2}\partial_x \tilde{u}_n^2 + g_0\left(\tilde{u}_n - \frac{1}{n}\right), \partial_t \tilde{u}_n^2\right)_{Q_T},\end{aligned}$$

The goal is to bound $(\partial_t \tilde{u}_n^2, \partial_t \tilde{u}_n^2)_{Q_T}$, however the left hand side is $(\partial_t \tilde{u}_n, \partial_t \tilde{u}_n^2)_{Q_T}$.

Note that by maximum principle, there exists some constant C_1 such that $\tilde{u}_n \in (\frac{1}{2n}, \frac{C_1}{2})$, therefore,

$$\left(\partial_t \tilde{u}_n^2\right)^2 = 4\tilde{u}_n^2\left(\partial_t \tilde{u}_n\right)^2 \le C_1 \cdot 2\tilde{u}_n\left(\partial_t \tilde{u}_n\right)^2 = C_1\left(\partial_t \tilde{u}_n\right)\left(\partial_t \tilde{u}_n^2\right).$$

Integrate both sides on Q_T,

$$\begin{aligned}\|\partial_t \tilde{u}_n^2\|_{L^2(Q_T)}^2 \le &C_1\left(\partial_t \tilde{u}_n, \partial_t \tilde{u}_n^2\right)_{Q_T} \\ = &C_1\left(-\left(\frac{\gamma}{4}, \left(\partial_x \tilde{u}_n^2(T)\right)^2\right)_{\Omega} + \left(\frac{\gamma}{4}, \left(\partial_x \tilde{u}_n^2(0)\right)^2\right)_{\Omega}\right) \\ &+ C_1\left(-\gamma\left(\partial_x \tilde{u}_n\right)^2 - \frac{\partial_x \gamma}{2}\partial_x \tilde{u}_n^2 + g_0\left(\tilde{u}_n - \frac{1}{n}\right), \partial_t \tilde{u}_n^2\right)_{Q_T} \\ \le &C_1\left(-\left(\frac{\gamma}{4}, \left(\partial_x \tilde{u}_n^2(T)\right)^2\right)_{\Omega} + \left(\frac{\gamma}{4}, \left(\partial_x \tilde{u}_n^2(0)\right)^2\right)_{\Omega}\right) \\ &+ C_1\|-\gamma\left(\partial_x \tilde{u}_n\right)^2 - \frac{\partial_x \gamma}{2}\partial_x \tilde{u}_n^2 + g_0\left(\tilde{u}_n - \frac{1}{n}\right)\|_{L^2(Q_T)} \cdot \|\partial_t \tilde{u}_n^2\|_{L^2(Q_T)},\end{aligned}$$

in which we used Hölder's inequality. Then use Young's inequality to bound the last term in the inequality above,

$$\begin{aligned}&\|-\gamma\left(\partial_x\tilde{u}_n\right)^2-\frac{\partial_x\gamma}{2}\partial_x\tilde{u}_n^2+g_0\left(\tilde{u}_n-\frac{1}{n}\right)\|_{L^2(Q_T)}\cdot\|\partial_t\tilde{u}_n^2\|_{L^2(Q_T)}\\ \leq&\frac{C_1}{2}\|-\gamma\left(\partial_x\tilde{u}_n\right)^2-\frac{\partial_x\gamma}{2}\partial_x\tilde{u}_n^2+g_0\left(\tilde{u}_n-\frac{1}{n}\right)\|_{L^2(Q_T)}^2+\frac{1}{2C_1}\|\partial_t\tilde{u}_n^2\|_{L^2(Q_T)}^2.\end{aligned}$$

It follows that

$$\begin{aligned}\|\partial_t\tilde{u}_n^2\|_{L^2(Q_T^*)}^2=\|\partial_t\tilde{u}_n^2\|_{L^2(Q_T)}^2\leq&2C_1\left(-\left(\frac{\gamma}{4},\left(\partial_x\tilde{u}_n^2(T)\right)^2\right)_\Omega+\left(\frac{\gamma}{4},\left(\partial_x\tilde{u}_n^2(0)\right)^2\right)_\Omega\right)\\&+C_1^2\left\|-\gamma\left(\partial_x\tilde{u}_n\right)^2-\frac{\partial_x\gamma}{2}\partial_x\tilde{u}_n^2+g_0\left(\tilde{u}_n-\frac{1}{n}\right)\right\|_{L^2(Q_T)}^2.\end{aligned}$$

By Theorem 6.3 and Proposition 6.5, the right hand side is uniformly bounded, hence $\|\partial_t\tilde{u}_n^2\|_{L^2(Q_T)}$ is uniformly bounded.

Meanwhile by Corollary 6.6, the spatial derivative $\|\partial_x\tilde{u}_n^2\|_{L^2(Q_T^*)}$ is also uniformly bounded, thus the result follows. □

The following lemma shows that convergence a.e. combined with uniform boundedness implies strong convergence.

Lemma 6.8 (**strong convergence from a.e. convergence**). *For a sequence v_n that is uniformly bounded in $L^4(Q_T)$, if v_n converges to $v\in L^4(Q_T)$ almost everywhere in Q_T, then v_n converges to v strongly in $L^2(Q_T)$.*

Proof. By Egorov's theorem, for any $\epsilon>0$, there exists a measurable set $S_\epsilon\subset Q_T$ such that $|S_\epsilon|\leq\epsilon$ and $v_n\to v$ uniformly in $Q_T\backslash S_\epsilon$. Therefore,

$$\begin{aligned}\int_{Q_T}|v_n-v|^2&=\int_{S_\epsilon}|v_n-v|^2+\int_{Q_T\backslash S_\epsilon}|v_n-v|^2\\&=\int_{Q_T}|v_n-v|^2\chi_{S_\epsilon}+\int_{Q_T}|v_n-v|^2\chi_{Q_T\backslash S_\epsilon}^2\end{aligned}$$

Apply Hölder's inequality on both terms,

$$\begin{aligned}\int_{Q_T} |v_n - v|^2 \leq & \||v_n - v|^2\|_{L^2(Q_T)} \cdot \|\chi_{S_\epsilon}\|_{L^2(Q_T)} + \left(\sup_{(x,t)\in Q_T\setminus S_\epsilon} |v_n - v|^2 \right) \cdot \|\chi_{Q_T\setminus S_\epsilon}\|_{L^1(Q_T)} \\ \leq & \|v_n - v\|^2_{L^4(Q_T)} \cdot \epsilon + C_1 \left(\sup_{(x,t)\in Q_T\setminus S_\epsilon} |v_n - v|^2 \right).\end{aligned}$$

Since $v_n \to v$ uniformly in $Q_T\setminus S_\epsilon$, taking the limit on both sides,

$$\limsup_{n\to\infty} \int_{Q_T} |v_n - v|^2 \leq \|v_n - v\|^2_{L^4(Q_T)} \cdot \epsilon.$$

As ϵ is arbitrary and $\|v_n - v\|_{L^4(Q_T)}$ is uniformly bounded, we conclude that v_n converges to v strongly in $L^2(Q_T)$.

□

Corollary 6.9. *For any given $q \in \mathbb{N}^+$, there exists a function $u \in L^q(0,T;W_0^{1,q}(\Omega^*))$ such that up to a subsequence,*

1. *$\tilde{u}_n$ converge to u strongly in $L^q(Q_T)$, for any $q \in \mathbb{N}^+$;*
2. *$\tilde{u}_n$ converge to u weakly in $L^q(0,T;W_0^{1,q}(\Omega^*))$, for any $q \in \mathbb{N}^+$;*
3. *$\partial_x\tilde{u}_n$ converge to $\partial_x u$ weakly in $L^2(Q_T)$.*

Proof. By compactness, Proposition 6.7 implies that, for any $q < \infty$, up to a subsequence, $\tilde{u}_n^2$ converge to some function y strongly in $L^q(Q_T)$. Up to a subsequence of that subsequence, $\tilde{u}_n$ converge to $\sqrt{y}$ almost everywhere in Q_T, Therefore by Lemma 6.8, $\tilde{u}_n$ converge to $u = \sqrt{y}$ strongly in $L^q(Q_T)$, for any $q < \infty$.

With inequality(6.19), use Banach-Alaoglu Theorem, there exists a function $u \in L^q(0,T;W_0^{1,q}(\Omega^*))$ such that up to a subsequence, $\tilde{u}_n$ converge to u weakly in $L^q(0,T;W_0^{1,q}(\Omega^*))$.

Since $\partial_x\tilde{u}_n$ is uniformly bounded in $L^2(Q_T)$, by Banach-Alaoglu Theorem, there exists a function $z \in L^2(Q_T)$ such that up to a subsequence, $\partial_x\tilde{u}_n$ converges to

z weakly in $L^2(Q_T)$. Since $\tilde{u}_n \to u$ a.e. in Q_T, it follows that $z = \partial_x u$ in the sense of distributions.

By taking subsequence of a subsequence, there must be a function u satisfying all the above conditions simultaneously. □

Now it remains to show that such a function u is indeed a weak solution to the problem. That requires $\partial_x \tilde{u}_n$ converging to $\partial_x u$ almost everywhere in Q_T, which will be elaborated in Theorem 6.13 of the next section. The proof of Theorem 6.13 will need a control of $\partial_x \left(\tilde{u}_n^{1-\theta}\right)$, therefore we present an auxiliary estimate as follows.

Proposition 6.10 (**Uniform weighted L^1 bound of $\partial_x \left(\tilde{u}_n^{1-\theta}\right)$**). *Let $\psi \in C_0^\infty(Q_T)$ be s.t. $\psi \geq 0$ in Q_T, then the sequence $\gamma\psi\tilde{u}_n^{-\theta}|\partial_x \tilde{u}_n|$ is uniformly bounded in $L^1(Q_T)$ for any $\theta \in (0, 1/2)$.*

Proof. Let $\psi \in C_0^\infty(Q_T)$ be s.t. $\psi \geq 0$ in Q_T. Test Equation(6.18) with $\psi\tilde{u}_n^{-\delta}$, where $\delta \in (0, 1)$. Since $\psi = 0$ on ∂Q_T, integrating by parts on x, the trace integral vanishes, it follows that,

$$\begin{aligned}
&\left(\partial_t \tilde{u}_n, \psi\tilde{u}_n^{-\delta}\right)_{Q_T} + \left(\gamma\tilde{u}_n\partial_x\tilde{u}_n, \partial_x\left(\psi\tilde{u}_n^{-\delta}\right)\right)_{Q_T} \\
&= \left(\partial_t \tilde{u}_n, \psi\tilde{u}_n^{-\delta}\right)_{Q_T} + \left(\gamma\tilde{u}_n\partial_x\tilde{u}_n, \tilde{u}_n^{-\delta}\left(\partial_x\psi\right)\right)_{Q_T} + \left(\gamma\tilde{u}_n\partial_x\tilde{u}_n, \psi\left(\partial_x\tilde{u}_n^{-\delta}\right)\right)_{Q_T} \\
&= \left(-\gamma\left(\partial_x\tilde{u}_n\right)^2, \psi\tilde{u}_n^{-\delta}\right)_{Q_T} + \left(-\frac{\partial_x\gamma}{2}\partial_x\tilde{u}_n^2, \psi\tilde{u}_n^{-\delta}\right)_{Q_T} + \left(g_0\left(\tilde{u}_n - \frac{1}{n}\right), \psi\tilde{u}_n^{-\delta}\right)_{Q_T}.
\end{aligned}$$

Thus, simplifying and rearranging each inner product term in the equation above to obtain

$$\begin{aligned}
&\left(\tilde{u}_n^{-\delta}\partial_t\tilde{u}_n, \psi\right)_{Q_T} + \left(\tilde{u}_n^{1-\delta}\partial_x\tilde{u}_n, \gamma\partial_x\psi\right)_{Q_T} + (-\delta)\left(\tilde{u}_n^{-\delta}\left(\partial_x\tilde{u}_n\right)^2, \gamma\psi\right)_{Q_T} \\
&= -\left(\tilde{u}_n^{-\delta}\left(\partial_x\tilde{u}_n\right)^2, \gamma\psi\right)_{Q_T} - \left(\tilde{u}_n^{1-\delta}\partial_x\tilde{u}_n, \psi\partial_x\gamma\right)_{Q_T} + \left(\tilde{u}_n^{-\delta}\left(\tilde{u}_n - \frac{1}{n}\right), \psi g_0\right)_{Q_T}.
\end{aligned}$$

Next, collecting the third term on the left hand side and the first term on the right hand side, we have

$$
\begin{aligned}
&(1-\delta)\left(\gamma\psi, \tilde{u}_n^{-\delta}\left(\partial_x \tilde{u}_n\right)^2\right)_{Q_T} \\
=&-\frac{1}{1-\delta}\left(\partial_t \tilde{u}_n^{1-\delta}, \psi\right)_{Q_T}-\frac{1}{2-\delta}\left(\partial_x(\gamma\psi), \partial_x \tilde{u}_n^{2-\delta}\right)_{Q_T}+\left(\psi g_0, \tilde{u}_n^{1-\delta}\right)_{Q_T}-\left(\psi g_0, \frac{1}{n}\tilde{u}_n^{-\delta}\right)_{Q_T}.
\end{aligned} \tag{6.22}
$$

The first three terms on the right hand side are apparently bounded. Indeed $\tilde{u}_n$ satisfies maximum principle and,

$$
\begin{aligned}
\left(\partial_t \tilde{u}_n^{1-\delta}, \psi\right)_{Q_T} &= -\left(\tilde{u}_n^{1-\delta}, \partial_t \psi\right)_{Q_T}, \\
\left(\partial_x(\gamma\psi), \partial_x \tilde{u}_n^{2-\delta}\right)_{Q_T} &= -\left(\partial_x^2(\gamma\psi), \tilde{u}_n^{2-\delta}\right)_{Q_T}.
\end{aligned}
$$

In addition, since $\tilde{u}_n = u_n + \frac{1}{n} \geq \frac{1}{n}$, in the last term

$$
\frac{1}{n}\tilde{u}_n^{-\delta} \leq n^{\delta-1} = \frac{1}{n^{1-\delta}} \leq 1.
$$

Therefore, the left hand side of (6.22) is uniformly bounded,

$$
\left(\gamma\psi, \tilde{u}_n^{-\delta}\left(\partial_x \tilde{u}_n\right)^2\right)_{Q_T} \leq C(\varphi_0, g_0, T, \psi, \delta). \tag{6.23}
$$

Finally, by Hölder's inequality, the L^1 norm of the sequence $\gamma\psi\tilde{u}_n^{-\delta/2}|\partial_x \tilde{u}_n|$ can be bounded as follows,

$$
\begin{aligned}
\|\gamma\psi\tilde{u}_n^{-\delta/2}|\partial_x \tilde{u}_n|\|_{L^1(Q_T)} &= \left((\gamma\psi)^{1/2}, \left(\gamma\psi\tilde{u}_n^{-\delta}\left(\partial_x \tilde{u}_n\right)^2\right)^{1/2}\right)_{Q_T} \\
&\leq \|(\gamma\psi)^{1/2}\|_{L^2(Q_T)} \cdot \|\left(\gamma\psi\tilde{u}_n^{-\delta}\left(\partial_x \tilde{u}_n\right)^2\right)^{1/2}\|_{L^2(Q_T)} \\
&= \|(\gamma\psi)^{1/2}\|_{L^2(Q_T)} \cdot \sqrt{\left(\gamma\psi, \tilde{u}_n^{-\delta}\left(\partial_x \tilde{u}_n\right)^2\right)_{Q_T}}.
\end{aligned}
$$

And so, by inequality (6.23), the sequence $\gamma\psi\tilde{u}_n^{-\theta}|\partial_x \tilde{u}_n|$ is uniformly bounded in $L^1(Q_T)$ for any $\theta = \frac{\delta}{2} \in (0, \frac{1}{2})$.

□

6.5 Convergence results

The aim of this section is to prove that the sequence $\partial_x u_n$ converge to $\partial_x u$ a.e. in Q_T, where we have adopted the techniques in the work of Abdellaoui et al. (2015). The roadmap is as follows:

1. Using Proposition 6.5 and Proposition 6.7 to prove Lemma 6.11.

2. Theorem 6.12 is a simple corollary of Lemma 6.11.

3. Combining Theorem 6.12 and Proposition 6.10 to prove Theorem 6.13.

Lemma 6.11. *If $\tilde{u}_n$ is the solution to Equation(6.18), then for any $s \in (0,1)$*

$$\lim_{n\to\infty} \int_{Q_T} \left[\gamma \tilde{u}_n \left(\partial_x \left(\tilde{u}_n - u\right)\right)^2\right]^s = 0.$$

Proof. Recall that $u \in L^2\left(0,T;W_0^{1,2}(\Omega)\right)$, introduce the time–regularization of $u(x,t)$ by Landes and Mustonen (1994),

$$u_\nu(x,t) = \exp(-\nu t)\varphi_0(x) + \nu \int_0^t \exp(-\nu(t-s))u(x,s)ds.$$

It is known that

1. $u_\nu(x,t)$ converge to $u(x,t)$ strongly in $L^2\left(0,T;W_0^{1,2}(\Omega)\right)$.

2. u_ν is the solution of the following problem,

$$\begin{cases} \dfrac{1}{\nu}\partial_t u_\nu + u_\nu = u \\ u_\nu(x,0) = \varphi_0(x) \end{cases} \tag{6.24}$$

Define a cut-off function T_ε as

$$T_\varepsilon(y) = \begin{cases} y, & y \in (-\varepsilon,\varepsilon) \\ \operatorname{sign}(y)\varepsilon, & \text{otherwise} \end{cases} \tag{6.25}$$

And define a non-negative function $J_\varepsilon(y)$, such that $J'_\varepsilon(y) = T_\varepsilon(y)$,

$$J_\varepsilon(y) = \begin{cases} -\varepsilon y - \dfrac{1}{2}\varepsilon^2, & y \in (-\infty,-\varepsilon) \\ \dfrac{1}{2}y^2, & y \in (-\varepsilon,\varepsilon) \\ \varepsilon y - \dfrac{1}{2}\varepsilon^2, & y \in (\varepsilon,\infty) \end{cases} \tag{6.26}$$

It takes two steps to prove that $\int_{Q_T}\left[\gamma\tilde{u}_n\left(\partial_x\left(\tilde{u}_n-u\right)\right)^2\right]^s$ converge to zero:

1. prove that $\int_{Q_T}\left[\gamma\tilde{u}_n\left(\partial_x\left(\tilde{u}_n-u\right)\right)^2\right]^s\chi\{|u_n-u_\nu|\le\varepsilon\}$ converge to zero
2. prove that $\int_{Q_T}\left[\gamma\tilde{u}_n\left(\partial_x\left(\tilde{u}_n-u\right)\right)^2\right]^s\chi\{|u_n-u_\nu|>\varepsilon\}$ converge to zero

For the first step, do the following decomposition

$$\begin{aligned}
&\int_{Q_T}\left[\gamma\tilde{u}_n\left(\partial_x\left(\tilde{u}_n-u\right)\right)^2\right]\chi\{|u_n-u_\nu|\le\varepsilon\}\\
=&\int_{\{|u_n-u_\nu|\le\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\left(\tilde{u}_n-u\right)\right)^2\\
=&\int_{\{|u_n-u_\nu|\le\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(\tilde{u}_n-u\right)-\int_{\{|u_n-u_\nu|\le\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x u\right)\partial_x\left(\tilde{u}_n-u\right)\\
=&\int_{\{|u_n-u_\nu|\le\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(\tilde{u}_n-u\right)\\
&-\int_{Q_T}\left[\gamma\left(\tilde{u}_n\chi\{|u_n-u_\nu|\le\varepsilon\}-u\chi\{|u-u_\nu|\le\varepsilon\}\right)\left(\partial_x u\right)\partial_x\left(\tilde{u}_n-u\right)\right]\\
&-\int_{Q_T}\left[\gamma u\chi\{|u-u_\nu|\le\varepsilon\}\left(\partial_x u\right)\partial_x\left(\tilde{u}_n-u\right)\right]\\
=&A_1+A_2+A_3.
\end{aligned}$$

Start first from A_2 and A_3, as their estimates are relatively simple and straightforward.

Indeed, by Hölder's inequality and Corollary 6.9, it follows that,

$$\begin{aligned}
A_2=&-\int_{Q_T}\left[\gamma\left(\tilde{u}_n\chi\{|u_n-u_\nu|\le\varepsilon\}-u\chi\{|u-u_\nu|\le\varepsilon\}\right)\left(\partial_x u\right)\partial_x\left(\tilde{u}_n-u\right)\right]\\
\le&C_1(\Omega,T)\|\tilde{u}_n\chi\{|u_n-u_\nu|\le\varepsilon\}-u\chi\{|u-u_\nu|\le\varepsilon\}\|_{L^2(Q_T)}\|\left(\partial_x u\right)\partial_x\left(\tilde{u}_n-u\right)\|_{L^2(Q_T)}\\
\le&C_2(\varphi_0,g_0,\Omega,T)\|\tilde{u}_n\chi\{|u_n-u_\nu|\le\varepsilon\}-u\chi\{|u-u_\nu|\le\varepsilon\}\|_{L^2(Q_T)},
\end{aligned}\tag{6.27}$$

also, the following term will converge to zero,

$$A_3=-\int_{Q_T}\gamma u\left(\partial_x u\right)\left(\partial_x\tilde{u}_n-\partial_x u\right)\chi\{|u-u_\nu|\le\varepsilon\}.\tag{6.28}$$

It remains to bound $A_1=\int_{\{|u_n-u_\nu|\le\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(\tilde{u}_n-u\right)$.

This estimate is performed by first testing Equation(6.18) with $T_\varepsilon(u_n - u_\nu)$, where T_ε is defined in Equation(6.25), to obtain,

$$\begin{aligned}
&\left(\partial_t \tilde{u}_n, T_\varepsilon(u_n - u_\nu)\right)_{Q_T} + \left(\gamma \tilde{u}_n \partial_x \tilde{u}_n, \partial_x \left(T_\varepsilon(u_n - u_\nu)\right)\right)_{Q_T} \\
&= \left(-\gamma \left(\partial_x \tilde{u}_n\right)^2, T_\varepsilon(u_n - u_\nu)\right)_{Q_T} + \left(-\frac{\partial_x \gamma}{2} \partial_x \tilde{u}_n^2, T_\varepsilon(u_n - u_\nu)\right)_{Q_T} + \left(g_0\left(\tilde{u}_n - \frac{1}{n}\right), T_\varepsilon(u_n - u_\nu)\right)_{Q_T}.
\end{aligned} \tag{6.29}$$

Since $|T_\varepsilon(u_n - u_\nu)| \leq \epsilon$, the right hand side of the above equation can be bounded as follows,

$$\text{RHS} \leq \varepsilon \left(\|\gamma \left(\partial_x \tilde{u}_n\right)^2\|_{L^1(Q_T)} + \|\frac{\partial_x \gamma}{2} \partial_x \tilde{u}_n^2\|_{L^1(Q_T)} + \|g_0\left(\tilde{u}_n - \frac{1}{n}\right)\|_{L^1(Q_T)} \right),$$

with the first term uniformly bounded by Proposition 6.5, the second one uniformly bounded by Corollary 6.6, and the last term by maximum principle. Consequently,

$$\left(\gamma \tilde{u}_n \partial_x \tilde{u}_n, \partial_x \left(T_\varepsilon(u_n - u_\nu)\right)\right)_{Q_T} \leq C_1(\varphi_0, g_0, \Omega, T)\varepsilon - \left(\partial_t \tilde{u}_n, T_\varepsilon(u_n - u_\nu)\right)_{Q_T}.$$

Since u_ν is a solution of Equation(6.24), $\partial_t u_\nu$ can be replaced with $\nu(u - u_\nu)$,

$$\begin{aligned}
\left(\partial_t \tilde{u}_n, T_\varepsilon(u_n - u_\nu)\right)_{Q_T} &= \left(\partial_t \left(u_n - u_\nu\right), T_\varepsilon(u_n - u_\nu)\right)_{Q_T} + \left(\partial_t u_\nu, T_\varepsilon(u_n - u_\nu)\right)_{Q_T} \\
&= \left(\partial_t \left(u_n - u_\nu\right), T_\varepsilon(u_n - u_\nu)\right)_{Q_T} + \nu \left(\left(u - u_\nu\right), T_\varepsilon(u_n - u_\nu)\right)_{Q_T} \\
&= \left(1, \partial_t J_\varepsilon(u_n - u_\nu)\right)_{Q_T} + \nu \left(\left(u - u_\nu\right), T_\varepsilon(u_n - u_\nu)\right)_{Q_T} \\
&= \left(1, J_\varepsilon(u_n(T) - u_\nu(T))\right)_\Omega - \left(1, J_\varepsilon(u_n(0) - u_\nu(0))\right)_\Omega \\
&\quad + \nu \left(\left(u - u_\nu\right), T_\varepsilon(u_n - u_\nu)\right)_{Q_T},
\end{aligned}$$

in which J_ε is defined in Equation(6.26) as the anti-derivative of T_ε.

Each term on the right hand side is bounded from below. Indeed, by definition of J_ε,

$$\left(1, J_\varepsilon(u_n(T) - u_\nu(T))\right)_\Omega \geq 0. \tag{6.30}$$

Since u_n and u_ν share the same initial condition, the second term is actually zero.

$$\left(1, J_\varepsilon(u_n(0) - u_\nu(0))\right)_\Omega = \left(1, J_\varepsilon(\varphi_0 - \varphi_0)\right)_\Omega = 0. \tag{6.31}$$

By the sign-keeping property of T_ε,

$$\begin{aligned}\nu\left(\left(u-u_\nu\right), T_\varepsilon(u_n-u_\nu)\right)_{Q_T} =&\nu\left(\left(u-u_\nu\right), T_\varepsilon(u-u_\nu-u+u_n)\right)_{Q_T}\\ =&\nu\left(\left(u-u_\nu\right), T_\varepsilon(u-u_\nu)\right)_{Q_T}+\nu\left(\left(u-u_\nu\right), T_\varepsilon(u_n-u)\right)_{Q_T}\\ \geq&\nu\left(\left(u-u_\nu\right), T_\varepsilon(u_n-u)\right)_{Q_T}.\end{aligned} \tag{6.32}$$

Therefore, combining inequalities (6.30), (6.31) and (6.32),

$$\begin{aligned}\left(\gamma\tilde{u}_n\partial_x\tilde{u}_n, \partial_x\left(T_\varepsilon(u_n-u_\nu)\right)\right)_{Q_T} \leq& C_1(\varphi_0, g_0, \Omega, T)\varepsilon-\left(\partial_t\tilde{u}_n, T_\varepsilon(u_n-u_\nu)\right)_{Q_T}\\ \leq& C_1(\varphi_0, g_0, \Omega, T)\varepsilon-\nu\left(\left(u-u_\nu\right), T_\varepsilon(u_n-u)\right)_{Q_T}.\end{aligned}$$

Consequently,

$$\begin{aligned}A_1=&\int_{\{|u_n-u_\nu|\leq\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(u_n-u\right)\\ =&\int_{\{|u_n-u_\nu|\leq\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(u_n-u_\nu\right)+\int_{\{|u_n-u_\nu|\leq\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(u_\nu-u\right)\\ =&\int_{Q_T}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(T_\varepsilon\left(u_n-u_\nu\right)\right)+\int_{\{|u_n-u_\nu|\leq\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(u_\nu-u\right)\\ \leq& C_1(\varphi_0, g_0, \Omega, T)\varepsilon-\nu\left(\left(u-u_\nu\right), T_\varepsilon(u_n-u)\right)_{Q_T}+\int_{\{|u_n-u_\nu|\leq\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(u_\nu-u\right).\end{aligned} \tag{6.33}$$

Putting together the inequalities (6.33), (6.27) and (6.28),

$$\begin{aligned}&\int_{Q_T}\left[\gamma\tilde{u}_n\left(\partial_x\left(\tilde{u}_n-u\right)\right)^2\right]\chi\{|u_n-u_\nu|\leq\varepsilon\}\\ =&A_1+A_2+A_3\\ \leq& C_1(\varphi_0, g_0, \Omega, T)\varepsilon-\nu\left(\left(u-u_\nu\right), T_\varepsilon(u_n-u)\right)_{Q_T}+\int_{\{|u_n-u_\nu|\leq\varepsilon\}}\gamma\tilde{u}_n\left(\partial_x\tilde{u}_n\right)\partial_x\left(u_\nu-u\right)\\ &+C_2(\varphi_0, g_0, \Omega, T)\|\tilde{u}_n\chi\{|u_n-u_\nu|\leq\varepsilon\}-u\chi\{|u-u_\nu|\leq\varepsilon\}\|_{L^2(Q_T)}\\ &-\int_{Q_T}\gamma u\left(\partial_x u\right)\left(\partial_x\tilde{u}_n-\partial_x u\right)\chi\{|u-u_\nu|\leq\varepsilon\}\\ =&B_1(n,\nu,\varepsilon).\end{aligned} \tag{6.34}$$

Since $\partial_x \tilde{u}_n$ converge to $\partial_x u$ weakly in $L^2(Q_T)$, the last term of B_1 converges to zero as n goes to infinity, therefore,

$$\lim_{\varepsilon \to 0^+} \limsup_{\nu \to \infty} \limsup_{n \to \infty} B_1(n, \nu, \varepsilon) = 0.$$

For the second step, consider $\int_{Q_T} \left[\gamma \tilde{u}_n \left(\partial_x \left(\tilde{u}_n - u\right)\right)^2\right]^s \chi\{|u_n - u_\nu| > \varepsilon\}$, using Hölder's inequality,

$$\begin{aligned}
&\int_{Q_T} \left[\gamma \tilde{u}_n \left(\partial_x \left(\tilde{u}_n - u\right)\right)^2\right]^s \chi\{|u_n - u_\nu| > \varepsilon\} \\
\leq & C_1(\Omega, T) \|(\partial_x u_n - \partial_x u)^{2s}\|_{L^{\rho'}(Q_T)} \\
& \cdot \left(\|\chi\{|u_n - u_\nu| > \varepsilon\} - \chi\{|u - u_\nu| > \varepsilon\}\|_{L^\rho(Q_T)} + \|\chi\{|u - u_\nu| > \varepsilon\}\|_{L^\rho(Q_T)}\right) \\
= & B_2(n, \nu, \varepsilon).
\end{aligned} \tag{6.35}$$

Taking the limit,

$$\lim_{\varepsilon \to 0^+} \limsup_{\nu \to \infty} \limsup_{n \to \infty} B_2(n, \nu, \varepsilon) = 0.$$

To summarize,

$$0 \leq \int_{Q_T} \left[\gamma \tilde{u}_n \left(\partial_x \left(\tilde{u}_n - u\right)\right)^2\right]^s \leq B_1(n, \nu, \varepsilon) + B_2(n, \nu, \varepsilon),$$

where B_1 and B_2 are on the right hand side of Equation (6.34) and (6.35). Consequently,

$$\lim_{n \to \infty} \int_{Q_T} \left[\gamma \tilde{u}_n \left(\partial_x \left(\tilde{u}_n - u\right)\right)^2\right]^s = 0.$$

□

Theorem 6.12. *The sequence $\partial_x \tilde{u}_n^2 = \partial_x (u_n + \frac{1}{n})^2$ converge to $\partial_x u^2$ strongly in $L^\sigma(Q_T)$ for all $\sigma \in (0, 2)$.*

Proof. Note that

$$\begin{aligned}
&\int_{Q_T} |\partial_x \tilde{u}_n^2 - \partial_x u^2|^{2s} \\
=&2^{2s} \int_{Q_T} |\tilde{u}_n \partial_x \tilde{u}_n - u \partial_x u|^{2s} \\
=&2^{2s} \int_{Q_T} |\tilde{u}_n \partial_x \tilde{u}_n - \tilde{u}_n \partial_x u + \tilde{u}_n \partial_x u - u \partial_x u|^{2s} \\
=&2^{2s} \int_{Q_T} |(\tilde{u}_n \partial_x \tilde{u}_n - \tilde{u}_n \partial_x u) + \partial_x u \, (\tilde{u}_n - u)|^{2s} \\
\leq& C \int_{Q_T} \left(|\tilde{u}_n \partial_x \tilde{u}_n - \tilde{u}_n \partial_x u|^{2s} + |\partial_x u \, (\tilde{u}_n - u)|^{2s} \right).
\end{aligned}$$

By Lemma 6.11 and Corollary 6.9, both terms converge to zero if $s \in (0,1)$. Let $\sigma = 2s$, then $\sigma \in (0,2)$. □

Theorem 6.13. *The sequence $\partial_x u_n$ converge to $\partial_x u$ a.e. in Q_T.*

Proof. Let $\psi \in C_0^\infty(Q_T)$ be s.t. $\psi \geq 0$ in Q_T. To prove convergence a.e., it is sufficient to show that for some $\alpha \in (0,1)$,

$$\lim_{n \to \infty} \int_{Q_T} |\partial_x u_n - \partial_x u|^\alpha \psi = 0.$$

Decompose the domain Q_T,

$$\begin{aligned}
\int_{Q_T} |\partial_x u_n - \partial_x u|^\alpha \psi &= \int_{\{u=0\}} |\partial_x u_n - \partial_x u|^\alpha \psi + \int_{\{u>0\}} |\partial_x u_n - \partial_x u|^\alpha \psi \\
&= \int_{\{u=0\}} |\partial_x u_n|^\alpha \psi + \int_{\{0<u\leq \frac{1}{m}\}} |\partial_x u_n - \partial_x u|^s \psi + \int_{\{u>\frac{1}{m}\}} |\partial_x u_n - \partial_x u|^s \psi \\
&= A_1 + A_2 + A_3.
\end{aligned} \tag{6.36}$$

Using Hölder's inequality to get the bound of A_2,

$$\begin{aligned}
A_2 &= \int_{\{0<u\leq \frac{1}{m}\}} |\partial_x u_n - \partial_x u|^s \psi \\
&\leq \| |\partial_x u_n - \partial_x u|^s \psi \|_{L^{2/s}(Q_T)} \| \chi_{\{0<u\leq \frac{1}{m}\}} \|_{L^{\frac{2}{2-s}}(Q_T)} \\
&\leq C \| \chi_{\{0<u\leq \frac{1}{m}\}} \|_{L^{\frac{2}{2-s}}(Q_T)}.
\end{aligned}$$

Note that $\|\chi_{\{0<u\leq\frac{1}{m}\}}\|_{L^{\frac{2}{2-s}}(Q_T)}$ can be arbitrarily small.

Next, by Theorem 6.12, it is known that $\partial_x\tilde{u}_n^2 \to \partial_x u^2$ strongly in $L^\sigma(Q_T)$ for all $\sigma < 2$, therefore A_3 converges to zero, in fact,

$$\begin{aligned} A_3 &= \int_{\{u>\frac{1}{m}\}} \frac{1}{|u|^s}|u\partial_x u_n - u\partial_x u|^s\psi \\ &= \int_{\{u>\frac{1}{m}\}} \frac{1}{|u|^s}|(u-\tilde{u}_n)\,\partial_x\tilde{u}_n + \frac{1}{2}\left(\partial_x\tilde{u}_n^2 - \partial_x u^2\right)|^s\psi \\ &\leq m^s \int_{Q_T} |(u-\tilde{u}_n)\,\partial_x\tilde{u}_n + \frac{1}{2}\left(\partial_x\tilde{u}_n^2 - \partial_x u^2\right)|^s\psi, \end{aligned}$$

and the limit follows from

$$\limsup_{n\to\infty} A_3(n) \leq m^s \limsup_{n\to\infty} \int_{Q_T} |(u-\tilde{u}_n)\,\partial_x\tilde{u}_n + \frac{1}{2}\left(\partial_x\tilde{u}_n^2 - \partial_x u^2\right)|^s\psi = 0.$$

Considering A_1 of Equation 6.36, since $u_n \to u$ strongly in $L^q(Q_T)$, by Egorov's Lemma, for every $\epsilon > 0$, there exists a measurable set E_ϵ such that $|E_\epsilon| \leq \epsilon$ and $u_n \to u$ uniformly in $Q_T\backslash E_\epsilon$.

$$\int_{\{u=0\}} |\partial_x u_n|^\alpha\psi = \int_{\{u=0\}\cap E_\epsilon} |\partial_x u_n|^\alpha\psi + \int_{\{u=0\}\cap Q_T\backslash E_\epsilon} |\partial_x u_n|^\alpha\psi.$$

The first term is bounded through Hölder's inequality,

$$\begin{aligned} \int_{\{u=0\}\cap E_\epsilon} |\partial_x u_n|^\alpha\psi &= \int_{Q_T} |\partial_x u_n|^\alpha\psi\chi_{\{u=0\}\cap E_\epsilon} \\ &\leq \int_{Q_T} |\partial_x u_n|^\alpha\psi\chi_{E_\epsilon} \leq C\||\partial_x u_n|^\alpha\|_{L^{1/\alpha}(Q_T)}\|\chi_{E_\epsilon}\|_{L^{1/(1-\alpha)}(Q_T)} \\ &\leq C\epsilon^{1-\alpha}. \end{aligned}$$

The second one uses the fact that for any $\mu > 0$, there exists N such that $|u_n - u| = |u_n| < \mu$ for all $n > N$ and for all $x \in \{u = 0\} \cap Q_T\backslash E_\epsilon$. In other words, for $n > N$,

$\{u=0\}\cap Q_T\backslash E_\epsilon$ is a subset of $\{u_n\le\mu\}\cap Q_T\backslash E_\epsilon$, hence the integral

$$\begin{aligned}
&\int_{\{u=0\}\cap Q_T\backslash E_\epsilon}|\partial_x u_n|^\alpha\psi\\
\le&\int_{\{u_n\le\mu\}\cap Q_T\backslash E_\epsilon}|\partial_x u_n|^\alpha\psi\\
\le&\left(\mu+\frac{1}{n}\right)^{\theta\alpha}\int_{\{u_n\le\mu\}\cap Q_T\backslash E_\epsilon}\left(\frac{|\partial_x u_n|}{(u_n+\frac{1}{n})^\theta}\right)^\alpha\psi\\
\le&\left(\mu+\frac{1}{n}\right)^{\theta\alpha}\int_{Q_T}\left(\frac{|\partial_x u_n|}{(u_n+\frac{1}{n})^\theta}\right)^\alpha\psi.
\end{aligned}$$

The boundedness of $\int_{Q_T}\left(\frac{|\partial_x u_n|}{(u_n+\frac{1}{n})^\theta}\right)^\alpha\psi$ is secured by Proposition 6.10. The result follows from taking $\mu\to 0$.

□

6.6 Existence of global weak solution

Using Equation(6.21) in another direction, we have the following lemma which is essential for the proof of Theorem 6.2.

Lemma 6.14. *If u_n are classical solutions to the problem(6.14) for $t\in[0,T]$, then the integral of $\partial_x u_n$ on $(0,T)$ is bounded as follows,*

$$\left\|\int_0^T\partial_x u_n(\cdot,t)dt\right\|_{L^\infty(\Omega)}\le C(\varphi_0,g_0,T)\sqrt{n}.$$

Proof. Let

$$y_n(x)=\int_0^T z_n(x,t)dt=\int_0^T\partial_x u_n(x,t)dt.$$

The goal is to bound $\|y_n(x)\|_{L^\infty(\Omega)}$.

Note that by definition

$$\int_a^b y_n(x)dx=\int_0^T\int_a^b\partial_x u_n(x,t)dxdt=\int_0^T\left(u_n(b,t)-u_n(a,t)\right)dt=0,$$

hence it is possible to use the Poincare's inequality if we can derive a bound for $\|\partial_x y_n\|_{L^2(\Omega)} = \left[\int_a^b (\partial_x y_n)^2 \, dx\right]^{\frac{1}{2}}$.

By Hölder's inequality,

$$|\partial_x y_n(x)|^2 = \left|\int_0^T \partial_x z_n(x,t)dt\right|^2 \leq \left(\int_0^T |\partial_x z_n(x,t)|dt\right)^2 \leq T \cdot \left(\int_0^T |\partial_x z_n(x,t)|^2 dt\right).$$

Integrate on Ω, the above inequality yields,

$$\|\partial_x y_n\|^2_{L^2(\Omega)} = \int_a^b |\partial_x y_n(x)|^2 dx \leq T \cdot \left(\int_a^b \int_0^T |\partial_x z_n(x,t)|^2 dt dx\right) = T \cdot \|\partial_x z_n\|^2_{L^2(Q_T)}. \tag{6.37}$$

To bound the right hand side, recall Equation(6.21) with $l = 0$,

$$\left(\gamma(u_n + \frac{1}{n})\partial_x z_n, \partial_x z_n\right)_\Omega = -\left(\partial_t z_n, z_n\right)_\Omega + \left(g_0 z_n, z_n\right)_\Omega + \left(u_n \partial_x g_0, z_n\right)_\Omega. \tag{6.38}$$

Since $u_n \geq 0$ and $\gamma(x) \geq \gamma_m$ on Ω, it follows that

$$\gamma_m \frac{1}{n}\left(\partial_x z_n, \partial_x z_n\right)_\Omega \leq \left(\gamma(u_n + \frac{1}{n})\partial_x z_n, \partial_x z_n\right)_\Omega.$$

Integrate on $(0,T)$ to obtain,

$$\left(\partial_x z_n, \partial_x z_n\right)_{Q_T} \leq \frac{n}{\gamma_m}\left[-\frac{1}{2}\left(z_n(T), z_n(T)\right)_\Omega + \frac{1}{2}\left(z_n(0), z_n(0)\right)_\Omega + \left(g_0 z_n, z_n\right)_{Q_T} + \left(u_n \partial_x g_0, z_n\right)_{Q_T}\right]$$

By Proposition 6.5 and the maximum principle, the above inequality yields,

$$\left(\partial_x z_n, \partial_x z_n\right)_{Q_T} \leq C_1(\varphi_0, g_0, T)n.$$

Combining the inequality above with inequality (6.37), we have

$$\|\partial_x y_n\|^2_{L^2(\Omega)} \leq C_1(\varphi_0, g_0, T)Tn.$$

Since y_n has zero mean, the Poincare's inequality renders

$$\|y_n\|^2_{H^1(\Omega)} \leq C_2(\varphi_0, g_0, T)n.$$

Thus by Sobolev's inequality we have

$$\|y_n\|_{L^\infty(\Omega)} \leq C(\varphi_0, g_0, T)\sqrt{n}.$$

□

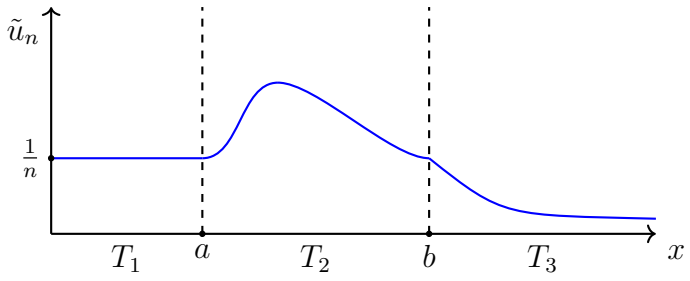


Figure 6.1: The lifted extension

We are now ready to prove the main result of the chapter, i.e. Theorem 6.2.

Proof. Note that any integral on Ω^* is a summation of three pieces,

$$
\begin{aligned}
(\tilde{u}_n, \partial_t \eta)_{Q_T^*} &= \int_0^T \int_0^a \frac{1}{n} (\partial_t \eta)\, dxdt + \left(u_n + \frac{1}{n}, \partial_t \eta\right)_{Q_T} + \int_0^T \int_b^{+\infty} \frac{1}{n} h\, (\partial_t \eta)\, dxdt, \\
(\gamma \tilde{u}_n \partial_x \tilde{u}_n, \partial_x \eta)_{Q_T^*} &= 0 + \left(\gamma(u_n + \frac{1}{n})\partial_x u_n, \partial_x \eta\right)_{Q_T} + \int_0^T \int_b^{+\infty} \gamma \left(\frac{1}{n} h\right) \left(\frac{1}{n} \partial_x h\right) (\partial_x \eta)\, dxdt, \\
\left(\gamma\, (\partial_x \tilde{u}_n)^2, \eta\right)_{Q_T^*} &= 0 + \left(\gamma\, (\partial_x u_n)^2, \eta\right)_{Q_T} + \int_0^T \int_b^{+\infty} \left[\gamma \left(\frac{1}{n} \partial_x h\right)^2 \eta\right] dxdt, \\
\left(\frac{\partial_x \gamma}{2} \partial_x \tilde{u}_n^2, \eta\right)_{Q_T^*} &= 0 + \left((\partial_x \gamma)\, (u_n + \frac{1}{n}) \partial_x u_n, \eta\right)_{Q_T} + \int_0^T \int_b^{+\infty} \left[\frac{\partial_x \gamma}{2} \left(\frac{1}{n^2} \partial_x h^2\right) \eta\right] dxdt, \\
(g_0 \tilde{u}_n, \eta)_{Q_T^*} &= \int_0^T \int_0^a \left(\frac{1}{n} g_0 \eta\right) dxdt + \left(g_0 \left(u_n + \frac{1}{n}\right), \eta\right)_{Q_T} + \int_0^T \int_b^{+\infty} \left(g_0 \frac{1}{n} h \eta\right) dxdt, \\
(\varphi_0, \eta(x,0))_{\Omega^*} &= 0 + (\varphi_0, \eta(x,0))_{\Omega} + 0.
\end{aligned}
$$

Consequently,

$$
\begin{aligned}
&- (\tilde{u}_n, \partial_t \eta)_{Q_T^*} + (\gamma \tilde{u}_n \partial_x \tilde{u}_n, \partial_x \eta)_{Q_T^*} + \left(\gamma\, (\partial_x \tilde{u}_n)^2, \eta\right)_{Q_T^*} + \left(\frac{\partial_x \gamma}{2} \partial_x \tilde{u}_n^2, \eta\right)_{Q_T^*} \\
&- (g_0 \tilde{u}_n, \eta)_{Q_T^*} - (\varphi_0, \eta(x,0))_{\Omega^*} \\
=& T_1 + T_2 + T_3,
\end{aligned}
\tag{6.39}
$$

where we collect the terms according to our partition of the domain, as illustrated in

Figure 6.1,

$$
\begin{aligned}
T_1 =& -\int_0^T \int_0^a \frac{1}{n}\left(\partial_t \eta\right) dxdt + 0 + 0 + 0 - \int_0^T \int_0^a \left(\frac{1}{n} g_0 \eta\right) dxdt - 0, \\
T_2 =& -\left(u_n + \frac{1}{n}, \partial_t \eta\right)_{Q_T} + \left(\gamma(u_n + \frac{1}{n})\partial_x u_n, \partial_x \eta\right)_{Q_T} + \left(\gamma\left(\partial_x u_n\right)^2, \eta\right)_{Q_T} \\
&+ \left(\left(\partial_x \gamma\right)(u_n + \frac{1}{n})\partial_x u_n, \eta\right)_{Q_T} - \left(g_0\left(u_n + \frac{1}{n}\right), \eta\right)_{Q_T} - \left(\varphi_0, \eta(x,0)\right)_{\Omega}, \\
T_3 =& -\int_0^T \int_b^{+\infty} \frac{1}{n} h\left(\partial_t \eta\right) dxdt + \int_0^T \int_b^{+\infty} \gamma\left(\frac{1}{n}h\right)\left(\frac{1}{n}\partial_x h\right)\left(\partial_x \eta\right) dxdt \\
&+ \int_0^T \int_b^{+\infty} \left[\gamma\left(\frac{1}{n}\partial_x h\right)^2 \eta\right] dxdt \\
&+ \int_0^T \int_b^{+\infty} \left[\frac{\partial_x \gamma}{2}\left(\frac{1}{n^2}\partial_x h^2\right)\eta\right] dxdt - \int_0^T \int_b^{+\infty} \left(g_0 \frac{1}{n} h\eta\right) dxdt - 0.
\end{aligned}
$$

We claim that up to a subsequence,

$$
\begin{aligned}
&\lim_{n\to\infty} \left\{ \begin{array}{l} -\left(\tilde{u}_n, \partial_t \eta\right)_{Q_T^*} + \left(\gamma \tilde{u}_n \partial_x \tilde{u}_n, \partial_x \eta\right)_{Q_T^*} + \left(\gamma\left(\partial_x \tilde{u}_n\right)^2, \eta\right)_{Q_T^*} \\ \quad + \left(\frac{\partial_x \gamma}{2}\partial_x \tilde{u}_n^2, \eta\right)_{Q_T^*} - \left(g_0 \tilde{u}_n, \eta\right)_{Q_T^*} - \left(\varphi_0, \eta(x,0)\right)_{\Omega^*} \end{array} \right\} \\
&= \begin{array}{l} -\left(u, \partial_t \eta\right)_{Q_T^*} + \left(\gamma u \partial_x u, \partial_x \eta\right)_{Q_T^*} + \left(\gamma\left(\partial_x u\right)^2, \eta\right)_{Q_T^*} \\ + \left(\frac{\partial_x \gamma}{2}\partial_x u^2, \eta\right)_{Q_T^*} - \left(g_0 u, \eta\right)_{Q_T^*} - \left(\varphi_0, \eta(x,0)\right)_{\Omega^*}. \end{array}
\end{aligned} \tag{6.40}
$$

The only non-trivial part is to prove that the third term on the left hand side converges to $(-\gamma(\partial_x u)^2, \eta)_{Q_T}$, for which we take the difference and use Hölder's inequality,

$$
\left(\gamma\left[\left(\partial_x \tilde{u}_n\right)^2 - \left(\partial_x u\right)^2\right], \eta\right)_{Q_T} \leq C\|\partial_x \tilde{u}_n - \partial_x u\|_{L^2(Q_T)} \cdot \|\partial_x \tilde{u}_n + \partial_x u\|_{L^2(Q_T)}.
$$

Since $\partial_x \tilde{u}_n$ converge to $\partial_x u$ a.e. in Q_T and $\partial_x \tilde{u}_n$ is uniformly bounded in $L^4(Q_T)$, by Lemma 6.8,

$$
\lim_{n\to\infty} \|\partial_x \tilde{u}_n - \partial_x u\|_{L^2(Q_T)} = 0.
$$

Now it remains to prove that

$$
\lim_{n\to+\infty} T_1 + T_2 + T_3 = 0.
$$

It can be easily verified that T_1 and T_3 goes to zero, given the definition of lifted extension $\tilde{u}_n$ in (6.15). For the second row, test Equation(6.14) with η and integrate by parts, it follows that

$$T_2 = -\left(\frac{1}{n}, \partial_t \eta\right)_{Q_T} + \gamma(b) \int_0^T \frac{1}{n} \left(\partial_x u_n(b,t)\right) \eta(b,t) dt - \gamma(a) \int_0^T \frac{1}{n} \left(\partial_x u_n(a,t)\right) \eta(a,t) dt - \frac{1}{n} \left(g_0, \eta\right)_{Q_T}.$$

By Lemma 6.14 the second and third term are in the order of $\mathcal{O}(\frac{1}{\sqrt{n}})$, and the other two terms are in the order of $\mathcal{O}(\frac{1}{n})$, therefore T_2 goes to zero.

To conclude, the limit function u satisfies the following weak form of the equation,

$$\begin{aligned} &- \left(u, \partial_t \eta\right)_{Q_T^*} + \left(\frac{\gamma}{2} \partial_x u^2, \partial_x \eta\right)_{Q_T^*} \\ &= \left(-\gamma (\partial_x u)^2, \eta\right)_{Q_T^*} + \left(-\left(\frac{\partial_x \gamma}{2}\right) \partial_x u^2, \eta\right)_{Q_T^*} + \left(g_0 u, \eta\right)_{Q_T^*} + \left(\varphi_0, \eta(x,0)\right)_{\Omega^*}. \end{aligned}$$

□

6.7 Appendix

According to Theorem 6.1 in Chapter V of Ladyzhenskaya et al. (1968), the following conditions (a) to (f) are sufficient for Theorem 6.4.

Recall that

$$\begin{aligned} a_1(x,t,u,p) &\coloneqq \gamma \mathcal{P}_n(u) p \\ \tilde{a}(x,t,u,p) &\coloneqq \gamma \mathcal{P}_n'(u) p^2 + \left(\partial_x \gamma\right) \mathcal{P}_n(u) p - g_0(x) u \\ A(x,t,u,p) &\coloneqq -g_0(x) u \end{aligned}$$

and

$$\varphi(x,t) \coloneqq \varphi_0(x) + \left[\gamma \mathcal{P}_n'(\varphi_0) \partial_x^2 \varphi_0 + g_0 \varphi_0\right] t \tag{6.41}$$

We will verify the conditions one by one.

1. For $(x,t) \in \overline{Q_T}$ and arbitrary u, the diffusion term is strictly coercive,

$$\frac{\partial a_1}{\partial p}(x,t,u,p) = \gamma \mathcal{P}_n(u) \geq \frac{\gamma_m}{2n} > 0,$$

and the reaction term has the following lower bound,

$$A(x,t,u,0)u = -g_0(x)u^2 \geq -\max(|g_0|)u^2.$$

2. For $(x,t) \in \overline{Q_T}$, when $|u| \leq M$, for arbitrary p, the operators are bounded in the following sense.

$$\frac{\partial a_1}{\partial p}(x,t,u,p) = \gamma \mathcal{P}_n(u) \leq \max(\gamma)(M+1),$$

and

$$\begin{aligned}
&\left(|a_1| + |\frac{\partial a_1}{\partial u}|\right)(1+|p|) + |\frac{\partial a_1}{\partial x}| + |\tilde{a}| \\
=&\left(\gamma \mathcal{P}_n(u)|p| + \gamma \mathcal{P}'_n(u)|p|\right)(1+|p|) + \left(\partial_x \gamma\right)\mathcal{P}_n(u)|p| + |\gamma \mathcal{P}'_n(u)p^2 + \left(\partial_x \gamma\right)\mathcal{P}_n(u)p - g_0(x)u| \\
\leq&\left(\gamma \mathcal{P}_n(u)|p| + \gamma \mathcal{P}'_n(u)|p|\right)(1+|p|) + \left(\partial_x \gamma\right)\mathcal{P}_n(u)|p| + \gamma \mathcal{P}'_n(u)p^2 + \left(\partial_x \gamma\right)\mathcal{P}_n(u)|p| + |g_0(x)u| \\
\leq&\mu(M, b, \max(|g_0|))(1+|p|)^2.
\end{aligned}$$

3. For $(x,t) \in \overline{Q_T}$, $|u| \leq M$ and $|p| \leq M_1$, the functions a_1, $\tilde{a}$, $\frac{\partial a_1}{\partial p}$, $\frac{\partial a_1}{\partial u}$, and $\frac{\partial a_1}{\partial x}$ are arbitrarily smooth in x, t, u and p, therefore they satisfy any Hölder continuity condition.

4. Note that

$$\begin{aligned}
\frac{\partial a_1}{\partial u} &= \gamma \mathcal{P}'_n(u)p, \\
\frac{\partial \tilde{a}}{\partial p} &= 2\gamma \mathcal{P}'_n(u)p + \left(\partial_x \gamma\right)\mathcal{P}_n(u), \\
\frac{\partial \tilde{a}}{\partial u} &= \gamma \mathcal{P}''_n(u)p^2 + \left(\partial_x \gamma\right)\mathcal{P}'_n(u)p - g_0(x).
\end{aligned}$$

For $(x,t) \in \overline{Q_T}$, $|u| \leq M$ and $|p| \leq M_1$, all the above terms are bounded by a constant $C(M, M_1, \mathcal{P}_n, g_0, \Omega)$.

In addition, neither $\tilde{a}$ nor a_1 depend on t, therefore condition (d) is satisfied.

5. By definition of φ in Equation(6.41), φ is arbitrarily smooth in $\overline{Q_T}$. In addition, for $x \in \partial\Omega$ and $t = 0$, the following identity holds,

$$\partial_t \varphi(x,t) = \gamma \mathcal{P}'_n(\varphi_0)\partial_x^2 \varphi_0 + g_0\varphi_0 = \gamma \mathcal{P}'_n(\varphi)\partial_x^2 \varphi + g_0\varphi.$$

6. It is trivial that the boundary $\partial\Omega$ satisfies any Hölder continuity condition.

Chapter 7: Summary

The quasilinear theory for particle-wave interaction in plasmas, as a weak turbulence limit of the Vlasov-Maxwell (Vlasov-Poisson) system, has been extensively studied and used in physics since it was proposed half a century ago. Nevertheless, there is barely any research from the numerical and analytical point of view. The work presented here overviewed the challenging problems in this field and attempted to contribute novel ideas for each of them.

The wellposedness of the quasilinear kinetic system has not gained attention until recent years. We presented the first result on solvability of the model. By invoking an observation proposed in the early days, we transformed the system into a single porous medium equation, and proved the existence of global weak solutions.

Motivated by the electron runaway problem in magnetic confinement fusion, we proposed a numerical solver for particle-wave interaction in non-uniform magnetized plasmas and presented numerical results. In tackling this example problem, we encountered and resolved a lot of difficulties, which are listed below.

The system is well-known for its high dimensionality. Even with axial symmetry, there are seven dimensions to be considered: one dimension for radial position, two dimensions for electron momentum, three dimensions for plasmon momentum, and one more dimension for time. By analyzing the structure of trilinear operators in the weak form, we proposed a novel approach to construct the sparse interaction tensors with double efficiency. Furthermore, our code is parallelized to accelerate.

Another challenge is the nonlinearity, for which we used Euler schemes for time discretization and discussed the conditions for positivity and stability in detail.

Concerning singular transition rates modeled by Dirac delta in the integro-differential formulations, we adopted the marching simplex algorithm and performed numerical integration on the discretized resonance manifold.

The multi-scale problem caused by rapid periodic advection of plasmons is resolved by trajectorial average, for which we proposed the connection-proportion algorithm to discretize the kernel (null space) of Nambu brackets. The trajectorial-averaged equation can thus be solved in a Galerkin approach.

Lastly and most importantly, the delicate structure of the kinetic system will in general be distorted after trajectorial average and discretization. To avoid that, we proposed three novel ideas, and combined them to obtain a structure-preserving solver. We proposed the unconditional conservative weak form, ensuring conservation regardless of the emission/absorption kernel. Then we demonstrate that our Galerkin approach for trajectorial average is structure-preserving. Based on that, we designed a special discrete directional differential operator which guarantees discrete conservation laws.

In the future, we will include the relativistic Landau collision operator and external electric field, to ultimately develop a solver for real-world application: kinetic simulation of electron runaway.

Error estimates for the proposed schemes will also be investigated.

Appendix A: Generalized inverse inequality

Theorem A.1. *If $\widetilde{Z} \in \mathcal{V}_h$ satisfies that*

$$\sum_R \left(\widetilde{Z}, V\right)_R = -\sum_R (f, \nabla \cdot V)_R + \sum_R \langle \widehat{f}, V^- \cdot \mathbf{n}^- \rangle_{\partial R}$$

for any $V \in \mathcal{V}_h$, then

$$\left(\widetilde{Z}, \widetilde{Z}\right) \lesssim \frac{1}{h^2} (f, f).$$

Proof. Note that

$$\begin{aligned}
\left(\widetilde{Z}, \nabla f\right)_R &= (\nabla f, \nabla f)_R + \langle \widehat{f} - f^-, (\nabla f)^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma} \\
\left(\widetilde{Z}, \widetilde{Z}\right)_R &= \left(\nabla f, \widetilde{Z}\right)_R + \langle \widehat{f} - f^-, (\widetilde{Z})^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma}
\end{aligned}$$

Sum up the above two lines and use Young's inequality twice to obtain

$$\begin{aligned}
\left(\widetilde{Z}, \widetilde{Z}\right)_R =& (\nabla f, \nabla f)_R + \langle \widehat{f} - f^-, (\nabla f)^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma} + \langle \widehat{f} - f^-, (\widetilde{Z})^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma} \\
\leq& (\nabla f, \nabla f)_R + \frac{C}{2} \langle (\nabla f)^- \cdot \mathbf{n}^-, (\nabla f)^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma} \\
&+ \frac{1}{C} \langle \widehat{f} - f^-, \widehat{f} - f^- \rangle_{\partial R/\Gamma} + \frac{C}{2} \langle (\widetilde{Z})^- \cdot \mathbf{n}^-, (\widetilde{Z})^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma}
\end{aligned}$$

Note that $\widetilde{Z}$ is a polynomial in R, hence

$$\sum_R \langle (\widetilde{Z})^- \cdot \mathbf{n}^-, (\widetilde{Z})^- \cdot \mathbf{n}^- \rangle_{\partial R/\Gamma} \leq \frac{C_1}{h} \left(\widetilde{Z}, \widetilde{Z}\right)$$

Let $C = \frac{2h}{C_1}$, it follows that

$$\left(\widetilde{Z}, \widetilde{Z}\right) \leq \frac{3}{2} (\nabla f, \nabla f) + \frac{C_1}{2h} \sum_R \langle \widehat{f} - f^-, \widehat{f} - f^- \rangle_{\partial R/\Gamma} + \frac{1}{2} \left(\widetilde{Z}, \widetilde{Z}\right)$$

Finally, inverse inequality and trace inequality gives

$$\left(\widetilde{Z}, \widetilde{Z}\right) \lesssim \frac{1}{h^2} (f, f).$$

□

Appendix B: Quasilinear theory: a weak turbulence model

Quasilinear theory in plasma physics is a model reduction of Vlasov-Maxwell (Poisson) system in weak turbulence regime. Weak turbulence is a concept originating in fluid dynamics, therefore we present a brief introduction of both models.

B.1 Weak turbulence model in fluid dynamics

Consider the Navier-Stokes equation for incompressible flow,

$$\frac{\partial \mathbf{v}}{\partial t} + \nabla \cdot (\mathbf{v} \otimes \mathbf{v}) = -\frac{1}{\rho}\nabla P + \nu \Delta \mathbf{v},$$
$$\nabla \cdot \mathbf{v} = 0.$$

Split the fields into steady parts and fluctuating parts,

$$\mathbf{v} = \overline{\mathbf{v}} + \delta \mathbf{v},$$
$$P = \overline{P} + \delta P,$$

where $\overline{\delta \mathbf{v}} = \mathbf{0}$ and $\overline{\delta P} = 0$.

The equations for the steady parts are the so-called RANS(Reynolds-averaged Navier–Stokes) equations,

$$\nabla \cdot (\overline{\mathbf{v}} \otimes \overline{\mathbf{v}}) + \nabla \cdot T_R = -\frac{1}{\rho}\nabla \overline{P} + \nu \Delta \overline{\mathbf{v}},$$
$$\nabla \cdot \overline{\mathbf{v}} = 0,$$

where $T_R \coloneqq \overline{\delta \mathbf{v} \otimes \delta \mathbf{v}}$ is called the Reynolds stress tensor. Meanwhile, the fluctuating parts satisfy:

$$\frac{\partial \delta \mathbf{v}}{\partial t} + \nabla \cdot (\overline{\mathbf{v}} \otimes \delta \mathbf{v}) + \nabla \cdot (\delta \mathbf{v} \otimes \overline{\mathbf{v}}) + \nabla \cdot (\delta \mathbf{v} \otimes \delta \mathbf{v}) - \nabla \cdot \left(\overline{\delta \mathbf{v} \otimes \delta \mathbf{v}}\right) = -\frac{1}{\rho}\nabla \delta P + \rho \nu \Delta \delta \mathbf{v},$$
$$\nabla \cdot \delta \mathbf{p} = 0.$$

It requires additional modeling of the nonlinear Reynolds stress term T_R to close the RANS equation, and has led to the creation of many different turbulence models.

B.2 Weak turbulence model in plasma physics

The same idea can be applied to the following Vlasov-Poisson system:

$$\frac{\partial f}{\partial t} + \nabla_p \mathcal{E}(\mathbf{p}) \cdot \nabla_x f - \mathbf{E}(\mathbf{x}, t) \cdot \nabla_p f = 0,$$
$$\nabla_x \cdot \mathbf{E} = 1 - \int f d\mathbf{v},$$

Split the particle *pdf* and the electric field,

$$f(\mathbf{x}, \mathbf{p}, t) = \overline{f}(\mathbf{p}, t) + \delta f(\mathbf{x}, \mathbf{p}, t),$$
$$\mathbf{E}(\mathbf{x}, t) = \overline{\mathbf{E}}(t) + \delta \mathbf{E}(\mathbf{x}, t),$$

where $\overline{g} = \frac{1}{|\Omega_x|} \int_{\Omega_x} g(\mathbf{x}) d\mathbf{x}$.

The equation for the slowly varying part is

$$\frac{\partial \overline{f}}{\partial t} - \overline{\delta \mathbf{E} \cdot \nabla_p \delta f} = \frac{\partial \overline{f}}{\partial t} - \nabla_p \cdot \left(\overline{\delta \mathbf{E} \delta f} \right) = 0,$$
$$0 = 1 - \int \overline{f} d\mathbf{p}. \tag{B.1}$$

The equation for the rapidly varying part reads,

$$\frac{\partial \delta f}{\partial t} + \nabla_p \mathcal{E}(\mathbf{p}) \cdot \nabla_x \delta f - \delta \mathbf{E} \cdot \nabla_p \overline{f} - \left(\delta \mathbf{E} \cdot \nabla_p \delta f - \overline{\delta \mathbf{E} \cdot \nabla_p \delta f} \right) = 0$$
$$\nabla_x \cdot \delta \mathbf{E} = 1 - \int \delta f d\mathbf{p} \tag{B.2}$$

Note that $\overline{\delta \mathbf{E} \delta f}$ in Equation(B.1) plays the same role as the Reynolds stress tensor in RANS. The model for it can be derived from Equation(B.2). Ultimately we will obtain a system for averaged *pdf* $\overline{f}(\mathbf{p}, t)$ and the spectral energy density of waves $W(\mathbf{k}, t) \propto \widehat{\delta \mathbf{E}}^* \widehat{\delta \mathbf{E}}$:

$$\partial_t \overline{f} = B(W, \overline{f}),$$
$$\partial_t W = H(\overline{f}, W).$$

Drop the bar over f and define $N(\mathbf{k}, t) = W(\mathbf{k}, t)/\hbar\omega(\mathbf{k})$, it is then identical to Equation(2.13).

Works Cited

Aurora diagram. https://scijinks.gov/aurora/auroradiagram.png. Accessed: 2023-11-21.

Schematic of a tokamak chamber and magnetic profile. https://www.researchgate.net/figure/Schematic-of-a-tokamak-chamber-and-magnetic-profile_fig1_274920774. Accessed: 2023-11-21.

Boumediene Abdellaoui, Ireneo Peral, and Magdalena Walias. Some existence and regularity results for porous media and fast diffusion equations with a gradient term. *Transactions of the American Mathematical Society*, 367(7):4757–4791, 2015.

Pavel Aleynikov and Boris Breizman. Stability analysis of runaway-driven waves in a tokamak. *Nuclear Fusion*, 55(4):043014, 2015.

Douglas N Arnold, Franco Brezzi, Bernardo Cockburn, and L Donatella Marini. Unified analysis of discontinuous galerkin methods for elliptic problems. *SIAM journal on numerical analysis*, 39(5):1749–1779, 2002.

Claude Bardos and Nicolas Besse. Diffusion limit of the vlasov equation in the weak turbulent regime. *Journal of Mathematical Physics*, 62(10):101505, 2021.

Nicolas Besse, Yves Elskens, DF Escande, and Pierre Bertrand. Validity of quasilinear theory: refutations and new numerical confirmation. *Plasma Physics and Controlled Fusion*, 53(2):025012, 2011.

Iwo Bialynicki-Birula and Philip J Morrison. Quantum mechanics as a generalization of nambu dynamics to the weyl-wigner formalism. *Physics Letters A*, 158(9):453–457, 1991.

Mihai Bostan. Transport equations with disparate advection fields. application to the gyrokinetic models in plasma physics. *Journal of Differential Equations*, 249(7):1620–1663, 2010.

Boris N Breizman, Pavel Aleynikov, Eric M Hollmann, and Michael Lehnen. Physics of runaway electrons in tokamaks. *Nuclear Fusion*, 59(8):083001, 2019.

Yingda Cheng, Irene M Gamba, Fengyan Li, and Philip J Morrison. Discontinuous galerkin methods for the vlasov–maxwell equations. *SIAM Journal on Numerical Analysis*, 52(2):1017–1049, 2014.

Daniele Antonio Di Pietro and Alexandre Ern. *Mathematical aspects of discontinuous Galerkin methods*, volume 69. Springer Science & Business Media, 2011.

Akio Doi and Akio Koide. An efficient method of triangulating equi-valued surfaces by using tetrahedral cells. *IEICE TRANSACTIONS on Information and Systems*, 74(1):214–224, 1991.

WE Drummond and D Pines. Non-linear stability of plasma oscillations. 1962.

Richard D Hazeltine and François L Waelbroeck. *The framework of plasma physics*. CRC Press, 2018.

Ross E Heath, Irene M Gamba, Philip J Morrison, and Christian Michler. A discontinuous galerkin method for the vlasov–poisson system. *Journal of Computational Physics*, 231(4):1140–1174, 2012.

Kun Huang and Irene M Gamba. Existence of global weak solutions to quasilinear theory for electrostatic plasmas. *arXiv preprint arXiv:2304.12430*, 2023.

Kun Huang, Michael Abdelmalik, Boris Breizman, and Irene M Gamba. A conservative galerkin solver for the quasilinear diffusion model in magnetized plasmas. *Journal of Computational Physics*, 488:112220, 2023a.

Kun Huang, Irene M Gamba, and Chi-Wang Shu. Structure-preserving solvers for particle-wave interaction in non-uniform magnetized plasmas. *in preparation*, 2023b.

AA Ivanov and LI Rudakov. Quasilinear relaxations dynamics of a collisionless plasma. *Sov. Phys.—JETP*, 24:1027, 1967.

Seong-Yeop Jeong, Daniel Verscharen, Robert T Wicks, and Andrew N Fazakerley. A quasi-linear diffusion model for resonant wave–particle instability in homogeneous plasma. *The Astrophysical Journal*, 902(2):128, 2020.

Allan N Kaufman. Resonant interactions between particles and normal modes in a cylindrical plasma. *The Physics of Fluids*, 14(2):387–397, 1971.

CF Kennel and F Engelmann. Velocity space diffusion from weak plasma turbulence in a magnetic field. *The Physics of Fluids*, 9(12):2377–2388, 1966.

Dmitrii Kiramov and Boris Breizman. Reduced quasilinear treatment of energetic electron instabilities in nonuniform plasmas. In *APS Division of Plasma Physics Meeting Abstracts*, volume 2021, pages JP11–068, 2021.

OA Ladyzhenskaya, VA Solonnikov, and NN Ural'tseva. Linear and quasilinear equations of parabolic type, transl. math. *Monographs, Amer. Math. Soc*, 23, 1968.

L Landau. On the vibrations of the electronic plasma. *Zhurnal eksperimentalnoi i teoreticheskoi fiziki*, 16(7):574–586, 1946.

LD Landau. Die kinetische gleichung fuer den fall coulombscher wechselwirkung. *Phys. Z. Sowjet. 10 (1936)*, 10:163, 1936.

Rüdiger Landes and Vesa Mustonen. On parabolic initial-boundary value problems with critical growth for the gradient. In *Annales de l'IHP Analyse non linéaire*, volume 11, pages 135–158, 1994.

I Lerche. Quasilinear theory of resonant diffusion in a magneto-active, relativistic plasma. *The Physics of Fluids*, 11(8):1720–1727, 1968.

Chang Liu, Eero Hirvijoki, Guo-Yong Fu, Dylan P Brennan, Amitava Bhattacharjee, and Carlos Paz-Soldan. Role of kinetic instability in runaway-electron avalanches and elevated critical electric fields. *Physical review letters*, 120(26): 265001, 2018.

Chohong Min and Frédéric Gibou. Geometric integration over irregular domains with application to level-set methods. *Journal of Computational Physics*, 226 (2):1432–1443, 2007.

Yoichiro Nambu. Generalized hamiltonian dynamics. *Physical Review D*, 7(8): 2405, 1973.

Joachim Nitsche. Über ein variationsprinzip zur lösung von dirichlet-problemen bei verwendung von teilräumen, die keinen randbedingungen unterworfen sind. In *Abhandlungen aus dem mathematischen Seminar der Universität Hamburg*, volume 36, pages 9–15. Springer, 1971.

Gergö Pokol, Tünde Fülöp, and Mietek Lisak. Quasi-linear analysis of whistler waves driven by relativistic runaway beams in tokamaks. *Plasma Physics and Controlled Fusion*, 50(4):045003, 2008.

VD Shapiro and VI Shevchenko. On the nonlinear theory of interaction between charged particle beams and a plasma in a magnetic field. *Zhur. Eksptl'. i Teoret. Fiz.*, 42, 1962.

Thomas H Stix. *Waves in plasmas*. Springer Science & Business Media, 1992.

Kip S Thorne and Roger D Blandford. *Modern classical physics: optics, fluids, plasmas, elasticity, relativity, and statistical physics*. Princeton University Press, 2017.

Juan Luis Vázquez. *The porous medium equation: mathematical theory*. Oxford University Press on Demand, 2007.

AA Vedenov, EP Velikhov, and RZ Sagdeev. Nonlinear oscillations of rarified plasma. *Nuclear Fusion*, 1(2):82, 1961.

Chenglong Zhang and Irene M Gamba. A conservative scheme for vlasov poisson landau modeling collisional plasmas. *Journal of Computational Physics*, 340: 470–497, 2017.

Chenglong Zhang and Irene M Gamba. A conservative discontinuous galerkin solver for the space homogeneous boltzmann equation for binary interactions. *SIAM Journal on Numerical Analysis*, 56(5):3040–3070, 2018.

Xiangxiong Zhang and Chi-Wang Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229(9):3091–3120, 2010.

Vita

Kun Huang, son of Xiaobing Huang and Xiaolan Liu, was born in Xinyu, Jiangxi Province, China, on March 25th, 1998. He received a Bachelor of Science degree in Theoretical and Applied Mechanics from Peking University in 2019. After that, he joined the Computational Science, Engineering, and Mathematics doctoral program at the University of Texas at Austin. He has been working under the advising of Professor Irene M. Gamba in the area of kinetic equations since 2020.

Address: k_huang@utexas.edu

This dissertation was typeset with LaTeX[†] by the author.

[†]LaTeX is a document preparation system developed by Leslie Lamport as a special version of Donald Knuth's TeX Program.