The Application of a High Order Semi-Lagrangian Scheme

By

Szu-Yi Chen

A dissertation submitted in partial fulfillment of

the requirements for the degree of

Doctor of Philosophy

(Electrical and Computer Engineering)

at the

UNIVERSITY OF WISCONSIN-MADISON

2015

Date of final oral examination: 11/20/2015

The dissertation is approved by the following members of the Final Oral Committee:

William Hitchon, Professor, Electrical and Computer Engineering

David Anderson, Professor, Electrical and Computer Engineering

Shi Jin, Professor, Mathematics

Irena Knezevic, Professor, Electrical and Computer Engineering

Amy Wendt, Professor, Electrical and Computer Engineering

Acknowledgments

Finally, I am graduating. It has been really hard to finish this degree. I would like to thank many people. First, I really thank my advisor, Prof. Hitchon. During these years, you always led me to the right way to think about my research. When I faced problems, you always patiently discussed things with me and gave me many valuable suggestions. When I made some mistakes, you always forgave me, taught me and encouraged me. I could not have made it without you. Thank you so much, Prof. Hitchon. You are the best advisor!

I would like to thank Profs. Jin, Wendt, and Knezevic for serving on my committee. You gave me so much good advice. I would also like to thank Prof. Anderson. You spent much time helping me prepare my defense. When I have any questions, you always teach me patiently. Thanks a lot.

In addition, I need to thank Prof. Walker. He helped me correct my thesis, he also gave me some useful suggestions for my thesis. I would like to thank my friends. We share our happiness and sadness. I will never forget the days that I spent with you. I extend my thanks to my colleagues, Yaman Güçlü, David Sirajuddin for their assistance. Special words of thanks to Yaman Güçlü who gave me many professional suggestions.

Home is always a safe haven with lots of love and support. Hence, my family is the most important. They support me spiritually as well as economically. When I encounter problems, they encourage me and advise me how to solve them. My family is the best in the world!

Abstract

A semi-Lagrangian numerical method is presented in our study, which is automatically mass conservative and positivity preserving. The Convected scheme (CS) refers to a 'forward-trajectory' semi-Lagrangian method for solving the Boltzmann equation and for plasma kinetic simulation. In this thesis, we explored the following issues: (1) developing a new version of the Convected scheme with high order accuracy and efficiency; (2) applying the Convected scheme to solve engineering problems such as plasma breakdown and quantum transport of electrons.

The basic CS has 2^{nd} order local truncation error (L.T.E.) in space, which is not accurate enough. Thus, a higher order CS is necessary, and a new method (the MF scheme) was developed and derived. The convergence analysis of the 1D continuity equation was first tested (for a stationary wave with time independent and a pulsating problem with time dependence) to ensure that our scheme is high order when the Courant number is greater than one. 1D breakdown simulations were then solved. Breakdown simulation should ideally permit time steps up to or greater than the Courant-Friedrichs-Lewy (CFL) limit and the numerical results proved the ability to accomplish this efficiently on a uniform mesh. Next, for the constant advection equation, an arbitrarily high order CS (the f_{22} scheme) was developed. Previously, the f_{22} scheme was used to solve 1D-1V Vlasov-Poisson systems such as linear Landau damping and the 'bump-on-tail' instability. Here, we extended the arbitrarily high order f_{22} scheme to other problems such as the Wigner equation and the Vlasov equation with a discontinuous potential. Moreover, the convergence analysis of the solution to the Wigner equation and the Vlasov equation with a discontinuous potential demonstrated that our scheme is efficient and accurate.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis describes research consisting of novel numerical simulations applied to a) plasma breakdown and b) the Wigner equation. The description of electrical breakdown in a gas (a phenomenon that we refer to as "gas breakdown", or simply "breakdown") is of considerable importance [1–8] and is also difficult to accurately achieve. Since the strongest growth in free electron density occurs at the front of the breakdown region, the spurious density caused by numerical diffusion can grow faster than the 'ideal' density and potentially exceed it. In addition, in the system of equations used here, the long time solution exhibits a discontinuity in the gradient of the density at the breakdown front. This makes breakdown a challenging test of a numerical scheme.

The Wigner equation has been applied in many fields, such as plasma, optics and semiconductor devices. In the past few decades, the Boltzmann transport equation was used to describe semiconductor devices, such as a metal-oxide-semiconductor field-effect transistor (MOSFET). However, when the size of the devices enters nano-scales, quantum effects need to be considered, and the Wigner equation becomes more important.

This chapter introduces the background. We explain the issues involved in Section 1.1. Next, we compare many different models and discuss why we believe that the

Wigner equation is the best candidate for quantum devices and the Convected scheme is a good tool for solving the advection equation. In Section 1.2, we introduce how we use a Convected Scheme (CS), combined with a splitting method, to describe plasma breakdown and solve the Wigner equation. The CS is a method of characteristics and provides an accurate method of solving plasma equations and the Boltzmann transport equation. Here, we want to apply the CS to solve the Wigner equation as well. The splitting method is a useful tool, which can split a vector field into a sum of two or more parts and these parts can be integrated more easily than the original one. Splitting methods include, for example, Lie splitting, Strang splitting and higher order splitting such as 4th order and 6th order splitting. In this document, we use Strang splitting (2nd order accurate) and 4th order splitting on the plasma breakdown case and Strang splitting on the Wigner equation to verify the scheme. Before starting to solve the Wigner equation, plasma breakdown is first solved by the CS and splitting method. By simulating plasma breakdown, we can learn how to use the CS and splitting method to predict accurately electrical properties. The Wigner equation is solved by analogous techniques.

1.1 Background

We now introduce the two areas where we have done simulations: plasma breakdown and quantum transport. Electrical breakdown in high pressure gases poses formidable challenges to numerical simulation. One of the principal problems is numerical diffusion, which allows electrons to travel ahead of the physical breakdown front: in this region of space the electric field has maximum intensity, the ionization level is highest, and even small amounts of spurious density can lead to qualitatively incorrect results.

On the other hand, in 1947, the bipolar point-contact transistor was invented at

Bell Telephone Laboratories by John Bardeen, Walter Brattain and William Shockley. In the 1950s and 1960s, the germanium transistor was more common. However, leakage current is easily generated with a germanium transistor. Now, silicon has been become the main material for semiconductors because silicon has superior physical and technological properties. During the past few decades, the development of semiconductors has grown quickly and they are now widely used, in applications such as automotive, medical, LED lighting and so on.

Due to the rapid development of semiconductor technology, device dimensions have decreased very rapidly. In order to improve the speed and functionality of integrated circuits (ICs), higher integration densities are necessary. In 1965, Gordon Moore predicted that transistor count would double every 18 months. Now, 16 nm technology has been produced, and in 2020, 7 nm technology will be introduced [9]. In the past few decades, devices were situated in the mesoscopic range, where electron transport can be described by the Boltzmann Transport Equation (BTE). The BTE is

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{x}} + \frac{\mathbf{F}}{m} \cdot \frac{\partial f}{\partial \mathbf{v}} = \left(\frac{\partial f}{\partial t} \right)_{collision} \tag{1.1}$$

where $\mathbf{F}(\mathbf{x}, \mathbf{v}, t)$ is the force field acting on each particle, and the term on the right-hand side is the *Boltzmann collision integral*. The BTE uses the distribution function, $f(\mathbf{r}, \mathbf{k}, t)$ to describe the probability of particles being at $(\mathbf{r}, \mathbf{k}, t)$. The main reason to use the BTE is that the distribution function f is a phase-space distribution function containing all of the information of interest about the carriers, such as their current and kinetic energy [11], and the phase-space distribution function is the most intuitive and efficient way to describe this information.

However, when the size of devices continues to shrink and enters the quantum region, the BTE has limitations. Because the BTE particle trajectories obey Newton's

laws, the position and momentum of the electrons can be exactly known at the same time. Since both the position and the momentum of a particle are arguments of the distribution function, apparently the quantum mechanical uncertainty principle $\Delta p \Delta r \geq \hbar$ is violated. Therefore, we need to find the proper equation to apply to nano-scale devices, and an efficient and intuitive model to solve this equation.

Instead of the Boltzmann transport equation (BTE), there are many choices such as the Schrödinger equation (SE), transfer-matrix (TM), density matrix (DM), Green's function (GF), Wigner function (WF) and so on. Each one has its own strengths and weaknesses. Comparing these methods, we believe that the Wigner function method may be the best candidate to simulate quantum devices for the reasons explained below [12].

The Wigner function was derived by Eugene Wigner in 1932. Recently, the Wigner function has become a popular tool and has been applied in many fields. Similar to the Boltzmann equation, the Wigner equation provides the quasi-distribution function from which we can obtain necessary information about the device, such as carrier density and current density. The main reason for using the Wigner equation is that the Wigner equation is analogous to the Boltzmann equation. Hence, the Wigner equation can adopt many common techniques used for the Boltzmann equation. Comparing with the NEGF and other models, the numerical complexity is much less for the Wigner equation.

Currently, Monte Carlo is one of the most widely used methods for solving the Wigner equation. The Monte Carlo method has been applied to the Boltzmann transport equation and has obtained successful results. Since 2002, the Monte Carlo approach has been extended to solve the WTE [13–16]. The Monte Carlo approach allows the study to include the whole scattering process. The benefit of using a Monte Carlo method to solve the Wigner equation is that the technique applied to the Boltzmann

equation can be used again.

In this work, we introduce a technique, called the Convected Scheme (CS). The term CS refers to a family of algorithms, most usually applied to the solution of the Boltzmann equation. In contrast to most finite difference schemes, the CS is not restricted by the Courant-Friedrichs-Lewy (CFL) criterion, so the CS can use a long time step which makes simulations more efficient and possibly more accurate. Moreover, the CS can reduce the numerical errors often associated with the Monte Carlo method. One purpose of this work is to demonstrate the advantages of CS scheme, both for breakdown and for broader classes of problems, such as the Wigner equation.

1.2 Conclusion

In this work, we extend a numerical method, the Convected Scheme or CS. The Convected Scheme (CS) is relatively intuitive and accurate and has been used for plasma kinetic simulation [17–21] and for solving the Boltzmann equation. Here, a new version of the CS (the MF scheme) is developed which has higher order accuracy in the presence of variable flow velocity, and still maintains the characteristics of the CS, including exact automatic conservation and positivity preservation. In addition, like earlier versions it can take time steps which exceed the CFL limit, achieving high efficiency. Further, we extend the CS to handle other equations, such as the Wigner equation and the Vlasov equation with a discontinuous potential, which are difficult and important to model.

Here, we employ two different CS schemes. One is the Moving Midpoint (MM) scheme, another is the Moving Face (MF) scheme. These two schemes will be discussed in detail later.

The overall approach is based on a technique (the CS) for solving advection equa-

tions which is combined with descriptions of other physical phenomena using a second (or higher) order splitting in time. The MF scheme is first tested for convergence in 1D breakdown simulations, which is sufficient to address the important behavior. Further, for the constant advection equation, an arbitrarily high order Convected scheme (AHOCS) was developed (the f_{22} scheme) which is based on the MM scheme. When the f_{22} scheme is employed on the constant advection equation, machine precision can be achieved rapidly. The Wigner equation contains a constant advection equation and a pseudo-differential operator. By taking advantage of the f_{22} scheme and the Fast Fourier transfer/inverse fast Fourier transform (FFT/IFFT), the Wigner equation can be solved by the CS. We solve the Wigner equation in 1D Landau damping and in the semi-classical regime (the scaled Planck constant (ε) is small). The solution demonstrates that our scheme is efficient. Finally, we combine the MF scheme and the f_{22} scheme to solve the Vlasov equation with a discontinuous potential to surpass the theoretical limit for finite differences. Encouraging results were obtained, exhibiting high accuracy and efficiency. All of these new contributions are described in this thesis.

This chapter contains the background and motivation for performing this work. An overview of the literature, such as the splitting method, the Wigner equation, Convected Scheme and other schemes, is introduced in Chapter 2. In Chapter 3, we present a simple analytic theory of breakdown in the frame of reference of the breakdown front. We develop a numerical scheme (the MF version of the Convected scheme), followed by results. We introduce an arbitrarily high order Convected scheme and employ this high order scheme to solve the Wigner equation and the Vlasov equation with a discontinuous potential in Chapter 4. Finally, we conclude this thesis in Chapter 5.

Chapter 2

Literature review

In this chapter, we review the relevant papers and theories that we will use in this document. In Section 2.1, we mention the implications for modeling of earlier work on breakdown, and difficulties encountered. In Section 2.2, we outline numerical methods for solving the Wigner equation, but especially those techniques for handling the problem of characteristics and the Boltzmann equation which can be adapted to the Wigner equation. The Wigner equation is the Fourier transform of the density matrix. Here, we introduce the Wigner function and the dynamic equation of the Wigner function. Finally, the Convected scheme will be introduced.

2.1 Plasma breakdown

Here, we review simulation of breakdown. While some work has been done on making the numerical scheme high order (which must not be confused with the use of high order moments of the fluid equations) and the difficulties due to numerical stiffness have been recognized, simulation of breakdown is still a relatively new topic. In [4, 5], the authors used a third order scheme to handle electron drift. They do not appear to

have coupled the electron drift to the other processes (ionization, updating the electric field) to create a scheme which is overall of high order, as discussed here. They remark that, as negative densities.

In [22], the authors present a splitting procedure consisting of *a*) Predictor step (chemical reaction + continuity equations are solved); *b*) Poisson solution; and *c*) Corrector step (same equations as predictor step). The corrector step is simply an iterated version of the predictor step using updated fields. The Poisson solution is (semi)-implicit but does not appear to involve the reactions described in the predictor step. The overall order of the scheme is not stated. The predictor and corrector steps are described using different forms of splitting, but it is stated that second order is not readily achieved in any case.

An important issue which is a focus of [23] is that the spatial scales involved in breakdown are very short indeed. Their response is a spatially adaptive solution in a standard Eulerian framework. As usual, the use of an explicit time integrator in the presence of small cells leads to unphysical time-step restrictions due to the Courant-Friedrichs-Lewy (CFL) stability limit.

Later we propose to overcome the CFL limit in breakdown simulations by using a simple semi-Lagrangian method: the Convected Scheme, or CS. Our numerical experiments amply demonstrate the efficiency of this approach on uniform meshes and we expect increased benefits when a simple refinement strategy is employed.

2.2 Wigner equation

Since its invention in 1932 by Wigner [24, 25], the Wigner equation has found applications in many physical fields, such as optics [26, 27] and chemistry [28].

Here, the Wigner transport equation (WTE) will be derived from the density matrix

as follows. Because the Wigner function describes a statistical ensemble of particles, we can use the density operator ρ to start the derivation of the WTE. The density matrix is defined in terms of the N-particle wave function [29–31]

$$\rho(z_1, z_2, t) = \sum_i p_i \psi_i(z_1) \psi_i^*(z_2) \tag{2.1}$$

where ψ_i is the wavefunction and p_i is the probability of state i.

In order to describe how the density operator changes in time, Liouville-von Neumann equation can be applied:

$$\frac{\partial \rho}{\partial t} = \frac{1}{i\hbar}[H, \rho] \tag{2.2}$$

H is the Hamiltonian and it can be expressed as:

$$H = \frac{-\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + V(z) \tag{2.3}$$

where m^* is the effective mass and V is the potential energy. Substituting Eq. (2.1) into Eq. (4.49), one can obtain:

$$\frac{\partial \rho(z_2, z_1)}{\partial t} = \frac{1}{i\hbar}\left\{\frac{-\hbar^2}{2m^*}\left[\frac{\partial^2}{\partial z_2^2} - \frac{\partial^2}{\partial z_1^2}\right]\rho(z_2, z_1) + [V(z_2) - V(z_1)]\rho(z_2, z_1)\right\} \tag{2.4}$$

For convenience in future, we need to transform z_1 and z_2 into new coordinates.

$$z_1 = z + \frac{z'}{2} \text{ and } z_2 = z - \frac{z'}{2} \tag{2.5}$$

Again, we transform the second derivative in Eq. (2.4) to the new coordinates according the chain rule.

$$\begin{cases} \dfrac{\partial}{\partial z} = \dfrac{\partial}{\partial z_1}\dfrac{\partial z_1}{\partial z} + \dfrac{\partial}{\partial z_2}\dfrac{\partial z_2}{\partial z} = \dfrac{\partial}{\partial z_1} + \dfrac{\partial}{\partial z_2} \\ \dfrac{\partial}{\partial z'} = \dfrac{\partial}{\partial z_1}\dfrac{\partial z_1}{\partial z'} + \dfrac{\partial}{\partial z_2}\dfrac{\partial z_2}{\partial z'} = \dfrac{1}{2}\dfrac{\partial}{\partial z_1} - \dfrac{1}{2}\dfrac{\partial}{\partial z_2}\dfrac{\partial z_2}{\partial z'} \end{cases}$$

Finally, we get the relationship between z, z' and z_1, z_2:

$$\frac{\partial}{\partial z}\frac{\partial}{\partial z'} = \frac{1}{2}\left[\frac{\partial^2}{\partial z_1^2} - \frac{\partial^2}{\partial z_2^2}\right] \tag{2.6}$$

Eq. (2.4) can be written in the new coordinates:

$$\frac{\partial \rho(z-\frac{z'}{2}, z+\frac{z'}{2})}{\partial t} = \frac{\hbar}{im}\frac{\partial}{\partial z}\frac{\partial}{\partial z'}\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) + \frac{1}{i\hbar}\left(V(z-\frac{z'}{2}) + V(z+\frac{z'}{2}))\right)\rho(z-\frac{z'}{2}, z+\frac{z'}{2}), \tag{2.7}$$

In order to connect the WF with the density matrix, we introduce the Fourier transform and inverse Fourier transform here.

$$f(\xi) = \frac{1}{\sqrt{2\pi}}\int f(x)e^{-ix\xi}dx,\ \ f(x) = \frac{1}{\sqrt{2\pi}}\int f(\xi)e^{ix\xi}d\xi. \tag{2.8}$$

The Wigner function is the Fourier transform of the density matrix, also called the Weyl-Wigner transform [30], i.e.

$$f(z,v) = \frac{m}{2\pi\hbar}\int dz' \exp(\frac{imv}{\hbar}z')\psi^*(z-\frac{z'}{2})\psi(z+\frac{z'}{2}). \tag{2.9}$$

On the other hand, the density matrix is the inverse Fourier transform of the WF.

$$\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) = \int dv' \exp(-\frac{imv'}{\hbar}z')f(z,v') \tag{2.10}$$

In order to get the WTE, which is the dynamic equation of the WF, the dynamic equation of the WF can written as:

$$\frac{\partial f}{\partial t} = \frac{m}{2\pi\hbar}\int dz' \exp(\frac{imv}{\hbar}z')\frac{\partial \rho}{\partial t} \tag{2.11}$$

substituting Eq. (2.7) into Eq. (2.11), we get:

$$\begin{aligned}\frac{\partial f}{\partial t} = &\frac{m}{2\pi\hbar}\int dz' \exp(\frac{imv}{\hbar}z')\frac{\hbar}{im}\frac{\partial}{\partial z}\frac{\partial}{\partial z'}\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) \\ &+ \frac{m}{2\pi\hbar}\int dz' \exp(\frac{imv}{\hbar}z')\frac{1}{i\hbar}\left(V(z-\frac{z'}{2})+V(z+\frac{z'}{2}))\right)\rho(z-\frac{z'}{2}, z+\frac{z'}{2})\end{aligned} \tag{2.12}$$

Now, we look at the first term the right hand side of Eq. (2.12),D_1 and use integration by parts to deal with it. According to the integration by parts at z', we know $\int uv' = uv - \int u'v$.

$$D_1 = \frac{1}{2i\pi}\frac{\partial}{\partial z}\int dz' \exp(\frac{imv}{\hbar}z')\frac{\partial}{\partial z'}\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) \tag{2.13}$$

where,

$$\int dz' \exp(\frac{imv}{\hbar}z')\frac{\partial}{\partial z'}\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) = \exp(\frac{imv}{\hbar}z')\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) - \int dz' \frac{imv}{\hbar}\exp(\frac{imv}{\hbar}z')\rho(z-\frac{z'}{2}, z+\frac{z'}{2})$$

and we assume that:

$$\lim_{x\to\pm\infty}\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) = 0$$

Therefore,

$$D_1 = -\frac{mv}{2\pi\hbar}\frac{\partial}{\partial z}\int dz' \exp(\frac{imv}{\hbar}z')\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) \tag{2.14}$$

Comparing with Eq. (2.9), Eq. (2.14) can be rewritten as:

$$D_1 = -v\frac{\partial f}{\partial z} \tag{2.15}$$

where $v(k)$ is the velocity. As for the second term on the right hand side of Eq. (2.12), D_2, this can be written according to the inverse Fourier transform as follows:

$$D_2 = \frac{m}{2\pi\hbar}\frac{1}{i\hbar}\int dz' \exp(\frac{imv}{\hbar}z')\left(V(z-\frac{z'}{2}) - V(z+\frac{z'}{2})\right)\rho(z-\frac{z'}{2}, z+\frac{z'}{2}), \tag{2.16}$$

where $\rho(z-\frac{z'}{2}, z+\frac{z'}{2}) = \int dv' \exp(-\frac{imv'}{\hbar}z')f(z,v')$. Eq. (2.16) can be rewritten as:

$$D_2 = \frac{m}{2i\pi\hbar^2}\int\int dz'dv' \exp(\frac{im(v-v')}{\hbar}z')\left(V(z-\frac{z'}{2}) - V(z+\frac{z'}{2})\right)f(z,v'), \tag{2.17}$$

Finally, the Wigner dynamic equation is:

$$\frac{\partial f(z,v,t)}{\partial t} + v\frac{\partial f(z,v,t)}{\partial z} - \frac{m}{2i\pi\hbar^2}\int\int dz'dv' \exp(\frac{im(v-v')}{\hbar}z')\left(V(z-\frac{z'}{2}) - V(z+\frac{z'}{2})\right)f(z,v',t) = 0 \tag{2.18}$$

After solving the WTE, f_w is found. Then, the interesting results are the density and current.

$$n(r,t) = \int dv f_w(r,v,t) \tag{2.19}$$

$$J(r,t) = \int dv\ v f_w(r,v,t) \tag{2.20}$$

2.2.1 Methods for solving the Wigner equation

In this section, we introduce the different methods for solving the Wigner equation. An efficient and accurate simulation is important for designing and developing the techniques. The Wigner equation is expressed in Eq. (2.18). Here we will review the varieties of methods which are applied to deal with the Wigner equation and the potential operator.

First, we review the finite difference method, which is the most intuitive method. We then discuss the Monte Carlo simulation. Finally, we introduce other methods.

2.2.1.1 Finite Difference Method

Many methods have been found for solving the WTE. Frensley first solved it by using the first-order upwind±downwind difference scheme to simulate a resonant tunneling diode (RTD) in 1987 [32, 33]. In this case, the WTE can be written as:

$$\frac{\partial f(x,k)}{\partial t} = -\frac{1}{\hbar}\sum_{k'} V(x,k-k') - \frac{\hbar k}{m\Delta_x} \times f(x+\Delta_x) - f(x)];\ k<0$$
$$\frac{\partial f(x,k)}{\partial t} = -\frac{1}{\hbar}\sum_{k'} V(x,k-k') - \frac{\hbar k}{m\Delta_x} \times [f(x) - f(x-\Delta_x) - f(x)];\ k>0$$

The potential operation can also be expressed in the discretized form:

$$V(x,k) = \frac{2}{N_k}\sum_{\xi} sin(k\xi)[v(x+\frac{1}{2}\xi) - v(x-\frac{1}{2}\xi)]$$

where Δ_x is the mesh space and N_k is the number of mesh points.

After that, the higher order upwind±downwind difference scheme was introduced [34–36]. On the other hand, in order to get an accurate electron distribution, the space charge also needs to be considered. The WTE is coupled with the Poisson equa-

tion [37, 38] for obtaining a self-consistent solution. Again, the finite difference scheme can be applied to discretize the Poisson equation.

In order to compare the result computed by the finite difference method with the exact solution and with the solution computed by the NEGF, [39] describes the simulation results of the resonant tunneling diode (RTD) at the different bias. From the results, they show that the solution of the NEGF is very similar to the exact solution. However, the solution of the Wigner equation approaches the exact solution only when the length of the contact is large. The authors think this is due to the inflow boundary condition of the Wigner equation being exact when the boundary condition is at infinity.

2.2.1.2 Monte Carlo

As we mentioned above, the Monte Carlo method is acknowledged as an important tool for solving the Boltzmann transport equation. For the Wigner equation, it is a phase-space description as is the Boltzmann transport equation. Therefore, it is possible to use analogous techniques which are used for the BTE applied to the Wigner equation. Unlike the BTE, where the distribution is always positive, the WTE contains a nonlocal term, and this nonlocal term may create some negativity, resulting in a quasi-distribution. This is the main difference between the BTW and the WTE. Therefore, in order to simulate quantum devices by the Monte Carlo method, a new term, the affinity (A_j), is introduced [30].

The affinity is a weighting given to the particle, and this weighting can be negative for a negative distribution. The affinity can be expressed as [30]:

$$\sum_{i \in M(x,k)} \frac{dA_i}{dt} = Q f_w$$

where $M(x,k)$ is a mesh of the phase-space.

Therefore, the affinity is updated at each step, and it allows the Monte Carlo method to be applied to the Wigner equation for the both positively and negatively valued Wigner function. In some papers [40, 41], the authors compare the solutions computed by the Wigner equation for RTD and MOSFET. When the authors used the Monte Carlo method to solve the Wigner equation, there is good agreement with the NEGF.

2.2.1.3 Gaussian beam

The Gaussian beam method was used by Helle in 1975 to solve the Schrödinger equation [42]. In [43], the author used a phase space Gaussian beam (PSGB) method for solving the Wigner equation. As the name implies, the Gaussian beam (GB) solution is the basic element for this method and it can be expressed as follows:

$$\phi^{\epsilon}(t, x_{,0}) = A(t,y)e^{iT(t,x,)/\epsilon} \tag{2.21}$$

where $y = y(t, y_0)$ is the center of the beam.

There are three steps to be implemented: 1) Use the sum of GBs to describe the initial value of the function and then get the initial values of the PSGBs, 2) Update the evolution equations for PSGB, and 3) Form an approximate solution of the Wigner equation by summing the solutions of the PSGBs. The truncation error of the Taylor expansion is used to gauge the accuracy of the beam. The main advantages of the PSGB are 1) computational cost can be reduced significantly, and 2) one can get an accurate approximation of the Wigner equation when the potential is discontinuous. When the potential is discontinuous, the interface conditions need to be applied to handle discontinuities. Hence, the PSGB can deal with the potential whether it is

continuous or not.

2.3 Convected Scheme

In this section, we will now introduce the Convected Scheme (CS). The CS provides a solution of continuity and kinetic equations [17, 44–47] which has a number of desirable characteristics, including: 1) it is efficient, 2) it is conservative, 3) it preserves positivity, and 4) it has low phase-error.

In contrast to most finite difference schemes, the CS is not restricted by the Courant-Friedrichs-Lewy (CFL) criterion, so the CS can use a long time step which makes simulations more efficient and possibly more accurate.

The CS solution method envisages a moving-cell (MC), $C(t)$, originating from C_0 at $t = t_0$. At the end of the time step, the particles in the moving cell are remapped back onto the Eulerian mesh as shown in Fig. 2.1 and move according to the Lagrangian trajectory $(\mathbf{x}(t),\ \mathbf{v}(t))$:

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{t}) \\ \\ \frac{d\mathbf{v}}{dt} = \frac{1}{m}\mathbf{F}(t, \mathbf{x}, \mathbf{v}) \end{cases} \tag{2.22}$$

with initial conditions $(\mathbf{x}_0, \mathbf{v}_0)$ at $t = t_0$.

Recently, Y. Güçlü [48] *et al.* applied small corrections $(\tilde{U})$ to the final position of the moving cells before remapping, in order to increase the spatial accuracy from 2nd to 4th order. The starting point is the 1D advection equation:

$$\frac{\partial n}{\partial t} + \frac{\partial (un)}{\partial x} = 0$$

where $n(x,t)$ is the density and $u(x,t)$ is the velocity. The normalized displacement of

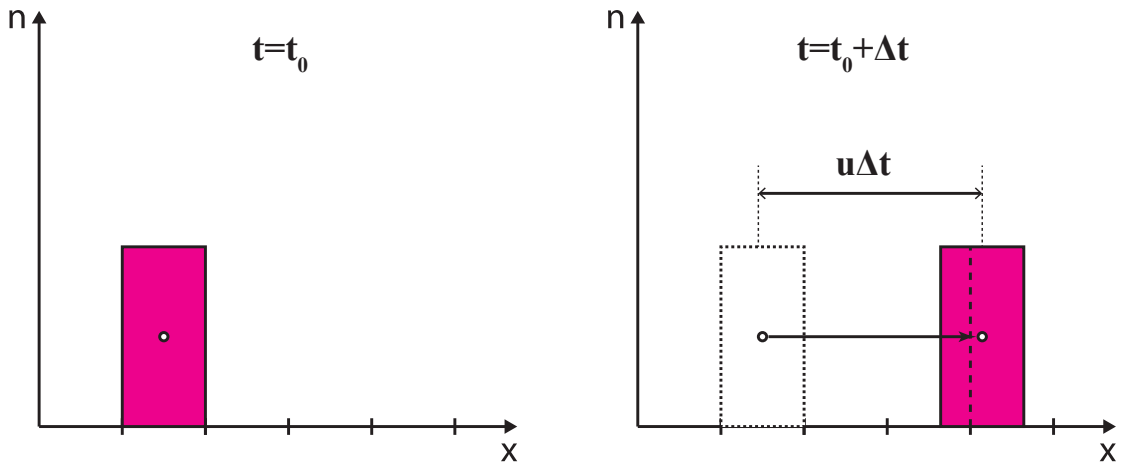

Figure 2.1: Ballistic operator of the moving center (MC) scheme. The right frame shows the density reallocated to cells, according to the MC scheme.

the moving cell coincides with the Courant number

$$U_0 := \left| \frac{u\Delta t}{\Delta x} \right|,$$

The 1D CS scheme can be rewritten as

$$n_{i+S}^{k+1} = \begin{cases} U_{i-1}^k n_{i-1}^k + (1 - U_i^k) n_i^k & \text{if } U \geq 0, \\ (1 + U_i^k) n_i^k - U_{i+1}^k n_{i+1}^k & \text{otherwise,} \end{cases} \tag{2.23}$$

Here, i denotes space, k means time and U is the fractional normalized displacement. If there is no correction, then $U(x,t) = U_0$. In order to get the high order CS, a small correction to the final position of the moving cell is employed. Hence, $U = U_0 + \tilde{U} = [u + \tilde{u}]\frac{\Delta t}{\Delta x}$, where $\tilde{U}$ is an 'anti-diffusive correction' [49, 50].

Later, in [51], Y. Güçlü *et al.* introduced a new improved scheme to calculate the product of (Un), which is a form of the high order derivatives for the constant advection equation, to extend the CS to arbitrarily high order of accuracy in time and space. In that paper, Y. Güçlü described two interpolations to construct the product of (Un) to

get a high order result.

One is polynomial interpolation, which uses the centered finite difference approximation to get the required derivatives. The updated density can be achieved by performing a local polynomial interpolation. Another is trigonometric interpolation, which is applied in cases with periodic boundary conditions, we call this scheme an f_{22} scheme. Because of the periodic boundary conditions, the high order derivatives can be efficiently computed by the fast Fourier transform (FFT) and the inverse fast Fourier transform (IFFT). For the new improved scheme, the orders of convergence are higher than for the old scheme. In addition, the f_{22} scheme can reach machine precision very quickly, so the efficiency can be improved greatly. Finally, the new scheme allows for coarser meshes and still retains accuracy for a long time. Therefore, the new improved scheme can be applied to problems with higher dimensionality and reduces memory requirements.

2.4 Summary

This chapter began with an overview of the development of the plasma breakdown. We introduced some earlier work on plasma breakdown, including what were used and what issues they faced. Following this, the Wigner equation was derived from the density matrix. The Wigner equation is the dynamic equation of the Wigner function. The Fourier transform and inverse Fourier transform connect the Wigner function with the density matrix. The Wigner function is the Fourier transform of the density matrix and the density matrix is the inverse Fourier transform of the Wigner function. Next, methods were introduced to solve the Wigner equation, such as a finite difference method and a Monte Carlo method. Finally, the Convected scheme was introduced. The Convected scheme solves the Continuity equation. It was then extended it to

solve the kinetic equation. Recently, improved schemes were developed which applied a small correction to the final position of the moving cells to improve accuracy and efficiency. These methods can achieve higher accuracy and still keep the desirable characteristics of the Convected scheme, including conservation, positivity and ease of implementation.

Chapter 3

Gas Breakdown

In this chapter, we focus on developing an approach for describing electrical breakdown. The essential feature of plasma is its ability to create, and respond to, electromagnetic fields. In most plasmas the production of electric fields depends critically on a fine balance between the densities of positive and negative species. It is essential therefore that a plasma simulation method conserve density, locally not just globally, and preserve positivity. We shall see below that breakdown simulations in particular naturally call for the resolution of space and time scales such that the Courant number is of order one. Here we present a scheme that achieves local conservation and positivity preservation in a natural way, and enables the use of Courant numbers throughout the range required for breakdown.

In order to solve the "breakdown equations", we propose a splitting scheme that combines a numerical solution of Poisson's equation and the plasma motion (handled by the CS) with the creation of ion-electron pairs due to ionization (handled analytically). The numerical error of the combined scheme is at least second order in the discreteness parameters Δx and Δt. As a result, accurate simulations are possible which require only moderate resolution in space and time.

One of the goals of this chapter is to develop a numerical solution of advection equations, based on the Convected Scheme, which applies analytic corrections to make the solution high order. The CS achieves positivity preservation by finding fractions of the initial cell which fall into each final cell, in this semi-Lagrangian scheme where density may travel across many mesh-spacings in a time step. The fractions themselves must be positive (or zero). Finding the final cells is straightforward in most cases, but in a small number of instances the list of final cells needs to be found accurately since small corrections will cause the entries in the list to jump. Handling these cases analytically is the greatest challenge in this work, and we shall devote some effort to demonstrating that good convergence is achieved by our method. So in summary, we propose and test a scheme for solving advection equations, which we believe is relatively simple (using a straightforward set of analytic expressions in the algorithm) and which has several desirable features (automatic local conservation, positivity, freedom from a CFL criterion).

In the chapter, we specify the model and the equations it employs, followed by a simple analytic theory of breakdown in the frame of reference of the breakdown front, and a discussion of the relevant time and space scales of breakdown in Section 3.1. In Section 3.2 we introduce the numerical scheme and derive expressions needed to achieve second and third orders of accuracy. Section 3.4 presents results of convergence studies for test cases involving only advection. Section 3.5 consists of results for breakdown, again emphasizing convergence, and Section 3.6 gives the conclusions.

3.1 Model Equations

The 1D breakdown problem can be described by the temporal evolution of the macroscopic quantities: $n_e(x,t)$, the electron number density [m^{-3}]; $n_i(x,t)$, the ion number

density [m^{-3}] and $E(x,t)$, the x component of the electric field [V/m].

Electron-ion pairs are produced by electron-impact ionization of the background gas; assuming that the breakdown event happens on the electron timescale, electrons are then able to drift in the electric field, while the slower ions remain at rest at the location where they were created.

For breakdown at atmospheric pressure, the electron mean free path for both elastic and inelastic collisions with the background gas is very small. Hence, we will assume that electrons are in local equilibrium with the electric field, so that all the kinetic coefficients (e.g. mobility and collision rates) will depend on the local electric field only. If we further neglect electron diffusion, the electron density satisfies an advection-reaction partial differential equation (PDE),

$$\frac{\partial n_e}{\partial t} + \frac{\partial (u n_e)}{\partial x} = \nu_{\mathrm{iz}}\, n_e, \tag{3.1}$$

where $u(E)$ is the electron drift velocity [m/s] along x and $\nu_{\mathrm{iz}}(E)$ is the ionization rate [s^{-1}], which will be discussed later. (Photoionization [52, 53] is potentially important but is beyond the scope of the present paper). On the other hand, the ion density satisfies an ordinary differential equation (ODE) of the form

$$\frac{\partial n_i}{\partial t} = \nu_{\mathrm{iz}}\, n_e, \tag{3.2}$$

while the instantaneous electric field can be determined by solving Gauss' law

$$\frac{\partial E}{\partial x} = \frac{q}{\epsilon_0}(n_e - n_i), \tag{3.3}$$

where $q < 0$ is the electron charge [C] and ϵ_0 is the vacuum permittivity [F/m], with the proper boundary conditions.

Eqs. (3.1), (3.2) and (3.3) are non-linearly and non-locally coupled to each other, and they should be solved simultaneously.

3.1.1 Energy-Based Formulation of The Source Term

The source term on the right-hand side of Eqs. (3.1) and (3.2) represents the production rate of electron-ion pairs in the unit volume [$\mathrm{m}^{-3}\mathrm{s}^{-1}$], due to electron impact ionization of the neutral gas background. Since the ionization process converts electron kinetic energy into atomic potential energy, and electrons in their turn gain kinetic energy from the electric field through Ohmic heating, we can conveniently reformulate the ionization rate in terms of energy-based quantities:

- The power per unit volume [$\mathrm{W/m}^3$] delivered to the electrons is $P_e = J \cdot E = qn_e uE$, where in general $uE < 0$, and hence the power is positive as expected;
- Being in local thermal equilibrium, the electrons dissipate electric power through elastic and inelastic collisions with the gas background; specifically, a fraction $\alpha(E)$ of the power goes into ionization, so that $P_{\mathrm{iz}} = \alpha P_e = \alpha\, qn_e uE$; clearly $0 < \alpha < 1$;
- Assuming the gas background is in the ground state, the creation of a single electron-ion pair requires an energy e_{iz} [J], that is the first ionization threshold, and hence the rate of creation of new electron-ion pairs in the unit volume is obtained by dividing the ionization power by the aforementioned energy, so that $S_{\mathrm{iz}} = P_{\mathrm{iz}}/e_{\mathrm{iz}} = \alpha\, qn_e uE/e_{\mathrm{iz}}$;
- By comparing the last formula with the source term in (3.1) we finally get

$$\nu_{\mathrm{iz}}(E) \;=\; \frac{q}{e_{\mathrm{iz}}}\,\alpha(E)u(E)E. \tag{3.4}$$

This last expression clarifies the relationship between the ionization frequency ν_{iz} and the drift velocity u, which are not independent quantities. This is especially important when experimental data are employed: setting $\alpha = 1$ in (3.4), which can only happen as $E \to \infty$, provides an upper bound on ν_{iz} that cannot be exceeded. Indeed, in order to ensure consistency between our parameters, in our analytical and numerical calculations we will always specify α and u, and then compute ν_{iz} using (3.4).

3.1.2 Analytic Theory

We assume that, after a first rapid transient, the velocity of the ionization front settles to a constant value v_f, while the electron and ion density profiles become stationary in a reference frame co-moving with the ionization front [54, 55]. Under this hypothesis, we look for a traveling-wave solution in the form

$$n_e(x,t) = \widetilde{n}_e(\xi), \qquad n_i(x,t) = \widetilde{n}_i(\xi), \qquad E(x,t) = \widetilde{E}(\xi),$$

where the change of coordinates $\xi = x - v_f t$ was used. Accordingly, Eqs. (3.1)- (3.3) become

$$\frac{d}{d\xi}\big[(u - v_f)\widetilde{n}_e\big] = \nu_{\text{iz}}\widetilde{n}_e, \tag{3.5a}$$

$$\frac{d}{d\xi}\big[-v_f\widetilde{n}_i\big] = \nu_{\text{iz}}\widetilde{n}_e, \tag{3.5b}$$

$$\frac{d\widetilde{E}}{d\xi} = \frac{q}{\epsilon_0}\big(\widetilde{n}_e - \widetilde{n}_i\big), \tag{3.5c}$$

Taking the difference of Eqs. (3.5a) and (3.5b) we get

$$\frac{d}{d\xi}\big[(u - v_f)\widetilde{n}_e + v_f\widetilde{n}_i\big] = 0,$$

which is simply a statement of the conservation of the current density $\widetilde{J}$ along the co-moving coordinate ξ:

$$\widetilde{J}(\xi) := q\big[(u - v_f)\widetilde{n}_e + v_f\widetilde{n}_i\big] \equiv constant.$$

We notice that $\widetilde{J}$ differs from the 'standard' current density $J = qun_e$ calculated in the laboratory reference frame, because of the term $qv_f(n_i - n_e)$ due to space-charge. Nevertheless, we know that behind the breakdown front the current is negligible ($J \simeq 0$), and the gas is quasi-neutral ($n_e \simeq n_i$); hence, it is reasonable to assume that $\widetilde{J} \equiv 0$, which yields a local relation between the electron and ion densities in the form

$$\frac{\widetilde{n}_i}{\widetilde{n}_e} = 1 - \frac{u}{v_f},$$

Substitution of the last equation into Gauss' law (3.5c) yields the expression

$$\widetilde{n}_e(\xi) = \frac{\epsilon_0 v_f}{qu}\frac{d\widetilde{E}}{d\xi}, \tag{3.6}$$

that will be used in the following derivation.

We now integrate both sides of Eq. (3.5a) from the location ξ^*, where the electric field is maximum ($E = E_{max}$), to a general position ξ. Integrating the left hand side gives

$$\int_{\xi^*}^{\xi} \frac{d}{d\xi'}\big[(u - v_f)\widetilde{n}_e\big]d\xi' = \big[(u - v_f)\widetilde{n}_e\big]_{\xi^*}^{\xi} = \big[u\big(\widetilde{E}(\xi)\big) - v_f\big]\widetilde{n}_e(\xi),$$

and integrating the right hand side gives

$$\begin{aligned}\int_{\xi^*}^{\xi} \nu_{\text{iz}}\widetilde{n}_e\, d\xi' &= \int_{\xi^*}^{\xi} \nu_{\text{iz}}\frac{\epsilon_0 v_f}{qu}\frac{d\widetilde{E}}{d\xi'}\, d\xi' = \frac{\epsilon_0 v_f}{q}\int_{E_{max}}^{\widetilde{E}(\xi)} \frac{\nu_{\text{iz}}(E')}{u(E')}\, dE' \\ &= -\frac{\epsilon_0 v_f}{e_{\text{iz}}}\int_{\widetilde{E}(\xi)}^{E_{max}} \frac{qu(E')E'\alpha(E')}{e_{\text{iz}}u(E')}\, dE' = -\frac{\epsilon_0 v_f}{e_{\text{iz}}}\int_{\widetilde{E}(\xi)}^{E_{max}} E'\alpha(E')\, dE'.\end{aligned}$$

By equating the two sides we get an expression for the electron density as a function of the local electric field $\widetilde{E}(\xi)$:

$$\widetilde{n}_e(\xi) = \frac{\epsilon_0}{e_{\rm iz}} \frac{v_f}{v_f - u\big(\widetilde{E}(\xi)\big)} \int_{\widetilde{E}(\xi)}^{E_{max}} E'\alpha(E')\,dE'. \tag{3.7}$$

A similar expression can be obtained for the ion density, as

$$\widetilde{n}_i(\xi) = \frac{\epsilon_0}{e_{\rm iz}} \int_{\widetilde{E}(\xi)}^{E_{max}} E'\alpha(E')\,dE'. \tag{3.8}$$

Eqs. (3.7) and (3.8) can be used to compute actual profiles once the functions $u(E)$ and $\alpha(E)$ are known, as well as the quantity v_f. For our preliminary tests, we will assume that the following simple relations hold:

$$u(E) = \mu E, \qquad \alpha(E) = \alpha_{max}\, E/E_{max}, \qquad v_f = (1+a)\mu E_{max}, \tag{3.9}$$

where $\mu(E)$ is the electron mobility (here assumed constant), $\alpha_{max} := \alpha(E_{max})$, and a is a small parameter used to adjust the velocity of the ionization front. Substituting the parameters from (3.9) into Eqs. (3.7) and (3.8) we obtain

$$\widetilde{n}_e(\xi) \;=\; \left(\frac{1}{2}\epsilon_0 E_{max}^2 \Big/ \frac{3}{2}\frac{e_{\rm iz}}{\alpha_{max}}\right) \left[1 - \left(\frac{\widetilde{E}}{E_{max}}\right)^3\right] \frac{1+a}{1+a-\widetilde{E}/E_{max}}, \tag{3.10}$$

$$\widetilde{n}_i(\xi) \;=\; \left(\frac{1}{2}\epsilon_0 E_{max}^2 \Big/ \frac{3}{2}\frac{e_{\rm iz}}{\alpha_{max}}\right) \left[1 - \left(\frac{\widetilde{E}}{E_{max}}\right)^3\right], \tag{3.11}$$

where the electric field $\widetilde{E}(\xi)$ is obtained by integrating Gauss' law, which assumes the

simple form

$$\frac{d\widetilde{E}}{d\xi} = \frac{q}{\epsilon_0}\left(\frac{1}{2}\epsilon_0 E_{max}^2 \Big/ \frac{3}{2}\frac{e_{\rm iz}}{\alpha_{max}}\right)\left[1-\left(\frac{\widetilde{E}}{E_{max}}\right)^3\right]\frac{\widetilde{E}/E_{max}}{1+a-\widetilde{E}/E_{max}}. \tag{3.12}$$

For breakdown of air at atmospheric pressure and standard temperature, we use $E_{max} = 4.999 \times 10^6\,\mathrm{V/m}$, $\alpha_{max} = 0.1$, $\mu = 0.01\,\mathrm{m^2/(Vs)}$ and $a = 0$. These expressions will be used to compare to numerical results in what follows. Results from the analytic solution are shown in Fig. 3.1.

3.1.3 Temporal and Spatial Scales of Interest

Here, we estimate the interesting scales for electric breakdown in air at atmospheric pressure; experimental evidence, as well as our analytical results, suggest that the electron density is of the order of 10^{18}–$10^{19}\,\mathrm{m^{-3}}$. The spatial and time scales for the cases are set by

1. The distance L_s over which the electric field is expected to be shielded, which is approximately $\epsilon_0 E_{max}/(nq)$, yielding $L_s \approx 500\,\mu\mathrm{m}$.

2. The distance L_{iz} over which an electron gains enough energy to create a new electron - which in our case is about ten times the distance in which the voltage changes by the ionization potential, since about $10\,\%$ of the energy goes into ionization. We estimate $L_{iz} \approx 40\,\mu\mathrm{m}$.

3. The dielectric relaxation time, $\tau_d \approx 10^{-8}\,\mathrm{s}$.

4. The rate of ionization, ν_{iz} is essentially the reciprocal of the time to travel a distance L_{iz}. The greatest speed is μE_{max} which is about $5 \times 10^4\,\mathrm{m/s}$. This gives $\nu_{iz} \approx 10^9\,\mathrm{Hz}$.

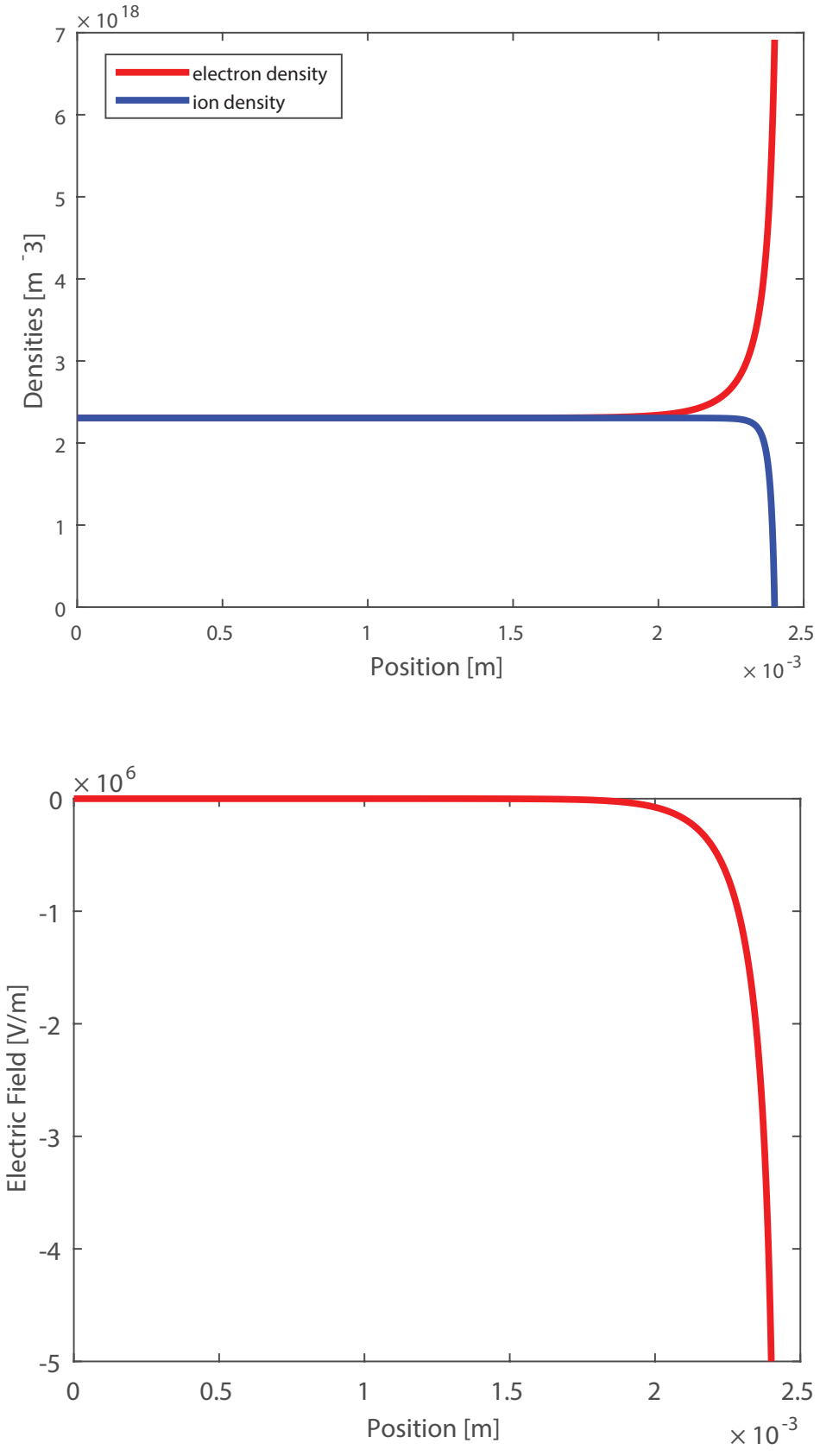

Figure 3.1: Analytic solution for electron and ion density (upper) and electric field (lower).

The length L_s and L_{iz} are each linked to a time scale (τ_d and $1/\nu_{iz}$, respectively) through the flow velocity, leading to a typical Courant number around one. In the cases which follow we examine a variety of Courant numbers, ranging from about 0.3 to of order 10.

3.2 Splitting Methods and Convected Scheme

In this section we introduce our scheme for solution of the continuity equation, which we refer to as the Convected Scheme. We first describe the Strang splitting which allows the overall model to be made second order [56]. A high-order splitting scheme is given below.

3.2.1 Strang Splitting

We now split the system of Eqs. (3.1), (3.2) and (3.3) into two separate systems:

A: Electron transport

$$\begin{cases} \dfrac{\partial n_e}{\partial t} = -\dfrac{\partial}{\partial x}[u(E)n_e] \\ \dfrac{\partial n_i}{\partial t} = 0 \end{cases} \quad \text{with } \frac{\partial E}{\partial x} = \frac{q}{\epsilon_0}(n_e - n_i)$$

B: Electron-impact ionization

$$\begin{cases} \dfrac{\partial n_e}{\partial t} = \nu_{\text{iz}}(E)n_e \\ \dfrac{\partial n_i}{\partial t} = \nu_{\text{iz}}(E)n_e \end{cases} \quad \text{with } \frac{\partial E}{\partial t} = 0$$

where the condition $\partial E/\partial t = 0$ used in System B is due to the fact that the ionization

process does not change the local charge density:

$$\frac{\partial}{\partial t}\left[q(n_e - n_i)\right] = 0$$

The Strang splitting procedure consists of evolving System A by half a timestep $\Delta t/2$, then System B by a full time-step Δt, and again System A by another $\Delta t/2$. This can be written in compact notation as

$$\begin{bmatrix} n_e(t+\Delta t) \\ n_i(t+\Delta t) \end{bmatrix} = e^{A\Delta t/2} \cdot e^{B\Delta t} \cdot e^{A\Delta t/2} \cdot \begin{bmatrix} n_e(t) \\ n_i(t) \end{bmatrix}.$$

We will only compute the solution at discrete time instants t_k, with a local separation in time $\Delta t_k := t_{k+1} - t_k$. Accordingly, we will use the notation $n_e^k \approx n_e(t_k)$ and $n_i^k \approx n_i(t_k)$ to indicate the numerical solution at time t_k. Further, the solutions at the intermediate stages of the Strang splitting procedure will be designated as n_e^* and n_e^{**}.

We point out that, even in the ideal case where systems A and B can both be solved exactly, one step of Strang splitting still introduces a local truncation error that is proportional to $(\Delta t_k)^3$. If all time-steps are equal to Δt, the global error in the solution after a fixed time T will be $O(\Delta t^2)$, and for this reason this is referred to as a 'second order' splitting. Overall this splitting procedure is adequate as long as the product $\Delta t\, \nu_{iz} \ll 1$; this is confirmed by the numerical results in Section 3.5.

In our case, System A will be solved approximately by using a Convected Scheme for n_e, while n_i remains unchanged. One complication is that the local advection velocity $u(E)$ is non-constant and non-homogeneous: it depends on the spatial distribution of the electron and ion densities in the domain, and on the instantaneous boundary conditions for Gauss' law.

On the other hand, system B can be solved analytically for both n_e and n_i: taking

a full-step Δt_k yields

$$n_e^{**} = n_e^* + \Delta n,$$

$$n_i^{**} = n_i^* + \Delta n,$$

where the density increment $\Delta n(x)$ due to ionization is

$$\Delta n = n_e^* \left\{ \exp[\nu_{\text{iz}}(E^*)\Delta t_k] - 1 \right\},$$

with the electric field $E^*(x)$ satisfying Gauss' law

$$\frac{\partial E^*}{\partial x} = \frac{q}{\epsilon_0}(n_e^* - n_i^*),$$

and boundary conditions defined at time $t_k + \Delta t_k/2$.

3.2.2 Higher-Order Operator Splitting and Higher Derivatives of Electric Field

A fourth-order accurate operator splitting was developed by Mclachlan [57]. The fourth-order splitting procedure consists of evolving System A and System B using several different time steps. This can be written in compact notation as

$$\begin{bmatrix} n_e(t+\Delta t) \\ n_i(t+\Delta t) \end{bmatrix} = \prod_{j=1}^{6} e^{Ac_j\Delta t} \cdot e^{Bd_j\Delta t} \cdot \begin{bmatrix} n_e(t) \\ n_i(t) \end{bmatrix},$$

$c_1 = c_6 = \frac{14-\sqrt{19}}{108}$, $c_2 = c_5 = \frac{20-7\sqrt{19}}{108}$, $c_3 = c_4 = \frac{5+2\sqrt{19}}{27}$, $d_1 = d_3 = d_5 = \frac{2}{5}$, $d_2 = d_4 = -\frac{1}{5}$, $d_6 = 0$.

In order to apply the MF scheme, we need to compute the first and second partial derivatives of the drift velocity $u(E)$, with respect to time t and to the spatial coordinate x. The first time derivative of the electric field is computed by means of Ampere's law,

which reads

$$\frac{\partial E}{\partial t} = -\frac{q}{\epsilon_0} u(E) n_e \tag{3.13}$$

and we differentiate the above to compute the second derivatives. For the second time derivative we obtain

$$\frac{\partial^2 E}{\partial t^2} = -\frac{q}{\epsilon_0} \left(u'(E) \frac{\partial E}{\partial t} n_e + u(E) \frac{\partial n_e}{\partial t} \right) \tag{3.14}$$

where Ampere's law is used again to compute $\partial E/\partial t$. $\partial n_e/\partial t$ is computed using the transport equation, Eq. (3.1). The second order mixed derivatives such as $\partial^2 E/\partial x \partial t$ are computed by taking a finite difference of $\partial E/\partial t$.

3.3 Convected Scheme and Advection

We will now introduce the CS and use it to solve the advection equation. The term CS refers to a family of algorithms, most usually applied to the solution of the Boltzmann equation. In contrast to most finite difference schemes, the CS is not restricted by the Courant-Friedrichs-Lewy (CFL) criterion, so the CS can use a long time step which makes simulations more efficient and possibly more accurate. Using the CS, the density in a cell at time $t + \Delta t$ is calculated by summing over different cells where particles started at time t. The fractions of the particles originating in a given initial cell which are placed in a final cell are determined by the fractional overlap of the initial cell, once it has been allowed to move within a time step, with the final cell.

The CS is used to solve by envisaging a moving-cell (MC) $C(t)$. At the end of the time step, the particles in the moving cell are remapped back onto the Eulerian mesh. In [48], the moving cell has a uniform density in physical space and

1) one can move the midpoint of the cell (Moving Midpoint, MM): each face has

the same velocity, or

2) one can move the faces of the cell (Moving Face, MF): the two spatial faces of a cell are moved independently.

We employ both of these, starting with the MM scheme because it is simpler and retains many desirable properties: it is efficient and conservative, it preserves positivity and has low phase-error.

However, the accuracy of the simplest version of the scheme is only second order in space, so the leading error will generate a strong numerical diffusion. As a result, a high-order remapping procedure is necessary, and as shown below it applies a small correction to the final position of the moving cell to reduce the undesired numerical diffusion [48].

3.3.1 High Order Remapping

The accuracy of the basic MM scheme is only second order in space, which is not enough for the proposed simulations. In order to increase the accuracy, two strategies could be employed:

1. a non-uniform density in the moving cell

2. a small correction to the final position of the moving cell.

A non-uniform density of the moving cell is more complex to implement than a small correction to the final position. Therefore, we derive the modified equation below which allows us to obtain a small correction to the final position of the moving cell. The derivation based on the Modified Equation summarizes results from [48] for the MM scheme, which we extend here to derive the MF scheme. This modification increases the order of the spatial accuracy from second to fourth. The main point is

when we apply this correction to the position, the operation of the CS is not changed, but it gives more accurate results.

The starting point is the 1D continuity equation:

$$\frac{\partial}{\partial t}n + \frac{\partial}{\partial x}(un) = 0 \tag{3.15}$$

where $n(x,t)$ is the density and $u(x,t)$ is the velocity.

The MM scheme is not limited by the CFL criterion: the normalized displacement of the moving cell coincides with the Courant number

$$C := \left|\frac{u\Delta t}{\Delta x}\right|, \tag{3.16}$$

and is not required to have a magnitude less than one. The displacement of the moving cell can be written as

$$C = m + U, \qquad m \in \mathbb{Z}, \qquad -1 < U < 1, \tag{3.17}$$

where m rounds the Courant number toward zero and U is the remainder or noninteger part of the Courant number. The 1D MM scheme can be rewritten as

$$n_{i+m}^{k+1} = \begin{cases} U_{i-1}^k n_{i-1}^k + (1 - U_i^k) n_i^k & \text{if } U \geq 0, \\ (1 + U_i^k) n_i^k - U_{i+1}^k n_{i+1}^k & \text{otherwise,} \end{cases} \tag{3.18}$$

Here, i denotes space and k means time. For the sake of clarity, in the present derivation we will only consider the case $C > 0$, the case $C < 0$ can be derived similarly, and hence Eq. (4.4) becomes

$$n_i^{k+1} = U_{i-m-1}^k n_{i-m-1}^k + (1 - U_{i-m}^k) n_{i-m}^k. \tag{3.19}$$

Then, we can use the third order Taylor expansion to expand Eq. (3.19) about point (x_i, t^k).

$$\begin{aligned} n_t\Delta t + (nC)_x\Delta x = &-n_{tt}\frac{\Delta t^2}{2} - n_{ttt}\frac{\Delta t^3}{6} + (nU)_{xx}\frac{(2m+1)\Delta x^2}{2} + n_{xx}\frac{m^2\Delta x^2}{2} \\ &- (nU)_{xxx}\frac{(3m^2+3m+1)\Delta x^3}{6} - n_{xxx}\frac{m^3\Delta x^3}{6} \end{aligned} \tag{3.20}$$

Derivatives of U are to be understood as the equivalent derivative of C, since in principle derivatives of U are discontinuous when m changes. Nevertheless we prefer to write expressions in terms of U.

In order to eliminate n_t, we assume

$$n_t\Delta t = -(nC)_x\Delta x + H\Delta t, \tag{3.21}$$

where H is the 1^{st} local truncation error (LTE). The expression for n_{tt} can be obtained by differentiating both sides of Eq. (3.21). During this process, n_t can by substituted for using Eq. (3.21):

$$n_{tt}\Delta t^2 = C(nC)_{xx}\Delta x^2 + C_x(nC)_x\Delta x^2 - (nC_t)_x\Delta x\Delta t - (CH)_x\Delta t\Delta x + H_t\Delta t^2 \tag{3.22}$$

The expression for n_{ttt} can be obtained by a similar process:

$$\begin{aligned} n_{ttt}\Delta t^3 = &-C^2(nC)_{xxx}\Delta x^3 + \left(2C_t\frac{\Delta t}{\Delta x} - 3CC_x\right)(nC)_{xx}\Delta x^3 \\ &+ \left(2C_{xt}\frac{\Delta t}{\Delta x} - C_x^2 - CC_{xx}\right)(nC)_x\Delta x^3 \\ &+ C(nC_t)_{xx}\Delta x^2\Delta t + C_x(nC_t)_x\Delta x^2\Delta t - (nC_{tt})_x\Delta x\Delta t^2 + H.O.T. \end{aligned} \tag{3.23}$$

Substituting Eqs. (3.21), (3.22) and (3.23) into Eq. (3.20), one gets:

$$\begin{aligned} n_t\Delta t + (nC)_x\Delta x = &\frac{\Delta x^2}{2}\left[(1+2m-C)\,(nC)_{xx} - C_x(nC)_x + (nC_t)_x\frac{\Delta t}{\Delta x} - n_{xx}m(m+1)\right] \\ &+\frac{\Delta x^3}{6}\left[\left(C^2-1\right)(nC)_{xxx} + \left(3CC_x - 2C_t\frac{\Delta t}{\Delta x}\right)(nC)_{xx}\right. \\ &+\left(C_x^2 + CC_{xx} - 2C_{xt}\frac{\Delta t}{\Delta x}\right)(nC)_x - C(nC_t)_{xx}\frac{\Delta t}{\Delta x} \\ &\left.- C_x(nC_t)_x\frac{\Delta t}{\Delta x} + (nC_{tt})_x\left(\frac{\Delta t}{\Delta x}\right)^2 - (nC)_{xxx}3m(m+1) + n_{xxx}m(m+1)(2m+1)\right] \\ &+\frac{\Delta t}{2}\left[(CH)_x\Delta x - H_t\Delta t\right] + O(\Delta x^3) \end{aligned} \tag{3.24}$$

where the *r.h.s* is the truncation error.

Comparing Eq. (3.24) with Eq. (3.21), the 1^{st} order part of the LTE, H_1 can be obtained:

$$H_1 = \frac{\Delta x^2}{2}\left[(1-C+2m)(nC)_{xx} - C_x(nC)_x + (nC_t)_x\frac{\Delta t}{\Delta x} - n_{xx}m(m+1)\right], \tag{3.25}$$

The 2^{nd} part of the LTE can be obtained by assuming $H \approx H_1$. The space and time derivatives of the LTE can thus be obtained:

$$\begin{aligned} H_x = &\frac{\Delta x^2}{2}\left[(1-C+2m)(nC)_{xxx} - 2C_x(nC)_{xx} - C_{xx}(nC)_x\right. \\ &\left.+(nC_t)_{xx}\frac{\Delta t}{\Delta x} - n_{xxx}m(m+1)\right] \end{aligned} \tag{3.26}$$

and

$$H_t = \frac{\Delta x^3}{2\Delta t^2}\left[\left(C^2 - C - 2mC + m(m+1)\right)(nC)_{xxx} + \left((3C-2)C_x - 4mC_x - 2C_t\frac{\Delta t}{\Delta x}\right)(nC)_{xx}\right.$$
$$+ \left((C-1)C_{xx} + C_x^2 - 2mC_{xx} - 2C_{xt}\frac{\Delta t}{\Delta x}\right)(nC)_x$$
$$\left.+(1 - C + 2m)(nC_t)_{xx}\frac{\Delta t}{\Delta x} - C_x(nC_t)_x\frac{\Delta t}{\Delta x} + (nC_{tt})_x\left(\frac{\Delta t}{\Delta x}\right)^2\right] \tag{3.27}$$

Substituting Eqs. (3.26) and (3.27) into Eq. (3.24), one gets:

$$n_t\Delta t + (nC)_x\Delta x = \frac{\Delta x^2}{2}\left[(1 + 2m - C)(nC)_x + (nC_t)\frac{\Delta t}{\Delta x} - n_x m(m+1)\right]_x$$
$$+\frac{\Delta x^3}{6}\left[\left(3C - 2C^2 - 1 + 6mC - \frac{9}{2}m(m+1)\right)(nC)_{xx}\right.$$
$$+\left(\frac{3}{2}C_x - 2CC_x + C_t\frac{\Delta t}{\Delta x} + 3mC_x\right)(nC)_x + \left(2C - \frac{3}{2} - 3m\right)(nC_t)_x\frac{\Delta t}{\Delta x}$$
$$\left.+ n_{xx}(2m^3 + 3m^2 + m) - \frac{3}{2}n_{xx}Cm(m+1) - \frac{1}{2}(nC_{tt})\left(\frac{\Delta t}{\Delta x}\right)^2\right]_x$$
$$+O(\Delta x^3) \tag{3.28}$$

In order to get the third order version of the MM scheme, we apply a small correction to the final position.

$$C = m + U_0 + U_1 + U_2 \quad \text{or} \quad u = u_0 + u_1 + u_2 \tag{3.29}$$

$u_0(x,t)$ is the physical velocity field, $U_1 \propto \Delta x$ and $U_2 \propto \Delta x^2$. Substituting Eq. (3.29) into Eq. (3.28) and comparing with the 'physical' continuity equation

$$n_t + (u_0 n)_x = O(\Delta x^3), \tag{3.30}$$

we find the corrections to the final position as

$$U_1 = \frac{\Delta x}{2}\left[(s_c - U_0)\,\frac{1}{n}(U_0 n)_x + U_{0t}\frac{\Delta t}{\Delta x} + mU_{0x}\right];\ \text{where } s_c \text{ is the sign of } C \quad (3.31)$$

and

$$U_2 = \frac{\Delta x^2}{6}\left[\frac{n_{xx}}{n}U_0 d_1 + \frac{n_x}{n}d_2 + U_{0xx}d_3 + U_{0x}^2 d_4 - 2U_{0x}U_{0t}\frac{\Delta t}{\Delta x} + U_{0xt}d_5 + U_{0tt}d_6\right]. \quad (3.32)$$

where

$$d_1 = U_0^2 - s_c\frac{3}{2}U_0 + \frac{1}{2}, \quad (3.33a)$$

$$d_2 = 3U_{0x}\left(U_0^2 - (m + s_c\frac{3}{2})U_0 + s_c\frac{m}{2} + \frac{1}{3}\right) + 3U_{0t}\frac{\Delta t}{\Delta x}\left(s_c\frac{1}{2} - U_0\right), \quad (3.33b)$$

$$d_3 = U_0^2 - \left(m + s_c\frac{3}{2}\right)U_0 + m^2 + s_c\frac{3m}{2} + \frac{1}{2}, \quad (3.33c)$$

$$d_4 = U_0 - 2m - s_c\frac{3}{2}, \quad (3.33d)$$

$$d_5 = \frac{\Delta t}{\Delta x}\left(2m + s_c\frac{3}{2} - U_0\right), \quad (3.33e)$$

$$d_6 = \frac{\Delta t^2}{\Delta x^2}, \quad (3.33f)$$

With U_1 and U_2, we can get fourth order accuracy in space without changing the remapping rule of the CS.

3.3.2 Moving Face (MF)

As we mentioned above, there are two methods to move the cells: 1) one can move the midpoint of the cell (Moving Midpoint, MM): each face has the same velocity, or 2) one can move the faces of the cell (Moving Face, MF): the two spatial faces of a cell are moved independently. The high order MM scheme was introduced above. We now

turn to the MF scheme. The MF scheme allows the faces of the initial cell to move to "final" positions at the end of a time step, as shown Fig. 3.2. The density in the initial cell is returned to the mesh in proportion to the area overlap, shown by shaded areas. Hence the fraction of the density going to each final cell is equal to the fraction of the area of the moving cell which overlaps that fixed mesh cell.

The displacement of the front face includes a ballistic part C_{bal}. C_{bal} can be found by taking the average C over the time step. First, we can use a Taylor expansion to expand $C(\Delta t)$.

$$C(t) = C(0) + C_t t + C_{tt}\frac{t^2}{2} + O(t^3) \tag{3.34}$$

then, taking the time average of C between 0 and Δt, we get

$$\overline{C} := \frac{1}{\Delta t}\int_0^{\Delta t} C dt = C + \frac{dC}{dt}\frac{\Delta t}{2} + \frac{d^2C}{dt^2}\frac{\Delta t^2}{6} \tag{3.35}$$

where

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \frac{\partial C}{\partial x}\frac{dx}{dt} = C_t + CC_x\frac{\Delta x}{\Delta t}.$$

Therefore,

$$\begin{aligned}\overline{C} = C &+ C_t\frac{\Delta t}{2} + CC_x\frac{\Delta x}{2} \\ &+ C_{tt}\frac{\Delta t^2}{6} + C_xC_t\frac{\Delta x\Delta t}{6} + 2C_{xt}C\frac{\Delta x\Delta t}{6} + C^2C_{xx}\frac{\Delta x^2}{6} + CC_x^2\frac{\Delta x^2}{6},\end{aligned} \tag{3.36}$$

and to third order

$$\begin{aligned}C_{bal} = C_0 &+ \frac{\Delta x}{2}\left(C_{0t}\frac{\Delta t}{\Delta x} + C_0C_{0x}\right) \\ &+ \frac{\Delta x^2}{6}\left(C_{0tt}\frac{\Delta t^2}{\Delta x^2} + C_{0x}C_{0t}\frac{\Delta t}{\Delta x} + 2C_0C_{0xt}\frac{\Delta t}{\Delta x} + C_0^2C_{0xx} + C_0C_{0x}^2\right).\end{aligned} \tag{3.37}$$

where $C_0 = C(x_i, t_k)$ is the local Courant number, and the small terms allow for the

velocity changing during the step. Here, C is not required to have a magnitude less than one: as usual, we can decompose C into integer and fractional parts as

$$C = (m + U), \quad m \in \mathbb{Z}, \quad -1 < U < 1. \tag{3.38}$$

where m rounds the Courant number toward zero and U is the remainder of the Courant number.

Finally, by means of a modified equation analysis based on the overlaps of the cell faces with the fixed cells (see below), we derive a second order "numerical" correction C_{nul} which resembles antidiffusive corrections found elsewhere. C_{nul} is added to C_{bal}.

$$C_{nul} = \frac{\Delta x}{2}\left[\frac{n_x}{n}U_0(s_c - U_0)\right] + \frac{\Delta x^2}{6}\left[\frac{n_{xx}}{n}U_0 d_1 + \frac{n_x}{n}d_2 + \frac{n_x^2}{n^2}U_0 d_3 + U_{0xx}d_4\right]. \tag{3.39}$$

where

$$d_1 = -2U_0^2 + s_c 3U_0 - 1, \tag{3.40a}$$

$$d_2 = U_{0x}\left(3U_0(-2U_0 + s_c\frac{3}{2} - m) + s_c\frac{3m}{2} - \frac{1}{2}\right) + 3U_{0t}\frac{\Delta t}{\Delta x}(s_c\frac{1}{2} - U_0)), \tag{3.40b}$$

$$d_3 = 3(\frac{1}{2} - s_c\frac{3}{2}U_0 + U_0^2), \tag{3.40c}$$

$$d_4 = 3(U_0^2 - \frac{1}{12} - s_c U_0). \tag{3.40d}$$

The equations above should be rewritten in terms of C_{bal}, after this is split into

$$C_{\mathrm{bal}} = m_{\mathrm{bal}} + U_{\mathrm{bal}}, \quad m \in \mathbb{Z}, \quad -1 < U < 1.$$

To first order we have

$$U_0 \approx U_{\mathrm{bal}} - C_{0t}\frac{\Delta t}{2} - C_0 C_{0x}\frac{\Delta x}{2}$$

and by substituting back into the equations (Eq. (3.39) and Eq. (3.40)) above we obtain

$$C_{1n} = \frac{\Delta x}{2}\frac{n_x}{n}\eta_0 \tag{3.41a}$$

$$C_{2n} = \frac{\Delta x^2}{6}\left[\left(\frac{n_{xx}}{n} - \frac{3}{2}\frac{n_x^2}{n^2}\right)\eta_1 + \frac{n_x}{n}C_{0x}\eta_2 + C_{0xx}\eta_3\right] \tag{3.41b}$$

where

$$\eta_0 = U_{\text{bal}}(s_c - U_{\text{bal}}) \tag{3.42a}$$

$$\eta_1 = U_{\text{bal}}\left(3\eta_0 + U_{\text{bal}}^2 - 1\right), \tag{3.42b}$$

$$\eta_2 = -1/2 + 3\eta_0, \tag{3.42c}$$

$$\eta_3 = -1/4 - 3\eta_0. \tag{3.42d}$$

Our procedure is the following:

1. Compute the ballistic displacement $\{C_{\text{bal}}\}_i$ at the cell centers using (3.37);
2. Using cubic interpolation, evaluate $\{C_{\text{bal}}\}_{i-\frac{1}{2}}$ at the cell faces; then decompose it as $\{C_{\text{bal}}\}_{i-\frac{1}{2}} = \{m_{\text{bal}}\}_{i-\frac{1}{2}} + \{U_{\text{bal}}\}_{i-\frac{1}{2}}$;
3. Using second order finite-differences, compute $\{n_x/n\}_{i-\frac{1}{2}}$, $\{n_{xx}/n\}_{i-\frac{1}{2}}$, $\{C_{0x}\}_{i-\frac{1}{2}}$ and $\{C_{0xx}\}_{i-\frac{1}{2}}$ also at the cell faces;
4. Compute the anti-diffusive corrections $\{C_{1n}\}_{i-\frac{1}{2}}$ and $\{C_{2n}\}_{i-\frac{1}{2}}$ using (3.41) and (3.42), and add those to the ballistic displacement;
5. Remap density from the moving cells onto the fixed mesh using the classical CS remapping rule.

The finite difference approximations that we use are

$$[C_{0x}]_i = \frac{[C_0]_{i+1} - [C_0]_{i-1}}{2\Delta x} + O(\Delta x^2) \tag{3.43a}$$

$$[C_{0xx}]_i = \frac{[C_0]_{i-1} - 2[C_0]_i + [C_0]_{i+1}}{\Delta x^2} + O(\Delta x^2) \tag{3.43b}$$

$$[C_{0tx}]_i = \frac{[C_{0t}]_{i+1} - [C_{0t}]_{i-1}}{2\Delta x} + O(\Delta x^2) \tag{3.43c}$$

$$[C_{\mathrm{bal}}]_{i-\frac{1}{2}} = \frac{-[C_{\mathrm{bal}}]_{i-2} + 9[C_{\mathrm{bal}}]_{i-1} + 9[C_{\mathrm{bal}}]_i - [C_{\mathrm{bal}}]_{i+1}}{16} + O(\Delta x^4) \tag{3.43d}$$

$$\left[\frac{n_x}{n}\right]_{i-\frac{1}{2}} = \frac{2}{\Delta x}\frac{n_i - n_{i-1}}{n_{i-1} + n_i + \varepsilon} + O(\Delta x^2) \tag{3.43e}$$

$$\left[\frac{n_{xx}}{n}\right]_{i-\frac{1}{2}} = \frac{2}{\Delta x^2}\frac{n_{i-2} - n_{i-1} - n_i + n_{i+1}}{n_{i-2} + n_{i-1} + n_i + n_{i+1} + \varepsilon} + O(\Delta x^2) \tag{3.43f}$$

$$[C_{0x}]_{i-\frac{1}{2}} = \frac{[C_0]_i - [C_0]_{i-1}}{\Delta x} + O(\Delta x^2) \tag{3.43g}$$

$$[C_{0xx}]_{i-\frac{1}{2}} = \frac{[C_0]_{i-2} - [C_0]_{i-1} - [C_0]_i + [C_0]_{i+1}}{2\Delta x^2} + O(\Delta x^2) \tag{3.43h}$$

where ε is a small positive number that avoids division by zero.

Therefore the normalized displacement averaged over a time step is $C = C_{bal} + C_{1n} + C_{2n}$ and in terms of the displacement defined in this way, the normalized displacement of the MF can be expressed as:

$$C_{MF} = C_{i\pm 1/2} = \{C_{bal}\}_{i\pm 1/2} + \{C_{1n}\}_{i\pm 1/2} + \{C_{2n}\}_{i\pm 1/2} \tag{3.44}$$

where the plus or minus are for the front or back face respectively, denoted C_f and C_b.

Then to find the fraction of the moving cell in the fixed cells f (in front) and b (behind), we first find the length of the moving cell, $1 + C_f - C_b$. The overlap of the moving cell with the fixed cells f and b is $C_f - m_f$ for the front cell where m_f is the index of cell f and the overlap is $(m_b + 1) - C_b$ for the back cell where m_b is the index of the back cell, respectively. The fractions in each cell are these overlaps divided by the length of the moving cell. (This is easily extended to cases where f and b are

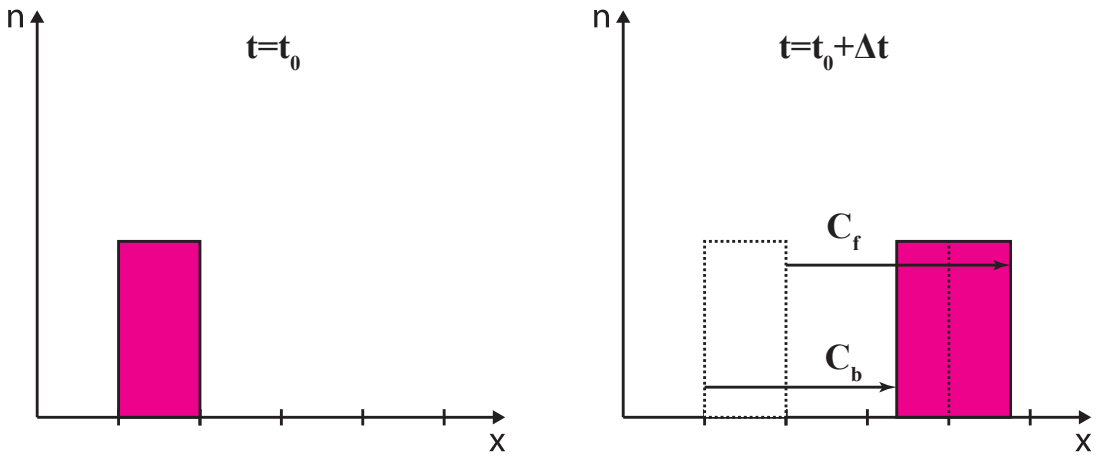

Figure 3.2: Ballistic operator of the moving face (MF) scheme. The right frame shows the density reallocated to cells, according to the MF scheme.

not adjacent.) A modified equation analysis, or direct comparison with the results of the MF scheme, confirms that this procedure, with the antidiffusion term C_{nul}, yields results correct to second order. The final expressions for the fractions are:

$$f_b = \frac{(m_b + 1) - C_b}{1 + C_f - C_b} \tag{3.45}$$

and

$$f_f = \frac{C_f - m_f}{1 + C_f - C_b} \tag{3.46}$$

These fractions sum to precisely one.

Provided that C varies smoothly, we encounter three situations with respect to m_f and m_b. Most frequently, $m_f = m_b$. We also find $m_f > m_b$ (but only $m_f = m_b + 1$ in practice) and $m_f = m_b - 1$. In the first two cases we find the overlaps using the formula above. In the last case, both faces are in the same final cell, and the entire density is put in that cell. If C is negative, a mirror-image set of rules applies. We describe the implementation details in Algorithm 1.

Algorithm 1 High order Convected Scheme with MF

1: Given the Courant parameter $C_i^k = u_i^k \Delta t/\Delta x$, compute the ballistic displacement $\{C_{bal}\}_i$ at the cell centers;

2: Using the cubic interpolation, evalute $\{C_{bal}\}_{i\pm1/2}$ at the cell faces, and decompose it as $\{C_{bal}\}_{i\pm1/2} = \{m_{bal}\}_{i\pm1/2} + \{U_{bal}\}_{i\pm1/2}$:

3: where $\{m_{bal}\}_{i\pm1/2} \in \mathbb{Z}, \quad -1 < \{U_{bal}\}_{i\pm1/2} < 1$;

4: Compute the anti-diffusive corrections $\{C_{1n}\}_{i\pm1/2}$ and $\{C_{2n}\}_{i\pm1/2}$ using (3.41) and (3.42), and add those to the ballistic displacement;

$$\begin{cases} C_{fi}^k = \{C_{bal}\}_{i+1/2} + \{C_{1n}\}_{i+1/2} + \{C_{2n}\}_{i+1/2} = (m_{fi}^k + U_{fi}^k) \\ C_{bi}^k = \{C_{bal}\}_{i+1/2} + \{C_{1n}\}_{i+1/2} + \{C_{2n}\}_{i+1/2} = (m_{bi}^k + U_{bi}^k) \end{cases},$$

where $m_{fi}^k,\ m_{bi}^k \in \mathbb{Z}$ and $-1 < U_{fi}^k,\ U_{bi}^k < 1$;

5: Depending on m_{fi}^k and m_{bi}^k, there are three situations 1) $m_{fi}^k = m_{bi}^k$, 2) $m_{fi}^k > m_{bi}^k$, 3) $m_{fi}^k = m_{bi}^k - 1$.

- $m_{fi}^k = m_{bi}^k$, using (3.45) and (3.46) to find the fractions, density remap according to the fractions. Cell (i) remaps only into cells $(i + m_{bi})$ and $(i + m_{fi} + 1)$:

$$\begin{cases} \Delta n_{i+m_{fi}+1}^{k+1} = n_i^k f_{fi} \\ \Delta n_{i+m_{bi}}^{k+1} = n_i^k f_{bi} \end{cases};$$

  Δn_i^k is the contribution to the density in cell i at time k, due to a single initial cell (the contribution from a cell to itself will be zero, if C_{fi} and C_{bi} are outside the cell).

- $m_{fi} > m_{bi}$, using (3.45) and (3.46) to find the fractions, density remap according to the fractions. Cell (i) remaps into cells from $(i + m_{bi})$ to $(i + m_{fi} + 1)$:

$$\begin{cases} \Delta n_{i+m_{fi}+1}^{k+1} = n_i^k f_{fi} \\ \Delta n_{i+m_{bi}+1:i+m_{fi}}^{k+1} = n_i^k \dfrac{1}{1 + C_{fi} - C_{bi}} \\ \Delta n_{i+m_{bi}}^{k+1} = n_i^k f_{bi} \end{cases};$$

- $m_{fi} = m_{bi} - 1$, both faces are in the same cell, and the entire density is put in that cell:

  $\Delta n_{i+m_{fi}+1}^{k+1} = n_i^k$.

3.4 Advection Tests

The purpose of this section is to assess the 'global error' of the scheme, i.e. the difference between the exact solution and the numerical solution at a certain final time T, and show that it converges as expected with respect to Δx and Δt.

Tables 3.1-3.5 show the convergence of different versions of the scheme (MM and MF) for different orders of accuracy of the CS. Runs are performed at fixed C_0 and at fixed Δt. In all cases the convergence is as expected. Tests for a flow field which varies with time have been done, but are not reported here, which also perform as expected.

3.4.1 Stationary Wave

In our first numerical test, we solve the 1D continuity equation in the presence of a spatially varying, but stationary, flow field $u(x)$. The 'stationary compression wave' problem [58] is defined by

$$\begin{aligned} &\frac{\partial n}{\partial t} + \frac{\partial (un)}{\partial x} = 0, \qquad && (x,t) \in [0, 2\pi] \times [0, T], \\ &u(x) = c + \sin(x), && n(x,0) = 1, \end{aligned} \tag{3.47}$$

where $T > 0$ is the final time, and $c \in \mathbb{R}$ is a constant satisfying $|c| > 1$. Since $u(x)$ and $n(x,0)$ are periodic with period 2π, so will be the solution $n(x,t)$. Further, because $u(x)$ does not change sign in the domain, the solution will remain strictly positive for all times, i.e. $n(x,t) > 0 \ \forall (x,t)$.

An analytical solution to Eq. (3.47) can be obtained through the method of characteristics, after a suitable change of variable [58]. The analytic solution for the original

quantity $n(x,t)$ can be written as

$$n(x,t) \;=\; n\big[x_0(x,t),0\big]\cdot\frac{c^2+\cos\big[bt+a(x,t)\big]-b\sin\big[bt+a(x,t)\big]}{c^2+\cos\big[a(x,t)\big]-b\sin\big[a(x,t)\big]}, \tag{3.48}$$

where

$$b := \sqrt{c^2-1}, \tag{3.49a}$$

$$a(x,t) := 2\arctan\left\{\frac{1+c\tan\big[x_0(x,t)/2\big]}{b}\right\}, \tag{3.49b}$$

$$x_0(x,t) := 2\arctan\left\{\frac{b}{c}\tan\left[\arctan\left(\frac{1+c\tan(x/2)}{b}\right)-\frac{bt}{2}\right]-\frac{1}{c}\right\}. \tag{3.49c}$$

For this test case we use $c=2$, which yields $b=\sqrt{3}$, and a final time $T=\pi/b$. The final solution obtained with the first and third order MM scheme is compared with the exact solution in Fig. 3.3 for different numbers N of subdivisions in space ($\Delta x = 2\pi/N$).

Now, we examine the convergence of the simulations with respect to the mesh size Δx and time step size Δt. For this purpose, we progressively reduce both quantities (refinement study) and calculate the relative L^2-norm of the error at the final time T, defined as

$$L^2_{\text{err}} := \sqrt{\frac{\sum_i \big[n_i - n(x_i)\big]^2}{\sum_i \big[n(x_i)\big]^2}}, \tag{3.50}$$

where $n(x_i)$ is the exact solution evaluated at the grid point $x=x_i$, and $n_i \approx n(x_i)$ is our numerical solution. Within each refinement study we keep a constant 'grid velocity' $u_G := \Delta x/\Delta t$; accordingly, all simulations therein employ the same Courant parameter $C_0(x) := u(x)\Delta t/\Delta x = u(x)/u_G$.

Assuming an algebraic order of convergence p, meaning that the error norm decreases with increasing number of cells N as $L^2_{\text{err}}(N) \propto N^{-p}$, we estimate p from two

successive error norms as

$$p(N_1, N_2) := \frac{\ln\left[L^2_{\text{err}}(N_1)/L^2_{\text{err}}(N_2)\right]}{\ln\left(N_1/N_2\right)}. \tag{3.51}$$

In our refinement analysis we progressively double the number of cells, and the formula above reduces to $p(N/2, N) = \log_2\left[L^2_{\text{err}}(N/2)/L^2_{\text{err}}(N)\right]$.

Table 3.1 reports the convergence results for the moving-midpoint (MM) version of the CS, for a grid velocity $u_G = \Delta x/\Delta t = 10$; this corresponds to a Courant number well below 1 throughout the domain, $C_0(x) = 0.2 + 0.1\sin(x)$. All three variations MM2/3/4 exhibit the expected order of convergence, i.e. the order of the truncation error minus one.

	MM2		MM3		MM4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
25	1.42×10^{-1}	–	2.55×10^{-2}	–	2.40×10^{-2}	–
50	8.73×10^{-2}	0.70	6.10×10^{-3}	2.06	4.33×10^{-3}	2.53
100	5.01×10^{-2}	0.80	1.43×10^{-3}	2.09	5.98×10^{-4}	2.86
200	2.72×10^{-2}	0.88	3.51×10^{-4}	2.02	7.64×10^{-5}	2.97
400	1.43×10^{-2}	0.93	8.79×10^{-5}	2.00	9.59×10^{-6}	2.99
800	7.33×10^{-3}	0.96	2.21×10^{-5}	2.00	1.20×10^{-6}	3.00
1600	3.72×10^{-3}	0.98	5.53×10^{-6}	2.00	1.50×10^{-7}	3.00
3200	1.87×10^{-3}	0.99	1.39×10^{-6}	2.00	1.84×10^{-8}	3.00

Table 3.1: Continuity equation, stationary wave: refinement analysis for the moving-midpoint (MM) version of the Convected Scheme (CS). We compare the 'standard' low-order implementation MM2 with its high-order variants MM3 and MM4. The table reports the L^2-norm of the error (difference of analytic and numerical solutions) at the final time, for progressively larger numbers of cells N. The algebraic order of convergence ('Order' column) is calculated as the base-2 logarithm of two successive error norms. All simulations employ the same Courant parameter $C_0(x) = 0.2 + 0.1\sin(x)$, which corresponds to a constant grid velocity $u_G = 10$.

Table 3.2 shows a refinement study identical to Table 3.1, but for the new moving-face (MF) version of the CS. Again, the order of convergence is as expected. We notice that the MF2 and MF3 schemes are performing slightly worse than their MM2(3)

counterparts, while the higher-order version MF4 performs consistently better than MM4.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
25	1.82×10^{-1}	—	2.89×10^{-2}	—	1.13×10^{-2}	—
50	1.04×10^{-1}	0.81	7.35×10^{-3}	1.97	1.42×10^{-3}	3.00
100	5.72×10^{-2}	0.86	1.81×10^{-3}	2.02	1.70×10^{-4}	3.06
200	3.04×10^{-2}	0.91	4.48×10^{-4}	2.01	2.11×10^{-5}	3.01
400	1.57×10^{-2}	0.95	1.12×10^{-4}	2.01	2.65×10^{-6}	2.99
800	8.01×10^{-3}	0.97	2.78×10^{-5}	2.00	3.33×10^{-7}	2.99
1600	4.05×10^{-3}	0.99	6.96×10^{-6}	2.00	4.17×10^{-8}	3.00
3200	2.03×10^{-3}	0.99	1.74×10^{-6}	2.00	5.22×10^{-9}	3.00

Table 3.2: Continuity equation, stationary wave: refinement analysis for the moving-face (MF) version of the Convected Scheme (CS). We compare the 'standard' low-order implementation MF2 with its high-order variants MF3 and MF4. As in Table 3.1, we report the number of cells N, the L^2-norm of the error, and the order of convergence. All simulations employ the same Courant parameter $C_0(x) = 0.2 + 0.1\sin(x)$, which corresponds to a constant grid velocity $u_G = 10$.

When the Courant parameter exceeds the value of 1 in the domain, the performance of the high-order MM-CS deteriorates sensibly, and the order of convergence cannot reach 2. This is shown in Table 3.3, where a constant grid velocity $u_G = 2.5$ is employed, corresponding to a Courant parameter $C_0(x) = 0.8 + 0.4\sin(x)$. In the same conditions the new MF-CS significantly outperforms the MM version, and reaches the expected order of convergence (see Table 3.4).

It is interesting to assess the behavior of the new MF-CS at higher values of the Courant number. For this purpose, we repeat our refinement study for a smaller grid velocity $u_G = 0.625$, which corresponds to $C_0(x) = 3.2 + 1.6\sin(x)$. The results are reported in Table 3.5 and they are discussed below.

In this section we examined convergence of the CS at fixed Courant number $C_0(x)$, which requires variation of Δx and Δt. In all cases the convergence of our best scheme, the MF4-CS, has an order of three over the range examined, as expected.

	MM2		MM3		MM4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
25	9.28×10^{-2}	–	1.84×10^{-2}	–	2.20×10^{-2}	–
50	5.25×10^{-2}	0.82	2.88×10^{-3}	2.67	3.24×10^{-3}	2.76
100	2.81×10^{-2}	0.90	7.40×10^{-4}	1.96	8.51×10^{-4}	1.93
200	1.46×10^{-2}	0.95	2.55×10^{-4}	1.54	2.94×10^{-4}	1.53
400	7.43×10^{-3}	0.97	9.45×10^{-5}	1.43	9.69×10^{-5}	1.60
800	3.76×10^{-3}	0.99	3.13×10^{-5}	1.60	3.24×10^{-5}	1.58
1600	1.89×10^{-3}	0.99	1.12×10^{-5}	1.48	1.13×10^{-5}	1.52
3200	9.48×10^{-4}	0.99	3.71×10^{-6}	1.60	3.75×10^{-6}	1.59

Table 3.3: Same as Table 3.1, but for a larger Courant parameter $C_0(x) = 0.8 + 0.4\sin(x)$, corresponding to a constant grid velocity $u_G = 2.5$. We notice that the MM3 and MM4 schemes cannot reach their nominal order of convergence (2 and 3 respectively) in this situation where $\max|C_0(x)| > 1$.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
25	1.11×10^{-1}	—	1.73×10^{-2}	—	9.58×10^{-3}	—
50	5.77×10^{-2}	0.94	3.62×10^{-3}	2.26	1.47×10^{-3}	2.70
100	2.99×10^{-2}	0.95	8.18×10^{-4}	2.14	1.94×10^{-4}	2.92
200	1.53×10^{-2}	0.97	1.96×10^{-4}	2.06	2.45×10^{-5}	2.98
400	7.73×10^{-3}	0.98	4.83×10^{-5}	2.02	3.07×10^{-6}	2.99
800	3.89×10^{-3}	0.99	1.20×10^{-5}	2.01	3.84×10^{-7}	3.00
1600	1.95×10^{-3}	1.00	3.00×10^{-6}	2.00	4.81×10^{-8}	3.00
3200	9.78×10^{-4}	1.00	7.48×10^{-7}	2.00	6.01×10^{-9}	3.00

Table 3.4: Same as Table 3.2, but for a larger Courant parameter $C_0(x) = 0.8 + 0.4\sin(x)$, corresponding to a constant grid velocity $u_G = 2.5$. Contrarily to the MM results in Table 3.3, all MF schemes exhibit the expected orders of convergence.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
25	3.07×10^{-1}	—	7.88×10^{-2}	—	3.11×10^{-2}	—
50	1.52×10^{-1}	1.01	1.75×10^{-2}	2.17	3.34×10^{-3}	3.22
100	7.58×10^{-2}	1.00	4.19×10^{-3}	2.07	3.88×10^{-4}	3.10
200	3.78×10^{-2}	1.00	1.03×10^{-3}	2.02	4.73×10^{-5}	3.04
400	1.89×10^{-2}	1.00	2.55×10^{-4}	2.01	5.84×10^{-6}	3.02
800	9.44×10^{-3}	1.00	6.35×10^{-5}	2.01	7.26×10^{-7}	3.01
1600	4.72×10^{-3}	1.00	1.58×10^{-5}	2.00	9.05×10^{-8}	3.00
3200	2.36×10^{-3}	1.00	3.96×10^{-6}	2.00	1.13×10^{-8}	3.00

Table 3.5: Same as Tables 3.2 and 3.4, but for a larger Courant parameter $C_0(x) = 3.2 + 1.6\sin(x)$, corresponding to a constant grid velocity $u_G = 0.625$. Again, all MF schemes exhibit the expected orders of convergence.

3.4.2 Pulsating solution

As a second numerical test, we now solve the 1D continuity equation in the presence of a time varying and inhomogeneous flow field $u(x,t)$. The 'pulsating solution' problem [58] is defined by

$$\begin{aligned} &\frac{\partial n}{\partial t} + \frac{\partial (un)}{\partial x} = 0, && (x,t) \in [0, 2\pi] \times [0, T], \\ &u(x,t) = B\cos(\omega t)\sin(x), && n(x,0) = 1, \end{aligned} \tag{3.52}$$

where B is the amplitude and ω is the angular frequency of oscillating velocity field. Equation (3.52) has analytical solution

$$n(x,t) = n_0\big[x_0(x,t), 0\big] \cdot \sigma(t) \frac{1 + \gamma(x,t)/\sigma^2(t)}{1 + \gamma(x,t)}, \tag{3.53}$$

where

$$\sigma(t) := \exp\big[-B\sin(\omega t)/\omega\big], \tag{3.54a}$$

$$\gamma(x,t) := \tan^2\big[x_0(x,t)/2\big], \tag{3.54b}$$

$$x_0(x,t) := 2\arctan\big[\tan(x/2)\sigma(t)\big]. \tag{3.54c}$$

In our numerical tests we choose $B = 1$ and $\omega = 1$ and we integrate until the final time $T = 3\pi$, corresponding to 1.5 cycles of the periodic solution. Table 3.6 reports convergence results for the MF2/3/4 schemes, obtained using a constant grid velocity $u_G := \Delta x/\Delta t = 0.4$; this corresponds to a Courant number $C_0(x,t)$ ranging in the interval $[-2.5, 2.5]$. We observe the expected order of convergence for all schemes.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
25	6.20×10^{-1}	—	6.63×10^{-2}	—	3.29×10^{-2}	—
50	3.87×10^{-1}	0.68	1.88×10^{-2}	1.81	4.33×10^{-3}	2.93
100	2.10×10^{-1}	0.88	4.75×10^{-3}	1.99	5.64×10^{-4}	2.94
200	1.08×10^{-1}	0.96	1.18×10^{-3}	2.01	7.11×10^{-5}	2.99
400	5.44×10^{-2}	0.99	2.94×10^{-4}	2.01	8.91×10^{-6}	3.00
800	2.73×10^{-2}	1.00	7.34×10^{-5}	2.00	1.12×10^{-6}	3.00
1600	1.37×10^{-2}	1.00	1.83×10^{-5}	2.00	1.39×10^{-7}	3.00
3200	6.84×10^{-3}	1.00	4.58×10^{-6}	2.00	1.74×10^{-8}	3.00

Table 3.6: 1D continuity equation, pulsating solution. Convergence analysis at a constant grid velocity $u_G := \Delta x/\Delta t = 0.4$.

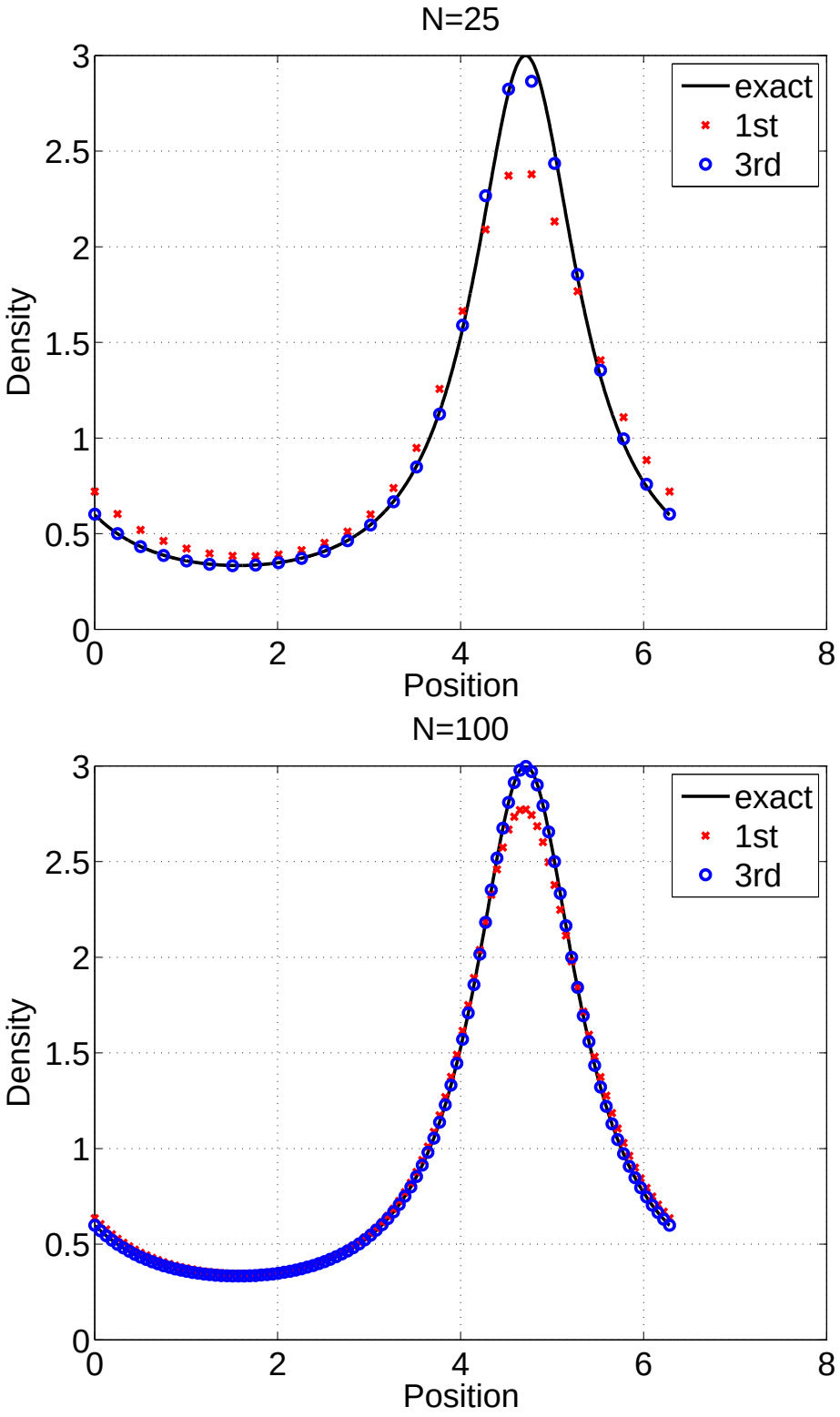

Figure 3.3: (MM version) The solution for different order versions of the CS for increasing number N of spatial subdivisions.

3.5 Numerical Experiments for 1D Breakdown

In this section we present results for the 1D gas breakdown problem, which is the subject of this paper. We solve the breakdown equations discussed in Section 3.1, using the same parameters as in the analytic theory of Section 3.1.2. According to the operator splitting strategy described in Section 3.2, we employ the Convected Scheme (CS) to describe the plasma flow, combined with the ionization process by means of a second or fourth order splitting.

All breakdown simulations make use of the moving-face (MF) version of the CS, which was described in Section 3.3. The MF scheme is seen to allow accurate simulations of breakdown with only moderately large numbers of spatial cells and of time steps. A uniform mesh is employed in all cases, and so the scheme appears to hold promise for a relatively simple approach to breakdown simulation in 2D and 3D.

In Section 3.5.1 we concisely report the common setup of our simulations, which includes the model equations, the domain size and time length, the free parameters, and the initial conditions employed. In Section 3.5.2 we show detailed convergence studies, carried out both at fixed C_0 and at fixed Δt, for the MF2/3/4 versions of the CS, coupled to 2nd- or 4th-order splitting.

3.5.1 Breakdown Test-Case: Problem Setup

The model equations were discussed in Section 3.1, and we report them here for completeness. The electron density $n_e(x,t)$ and ion density $n_i(x,t)$ [m^{-3}] satisfy the equa-

tions

$$\begin{aligned} \frac{\partial n_e}{\partial t} + \frac{\partial (u n_e)}{\partial x} &= \nu_{\text{iz}}\, n_e, \qquad && (x,t) \in [0,L] \times [0,T], \\ \frac{\partial n_i}{\partial t} &= \nu_{\text{iz}}\, n_e, && n_e(x,0) = n_i(x,0) = n_0(x), \end{aligned} \tag{3.55}$$

which are non-linearly coupled through the drift velocity $u(E)$ [m/s] and the ionization rate $\nu_{\text{iz}}(E)$ [s^{-1}]. The local electric field $E(x,t)$ [V/m] satisfies a 1D Gauss law with $E(0,t) = E_{\max}$ as boundary condition, which yields

$$E(x,t) = E_{\max} - \frac{q}{\varepsilon_0} \int_0^x \Big[n_i(x',t) - n_e(x',t) \Big] dx'. \tag{3.56}$$

In order to compare with the analytic solution derived in Section 3.1.2, we employ the same parameters used there:

$$\begin{aligned} u(E) &= \mu E_{\max} \left(\frac{E}{E_{\max}} \right), \\ \nu_{\text{iz}}(E) &= \alpha_{\max} \frac{q \mu E_{\max}^2}{e_{\text{iz}}} \left(\frac{|E|}{E_{\max}} \right)^3. \end{aligned} \tag{3.57}$$

The constants that appear in (3.55), (3.56) and (3.57) are defined in Table 3.7 below.

Symbol	Value	SI unit	Description
L	2.4×10^{-3}	m	Domain size
T	1.6×10^{-8}	s	Simulation length
$E_{\max}$	4.999×10^{6}	V/m	Maximum electric field
q	$-1.602176565 \times 10^{-19}$	C	Electron charge
ε_0	$8.854187817 \times 10^{-12}$	F/m	Vacuum permittivity
μ	-0.01	$\text{m}^2/(\text{Vs})$	Electron mobility
$\alpha_{\max}$	0.1	–	Maximum ionization efficiency
e_{iz}	$3.204353130 \times 10^{-18}$	J	Ionization energy (= 20 eV)

Table 3.7: Constants employed in all breakdown simulations, as they appear in (3.55), (3.56) and (3.57).

The initial conditions for the electron and ion densities in (3.55) are given by a

narrow Gaussian profile with low amplitude:

$$n_0(x) = A\left[\delta + e^{-\left(\frac{x-c}{\sigma}\right)^2}\right],$$
$$A = 6.875 \times 10^{16}\,\mathrm{m}^{-3}, \quad \delta = 1 \times 10^{-16}, \quad c = 2 \times 10^{-4}\,\mathrm{m}, \quad \sigma = 0.5 \times 10^{-4}\,\mathrm{m}, \tag{3.58}$$

where $A(1+\delta)$ is the peak density, $A\delta$ is the minimum density, c is the x position of the peak and σ is the Gaussian width.

3.5.2 Breakdown Test-Case: Refinement Analysis

The aim of this section is to assess the accuracy and efficiency of the various numerical schemes discussed in this paper, when applied to the 1D breakdown problem described in Section 3.5.1.

The evolution of electron density, ion density, electric field and charge density are shown in Fig. 3.4. From Fig. 3.4, we can know the numerical solution is going to approach the analytic solution when time is infinite.

When the numerical error is calculated over the whole domain, two regions where the electron density has very large slopes dominate the errors. These are at the extreme front and extreme rear of the plasma. The front of the plasma is the focus of this paper and of previous papers, and while the behavior of the rear may be of some interest, from a physical point of view it is not clear why one would wish to study it. Further, it would require a separate analytic development to understand the processes taking place there which has not been done. We suspect the behavior at the rear will be dominated by the initial conditions, which again makes the behavior of less interest. We thus exclude the far rear of the plasma from our error calculations.

For the convergence test we use successively increased numbers N of spatial subdi-

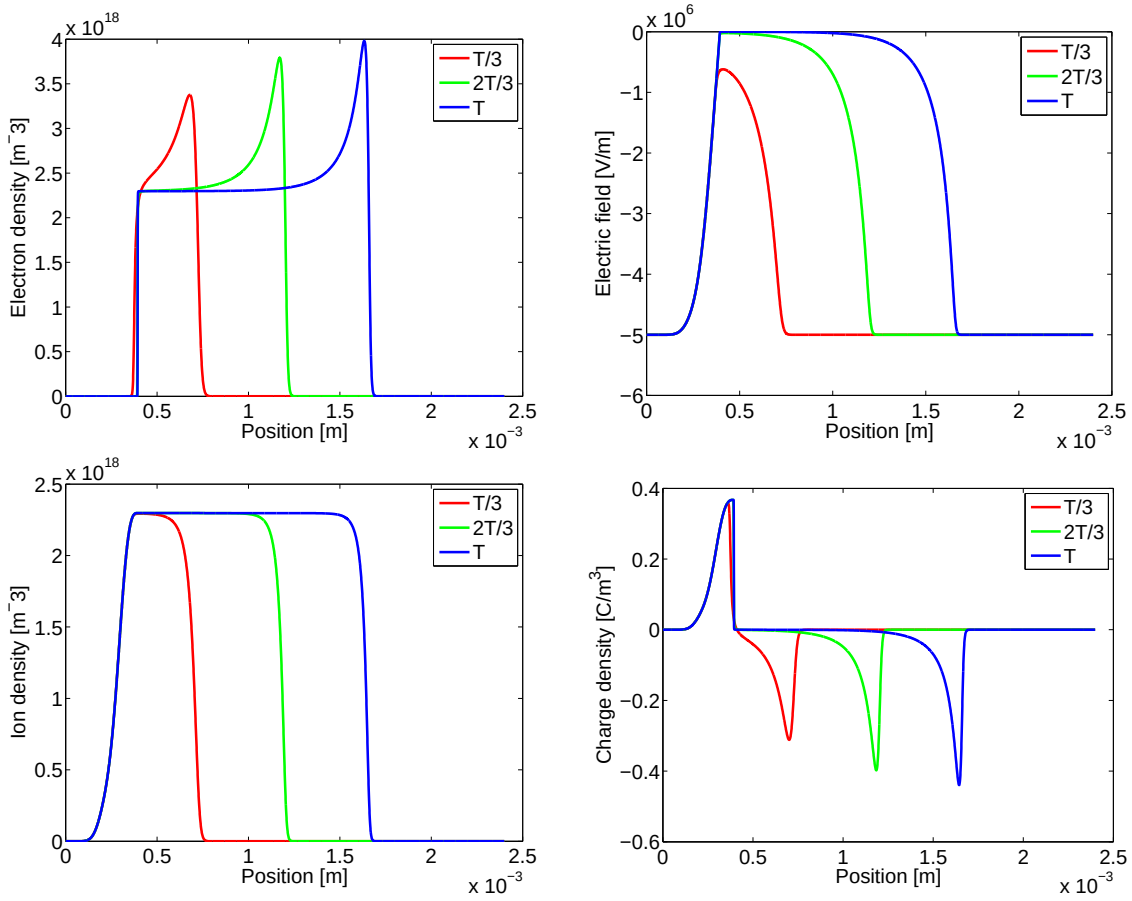

Figure 3.4: Breakdown simulation: spatial profiles of the electron density (upper left), ion density (lower left), electric field intensity (upper right) and charge density (lower right) at different times $t \in [T/3, 2/3\,T, T]$ with $T = 2.4 \times 10^{-8}$ s. These profiles are very similar to those obtained analytically (Fig. 3.1), although the analytic results apply to the steady state at very long times.

visions; the cell size is $\Delta x = L/N$, while the time-step size Δt is either kept constant, or is reduced at the same rate of Δx (constant grid velocity $u_G := \Delta x/\Delta t$). We then compare the numerical results for a given value of N with a 'reference solution, n_{rs}' computed with $N = 64000$. Each of the convergence tables presents results with the three different versions MF2/3/4. The L^2-norm of the error is computed as follows:

$$Error(x) = |n_e(x) - n_{rs}(x)|, \;\; L^2 = \frac{\sqrt{\sum_x (Error(x)^2) \times \Delta x}}{\sqrt{\sum_x (n_{rs}(x)^2) \times \Delta x}}$$

In Tables 3.8 and 3.9 we perform runs at fixed grid velocity $u_G = 9375$; this choice leads to roughly the same Courant parameter profile $C_0(x,t)$ used in all simulations, which approximately ranges between 0 and 2.67. Table 3.8 uses 2nd-order Strang's splitting. Here MF4 performs only marginally better than MF3, and both schemes show an order of convergence around 2. Therefore, we think that the splitting error is dominant in those simulations. Table 3.9 uses 4th-order McLachlan's splitting [57]. Here the very diffusive MF2 scheme performs more poorly than before, because of the higher number of stages. MF2 and MF3 perform significantly better than before, but not quite enough to compensate the additional work. MF3 shows an order of convergence that decreases toward 2, suggesting that the error is dominated by the 4th-order time splitting error for large values of Δt, and by the 2nd-order advection error when Δt gets smaller. MF4 shows an order of convergence consistently around 3, suggesting that the 3rd-order advection error is dominating in this case.

In Tables 3.10 and 3.11 we perform runs at fixed time-step size $\Delta t = 4 \times 10^{-12}$ s. As before, 2nd-order Strang's splitting is used in Table 3.10, and 4th-order McLachlan's splitting in Table 3.11. The amplitude of the Courant number profile grows linearly with N, with a peak value that grows from approximately. We expect that only the advection error should effect the convergence rate, so that the order of convergence should match the order of the local truncation error: in fact, in both tables we observe 2nd-order convergence for MF2, 3rd-order for MF3 and 4th-order for MF4. We also notice that McLachlan's splitting enhances the accumulation of advection errors, since the sum of the absolute values of the sub-steps is larger than Δt.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
2000	4.94×10^{-2}	–	9.06×10^{-4}	–	7.59×10^{-4}	–
4000	2.45×10^{-2}	1.01	2.21×10^{-4}	2.03	$1.1.83 \times 10^{-4}$	2.05
8000	1.15×10^{-2}	1.08	5.41×10^{-5}	2.03	4.48×10^{-5}	2.03
16000	4.98×10^{-3}	1.21	1.28×10^{-5}	2.08	1.06×10^{-5}	2.08
32000	1.66×10^{-3}	1.58	2.56×10^{-6}	2.33	2.12×10^{-6}	2.32

Table 3.8: Breakdown simulation: refinement analysis at fixed $\Delta x/\Delta t$ ratio. The moving-face (MF) version of the Convected Scheme (CS) is combined with 2nd-order Strang's splitting. The table reports the relative L^2 norm of the error in the electron density at the final time, with respect to a reference solution. All simulations employ a constant grid velocity $u_G = \Delta x/\Delta t = 9375$, and therefore have the same nominal profile of the Courant parameter $C_0(x,t)$, which approximately ranges between 0 and 2.67.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
2000	1.33×10^{-1}	–	5.37×10^{-4}	–	3.42×10^{-4}	–
4000	7.06×10^{-2}	0.92	1.04×10^{-4}	2.37	4.33×10^{-5}	2.98
8000	3.42×10^{-2}	1.04	2.21×10^{-5}	2.22	5.42×10^{-6}	3.00
16000	1.49×10^{-2}	1.19	4.89×10^{-6}	2.18	6.71×10^{-7}	3.02
32000	5.02×10^{-3}	1.57	9.42×10^{-7}	2.37	7.46×10^{-8}	3.17

Table 3.9: Breakdown simulation: refinement analysis at fixed $\Delta x/\Delta t$ ratio. Same as Table 3.8, but using 4th-order McLachlan's splitting [57].

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
2000	4.27×10^{-1}	–	1.57×10^{-2}	–	1.74×10^{-3}	–
4000	1.90×10^{-1}	0.70	2.80×10^{-3}	2.48	1.63×10^{-4}	3.42
8000	8.92×10^{-2}	1.09	2.81×10^{-4}	3.32	1.31×10^{-5}	3.63
16000	2.30×10^{-2}	1.96	3.40×10^{-5}	3.04	7.86×10^{-7}	4.06
32000	4.60×10^{-3}	2.32	5.23×10^{-6}	2.70	4.78×10^{-8}	4.04

Table 3.10: Breakdown simulation: refinement analysis at fixed Δt. The moving-face (MF) version of the Convected Scheme (CS) is combined with 2nd-order Strang's splitting. The table reports the relative L^2 norm of the error in the electron density at the final time, with respect to a reference solution. All simulations employ a constant time-step size $\Delta t = 4 \times 10^{-12}$ s.

	MF2		MF3		MF4	
N	L^2 error	Order	L^2 error	Order	L^2 error	Order
2000	3.60×10^{-1}	–	1.55×10^{-2}	–	2.40×10^{-3}	–
4000	2.39×10^{-1}	0.58	2.74×10^{-3}	2.50	2.42×10^{-4}	3.31
8000	1.26×10^{-1}	0.93	2.59×10^{-4}	3.40	2.21×10^{-5}	3.45
16000	4.21×10^{-2}	1.58	3.79×10^{-5}	2.77	1.71×10^{-6}	3.67
32000	1.17×10^{-2}	1.85	4.83×10^{-6}	2.97	1.32×10^{-7}	3.71

Table 3.11: Breakdown simulation: refinement analysis at fixed Δt. Same as Table 3.10, but using 4th-order McLachlan's splitting [57].

3.6 Summary

The physical problem of breakdown and its simulation was studied in this chapter. A method is presented which naturally achieves exact local density conservation and positivity preservation, essential for almost any plasma simulation, which allows Courant numbers across the range called for by the physical scales present ($C \gtrsim 1$) and gives third order accuracy in the breakdown simulations, using relatively simple analytic expressions to update the density. The scheme used here is based on an accurate semi-Lagrangian numerical description of fluid advection, and was combined by means of Strang or higher-order splitting with solution of Poisson's equation, as well as the rate equation for ionization.

This example is of great practical importance, but it is also a very challenging simulation task, because analytic results show that over time the electron density grows ever steeper at the front of the advancing plasma. In practical examples the discharge is of finite duration, so it should always be possible to define a mesh which resolves the electron density at the front. In our simulation results, the most severe errors were in fact at the back of the plasma, where another very steep region arose. This region is of less interest and so we have not analyzed it but excluded it from the error calculation in some of the cases presented. (If we were to extend the model to include other physical

processes, there is reason to believe the very sharp jump would also be mitigated. The most obvious examples of this are diffusion of density and photoionization, which would likely be handled effectively as part of the splitting scheme.)

We find second or third order convergence, with respect to the mesh size Δx and the time step Δt. High order splitting and the high order CS were required for third order convergence. Strang splitting gave a lower order of convergence, but unless the mesh was very highly refined, Strang splitting actually required less work for a given accuracy. Most importantly, Δt can be chosen based on physical considerations rather than stability requirements. The new scheme, and in particular its 'moving face' version, shows promise for very efficient simulations.

It is likely that for long enough times, the convergence would become less good because the shape of the breakdown front gets closer to the analytic result, which is singular. This can be partially mitigated with a finer mesh, if necessary. This singularity does complicate the issue of what is the optimal order of the scheme, since the advantage of high order in space is in part the ability to use a coarser mesh. As a result, it is not clear if schemes which are nominally of higher order than the one presented here would achieve significant advantage in practice.

Chapter 4

Arbitrarily High Order Convected Scheme and Its Application

In this chapter, we will introduce an arbitrarily high order Convected Scheme (AHOCS) developed by Güçlü *et al.* [51]. In [51], the author first showed that an AHOCS can successfully solve the constant advection equation. Because this scheme is a very high order scheme, machine precision is reached very soon. Moreover, there is no limitation due to the CFL for the AHOCS. Hence, the result is accurate and found efficiently by using the arbitrarily high order Convected Scheme. Further, the author employed this scheme to solve equations such as the Vlasov equation by combining with Strang splitting or other high order splitting methods. Here, we expand and employ this arbitrarily high order Convected Scheme in other cases, such as the Wigner equation and the Vlasov equation with a discontinuous potential.

In section 4.1, we review the derivation of the AHOCS. We then describe how we solve the Wigner equation by using the AHOCS and FFT/IFFT in section 4.2. Then, we choose the two stream instability problem to verify the convergence. We introduce the solution of Schrödinger equation due to [59] in the semiclassical region and its

numerical discretization in section 4.3.1. In this section, we use our scheme to solve the Wigner equation and compare the numerical solution with the numerical solution of the Schrödinger equation to demonstrate its higher efficiency. In section 4.4, where we consider a potential which is discontinuous, the interface condition is introduced. The combination of the AHOCS and the first order MF scheme can be used and leads to a higher convergence order.

4.1 Arbitrarily High Order Convected Scheme for The Constant Advection Equation

In section 3.3, we introduced the CS and use it to solve the advection equation. The CS is used to solve by envisaging a moving-cell (MC) $C(t)$. At the end of the time step, the particles in the moving cell are remapped back onto the Eulerian mesh. In [48], the moving cell has a uniform density in physical space and

1) one can move the midpoint of the cell (Moving Midpoint, MM): each face has the same velocity, or

2) one can move the faces of the cell (Moving Face, MF): the two spatial faces of a cell are moved independently.

The accuracy of the simplest version of the scheme is only second order in space, so the leading error will generate a strong numerical diffusion. Therefore, a high-order remapping procedure is necessary, and we also developed the higher order MF scheme for the advection equation in Chapter 3.

When the advection equation has constant velocity such as Eq. (4.1), the new AHOCS has been developed [51].

$$\frac{\partial n}{\partial t} + u\frac{\partial n}{\partial x} = 0 \tag{4.1}$$

where $n(x,t)$ is the density and $u(x,t)$ is the velocity. In that paper, the author starts by Taylor expanding the exact solution as a series of spatial derivatives and equates this series to the expanded form of CS. If the expanded exact solution has $(N-1)$ smooth spatial derivatives, then the scheme will match the exact solution up to a local truncation error $O(\Delta x^N)$.

Before we show the derivation of the AHOCS, we review some parameters and equations from Chapter 3.

The starting point is the 1D constant advection equation:

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

The CS scheme is not limited by the CFL criterion: the normalized displacement of the moving cell coincides with the Courant number

$$C_0 := \left|\frac{u\Delta t}{\Delta x}\right|, \tag{4.2}$$

The displacement of the moving cell can be written as

$$C_0 = S + \alpha, \qquad S \in \mathbb{Z}, \qquad -1 < \alpha < 1, \tag{4.3}$$

where S is the integer part and α is the remainder, or noninteger part, of the Courant number. The 1D CS scheme can be rewritten as

$$n_{i+S}^{k+1} = \begin{cases} U_{i-1}^k n_{i-1}^k + (1-U_i^k)n_i^k & \text{if } U \geq 0, \\ (1+U_i^k)n_i^k - U_{i+1}^k n_{i+1}^k & \text{otherwise,} \end{cases} \tag{4.4}$$

Here, i denotes space, k means time and U is the fractional normalized displacement. If there is no correction, then $U(x,t) = \alpha$. In order to get the high order CS, a

small correction to the final position of the moving cell is employed. Hence, $U = [u + \tilde{u}]\frac{\Delta t}{\Delta x} - S = \alpha + \tilde{\alpha}$, where $\tilde{u}$ is an 'anti-diffusive correction' [49, 50].

The exact solution of Eq. (4.1) is $n(x, t + \Delta t) = n(x - u\Delta t, t)$ which means that at time $t + \Delta t$, the solution is equal to the solution at time t, but shifted by a distance $u\Delta t$ in space. By combining with Eq. (4.3), the exact solution can be reformulated as:

$$n(x + S\Delta x, t + \Delta t) = n(x - \alpha\Delta x, t) \tag{4.5}$$

By Taylor expanding the right hand side of the analytic solution in Eq. (4.5), about the point (x, t) for $N - 1$ terms, Eq. (4.5) can be rewritten as:

$$n(x + S\Delta x, t + \Delta t) = n(x, t) + \left(\sum_{p=1}^{N-1} (-\alpha)^p \frac{(\Delta x)^p}{p!} \frac{\partial^p}{\partial x^p} \right) n(x, t) + O(\Delta x^N). \tag{4.6}$$

For the sake of clarity, in the present derivation we will only consider the case $\alpha \geq 0$ and hence Eq (4.4) becomes:

$$n(x + S\Delta x) = n(x, t) - \big(U(x, t)n(x, t) - U(x - \Delta x, t)n(x - \Delta x, t)\big) \tag{4.7}$$

Similarly, Taylor expanding apply on Eq. (4.7) about the point (x, t) for $(N-1)$ terms as well:

$$n_{CS}(x + S\Delta x, t + \Delta t) = n(x, t) + \left(\sum_{p=1}^{N-1} (-\alpha)^p \frac{(\Delta x)^p}{p!} \frac{\partial^p}{\partial x^p} \right) U(x, t)n(x, t) + O(\Delta x^N). \tag{4.8}$$

As we mentioned before, we want to find the anti-diffusive Courant parameter $\tilde{\alpha}$ to correct the position and achieve the high order scheme. When we add a small correction, the solution of the CS can achieve only a local truncation error (LTE) of

the order of $(\Delta x)^N$, i.e.

$$\varepsilon(x,t) := n(x + S\Delta x, t + \Delta x) - n_{CS}(x + S\Delta x, t + \Delta x) = (\Delta x)^N$$

After comparing with Eq. (4.6) and Eq. (4.8), the LTE can be rewritten as:

$$\left(\sum_{p=1}^{N-1}(-\alpha)^p\frac{(\Delta x)^p}{p!}\frac{\partial^p}{\partial x^p}\right)n(x,t) - \left(\sum_{p=1}^{N-1}(-1)^p\frac{(\Delta x)^p}{p!}\frac{\partial^p}{\partial x^p}\right)U(x,t)n(x,t) = O(\Delta x^N). \tag{4.9}$$

Now, we choose a polynomial ansatz to express the product Un:

$$U(x,t)n(x,t) = \sum_{q=0}^{N-2}(-1)^q\beta_q(\alpha)(\Delta x)^q\frac{\partial^q n(x,t)}{\partial x^q}, \tag{4.10}$$

where the $(N-1)$ coefficients $\beta_q(\alpha)$ are unknown functions and need to be determined from Eq. (4.9). Substituting Eq. (4.10) into the second term of Eq. (4.9), we obtain:

$$\begin{aligned}
\sum_{p=1}^{N-1}(-1)^p\frac{(\Delta x)^p}{p!}\frac{\partial^p(Un)}{\partial x^p} &= \sum_{p=1}^{N-1}(-1)^p\frac{(\Delta x)^p}{p!}\frac{\partial^p}{\partial x^p}\left(\sum_{q=0}^{N-2}(-1)^q\beta_q(\alpha)(\Delta x)^q\frac{\partial^q n(x,t)}{\partial x^q}\right)\\
&= \sum_{p=1}^{N-1}\sum_{q=0}^{N-2}(-1)^{p+q}(\Delta x)^{p+q}\frac{\beta_q}{p!}\frac{\partial^{p+q}n(x,t)}{\partial x^{p+q}}\\
&= \sum_{p=1}^{N-1}\sum_{q=0}^{N-2}(-1)^r(\Delta x)^r\frac{\beta_q}{(r-q)!}\frac{\partial^r n(x,t)}{\partial x^r}\\
&= \sum_{r=1}^{N-1}(-1)^r\left[\sum_{q=0}^{r-1}\frac{\beta_q}{(r-q)!}\right](\Delta x)^r\frac{\partial^r n(x,t)}{\partial x^r}
\end{aligned}$$

After replacing the index r with the original p and substituting this expression into Eq. (4.9), we get $(N-1)$ linear equations:

$$\sum_{q=0}^{p-1}\frac{\beta_q}{(p-q)!} = \frac{\alpha^p}{p!},\ p = 1, 2, ...(N-1) \tag{4.11}$$

Finally, we obtain the $(N-1)$ coefficients $\beta_q(\alpha)$ for any given α:

$$\begin{aligned} \beta_0(\alpha) &= \alpha, \\ \beta_p(\alpha) &= \frac{\alpha^{p+1}}{(p+1)!} - \sum_{q=0}^{p-1} \frac{\beta_q(\alpha)}{(p+1-q)!}, \text{ for } p > 1 \end{aligned} \tag{4.12}$$

Eq. (4.12) can also be expressed explicitly as follows (here shown for $N = 6$):

$$\begin{bmatrix} \frac{1}{1!} & 0 & 0 & 0 & 0 \\ \frac{1}{2!} & \frac{1}{1!} & 0 & 0 & 0 \\ \frac{1}{3!} & \frac{1}{2!} & \frac{1}{1!} & 0 & 0 \\ \frac{1}{4!} & \frac{1}{3!} & \frac{1}{2!} & \frac{1}{1!} & 0 \\ \frac{1}{5!} & \frac{1}{4!} & \frac{1}{3!} & \frac{1}{2!} & \frac{1}{1!} \end{bmatrix} \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \frac{\alpha^1}{1!} \\ \frac{\alpha^2}{2!} \\ \frac{\alpha^3}{3!} \\ \frac{\alpha^4}{4!} \\ \frac{\alpha^5}{5!} \end{bmatrix}$$

which can be inverted to give:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} \frac{B_0}{0!} & 0 & 0 & 0 & 0 \\ \frac{B_1}{1!} & \frac{B_0}{0!} & 0 & 0 & 0 \\ \frac{B_2}{2!} & \frac{B_1}{1!} & \frac{B_0}{0!} & 0 & 0 \\ \frac{B_3}{3!} & \frac{B_2}{2!} & \frac{B_1}{1!} & \frac{B_0}{0!} & 0 \\ \frac{B_4}{4!} & \frac{B_3}{3!} & \frac{B_2}{2!} & \frac{B_1}{1!} & \frac{B_0}{0!} \end{bmatrix} \cdot \begin{bmatrix} \frac{\alpha^1}{1!} \\ \frac{\alpha^2}{2!} \\ \frac{\alpha^3}{3!} \\ \frac{\alpha^4}{4!} \\ \frac{\alpha^5}{5!} \end{bmatrix} \tag{4.13}$$

where B is a Bernoulli number and $B_0 = 1$, $B_1 = -\frac{1}{2}$, $B_2 = \frac{1}{6}$, $B_3 = 0$, $B_4 = -\frac{1}{30}$. After simplifying Eq. (4.13), we obtain a general form expressing $\beta_q(\alpha)$ in terms of Bernoulli numbers:

$$\beta_p(\alpha) = \sum_{q=0}^{p} \frac{B_q}{q!} \frac{\alpha^{p+1-q}}{(p+1-q)!}, \tag{4.14}$$

According to the derivation by Jacob Bernoulli in 1713[60], Bernoulli numbers are

defined as polynomials in terms of sums of the powers of consecutive integers,

$$\sum_{k=0}^{m-1} k^n = \sum_{k=0}^{n} \frac{B_k}{k!} \frac{n!}{(n+1-k)!} m^{n+1-k} = \frac{B_{n+1}(m) - B_{n+1}(0)}{n+1}, \tag{4.15}$$

Finally, the coefficient $\beta_q(\alpha)$ can be expressed in terms of Bernoulli polynomials:

$$\beta_q(\alpha) = \frac{B_{p+1}(\alpha) - B_{p+1}(0)}{(p+1)!} \tag{4.16}$$

The case $\alpha < 0$ can be derived similarly. The final general form for the coefficient $\beta_q(\alpha)$, for any value of α, is

$$\begin{aligned} &\beta_0(\alpha) = \alpha, \\ &\beta_q(\alpha) = \frac{B_{p+1}(<\alpha>) - B_{p+1}(0)}{(p+1)!} \ (p > 1) \end{aligned} \tag{4.17}$$

Here $<\alpha>:=\mod(\alpha, 1)$, which means that $<\alpha>=\alpha$ for $0 \leq \alpha < 1$ and $<\alpha>= 1+\alpha$ for $-1 < \alpha < 0$.

When the computational region is periodic, we can use the discrete Fourier transform to deal with the derivatives in Eq. (4.10). Here, the fast Fourier transform (FFT) and inverse fast Fourier transform (IFFT) can be applied for a large order of accuracy N to speed up the simulation. The function (Un) can be solved for as follows:

$$\begin{aligned} U(x,t)n(x,t) &= \mathscr{F}^{-1}\left\{ \mathscr{F}\left[\sum_{q=0}^{N-2} (-1)^q \beta_q(\alpha) (\Delta x)^q \frac{\partial^q n(x,t)}{\partial x^q} \right] \right\} \\ &= \mathscr{F}^{-1}\left\{ \sum_{q=0}^{N-2} (-1)^q \beta_q(\alpha) (\Delta x)^q (j\xi\Delta x)^q \mathscr{F}[n] \right\} \end{aligned} \tag{4.18}$$

The following algorithm is the detailed implementation of the arbitrarily high order CS in Algorithm 2. In Algorithm 2, we employ a limiter for ensuring positivity preservation

and a filter (Appendix A) for obtaining a stable numerical scheme. The positivity preservation can be obtained by limiting the normalized displacement U_i^k. For $\alpha \geq 0$, $U_i^k \geq 0$ and $(1 - U_i^k) \geq 0$ are necessary to keep positivity preservation. On the other hand, for $\alpha < 0$, $U_i^k < 0$ and $(1 + U_i^k) > 0$ are necessary. Hence, we can combine these two conditions, for either positive or negative α, along with the general form to obtain:

$$\begin{aligned} &\text{if } \alpha \geq 0: \ 0 \leq U_i^k \leq 1 \text{ or } 0 \leq [nU]_i^k \leq n_i^k, \\ &\text{if } \alpha < 0: \ -1 \leq U_i^k \leq 0 \text{ or } -n_i^k \leq [nU]_i^k \leq 0, \end{aligned} \tag{4.19}$$

Here, (Un) is 'upwind fluxes'. Finally, the high-order flux Γ_i is obtained from adding the high-order correction in Eq. (4.10), the low order flux (αn), and the limiter as follows:

$$\begin{aligned} &\text{if } \alpha \geq 0: \ [nU]_i^k = \min(\max(0, \Gamma_i), n_i^k), \\ &\text{if } \alpha < 0: \ [nU]_i^k = \min(\max(-n_i^k, \Gamma_i), 0), \end{aligned} \tag{4.20}$$

In Algorithm 2, both the spatial index i and frequency index r range from 0 to $(N_x - 1)$, where N_x is the number of mesh subdivisions. Moreover, the scalar quantity $\omega := \exp(2\pi j/N_x)$ represents the N_xth primitive root of unity.

4.1.1 Numerical Test

The purpose of this section is to verify mass conservation, positivity preservation, and the expected order of the higher order CS for the constant advection equation.

Example 4.1.1. The aforementioned schemes are applied to the solution of the 1D constant advection equation:

$$\frac{\partial n}{\partial t} + \frac{\partial n}{\partial x} = 0, \ x \in [-0.5, 0.5], \ t \in [0, T],$$

with the periodic boundary condition $n(-0.5, t) \equiv n(0.5, t)$.

Algorithm 2 High order Convected Scheme with filtered trigonometric interpolation.[51]

1: Compute the normalized wave-number $\xi_r \Delta x \in [-\pi, \pi]$ supported by the mesh, and sample the filter $\hat{K}(\cdot)$ at those locations.

2:
$$\xi_r \Delta x = \begin{cases} 2\pi r/N_x; \text{ for} r \leq N_x, \\ 2\pi(r - N_x)/N_x; \text{ otherwise.} \end{cases} \quad \hat{K_r} = \hat{K}(\xi_r \Delta x), \ \forall r;$$

3: Given the Courant parameter $C := u\Delta t/\Delta x$, decompose it into integer and fractional parts as :

4: $C = S + \alpha$, $S \in \mathbb{Z}$, $\alpha \in [0, 1) \subset \mathbb{R}$

5: Given α, compute the correction polynomials $\beta_q(\alpha)$ according to Eq. (4.17) to achieve a nominal order of accuracy N:

$c_q = (-1)^q \beta_q(\alpha), \ q \in 1, 2, ..., N-2;$

6: Compute the discrete Fourier transform of the solution n_i^k at time t_k, using an FFT algorithm

$\hat{n_r} = \sum_{i=0}^{N_x - 1} n_i^k \omega^{-ir}, \ \forall r;$

7: Reduce roundoff noise by suppressing the modes below a threshold ε (in double precision we choose $\varepsilon = 2 \times 10^{-15}$):

8: $A = \max |\hat{n_r}|$, if $|\hat{n_r}| \leq A\varepsilon$ set $\hat{n_r} = 0$,

9: Compute the Fourier coefficients of the filtered high-order flux corrections:

$\hat{H_r} = \left(\sum_{q=1}^{N-2} c_q \cdot (j\xi_r \Delta x)^q \right) \hat{K_r} \hat{n_r}, \ \forall r;$

10: Compute the inverse discrete Fourier transform of $\hat{H_r}$ using an IFFT algorithm, and compute the high-order fluxes, Γ_i while enforcing positivity of the fractions:

$H_i = \frac{1}{N_x} \sum_{r=0}^{N_x - 1} \hat{H_r} \omega^{ir}, \ \Gamma_i = \alpha n_i^k + H_i$

11: Obtain the solution n_i^{k+1} at time $t_{k+1} = t_k + \Delta t$, according to the fractions, U. Cell (i) remaps into cells $(i + S + 1)$ and $(i + S)$ for $C \geq 0$ and into cells $(i + S - 1)$ and $(i + S)$ for $C < 0$:

$$\begin{cases} \Delta n_{i+S+1}^{k+1} = n_i^k \\ \Delta n_{i+S}^{k+1} = n_i^k - [Un]_i^k. \end{cases} \text{ for } \alpha \geq 0;$$

$$\begin{cases} \Delta n_{i+S-1}^{k+1} = -[Un]_i^k \\ \Delta n_{i+S}^{k+1} = n_i^k + [Un]_i^k. \end{cases} \text{ for } \alpha < 0;$$

The initial condition $n(x, t = 0) = n_0(x)$ is given by the symmetric superposition of three Gaussian bells:

$$n_0(x) = 0.5e^{-(\frac{x+0.2}{0.03})^2} + e^{-(\frac{x}{0.06})^2} + e^{-(\frac{x-0.2}{0.03})^2} \tag{4.21}$$

In this simulation, we choose a constant CFL parameter $C = u\Delta t/\Delta x = \Delta t/\Delta x = 0.32$, $u = 1$. Therefore, the number of subdivisions in the spatial domain is a multiple of 32 (i.e. $N_x = 32k$ with $k \in \mathbb{N}$) and the number of subdivisions in the time domain is a multiple of 100 (i.e. $N_t = 32k$ with $k \in \mathbb{N}$).

$\boldsymbol{N_x}$	L^2**-norm**	**Order**	**Min value**
32	5.56×10^{-2}	—	0
64	6.01×10^{-4}	6.53	0
128	2.45×10^{-10}	22.22	0
256	1.46×10^{-15}	17.35	0
512	1.60×10^{-15}	$(m.p.)$	0
1024	1.17×10^{-15}	$(m.p.)$	0
2048	8.17×10^{-16}	$(m.p.)$	0

Table 4.1: Constant Advection equation: refinement analysis for the f_{22} version of the Convected Scheme (CS) for the Example 4.1.1. The table reports the L^2-norm of the error (difference of analytic and numerical solutions) at the final time ($T = 1$), for progressively larger numbers of cells N_x. The algebraic order of convergence ('Order' column) is calculated as the base-2 logarithm of two successive error norms. All simulations employ the same Courant parameter $C_0 = 0.32$.

Table 4.1 shows the normalized L^2-norm of the error, Eq. (4.22) and the minimum value of the solution. The f_{22} scheme can reach the machine precision quickly, and the order of convergence is as we expected. Because the positivity limiter Eq. (4.20) is employed, we can see that the solutions are positive or equal to zero everywhere. The final solution obtained with the f_{22} scheme is shown in Fig. 4.1, for different numbers N_x of subdivisions in space ($\Delta x = 1/N_x$). Here, the analytic solution is $n(x, t + \Delta t) \equiv n(x - \Delta t, t)$.

$$L^2_{\text{err}} := \sqrt{\frac{\sum_i \left[n_i - n(x_i)\right]^2}{\sum_i \left[n(x_i)\right]^2}}, \tag{4.22}$$

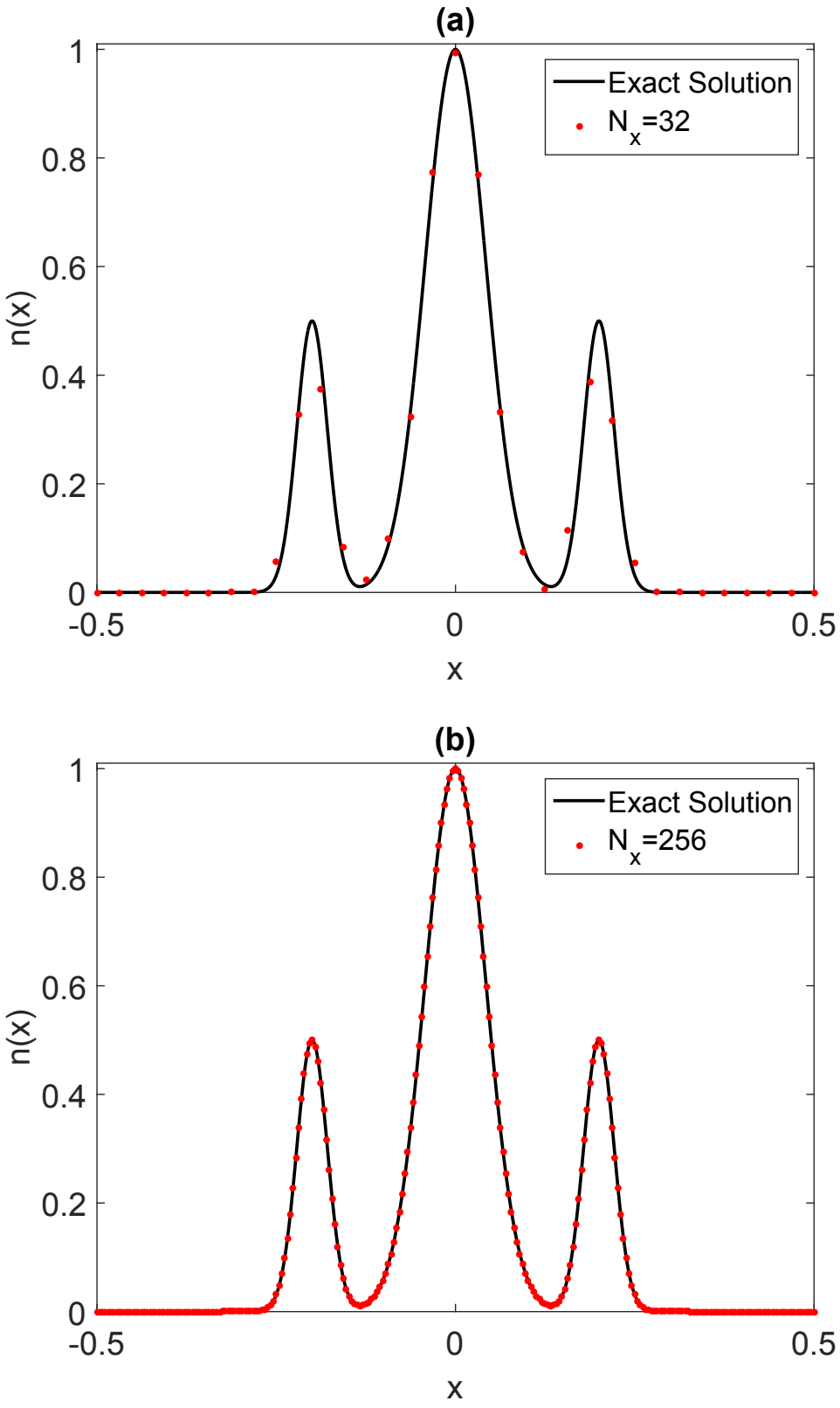

Figure 4.1: The solution from the f_{22} scheme of the CS for increasing number N_x of spatial subdivisions for the Example 4.1.1. (a) $N_x = 32$, (b) $N_x = 256$

4.2 Wigner-Poisson equation for 1D Landau damping

In Chapter 3, we employed different splitting schemes to allow the overall model to be made second or higher order. In this section, we employ the splitting scheme again to split the Wigner equation into two separate systems and solve the Wigner equation step by step. Here, we mainly focus on the Strang splitting method. An operator splitting method for the Wigner-Poisson problem was first introduced by Arnold and Ringho [61]. After splitting the Wigner equation, it consists of a transport term and a non-linear pseudo-differential operator. With the Poisson equation, the Wigner-Poisson equation is used to describe electrons in a self-consistent electrostatic field. Then, we use the Wigner-Poisson system with an operator splitting scheme to study the Landau damping phenomenon (nonlinear Landau damping and two stream instability problem) of a quantum system and other quantum devices [62–64]. In section 2.2, the Wigner equation is described by Eq. (2.18). Here, in order to describe the Wigner equation compactly, Eq. (2.18) can be rewritten as

$$\frac{\partial f(z,v,t)}{\partial t} + v\frac{\partial f(z,v,t)}{\partial z} + \Theta[V]f = 0 \tag{4.23}$$

where

$$\Theta[V]f = -\frac{m}{2i\pi\hbar^2}\int\int dz'dv' \exp(\frac{im(v-v')}{\hbar}z')\left(V(z-\frac{z'}{2}) - V(z+\frac{z'}{2})\right) f(z,v',t).$$

We then split the system, Eq. (4.23), into two separate systems (Eq.(4.24) and Eq.(4.25)) and use Strang splitting to update the solution from time t to time $t + \Delta t$:

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial z} = 0, \tag{4.24}$$

and

$$\frac{\partial f}{\partial t} + \Theta[V]f = 0. \tag{4.25}$$

1. $\Delta t/2$ step on Eq. (4.24);
2. Compute $n_e = \int f dv$, solving $\nabla_x^2 V = \frac{q_e}{\varepsilon}(n_e - n_0)$;
3. Δt step on Eq. (4.25);
4. $\Delta t/2$ step on Eq. (4.24);

Steps 1 and 4 use the transport equation. Because the advection velocity v does not depend on the independent variables x and t, $\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial z} = 0$ is a constant advection equation. Step 2 is the Poisson equation. Here, n_e is the electron density and n_0 is the background ion density. q_e is the electron charge, $q_e = -1.6 \times 10^{-19}$(C). The Poisson solver is described in Appendix B. After step 1, we need to solve the Poisson equation to obtain the instantaneous $V(z)$ for step 3. Step 3 is to solve the pseudo-differential equation. Because the Wigner potentials, $V(z - \frac{z'}{2})$ and $V(z + \frac{z'}{2})$, are nonlocal, they are constructed using cubic splines. Here, we will use the FFT/IFFT to deal with the pseudo-differential equation. The detailed derivation is shown below.

As mentioned before, step 1 and step 4 solve the constant advection equation. In Section 4.1, we showed that the f_{22} scheme shows excellent performance in solving the constant advection equation. In that section, we found that the f_{22} scheme reaches machine precision very soon and its convergence order is very high. Hence, in this

section, we use this scheme again to solve the transport equation in the Wigner equation and we expect that the result will have good performance as well.

Here, we need to note one thing about the use of the limiter for solving the Eq. (4.24). In Section 4.1, we employed the limiter to preserve positivity. However, the Wigner distribution can and normally does go negative ($n_i^k < 0$) because of the nonlocal potential. Hence, we need to add two more conditions for the limiters. Eq. (4.19) and Eq. (4.20) are corrected as follows:

$$\begin{aligned}
&\text{if } \alpha \geq 0 \text{ and } n_i^k \geq 0: \ 0 \leq U_i^k \leq 1 \text{ or } 0 \leq [nU]_i^k \leq n_i^k, \\
&\text{if } \alpha < 0 \text{ and } n_i^k \geq 0: \ -1 \leq U_i^k \leq 0 \text{ or } -n_i^k \leq [nU]_i^k \leq 0, \\
&\text{if } \alpha \geq 0 \text{ and } n_i^k < 0: \ 0 \leq U_i^k \leq 1 \text{ or } -n_i^k \leq [nU]_i^k \leq 0, \\
&\text{if } \alpha < 0 \text{ and } n_i^k < 0: \ -1 \leq U_i^k \leq 0 \text{ or } 0 \leq [nU]_i^k \leq n_i^k,
\end{aligned} \tag{4.26}$$

$$\begin{aligned}
&\text{if } \alpha \geq 0 \text{ and } n_i^k \geq 0: \ [nU]_i^k = \min(\max(0, \Gamma_i), n_i^k), \\
&\text{if } \alpha < 0 \text{ and } n_i^k \geq 0: \ [nU]_i^k = \min(\max(-n_i^k, \Gamma_i), 0), \\
&\text{if } \alpha \geq 0 \text{ and } n_i^k < 0: \ [nU]_i^k = \min(\max(n_i^k, \Gamma_i), 0), \\
&\text{if } \alpha < 0 \text{ and } n_i^k < 0: \ [nU]_i^k = \min(\max(0, \Gamma_i), -n_i^k),
\end{aligned} \tag{4.27}$$

When the v-space is periodic, Eq. (4.25) will be discretized in space by the spectral method and integrated in time exactly [62, 65].

$$\mathscr{F}\{\frac{\partial f(z,v,t)}{\partial t} = \Theta[V]f\} \ \Rightarrow \ \frac{\partial F(z,\lambda,t)}{\partial t} = \hat{\Theta}[V]f \tag{4.28}$$

where,

$$\begin{cases}
F(z,\lambda,t) = \mathscr{F}\{f(z,v,t)\} = \int f(z,v,t)\exp(-i\lambda v)dv \\
\hat{\Theta}[V]f = \mathscr{F}\{\Theta[V]f\} = \int \Theta[V]f\exp(-i\lambda v)dv
\end{cases}$$

Putting Eq. (2.17) into Eq. (4.28),

$$\begin{aligned}\hat{\Theta}[V]f &= \frac{m}{2i\pi\hbar^2}\int\int\int dz'dv'dv\exp\left(i(\frac{mz'}{\hbar}-\lambda)v\right)\exp(-\frac{imv'}{\hbar}z')\left(V(z-\frac{z'}{2})-V(z+\frac{z'}{2})\right)f(z,v')\\ &= \frac{m}{2i\pi\hbar^2}\int\int dz'dv'\exp(-\frac{imv'}{\hbar}z')\left(V(z-\frac{z'}{2})-V(z+\frac{z'}{2})\right)\left[\int\exp\left(iv(\frac{mz'}{\hbar}-\lambda)\right)dv\right]f(z,v')\end{aligned} \tag{4.29}$$

Here, we introduce an expression for the δ-function: $\delta(x-\alpha)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{ip(x-\alpha)}dp$ and Eq. (4.29) can be rewritten as

$$\begin{aligned}\hat{\Theta}[V]f &= \frac{m}{2i\pi\hbar^2}\int\int dz'dv'\exp(-\frac{imv'}{\hbar}z')\left(V(z-\frac{z'}{2})-V(z+\frac{z'}{2})\right)\left[2\pi\delta(\frac{mz'}{\hbar}-\lambda)\right]f(z,v')\\ &= \frac{m}{i\hbar^2}\int\int dz'dv'\exp(-\frac{imv'}{\hbar}z')\left(V(z-\frac{z'}{2})-V(z+\frac{z'}{2})\right)\left[\delta(\frac{mz'}{\hbar}-\lambda)\right]f(z,v')\end{aligned} \tag{4.30}$$

In order to simplify Eq. (4.30), we substitute $\frac{mz'}{\hbar}$ by X, so that $z'=\frac{\hbar}{m}X$ and $dz'=\frac{\hbar}{m}dX$.

$$\begin{aligned}\hat{\Theta}[V]f &= \frac{1}{i\hbar}\int\int dXdv'\exp(-iv'X)\left(V(z-\frac{\hbar}{2m}X)-V(z+\frac{\hbar}{2m}X)\right)\left[\delta(X-\lambda)\right]f(z,v',t)\\ &= \frac{1}{i\hbar}\int dv'\exp(-iv'\lambda)\left(V(z-\frac{\hbar}{2m}\lambda)-V(z+\frac{\hbar}{2m}\lambda)\right)f(z,v',t)\\ &= \frac{1}{i\hbar}\left(V(z-\frac{\hbar}{2m}\lambda)-V(z+\frac{\hbar}{2m}\lambda)\right)\int dv'\exp(-iv'\lambda)f(z,v',t)\\ &= \frac{1}{i\hbar}\left(V(z-\frac{\hbar}{2m}\lambda)-V(z+\frac{\hbar}{2m}\lambda)\right)F(z,\lambda,t),\end{aligned} \tag{4.31}$$

Eq. (4.28) can be written as

$$\begin{aligned}\frac{\partial F(z,\lambda,t)}{\partial t} &= \frac{1}{i\hbar}\left(V(z-\frac{\hbar}{2m}\lambda)-V(z+\frac{\hbar}{2m}\lambda)\right)F(z,\lambda,t)\\ &= \frac{q_e}{i\hbar}\left(\phi(z-\frac{\hbar}{2m}\lambda)-\phi(z+\frac{\hbar}{2m}\lambda)\right)F(z,\lambda,t),\end{aligned} \tag{4.32}$$

Some physical parameters are introduced here. V is the potential energy and ϕ is the electric potential, $V = q_e\phi$. Eq. (4.32) can solved analytically:

$$F(z,\lambda,t) = F(z,\lambda,0)e^{\frac{q_e}{i\hbar}\left(\phi(z-\frac{\hbar}{2m}\lambda)-\phi(z+\frac{\hbar}{2m}\lambda)\right)\Delta t}, \tag{4.33}$$

Finally, $f(z,v,t)$ can be found by taking an inverse Fourier transform of $F(z,\lambda,t)$, $f(z,v,t) = \mathscr{F}^{-1}\{F(z,\lambda,t)\}$.

4.2.1 Numerical Test

In this section, we solve the Wigner-Poisson equation for two stream instability problems with quantum effects. The two stream instability is a very common instability in plasma physics. Here, we use the same example in [62] to verify our method in solving the Wigner equation.

Example 4.2.1. The initial condition for the two stream instability is

$$f(x,v,0) = \frac{1}{6\sqrt{\pi}}(1+5v^2)\exp(-v^2/2)\left(1+\alpha\cos(k_0x)\right),\ x\in[-L,L],\ v\in[-V,V], \tag{4.34}$$

with $\alpha = 1\times10^{-3}$, $k_0 = \pi/L$, $L = 15.39$ and $V = 6$. The quantum dispersion relation is given by [62]

$$\begin{aligned}\frac{\hbar k^3}{\omega_p^2} = &\frac{1}{6\sqrt{2}}\left(1+5(\frac{\omega}{k}+\frac{\hbar k}{2m})^2\right)Z\left\{\frac{1}{\sqrt{2}}(\frac{\omega}{k}+\frac{\hbar k}{2m})\right\}\\ &-\frac{1}{6\sqrt{2}}\left(1+5(\frac{\omega}{k}-\frac{\hbar k}{2m})^2\right)Z\left\{\frac{1}{\sqrt{2}}(\frac{\omega}{k}-\frac{\hbar k}{2m})\right\}+\frac{5\hbar k}{6m}.\end{aligned} \tag{4.35}$$

where ω_p is the plasma frequency and Z is defined as

$$Z(\xi) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\frac{\exp(-x^2)}{x-\xi}dx.$$

Once k and $\hbar$ are given, the growth rate of the two stream instability can be obtained from Eq. (4.35). In this section, we choose $\hbar = 0$, 8 and 15 to be the test values. For $\hbar = 8$ and $\hbar = 15$, the theoretical values of growth rate (γ) are $\gamma = 000757$ and $\gamma = -0.0029$, respectively.

According to the theoretical values of the growth rate, we find the instability gradually fades out for larger $\hbar$, and then the plasma oscillates with a fixed frequency. Hence, in Fig. 4.2, for $\hbar = 0$, the starting point for the instability is around $t = 20$. Comparing this result to the one for $\hbar = 8$, where the starting point for the instability is around at $t = 40$, we see that it occurs at a later time than for $\hbar = 0$. When $\hbar = 15$, there is no instability, even when we run the simulation until $t = 200$. Finally, the growth rate of our numerical solutions show good agreement with the theoretical results.

We can see the evolution of the distribution in Figs. 4.3- 4.5. From these three figures, $\hbar = 0$ shows instability earlier than $\hbar = 8$. For larger $\hbar$, such as $\hbar = 15$, there is no instability. These results correspond to Fig. 4.2.

Now, we calculate the L^2-norm error of the distribution, $f(x, v, t)$, by Eq. (4.36), and check the convergence for the classical and quantum cases in Tables 4.2– 4.4. For each $\hbar$, we computed the numerical solution with a fine mesh, $N_x = N_v = 512$, and a small time step, $\Delta t = 3.90625 \times 10^{-3}$, and refer to this as the "exact" solution, $f_{ex}(x, v)^{\hbar}$. First, we can see in these tables, for fixed N_x and N_v and varying Δt, the order of two that we expect because we are using the Strang splitting which is 2nd order in time. Moreover, for fixed Δt and varying N_x and N_v, we see that the error does not decrease for different numbers of subdivisions N_x and N_v. These results verify that our f_{22} scheme converges quickly, as mentioned in Section 4.1, and also show its

high efficiency.

$$Error(x,v) = |f^{\hbar}(x,v) - f_{ex}(x,v)^{\hbar}|,$$
$$L^2_{normalization} = \frac{\sqrt{\sum_{x,v}\{Error^2 \times \Delta x \times \Delta v\}}}{\sqrt{\sum_{x,v}\{f_{ex}(x,v)^2 \times \Delta x \times \Delta v\}}} \tag{4.36}$$

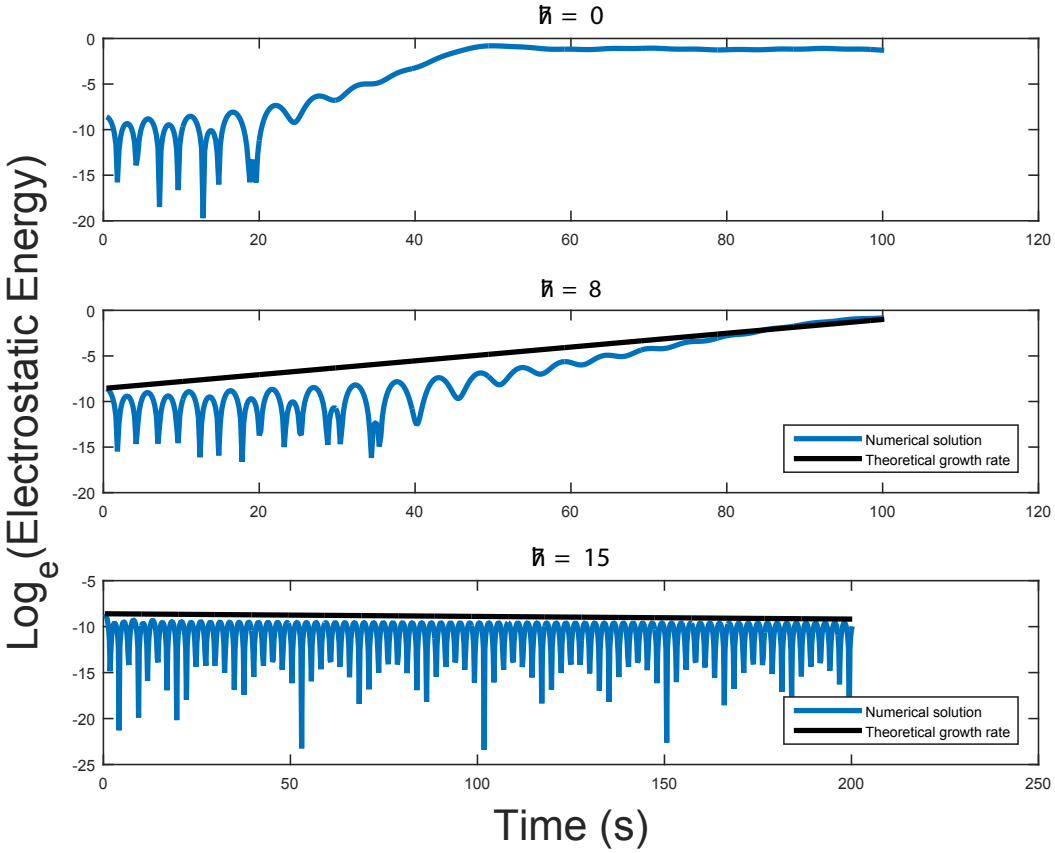

Figure 4.2: The figures for the two stream instability with the different $\hbar$. Here $\hbar$ is 0, 8 and 15, respectively. When the quantum parameter $\hbar$ becomes larger, the plasma oscillation shows a linear damping behavior because the quantum effect overcomes the instability. [62]

4.2.2 Summary

In this section, we employed the f_{22} scheme and FFT/IFFT to solve the Wigner equation in x and v space for two stream instability problems. If the quantum parameter

	$N_x = N_v = 64$		$N_x = N_v = 128$		$N_x = N_v = 256$	
Δt	L^2 error	Order	L^2 error	Order	L^2 error	Order
0.5	3.17×10^{-5}	–	3.18×10^{-5}	–	3.18×10^{-5}	–
0.25	7.46×10^{-6}	2.09	7.48×10^{-6}	2.09	7.49×10^{-6}	2.09
0.125	1.84×10^{-6}	2.02	1.84×10^{-6}	2.02	1.84×10^{-6}	2.02
0.0625	4.56×10^{-7}	2.01	4.57×10^{-7}	2.01	4.58×10^{-7}	2.01
0.03125	1.12×10^{-8}	2.02	1.13×10^{-8}	2.02	1.13×10^{-8}	2.02

Table 4.2: Two stream instability simulation: The error in the L^2-norm at $T = 2.5$ with $\hbar = 0$. The f_{22} version of the Convected Scheme (CS) is combined with 2nd-order Strang splitting. The table reports the relative L^2 norm of the error in the electron distribution at the final time, with respect to a reference solution.

	$N_x = N_v = 64$		$N_x = N_v = 128$		$N_x = N_v = 256$	
Δt	L^2 error	Order	L^2 error	Order	L^2 error	Order
0.5	2.48×10^{-5}	–	2.48×10^{-5}	–	2.49×10^{-5}	–
0.25	5.87×10^{-6}	2.08	5.88×10^{-6}	2.08	5.89×10^{-6}	2.08
0.125	1.45×10^{-6}	2.02	1.45×10^{-6}	2.02	1.45×10^{-6}	2.02
0.0625	3.59×10^{-7}	2.01	3.60×10^{-7}	2.01	3.61×10^{-7}	2.01
0.03125	8.80×10^{-8}	2.06	8.87×10^{-8}	2.02	8.90×10^{-8}	2.02

Table 4.3: Two stream instability simulation: The error in the L^2-norm at $T = 2.5$ with $\hbar = 8$. The f_{22} version of the Convected Scheme (CS) is combined with 2nd-order Strang splitting. The table reports the relative L^2 norm of the error in the electron distribution at the final time, with respect to a reference solution.

	$N_x = N_v = 64$		$N_x = N_v = 128$		$N_x = N_v = 256$	
Δt	L^2 error	Order	L^2 error	Order	L^2 error	Order
0.5	1.85×10^{-5}	–	1.85×10^{-5}	–	1.86×10^{-5}	–
0.25	4.41×10^{-6}	2.07	4.43×10^{-6}	2.07	4.43×10^{-6}	2.07
0.125	1.09×10^{-6}	2.02	1.09×10^{-6}	2.02	1.09×10^{-6}	2.02
0.0625	2.69×10^{-7}	2.02	2.71×10^{-7}	2.01	2.72×10^{-7}	2.01
0.03125	6.69×10^{-8}	2.01	6.66×10^{-8}	2.02	6.70×10^{-8}	2.02

Table 4.4: Two stream instability simulation: The error in the L^2-norm at $T = 2.5$ with $\hbar = 15$. The f_{22} version of the Convected Scheme (CS) is combined with 2nd-order Strang splitting. The table reports the relative L^2 norm of the error in the electron distribution at the final time, with respect to a reference solution.

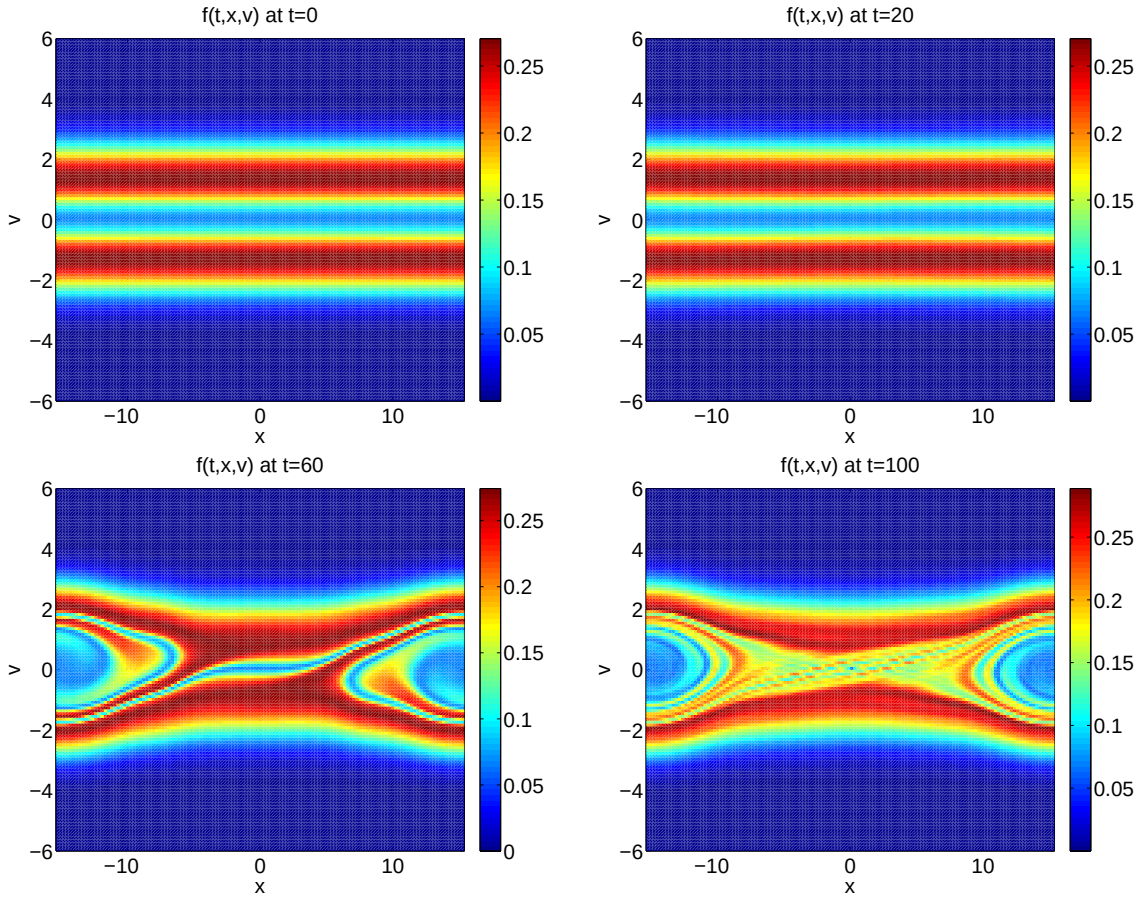

Figure 4.3: (Two stream instability simulation) Classical phase space distribution evolution for $\hbar = 0$. When instability appears, a hole structure forms in the distribution. The starting point for the instability is around $t = 40$. [62]

$\hbar$ becomes larger, the plasma oscillation shows a linear damping behavior because the quantum effect overcomes the instability. We can determine whether there is instability or not from the appearance of the density distribution. If instability exists, we find a set of connected vortices, or holes. Our results have good agreement with the theoretical values.

Finally, we calculate the L^2-norm of the error. The errors stop decreasing at certain values of N_x or N_v. This means that we can use a bigger mesh to obtain the same error as the smaller mesh, and achieve higher efficiency.

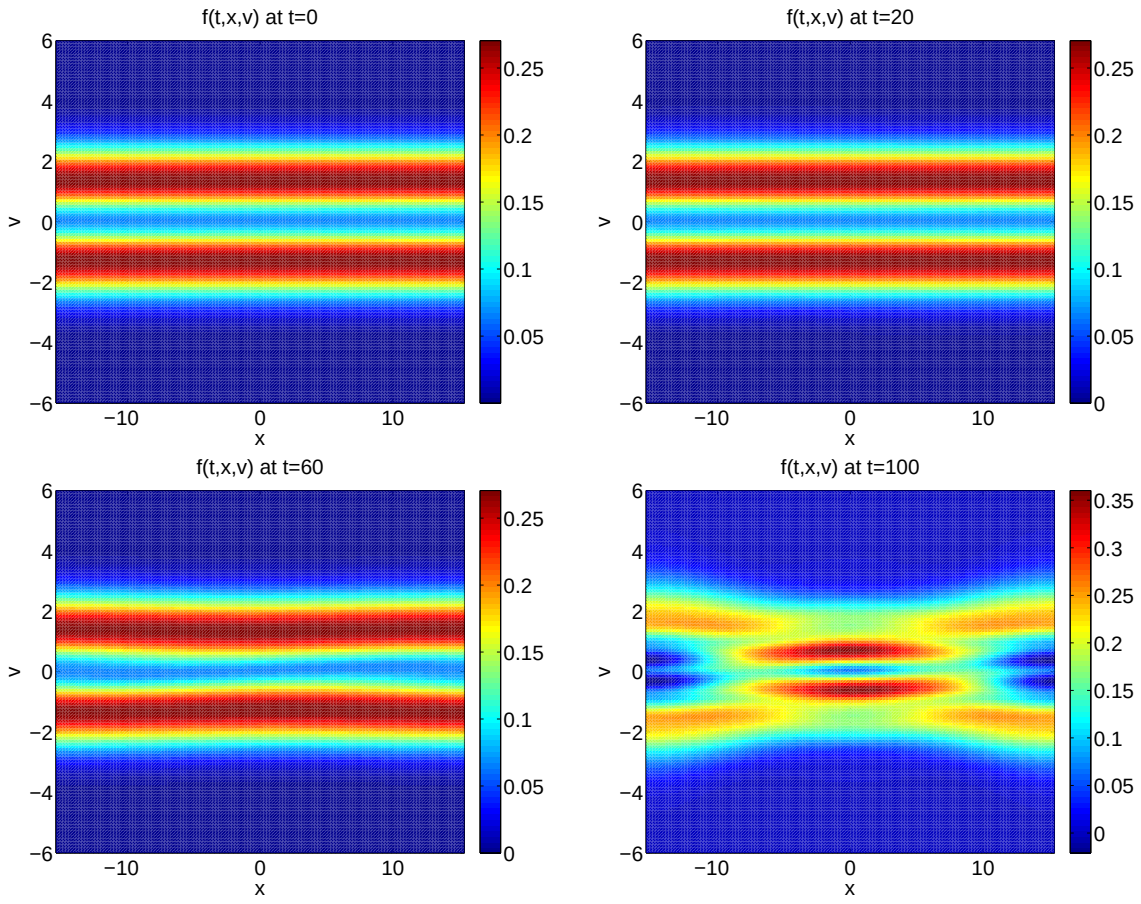

Figure 4.4: (Two stream instability simulation) Quantum phase space distribution evolution for $\hbar = 8$. When instability appears, a hole structure forms in the distribution. The starting point for the instability is around $t = 60$ which is longer than $\hbar = 0$. [62]

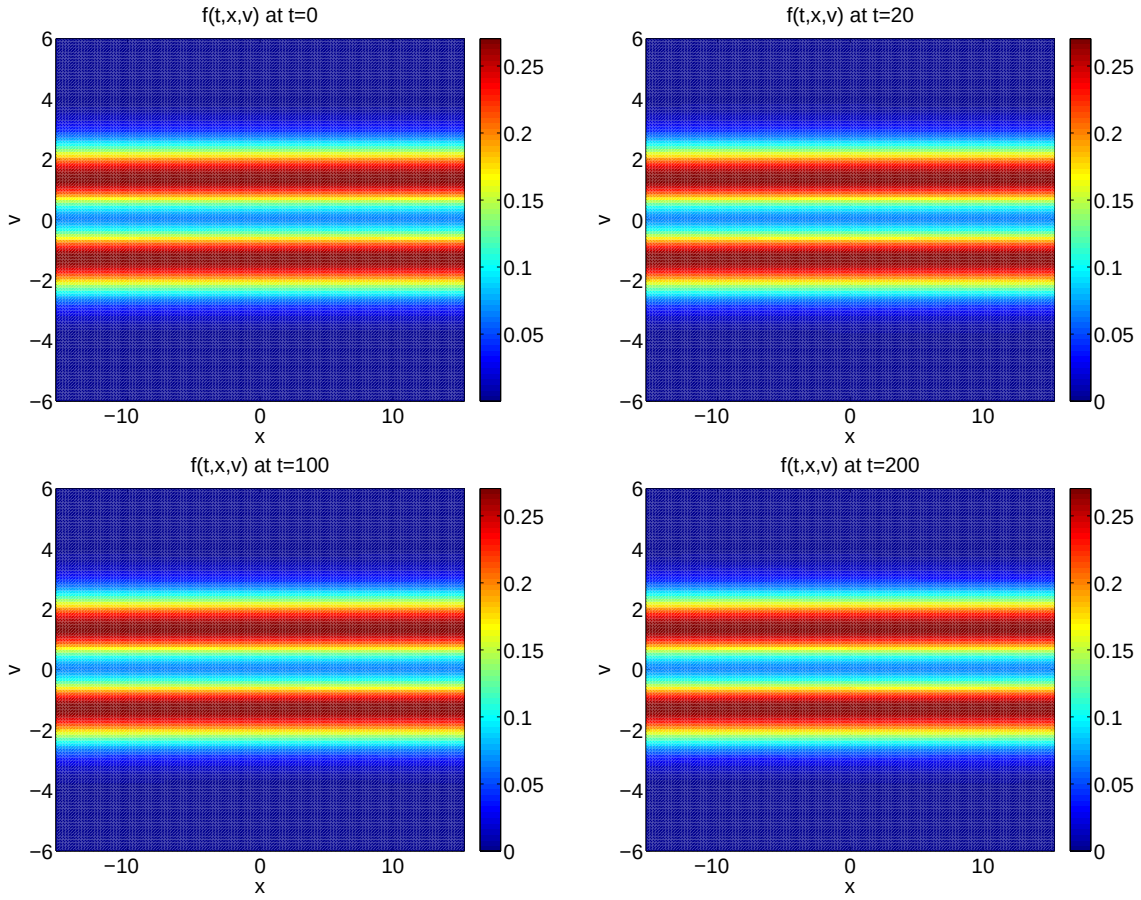

Figure 4.5: (Two stream instability simulation) Quantum phase space distribution evolution for $\hbar = 15$. For $\hbar = 15$, there is no formation of hole structure. The instability is overcome by the quantum effect. [62]

4.3 Wigner equation and Schrödinger Equation in The Semi-classical Regime

The Schrödinger equation can also be applied to many other problems in solid physics [59, 66, 67]. In this section, we compare our Wigner approach to the linear Schrödinger equation in the semi-classical regime, where the Planck constant (ε) is small. The solution of the Schrödinger equation oscillates with a wavelength of $O(\varepsilon)$. In order to obtain the correct physical observables for small ε, the spatial mesh size Δx and the time step Δt must be on the order of $O(\varepsilon)$. This means that Δx and Δt should be of the same order of the magnitude of ε and then the numerical solution is going to converge. Because of this, we use the phrase "mesh strategy" to represent setting $\Delta x = O(\varepsilon)$ and $\Delta t = O(\varepsilon)$.

In Section 2.9, we derived the Wigner equation from the Schrödinger equation, and we successfully developed the high efficiency scheme for solving the Wigner equation in Section 4.2. In this section, we first verify the mesh strategy for the Schrödinger equation in the semi-classical region, then we want to duplicate this good performance to solve the Wigner equation to overcome the mesh strategy for the Schrödinger equation in the semi-classical region.

4.3.1 Schrödinger Equation in The Semi-classical Regime

In this section, we introduce the method [59] to solve the Schrödinger equation in the semi-classical regime and discuss the "mesh strategy" for the solution of the Schrödinger

equation. The Schrödinger equation can be written as:

$$\begin{aligned} &\frac{\partial u}{\partial t} - i\frac{\varepsilon}{2}\frac{\partial^2 u}{\partial x^2} + \frac{i}{\varepsilon}V(x)u = 0,\ t \in R, x \in R \\ &u(x, t = 0) = u_0(x),\ x \in R \end{aligned} \tag{4.37}$$

where $\varepsilon \in (0, 1]$ denotes a small semi-classical parameter (the scaled Planck's constant).

In the absence of $V(x)$, a solution of the Schrödinger equation can be shown as [59],

$$u(t, x) = \exp\left(\frac{i}{\varepsilon}\left(\xi \cdot x - \frac{t}{2}|\xi|^2\right)\right),$$

where ξ is the wave vector. We see that u oscillates with a frequency $1/\varepsilon$ in time and space. With a small scale Planck constant, $\varepsilon \to 0$, the convergence of the wave function is hard to achieve.

In order to overcome this problem, it is necessary to use a tiny time and space discretization (Δt and Δx) for small ε [68, 69]. In [59], Bao *et al.* used a high-order scheme. Nevertheless, convergence analysis still confirms the mesh strategy, Eq. (4.38).

$$\Delta t = O(\varepsilon),\ \Delta x = O(\varepsilon) \tag{4.38}$$

Here, the author first uses Strang splitting to decompose the Schrödinger equation into two parts:

$$\frac{\partial u}{\partial t} - i\frac{\varepsilon}{2}\frac{\partial^2 u}{\partial x^2} = 0 \tag{4.39}$$

and

$$\frac{\partial u}{\partial t} + \frac{i}{\varepsilon}V(x)u = 0 \tag{4.40}$$

Eq. (4.39) will be solved using the following steps. First, taking the Fourier trans-

form of Eq. (4.39), we obtain:

$$\mathscr{F}\{\frac{\partial u(x,t)}{\partial t} - i\frac{\varepsilon}{2}\frac{\partial^2 u(x,t)}{\partial x^2} = 0\} \;\Rightarrow\; \frac{\partial \tilde{u}(X,t)}{\partial t} - i\frac{\varepsilon}{2}(i\xi)^2\tilde{u}(X,t) = 0 \tag{4.41}$$

where ξ is the conjugate of x and

$$\tilde{u}(X,t) = \mathscr{F}\{u(x,t)\} = \int u(x,t)\exp(-iXx)dx$$

Eq. (4.41) can be integrated in time exactly:

$$\frac{\partial \tilde{u}(X,t)}{\partial t} - i\frac{\varepsilon}{2}(j\xi)^2\tilde{u}(X,t) = 0 \;\Rightarrow\; \tilde{u}(X,\Delta t) = \tilde{u}(X,0)e^{-i\frac{\varepsilon}{2}\xi^2\Delta t}$$

Finally, $u(x,t)$ can be found from the inverse Fourier transform of $\tilde{u}(X,t)$:

$$u(x,t) = \mathscr{F}^{-1}\{\tilde{u}(X,t)\}$$

On the other hand, the ODE, Eq. (4.40) can be rewritten as below and will be then solved exactly:

$$\frac{\partial u}{u} + \frac{i}{\varepsilon}V(x)\partial t = 0 \;\Rightarrow\; u(\Delta t) = u(0)e^{-iV(x)\Delta t/\varepsilon}.$$

Let U_i^k be the approximation of $u(x_i,t_k)$ which is the solution at time $t = t_k = nk$. From time $t = t_k$ to time $t = t_{k+1}$, Eq. (4.37) can be split according to the Strang

method:

$$
\begin{aligned}
U_j^{\varepsilon,*} &= e^{-iV(x)\Delta t/2\varepsilon} U_j^{\varepsilon,k}, \ j = 0, 1, 2, ..., (M-1) \\
U_j^{\varepsilon,**} &= \frac{1}{M} \sum_{l=-M/2}^{M/2-1} e^{-i\varepsilon\Delta t\mu_l^2/2} \hat{U_l^{\varepsilon,*}} e^{i\mu_l(x-a)} \quad j = 0, 1, 2, ..., (M-1), \\
U_j^{k+1} &= e^{-iV(x)\Delta t/2\varepsilon} U_j^{\varepsilon,**}, \ j = 0, 1, 2, ..., (M-1),
\end{aligned}
\tag{4.42}
$$

where $\Delta x = (b-a)/M$, M is the number of mesh subdivisions, $x \in [a, b]$ and $\hat{U_l^{\varepsilon,*}}$ are the Fourier coefficients of $U^{\varepsilon,*}$.

4.3.2 Numerical test

In this section, we will describe the numerical solution of the Schrödinger equation with this Strang splitting, and verify the mesh strategy, Eq. (4.38). Then, the Wigner equation will be decomposed into two parts, also by Strang splitting. We employ the f_{22} scheme and FFT/IFFT to solve the Wigner equation and we show that an improved result can be achieved, compared to the numerical solution of the Schrödinger equation.

Example 4.3.1. The initial condition is taken as

$$
u^{\varepsilon}(x, t=0) = u_o^{\varepsilon}(x) = \left(e^{-25(x-0.5)^2}\right) e^{i(x+1)/\varepsilon} \tag{4.43}
$$

Let $V(x) = \frac{x^2}{2}$ and the problem be on the interval $[-2, 2]$ with periodic boundary. Here, we choose two different ε to check the convergence rate. For each ε, Bao *et al.* [59] compute the numerical solution with a very fine mesh, $\Delta x = \frac{1}{32768}$, and a very small time step ,$\Delta t = 0.00001$, and refer to this as the "exact" solution, u_{ex}^{ε}.

As shown in Tables 4.5 and 4.6, it is obvious that the error of the L^2-norm converges for the mesh strategy, $\Delta t = O(\varepsilon)$ and $\Delta x = O(\varepsilon)$. Taking $\varepsilon = 0.04$ for example, in Table 4.5, it is obvious that the numerical solution starts to converge at $\Delta t = 0.01$

	$\Delta x = \frac{1}{4}$	$\Delta x = \frac{1}{16}$	$\Delta x = \frac{1}{64}$	$\Delta x = \frac{1}{256}$	$\Delta x = \frac{1}{1024}$
$\Delta t = 0.16$	1.52×10^{0}	1.14×10^{-2}	1.14×10^{-2}	1.14×10^{-2}	1.14×10^{-2}
$\Delta t = 0.04$	1.53×10^{0}	7.14×10^{-4}	7.08×10^{-4}	7.08×10^{-4}	7.08×10^{-4}
$\Delta t = 0.01$	1.53×10^{0}	1.12×10^{-4}	4.23×10^{-5}	4.23×10^{-5}	4.23×10^{-5}
$\Delta t = 0.0025$	1.53×10^{0}	1.04×10^{-4}	2.76×10^{-6}	2.76×10^{-6}	2.76×10^{-6}
$\Delta t = 0.000625$	1.53×10^{0}	1.04×10^{-4}	1.73×10^{-7}	1.73×10^{-7}	1.73×10^{-7}

Table 4.5: The Schrödinger equation simulation in the semi-classical regime. The table reports the relative L^2 norm of the error in the position density at the final time ($t = 0.64$) and $\varepsilon = 0.04$, with respect to a reference solution. [59]

	$\Delta x = \frac{1}{64}$	$\Delta x = \frac{1}{256}$	$\Delta x = \frac{1}{1024}$	$\Delta x = \frac{1}{4096}$
$\Delta t = 0.16$	1.41×10^{0}	1.99×10^{-2}	1.99×10^{-2}	1.99×10^{-2}
$\Delta t = 0.04$	1.41×10^{0}	1.28×10^{-3}	1.28×10^{-3}	1.28×10^{-3}
$\Delta t = 0.01$	1.41×10^{0}	7.73×10^{-5}	7.73×10^{-5}	7.73×10^{-5}
$\Delta t = 0.0025$	1.41×10^{0}	4.83×10^{-6}	4.83×10^{-6}	4.83×10^{-6}
$\Delta t = 0.000625$	1.41×10^{0}	3.02×10^{-7}	3.02×10^{-7}	3.02×10^{-7}

Table 4.6: The Schrödinger equation simulation in the semi-classical regime. The table reports the relative L^2 norm of the error in the position density at the final time ($t = 0.64$) and $\varepsilon = 0.0025$, with respect to a reference solution. [59]

and $\Delta x = 1/64$ and they are of the same order of magnitude of $\varepsilon = 0.04$. Here, the L^2-norm is defined as:

$$Error(x) = |n^{\varepsilon}(x) - n_{ex}(x)^{\varepsilon}|, \;\; L^2 = \frac{\sqrt{\sum_x (Error(x)^2) \times \Delta x}}{\sqrt{\sum_x (n_{ex}(x)^2) \times \Delta x}} \tag{4.44}$$

where $n(x,t)$ can be defined by the wave function as follows [59, 70]:

$$n(x,t) = |u^{\varepsilon}(x,t)|^2 \tag{4.45}$$

In Section 4.2.1, the results show excellent performance when solving the Wigner equation, because we can get the same error using coarser meshes. This means the f_{22} scheme can achieve convergence rapidly, and then the error does not decrease when the number of mesh divisions increases. Because the Wigner equation can be

derived from the Schrödinger equation using the Wigner transform [70], we want to use this advantage to solve the Wigner equation by using the same initial condition as Example 4.3.2 without a converging mesh strategy, $\Delta t = O(\varepsilon)$ and $\Delta x = O(\varepsilon)$.

Before solving the Wigner equation, it is necessary to get the initial condition for the Wigner equation from the Fourier transform of the density matrix, Eq. (2.9). The initial condition for the Wigner equation is as follows:

$$f(x, v, t = 0) = \frac{1}{\sqrt{50\pi}} \frac{1}{\hbar} e^{-50(x-0.5)^2} e^{-\frac{(v-1)^2}{50\hbar^2}}$$

For the Schrödinger equation, $u(x,t)^\varepsilon$ only has one space variable, x. But, for the Wigner equation, $f(x, v, t)$ has two space variables, x and v. Therefore, when we solve the Wigner equation, we need to decide the region and mesh size of v-space first. In order to decide the mesh size, we can plot the initial condition for the position density, $n(x,t)$ from the wave function and the Wigner function to compare them. We let $n_w(x,t)$ be $n(x,t)$ from the wave function and $n_{wi}(x,t)$ be $n(x,t)$ from the Wigner function.

$$n(x,t) = |u^\varepsilon(x,t)|^2 = \int f(x,v,t)dv$$

For an acceptable region and mesh size, their initial conditions, $n_w(x,t)$ and $n_{wi}(x,t)$, must be identical. The mesh size is important for solving the Wigner equation. If the mesh size is too big, the initial conditions for two functions may be quite different and will cause different final results. Therefore, we need to decide the biggest mesh size according to the initial conditions. From Fig. 4.6, it is obvious that for a smaller ε, the initial condition $f(x, v, 0)$ is less smooth in v-space. This is because $f(x, v, 0)$ is a δ-function in v-space when $\varepsilon \to 0$, so $f(x, v, 0)$ is sharper in v-space when ε becomes smaller.

Because $f(x, v, 0)$ is sharper in v-space for smaller ε, we need to use a smaller Δv

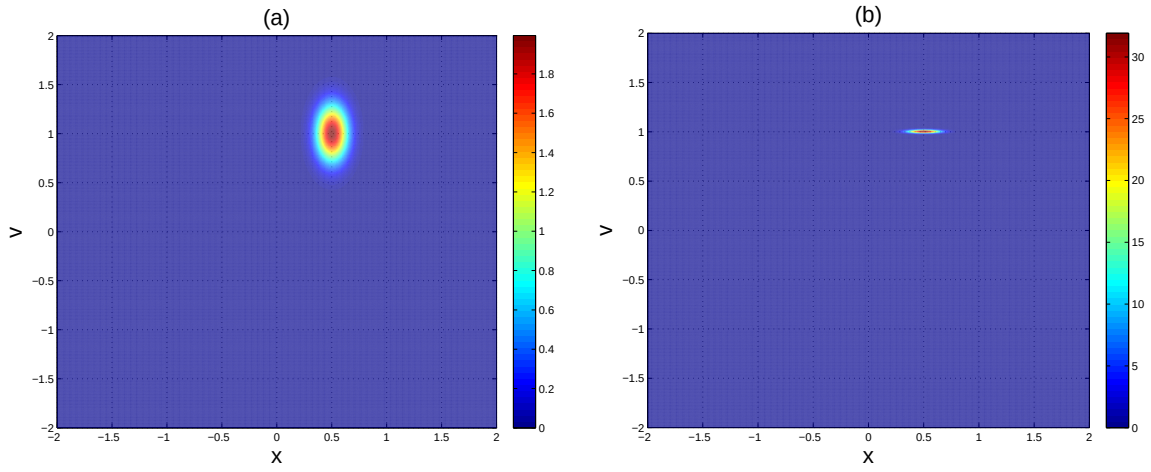

Figure 4.6: The initial condition of the Wigner function. The distribution density, $f(x, v, t = 0)$ for different ε. (a) $\varepsilon = 0.04$, (b) $\varepsilon = 0.0025$

for a smaller ε to get the same initial solution of the position density. Comparing Fig. 4.7 with Fig. 4.8, we can find that we can use a larger Δv to get the same initial solution of the position density for a larger ε.

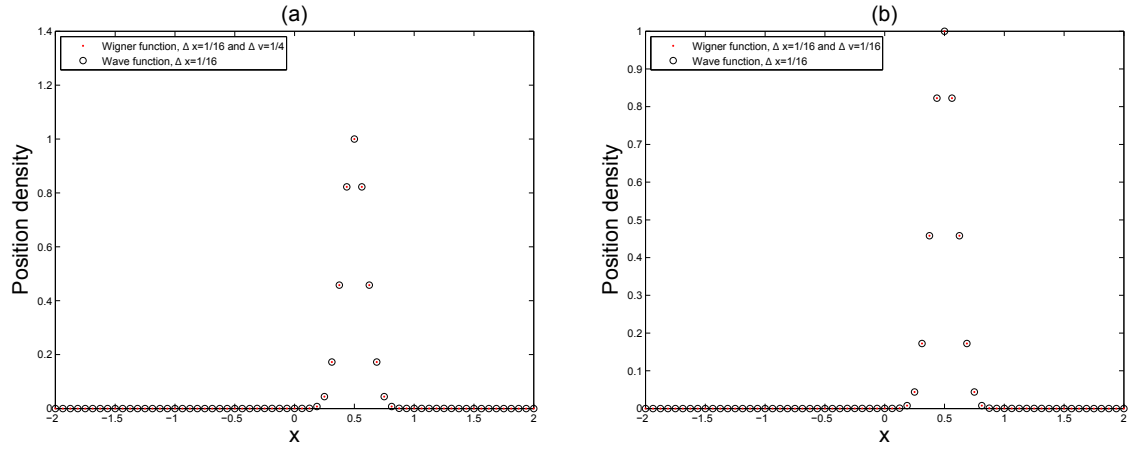

Figure 4.7: The initial condition of the Wigner function. The position density, $n(x, t = 0)$, for different mesh size $\Delta x = 1/16$ ($\varepsilon = 0.04$). (a) $\Delta v = 1/4$, (b) $\Delta v = 1/16$

As for the region, because we use the Fourier transform and f_{22} scheme to solve the Wigner equation, it is necessary to make sure that the distribution, $f(x, v, t)$, is periodic in x and v-space. We need to choose a large enough region for v-space and

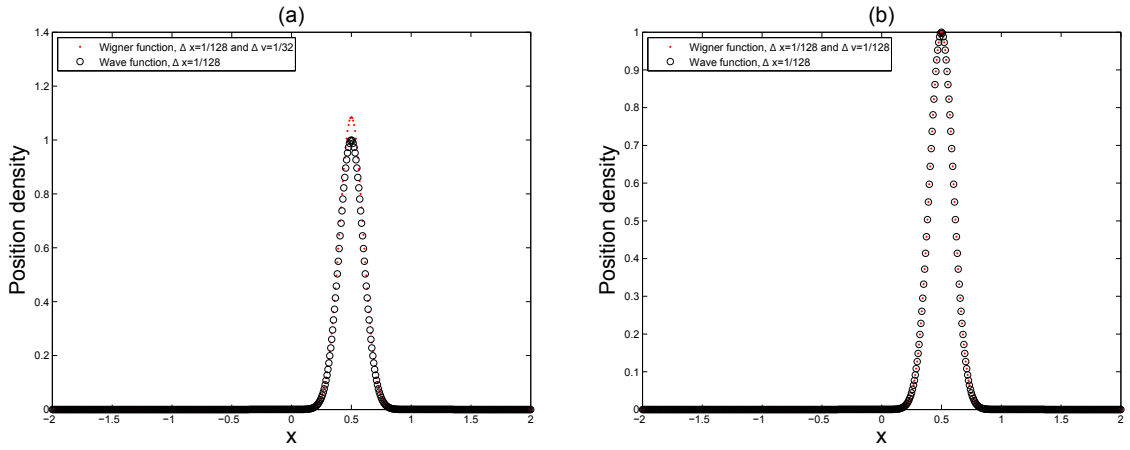

Figure 4.8: The initial condition of the Wigner function. The position density, $n(x, t = 0)$, for different mesh size $\Delta x = 1/128$ ($\varepsilon = 0.0025$). (a) $\Delta v = 1/32$, (b) $\Delta v = 1/256$

$f(x, v, t)$ must be almost zero at the boundary at any time.

	$\Delta x = \frac{1}{16}$	$\Delta x = \frac{1}{32}$	$\Delta x = \frac{1}{64}$	$\Delta x = \frac{1}{256}$
$\Delta t = 0.16$	7.88×10^{-3}	7.88×10^{-3}	7.88×10^{-3}	7.88×10^{-3}
$\Delta t = 0.04$	4.91×10^{-4}	4.91×10^{-4}	4.91×10^{-4}	4.91×10^{-4}
$\Delta t = 0.01$	3.07×10^{-5}	3.07×10^{-5}	3.07×10^{-5}	3.07×10^{-5}
$\Delta t = 0.0025$	2.04×10^{-6}	2.06×10^{-6}	2.06×10^{-6}	2.06×10^{-6}
$\Delta t = 0.000625$	6.75×10^{-7}	6.79×10^{-7}	6.79×10^{-7}	6.78×10^{-7}

Table 4.7: The Wigner equation simulation in the semi-classical regime. The table reports the relative L^2 norm of the error in the position density at the final time ($t = 0.64$) and $\varepsilon = 0.04$, with respect to a reference solution.

	$\Delta x = \frac{1}{32}$	$\Delta x = \frac{1}{64}$	$\Delta x = \frac{1}{128}$	$\Delta x = \frac{1}{256}$
$\Delta t = 0.16$	1.25×10^{-2}	1.38×10^{-2}	1.38×10^{-2}	1.38×10^{-2}
$\Delta t = 0.04$	3.92×10^{-3}	8.61×10^{-4}	8.61×10^{-4}	8.61×10^{-4}
$\Delta t = 0.01$	4.78×10^{-2}	5.36×10^{-5}	5.38×10^{-5}	5.38×10^{-5}
$\Delta t = 0.0025$	3.89×10^{-3}	3.76×10^{-6}	3.36×10^{-6}	3.36×10^{-6}
$\Delta t = 0.000625$	3.67×10^{-3}	1.69×10^{-6}	2.10×10^{-7}	2.10×10^{-7}

Table 4.8: The Wigner equation simulation in the semi-classical regime. The table reports the relative L^2 norm of the error in the position density at the final time ($t = 0.64$) and $\varepsilon = 0.0025$, with respect to a reference solution.

From Table 4.7, the result shows good performance because the error can be con-

trolled at bigger Δx. For the present scheme, when Δx and Δv are smaller than 1/16, the error is not going to decrease and achieves convergence. Comparing with the result, in Table 4.5 [59], we can use a bigger Δx and the error is of the same order as their result. Here, the L^2-norm is also calculated using Eq. (4.44), but $n(x,t)$ is calculated as follows:

$$n(x,t) = \int f(x,v,t)dv \tag{4.46}$$

Again, in Table 4.8, when Δx and Δv are smaller than 1/128, the error is not going to decrease and achieves convergence. It is an improvement over than the result in Table 4.6, as we can use bigger Δx and the error is of the same order.

4.4 Vlasov equation with discontinuous potentials

In this section, we use our AHOCS method to study the Vlasov equation with a discontinuous potential. This problem arises in the phase space description of geometrical optics [71–75]. Due to the discontinuous potential, where the L^1norm convergence rate for the finite difference method applied to the linear advection problem is at most order one half. Hence, we want to solve the Vlasov equation with a discontinuous potential to achieve a higher order of convergence than the estimated theory, where the L^1-norm is at most halfth order by using the finite difference method.

In section 4.1, we discussed the AHOCS for the constant advection equation. The AHOCS also can be applied to the Vlasov equation, Eq. (4.47) [51].

$$\frac{\partial f}{\partial t} + v\frac{\partial f}{\partial x} - \frac{dV}{dx}\frac{\partial f}{\partial v} = 0 \tag{4.47}$$

where $f(t,x,v)$ is the density distribution of a classical particle at position x, time t

and velocity v. $V(x)$ is the potential. The Vlasov equation is a different formulation of Newton's second law [76]:

$$\frac{dx}{dt} = v, \ \frac{dv}{dt} = \frac{dV}{dx} \tag{4.48}$$

which is a Hamiltonian system with the Hamiltonian:

$$H = \frac{1}{2}|v|^2 + V(x). \tag{4.49}$$

When $V(x)$ is smooth, the initial value problem of Eq. (4.48) is well-posed. However, when the potential V is discontinuous, its derivative, dV/dx, will be infinite at the point of discontinuity. Therefore, it is impossible to solve the differential equation, or Eq. (4.47), at that point.

In order to solve the Vlasov equation with discontinuous potentials, it is necessary to introduce an interface condition. The discontinuities in the potential correspond to potential barriers, at which incoming waves can be partially transmitted and reflected [77]. Then, the interface condition, which includes both reflection and transmission, lets us solve the differential equation with an infinite discontinuity point.

To describe the reflection/transmission at the potential barrier, we introduce two possible cases (total reflection and total transmission) and then derive the interface condition [78, 79]. Let a particle have momentum, $p > 0$, and meet a potential barrier at $x = X_d$ as shown in Fig. 4.9:

- Total reflection:

  If the particle does not have enough energy to overcome the potential barrier, then the particle is going to reflect at $x = X_d$ with a momentum $-p$. Hence, the density distribution f, $f(X_d^-, -p)$, after reflecting having entered with an

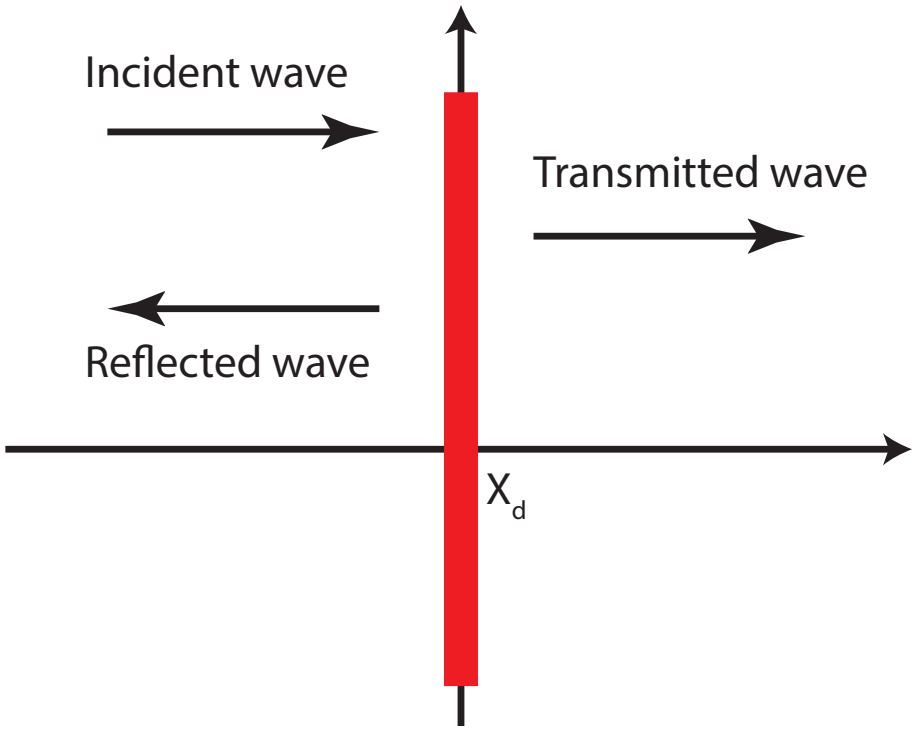

Figure 4.9: The particle is incident on the potential barrier. One condition leads to reflection and another condition to transmission.

incident momentum p is Eq. (4.50).

$$f(X_d^-, -p) = f(X_d^-, p) \tag{4.50}$$

where '$-$' means the left limit of the potential barrier.

- Total transmission:

  If the particle has enough energy to overcome the potential barrier, then the particle is going to transmit through $x = X_d^-$ and reach $x = X_d^+$. For this case, a particle is shown leaving of the interface $x = X_d$ with a momentum p^*, where $p^* = \sqrt{p^2 + 2V^- - 2V^+}$. Hence, the density distribution f, $f(X_d^+, p^*)$, after transmitting having entered with an incident momentum p is $f(X_d^+, p^*) = f(X_d^-, p)$, where '$+$' means the right limit of the potential barrier.

In order to obtain the general form of the interface condition including the partial reflection and partial transmission, we try to impose that the density distribution after transmission has the same form as the reflection, $f(X_d^-, -p)$, in Fig. 4.10 (a). This means we need to find an incident momentum $-p^*$ and transmitted momentum $-p$, according to Hamiltonian preservation, $p^* = -\sqrt{(-p)^2 + 2V^- - 2V^+}$, in Fig. 4.10 (b). Hence, the density distribution f after transmitting having entered with an incident momentum $-p^*$ is Eq. (4.51).

$$f(X_d^-, -p) = f(X_d^+, -p^*) \tag{4.51}$$

Finally, we combine the results above and write the interface condition in the following general way:

$$f(X_d^-, -p) = R(p)f(X_d^-, p) + T(p)f(X_d^+, -p^*) \tag{4.52}$$

It is obvious that Eq. (4.52) becomes Eq. (4.50) and Eq. (4.51) under total reflection ($R = 1$) and total transmission ($T = 1$), respectively.

Now, the general form for the interface condition including $p > 0$ and $p < 0$ is written as follows [78, 79]:

$$\begin{cases} f(x_1, p) = R(-p)f(x_1, -p) + T(-p)f(x_2, p_{x_2}(p)), \ p < 0 \\ f(x_2, p) = R(-p)f(x_2, -p) + T(-p)f(x_1, p_{x_1}(p)), \ p > 0 \end{cases} \tag{4.53}$$

where $p_{x_2}(p) = -\sqrt{p^2 + 2V_1 - 2V_2}$ and $p_{x_1}(p) = \sqrt{p^2 - 2V_1 + 2V_2}$.

4.4.1 First Order in Classical Mechanics

Here, we first provide the details for the classical situation. In classical mechanics, a particle will either be reflected ($R = 1$) or cross the potential barrier ($R = 0$). In this section, we first introduce the method which looks back from a final cell to an

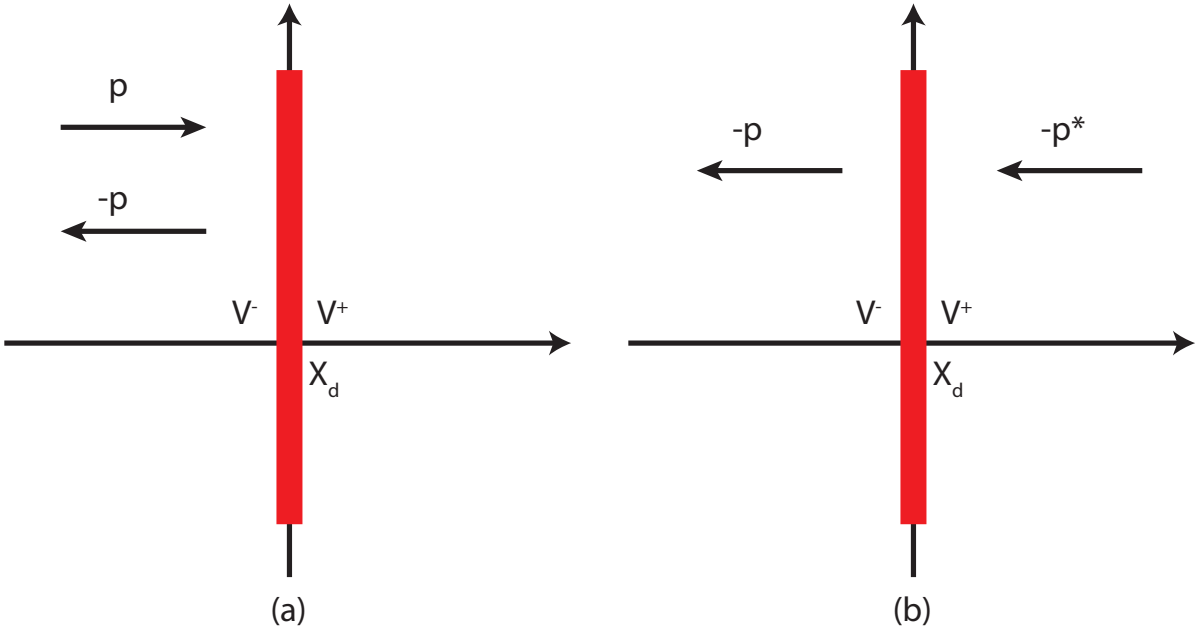

Figure 4.10: (a) Reflection: the particle with incident momentum p is reflected with momentum $-p$, (b) Transmission: the particle with incident momentum $-p^*$ is transmitted with momentum $-p$.

interpolated initial cell in [76]. Then, in order to employ the AHOCS on the equation to achieve the higher order, we introduce how the CS can look forward from an initial cell to several final cells.

If a particle has enough energy to overcome the height of the potential barrier, then this particle is going to cross the potential barrier with a changed momentum. During this process, the Hamiltonian $H = \frac{1}{2}p^2 + V$ should be preserved across the potential barrier:

$$\frac{1}{2}(v^+)^2 + V^+ = \frac{1}{2}(v^-)^2 + V^- \tag{4.54}$$

where the superscripts $\pm$ indicate the right and left limits of the quantity at the potential barrier, and $m = 1$.

The quantum case, including partial reflection and transmission, is very similar and will be introduced later. In [76], Jin *et al.* first introduce the formula of classical

mechanics for evaluating numerical fluxes of first order. In the following description, i is the space index , j is the velocity index and k is the time index.

The semidiscrete scheme for the Vlasov equation is:

$$\frac{f_{i,j}^{k+1}-f_{i,j}^{k}}{\Delta t}+v_j\frac{f_{i+1/2,j}^{-}-f_{i-1/2,j}^{+}}{\Delta x}-\frac{V_{i+1/2,j}^{-}-V_{i-1/2,j}^{+}}{\Delta x}\frac{f_{i,j+1/2}-f_{i,j-1/2}}{\Delta v}=0 \quad (4.55)$$

The superscripts '$+$' and '$-$' mean the right and left limit, respectively. Assume the potential barrier is at $x_{i-1/2}$ for $v_j>0$ and $x_{i+1/2}$ for $v_j<0$.

In [76], the main goal is to obtain $f_{i+1/2,j}^{-}$ for $v_j<0$ and $f_{i-1/2,j}^{+}$ for $v_j>0$ in Eq. (4.55). Once $f_{i+1/2,j}^{-}$ and $f_{i-1/2,j}^{+}$ are known, we can put them into Eq. (4.55) to get the updated density $f_{i,j}^{k+1}$. Here, we take two types of potential and $v_j>0$ for example. The case for v_j can be treated similarly and the detailed algorithm is described in Algorithm 3: .

- Continuous potential in Fig. 4.11 (a):

  ✓ $V_{i-1/2,j}^{+}=V_{i-1/2,j}^{-}$ for $v_j>0$:

  $f_{i+1/2,j}^{-}=f_{i,j}$ using the upwind scheme.

  1. $V_{i-1/2,j}^{+}=V_{i-1/2,j}^{-}\Rightarrow f_{i-1/2,j}^{+}=f_{i-1/2,j}^{-}$
  2. $f_{i-1/2,j}^{-}=f_{i-1,j}$ using the upwind scheme.
  3. Finally, $f_{i-1/2,j}^{+}=f_{i-1,j}$
  4. Eq. (4.55) can be rewritten:

$$\frac{f_{i,j}^{k+1}-f_{i,j}^{k}}{\Delta t}+v_j\frac{f_{i,j}^{k}-f_{i-1,j}^{k}}{\Delta x}=0$$

  or

$$f_{i,j}^{k+1}=f_{i,j}^{k}(1-\frac{v_j\Delta t}{\Delta x})+\frac{v_j\Delta t}{\Delta x}f_{i-1,j}^{k}=0$$

where $\frac{V^-_{i+1/2,j}-V^+_{i-1/2,j}}{\Delta x} = 0$ if the discontinuous barrier is only at $x = x_{i-1/2}$.

- Discontinuous potential:

  ✓ $V^+_{i-1/2,j} > V^-_{i-1/2,j}$ for $v_j > 0$:

  $f^-_{i+1/2,j} = f_{i,j}$ using the upwind scheme.

  * Reflection, Fig. 4.11 (b):

    1. Compute $f^+_{i-1/2,j}$ according to the interface condition, Eq. (4.53): $f^+_{i-1/2,j} = f^+_{i-1/2,m}$ where $v_m = -v_j$.
    2. $f^+_{i-1/2,m} = f_{i,m}$ using the upwind scheme.
    3. Finally, $f^+_{i-1/2,j} = f_{i,m}$
    4. Eq. (4.55) can be rewritten:

$$\frac{f^{k+1}_{i,j} - f^k_{i,j}}{\Delta t} + v_j \frac{f_{i,j} - f_{i,m}}{\Delta x} = 0$$

    or

$$f^{k+1}_{i,j} = f^k_{i,j}(1 - \frac{v_j \Delta t}{\Delta x}) + \frac{v_j \Delta t}{\Delta x} f_{i,m} = 0$$

  * Transmission, Fig. 4.11 (c):

    1. Compute the transmitted velocity, v^-: $v^- = \sqrt{(v_j)^2 + 2(V^+_{i-1/2} - V^-_{i-1/2})}$, where $v_m \le v^- < v_{m+1}$ for some m.
    2. Compute $f^+_{i-1/2,j}$ according to the interface condition, Eq. (4.53): $f^+_{i-1/2,j} = \frac{v_{m+1}-v^-}{\Delta v} f^-_{i-1/2,m} + \frac{v^- - v_m}{\Delta v} f^-_{i-1/2,m+1}$
    3. $f^-_{i-1/2,m} = f_{i-1,m}$ and $f^-_{i-1/2,m+1} = f_{i-1,m+1}$ using the upwind scheme.
    4. Finally, $f^+_{i-1/2,j} = \frac{v_{m+1}-v^-}{\Delta v} f_{i-1,m} + \frac{v^- - v_m}{\Delta v} f_{i-1,m+1}$

5. Eq. (4.55) can be rewritten:

$$\frac{f_{i,j}^{k+1}-f_{i,j}^{k}}{\Delta t}+v_j\frac{f_{i,j}-\frac{v_{m+1}-v^-}{\Delta v}f_{i-1,m}+\frac{v^--v_m}{\Delta v}f_{i-1,m+1}}{\Delta x}=0$$

or

$$f_{i,j}^{k+1}=f_{i,j}^{k}(1-\frac{v_j\Delta t}{\Delta x})+\frac{v_j\Delta t}{\Delta x}\left(\frac{v_{m+1}-v^-}{\Delta v}f_{i-1,m}-\frac{v^--v_m}{\Delta v}f_{i-1,m+1}\right)=0$$

In order to combine the present higher order scheme and achieve mass conservation, it is necessary to update the density or distribution function by following the initial cells. We used the MC to denote the moving cell as mentioned in Section 3.3. As a result, we are tracking the MC to see where it is going. At the end of the time step, the particles in the MC are remapped back onto the overlapped fixed cells according to the remapping rule [48]. The remapping rule for a continuous potential barrier is shown in Fig. 4.12 and Eq. (4.56).

$$\Delta n_i=\oint_{C_i\cap MC}n_{MC}(x)\,dx=n_{MC}(x)\oint_{C_i\cap MC}dx=n_{MC}V_{C_i\cap MC}=N_{MC}F_i \qquad (4.56)$$

Here, we are assuming that n_{MC} is constant in the MC and the symbol V is used to denote the volumes, and F_i is called the overlapping fraction. Therefore, in Fig. 4.12, U_0 is the fraction, $U_0=\frac{v\Delta t}{\Delta x}$, and v is the velocity of the MC. At the end of the time step ($t_{k+1}=t_k+\Delta t$), the particles in the MC ($f_{i,j}$) are remapped back onto the overlapped fixed cells according to the remapping rule for $v>0$ or $v<0$.

Next, we introduce the two cases for a discontinuous potential barrier from classical mechanics. The first case has $V^+>V^-$ in Fig. 4.13 (a), while the second case has $V^+<V^-$ in Fig. 4.13 (b). Here, we only consider $v_j>0$. The case of $v_j<0$ is treated similarly. It should be noted that v_j is the center velocity of the initial cell.

Algorithm 3 The detailed algorithm to generate the numerical flux is given below. The superscripts '$+$' and '$-$' mean the right and left limit, respectively. [76]

1: $v_j > 0$ and potential barrier is at $x_{i-1/2}$: $f^-_{i-1/2,j} = f_{i-1,j}$

- $V^-_{i-1/2} = V^+_{i-1/2}$:
  $f^+_{i-1/2,j} = f^-_{i-1/2,j} = f_{i-1,j}$.
- $V^-_{i-1/2} < V^+_{i-1/2}$:
  - Compute the incident velocity, v^-:
    $v^- = \sqrt{(v_j)^2 - 2(V^-_{i-1/2} - V^+_{i-1/2})}$, where $v_m \le v^- < v_{m+1}$ for some integer m.
  - Use the interface condition to compute $f^+_{i-1/2,j}$:
    $f^+_{i-1/2,j} = \frac{v_{m+1}-v^-}{\Delta v} f_{i-1,m} + \frac{v^- - v_m}{\Delta v} f_{i-1,m+1}$
- $V^-_{i-1/2} > V^+_{i-1/2}$:
  - Compute the incident velocity, v^-:
    $v^- = \sqrt{(v_j)^2 - 2(V^-_{i-1/2} - V^+_{i-1/2})}$, where $v_m \le v^- < v_{m+1}$ for some integer m.
  1. $v_j < \sqrt{2(V^-_{i-1/2} - V^+_{i-1/2})}$: Reflection
     - Use the interface condition to compute $f^+_{i-1/2,j}$:
       $f^+_{i-1/2,j} = f_{i,m}$ where $v_m = -v_j$.
  2. $v_j > \sqrt{2(V^-_{i-1/2} - V^+_{i-1/2})}$: Transmission
     - Use the interface condition to compute $f^+_{i-1/2,j}$:
       $f^+_{i-1/2,j} = \frac{v_{m+1}-v^-}{\Delta v} f_{i-1,m} + \frac{v^- - v_m}{\Delta v} f_{i-1,m+1}$

2: $v_j < 0$ and and potential barrier is at $x_{i+1/2}$: $f^+_{i+1/2,j} = f_{i+1,j}$

- $V^-_{i+1/2} = V^+_{i+1/2}$:
  $f^-_{i+1/2,j} = f^+_{i+1/2,j}$.
- $V^-_{i+1/2} < V^+_{i+1/2}$:
  - Compute the incident velocity, v^+:
    $v^+ = -\sqrt{(v_j)^2 - 2(V^+_{i+1/2} - V^-_{i+1/2})}$, where $v_m \leq v^+ < v_{m+1}$ for some integer m.
  1. $v_j < \sqrt{2(V^+_{i+1/2} - V^-_{i+1/2})}$: Reflection
     - Use the interface condition to compute $f^-_{i+1/2,j}$:
       $f^-_{i+1/2,j} = f_{i+1,m}$ where $v_m = -v_j$.
  2. $v_j > \sqrt{2(V^+_{i+1/2} - V^-_{i+1/2})}$: Transmission
     - Use the interface condition to compute $f^-_{i+1/2,j}$:
       $f^-_{i+1/2,j} = \frac{v_{m+1}-v^+}{\Delta v} f_{i+1,m} + \frac{v^+-v_m}{\Delta v} f_{i+1,m+1}$.
- $V^-_{i+1/2} > V^+_{i+1/2}$:
  - Compute the incident velocity, v^+:
    $v^+ = -\sqrt{(v_j)^2 - 2(V^+_{i+1/2} - V^-_{i+1/2})}$, where $v_m \leq v^+ < v_{m+1}$ for some m.
  - Use the interface condition to compute $f^-_{i+1/2,j}$:
    $f^-_{i+1/2,j} = \frac{v_{m+1}-v^+}{\Delta v} f_{i+1,m} + \frac{v^+-v_m}{\Delta v} f_{i+1,m+1}$.

When the MC crosses the discontinuous potential, we track the two faces (top and bottom) of the MC in velocity space, v_{top} and v_{bot} to find the fraction accurately according to the Hamiltonian preserving scheme in Eq. (4.54). After crossing the discontinuous potential, these two faces of the MC are going to new positions, v'_{top} and v'_{bot}, respectively. Again, at the end of the time step, the particles in the MC are remapped back onto the overlapped fixed cells according to the remapping rule [48].

As shown in Fig. 4.14, we classify the final position of the MC faces into the following two situations:

- v'_{top} and v'_{bot} are in the same cell, as shown in Fig. 4.14 (a):
  The fraction U_0 of particles in the MC crossing the potential barrier go to that cell.

- v'_{top} and v'_{bot} are not in the same cell, as shown in Fig. 4.14 (b). There are two methods to remap the particles as below:

  - Method 1: We divide the velocity between v_{top} and v_{bot} into N uniform divisions. Then, the new position of these divisions after crossing the barrier can be found from Eq. (4.54). Again, the density in each division can be remapped onto the overlapped fixed cell using the usual remapping rule, in Fig. 4.15 (a).
  - Method 2: We choose the values of v_1 and v_2 so that their transmitted velocities, v'_1 and v'_2, are at velocity grid points between v'_{top} and v'_{bot}. Then, the density in each division can be remapped onto the overlapped fixed cell using the usual remapping rule.

The detailed algorithm 4 is described as follows. Here, we focus on the initial position of $f_{i,j}$ and $v_j > 0$. As for the case where v'_{top} and v'_{bot} are not in the same cell, we only show the algorithm and results for method 1 because method 1 and method 2 give similar results. The case of $v_j < 0$ is treated similarly. There are two types of potential which a particle encounters:

- Continuous potential in Fig. 4.16(a):

  ✓ $v_j > 0$: $f_{i,j}$ goes to $(i+1, j)$ and (i, j).

  ✓ $v_j < 0$: $f_{i,j}$ goes to $(i-1, j)$ and (i, j).

- Discontinuous potential:

  – Reflection in Fig. 4.16 (b):

    $f_{i,j}$ is going to $(i, -j)$.

  – Transmission in Fig. 4.16 (c):

    ✓ $v_j > 0$: $f_{i,j}$ goes to $(i+1, m)$.

    ✓ $v_j < 0$: $f_{i,j}$ goes to $(i-1, m)$.

    v_m may not be a grid point, $v_p \leq v_m < v_{p+1}$ for some p and m.

When we use the initial cells and track their final position, we can apply the higher order CS on the discontinuous potential with quantum effects. In the next Section 4.4.1.1, the numerical results are shown for these two schemes first. Later, we will show the algorithm and results that include quantum effects and are higher order.

4.4.1.1 Numerical test

In this section, we present a numerical example which was already used in [76]. The scheme is described in Algorithm 3. As mentioned before, the scheme described in [76]

Algorithm 4 The detailed algorithm of CS on the discontinuous potential (classical mechanics), $v_j (v_j > 0)$ is the initial velocity, $v_{top} = v_j + \Delta v/2$, $v_{bot} = v_j - \Delta v/2$ and $v'_{top} = \sqrt{v_{top}^2 + 2(V^-_{i+1/2} - V^+_{i+1/2})}$, $v'_{bot} = \sqrt{v_{bot}^2 + 2(V^-_{i+1/2} - V^+_{i+1/2})}$

1: Compute the Courant parameter: $U_0 = v_j \Delta t/\Delta x$.

2: $V^-_{i+1/2} = V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells from (i, j) to $(i + 1, j)$:
  $\Delta f^{k+1}_{i+1,j} = U_0 f^k_{i,j}$ and $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$.

3: $V^-_{i+1/2} < V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells (i, j) and $(i + 1, m)$ (transmission) or $(i, -j)$ (reflection):
  $\Delta f^{k+1}_{i+1,m} = U_0 f^k_{i,j}$ (transmission) or $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j}$ (reflection).
  and
  $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$.

1. $v_{top} < V^+_{i+1/2}$ and $v_{bot} < V^+_{i+1/2}$: Reflection
   $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j}$
2. $v_{top} > V^+_{i+1/2}$ and $v_{bot} > V^+_{i+1/2}$: Transmission
   (a) v'_{top} and v'_{bot} are at the same cell $(i + 1, m)$:
   $\Delta f^{k+1}_{i+1,m} = U_0 f^k_{i,j}$
   (b) v'_{top} and v'_{bot} are not at same cell:
   i. Divide the space between v_{bot} and v_{top} into N divisions:
   $v_b = v_{bot} : \frac{v_{top} - v_{bot}}{N} : v_{top}$
   ii. Find the final position of v_b:
   $v'_b = \sqrt{v_b^2 + 2(V^-_{i+1/2} - V^+_{i+1/2})}$
   iii. Remap according to remapping rule:
   v'_b are at the cell $(i + 1, m_1)$,..., cell $(i + 1, m_N)$
   $\Delta f^{k+1}_{i+1,m1} = \frac{U_0 f^k_{i,j}}{N}$, ..., $\Delta f^{k+1}_{i+1,m_N} = \frac{U_0 f^k_{i,j}}{N}$
3. $v_{top} > V^+_{i+1/2}$ and $v_{bot} < V^+_{i+1/2}$: partial reflection and partial transmission.
   (a) Compute the rate of the particles to reflect $(v < V^+_{i+1/2})$ and transmit$(v > V^+_{i+1/2})$:
   Reflection rate: $\frac{V^+_{i+1/2} - v_{bot}}{\Delta v}$
   Transmission rate: $\frac{v_{top} - V^+_{i+1/2}}{\Delta v}$
   (b) Reflection: $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j} \frac{V^+_{i+1/2} - v_{bot}}{\Delta v}$
   (c) Transmission: similar to (a) and (b), but $v'_{bot} = 0$

4: $V^-_{i+1/2} > V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells (i, j) and $(i + 1, m)$ (transmission):
  $\Delta f^{k+1}_{i+1,m} = U_0 f^k_{i,j}$ (transmission)
  and
  $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$.
- Repeat step 2(a) (b)

focuses on the final position of the cell. Therefore, in Algorithm 3, the distribution functions, f, are the final ones and we need to know which cells are going to that final position. On the contrary, in Algorithm 4, the distribution functions, f, are the initial ones and we need to know the final position the initial ones are going to.

Here, we calculate the numerical solution using these two different schemes and verify that they give very similar results.

Example 4.4.1. A 1D problem with an exact L^∞-solution. Consider the 1D Vlasov equation

$$f_t + v f_x - V_x f_v = 0, \ (x, v) \in [-1.5, 1.5] \times [-1.5, 1.5] \tag{4.57}$$

with a discontinuous potential barrier given by

$$V(x) = \begin{cases} 0.2, \ x < 0 \\ 0, \ x > 0 \end{cases}$$

The initial data is given by

$$f(x, v, 0) = \begin{cases} 1, \ x \le 0, v > 0, \sqrt{x^2 + v^2} < 1 \\ 1, \ x \ge 0, v < 0, \sqrt{x^2 + v^2} < 1 \\ 0, \text{otherwises}; \end{cases} \tag{4.58}$$

as shown in Fig. 4.18 (a). The exact solution at $t = 1$ is given by [76]

$$f(x, v, t = 1) = \begin{cases} 1,\ x \geq 0,\ v < \sqrt{0.4},\ v > x; \\ 1,\ x \geq 0,\ v < 0,\ x < 1,\ v < \dfrac{x - \sqrt{2 - x^2}}{2}; \\ 1,\ x \leq 0,\ v < x,\ v > -\sqrt{0.6},\ x < \left(1 - \dfrac{\sqrt{0.6 - v^2}}{\sqrt{v^2 + 0.4}}\right) v; \\ 1,\ x \leq 0,\ v > 0, x > -1,\ v < \dfrac{x + \sqrt{2 - x^2}}{2}; \\ 1,\ x \geq 0,\ v > \sqrt{0.4},\ v > x,\ v < \sqrt{1.4},\ x > \left(1 - \dfrac{\sqrt{1.4 - v^2}}{\sqrt{v^2 - 0.4}}\right) v; \\ 0,\ \text{otherwises}; \end{cases} \tag{4.59}$$

as shown in Fig. 4.18 (b).

The numerical solutions computed with a different mesh using Algorithm 3 and Algorithm 4 are shown in Figure 4.19. They show good agreement with the exact solution shown in Fig. 4.18. Table 4.9 compares the L^1-norm of the numerical solutions by using scheme I (Algorithm 3) and scheme II (Algorithm 4) with different meshes. In the table, the error is the difference between the analytic and numerical solutions.

From the table, the convergence order of the numerical solutions appears to be around 0.5 for both schemes. These results agree with well-established theory [80]. In [76, 80, 81], the L^1-norm for a finite difference scheme for a discontinuous solution of a linear equation is at most half-order. We have found that scheme I and scheme II have very similar results. In the next section, we will expand scheme II to include quantum effects and to a higher order.

4.4.2 High Order Scheme in Quantum Mechanics

In this section we apply the AHOCS to the region with the continuous potential in quantum mechanics. Then, we compare the solution which be obtained by first order

$N_x \times N_v$	Scheme I	Order	Scheme II	Order
50×50	5.66×10^{-2}	—	5.64×10^{-2}	—
100×100	3.99×10^{-2}	0.50	3.99×10^{-2}	0.50
200×200	2.86×10^{-2}	0.48	2.86×10^{-3}	0.48

Table 4.9: Vlasov equation with discontinuous potential (classical mechanics). The table reports the L^1-norm of the error (difference of analytic and numerical solutions) at the final time ($T = 1$), for progressively larger numbers of cells N_x and N_v. The algebraic order of convergence ('Order' column) is calculated as the base-2 logarithm of two successive error norms. Scheme I is using Algorithm 3 and scheme II is using Algorithm 4.

CS to the solution gotten by AHOCS.

In classical mechanics, a particle will either be reflected ($R = 1$) or cross the potential barrier ($R = 0$). If a particle has enough energy to overcome the potential barrier, then this particle is going to cross the potential barrier with changed momentum. During this process, the Hamiltonian $H = \frac{1}{2}p^2 + V$ should be preserved across the potential barrier. On the contrary, if a particle does not have enough energy to overcome the height of the potential barrier, then this particle is going to reflect with the same magnitude of momentum, but changed direction.

However, in quantum mechanics, a particle actually behaving as a matter wave has a finite probability that it will penetrate the barrier and continue its travel as a wave on the other side. Therefore, unlike in classical mechanics where the particles are either reflected or transmitted, in quantum mechanics the particles could be partially transmitted and partially reflected. Because of this, the algorithm of the CS for the discontinuous potential needs to consider the reflection and transmission at the same time. The detailed algorithm for the first order scheme is described in Algorithm 5.

In Section 4.1, we mentioned that we can correct the position of the final cell, so the scheme can achieve arbitrarily high order. We call this high order scheme an f_{22} scheme. Now, we want to use an f_{22} scheme on the case with discontinuous

Algorithm 5 The detailed algorithm of CS (first order) on the discontinuous potential (quantum mechanics). $v_j > 0$ is the initial velocity, $v_{top(bot)} = v_j \pm \Delta v/2$ and $v'_{top(bot)} = \sqrt{v^2_{top(bot)} + 2(V^-_{i+1/2} - V^+_{i+1/2})}$

1: Compute the Courant number: $U_0 = v_j \Delta t/\Delta x$
2: $V^-_{i+1/2} = V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells from (i, j) to $(i + 1, j)$:
  $\Delta f^{k+1}_{i+1,j} = U_0 f^k_{i,j}$ and $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$.

3: $V^-_{i+1/2} < V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells (i, j) and $(i + 1, m)$ (transmission) or $(i, -j)$ (reflection):
  $\Delta f^{k+1}_{i+1,m} = U_0 f^k_{i,j}$ (transmission) or $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j}$ (reflection).
  and
  $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$.

1. $v_{top} < V^+_{i+1/2}$ and $v_{bot} < V^+_{i+1/2}$: Reflection
   $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j}$
2. $v_{top} > V^+_{i+1/2}$ and $v_{bot} > V^+_{i+1/2}$: Partial reflection and partial transmission
   $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$
   (a) v'_{top} and v'_{bot} are at the same cell $(i + 1, m)$:
   $\Delta f^{k+1}_{i+1,m} = T(v_j) U_0 f^k_{i,j}$ and $\Delta f^{k+1}_{i,-j} = R(v_j) U_0 f^k_{i,j}$
   (b) v'_{top} and v'_{bot} are not at same cell:
   i. Divide the space between v_{bot} and v_{top} into N divisions:
   $v_b = v_{bot} : \frac{v_{top} - v_{bot}}{N} : v_{top}$
   ii. Find the final position of v_b:
   $v'_b = \sqrt{v^2_b + 2(V^-_{i+1/2} - V^+_{i+1/2})}$
   iii. Remap according to remapping rule:
   v'_b are at the cell $(i + 1, m_1)$,..., cell $(i + 1, m_N)$
   $\Delta f^{k+1}_{i+1,m1} = \frac{T(v_j) U_0 f^k_{i,j}}{N}$, ..., $\Delta f^{k+1}_{i+1,m_N} = \frac{T(v_j) U_0 f^k_{i,j}}{N}$
   $\Delta f^{k+1}_{i,-j} = \frac{R(v_j) U_0 f^k_{i,j}}{N}$, ..., $\Delta f^{k+1}_{i,-j} = \frac{R(v_j) U_0 f^k_{i,j}}{N}$
3. $v_{top} > V^+_{i+1/2}$ and $v_{bot} < V^+_{i+1/2}$
   (a) Compute the rate of the particles to reflect ($v < V^+_{i+1/2}$) and transmit($v > V^+_{i+1/2}$):
   Reflection rate: $\frac{V^+_{i+1/2} - v_{bot}}{\Delta v}$
   Transmission rate: $\frac{v_{top} - V^+_{i+1/2}}{\Delta v}$
   (b) Reflection:
   $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j} \frac{V^+_{i+1/2} - v_{bot}}{\Delta v}$
   (c) Transmission: similar to 2(a) and 2(b), but $v'_{bot} = 0$

4: $V^-_{i+1/2} > V^+_{i+1/2}$: $\Delta f^{k+1}_{i,j} = (1 - U_0) f^k_{i,j}$

1. Density remap according to the fractions. Cell (i, j) remaps into cells (i, j) and $(i + 1, m)$ (transmission) or $(i, -j)$ (reflection):
   $\Delta f^{k+1}_{i+1,m} = U_0 f^k_{i,j}$ (transmission) or $\Delta f^{k+1}_{i,-j} = U_0 f^k_{i,j}$ (reflection).
2. Repeat 2(a) and 2(b)

potential. However, the f_{22} scheme can only be used on the continuous part. For the discontinuous boundary, we still use first order MF scheme. We hope that we can use the higher order part in the large region to partially compensate for the low order part in the small region.

In order to use a higher order scheme, it is necessary to use $[U f^k_{i,j}]$ from Eq. (4.20) to replace the $[U_0 f^k_{i,j}]$ in Algorithm 5. The detailed algorithm for the CS (high order) with a discontinuous potential is described in Algorithm 6.

4.4.2.1 Reflection and Transmission

As mentioned before, it is necessary to calculate of the reflection coefficient and transmission coefficient.

For a time-independent potential, the wavefunction can be factorized as $\Psi(x, t) = e^{-iEt/\hbar}\psi$, where $\psi(x)$ is obtained from the Schrödinger equation,

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V(x)\right]\psi(x) = E\psi(x), \tag{4.60}$$

and where E denotes the energy of the particles. Consider the influence of a potential step on the propagation of a beam of particles. A beam of particles with kinetic energy, $E = p^2/2m$, moving from left to right is incident on a potential step of height V_0 at position $x = 0$. If the beam has unit amplitude, the reflected and transmitted

Algorithm 6 The detailed algorithm of CS (High order) on the discontinuous potential (quantum mechanics). $v_j(v_j > 0)$ is the initial velocity, $v_{top(bot)} = v_j \pm \Delta v/2$ and $v'_{top(bot)} = \sqrt{v^2_{top(bot)} + 2(V^-_{i+1/2} - V^+_{i+1/2})}$

1: Given a higher order correction, $[Uf]^k_{i,j} = \min(\max(0, \Gamma_i), n^k_i)$, from Eq. (4.20)

2: $V^-_{i+1/2} = V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells from (i, j) to $(i + 1, j)$:
  $\Delta f^{k+1}_{i+1,j} = [Uf]^k_{i,j}$ and $\Delta f^{k+1}_{i,j} = f^k_{i,j} - [Uf]^k_{i,j}$.

3: $V^-_{i+1/2} < V^+_{i+1/2}$:

- Density remap according to the fractions. Cell (i, j) remaps into cells (i, j) and $(i + 1, m)$ (transmission) or $(i, -j)$ (reflection):
  $\Delta f^{k+1}_{i+1,m} = [Uf]^k_{i,j}$ (transmission) or $\Delta f^{k+1}_{i,-j} = [Uf]^k_{i,j}$ (reflection).
  and
  $\Delta f^{k+1}_{i,j} = f^k_{i,j} - [Uf]^k_{i,j}$.

1. $v_{top} < V^+_{i+1/2}$ and $v_{bot} < V^+_{i+1/2}$: Reflection
   $\Delta f^{k+1}_{i,-j} = [Uf]^k_{i,j}$
2. $v_{top} > V^+_{i+1/2}$ and $v_{bot} > V^+_{i+1/2}$: Partial reflection and partial transmission
   $\Delta f^{k+1}_{i,j} = f^k_{i,j} - [Uf]^k_{i,j}$
   (a) v'_{top} and v'_{bot} are at the same cell $(i + 1, m)$:
   $\Delta f^{k+1}_{i+1,m} = T(v_j)[Uf]^k_{i,j}$ and $\Delta f^{k+1}_{i,-j} = R(v_j)[Uf]^k_{i,j}$
   (b) v'_{top} and v'_{bot} are not at same cell:
   i. Divide the space between v_{bot} and v_{top} into N divisions:
   $v_b = v_{bot} : \frac{v_{top} - v_{bot}}{N} : v_{top}$
   ii. Find the final position of v_b:
   $v'_b = \sqrt{v^2_b + 2(V^-_{i+1/2} - V^+_{i+1/2})}$
   iii. Remap according to remapping rule:
   v'_b are at the cell $(i + 1, m_1)$,..., cell $(i + 1, m_N)$
   $\Delta f^{k+1}_{i+1,m1} = \frac{T(v_j)[Uf]^k_{i,j}}{N}$, ..., $\Delta f^{k+1}_{i+1,m_N} = \frac{T(v_j)[Uf]^k_{i,j}}{N}$
   $\Delta f^{k+1}_{i,-j} = \frac{R(v_j)[Uf]^k_{i,j}}{N}$, ..., $\Delta f^{k+1}_{i,-j} = \frac{R(v_j)[Uf]^k_{i,j}}{N}$
3. $v_{top} > V^+_{i+1/2}$ and $v_{bot} < V^+_{i+1/2}$
   (a) Compute the rate of the particles to reflect ($v < V^+_{i+1/2}$) and transmit($v > V^+_{i+1/2}$):
   Reflection rate: $\frac{V^+_{i+1/2} - v_{bot}}{\Delta v}$
   Transmission rate: $\frac{v_{top} - V^+_{i+1/2}}{\Delta v}$
   (b) Reflection:
   $\Delta f^{k+1}_{i,-j} = [Uf]^k_{i,j} \frac{V^+_{i+1/2} - v_{bot}}{\Delta v}$
   (c) Transmission: similar to 2(a) and 2(b), but $v'_{bot} = 0$

4: $V^{-}_{i+1/2} > V^{+}_{i+1/2}$:
5: $\Delta f^{k+1}_{i,j} = f^{k}_{i,j} - [Uf]^{k}_{i,j}$

1. Density remap according to the fractions. Cell (i,j) remaps into cells (i,j) and $(i+1,m)$ (transmission) or $(i,-j)$ (reflection):
   $\Delta f^{k+1}_{i+1,m} = [Uf]^{k}_{i,j}$ (transmission) or $\Delta f^{k+1}_{i,-j} = [Uf]^{k}_{i,j}$ (reflection).
2. Repeat 2(a) and 2(b)

amplitudes are set by $\hat{R}$ and $\hat{T}$. The corresponding wavefunction is given by

$$\begin{cases} \psi(x) = e^{i\kappa_1 x/\hbar} + re^{-i\kappa_1 x/\hbar}, \ x < 0 \\ \psi(x) = te^{i\kappa_1 x/\hbar}, \ x > 0 \end{cases}$$

where $\kappa_1 = \sqrt{p^2}$ and $\kappa_2 = \sqrt{p^2 - 2mV_0}$.

Assuming $E > V_0$ and applying the continuity conditions ψ and $d\psi/dx$ at the step $(x = 0)$, one obtains the relations $1 + \hat{R} = \hat{T}$ and $i\kappa_1(1 - \hat{R}) = i\kappa_2\hat{T}$ leading to the reflection and transmission amplitudes

$$\hat{R} = \frac{\kappa_1 - \kappa_2}{\kappa_1 + \kappa_2}, \ \hat{T} = \frac{2\kappa_1}{\kappa_1 + \kappa_2}, \tag{4.61}$$

The reflection, R, and transmission, T, are defined as

$$R = |\hat{R}^2| = \left(\frac{\kappa_1 - \kappa_2}{\kappa_1 + \kappa_2}\right)^2, \ T = (\kappa_2/\kappa_1)|\hat{T}^2| = \frac{4\kappa_1\kappa_2}{(\kappa_1 + \kappa_2)^2}. \tag{4.62}$$

For $E < V_0$, $\kappa_2 = \sqrt{p^2 - 2mV_0} = ik_2$ becomes pure imaginary. The wavefunction, $\psi(x) = te^{i\kappa_2 x/\hbar} = te^{k_2 x/\hbar}$, decays evanescently. Because κ_2 is pure imaginary, the reflection (R) is described by Eq. (4.63) and the beam is completely reflected from the barrier.

$$R = |r^2| = \left(\frac{\kappa_1 - ik_2}{\kappa_1 + ik_2}\right)\left(\frac{\kappa_1 + ik_2}{\kappa_1 - ik_2}\right) = 1, \ T = 0. \tag{4.63}$$

Similarly, the reflection and transmission of a beam of particles moving from left to right can be described by the symmetric model, and we obtain

$$R = |r^2| = \left(\frac{\kappa_2 - \kappa_1}{\kappa_1 + \kappa_2}\right)^2,\ T = (\kappa_2/\kappa_1)|t^2| = \frac{4\kappa_1\kappa_2}{(\kappa_1 + \kappa_2)^2}. \tag{4.64}$$

where $\kappa_1 = \sqrt{p^2}$ and $\kappa_2 = \sqrt{p^2 - 2mV_0}$.

Here, the reflection and transmission by the step potential were described in terms of velocity. From classical mechanics, the Hamiltonian remains constant across the potential barrier. Therefore, if a beam is moving from right to left or from left to right, and the potential is like that in Fig. 4.20, then the relationship between the incident velocity v_i and the transmitted velocity v_t can be obtained according to the Hamiltonian-preserving schemes as follows:

$$\begin{cases} v_i^2 + 0 = v_t^2 + 2mV_0;\ \text{a beam move from right to left} \\ v_i^2 + 2mV_0 = v_t^2 + 0;\ \text{a beam move from left to right} \end{cases} \tag{4.65}$$

As mentioned above, $\kappa_1 = \sqrt{p^2}$ and $\kappa_2 = \sqrt{p^2 - 2mV_0}$. Now, if a beam is moving from right to left, then p in κ_1 and κ_2 are incident momenta. As for a beam which is moving from left to right, p in κ_1 and κ_2 are transmitted momenta. Finally, the reflection and transmission coefficients of a beam can be expressed in terms of these momenta (or velocities):

$$R = \left(\frac{v_i - v_t}{v_i + v_t}\right)^2,\ T = \frac{4v_iv_t}{(v_i + v_t)^2}. \tag{4.66}$$

4.4.2.2 Numerical tests

The purpose of this section is to verify the AHOCS for solving the Vlasov equation with a discontinuous potential in quantum mechanics.

Example 4.4.1. Vlasov equation with step function:

$$f_t + vf_x - V_x f_v = 0,\ (x,v) \in [-1,1] \times [-3,3] \tag{4.67}$$

with a discontinuous potential barrier given by

$$V(x) = \begin{cases} 0,\ x < 0 \\ 0.5,\ x > 0 \end{cases}$$

The initial data is given by

$$f(x,p,0) = \frac{1}{2\pi\sigma_x\sigma_p} \exp(-\frac{(x-x_0)^2}{2\sigma_x^2}) \exp(-\frac{(p-p_0)^2}{2\sigma_p^2}) \tag{4.68}$$

with $\sigma_x = \sigma_p = 0.05$, $x_0 = -0.5$, and $p_0 = 1$. The exact solution is given as follows letting $\Omega(t) = \{(x,p) | x < 0 \text{ and } x - pt < 0,\ \text{or} x > 0 \text{ and } x - pt > 0\}$ [76].

$$f(x,p,t) = \begin{cases} f(x-pt,p,0);\ (x,p) \in \Omega(t) \\ T(p) \cdot f(\frac{q}{p}x - qt, q, 0) + R(p) \cdot f(-x+pt, -p, 0);\ \text{otherwise} \end{cases} \tag{4.69}$$

Here p is the transmitted velocity and q is the incident velocity. In Section 4.4.2.1, we derived the reflection and transmission coefficients from [76] in terms of incident velocity and transmitted velocity in Eq. (4.66) and Fig. 4.21.

As we mentioned in Section 4.4.1, Jin *et al.* [76] focuses on the final position of the cell. Therefore, T and R in Eq. (4.69) are both functions of the transmitted velocity, p. Therefore, we need to write q as a function of p [76].

From Eq. (4.66) and Eq. (4.70), the reflection coefficient is given by

$$\begin{cases} q = \sqrt{p^2+1}, \text{ for } p > 0, \\ q = -\sqrt{p^2-1}, \text{ for } p \leq 0. \end{cases} \tag{4.70}$$

From Eq. (4.66) and Eq. (4.70), the reflection coefficient is given by

$$R(p) = \begin{cases} \left(\dfrac{\sqrt{p^2+1}-p}{\sqrt{p^2+1}+p}\right)^2 & p > 0, \\ \left(\dfrac{-\sqrt{p^2-1}-p}{-\sqrt{p^2-1}+p}\right)^2 & p \leq 0. \end{cases} \tag{4.71}$$

Note that when $p \in [-1, 0]$, q is imaginary and $R = 1$. Finally, we can get T according to the equation $T = 1 - R$ from Eq. (4.71).

Once we get $R(p)$ and $T(p)$, we can put them into Eq. (4.69) to obtain the exact solution, in Fig. 4.22.

However, in Algorithm 5 and in 6, we need to track the initial cells and see where they are. Therefore, T and R need to be functions of the incident velocity, q. Consequently, we also need to write p as a function of q.

$$\begin{cases} p = \sqrt{q^2-1}, \text{ for } q > 0, \\ p = -\sqrt{q^2+1}, \text{ for } q \leq 0. \end{cases} \tag{4.72}$$

Again, from Eq. (4.66) and Eq. (4.72), the reflection coefficient is given by

$$R(p) = \begin{cases} \left(\dfrac{q-\sqrt{q^2-1}}{q+\sqrt{q^2-1}}\right)^2 & q > 0, \\ \left(\dfrac{q+\sqrt{q^2+1}}{q-\sqrt{q^2+1}}\right)^2 & q \leq 0. \end{cases} \tag{4.73}$$

Note that when $q \in [0,1]$, p is imaginary and $R = 1$. $R(q)$ and $T(q)$ are shown in Fig. 4.23.

The final solutions obtained with the low and high order CS scheme are compared with the exact solution in Fig. 4.24. From the Figure, it is obvious that the high order scheme provides better agreement with the exact solution than the low order scheme. Table 4.10 shows the L^1-norm. We find that if we did not do any modification of the CS, it is a first order scheme. Therefore, the result of the low order scheme is limited by the discontinuous potential and the order is less than one. On the other hand, when we apply the high order CS to the continuous potential region, although the discontinuous barrier is still only handled to first order, the other high order part can compensate for the first order part to improve the overall order. Therefore, the order can surpass the estimated theory which applies to the finite difference method. In addition, we see that the error for the high order scheme is also more than 10 times less than the error for the low order scheme.

$N_x \times N_v$	Low order	Order	High order	Order
200×200	4.11×10^{-1}	—	6.56×10^{-2}	—
400×400	2.67×10^{-1}	0.62	2.04×10^{-2}	1.68
800×800	1.52×10^{-1}	0.81	9.95×10^{-3}	1.04
1600×1600	8.55×10^{-2}	0.83	3.83×10^{-3}	1.38

Table 4.10: Vlasov equation with discontinuous potential (quantum mechanics) for the Example 4.4.1. The table reports the L^1-norm of the error (difference of analytic and numerical solutions) at the final time ($T = 1$), for progressively larger numbers of cells N_x and N_v. The algebraic order of convergence ('Order' column) is calculated as the base-2 logarithm of two successive error norms. The low order scheme uses Algorithm 5 and the high order scheme uses Algorithm 6.

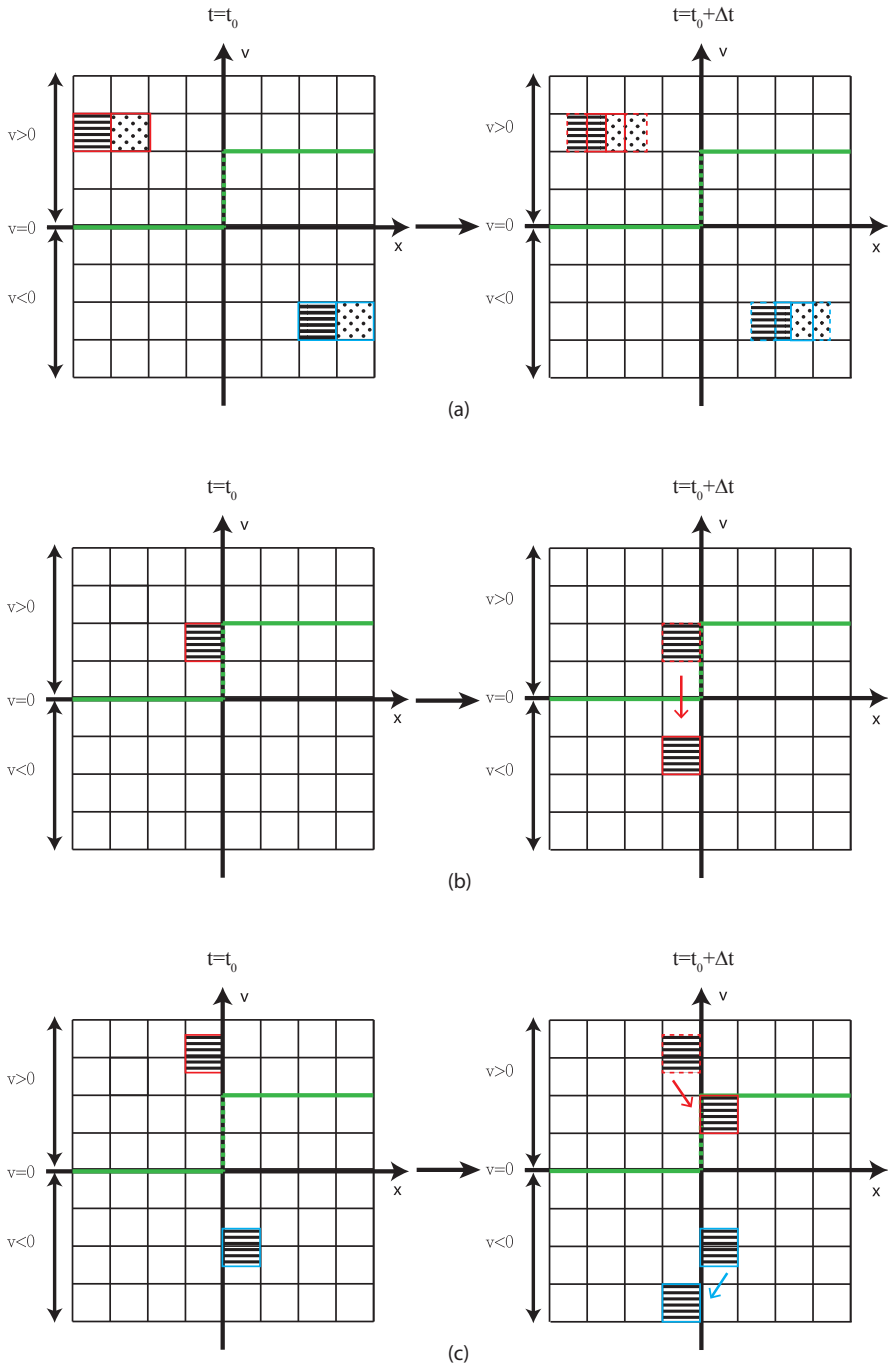

Figure 4.11: Illustration of how the method of characteristics (MOC) is applied at the boundary; In [76], the MOC looks back to several initial cells.(a) For continuous potential, velocity does not change during the advection step. The distribution function $f_{i,j}$ at $t = t_0+\Delta t$ is from $f_{i,j}$ and $f_{i-1,j}$ or from $f_{i,j}$ and $f_{i+1,j}$ at $t = t_0$. For discontinuous potential, a boundary condition is used to change velocity at the boundary, $x = x_d$. (b)Reflection: the distribution function $f_{i,j}$ at $t = t_0 + \Delta t$ is from $f_{-i,j}$ at $t = t_0$, (c) Transmission: the final distribution function $f_{i,j}$ at $t = t_0 + \Delta t$ is from $f_{i-1,m}$ or from $f_{i+1,m}$ at $t = t_0$.

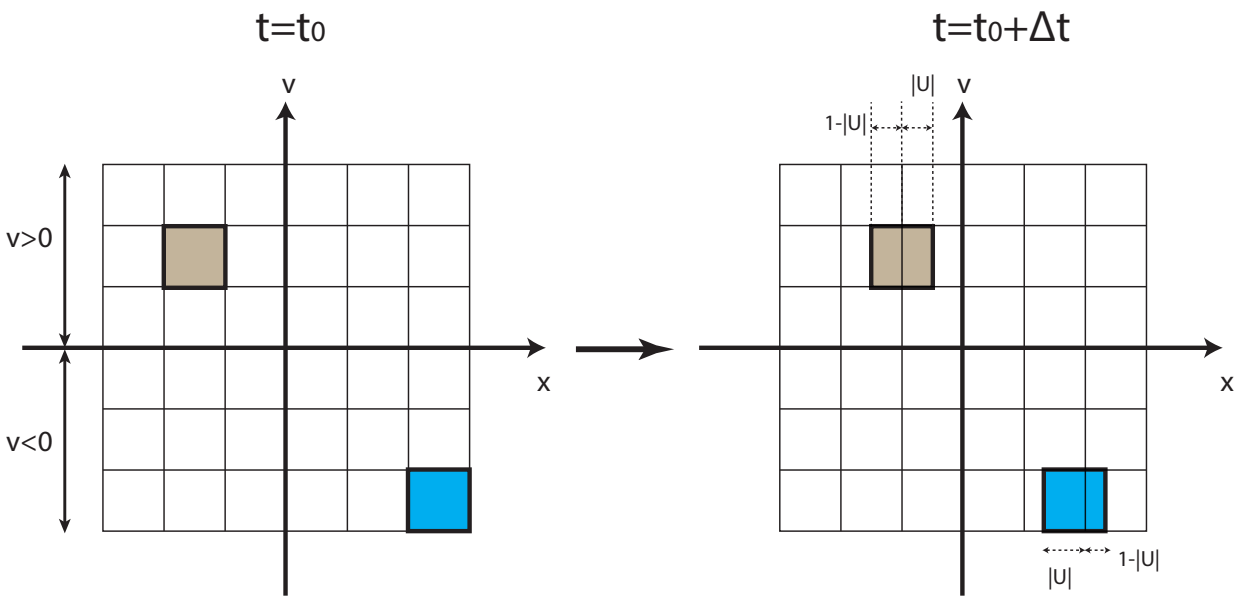

Figure 4.12: Ballistic operator of the moving cell for both $v > 0$ and $v < 0$. The left frame shows the initial positions of two MCs. The right frame shows the densities reallocated to cells, according to the remapping rule.

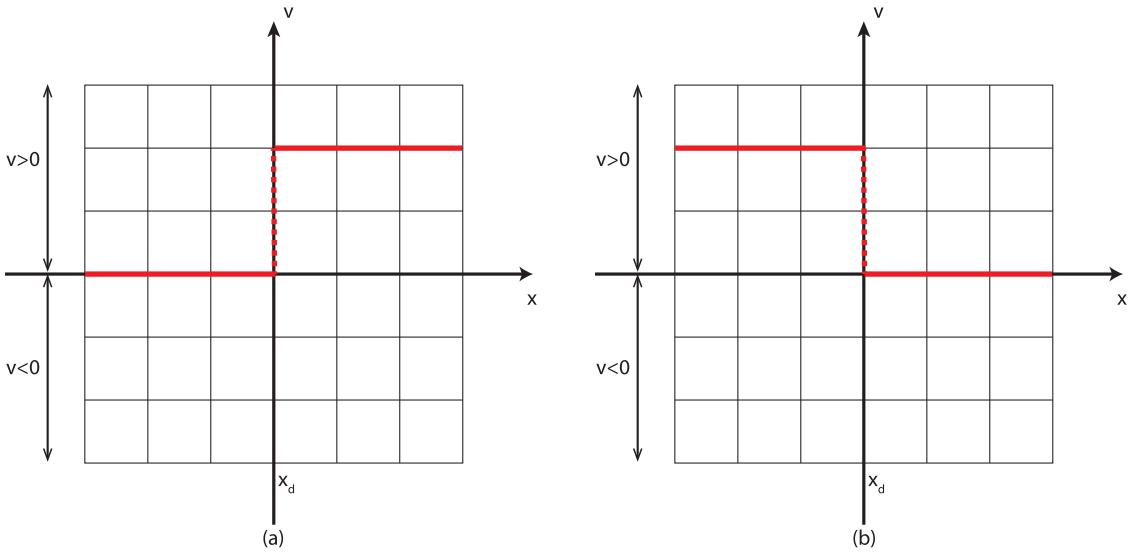

Figure 4.13: Discontinuous potential barrier. (a) $V^- < V^+$, (b) $V^- > V^+$. Here, the superscripts '+' and '-' mean the right and left limit of the barrier, x_d.

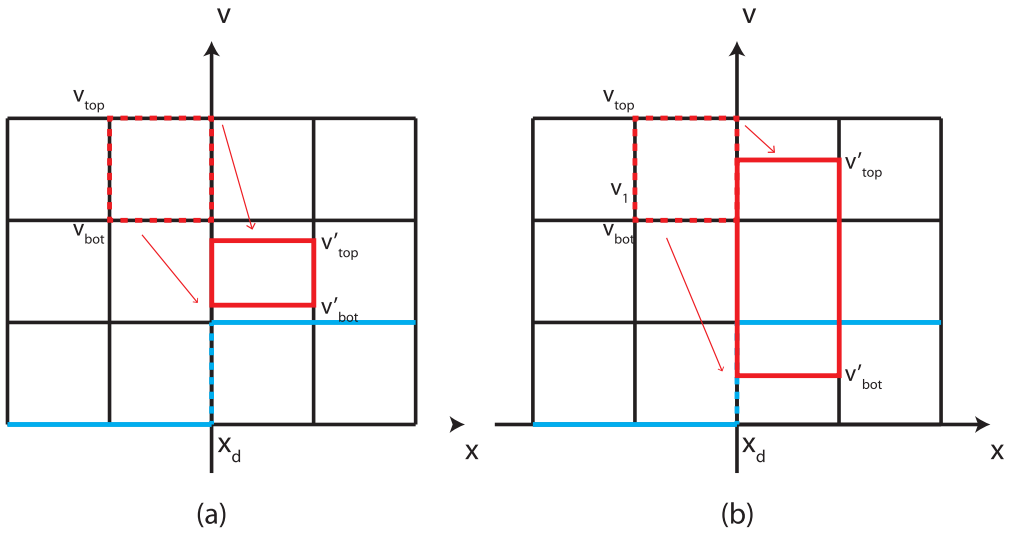

Figure 4.14: Ballistic operator of the moving cell across the discontinuous potential barrier ($V^- < V^+$). (a) both the velocity faces of the initial cell go to the same cell, (b) the velocity faces of the initial cell go to different cells. The superscripts '+' and '-' mean the right and left limit of the barrier, x_d.

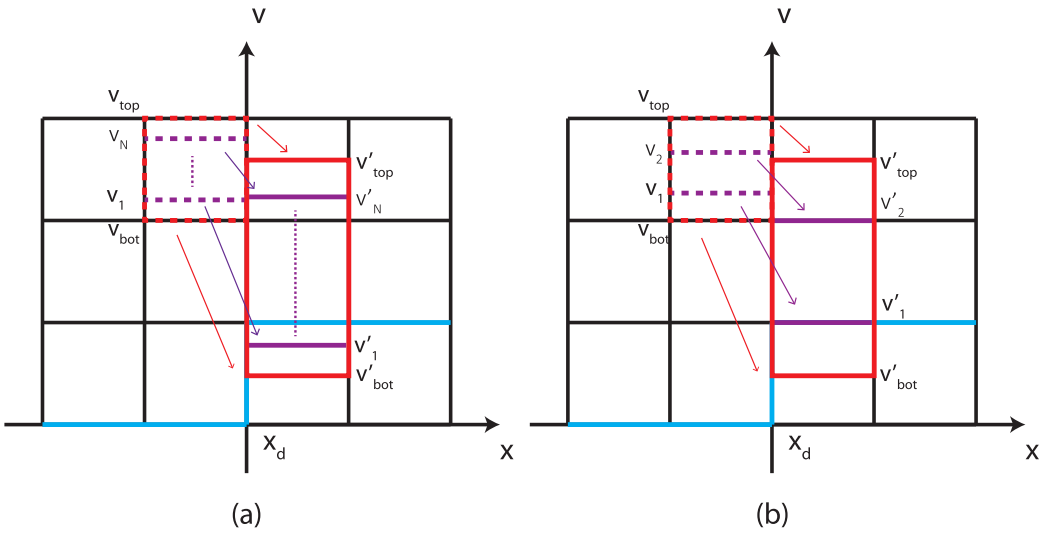

Figure 4.15: v'_{top} and v'_{bot} are the transmitted velocities corresponding to v_{top} and v_{bot}. If v'_{top} and v'_{bot} are not in the same cell, there are two methods to deal with the remapping rule: (a) Method 1, the initial cell is divided into many smaller cells which are remapped separately;(b) Method 2, velocities (such as v_1 and v_2) are found within the initial cell which map onto velocity cell boundaries at the final location. The initial cell is split, at v_1 and v_2, into new cells which are remapped separately. The superscripts '+' and '-' mean the right and left limit of the barrier, x_d.

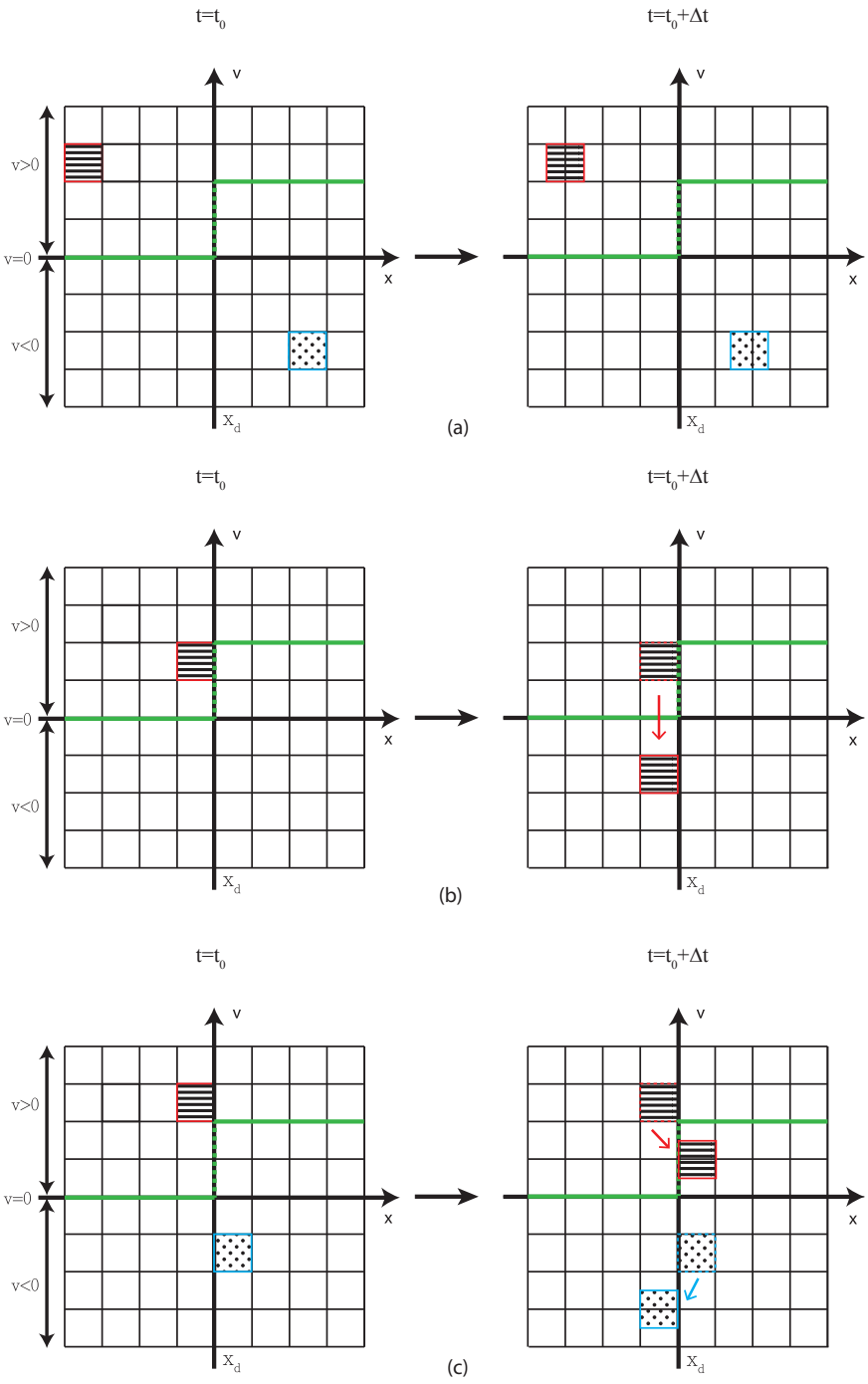

Figure 4.16: Illustration of how the method of characteristics (MOC) is applied at the boundary; a MOC such as the CS looks forward from an initial cell to several final cells.(a) For the continuous potential, the initial distribution function $f_{i,j}$ goes to (i, j) and $(i + 1, j)$ or to (i, j) and $(i - 1, j)$. For the discontinuous potential the boundary is at $x = x_d$, (b) Reflection: the initial distribution function $f_{i,j}$ goes to $(i, -j)$, (c) Transmission: the initial distribution function $f_{i,j}$ goes to $(i + 1, m)$ or to $(i - 1, m)$.

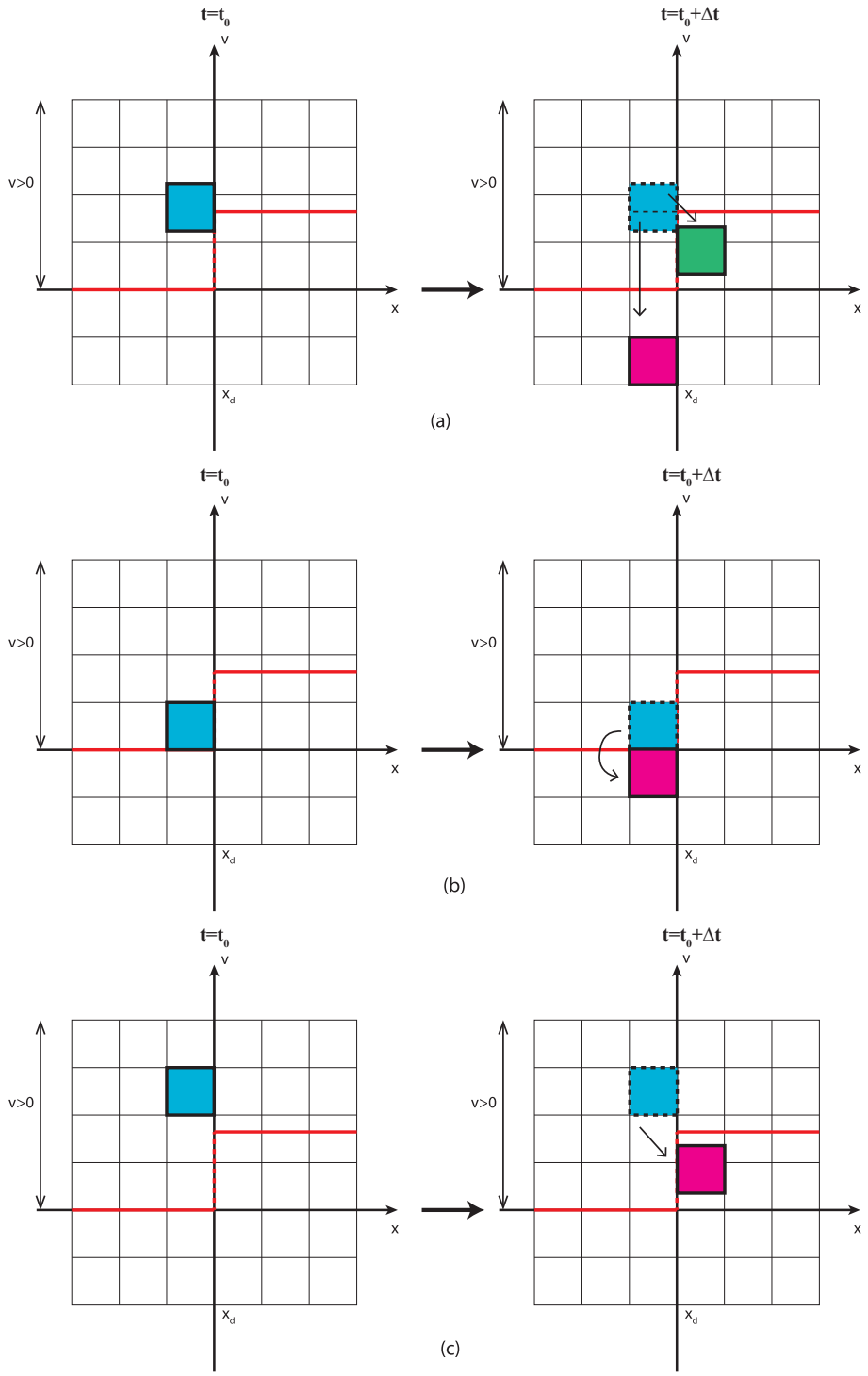

Figure 4.17: Illustration of how the method of characteristics (MOC) is applied at the boundary; a MOC such as the CS looks forward from an initial cell to several final cells. (a) some particles can transmit across the barrier, while some particles are reflected, (b) particles do not have enough energy to cross the barrier, they are reflected, (c) particles have enough energy to cross the barrier, they are transmitted.

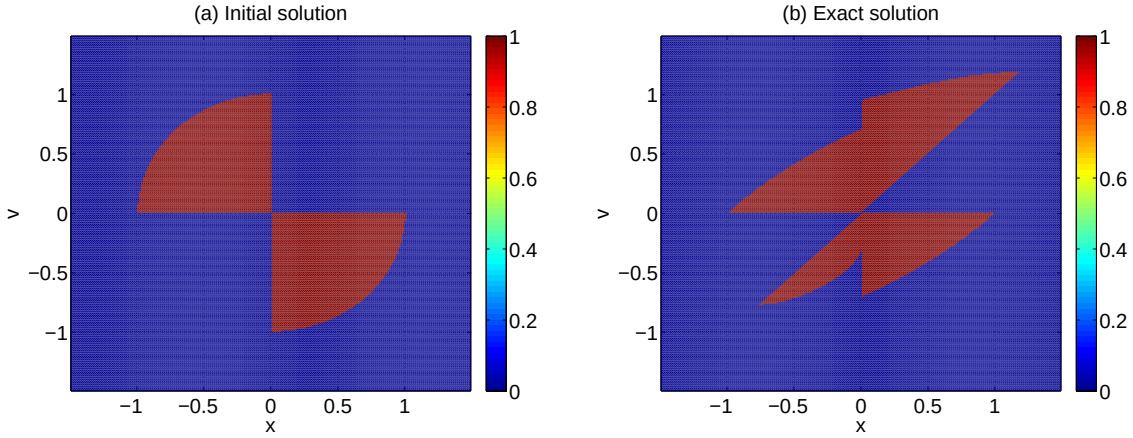

Figure 4.18: Vlasov equation with discontinuous potential (classical mechanics). The figure shows the (a) initial condition, and (b) exact solution of the density distribution for $N_x = N_v = 200$

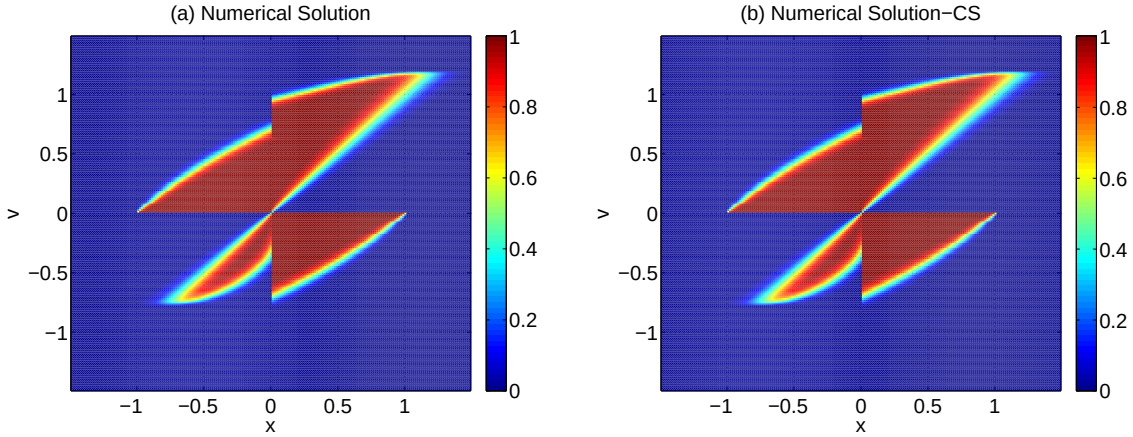

Figure 4.19: Vlasov equation with discontinuous potential (classical mechanics). The numerical solution for two algorithms for the Example 4.4.1. (a) Algorithm 3, and (b) Algorithm 4, for the numerical example in section 4.4.1.1. $N_x = N_v = 200$

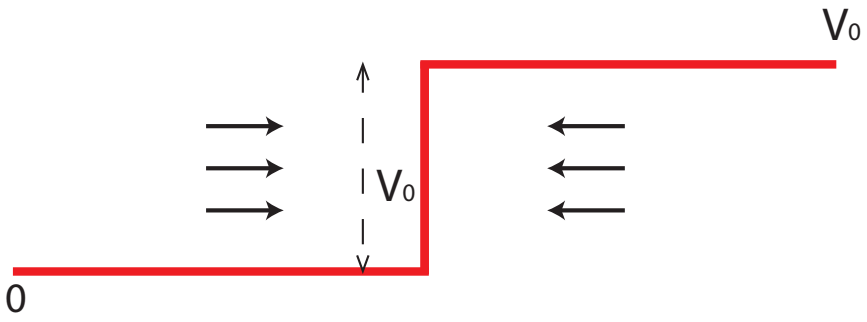

Figure 4.20: A schematic representation of a potential step, particles are moving from right to left and from left to right.

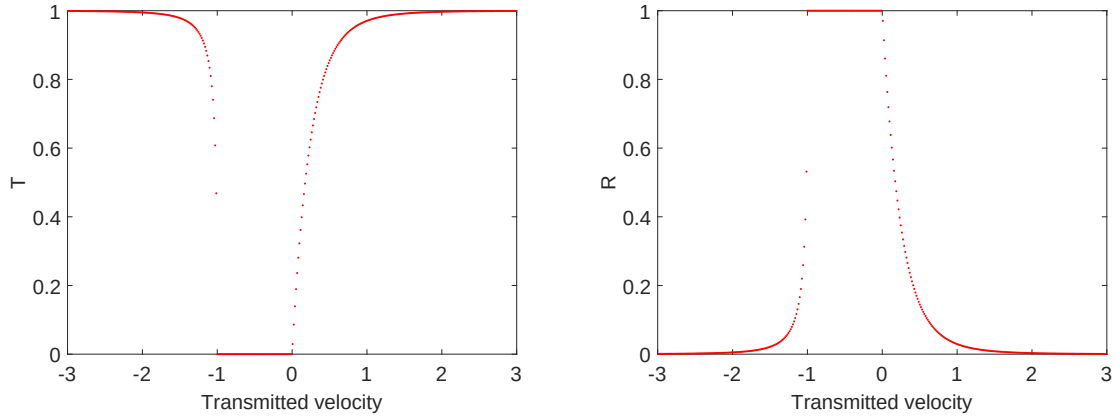

Figure 4.21: The reflection coefficient (R) and transmission coefficient (T) for the Example 4.4.1. R and T are functions of transmitted velocity, p.

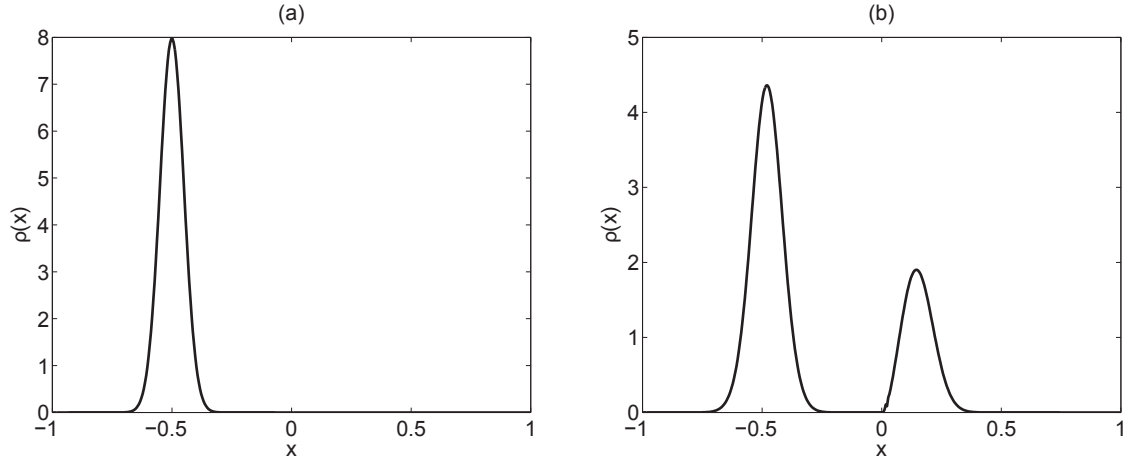

Figure 4.22: Vlasov equation with discontinuous potential (quantum mechanics). The figure shows the (a) initial condition, and (b) exact solution of the position density.

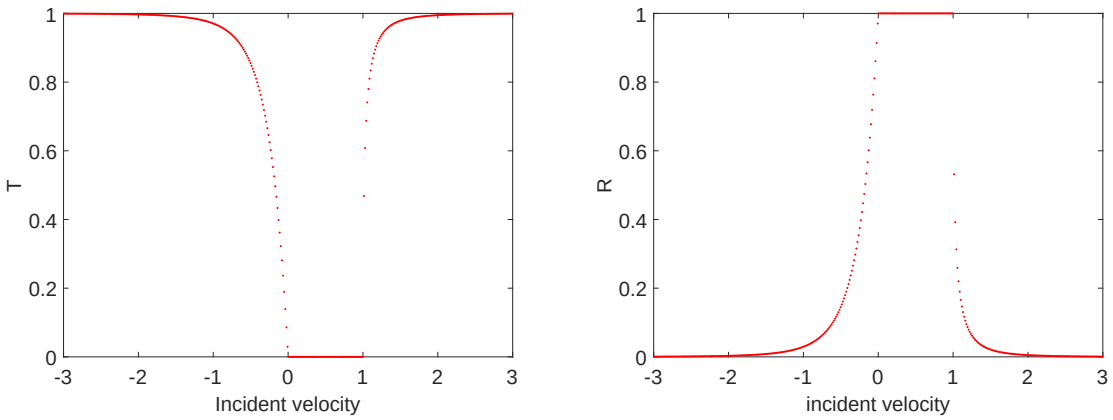

Figure 4.23: The reflection coefficient (R) and transmission coefficient (T) for the Example 4.4.1. R and T are functions of incident velocity, q

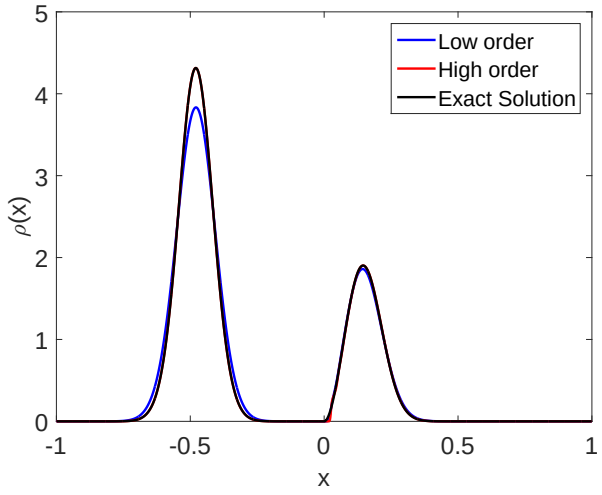

Figure 4.24: Vlasov equation with discontinuous potential (quantum mechanics). The final solutions obtained with the low and high order CS scheme are compared with the exact solution.

4.5 Summary

In this chapter, we mainly discussed the arbitrarily high order CS (f_{22} scheme) as developed for the constant advection equation and its applications. The f_{22} scheme has many desirable properties: it is efficient and conservative, it preserves positivity and has low phase-error. When we use the f_{22} scheme to solve the constant advection equation, machine precision can be reached very quickly.

We used this scheme on the Wigner equation and the Vlasov equation with a discontinuous potential, where we also got good performance. For the Wigner equation, the results show better efficiency because the method did not require the "mesh strategy", $\Delta t = O(\varepsilon)$ and $\Delta x = O(\varepsilon)$. This means that we can use a bigger mesh to get the same results as the reference paper.

As for the Vlasov equation with discontinuous potential, we used the f_{22} scheme for updating the density for the region with continuous potential and used the first order MF scheme to deal with the discontinuous potential region. Although it is only first order at the boundary, the high order can compensate to some extent for this low order. Accordingly, we can get higher order, compared to the theoretical limit and smaller error for the same number of subdivisions.

Chapter 5

Conclusion

In the past the Convected scheme had been mainly employed to solve the continuity equation and the Boltzmann equation and further applied in many fields, such as plasma and semiconductors. The CS provides exact automatic conservation and positivity preservation, and it is not restricted by the CFL limit. In this work, we improved the CS to obtain higher order accuracy, in the case where the advection velocity is variable. We then applied the CS to solve the plasma breakdown problem which is an important and difficult engineering task. Further, the CS had been extended to solve other equations, such as the Wigner equation, which has widely been used in modeling semiconductor nano-structures, and the Vlasov equation with a discontinuous potential which can be applied in geometrical optics. The new versions of the CS allow us to solve the appropriate equations and obtain good results.

The CS is used to solve by envisaging a moving-cell (MC) and the moving cell is of two different types:

1) one can move the midpoint of the cell (Moving Midpoint, MM), with each face having the same velocity, or

2) one can move the faces of the cell (Moving Face, MF), the two spatial faces of a

cell being moved independently.

Here, our first contribution was to develop an efficient MF scheme with high order accuracy. For this newly developed MF scheme, there is no CFL limit. This means that we can simulate with a bigger Courant number ($C \gtrsim 1$) to increase efficiency. Moreover, when we use the high order MF scheme, it gives third order accuracy in the advection equation. Further, when the high order MF scheme with high order splitting is used, third or higher order accuracy can be obtained in breakdown simulations.

Next, an arbitrarily high order f_{22}-scheme was employed for solving the constant advection equation. The Wigner equation can be split into a constant advection equation and a pseudo-differential operator. In this thesis, we take advantage of this high order scheme to solve the Wigner equation with FFT/IFFT. Because the f_{22}-scheme can reach machine precision rapidly, when we employ this scheme to solve the Wigner equation for quantum plasma and in a semi-classical region (with a scale for which Planck's constant ε is small), we can also get good performance. We were able to use a bigger mesh to obtain the same error scale as the smaller mesh, and thereby achieve high efficiency.

Later, the f_{22}-scheme was used on the Vlasov equation with a discontinuous potential. In the past, the convergence rate, using the finite difference method to solve with a discontinuous potential was limited to 0.5. Here, we used the f_{22}-scheme on the continuous part and the MF scheme for the discontinuous part to solve the discontinuous potential case. The convergence rate is higher than the theoretical limit of 0.5 applicable to solving the advection equation with the finite difference method, and the error for the high order scheme is also more than 10 times less than the error for the low order finite difference method.

Bibliography

[1] S Kanazawa, M Kogoma, T Moriwaki, and S Okazaki. Stable glow plasma at atmospheric pressure. *Journal of Physics D: Applied Physics*, 21(3):838–840, 1988.

[2] Françoise Massines, Ahmed Rabehi, Philippe Decomps, Rami Ben Gadri, Pierre Ségur, and Christian Mayoux. Experimental and theoretical study of a glow discharge at atmospheric pressure controlled by dielectric barrier. *Journal of Applied Physics*, 83(6):2950–2957, 1998.

[3] S Manolache, E B Somers, A C L Wong, V Shamamian, and F Denes. Dense Medium Plasma Environments: A New Approach for the Disinfection of Water. *Environmental Science & Technology*, 35(4):3780–3785, 2001.

[4] R.J. Carman and R.P. Mildren. Computer modeling of electrical breakdown in a pulsed dielectric barrier discharge in xenon. *Plasma Science, IEEE Transactions on*, 30(1):154–155, 2002.

[5] R J Carman and R P Mildren. Computer modelling of a short-pulse excited dielectric barrier discharge xenon excimer lamp (λ˜172 nm). *Journal of Physics D: Applied Physics*, 36(2):19–33, 2003.

[6] Ulrich Kogelschatz. Dielectric-barrier discharges: Their history, discharge physics, and industrial applications. *Plasma Chemistry and Plasma Processing*, 23(1):1–46, 2003.

[7] W N G Hitchon and C Wichaidit. The suitability of discretized fluid equations to describe breakdown at atmospheric pressure. *Journal of Physics A: Mathematical and General*, 38(30):6841, 2005.

[8] X. T. Deng, J. J. Shi, and M. G. Kong. Protein destruction by a helium atmospheric pressure glow discharge: Capability and mechanisms. *Journal of Applied Physics*, 101(7):074701–074701–9, 2007.

[9] *International Technology Roadmap for Semiconductors, 2012 Edition.*

[10] Vladimir V. Mitin, VIatcheslav A. Kochelap, and Michael A. Stroscio. *Quantum Heterostructures: Microelectronics and Optoelectronics.* Cambridge University Press, New York, 1999.

[11] Mark Lundstrom. Fundamentals of carrier transport, 2nd edn. *Measurement Science and Technology*, 13(2):230, 2002.

[12] Bryan A. Biegel. *Quantum electronic devices simulation*. PhD thesis, Stanford University, 1997.

[13] L. Shifren, C. Ringhofer, and D.K. Ferry. A wigner function-based quantum ensemble monte carlo study of a resonant tunneling diode. *Electron Devices, IEEE Transactions on*, 50(3):769–773, 2003.

[14] D. Querlioz, J. Saint-Martin, V.-N. Do, A. Bournel, and P. Dollfus. A study of quantum transport in end-of-roadmap dg-mosfets using a fully self-consistent wigner monte carlo approach. *Nanotechnology, IEEE Transactions on*, 5(6):737–744, 2006.

[15] F Rossi, C Jacoboni, and M Nedjalkov. A monte carlo solution of the wigner transport equation. *Semiconductor Science and Technology*, 9(5S):934, 1994.

[16] V. Sverdlov, E. Ungersboeck, H. Kosina, and S. Selberherr. Current transport models for nanoscale semiconductor devices. *Materials Science and Engineering: R: Reports*, 58(6):228 – 270, 2008.

[17] J. Feng and W. N. G. Hitchon. Self-consistent kinetic simulation of plasmas. *Phys. Rev. E*, 61:3160–3173, Mar 2000.

[18] A.J. Christlieb, W.N.G. Hitchon, and E.R. Keiter. A computational investigation of the effects of varying discharge geometry for an inductively coupled plasma. *Plasma Science, IEEE Transactions on*, 28(6):2214–2231, 2000.

[19] W.N.G. Hitchon, G.J. Parker, and J.E. Lawler. Physical and numerical verification of discharge calculations. *Plasma Science, IEEE Transactions on*, 21(2):228–238, 1993.

[20] W.N.G. Hitchon, Timothy J. Sommerer, and J.E. Lawler. A self-consistent kinetic plasma model with rapid convergence. *Plasma Science, IEEE Transactions on*, 19(2):113–121, 1991.

[21] G. J. Parker, W. N. G. Hitchon, and J. E. Lawler. Numerical solution of the boltzmann equation in cylindrical geometry. *Phys. Rev. E*, 50:3210–3219, Oct 1994.

[22] Balaji Jayaraman and Wei Shyy. Modeling of dielectric barrier discharge-induced fluid dynamics and heat transfer. *Progress in Aerospace Sciences*, 44(3):139 – 191, 2008.

[23] Max Duarte, Zdeněk Bonaventura, Marc Massot, Anne Bourdon, Stéphane Descombes, and Thierry Dumont. A new numerical strategy with space-time adaptivity and error control for multi-scale streamer discharge simulations. *Journal of Computational Physics*, 231(3):1002 – 1019, 2012.

[24] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, Jun 1932.

[25] M. Hillery, R.F. O'Connell, M.O. Scully, and E.P. Wigner. Distribution functions in physics: Fundamentals. *Physics Reports*, 106(3):121 – 167, 1984.

[26] M.J. Bastiaans. The wigner distribution function applied to optical signals and systems. *Optics Communications*, 25(1):26 – 30, 1978.

[27] M. J. Bastiaans. Wigner distribution function and its application to first-order optics. *J. Opt. Soc. Am.*, 69(12):1710–1716, Dec 1979.

[28] Kurusch Ebrahimi-Fard and JoséM. Gracia-Bondía. Harmonium as a laboratory for mathematical chemistry. *Journal of Mathematical Chemistry*, 50(3):440–454, 2012.

[29] Robert G. Parr Chengteh Lee Robert C. Morrison, Weitao Yang. Approximate density matrices and wigner distribution functions from density, kinetic energy density, and idempotency constraints. *International Journal of Quantum Chemistry*, 38:819–830, 1990.

[30] Damien Querlioz; Philippe Dollfus. *The Wigner Monte Carlo method for nanoelectronic devices : a particle description of quantum transport and decoherence*. London : ISTE ; Hoboken, NJ : Wiley, 2010.

[31] Fausto Rossi. *Theory of Semiconductor Quantum Devices : Microscopic Modeling and Simulation Strategies*. Berlin, Heidelberg : Springer-Verlag Berlin Heidelberg, 2011.

[32] William R. Frensley. Wigner-function model of a resonant-tunneling semiconductor device. *Phys. Rev. B*, 36:1570–1580, Jul 1987.

[33] William R. Frensley. Boundary conditions for open quantum systems driven far from equilibrium. *Rev. Mod. Phys.*, 62:745–791, Jul 1990.

[34] K. L. Jensen and F. A. Buot. Numerical aspects on the simulation of i-v characteristics and switching times of resonant tunneling diodes. *Journal of Applied Physics*, 67(4):2153–2155, 1990.

[35] Kyoung-Youm Kim and Byoungho Lee. On the high order numerical calculation schemes for the wigner transport equation. *Solid-State Electronics*, 43(12):2243 – 2245, 1999.

[36] Y. Yamada, H. Tsuchiya, and Matsuto Ogawa. Quantum transport simulation of silicon-nanowire transistors based on direct solution approach of the wigner transport equation. *Electron Devices, IEEE Transactions on*, 56(7):1396–1401, 2009.

[37] N. C. Kluksdahl, A. M. Kriman, D. K. Ferry, and C. Ringhofer. Self-consistent study of the resonant-tunneling diode. *Phys. Rev. B*, 39:7720–7735, Apr 1989.

[38] William R. Frensley. Effect of inelastic processes on the self-consistent potential in the resonant-tunneling diode. *Solid-State Electronics*, 32(12):1235 – 1239, 1989.

[39] Haiyan Jiang, Wei Cai, and Raphael Tsu. Accuracy of the frensley inflow boundary condition for wigner equations in simulating resonant tunneling diodes. *Journal of Computational Physics*, 230(5):2031 – 2044, 2011.

[40] P. Dollfus, D. Querlioz, J. Saint-Martin, V.-N. Do, and A. Bournel. Wigner monte carlo approach to quantum transport in nanodevices. pages 277–280, 2008.

[41] Damien Querlioz, Huu-Nha Nguyen, Jérôme Saint-Martin, Arnaud Bournel, Sylvie Galdin-Retailleau, and Philippe Dollfus. Wigner-boltzmann monte carlo approach to nanodevice simulation: from quantum to semiclassical transport. *Journal of Computational Electronics*, 8(3-4):324–335, 2009.

[42] Eric J. Heller. Time-dependent approach to semiclassical dynamics. *The Journal of Chemical Physics*, 62(4):1544–1555, 1975.

[43] Dongsheng Yin, Min Tang, and Shi Jin. The Gaussian beam method for the Wigner equation with discontinuous potentials. *Inverse Probl. Imaging*, 7(3):1051–1074, 2013.

[44] W.N.G. Hitchon, D.J. Koch, and J.B. Adams. An efficient scheme for convection-dominated transport. *Journal of Computational Physics*, 83(1):79 – 95, 1989.

[45] D. J. Koch and W. N. G. Hitchon. The effects of collisions on the plasma presheath. *Physics of Fluids B: Plasma Physics (1989-1993)*, 1(11):2239–2243, 1989.

[46] Yasushi Matsunaga, Tadatsugu Hatori, and Tomokazu Kato. Kinetic simulation of nonlinear phenomena of an ion acoustic wave in gas discharge plasma with convective scheme. *Physics of Plasmas (1994-present)*, 8(3):1057–1069, 2001.

[47] D A Fixel and W N G Hitchon. Kinetic investigation of electron-ectron scattering in nanometer-scale metal- oxide-semiconductor field-effect transistors. *Semiconductor Science and Technology*, 23(3):035014, 2008.

[48] Y Güçlü and W N G Hitchon. A high order cell-centered semi-Lagrangian scheme for multi-dimensional kinetic simulations of neutral gas flows. *Journal of Computational Physics*, 231(12):3289–3316, 2012.

[49] Piotr K Smolarkiewicz. A fully multidimensional positive definite advection transport algorithm with small implicit diffusion. *Journal of Computational Physics*, 54(2):325 – 362, 1984.

[50] Piotr K Smolarkiewicz and Wojciech W Grabowski. The multidimensional positive definite advection transport algorithm: nonoscillatory option. *Journal of Computational Physics*, 86(2):355 – 375, 1990.

[51] Y Güçlü, A.J. Christlieb, and W N G Hitchon. Arbitrarily high order convected scheme solution of the vlasov-poisson system. *Journal of Computational Physics*, 270:711 – 752, 2014.

[52] A A Kulikovsky. The role of photoionization in positive streamer dynamics. *Journal of Physics D: Applied Physics*, 33(12):1514, 2000.

[53] Alejandro Luque, Ute Ebert, Carolynne Montijn, and Willem Hundsdorfer. Photoionization in negative streamers: Fast computations and two propagation modes. *Applied Physics Letters*, 90(8):081501, 2007.

[54] W N G Hitchon. The time history of breakdown. *Journal of Physics D: Applied Physics*, 41(5):222002, 2008.

[55] W N G Hitchon. The structure of the breakdown front. *Physics Letters A*, 373(6):773–775, 2009.

[56] Hongen Jia and Kaitai Li. A third accurate operator splitting method. *Mathematical and Computer Modelling*, 53(7):387–396, 2011.

[57] Robert I. Mclachlan and R. I. Mclachlan. On the numerical integration of ordinary differential equations by symmetric composition methods. *SIAM J. Sci. Comput*, 16:151–168, 1995.

[58] R McDermott. Analytical solutions to the continuity equation. *NIST Technical Note 1488*, (13), 2007.

[59] Weizhu Bao, Shi Jin, and Peter A. Markowich. On time-splitting spectral approximations for the schrödinger equation in the semiclassical regime. *Journal of Computational Physics*, 175(2):487 – 524, 2002.

[60] Eric W. Weisstein. *CRC Concise Encyclopedia of Mathematics*. CRC Press; 2 edition, 2002.

[61] A. Arnold and C. Ringhofer. An operator splitting method for the wigner-poisson problem. *SIAM Journal on Numerical Analysis*, 33(4):1622–1643, 1996.

[62] Nam-Duk Suh, Marl R Feix, and Pierre Bertrand. Numerical simulation of the quantum liouville-poisson system. *Journal of Computational Physics*, 94(2):403 – 418, 1991.

[63] D. Querlioz, J. Saint-Martin, K. Huet, A. Bournel, V. Aubry-Fortuna, C. Chassat, S. Galdin-Retailleau, and P. Dollfus. On the ability of the particle monte carlo technique to include quantum effects in nano-mosfet simulation. *Electron Devices, IEEE Transactions on*, 54(9):2232–2242, Sept 2007.

[64] V. Sverdlov, A. Gehring, H. Kosina, and S. Selberherr. Quantum transport in ultra-scaled double-gate mosfets: A wigner function-based monte carlo approach. *Solid-State Electronics*, 49(9):1510 – 1515, 2005. Special Issue: Papers Selected from the {EUROSOI} Workshop Granada, 19-21 January 2005.

[65] Paul F. Zweifel C. Sean Bohun, Reinhard Illner. Some remarks on the wigner transform and the wigner-poisson system. *Le Matematiche*, 46(1):429–438, 1991.

[66] Uri Peskin and Nimrod Moiseyev. The solution of the time?dependent schrödinger equation by the (t,t') method: Theory, computational algorithm and applications. *The Journal of Chemical Physics*, 99(6):4590–4596, 1993.

[67] Shi Jin, Peter Markowich, and Christof Sparber. Mathematical and computational methods for semiclassical schrödinger equations. *Acta Numerica*, 20:121–209, 5 2011.

[68] Peter A. Markowich, Paola Pietra, and Carsten Pohl. Numerical approximation of quadratic observables of schrödinger-type equations in the semi-classical limit. *Numerische Mathematik*, 81(4):595–630, 1999.

[69] Peter A. Markowich, Paola Pietra, Carsten Pohl, and Hans Peter Stimming. A wigner-measure analysis of the dufort–frankel scheme for the schrödinger equation. *SIAM Journal on Numerical Analysis*, 40(4):1281–1310, 2002.

[70] Dongsheng Yin, Min Tang, and Shi Jin. The gaussian beam method for the wigner equation with discontinuous potentials. 43, 2013.

[71] Shi Jin and Xin Wen. Hamiltonian-preserving schemes for the liouville equation with discontinuous potentials. *Comm. Math. Sci*, 3:285–315, 2006.

[72] Shi Jin and Xin Wen. Hamiltonian-preserving schemes for the liouville equation of geometrical optics with discontinuous local wave speeds. *Journal of Computational Physics*, 214(2):672 – 697, 2006.

[73] Björn Engquist, Olof Runborg, and Anna-Karin Tornberg. High-frequency wave propagation by the segment projection method. *Journal of Computational Physics*, 178(2):373 – 390, 2002.

[74] S. Fomel and J. A. Sethian. Fast-phase space computation of multiple arrivals. *Proceedings of the National Academy of Sciences of the United States of America*, 11:7329–7334, 2002.

[75] Shi Jin, Hailiang Liu, Stanley Osher, and Richard Tsai. Computing multi-valued physical observables for the high frequency limit of symmetric hyperbolic systems. *Journal of Computational Physics*, 210(2):497 – 518, 2005.

[76] Shi Jin and Xin Wen. Hamiltonian-preserving schemes for the liouville equation with discontinuous potentials. *Communications in Mathematical Sciences*, 3(3):285–315, 2005.

[77] Dongming Wei, Shi Jin, Richard Tsai, and Xu Yang. A level set method for the semiclassical limit of the schrödinger equation with discontinuous potentials. *Journal of Computational Physics*, 229(19):7440 – 7455, 2010.

[78] Naoufel Ben Abdallah. A hybrid kinetic-quantum model for stationary electron transport. *Journal of Statistical Physics*, 90(3-4):627–662, 1998.

[79] N. Ben Abdallah, P. Degond, and I. M. Gamba. Coupling one-dimensional time-dependent classical and quantum transport models. *Journal of Mathematical Physics*, 43(1):1–24, 2002.

[80] Tao Tang and Zhen Huan Teng. The sharpness of Kuznetsov's $O(\sqrt{\Delta x})L^1$-error estimate for monotone difference schemes. *Mathematics of Computation*, 64(210):581–589, April 1995.

[81] N. N. Kuznetsov. On stable methods for solving non-linear first order partial differential equations in the class of discontinuous functions, topics in numerical analysis. *Academic Press*, 1977.

[82] Yuhui Sun, Y.C. Zhou, Shu-Guang Li, and G.W. Wei. A windowed fourier pseudospectral method for hyperbolic conservation laws. *Journal of Computational Physics*, 214(2):466 – 490, 2006.

[83] John P Boyd. *Chebyshev and Fourier Spectral Methods*. Berlin: Springer-Verlag, 1999.

[84] Bengt Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation*, 51(184):699–706, October 1988.

Appendix A

Fourier Filter for the spectral-CS

In Section4.1, the spectral CS is used to find the high order corrections by using FFT/IFFT to find the high order derivatives. In order to obtain a stable numerical scheme, we multiply the Fourier spectrum of the high-order corrections by the Fourier spectrum of the regularized Shannon kernel (RSK) described in [51, 82]:

$$K(x) = \frac{\sin(\pi X)}{\pi X} \exp(-\frac{X^2}{2\sigma^2}),\ X \in \mathbb{R} \tag{A.1}$$

The filter has nothing to do with the dimension of the domain, because it is based on a regularization of the Shannon kernel, which is for an infinite domain. We can just imagine a function of an arbitrary variable X on an infinite domain.

Now, because of the exponential, this function will decrease very fast toward plus and minus infinity, and therefore we can consider a small interval. Here, the interval is symmetric interval $[-W, W]$ where $W = 9\sigma$ and $\sigma = 4$.

Further, since $K(X)$ is band-limited, with a Fourier spectrum contained between $-\pi$ and π, we can recover its spectrum using a sampling interval of 1. But, $K(X)$ is zero at all integer values of X, therefore we sample it at half-integer values instead, such as $X_i = i + \frac{1}{2}$ with $i \in \mathbb{N}$.

After computing the discrete Fourier transform of $2W$ sample by $K(X)$, we can obtain the $\hat{K_r} \in \mathbb{R}$ of the $2W$ Fourier modes. Finally, for a given ξ_r and $\hat{K_r}$, we can find the $\hat{K}(\xi \Delta x)$ by using a cubic spline over the whole internal $[-\pi, \pi]$.

Appendix B

Poisson solver 1

Section 4.1, the spectral CS is to find the high order corrections by using FFT/IFFT when the space domain is periodic. Since the domain is periodic and the mesh is uniform, this is also called pseudo-spectral Fourier method[83]. Here, we can use the same method to solve the Poisson equation, Eq. (B.1)

$$\frac{\partial E}{\partial x} = \frac{q}{\epsilon_0}(n_e - n_i),\ x \in [a, b], \tag{B.1}$$

For the sake of brevity, we refer the right hand side of Eq. (B.1) as the net charge density $\rho(x)$, Eq. (B.3).

$$\frac{\partial E}{\partial x} = \rho(x),\ x \in [a, b], \tag{B.2}$$

In the following description, both the spatial index i and frequency index r range from 0 to $(N_x - 1)$, where N_x is the number of mesh subdivisions. $E(x)$(or E_i) and $\rho(x)$(or ρ_i)

are approximated by the trigonometric polynomials $E_T(x)$ and $\rho_T(x)$.

$$\begin{aligned} E_T(x) &= \frac{1}{N_x}\sum_{r=0}^{N_x-1}\left\{\sum_{r=0}^{N_x-1} E_i\omega^{-ir}\right\}\omega^{ir}, \\ \rho_T(x) &= \frac{1}{N_x}\sum_{r=0}^{N_x-1}\left\{\sum_{r=0}^{N_x-1} \rho_i\omega^{-ir}\right\}\omega^{ir}, \end{aligned} \tag{B.3}$$

Since the mesh is uniform, the coefficients of the trigonometric polynomials can be obtained by the discrete Fourier transform (DFT) of the grid values. Moreover, the scalar quantity $\omega := \exp(2\pi j/N_x)$ represents the N_x-th primitive root of unity. The detailed algorithm is as follows:

1. Take the Fourier transform on Eq. (B.3)

$$\mathscr{F}\left\{\frac{\partial E}{\partial x} = \rho(x)\right\} \;\Rightarrow\; (j\xi_r)\hat{E}_r = \hat{\rho}_r$$

   Here, ξ_r is the Fourier conjugate of x, $\hat{E}_r = \mathscr{F}\{E_i\}$ and $\hat{\rho}_r = \mathscr{F}\{\rho_i\}$.

$$\xi_r = \begin{cases} 2\pi r/L; \text{ for } r \le N_x/2, \\ 2\pi(r-N_x)/L; \text{ otherwise.} \end{cases} ; L = (b-a),$$

2. Computer the discrete Fourier transform of $\{\rho_i\}$ using an FFT algorithm

$$\hat{\rho}_r = \sum_{i=0}^{N_x-1} \rho_i\omega^{-ir}$$

3. Computer the Fourier coefficients of the electric field according to the following equation:

$$\hat{E}_0 = 0,\; \hat{E}_r = \frac{\hat{\rho}_r}{j\xi_r} \text{ for } r$$

4. Computer the inverse discrete Fourier transform of $\{\hat{E}_r\}$ using an IFFT algorithm

$$E_i = \frac{1}{N_x} \sum_{r=0}^{N_x-1} \hat{E}_r \omega^{ir}$$

Appendix C

Poisson solver 2

In this section, we describe how to accurately solve the Poisson equation in Section 3.2.1:

$$\frac{\partial E}{\partial x} = \frac{q}{\epsilon_0}(n_e - n_i), \tag{C.1}$$

In order to have a compact expression, we define $q_t = \frac{q}{\epsilon_0}(n_e - n_i)$. The first spatial derivative of the electric field is computed by taking with a fourth-order accuracy backward finite difference [84]:

$$\left(\frac{\partial E}{\partial x}\right)_{x=x_i} = a_1 E(i) + a_2 E(i-1) + a_3 E(i-2) + a_4 E(i-3) + a_5 E(i-4), \tag{C.2}$$

where $a_1 = \frac{25}{12\Delta x}$, $a_2 = \frac{-4}{\Delta x}$, $a_3 = \frac{3}{\Delta x}$, $a_4 = \frac{-4}{3\Delta x}$, $a_5 = \frac{1}{4\Delta x}$ and $a_1 + a_2 + a_3 + a_4 + a_5 = 1$.

Therefore, the Poisson equation can be rewritten this form:

$$a_1 E(i) + a_2 E(i-1) + a_3 E(i-2) + a_4 E(i-3) + a_5 E(i-4) = q_t(i) \tag{C.3}$$

Eq. (C.3) can be written in discrete form:

$$\begin{cases} a_1E(5)+a_2E(4)+a_3E(3)+a_4E(2)+a_5E(1)=q_t(5) \\ a_1E(6)+a_2E(5)+a_3E(4)+a_4E(3)+a_5E(2)=q_t(6) \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad . \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad . \\ a_1E(i-1)+a_2E(i-2)+a_3E(i-3)+a_4E(i-4)+a_5E(i-5)=q_t(i-1) \\ a_1E(i)+a_2E(i-1)+a_3E(i-2)+a_4E(i-3)+a_5E(i-4)=q_t(i) \end{cases} \tag{C.4}$$

Adding the right hand side and left hand side of Eq. (C.4) respectively and using $a_1+a_2+a_3+a_4+a_5=1$, we can get

$$\begin{aligned} &-a_1E(4)-(a_1+a_2)E(3)+(a_4+a_5)E(2)+a_5E(1)+a_1E(i)+(a_1+a_2)E(i-1) \\ &-(a_4+a_5)E(i-2)-a_5E(i-3)=Q(i), \end{aligned} \tag{C.5}$$

where $Q(i)=q_t(5)+q_t(6)+...+q_t(i-1)+q_t(i)$. Here, we assume $E(1)$, $E(2)$, $E(3)$ and $E(4)$ are constant and $Q(1)$, $Q(2)$, $Q(3)$ and $Q(4)$ are zero. We thus have an alternative form of the equation, in terms of the total charge $Q(i)$.

For the corrections to the velocity, n_x/n and n_{xx}/n which will be approximated using the finite difference expressions and are discretized as:

$$\left(\frac{n_x}{n}\right)_{x_i}=\frac{n_{i+1}-n_{i-1}}{2\Delta x}\frac{1}{n_i} \tag{C.6}$$

and

$$\left(\frac{n_{xx}}{n}\right)_{x_i}=\frac{4}{\Delta x^2}\frac{n_{i+1}-2n_i+n_{i-1}}{n_{i+1}+2n_i+n_{i-1}} \tag{C.7}$$

Appendix D

Poisson solver 3

Here, we are introducing a Poisson solver based on the Simpson's rule. The backward differentiation formula (BDF) is a family of implicit methods for the numerical integration of ordinary differential equations. Now, we are applying the 4th-order BDF, which is a multi-step method used for integrating in time stiff ODEs of this form:

$$\frac{dy}{dt} = f(y),$$

In our case we are simply doing a quadrature, because our equation is in the form

$$\frac{\partial E(x)}{\partial x} = \frac{q}{\epsilon_0}(n_e(x) - n_i(x)), \tag{D.1}$$

Again, in order to have a compact expression, we define $q_t(x) = \frac{q}{\epsilon_0}(n_e(x) - n_i(x))$. and Eq. (D.1) can be rewritten as a definite integral:

$$E(x) = E(0) + \int_0^x q_t(x')dx'.$$

In order to compute the integral of $q_t(x)$ on a uniform domain one can use the so-called Newton-Cotes formulas, the most famous of which are the trapezoidal rule (2-nd order) and the Simpson's rule (4-th order). Here, $E(0) = E_{max}$ and we apply the trapezoidal rule for $E(\Delta x)$ and the Simpson's rule for the rest points:

$$E(\Delta x) = E(0) + \frac{\Delta x}{2}(q_t(x = 0) + q_t(\Delta x));$$

and

$$E(x + \Delta x) = E(x) + \frac{x}{6}\left\{q_t(x + \Delta x) + 4q_t(\frac{\Delta x}{2}) + q_t(x)\right\}, x > \Delta x.$$

where $q_t(\frac{\Delta x}{2})$ can be obtained according to cubic interpolation:

$$q_t(\frac{\Delta x}{2}) = (-q_t(x + 2\Delta x) + 9q_t(x + \Delta x) + 9q_t(x) - q_t(x - \Delta x))/16;$$