UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

FINITE DIFFERENCE METHOD ON 1D AND 2D RIEMANN PROBLEMS FOR EULER EQUATIONS WITH DIFFERENT STATE OF EQUATIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

LING JIN
Norman, Oklahoma
2022

FINITE DIFFERENCE METHOD ON 1D AND 2D RIEMANN PROBLEMS FOR EULER EQUATIONS WITH DIFFERENT STATE OF EQUATIONS

A DISSERTATION APPROVED FOR THE
DEPARTMENT OF MATHEMATICS

BY THE COMMITTEE CONSISTING OF

Dr. Ying Wang, Chair

Dr. John Albert

Dr. Alexander Grigo

Dr. Nikola Petrov

Dr. Jessica Ruyle

Acknowledgements

There are so many people whom I owe a debt of gratitude that I hardly know where to begin. I suppose it is proper to start by thanking the primary mathematical contributors. I will start by thanking my advisor Dr. Ying Wang. Without you, I would not have started this adventure, let alone stand here today. I really appreciate all of your help over the past seven years. I would also like to thank the remainder of my committee: Dr. John Albert, Dr. Alexander Grigo, Dr. Nikola Petrov and Dr. Jessica Ruyle. Thank you all for serving on my committee and being patient with me and all kinds of deadlines and signatures. I would also like to especially thank Dr.Ruyle, my committee member from outside of the department, for taking such an interest in my dissertation topic. I would, in particular, like to thank Professors Wayne and Sepideh Stewarts for mentally supporting me when I would drop by your offices and never making me feel unwelcome, as well as our conversations not related to mathematics. Next I would like to thank Dr. Ameya Pitale and Dr. Deborah Moore-Russo for arranging teaching assignments so I could stay with my husband in Little Rock after we got married. At last, we really would like to express my great thanks to Dr. Qi Wang from the Southwestern University of Finance and Economics. Without you, I would not have the opportunity to pursue further education at OU, let alone all of the wonderful experiences ever since.

Next I would like to thank my family. To my parents and grandparents, you have all encouraged me to pursue my education and have been there to support me throughout this long journey that I had never imagined that I would go on. It means a lot to me that you have always been there for me and I am truly blessed to have the best family that anyone could ask for. I would also like to thank my husband, Andrew, who has been the closest by during my hard times and good times. I did not feel at home here

until you came into my life. I could have never made it this far without you. To my uncles, Guoqing and Guoli, I would like to thank you for your support as well. I have really enjoyed our conversations. I would also like to thank our cats Kitcat, Little Paw, and Oreo for adding more happiness to my life and relieving so much stress. It is a wonder how I made it through four years of grad school without you all.

I would like to thank my friends who have been supportive. Let me start with Wanru, my roommate when I just came to the U.S.. It was such a stress relief that we could talk in Chinese in the evening. We are so different, at the same time, so much alike. We have helped each other find the best of ourselves. I also would love to thank my friend, Xuehui, who went abroad with me in the same year but to Canada. Xuehui has always motivated me to study English since middle school, and to finish my degree ever since she received her PhD in Psychology from the University of Alberta. I also want to thank Shaoyun and his wife, Yulin, for all of the home-feeling parties. Shaoyun is going to be a math professor at Xiamen University in China. I wish him all of the best in his future journey!

Contents

Abstract

The compressible Euler equations of gas dynamics are governed by the following system

$$\begin{cases} \rho_t + \nabla \cdot (\rho \vec{u}) = 0, \\ (\rho \vec{u})_t + \nabla \cdot (\rho \vec{u} \otimes \vec{u}) + \nabla p = f\rho, \\ E_t + \nabla \cdot [(E+p)\vec{u}] = f\rho\vec{u}, \end{cases} \tag{-1.0.1}$$

where ρ is the density function, $\vec{u}$ is the velocity field, E is the total energy function, p is the pressure function and f is an external force. It is well known that the finite difference method, the finite volume method and the discontinuous Galerkin method are popular in solving the above system. Many high order of convergence results and the positivity preserving properties have been developed on the finite volume method and the discontinuous Galerkin method. [ZS10] did some pioneering work by extending the good features to the finite difference method as a modification to the finite volume method. The equation of state in [ZS10] was the common perfect gas equation

$$p = (\gamma - 1)(E - \rho u^2/2), \tag{-1.0.2}$$

where $\gamma = 1.4$.

In this dissertation, we focused on the finite difference method introduced in [ZS10] and applied the scheme on a one dimensional system with the Chaplygin gas as the equation of state

$$p = -\frac{1}{\rho}, \tag{-1.0.3}$$

and a two dimensional system with zero-pressure, i.e.,

$$p \equiv 0. \tag{-1.0.4}$$

In particular, the second system produces vacuum density areas which sometimes cause the approximated velocity in those areas to blow up. And the maximum velocity value from each time step determines the following time step size

$$\Delta t = CFL \times \min\left\{\frac{\Delta x}{\max_j\{u^n\}}, \frac{\Delta y}{\max_j\{u^n\}}\right\}. \tag{-1.0.5}$$

u or v blowing up will cause Δt to vanish, and therefore the approximation will fail. [ZS10] introduced a convex set to maintain the positivity of the density function and the boundedness of the velocity field

$$G = \left\{\left(\rho, \rho u, \rho v\right)^T | \rho > 0, u^2 + v^2 \leq S^2\right\}, \tag{-1.0.6}$$

and a limiter to preserve the future values $\left(\rho^{n+1}, (\rho u)^{n+1}, (\rho v)^{n+1}\right)^T$ inside G, given that the current values $\left(\rho^n, (\rho u)^n, (\rho v)^n\right)^T$ are inside G. We, in addition to the limiter, introduced a global condition that depends on the density level on each grid point to determine whether the limiter is or not applied at the location. This global condition greatly improved the accuracy of the scheme.

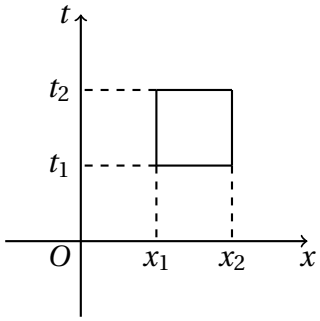
Chapter 0

Introduction

In computational fluid dynamics, the compressible Euler equations are one of the most commonly used partial differential equation system to describe the motions of some certain fluid with zero viscosity. Leonhard Euler introduced the first two equations of the system in [Eul57] in 1757. The first equation describes the conservation law of the density $\rho(x,t)$

$$\rho_t + (\rho u)_x = 0 \tag{0.0.1}$$

where $u(x,t)$ is the velocity function in one dimensional spatial space. This equation is later known as the continuity equation. It can be derived from the integral form. Consider the test domain $[x_1, x_2] \times [t_1, t_2]$



The mass between x_1 and x_2 is

$$\int_{x_1}^{x_2} \rho(x,t)\,dx. \tag{0.0.2}$$

The rate of change of the mass between x_1 and x_2 with respect to time t is

$$\frac{d}{dt}\int_{x_1}^{x_2} \rho(x,t)\,dx. \tag{0.0.3}$$

The total change of the mass between x_1 and x_2 from t_1 to t_2 is

$$\int_{t_1}^{t_2} \Big[\frac{d}{dt}\int_{x_1}^{x_2} \rho(x,t)\,dx\Big]\,dt. \tag{0.0.4}$$

The flux of the fluid going inside the region $[x_1,x_2]$ at time t is $\rho(x_1,t)u(x_1,t)$, and the flux going outside is $\rho(x_2,t)u(x_1,t)$. The total net change of the mass from t_1 to t_2 is

$$\int_{t_1}^{t_2} \Big[\rho(x_1,t)u(x_1,t) - \rho(x_2,t)u(x_1,t)\Big]\,dt. \tag{0.0.5}$$

(0.0.4) and (0.0.5) describe the same quantity: net change of the mass in the test domain $[x_1,x_2]\times[t_1,t_2]$, so we obtain the following identity

$$\int_{t_1}^{t_2} \Big[\frac{d}{dt}\int_{x_1}^{x_2} \rho(x,t)\,dx\Big]\,dt = \int_{t_1}^{t_2} \Big[\rho(x_1,t)u(x_1,t) - \rho(x_2,t)u(x_1,t)\Big]\,dt. \tag{0.0.6}$$

Let's consider smooth density and velocity functions here. We rewrite the above equation as follows

$$\int_{t_1}^{t_2}\int_{x_1}^{x_2} \Big[\frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}(\rho u)(x,t)\Big]\,dx\,dt = 0. \tag{0.0.7}$$

Since (0.0.7) holds for all test domains, we obtain the continuity equation

$$\frac{\partial}{\partial t}\rho(x,t) + \frac{\partial}{\partial x}(\rho u)(x,t) = 0. \tag{0.0.8}$$

Omitting (x,t) gives us the simple formula in (0.0.1). More generally, we always use a quantity times the velocity to represent the quantity advecting with the flow. The second equation in the compressible Euler equations describes the momentum ρu and works out in a similar way. The first part is the rate of change of the momentum with respect to time $(\rho u)_t$. The second part is the total flux, consisting of the flux caused by advection ρu^2 and the pressure p between molecules. What's different is that the right hand has some source term, representing an external force f working on the flow. It is also known as the balance law of the momentum

$$(\rho u)_t + (\rho u^2 + p)_x = f\rho. \tag{0.0.9}$$

When the external force is a constant, such as the gravity of the flow, we have the following form

$$(\rho u)_t + (\rho u^2 + p)_x = \beta\rho, \tag{0.0.10}$$

where β is a constant. During the second half of the 19th century, the third equation, the balance law for the energy E, was included in the compressible Euler equations. It starts with the rate of change of the total energy E_t. The total flux consists of the flux from advection Eu and the transport caused by pressure pu. Also taking the external force f into account, we come to the balance law for the energy

$$E_t + [(E+p)u]_x = f\rho u. \tag{0.0.11}$$

Combining the above three equations in a system, we have obtained the compressible Euler equations describing the motion of some fluid.

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + p) = f\rho, \\ E_t + [(E+p)u]_x = f\rho u, \end{cases} \tag{0.0.12}$$

Suppose the external force f is known, we have come to four variables but three equations. Depending on different types of fluids, we apply different choices of the pressure equation, known as the equation of state. For example, when

$$p = (\gamma - 1)(E - \rho u^2/2), \tag{0.0.13}$$

with $\gamma = 1.4$, a perfect gas is introduced. A perfect fluid is assumed to have zero viscosity and thermal conductivity. These assumptions greatly simplify the process of mathematically modeling. $\gamma > 1$ gives a gas. The vector form of (0.0.12) is as follows

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x = \mathbf{S}(\mathbf{w}), \tag{0.0.14}$$

where

$$\mathbf{w} = \begin{pmatrix} \rho \\ m \\ E \end{pmatrix}, \ \mathbf{f}(\mathbf{w}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ (E+p)u \end{pmatrix} \text{ and } \mathbf{S}(\mathbf{w}) = \begin{pmatrix} 0 \\ f\rho \\ f\rho u \end{pmatrix}. \tag{0.0.15}$$

When $f \equiv 0$, the system describes the three conservation laws for the density, the momentum and the energy

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x = \mathbf{0}. \tag{0.0.16}$$

The conservation laws with the perfect gas is well studied analytically and numerically. There are plenty of papers analyzing the exact solution behaviors. [CM14] considered the classical compressible Euler's Equations in three space dimensions with an arbitrary equation of state, and whose initial data corresponds to a constant state outside a sphere. [LPR99] introduced a central Weighted Essential Non-Oscillatory (WENO) scheme on the the system (0.0.14) over Sod's initial data for the shock tube problem [Sod78], Lax's initial data for the shock tube problem [Lax54] and the double blast wave setting by Woodward and Colella [WC84]. [DLY18] solve the three-dimensional system with the finite volume method with WENO spatial reconstruction. It is well-known facts that this system has three distinct eigenvalues, thus, it is a strictly hyperbolic system. The system produces non-smooth solutions even with smooth initial conditions. These non-smooth solutions include Delta shock solutions, rarefaction wave solutions, and contact discontinuity solutions. Numerous papers have been studying the system and improving the numerical order of convergene. In particular, both the finite volume method and the discontinuous Galerkin (DG) method were in [ZS10], [ZS11b], and [ZS11a]. They not only preserve uniform high order of convergence, but also positivity of density and pressure. Furthermore, [ZS11b] pointed out that the finite volume method and the discontinuous Galerkin method for the compressible Euler equations with a gravitational source term would preserve the positivity of density and pressure functions under certain CFL conditions. Zhang and Shu in [ZS12] further developed the positivity preserving property with the finite difference WENO method. On one hand, the big advantage of the finite difference method is to greatly reduce the computational cost, especially for problems with higher spatial dimensions than one. On the other hand, it is most natural to choose the finite difference method when the target system has a nonzero right hand side. It is well known that the finite volume

method and the discontinuous method are based the cell average values instead of point values. We can apply Gaussian quadrature techniques to approximate the right hand side but will require approximations to future values first.

In this dissertation, we extend the finite difference method described in [ZS12] to solve two compressible Euler equations. The first project is the one dimensional Chaplygin gas with a constant external force [PHW18]

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + p)_x = \beta\rho, \\ E_t + [(E+p)u]_x = \beta\rho u, \end{cases} \tag{0.0.17}$$

where

$$p = -1/\rho \tag{0.0.18}$$

and β, denoting the external force to the system (0.0.17), is a constant. The above equation of state is used to characterize a perfect fluid called the Chaplygin gas. The system (0.0.17) was introduced to approximate the lifting force on a wing of an airplane in aerodynamics [Cha02], [Tsi39], [vK48]. The Chaplygin gas possesses a negative pressure and describes a transition from a decelerated cosmological expansion to a stage of cosmic acceleration. The Chaplygin gas was also advertised as a possible model of dark energy in [BTV02], [GKMP05], and [Set07]. Studies over the generalized Chaplygin gas [BBS02] and [MSON$^+$21] introduce two positive real parameters A and $\alpha \leq 1$ into the equation of state

$$p = -\frac{A}{\rho^\alpha}. \tag{0.0.19}$$

The generalized Chaplygin gas system has been extensively studied as quartessence prototype under dark matter and dark energy. [FA18] extended the generalized Chap-

lygin gas models by introducing one more positive real parameter β, where $\beta = 1$ or $\beta = (1+\alpha)/(2\alpha)$. The system was analyzed through Lagrangian formulation. In [PHW18], Pang, Hu and Wang derived two kinds of analytical soluitons, one consisting of contact discontinuity solutions, and the other delta shock solutions by analyzing the generalized Rankine-Hugoniot relation and entropy condition. They further pointed out that the solutions to the system (0.0.17) are no longer self-similar as the constant external force was introduced. The second project is the two dimensional pressureless system

$$\begin{cases} \rho_t + (\rho u)_x + (\rho v)_y = 0, \\ (\rho u)_t + (\rho u^2)_x + (\rho uv)_y = 0, \\ (\rho v)_t + (\rho uv)_x + (\rho v^2)_y = 0, \end{cases} \tag{0.0.20}$$

The one dimensional pressureless system is as follows

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2)_x = 0. \end{cases} \tag{0.0.21}$$

The 1D system (0.0.21) is well-studied analytically. e.g. [Bou94] conducted a study of priori estimates and defined a notion of measure solution and solve the system for a few examples of initial data, especially the Riemann problems. [BJ99] interpreted the system (0.0.21) as two nonlinearly coupled linear equations and proved the existence of solutions for the Cauchy problem and the uniqueness under optimal conditions on initial data. [BG98] analyzed the system (0.0.21) with the sticky particle models. [CL03] discusses the phenomena of concentration and cavitation and the formalation of delta shocks and vacuum states in solutions to the system (0.0.21) as the pressure vanishes. [ERS96] showed the global weak solutions can be constructed explicitly using the ini-

tial data by a procedure analogous to the Lax-Oleinik variational principle for scalar conservation laws. However, the weak solution is not unique among weak solutions satisfying the standard entropy condition. There are a few numerical methods that have been conducted over the system (0.0.21). [BJL03] designed a first-order and a second-order methods based on kinetic approximations for the one dimensional system. [BM12] added an artificial viscosity term ϵu_{xx} on the right hand side in the momentum equation to make it a diffusive system. A significant drawback of their scheme is that it does not ensure the exact conservation of the total momentum, since it involves a scheme on the velocity and not on the momentum. [CKR07] proposed a new sticky particle method and proved rigorously that their particle approximation satisfies the original system of pressureless gas dynamics in a week sense, but only within a certain residual. [BBT06] implemented the relaxation schemes on both systems (0.0.21) and (0.0.20). But they introduced one more equation into the system and led to more computational cost. The pressureless systems, both (0.0.21) and (0.0.20) are not strictly hyperbolic and produce vacuum solutions that require special treatments to preserve the positivity of the density function. Yang, Wei and Shu in [YWS13] introduced convex sets for both one- and two dimensional systems using the discontinuous Galerkin method. The density is nonnegative and the velocity is bounded in the convex set. They further introduced how to set up limiters to maintain future values inside the convex set, given that the current values are inside the convex set. This is one of the few works that focus on the numerical aspect of the two dimensional pressureless system of gas dynamics. To the best knowledge of the author, no work has been done with the finite difference method on the two dimensional pressureless system (0.0.20). We are interested in the finite difference WENO method because [ZS10] points out that it has smaller memory cost compared with the discontinuous Galerkin methods and

smaller computational cost compared both to the finite volume WENO schemes and to DG schemes for multi-dimensional problems. [CSA94] gave a detailed comparison in the context of Essential Non-Oscilatory (ENO) schemes [Shu90], [Shu99], [SZE$^+$92], etc.

This dissertation is organized as follows. In Chapter 1, we discussed the finite difference WENO scheme with TVD RK4 time integration on the Chaplygin gas system (1.0.4) including the source term. We further tested the numerical accuracy of the scheme over conservation laws with the perfect gas governed by (0.0.13) and gave the numerical results with the initial conditions introduced by Pang in [Pan17] for the system with the right hand side being $\vec{0}$. In Chapter 2, we introduced the two dimensional pressureless system. We further applied a convex set to maintain the positivity for the density function and the boundedness of the velocity field through a limiter. A global condition depending on the size of the density function was introduced to improve the order of accuracy of the scheme. TVD RK3 time integration was required for the limiter to preserve a convex combination of forward Euler methods. In the last section of Chapter 2, 23 numerical cases were displayed based on the 23 cases of diagrams given in [Pan19].

Chapter 1

Chaplygin Gas

We focus on the conservation law system

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + p)_x = 0, \\ (\rho u^2/2 + H)_t + [(\rho u^2/2 + H + p)u]_x = 0, \end{cases} \tag{1.0.1}$$

with a perfect fluid characterized by the equation of state

$$p = -\frac{1}{\rho}, \tag{1.0.2}$$

which is called the Chaplygin gas. Pang in [Pan17] introduced the above system with the following initial conditions

$$(\rho, u, H) = \begin{cases} (\rho_-, u_-, H_-), \ x < 0, \\ (\rho_+, u_+, H_+), \ x > 0, \end{cases} \tag{1.0.3}$$

where $\rho_i > 0, u_i$, and $H_i > 0, i = -, +$, are constants. Pang [Pan17] applied the method of characteristic analysis on the physically relevant regions. In particular, when $u_+ + 1/\rho_+ > u_- - 1/\rho_-$, a Riemann solution with two contact discontinuities in the density function ρ and the velocity function u, and three discontinuities in the internal energy function H and the total energy function E were obtained. In the other case where $u_+ + 1/\rho_+ \leq u_- - 1/\rho_-$, a Dirac Delta shock function is obtained in both the density function ρ and the internal energy function H. The velocity function maintains a jump solution. Pang [Pan17] further applied Nessyahu-Tadmor scheme [LPR02] to give the numerical illustrations. Nessyahu-Tadmor scheme is a second order central method. In Pang's later work [PHW18], a constant external force β was introduced to the system on the right hand side

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 - 1/\rho)_x = \beta\rho, \\ (\rho u^2/2 + H)_t + [(\rho u^2/2 + H - 1/\rho)u]_x = \beta\rho u, \end{cases} \tag{1.0.4}$$

where β is a constant. (1.0.4) has the same two kinds of exact solutions with the same conditions. The main difference is the discontinuity locations, which depend on the parameter β as well. This time, they didn't provide any numerical results. Our first project is to focus on the numerical methods on the above problem. We will pursue higher order methods with WENO reconstruction in this dissertation.

1.1 Finite Difference WENO Method

Consider the vector form the system

$$\mathbf{w}_t + [\mathbf{f}(\mathbf{w})]_x = \mathbf{S}(\mathbf{w}), \tag{1.1.1}$$

where

$$\mathbf{w} = \begin{pmatrix} \rho \\ \rho u \\ E \end{pmatrix}, \ \mathbf{f}(\mathbf{w}) = \begin{pmatrix} \rho u \\ \rho u^2 - 1/\rho \\ (E - 1/\rho)u \end{pmatrix}, \text{ and } \mathbf{S}(\mathbf{w}) = \begin{pmatrix} 0 \\ \beta\rho \\ \beta\rho u \end{pmatrix}. \tag{1.1.2}$$

Given the numerical domain $a \le x \le b, 0 \le t \le T$, we take the uniform mesh for the finite difference WENO scheme $\Delta x = (b-a)/N$. And the grid points are

$$a = x_1 < x_2 < \cdots < x_{j-1} < x_j < x_{j+1} < \cdots < x_N < x_{N+1} = b, \tag{1.1.3}$$

with

$$x_j = a + (j-1)\Delta x. \tag{1.1.4}$$

for all $j = 1, 2, \cdots, N+1$. Let I_j denote the cell around x_j

$$I_j = (x_{j-1/2}, x_{j+1/2}). \tag{1.1.5}$$

Let's rewrite $\mathbf{f}(\mathbf{w})$ in (1.1.2) in terms of $\rho, m = \rho u, E$ as follows

$$\mathbf{f}(\mathbf{w}) = \begin{pmatrix} m \\ (m^2 - 1)/\rho \\ (E - 1/\rho)m/\rho \end{pmatrix}, \tag{1.1.6}$$

and find the Jacobian matrix of $\mathbf{f}(\mathbf{w})$ with respect to ρ, m, E. The following expression (1.1.7) is the simplified version.

$$\mathbf{f}'(\mathbf{w}) = \begin{pmatrix} 0 & 1 & 0 \\ 1/\rho^2 - u^2 & 2u & 0 \\ 2u/\rho^2 - Eu/\rho & (E-1/\rho)/\rho & u \end{pmatrix}. \tag{1.1.7}$$

The above Jacobian matrix (1.1.7) has three distinct real eigenvalues when $\rho > 0$

$$\lambda_1 = u - 1/\rho,\ \lambda_2 = u, \text{ and } \lambda_3 = u + 1/\rho. \tag{1.1.8}$$

The three linearly independent right eigenvectors form a matrix

$$R(\mathbf{w}) = \begin{pmatrix} -\rho^2 & 0 & \rho^2 \\ -\rho^2 u + \rho & 0 & \rho^2 u + \rho \\ \rho u - E\rho + 1 & 1 & \rho u + E\rho - 1 \end{pmatrix} \tag{1.1.9}$$

The corresponding left eigenvector matrix of $\mathbf{f}'(\mathbf{w})$ is

$$L(\mathbf{w}) = R^{-1}(\mathbf{w}) = \frac{1}{2\rho^2} \begin{pmatrix} -(\rho u + 1) & \rho & 0 \\ 2(\rho^2 u^2 - E\rho + 1) & -2\rho^2 u & 2\rho^2 \\ -(\rho u - 1) & \rho & 0 \end{pmatrix} \tag{1.1.10}$$

These matrices will be used to project split fluxes onto the local characteristic fields. We now apply the semi-discrete finite difference method on (1.1.1) over I_j, and obtain the following form

$$\frac{d}{dt}\mathbf{w}_j(t) = -\frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j+1/2} - \hat{\mathbf{f}}_{j-1/2}\right) + \mathbf{S}_j. \tag{1.1.11}$$

where the numerical flux $\hat{\mathbf{f}}_{j+1/2}$ is the high order approximation to $\mathbf{f}$ at $x = x_{j+1/2}$ and $\mathbf{S}_j = \mathbf{S}(\mathbf{w}_j)$. $\hat{\mathbf{f}}_{j+1/2}$ is obtained from Lax-Friedrich flux splitting method described in Section 1.1.2 and the WENO reconstruction procedure in Section 1.1.3. Before getting to the spatial portion, let's discuss the time integration method we used in this dissertation.

1.1.1 TVD RK4 Time Integration

We apply the adaptive time step size in our algorithm. At each time level $t = t_n$, we find the largest characteristic speed among (1.1.8) caused by advection by

$$\alpha^n = \max_j \left\{ |u_j^n| + 1/\rho_j^n \right\}. \tag{1.1.12}$$

The time step size used for the next loop is defined as

$$\Delta t = CFL \times \frac{\Delta x}{a^n}. \tag{1.1.13}$$

As the flows have two possible directions, positive and negative along the x axis, we theoretically take $CFL \leq 1/2$. In our simulations, $CFL = 1/2$ is taken for the one dimensional linear cases, and $CFL = 0.475$ is taken for one dimensional system cases in Section 1.2. After determining the time step size at each time level $t = t_n$, we can apply the TVD-RK4 [SO89], [LOC94] method on the semi-discrete finite difference scheme we obtained previously in (1.1.11)

$$\frac{d}{dt}\mathbf{w}_j(t) = -\frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j+1/2} - \hat{\mathbf{f}}_{j-1/2}\right) + \mathbf{S}_j. \tag{1.1.14}$$

Denote the right hand side as an opertation $\mathbf{L}$ of $\mathbf{w}_j$ at some time level

$$\mathbf{L} = -\frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j+1/2} - \hat{\mathbf{f}}_{j-1/2}\right) + \mathbf{S}_j. \tag{1.1.15}$$

We start with the variable values at the current time, $\mathbf{w}_j^n$, and the four-stage total variation diminishing Runge-Kutta method, [SO88], [SO89], [LOC94], will give a 4th order accuracy in time t

$$\begin{aligned}
\mathbf{w}_j^{(0)} =& \mathbf{w}_j^n, \\
\mathbf{w}_j^{(1)} =& \mathbf{w}_j^{(0)} + \frac{\Delta t}{2} L(\mathbf{w}_j^{(0)}), \\
\mathbf{w}_j^{(2)} =& \frac{1}{2}\left(\mathbf{w}_j^{(0)} + \mathbf{w}_j^{(1)}\right) - \frac{\Delta t}{4} L(\mathbf{w}^{(0)}) + \frac{\Delta t}{2} L(\mathbf{w}^{(1)}), \\
\mathbf{w}_j^{(3)} =& \frac{1}{9}\left(\mathbf{w}_j^{(0)} + 2\mathbf{w}_j^{(1)} + 6\mathbf{w}_j^{(2)}\right) - \frac{\Delta t}{9} L(\mathbf{w}^{(0)}) - \frac{\Delta t}{3} L(\mathbf{w}^{(1)}) + \Delta t L(\mathbf{w}^{(2)}), \\
\mathbf{w}_j^{n+1} =& \frac{1}{3}\left(\mathbf{w}_j^{(1)} + \mathbf{w}_j^{(2)} + \mathbf{w}_j^{(3)}\right) + \frac{\Delta t}{6} L(\mathbf{w}^{(1)}) + \frac{\Delta t}{6} L(\mathbf{w}^{(3)}).
\end{aligned} \tag{1.1.16}$$

1.1.2 Flux Splitting Method for High Order Numerical Fluxes

Our next goal is to develop the numerical flux at each interface $x_{j+1/2}$ introduced in [ZS10]. Before we review the procedures, let's study the definition of a monotone numerical flux. Denote the numerical flux at $x_{j+1/2}$ as

$$\hat{\mathbf{f}}_{j+1/2} = \mathbf{h}\left(\mathbf{w}_{j+1/2}^-, \mathbf{w}_{j+1/2}^+\right). \tag{1.1.17}$$

The two argument function $\mathbf{h}(\mathbf{a}, \mathbf{b})$ is a monotone flux if [Shu98]

- $\mathbf{h}(\mathbf{a}, \mathbf{b})$ is consistent with the physical flux $\mathbf{f}$, that is, $\mathbf{h}(\mathbf{a}, \mathbf{a}) = \mathbf{f}(\mathbf{w})$;
- $\mathbf{h}(\mathbf{a}, \mathbf{b})$ is Lipschitz continuous function in both arguments;
- $\mathbf{h}(\mathbf{a}, \mathbf{b})$ is a nondecreasing function in $\mathbf{a}$ and anonincreasing function in $\mathbf{b}$, i.e.,

$\mathbf{h}(\uparrow,\downarrow)$.

The definition of a monotone flux inspired us to do flux splitting. In the end, we will define a monotone flux

$$\mathbf{h}(\mathbf{a},\mathbf{b}) = \mathbf{h}^{+}(\mathbf{a}) - \mathbf{h}^{-}(\mathbf{b}). \tag{1.1.18}$$

Given the point values $\mathbf{w}_j^n$ at time level n, we evaluate the flux functions $\mathbf{f}(\mathbf{w})$, and apply the Lax-Friedrichs flux splitting [Shu98], [ZS12]

$$\mathbf{f}^{\pm}(\mathbf{w}_j^n) = \frac{1}{2}\Big(\mathbf{w}_j^n \pm \frac{\mathbf{f}(\mathbf{w}_j^n)}{\alpha^n}\Big) \tag{1.1.19}$$

where

$$\alpha^n = \max_j\{|u_j^n| + 1/\rho_j^n\}. \tag{1.1.20}$$

gives the largest characteristic speed. Flux splitting method is a natural consequence of regarding a fluid as an ensemble of particles [Van91]. Measured along the x axis, some particles move forward, and others backward. This automatically splits the fluxes of mass, momentum and energy into two parts heading in opposite directions. With the step above, we split the flux $\mathbf{f}(\mathbf{w}_j^n)$ inside the cell I_j into two fluxes, $\mathbf{f}^{+}(\mathbf{w}_j^n) \geq 0$ and $\mathbf{f}^{-}(\mathbf{w}_j^n) \geq 0$, heading in the positive and negative directions of x axis, respectively. The supercripts $\pm$ denote the directions of the split flux at the grid point x_j.

[Shu98] and [ZS12] pointed out that for every time level t_n, there exists a vector of functions $\mathbf{h}^{\pm}$ where the cell average values of $\mathbf{h}^{\pm}$ over I_j are equal to $\mathbf{f}^{\pm}(\mathbf{w}_j^n)$ respectively. Let $\overline{\mathbf{h}}_{\pm i}^n = \mathbf{f}^{\pm}(\mathbf{w}_i^n), i = j-2,\cdots,j+3$ denote the stencil points for the fluxes $\hat{\mathbf{f}}_{j+1/2}$. The subscripts $\pm$ in $\overline{\mathbf{h}}_{\pm i}^n$ denote the directions of the split fluxes like the supercripts in $\mathbf{f}^{\pm}(\mathbf{w}_i^n)$. Transform all the cell averages $\overline{\mathbf{h}}_{\pm i}^n$ to the local characteristic fields by setting

$$\overline{\mathbf{H}}_{\pm i}^n = L_{j+1/2}\overline{\mathbf{h}}_{\pm i}^n, \tag{1.1.21}$$

where

$$L_{j+1/2} = L\Big(\frac{\mathbf{w}_j + \mathbf{w}_{j+1}}{2}\Big) \tag{1.1.22}$$

was defined in (1.1.10). This step locally decouples the three components of the split fluxes $\overline{\mathbf{h}}^n_{\pm i}$. In numerical analysis, we refer to this step as solving the system characteristicwise. When the Jacobian matrix of the flux function $\mathbf{f}(\mathbf{w})$ is not a constant matrix, the matrices R, L and Λ are now dependent of $\mathbf{f}(\mathbf{w})$. [Shu98] pointed out that we must "freeze" them locally to carry out a similar procedure as in the constant coefficient case. Thus, in order to compute the flux at the interface $x_{j+1/2}$, we will need an approximated Jacobian matrix $\mathbf{f}(\mathbf{w})$ at $\mathbf{w}_{j+1/2}$. The valuation of $\mathbf{f}'(\mathbf{w})$ at the arithmetic mean

$$\mathbf{w}_{j+1/2} = \frac{1}{2}(\mathbf{w}_j + \mathbf{w}_{j+1}) \tag{1.1.23}$$

is applied in the Chaplygin gas system (1.1.1) since the Roe average (1.2.14) does not exist for the Chaplygin gas system. [Shu98] pointed out that the simple mean choice does not affect the high order accuracy of the scheme. On the contrary, solving the system componentwise refers to treating the equations as independent PDEs and solving them individually. This approach is not as robust as the characteristic decomposition, but is essential to non-strictly hyperbolic systems where right-eigenvalue matrix R is singular. The pressureless system (2.0.1) which will be discussed in Chapter 2 is one of the examples.

For each cell I_j, perform the WENO reconstruction from (1.1.28) - (1.1.42) described in Section 1.1.3 on each component of $\overline{\mathbf{H}}^n_{+i}$, $i = j-2, \cdots, j+2$ to obtain approximations of the point value of the function $\mathbf{H}_+$ at the point $x_{j+\frac{1}{2}}$ and denote them as $(\mathbf{H}_+)^-_{j+\frac{1}{2}}$. The + sign inside the parentheses denotes the flux in the cell I_j heading in the positive direction of x-axis before projection, while the $-$ sign outside the parentheses denotes

that such a flux in the cell I_j heads to the interface at $x_{j+1/2}$ from the left.

For each cell I_{j+1}, perform the WENO reconstruction from (1.1.43) - (1.1.56) described in Section 1.1.3 on each component of $\overline{\mathbf{H}}^n_{-i}, i = j-1, \cdots, j+3$ to obtain approximations of the point value of the function $\mathbf{H}_-$ at the point $x_{j+\frac{1}{2}}$ and denote them as $(\mathbf{H}_-)^+_{j+\frac{1}{2}}$. The $-$ sign inside the parentheses denotes the flux in the cell I_{j+1} heading in the negative direction of x-axis before projection, while the + sign outside the parentheses denotes that such a flux in the cell I_{j+1} heads to the interface at $x_{j+1/2}$ from the right.

Project $(\mathbf{H}_\pm)^\mp_{j+1/2}$ values back to the original space to obatin $(\mathbf{h}_\pm)^\mp_{j+1/2}$ values

$$(\mathbf{h}_\pm)^\mp_{j+1/2} = R_{j+1/2}(\mathbf{H}_\pm)^\mp_{j+1/2} \tag{1.1.24}$$

where

$$R_{j+1/2} = R\Big(\frac{\mathbf{w}_j + \mathbf{w}_{j+1}}{2}\Big) \tag{1.1.25}$$

is defined in (1.1.9). [Shu98] pointed out that the simple mean choice does not affect the high order accuracy of our scheme. For the componentwise approach, the projection steps (1.1.21) and (1.1.24) will be omitted. Instead, the WENO reconstruction steps will be performed on $\overline{\mathbf{h}}^n_{\pm i}$ directly.

We form the flux by combining left- and right-going fluxes at the interface $x_{j+1/2}$

$$\hat{\mathbf{f}}_{j+\frac{1}{2}} = \alpha[(\mathbf{h}_+)^-_{j+\frac{1}{2}} - (\mathbf{h}_-)^+_{j+\frac{1}{2}}]. \tag{1.1.26}$$

$(\mathbf{h}_+)^-_{j+\frac{1}{2}}$ and $(\mathbf{h}_-)^+_{j+\frac{1}{2}}$ are $O(\Delta x^4)$ [LOC94] approximations to the positive and negative fluxes passing $x_{j+1/2}$, respectively, because of (1.1.41) and (1.1.55). One more Δx occurs during the subtraction. Therefore, $\hat{\mathbf{f}}_{j+\frac{1}{2}}$ is a fifth-order approximation to the net flux at $x_{j+1/2}$. Now we have obtained the flux terms in the semi-discrete scheme in

(1.1.11)

$$\frac{d}{dt}\mathbf{w}_j(t) = -\frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j+\frac{1}{2}} - \hat{\mathbf{f}}_{j-\frac{1}{2}}\right) + \mathbf{S}_j. \tag{1.1.27}$$

With the choice of the source term introduced in Section 1.1.4, the right hand side of (1.1.27) is a fourth order approximation to $-\mathbf{f}(\mathbf{w})_x + \mathbf{s}(\mathbf{w})$. In the following section 1.1.3, we will introduce how to obtain $(\mathbf{h}_\pm)^\mp_{j+1/2}$ with the WENO spatial reconstruction. The term $\mathbf{S}(\mathbf{w}_j)$ will be introduced in Section 1.1.4. The right hand side of the semi-discrete scheme above gives a 4th order approximation to $-\mathbf{f}(\mathbf{w})_x + \mathbf{s}(\mathbf{w})$. It will then be solved with TVD-RK4 time integration method introduced in (1.1.1) with 4th order of accuracy in time.

1.1.3 WENO Spatial Reconstruction

We apply the WENO reconstruction method introduced in [LOC94]. Note that $\overline{\mathbf{h}}_{+j}$ is a flux heading in the positive direction of x axis. We apply the WENO reconstruction on the projected flux values $\overline{\mathbf{H}}^n_{+j}$. For each cell I_j, to find $(\mathbf{H}_+)^-_{j+\frac{1}{2}}$, we utilize the stencil $\overline{\mathbf{H}}^n_{+i}, i = j-2, j-1, j, j+1, j+2$ to obtain three quadratic interpolations centered at x_{j-1}, x_j, and x_{j+1}, respectively, with $\overline{\mathbf{H}}$ values

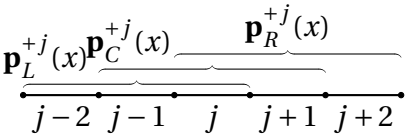
$$\begin{aligned}\mathbf{p}_L^{+j}(x) = &\frac{\overline{\mathbf{H}}^n_{+j} - 2\overline{\mathbf{H}}^n_{+,j-1} + \overline{\mathbf{H}}^n_{+,j-2}}{2\Delta x^2}(x - x_{j-1})^2 + \frac{\overline{\mathbf{H}}^n_{+j} - \overline{\mathbf{H}}^n_{+,j-2}}{2\Delta x}(x - x_{j-1}) \\ &+\overline{\mathbf{H}}^n_{+,j-1} - \frac{\overline{\mathbf{H}}^n_{+j} - 2\overline{\mathbf{H}}^n_{+,j-1} + \overline{\mathbf{H}}^n_{+,j-2}}{24},\end{aligned} \tag{1.1.28}$$

$$\begin{aligned}\mathbf{p}_C^{+j}(x) = &\frac{\overline{\mathbf{H}}^n_{+,j+1} - 2\overline{\mathbf{H}}^n_{+j} + \overline{\mathbf{H}}^n_{+,j-1}}{2\Delta x^2}(x - x_j)^2 + \frac{\overline{\mathbf{H}}^n_{+,j+1} - \overline{\mathbf{H}}^n_{+,j-1}}{2\Delta x}(x - x_j) \\ &+\overline{\mathbf{H}}^n_{+j} - \frac{\overline{\mathbf{H}}^n_{+,j+1} - 2\overline{\mathbf{H}}^n_{+j} + \overline{\mathbf{H}}^n_{+,j-1}}{24},\end{aligned} \tag{1.1.29}$$

and

$$\begin{aligned}\mathbf{p}_R^{+j}(x) = &\frac{\overline{\mathbf{H}}_{+,j+2}^n - 2\overline{\mathbf{H}}_{+,j+1}^n + \overline{\mathbf{H}}_{+j}^n}{2\Delta x^2}(x - x_{j+1})^2 + \frac{\overline{\mathbf{H}}_{+,j+2}^n - \overline{\mathbf{H}}_{+j}^n}{2\Delta x}(x - x_{j+1}) \\ &+\overline{\mathbf{H}}_{+,j+1}^n - \frac{\overline{\mathbf{H}}_{+,j+2}^n - 2\overline{\mathbf{H}}_{+,j+1}^n + \overline{\mathbf{H}}_{+j}^n}{24}.\end{aligned} \tag{1.1.30}$$

The following diagram shows the stencils for the three polynomials above.



Let's use $\mathbf{p}_L^{+j}(x)$ as the example to explain how to find the above three polynomials. The desired $\mathbf{p}_L^{+j}(x)$ need to be exact when comes to the cell average values on the three cells in the stencil, i.e.,

$$\begin{cases} \frac{1}{\Delta x}\int_{I_{j-2}} \mathbf{p}_L^{+j}(x)dx = \overline{\mathbf{H}}_{+,j-2}^n, \\ \frac{1}{\Delta x}\int_{I_{j-1}} \mathbf{p}_L^{+j}(x)dx = \overline{\mathbf{H}}_{+,j-1}^n, \\ \frac{1}{\Delta x}\int_{I_j} \mathbf{p}_L^{+j}(x)dx = \overline{\mathbf{H}}_{+,j}^n, \end{cases} \tag{1.1.31}$$

The three indicators of smoothness inside the cell I_j are

$$\mathbf{IS}_L^j = \frac{1}{2}\left((\overline{\mathbf{H}}_{+,j-1}^n - \overline{\mathbf{H}}_{+,j-2}^n)^2 + (\overline{\mathbf{H}}_{+j}^n - \overline{\mathbf{H}}_{+,j-1})^2\right) + (\overline{\mathbf{H}}_{+j}^n - 2\overline{\mathbf{H}}_{+,j-1}^n + \overline{\mathbf{H}}_{+,j-2}^n)^2. \tag{1.1.32}$$

$$\mathbf{IS}_C^j = \frac{1}{2}\left((\overline{\mathbf{H}}_{+j}^n - \overline{\mathbf{H}}_{+,j-1}^n)^2 + (\overline{\mathbf{H}}_{+,j+1}^n - \overline{\mathbf{H}}_{+j}^n)^2\right) + (\overline{\mathbf{H}}_{+,j+1}^n - 2\overline{\mathbf{H}}_{+j}^n + \overline{\mathbf{H}}_{+,j-1}^n)^2. \tag{1.1.33}$$

and

$$\mathbf{IS}_R^j = \frac{1}{2}\left((\overline{\mathbf{H}}_{+,j+1}^n - \overline{\mathbf{H}}_{+j}^n)^2 + (\overline{\mathbf{H}}_{+,j+2}^n - \overline{\mathbf{H}}_{+,j+1}^n)^2\right) + (\overline{\mathbf{H}}_{+,j+2}^n - 2\overline{\mathbf{H}}_{+,j+1}^n + \overline{\mathbf{H}}_{+j}^n)^2. \tag{1.1.34}$$

The reconstruction solution $\mathbf{R}_j^+(x)$ over the cell I_j will be a convex combination of $\mathbf{p}_L^{+j}(x)$, $\mathbf{p}_C^{+j}(x)$, and $\mathbf{p}_R^{+j}(x)$ as follows

$$\mathbf{R}_j^+(x) = \frac{\alpha_L^{+j}}{\alpha_L^{+j}+\alpha_C^{+j}+\alpha_R^{+j}}\mathbf{p}_L^{+j}(x) + \frac{\alpha_C^{j}}{\alpha_L^{+j}+\alpha_C^{+j}+\alpha_R^{+j}}\mathbf{p}_C^{+j}(x) + \frac{\alpha_R^{+j}}{\alpha_L^{+j}+\alpha_C^{+j}+\alpha_R^{+j}}\mathbf{p}_R^{+j}(x) \tag{1.1.35}$$

where

$$\alpha_L^{+j} = \frac{C_L^+}{(\epsilon+\mathbf{IS}_L^j)^3} \tag{1.1.36}$$

$$\alpha_C^{+j} = \frac{C_C^+}{(\epsilon+\mathbf{IS}_C^j)^3} \tag{1.1.37}$$

and

$$\alpha_R^{+j} = \frac{C_R^+}{(\epsilon+\mathbf{IS}_R^j)^3} \tag{1.1.38}$$

and

$$C_L^+ = \frac{1}{12}, C_C^+ = \frac{1}{2}, \text{ and } C_R^+ = \frac{1}{4} \tag{1.1.39}$$

are linear weights to calculate $(\mathbf{H}_+)^-_{j+\frac{1}{2}}$. Let $(\mathbf{H}_+)_j^n(x)$ denote the actual function whose cell averages are $\overline{\mathbf{H}}_{+,i}^n$, $i = j-2,\cdots,j+2$, and then [LOC94]

$$\mathbf{R}_j^+(x) = (\mathbf{H}_+)_j^n(x) + O(\Delta x^3), \tag{1.1.40}$$

for any $x \in [x_{j-1/2}, x_{j+1/2})$, and

$$\mathbf{R}_j^+(x_{j+1/2}) = (\mathbf{H}_+)_j^n(x_{j+1/2}) + O(\Delta x^4). \tag{1.1.41}$$

Therefore, $(\mathbf{H}_+)^-_{j+1/2} = (\mathbf{H}_+)^n_j(x_{j+1/2})$ is obtained by taking a convex combination of $\mathbf{p}^j_L(x), \mathbf{p}^j_C(x)$ and $\mathbf{p}^j_R(x)$ evaluated at $x_{j+1/2}$

$$\begin{aligned}(\mathbf{H}_+)^-_{j+1/2} =& \frac{\alpha^{+j}_L}{\alpha^{+j}_L+\alpha^{+j}_C+\alpha^{+j}_R}\mathbf{p}^{+j}_L(x_{j+1/2}) + \frac{\alpha^{+j}_C}{\alpha^{+j}_L+\alpha^{+j}_C+\alpha^{+j}_R}\mathbf{p}^{+j}_C(x_{j+1/2}) \\ &+ \frac{\alpha^{+j}_R}{\alpha^{+j}_L+\alpha^{+j}_C+\alpha^{+j}_R}\mathbf{p}^{+j}_R(x_{j+1/2}).\end{aligned} \tag{1.1.42}$$

To calculate $(\mathbf{H}_-)^+_{j+\frac{1}{2}}$, we focus on the cell I_{j+1} and utilize the stencil $\overline{\mathbf{H}}^n_{-i}, i = j-1, j, j+1, j+2, j+3$ to obtain the three quadratic polynomials in the cell I_{j+1}. The three polynomials needed are

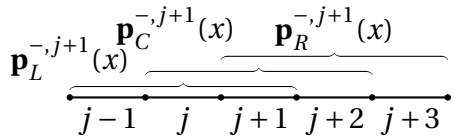
$$\begin{aligned}\mathbf{p}^{-,j+1}_L(x) =& \frac{\overline{\mathbf{H}}^n_{-,j+1}-2\overline{\mathbf{H}}^n_{-j}+\overline{\mathbf{H}}^n_{-,j-1}}{2\Delta x^2}(x-x_j)^2 + \frac{\overline{\mathbf{H}}^n_{-,j+1}-\overline{\mathbf{H}}^n_{-,j-1}}{2\Delta x}(x-x_j) \\ &+\overline{\mathbf{H}}^n_{-j} - \frac{\overline{\mathbf{H}}^n_{-,j+1}-2\overline{\mathbf{H}}^n_{-,j}+\overline{\mathbf{H}}^n_{-,j-1}}{24},\end{aligned} \tag{1.1.43}$$

$$\begin{aligned}\mathbf{p}^{-,j+1}_C(x) =& \frac{\overline{\mathbf{H}}^n_{-,j+2}-2\overline{\mathbf{H}}^n_{-,j+1}+\overline{\mathbf{H}}^n_{-j}}{2\Delta x^2}(x-x_{j+1})^2 + \frac{\overline{\mathbf{H}}^n_{-,j+2}-\overline{\mathbf{H}}^n_{-,j}}{2\Delta x}(x-x_{j+1}) \\ &+\overline{\mathbf{H}}^n_{-,j+1} - \frac{\overline{\mathbf{H}}^n_{-,j+2}-2\overline{\mathbf{H}}^n_{-,j+1}+\overline{\mathbf{H}}^n_{-,j}}{24},\end{aligned} \tag{1.1.44}$$

and

$$\begin{aligned}\mathbf{p}^{-,j+1}_R(x) =& \frac{\overline{\mathbf{H}}^n_{-,j+3}-2\overline{\mathbf{H}}^n_{-,j+2}+\overline{\mathbf{H}}^n_{-,j+1}}{2\Delta x^2}(x-x_{j+2})^2 + \frac{\overline{\mathbf{H}}^n_{-,j+2}-\overline{\mathbf{H}}^n_{-,j+1}}{2\Delta x}(x-x_{j+2}) \\ &+\overline{\mathbf{H}}^n_{-,j+2} - \frac{\overline{\mathbf{H}}^n_{-,j+3}-2\overline{\mathbf{H}}^n_{-,j+2}+\overline{\mathbf{H}}^n_{-,j+1}}{24}.\end{aligned} \tag{1.1.45}$$

The superscripts $-, j+1$ denote the fluxes in the cell of I_{j+1} heading in the negative direction of x axis. The following diagram shows the stencils for each polynomial obtained above.



The three corresponding indicators of smoothness inside the cell I_{j+1} are

$$\mathbf{IS}_L^{j+1} = \frac{1}{2}\Big((\overline{\mathbf{H}}_{-,j}^n - \overline{\mathbf{H}}_{-,j-1}^n)^2 + (\overline{\mathbf{H}}_{-,j+1}^n - \overline{\mathbf{H}}_{-,j}^n)^2\Big) + (\overline{\mathbf{H}}_{-,j+1}^n - 2\overline{\mathbf{H}}_{-,j}^n + \overline{\mathbf{H}}_{-,j-1}^n)^2. \quad (1.1.46)$$

$$\mathbf{IS}_C^{j+1} = \frac{1}{2}\Big((\overline{\mathbf{H}}_{-,j+1}^n - \overline{\mathbf{H}}_{-j}^n)^2 + (\overline{\mathbf{H}}_{-,j+2}^n - \overline{\mathbf{H}}_{-,j+1}^n)^2\Big) + (\overline{\mathbf{H}}_{-,j+2}^n - 2\overline{\mathbf{H}}_{-,j+1}^n + \overline{\mathbf{H}}_{-,j}^n)^2. \quad (1.1.47)$$

and

$$\mathbf{IS}_R^{j+1} = \frac{1}{2}\Big((\overline{\mathbf{H}}_{-,j+2}^n - \overline{\mathbf{H}}_{-,j+1}^n)^2 + (\overline{\mathbf{H}}_{-,j+3}^n - \overline{\mathbf{H}}_{-,j+2}^n)^2\Big) + (\overline{\mathbf{H}}_{-,j+3}^n - 2\overline{\mathbf{H}}_{-,j+2}^n + \overline{\mathbf{H}}_{-,j+1}^n)^2. \quad (1.1.48)$$

The reconstruction solution $\mathbf{R}_{j+1}^-(x)$ will be a convex combination of $\mathbf{p}_L^{-,j+1}(x)$, $\mathbf{p}_C^{-,j+1}(x)$, and $\mathbf{p}_R^{-,j+1}(x)$ as follows

$$\begin{aligned}\mathbf{R}_{j+1}^-(x) = &\frac{\alpha_L^{-,j+1}}{\alpha_L^{-,j+1} + \alpha_C^{-,j+1} + \alpha_R^{-,j+1}}\mathbf{p}_L^{-,j+1}(x) + \frac{\alpha_C^{-,j+1}}{\alpha_L^{-,j+1} + \alpha_C^{-,j+1} + \alpha_R^{-,j+1}}\mathbf{p}_C^{-,j+1}(x) \\ &+ \frac{\alpha_R^{-,j+1}}{\alpha_L^{-,j+1} + \alpha_C^{-,j+1} + \alpha_R^{-,j+1}}\mathbf{p}_R^{-,j+1}(x)\end{aligned} \quad (1.1.49)$$

where

$$\alpha_L^{-,j+1} = \frac{C_L^-}{(\epsilon + \mathbf{IS}_L^{j+1})^3} \quad (1.1.50)$$

$$\alpha_C^{-,j+1} = \frac{C_C^-}{(\epsilon + \mathbf{IS}_C^{j+1})^3} \quad (1.1.51)$$

and

$$\alpha_R^{-,j+1} = \frac{C_R^-}{(\epsilon + \mathbf{IS}_R^{j+1})^3} \tag{1.1.52}$$

and

$$C_L^- = \frac{1}{4}, C_C^- = \frac{1}{2}, \text{ and } C_R^- = \frac{1}{12} \tag{1.1.53}$$

are the updated set of linear coefficients for fluxes heading in the negative direction of x axis. Let's use $(\mathbf{H}_-)_{j+1}^n(x)$ denote the actual function whose cell averages are $\overline{\mathbf{H}}_{-,i}^n$, $i = j-1, \cdots, j+3$, and then we have [LOC94]

$$\mathbf{R}_{j+1}^-(x) = (\mathbf{H}_-)_{j+1}^n(x) + O(\Delta x^3), \tag{1.1.54}$$

for any $x \in (x_{j+1/2}, x_{j+3/2}]$, and

$$\mathbf{R}_{j+1}^-(x_{j+1/2}) = (\mathbf{H}_-)_{j+1}^n(x_{j+1/2}) + O(\Delta x^4). \tag{1.1.55}$$

Therefore, $(\mathbf{H}_-)_{j+1/2}^+ = (\mathbf{H}_-)_{j+1}^n(x_{j+1/2})$ is obtained by

$$\begin{aligned}(\mathbf{H}_-)_{j+1/2}^+ = &\frac{\alpha_L^{-,j+1}}{\alpha_L^{-,j+1} + \alpha_C^{-,j+1} + \alpha_R^{-,j+1}} \mathbf{p}_L^{-,j+1}(x_{j+1/2}) \\ &+ \frac{\alpha_C^{-,j+1}}{\alpha_L^{-,j+1} + \alpha_C^{-,j+1} + \alpha_R^{-,j+1}} \mathbf{p}_C^{-,j+1}(x_{j+1/2}) + \frac{\alpha_R^{-,j+1}}{\alpha_L^{-,j+1} + \alpha_C^{-,j+1} + \alpha_R^{-,j+1}} \mathbf{p}_R^{-,j+1}(x_{j+1/2}).\end{aligned} \tag{1.1.56}$$

1.1.4 Source Term

We have discussed the first term on the right hand side of (1.1.11). Let's continue with the source term $\mathbf{S}_j$. Because the finite difference method is to iterate the point values

at each cell center x_j, we simply evaluate $\mathbf{S}(\mathbf{w})$ at x_j, i.e.,

$$\mathbf{S}(\mathbf{w}_j^n) = \begin{pmatrix} 0 \\ \beta\rho_j^n \\ \beta\rho_j^n u_j^n \end{pmatrix} \tag{1.1.57}$$

at each time step t_n. [ZS12] pointed out that this choice would not affect the high order of accuracy of the scheme.

1.2 Numerical Results

1.2.1 Accuracy Tests

[LOC94] provided us with a finite volume WENO scheme with many examples for testing the order of convergence. The scheme is theoretically 4th order due to the application of RK4 time integration. They obtained better than 4th order results when the initial conditions are sine functions. We ran our finite difference WENO method on the same initial value problems and parameter settings to compared our results with those from [LOC94], and came to the conclusion that the accuracy is consistent. The first five examples are over linear equation

$$u_t + u_x = 0 \tag{1.2.1}$$

with different initial conditions. The first two cases, where the initial conditions are sine functions, gave a good order of convergence higher than 5.

Case 1. Linear Equation

$$u_t + u_x = 0, \text{ with } u_0(x) = \sin(\pi x) \tag{1.2.2}$$

over $[-1,1]$, the final time $T = 1$ and $CFL = 0.5$.

[LOC94] calculated the errors and order of convergence in both L_1 and L_∞ norms. Linear equation together with regular sine function is the ideal case to test for accuracy. The theoretically 4th order scheme gave an an average 5th order result. Then spatial 5th order reconstruction plays a more crucial role in determining the accuracy of the scheme than the order of time integration. As is shown in the above table, the order of convergence increases as grid points get finer. This means we can achieve highly accurate results with lower order schemes by refining the grid points inside the numerical domain. In addition, we noticed that the errors near local extreme values are larger than those in other regions. This is a common issue in many numerical schemes [Kri17], [HW07], etc.

Table 1.1: Case 1.

N	L_1 error	L_1 order	L_∞ error	L_∞ order
80	$4.8526e-05$	-	$9.8817e-05$	-
160	$2.1935e-06$	4.4675	$7.6915e-06$	3.6834
320	$3.9508e-08$	5.7949	$1.3395e-07$	5.8435
640	$5.2566e-10$	6.2319	$1.0750e-09$	6.9613

Case 2. Linear Equation

$$u_t + u_x = 0, \text{ with } u_0(x) = \sin^4(\pi x) \tag{1.2.3}$$

over $[-1,1]$, the final time $T = 1$ and $CFL = 0.5$.

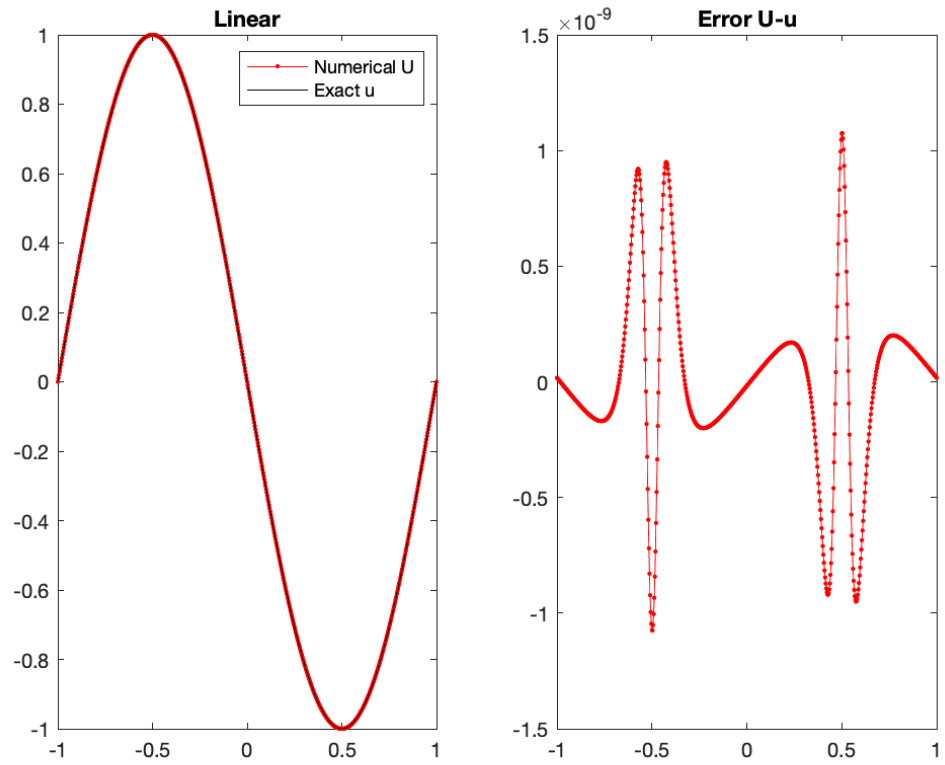


Figure 1.1: Case 1. The number of total grid points is $N = 640$. The left figure gave the plots of the numerical solution U in red-dot-line pattern and the exact solution u in black-line pattern. The right figure gave the point-wise errors between the numerical solution U and the exact solution u.

Case 2. obtained about 1 order lower accuracy than Case 1, due to the higher frequency of $\sin^4(\pi x)$ than $\sin(\pi x)$. Higher frequency of a trigonometric function leads to more local extreme values in the same computational domain. It is well known that order of accuracy of any scheme decreases near extreme values.

Table 1.2: Case 2.

N	L_1 error	L_1 order	L_∞ error	L_∞ order
80	$4.3686e-03$	-	$6.8861e-03$	-
160	$2.2726e-04$	4.2648	$3.8452e-04$	4.1625
320	$7.5479e-06$	4.9121	$3.1863e-05$	3.5931
640	$1.2821e-07$	5.8795	$6.5247e-07$	5.6098

Case 3. Linear Equation

$$u_t + u_x = 0, \text{ with } u_0(x) = \begin{cases} 1, & -\frac{1}{5} \le x \le \frac{1}{5} \\ 0, & \text{otherwise} \end{cases} \tag{1.2.4}$$

over $[-1,1]$, the final time $T = 0.5$ and $CFL = 0.5$.

Case 3. shows that the order of accuracy decreases dramatically if there are contact discontinuities in the solution, even when the equation is linear. This information is important to the Chaplygin gas system. Later in this section when we calculate the order of convergence for the Chaplygin gas system, we will need to take off the contact discontinuity portions and use the rest of points to determine the order of accuracy.

Table 1.3: Case 3.

N	L_1 error	L_1 order	L_∞ error	L_∞ order
80	$6.8589e-02$	-	$6.2865e-01$	-
160	$4.0106e-02$	0.7742	$6.0412e-01$	0.0574
320	$2.3410e-02$	0.7767	$5.8553e-01$	0.0451
640	$1.3599e-02$	0.7836	$5.7173e-01$	0.0344

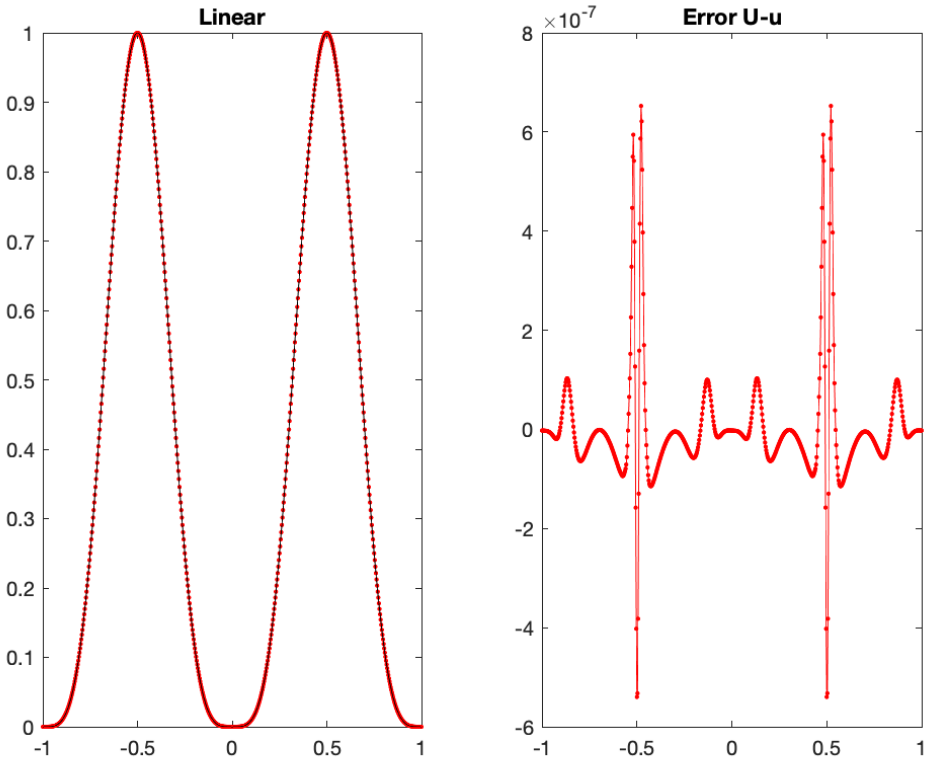


Figure 1.2: Case 2. The number of total grid points is $N = 640$. The left figure gave the plots of the numerical solution U in red-dot-line pattern and the exact solution u in black-line pattern. The right figure gave the point-wise errors between the numerical solution U and the exact solution u.

Case 4. Linear Equation

$$u_t + u_x = 0, \text{ with } u_0(x) = \begin{cases} (1 - (\frac{10}{3}x)^2)^{\frac{1}{2}}, & -\frac{3}{10} \le x \le \frac{3}{10} \\ 0, \text{ otherwise} \end{cases} \tag{1.2.5}$$

over $[-1,1]$, the final time $T = 0.5$ and $CFL = 0.5$.

Case 4. obtained better L_1 order than Case 3. because the initial condition is differentiable. L_∞ order did not get improved due to the two corners between the horizontal line and taller portion.

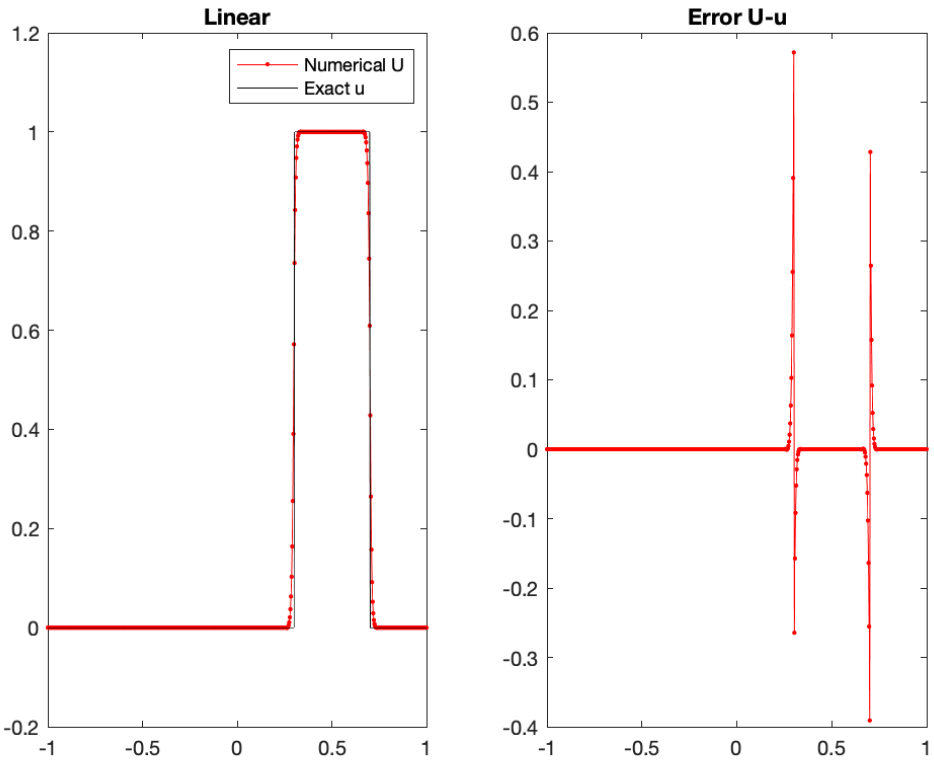


Figure 1.3: Case 3. The number of total grid points is $N = 640$. The left figure gave the plots of the numerical solution U in red-dot-line pattern and the exact solution u in black-line pattern. The right figure gave the point-wise errors between the numerical solution U and the exact solution u.

Case 5. Linear Equation

$$u_t + u_x = 0, \text{ with } u_0(x) = e^{-300x^2} \tag{1.2.6}$$

over $[-1,1]$, the final time $T = 0.5$ and $CFL = 0.5$.

Case 5. has the initial condition smooth, but the order of accuracy is still worse than a regular sine function like what is shown in Case 1. We concluded that the pointy ex-

Table 1.4: Case 4.

N	L_1 error	L_1 order	L_∞ error	L_∞ order
80	$2.3296e-02$	-	$1.4077e-01$	-
160	$1.0405e-02$	1.1629	$1.1151e-01$	0.3361
320	$4.5597e-03$	1.1902	$8.7347e-02$	0.3524
640	$1.9515e-03$	1.2244	$6.7685e-02$	0.3679

treme value could also decrease the order. This information is important to the Chaplygin gas system as well. We will need to take off points near the delta shock location and consider of the error from the rest of the points.

Table 1.5: Case 5.

80	$1.5841e-02$	-	$1.0979e-01$	-
160	$2.8851e-03$	2.4570	$2.8137e-02$	1.9643
320	$3.7606e-04$	2.9396	$3.9767e-03$	2.8228
640	$2.8221e-05$	3.7361	$6.3522e-04$	2.6462

Case 6. Euler Equations of Gas Dynamics: Characteristicwise

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + p)_x = 0, \\ E_t + [(E+p)u]_x = 0, \end{cases} \tag{1.2.7}$$

where

$$p = (\gamma - 1)(E - \rho u^2/2), \text{ and } \gamma = 1.4. \tag{1.2.8}$$

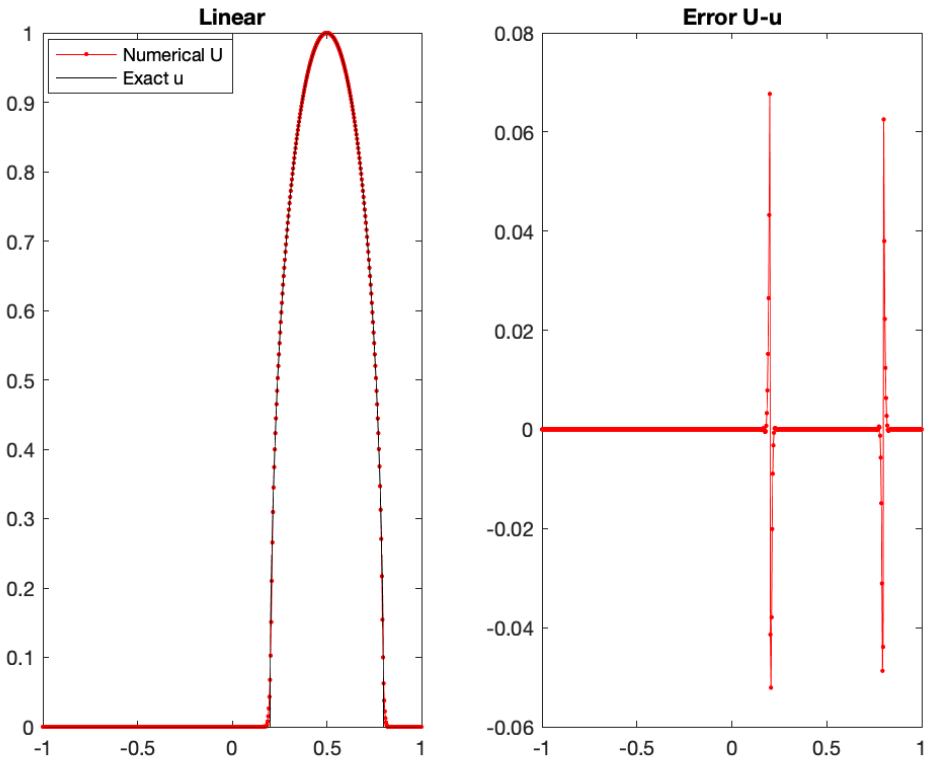


Figure 1.4: Case 4. The number of total grid points is $N = 640$. The left figure gave the plots of the numerical solution U in red-dot-line pattern and the exact solution u in black-line pattern. The right figure gave the point-wise errors between the numerical solution U and the exact solution u.

The initial conditions are

$$\begin{cases} \rho_0(x) = 1 + 0.2\sin(\pi x), \\ u_0(x) = 1, \\ p_0(x) = 1, \end{cases} \tag{1.2.9}$$

We calculate

$$E_0(x) = p_0(x)/(\gamma - 1) + \rho_0(x)u_0^2(x)/2. \tag{1.2.10}$$

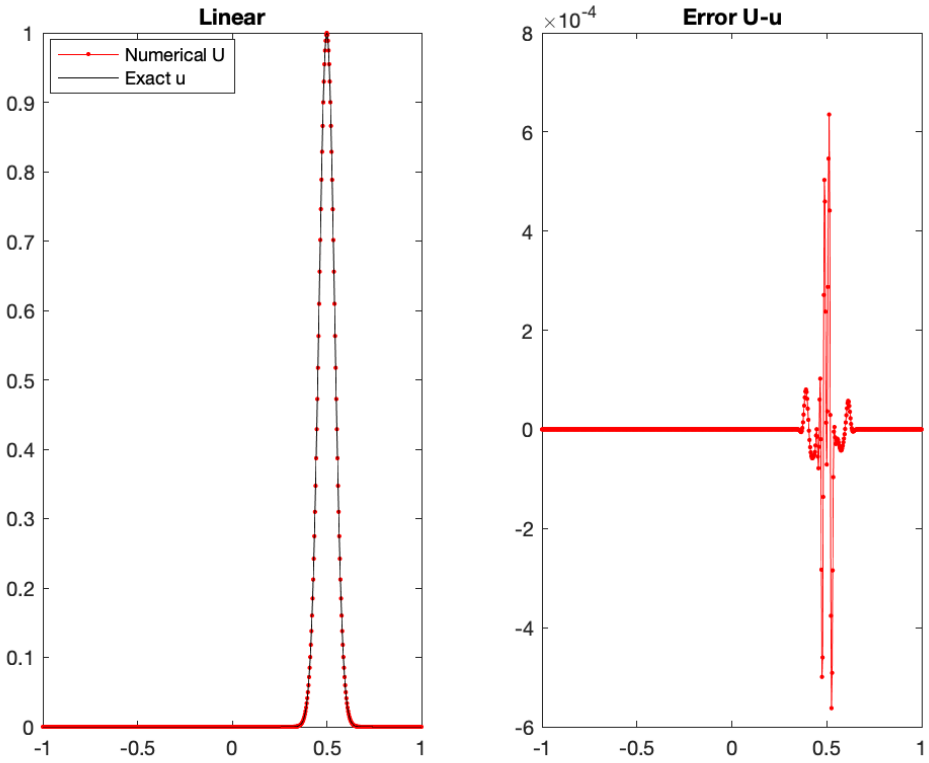


Figure 1.5: Case 5. The number of total grid points is $N = 640$. The left figure gave the plots of the numerical solution U in red-dot-line pattern and the exact solution u in black-line pattern. The right figure gave the point-wise errors between the numerical solution U and the exact solution u.

The exact solution

$$\begin{cases} \rho(x,t) = 1 + 0.2\sin(\pi(x-t)), \\ u(x,t) = 1, \\ p(x,t) = 1, \end{cases} \tag{1.2.11}$$

and

$$E(x,t) = p(x,t)/(\gamma-1) + \rho(x,t)u^2(x,t)/2. \tag{1.2.12}$$

The numerical domain is $[-1,1]$, the final time $T = 1$, and $CFL = 0.475$.

We apply the characteristicwise approach with the Jacobian matrix being the Roe matrix [Roe81]

$$\hat{A} = \mathbf{f}'(\mathbf{w}_{j+1/2}) = \begin{pmatrix} 0 & 1 & 0 \\ (\gamma-1)\hat{H}_{j+1/2} - \hat{u}^2_{j+1/2} - \hat{c}^2_{j+1/2} & (3-\gamma)\hat{u}_{j+1/2} & \gamma - 1 \\ \hat{u}_{j+1/2}/2[(\gamma-3)\hat{H}_{j+1/2} - \hat{c}^2_{j+1/2}] & \hat{H}_{j+1/2} - (\gamma-1)\hat{u}^2_{j+1/2} & \gamma\hat{u}_{j+1/2} \end{pmatrix} \tag{1.2.13}$$

where the Roe averages are defined as

$$\begin{aligned} \hat{u}_{j+1/2} &= \frac{\sqrt{\rho_j}u_j + \sqrt{\rho_{j+1}}u_{j+1}}{\sqrt{\rho_j} + \rho_{j+1}}, \\ \hat{H}_{j+1/2} &= \frac{\sqrt{\rho_j}H_j + \sqrt{\rho_{j+1}}H_{j+1}}{\sqrt{\rho_j} + \rho_{j+1}}, \\ \hat{c}^2_{j+1/2} &= (\gamma-1)\left(\hat{H}_{j+1/2} - \frac{\hat{u}^2_{j+1/2}}{2}\right) \end{aligned} \tag{1.2.14}$$

for compressible Euler equations of gas dynamics with the equation of state $p = (\gamma - 1)(E - \rho u^2/2)$. In particular,

$$H = \frac{E+p}{\rho} \tag{1.2.15}$$

is the total enthalpy. With the settings in (1.2.14), the Jacobian matrix $\hat{A} = \mathbf{f}'(\mathbf{w}_{j+1/2})$ satisfys the mean value theorem

$$\mathbf{f}(\mathbf{w}_{j+1}) - \mathbf{f}(\mathbf{w}_j) = \hat{A}(\mathbf{w}_{j+1} - \mathbf{w}_j). \tag{1.2.16}$$

Even though the Roe average seems to be a better approximation of the variable value at the interfaces, [Shu98] pointed out that the arithmetic mean gives a similarly satisfying result. After all, it is not guaranteed to be able to find the Roe average for every system. The Chaplygin gas system we discussed in Chapter 1 is one of the example where the Roe average does not exist. Therefore, the arithmetic mean will be applied.

Case 6. gave the accuracy of our finite difference WENO method over a conservation law system with the equation of state describing a perfect gas. We applied the characteristic approach described in Chapter 1. The order of accuracy is incredibly good. The order of convergence achieved almost 7th order when the number of grid points increases to $N+1 = 641$. The pressure function, as a variable needing manipulations over the three target variables, may lead to bigger errors. As is shown in the figure that the L_∞ error of the pressure function is less than 3×10^{-13}.

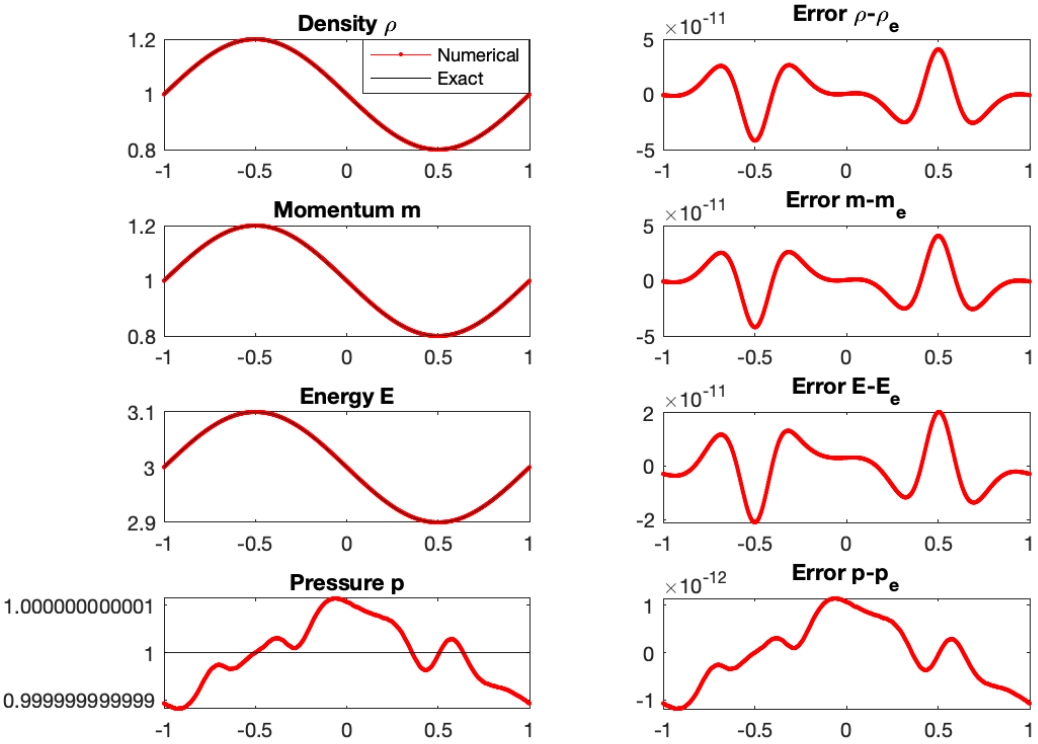


Figure 1.6: Case 6. The number of total grid points is $N = 640$. The left figures gave the plots of the numerical solution **U** in red-dot-line pattern and the exact solution **u** in black-line pattern. The right figures gave the point-wise errors between the numerical solution **U** and the exact solution **u**.

Table 1.6: Case 6. Euler Equations with $p = (\gamma - 1)(E - \frac{1}{2}\rho u^2)$ - Characteristicwise

Energy ρ				
N	L_1 error	L_1 order	L_∞ error	L_∞ order
80	$1.1652e-05$	-	$2.3952e-05$	-
160	$2.3089e-07$	5.6573	$5.8012e-07$	5.3677
320	$2.7911e-09$	6.3702	$5.0922e-09$	6.8319
640	$2.6777e-11$	6.7037	$4.1874e-11$	6.9261
Momentum ρu				
80	$1.1535e-05$	-	$2.3788e-05$	-
160	$2.2844e-07$	5.6581	$5.7430e-07$	5.3723
320	$2.7797e-09$	6.3607	$5.1889e-09$	6.7902
640	$2.6510e-11$	6.7122	$4.2063e-11$	6.9467
Total Energy E				
80	$5.5508e-06$	-	$1.2081e-05$	-
160	$1.1364e-07$	5.6101	$2.8603e-07$	5.4005
320	$1.4405e-09$	6.3018	$2.6560e-09$	6.7508
640	$1.4513e-11$	6.6331	$2.1125e-11$	6.9741

Case 7. Euler Equations of Gas Dynamics: Componentwise

In Case 7., we work on the same system as Case 6. with the same initial conditions, but with a slightly different scheme approach. Without projecting the fluxes $\overline{\mathbf{h}}^n_{\pm i}$onto the characteristic space to obtained $\overline{\mathbf{H}}^n_{\pm i}$ described in (1.1.21), we apply the WENO reconstruction on the original fluxes $\overline{\mathbf{h}}^n_{\pm i}$. The concluded numerical flux values, $(\mathbf{h}_+)^-_{j+1/2}$ and $(\mathbf{h}_+)^-_{j-1/2}$ at the interface $x_{j+1/2}$ will be fed into the overall numerical flux described in (1.1.26) without needing to project $(\mathbf{H}_+)^-_{j+1/2}$ and $(\mathbf{H}_+)^-_{j-1/2}$ back to $(\mathbf{h}_+)^-_{j+1/2}$ and $(\mathbf{h}_+)^-_{j-1/2}$ described in (1.1.24). Case 7. is essential to our dissertation. After comparing the numerical and the order of accuracy results, we came to the conclusion that both componentwise and chacteristicwise approches gave excellent numerical simulations to the conservation law system. Since the two dimensional pressureless system in Chapter 2 does not allow characteristicwise approach, we will be utilizing the componentwise approach.

Table 1.7: Case 7. Euler Equations with $p = (\gamma - 1)(E - \frac{1}{2}\rho u^2)$ - Componentwise

Density ρ				
N	L_1 error	L_1 order	L_∞ error	L_∞ order
80	$1.3744e-05$	-	$2.7705e-05$	-
160	$2.7644e-07$	5.6357	$7.1275e-07$	5.2806
320	$3.3463e-09$	6.3682	$6.0203e-09$	6.8874
640	$3.1643e-11$	6.7246	$4.9482e-11$	6.9268
Momentum ρu				
80	$1.3737e-05$	-	$2.7617e-05$	-
160	$2.7679e-07$	5.6331	$7.1191e-07$	5.2777
320	$3.3476e-09$	6.3695	$6.0188e-09$	6.8861
640	$3.1705e-11$	6.7223	$4.9484e-11$	6.9264
Total Energy E				
80	$7.0160e-06$	-	$1.3914e-05$	-
160	$1.4232e-07$	5.6234	$3.6123e-07$	5.2675
320	$1.6918e-09$	6.3945	$2.9009e-09$	6.9603
640	$1.6383e-11$	6.6902	$2.4478e-11$	6.8889

Theoretically, a numerical scheme for a system of equations is more accurate with a characteristic decomposition compared with simply iterating each equation individually. Comparing Case 6. and Case 7 above, characteristicwise method obtains slightly smaller errors than componentwise method and similar orders of convergence. This greatly increased our confidence in the following project in Chapter 2, as the pressureless system is not strictly hyperbolic, and therefore can't be characteristicwise solved.

1.2.2 Numerical Results to the Chaplygin Gas

In this section, the characteristicwise approach is applied with the arithmetic mean defined in (1.1.23). Pang illustrated the numerical solutions to the zero external force

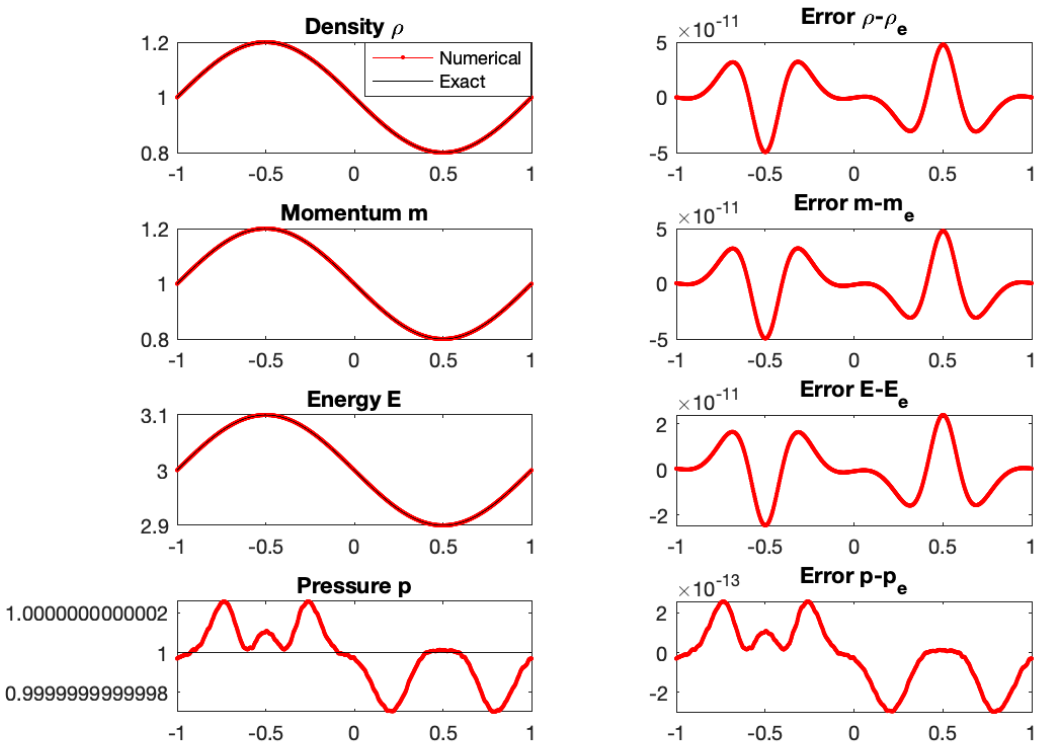


Figure 1.7: Case 7. The number of total grid points is $N = 640$. The left figures gave the plots of the numerical solution $\mathbf{U}$ in red-dot-line pattern and the exact solution $\mathbf{u}$ in black-line pattern. The right figures gave the point-wise errors between the numerical solution $\mathbf{U}$ and the exact solution $\mathbf{u}$.

(1.0.1), i.e.,

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + p)_x = 0, \\ (\rho u^2/2 + H)_t + [(\rho u^2/2 + H + p)u]_x = 0, \end{cases} \tag{1.2.17}$$

where

$$p = -1/\rho. \tag{1.2.18}$$

on the numerical domain $[-2,2]$ at the final time $T = 0.3$ with the following initial data in [Pan17]. The physically relevant region is

$$\left\{(\rho,u,H)\middle|\rho>0, H\geq 1/(2\rho), u\in\mathbb{R}\right\}. \tag{1.2.19}$$

Case 1 A contact discontinuity solution is achieved when $u_+ + \rho_+^{-1} > u_- - \rho_-^{-1}$.

$$\rho_- = 0.53, u_- = -0.33, H_- = 1.13, \rho_+ = 0.54, u_+ = 0.47, H_+ = 1.36 \tag{1.2.20}$$

Case 2 A shock solution is achieved when $u_+ + \rho_+^{-1} \leq u_- - \rho_-^{-1}$.

$$\rho_- = 1.23, u_- = 2.33, H_- = 1.33, \rho_+ = 7.34, u_+ = -0.87, H_+ = 8.36 \tag{1.2.21}$$

We set $\beta = 0.2$ for the Chaplygin gas system with a constant external force. Here the choice for β is arbitrary as long as no discontinuities move outside the numerical domain. [PHW18] provided with the exact solutions for both cases as well. For the contact discontinuity case, the exact solutions are

$$(\rho,u,H)(t,x) = \begin{cases} (\rho_-, u_- + \beta t, H_-), & x < x_1, \\ (\rho_{*1}, v_{*1} + \beta t, H_{*1}), & x_1 \leq x \leq x_2, \\ (\rho_{*2}, u_{*2} + \beta t, H_{*2}), & x_2 < x \leq x_3, \\ (\rho_+, u_+ + \beta t, H_+), & x > x_3, \end{cases} \tag{1.2.22}$$

where

$$\begin{cases} v_{*1} = v_{*2} = v_* = (u_+ + 1/\rho_+ + u_- - 1/\rho_-)/2, \\ \rho_{*1} = \rho_{*2} = \rho_* = 2/(u_+ + 1/\rho_+ - u_- + 1/\rho_-), \\ H_{*1} = 1/2/\rho_{*1} + \rho_{*1}(H_- - 1/2/\rho_-)/\rho_-, \\ H_{*2} = 1/2/\rho_{*2} + \rho_{*2}(H_+ - 1/2/\rho_+)/\rho_+. \end{cases} \tag{1.2.23}$$

The discontinuity locations are given by

$$\begin{cases} x_1(t) = \quad (u_- - 1/\rho_-)\,t + \beta t^2/2, \\ x_2(t) = \quad v_*\,t + \beta t^2/2, \\ x_3(t) = \quad (u_+ + 1/\rho_+)\,t + \beta t^2/2. \end{cases} \tag{1.2.24}$$

For the delta shock case, the exact solutions are

$$(\rho, u, H)(t,x) = \begin{cases} (\rho_-, u_-, H_-), & x < x(t), \\ (w(t)\delta(x - x(t)), v_\delta(t) + \beta t, h(t)\delta(x - x(t))), & x = x(t), \\ (\rho_+, u_+, H_+), & x > x(t), \end{cases} \tag{1.2.25}$$

where

$$
\begin{cases}
v_\delta(t) = \begin{cases} (u_- + u_+)\,t/2, & \rho_- = \rho_+ \\ \frac{\rho_+ u_+ - \rho_- u_- + \sqrt{\rho_-\rho_+\left((u_- - u_+)^2 - (1/\rho_- - 1/\rho_+)^2\right)}}{\rho_+ - \rho_-}, & \rho_- \neq \rho_+ \end{cases} \\
u_\delta(t) = v_\delta(t)\,t + \beta t^2/2, \\
w(t) = \big((\rho_+ - \rho_-)\,v_\delta(t) - \rho_+ u_+ + \rho_- u_-\big)\,t, \\
h(t) = \big(\big(u_- - v_\delta(t)\big)\big(\rho_-(u_- - v_\delta(t))^2/2 + H_- - 1/\rho_-\big) \\
\qquad\qquad + \big(v_\delta(t) - u_+\big)\big(\rho_+\big(v_\delta(t) - u_+\big)^2/2 + H_+ - 1/\rho_+\big)\big)\,t.
\end{cases}
\tag{1.2.26}
$$

As we can see that both cases have similar patterns as Pang's [Pan17] simulations for the case where the right hand sides of the system are zeros, as the external force β only contributes a little to the flow velocity. By taking $\beta = 0.2$ and the final time $T = 0.3$, the flow has a slightly bigger velocity to the right. We were careful to choose parameters that no important patterns would escape the numerical domain. Therefore, the two figures should look close.

We ran the shock case over $[-10, 10]$ until $T = 5$ to see the delta shock behavior actually keeps increasing in height. To be comparable with the previous result, we kept the same step size Δx, so the number of grid points is now $N + 1 = 2001$. We removed points within 0.1 of the shock location to calculate the errors and orders of accuracy.

Table 1.8: **Constant external force, $\beta = 0.2$.Chaplygin Contact Discontinuity with $|x - x_{cd}| \leq 0.1$ removed.**

Density ρ				
N	L_1 error	L_1 order	L_∞ error	L_∞ order
100	$1.1797e-03$	-	$9.1501e-03$	-
200	$9.1723e-05$	3.6850	$1.6412e-03$	2.4790
400	$2.2247e-06$	5.3656	$2.9372e-05$	5.8042
800	$1.7852e-07$	3.6395	$2.6422e-06$	3.4746
Momentum ρu				
100	$2.7066e-03$	-	$2.1802e-02$	-
200	$2.1333e-04$	3.6653	$3.9419e-03$	2.4675
400	$5.3014e-06$	5.3306	$7.0585e-05$	5.8034
800	$3.0528e-07$	4.1182	$6.6683e-06$	3.4040
Total Energy E				
100	$4.2791e-03$	-	$1.8216e-02$	-
200	$1.4007e-03$	1.6112	$6.9706e-03$	1.3858
400	$7.0066e-04$	0.9993	$3.0977e-03$	1.1701
800	$3.8790e-04$	0.8530	$1.1552e-03$	1.4231

Table 1.9: **Constant external force, $\beta = 0.2$. Chaplygin Shock with $|x - x_{shock}| \leq 0.05$ removed.**

Density ρ				
N	L_1 error	L_1 order	L_∞ error	L_∞ order
100	1.1111	-	18.1293	-
200	$1.1816e-01$	3.2332	4.1401	2.1306
400	$2.0742e-03$	5.8321	$1.5811e-01$	4.7107
800	$1.3039e-06$	10.6355	$1.7201e-04$	9.8442
Momentum ρu				
100	$3.3121e-01$	-	3.3575	-
200	$9.1975e-02$	1.8484	3.2183	0.0611
400	$1.2761e-03$	6.1714	$8.4590e-02$	5.2497
800	$9.8252e-07$	10.3430	$1.2955e-04$	9.3508
Total Energy E				
100	2.2601	-	37.2105	-
200	$2.5967e-01$	3.1216	9.6217	1.9513
400	$3.2656e-03$	6.3132	$2.4854e-01$	5.2748
800	$1.4188e-06$	11.1685	$1.9888e-04$	10.2874

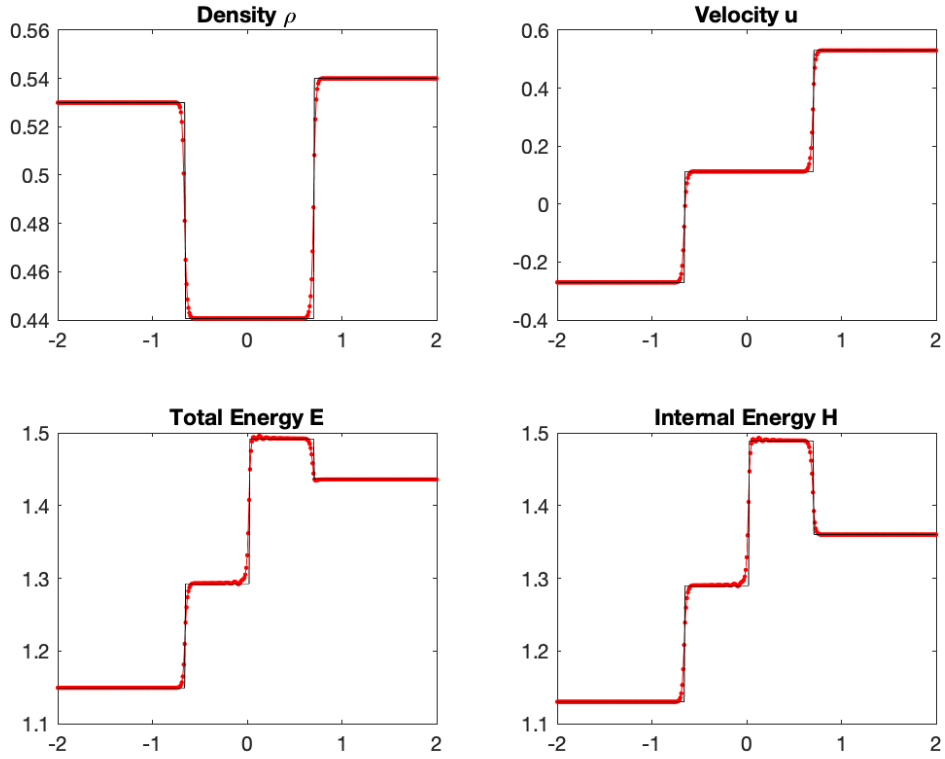


Figure 1.8: Constant external force, $\beta = 0.2$. Contact Discontinuity Solution with total number of grid points $N+1 = 401$. The final time $T = 0.3$. We see that there are two contact discontinuities in the density function ρ and the velocity function. There are three contact discontinuities in the internal energy function H and the total energy function E.

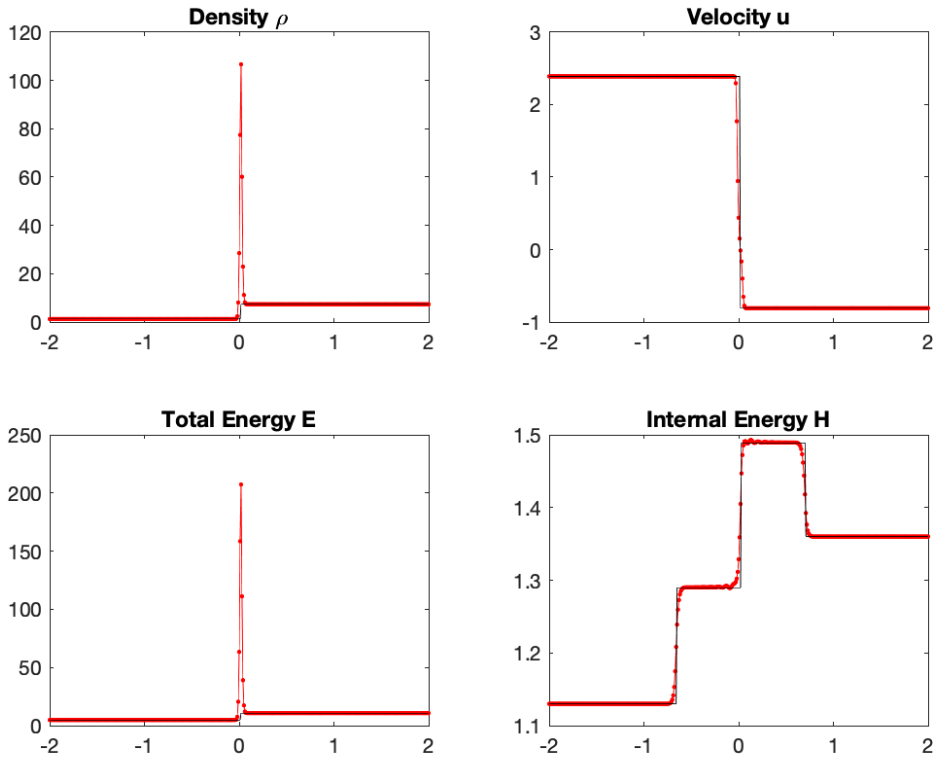


Figure 1.9: Constant external force, $\beta = 0.2$. Shock Solution with total number of grid points $N+1 = 401$. The final time $T = 0.3$. We see that there is a delta shock in the density function ρ, the internal energy function H and the total energy E. The velocity function u maintains a contact discontinuity.

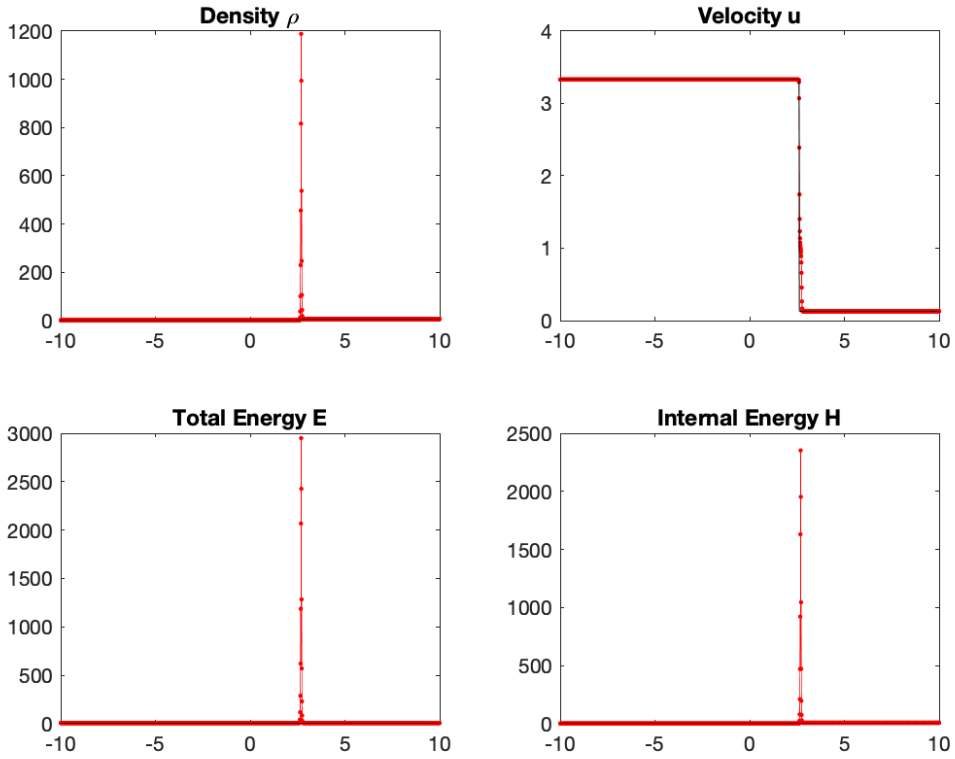


Figure 1.10: Constant external force, $\beta = 0.2$. Shock Solution with total number of grid points $N+1 = 2001$. The final time $T = 5$.

Table 1.10: **Constant external force, $\beta = 0.2$. Chaplygin Shock with $|x - x_{shock}| \leq 0.15$ removed.**

Density ρ				
N	L_1 error	L_1 order	L_∞ error	L_∞ order
500	5.3276	-	87.4650	-
1000	1.2059	2.1434	41.2943	1.0828
2000	$1.2234e-02$	6.6231	$9.2060e-01$	5.4872
4000	$7.9788e-08$	17.2263	$1.1346e-05$	16.3082
8000	$9.6128e-12$	13.0189	$4.8601e-12$	21.1546
Momentum ρu				
500	4.3521	-	75.4585	-
1000	$8.9045e-01$	2.2891	32.4080	1.2193
2000	$5.5040e-03$	7.3379	$4.3015e-01$	6.2354
4000	$2.8093e-08$	17.5799	$4.0764e-06$	16.6872
8000	$6.7887e-11$	8.6929	$7.4295e-12$	19.0656
Total Energy E				
500	11.6632	-	197.9859	-
1000	2.4887	2.2285	88.3179	1.1646
2000	$1.9407e-02$	7.0027	1.4894	5.8899
4000	$8.8316e-08$	17.7455	$1.2969e-05$	16.8094
8000	$6.8742e-11$	10.3273	$1.0969e-11$	20.1732

Chapter 2

Pressureless Euler Equations

The two dimensional zero-pressure Euler equations are governed by the following system

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x + \mathbf{g}(\mathbf{w})_y = \mathbf{0},\ t > 0,\ (x, y) \in \mathbb{R}^2, \tag{2.0.1}$$

$$\mathbf{w} = \begin{pmatrix} \rho \\ m \\ n \end{pmatrix},\ \mathbf{f}(\mathbf{w}) = \begin{pmatrix} \rho u \\ \rho u^2 \\ \rho uv \end{pmatrix},\ \text{and } \mathbf{g}(\mathbf{w}) = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 \end{pmatrix}$$

with

$$m = \rho u,\ n = \rho v,$$

where ρ is the density function and (u, v) is the velocity field. The above system (2.0.1) models the motion of free particles sticking together under collision and the formation of large-scale structures in the universe [AH94], [SZ89]. Note that we did not include the energy equation

$$E_t + (Eu)_x + (Ev)_y = 0 \tag{2.0.2}$$

in our system, because it does not affect the first three equations. Pang in [Pan19] studied the above system (2.0.1) with the energy equation (2.0.2) when the initial conditions were constants in each quadrant. Pang gave 23 diagrams to illustrate the density distributions under different initial data. The corresponding constraints regarding the initial data ρ_0, u_0 and v_0 were provided. There was no E_0 involved as it does not affect the density function ρ. [Pan19] did not provide solutions or diagrams regarding the velocity functions, u, v, or the energy function, E. In addition, we will be focusing on preserving the positivity property of the density function, ρ, and the maximum principle property of the velocity functions, u and v, so we decided not to include the energy equation. It can be solved easily especially in the componentwise approach. We simply add one more variable and one more function in the flux. Our goal in this chapter is to numerical solve the 23 cases and preserve the positivity of the density function, ρ and the maximum principle of the velocity functions, u and v. We apply the rectangular spatial discretization. Consider the domain $[a,b]\times[c,d]$, and the uniform step sizes

$$\Delta x = \frac{b-a}{N_x}, \text{ and } \Delta y = \frac{d-c}{N_y}. \tag{2.0.3}$$

The grid points are

$$(x_j, y_k) = \Big(a + (j-1)\Delta x, c + (k-1)\Delta y\Big) \tag{2.0.4}$$

for $j = 1,2,\cdots N_x+1$ and $k = 1,2,\cdots N_y+1$. The cell I_{jk} around (x_j, y_k) is

$$I_{jk} = (x_{j-1/2}, x_{j+1/2}) \times (y_{k-1/2}, y_{k+1/2}). \tag{2.0.5}$$

The system (2.0.1) is weakly hyperbolic, as the Jacobian matrices $\mathbf{f}'(\mathbf{w})$ and $\mathbf{g}'(\mathbf{w})$ do not have three linearly independent eigenvectors. Let's rewrite $\mathbf{f}(\mathbf{w})$ and $\mathbf{g}(\mathbf{w})$ in terms of ρ,

m and n as follows

$$\mathbf{f}(\mathbf{w}) = \begin{pmatrix} m \\ m^2/\rho \\ mn/\rho \end{pmatrix}, \text{ and } \mathbf{g}(\mathbf{w}) = \begin{pmatrix} n \\ mn/\rho \\ n^2/\rho \end{pmatrix} \tag{2.0.6}$$

and calculate the Jacobian matrices of the system by taking partial directives of $\mathbf{f}(\mathbf{w})$ and $\mathbf{g}(\mathbf{w})$ with respect to ρ, m and n. The matrices below are simplified.

$$\mathbf{f}'(\mathbf{w}) = \begin{pmatrix} 0 & 1 & 0 \\ -u^2 & 2u & 0 \\ -uv & v & u \end{pmatrix} \text{ and } \mathbf{g}'(\mathbf{w}) = \begin{pmatrix} 0 & 0 & 1 \\ -uv & v & u \\ -v^2 & 0 & 2v \end{pmatrix} \tag{2.0.7}$$

They are singular. $\mathbf{f}'(\mathbf{w})$ has a repeated eigenvalue u with multiplicity 3 and $\mathbf{g}'(\mathbf{w})$ has a repeated eigenvalue v with multiplicity 3. Both of the two matrices has only two linearly independent eigenvectors. This disadvantage leads to a failure of applying any characteristicwise numerical method. Some papers dealt with this issue by adding an artificial term in the flux functions to make the system strictly hyperbolic, i.e., to make the Jacobian matrices have three linearly independent eigenvectors. [Lev04] studied the one dimensional pressureless system (0.0.21) and discussed the technique of adding $(a^2\rho)_x$ in the second equation. The obtained system below is thus a strictly hyperbolic system with two distinct eigenvalues, $u-a$ and $u+a$, and two corresponding linearly independent eigenvectors.

$$\begin{cases} \rho_t + (\rho u)_x = 0, \\ (\rho u)_t + (\rho u^2 + a^2\rho)_x = 0, \end{cases} \tag{2.0.8}$$

where a is the sound speed. [Lev04] further found a weak solution to the one dimensional pressureless system (0.0.21) by assuming it is a solution to the above system

(2.0.8) as a vanishes. The same technique can be applied in the two dimensional pressureless system by adding the same flux term in the third equation

$$\begin{cases} \rho_t + (\rho u)_x + (\rho v)_y = 0, \\ (\rho u)_t + (\rho u^2 + a^2\rho)_x + (\rho uv)_y = 0, \\ (\rho v)_t + (\rho uv)_x + (\rho v^2 + a^2\rho)_y = 0. \end{cases} \tag{2.0.9}$$

The above system has three distinct eigenvalues $u-a$, u, and $u+a$ for $\mathbf{f}'(\mathbf{w})$, and three distinct eigenvalues $v-a$, v and $v+a$ for $\mathbf{g}'(\mathbf{w})$. Three distinct eigenvalues lead to three linearly independent eigenvectors. [Lev04] gave a finite volume method on the one dimensional pressureless system (0.0.21) without the parameter a and replaced the initial density values by $\rho = 10^{-20}$. The negative undershoots observed during the computation were replaced by $\rho = 10^{-20}$ as well. In this chapter, we will solve the pressureless system (2.0.1) with the finite difference WENO method componentwise and introduce a more robust method to preserve the positivity of the density function, ρ, and the boundedness of the velocities, u and v. Our method was inspired by the limiters introduced in [ZS12], [YWS13]. We did not utilize the parameter a because it violates the convex set designed in Section 2.1.1. Moreover, from our experience in Case 6. and Case 7. from the previous chapter, the componentwise approach is more straight forward to code, and more efficient than the characteristic approach when comes to computational and memory costs, as no matrix multiplication is involved. It turned out that both approaches gave similarly good results on the errors and orders of accuracy.

2.1 Finite Difference WENO Method

For any time step $t = t_n$, we apply the Lax-Friedrichs flux splitting method [Shu98], [ZS12]

$$\mathbf{f}^{\pm}(\mathbf{w}^n_{jk}) = \frac{1}{2}\Big(\mathbf{w}^n_{jk} \pm \frac{\mathbf{f}(\mathbf{w}^n_{jk})}{\alpha_x}\Big),\ \alpha_x = \max_{j,k}|u^n_{jk}|, \tag{2.1.1}$$

$$\mathbf{g}^{\pm}(\mathbf{w}^n_{jk}) = \frac{1}{2}\Big(\mathbf{w}^n_{jk} \pm \frac{\mathbf{g}(\mathbf{w}^n_{jk})}{\alpha_y}\Big),\ \alpha_y = \max_{j,k}|v^n_{jk}|. \tag{2.1.2}$$

$$\tag{2.1.3}$$

The above splitting process regards the fluid at (x_j, y_k) as an ensemble of particles [Van91]. Measured along the x axis, some particles move to the right to form the flow $\mathbf{f}^+(\mathbf{w}^n_{jk})$, others to the left to form the flow $\mathbf{f}^-(\mathbf{w}^n_{jk})$. Along the y axis, we also have the flow $\mathbf{g}^+(\mathbf{w}^n_{jk})$ heading up and the flow $\mathbf{g}^-(\mathbf{w}^n_{jk})$ heading down.
[Shu98] and [ZS10] pointed out that at each time level $t = t_n$, there exists two vectors of functions $\mathbf{h}_\pm$ whose cell average values inside the cell

$$I_{jk} = [x_{j-1/2}, x_{j+1/2}] \times [y_{k-1/2}, y_{k+1/2}] \tag{2.1.4}$$

are $\mathbf{f}^{\pm}(\mathbf{w}^n_{jk})$.That is to say, $(\overline{\mathbf{h}}_\pm)^n_{jk} = \mathbf{f}^{\pm}(\mathbf{w}^n_{jk})$. The $\pm$ signs in $\mathbf{h}$ correspond to those in $\mathbf{f}$. They denote the directions along the x axis which the fluxes are heading to at the grid point x_{jk}. For each j, we perform the one dimensional WENO reconstruction introduced in [LOC94] on each component of $(\overline{\mathbf{h}}_+)^n_{ik}$, $i = j-2, \cdots, j+2$ in the x-direction to obtain approximations of the point values $(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}$, where the + sign inside the parentheses denotes the flux is heading to the right, and the $-$ sign denote that the approximated value at the interface $(x_{j+1/2}, y_k)$ is based on the polynomial reconstructed in the cell I_{jk} from the left hand side. We continue applying the one dimensional WENO

reconstruction on each component of $(\overline{\mathbf{h}}_-)^n_{ik}$, $i = j-1,\cdots,j+3$ to obtain $(\mathbf{h}_-)^{+,n}_{j+\frac{1}{2},k}$. This time, the flux is heading to the left from the cell $I_{j+1,k}$ from the right. Therefore, we have − sign inside the parentheses and + outside. The superscripts − and + denote approximations within the cells $I_{j,k}$ and $I_{j+1,k}$, respectively.

We apply the same WENO reconstruction process along the y axis. [Shu98] and [ZS10] proved the existence of two vectors of functions $\mathbf{q}_\pm$ whose cell averages values inside the cell I_{jk} are $\mathbf{g}^\pm(\mathbf{w}^n_{jk})$. We identify $(\overline{\mathbf{q}}_\pm)^n_{jk} = \mathbf{g}^\pm(\mathbf{w}^n_{jk})$. For each k, we perform the WENO reconstruction again on the WENO stencil $(\overline{\mathbf{q}}_\pm)^n_{ji}$, $i = k-2,\cdots,k+2$ in the y-direction to obtain $(\mathbf{q}_+)^{-,n}_{j,k+\frac{1}{2}}$, where the + sign inside the parentheses denotes the flux is heading upwards, and the − sign denote that the approximated value at the interface $(x_j, y_{k+1/2})$ is based on the polynomial reconstructed in the cell $I_{j,k}$ from below. We next perform the WENO reconstruction on the stencil $(\overline{\mathbf{q}}_-)^{+,n}_{ji}$, i = $k-1,\cdots,k+3$ to estimate $(\mathbf{q}_-)^{+,n}_{j,k+\frac{1}{2}}$. This time, the flux is heading down to the interface $(x_j, y_{k+1/2})$ from the cell $I_{j,k+1}$ from above. Thus we have the − sign inside the parentheses and the + sign outside.

Form the fluxes by

$$\hat{\mathbf{f}}_{j+\frac{1}{2},k} = \alpha_x\Big((\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k} - (\mathbf{h}_-)^{+,n}_{j+\frac{1}{2},k}\Big) \tag{2.1.5}$$

$$\hat{\mathbf{g}}_{j,k+\frac{1}{2}} = \alpha_y\Big((\mathbf{q}_+)^{-,n}_{j,k+\frac{1}{2}} - (\mathbf{q}_-)^{+,n}_{j,k+\frac{1}{2}}\Big) \tag{2.1.6}$$

In this dissertation, we applied the semi-discrete scheme

$$\frac{d}{dt}\mathbf{w}_{jk}(t) = -\frac{1}{\Delta x}\Big(\hat{\mathbf{f}}_{i+\frac{1}{2},j} - \hat{\mathbf{f}}_{i-\frac{1}{2},j}\Big) - \frac{1}{\Delta y}\Big(\hat{\mathbf{g}}_{j,k+\frac{1}{2}} - \hat{\mathbf{g}}_{j,k-\frac{1}{2}}\Big) \tag{2.1.7}$$

together with TVD-RK3 [SO89], [LOC94] introduced in (2.1.4) to achieve 3th order accuracy in in time. The right hand side of the semi-discrete scheme above gives a 4th order approximation to $-\mathbf{f}(\mathbf{w})_x - \mathbf{g}(\mathbf{w})_y$. Theoretically, the TVD RK3 time integration

can only give a 3rd order scheme. But in practice, the spatial reconstruction order is more crucial. We will see that in Section 2.2.1 that our method achieves higher than 4th order of accuracy even though we only apply RK3 as time integration.

2.1.1 Convex Set

So far, this two dimensional finite difference method is very much alike the one dimensional method we used in Chapter 1, except that there is one more dimension in WENO spatial reconstruction along the y axis. The key difference is that the pressureless system produces vacuum areas where the density function $\rho = 0$. Analytically, the velocity field (u, v) is undefined in those areas. But in numerical analysis, we cannot accept the density to be zero. Otherwise, the program cannot continue running as we need to calculate the velocities u and v in each time step to determine the next time step size. [Hu00] pointed out that for pressureless Euler equations, $\rho(x, y, t)$ could be unbounded on some planes, also known as shock waves, while $u(x, y, t)$ and $v(x, y, t)$ are always bounded on $\mathbb{R}^2 \times \mathbb{R}_+$. We thus define the following admissible set

$$G = \left\{ \mathbf{w} = (\rho, \ \ m, \ \ n)^T : \rho \geq 0, m^2 + n^2 \leq S^2 \rho^2 \right\} \tag{2.1.8}$$

where

$$S > 0, \text{ and } S^2 = \max_{x,y} \left\{ u^2(x, y, 0) + v^2(x, y, 0) \right\}$$

with $(u, v)(x, y, 0)$ being the initial velocity field. This convex set preserves the positivity of the density function and the maximum principle of the velocity functions. Note that this convex set is modified from that in [YWS13]. From the definition of S, it is clear that

$$\mathbf{w}_{jk}^0 \in G, \forall\, j, k. \tag{2.1.9}$$

Our goal is to maintain

$$\mathbf{w}_{jk}^{n+1} \in G, \forall\, j,k, \tag{2.1.10}$$

given that

$$\mathbf{w}_{jk}^{n} \in G, \forall\, j,k,\ n>0. \tag{2.1.11}$$

At any time level $t=t_n$, for every j and k, the following lemma holds.

Lemma 2.1.1. *If* $\mathbf{w}_{jk}^{n} \in G$, *then* $\mathbf{f}^{\pm}(\mathbf{w}_{jk}^{n}) \in G$ *and* $\mathbf{g}^{\pm}(\mathbf{w}_{jk}^{n}) \in G$.

Proof. Let's look at $\mathbf{f}^{\pm}(\mathbf{w}_{jk}^{n})$ first. We denote

$$\mathbf{f}^{\pm}(\mathbf{w}_{jk}^{n}) = \begin{pmatrix} \rho_f^{\pm} \\ m_f^{\pm} \\ n_f^{\pm} \end{pmatrix}_{jk}^{n} = \frac{\rho_{jk}^{n}}{2} \begin{pmatrix} 1 \pm \frac{u_{jk}^{n}}{\alpha_x} \\ u_{jk}^{n}(1 \pm \frac{u_{jk}^{n}}{\alpha_x}) \\ v_{jk}^{n}(1 \pm \frac{u_{jk}^{n}}{\alpha_x}) \end{pmatrix} \tag{2.1.12}$$

Because

$$\left((m_f^{\pm})_{jk}^{n}\right)^2 + \left((n_f^{\pm})_{jk}^{n}\right)^2 = \frac{(\rho_{jk}^{n})^2}{4}(1 \pm \frac{u_{jk}^{n}}{\alpha_x})^2\left((u_{jk}^{n})^2 + (v_{jk}^{n})^2\right) \le S^2\left((\rho_f^{\pm})_{jk}^{n}\right)^2 \tag{2.1.13}$$

We have proven that $\mathbf{f}^{\pm}(\mathbf{w}_{jk}^{n}) \in G$.

The same strategy applies to $\mathbf{g}^{\pm}(\mathbf{w}_{jk}^{n})$. Let's denote

$$\mathbf{g}^{\pm}(\mathbf{w}_{jk}^{n}) = \begin{pmatrix} \rho_g^{\pm} \\ m_g^{\pm} \\ n_g^{\pm} \end{pmatrix}_{jk}^{n} = \frac{\rho_{jk}^{n}}{2} \begin{pmatrix} 1 \pm \frac{v_{jk}^{n}}{\alpha_y} \\ u_{jk}^{n}(1 \pm \frac{v_{jk}^{n}}{\alpha_y}) \\ v_{jk}^{n}(1 \pm \frac{v_{jk}^{n}}{\alpha_y}) \end{pmatrix} \tag{2.1.14}$$

Becasue

$$\left((m_g^{\pm})_{jk}^n\right)^2+\left((n_g^{\pm})_{jk}^n\right)^2=\frac{(\rho_{jk}^n)^2}{4}(1\pm\frac{v_{jk}^n}{\alpha_y})^2\left((u_{jk}^n)^2+(v_{jk}^n)^2\right)\leq S^2\left((\rho_g^{\pm})_{jk}^n\right)^2 \tag{2.1.15}$$

We have proven that $\mathbf{g}^{\pm}(\mathbf{w}_{jk}^n)\in G$.

□

2.1.2 Convex Combination

Denoting $\lambda_x=\frac{\Delta t}{\Delta x}$ and $\lambda_y=\frac{\Delta t}{\Delta y}$, the fluxes (2.1.5) and (2.1.6) imply the following forward Euler method

$$\begin{aligned}\mathbf{w}_{jk}^{n+1}=&\mathbf{w}_{jk}^n-\lambda_x\left(\hat{\mathbf{f}}_{j+\frac{1}{2},k}-\hat{\mathbf{f}}_{j-\frac{1}{2},k}\right)-\lambda_y\left(\hat{\mathbf{g}}_{j,k+\frac{1}{2}}-\hat{\mathbf{g}}_{j,k-\frac{1}{2}}\right)\\=&\frac{1}{2}\left((\overline{\mathbf{h}}_+)_{jk}^n+(\overline{\mathbf{h}}_-)_{jk}^n+(\overline{\mathbf{q}}_+)_{jk}^n+(\overline{\mathbf{q}}_-)_{jk}^n\right)\\&-\alpha_x\lambda_x\left((\mathbf{h}_+)_{j+1/2,k}^{-,n}-(\mathbf{h}_-)_{j+1/2,k}^{+,n}+(\mathbf{h}_+)_{j-1/2,k}^{-,n}-(\mathbf{h}_-)_{j-1/2,k}^{+,n}\right)\\&-\alpha_y\lambda_y\left((\mathbf{q}_+)_{j,k+1/2}^{-,n}-(\mathbf{q}_-)_{j,k+1/2}^{+,n}+(\mathbf{q}_+)_{j,k-1/2}^{-,n}-(\mathbf{q}_-)_{j,k-1/2}^{+,n}\right)\\=&\mathbf{H}^+ +\mathbf{H}^- +\mathbf{Q}^+ +\mathbf{Q}^-,\end{aligned} \tag{2.1.16}$$

where

$$\mathbf{H}^+=\frac{1}{2}(\overline{\mathbf{h}}_+)_{jk}^n-\alpha_x\lambda_x\left((\mathbf{h}_+)_{j+1/2,k}^{-,n}-(\mathbf{h}_+)_{j-1/2,k}^{-,n}\right) \tag{2.1.17}$$

$$\mathbf{H}^-=\frac{1}{2}(\overline{\mathbf{h}}_-)_{jk}^n-\alpha_x\lambda_x\left((\mathbf{h}_-)_{j+1/2,k}^{+,n}-(\mathbf{h}_-)_{j-1/2,k}^{+,n}\right) \tag{2.1.18}$$

$$\mathbf{Q}^+=\frac{1}{2}(\overline{\mathbf{q}}_+)_{jk}^n-\alpha_y\lambda_y\left((\mathbf{q}_+)_{j,k+1/2}^{-,n}-(\mathbf{q}_+)_{j,k-1/2}^{-,n}\right) \tag{2.1.19}$$

$$\mathbf{Q}^-=\frac{1}{2}(\overline{\mathbf{q}}_-)_{jk}^n-\alpha_y\lambda_y\left((\mathbf{q}_-)_{j,k+1/2}^{+,n}-(\mathbf{q}_-)_{j,k-1/2}^{+,n}\right) \tag{2.1.20}$$

The finite difference WENO scheme is now written as the sum of four finite volume schemes [ZS12]. Let's study (2.1.17) first. For a fixed k, the one dimensional WENO

Table 2.1: 4-point Gauss-Lobatto quadrature rule on $[x_{j-1/2}, x_{j+1/2}]$

α	$\hat{x}_j^\alpha$	$\hat{\omega}_\alpha$
1	$x_{j-1/2}$	$\frac{1}{6}$
2	$x_{j-1/2} + \left(1 - \frac{1}{\sqrt{5}}\right)\frac{\Delta x}{2}$	$\frac{5}{6}$
3	$x_{j+1/2} - \left(1 - \frac{1}{\sqrt{5}}\right)\frac{\Delta x}{2}$	$\frac{5}{6}$
4	$x_{j+1/2}$	$\frac{1}{6}$

spatial reconstruction along the x axis gives a quadratic function $\mathbf{p}_j^{k,H^+}(x)$ defined in (1.1.35) over $[x_{j-1/2}, x_{j+1/2}]$ satisfying

$$\mathbf{p}_j^{k,H^+}(x_{j+1/2}) = (\mathbf{h}_+)_{j+1/2,k}^{-,n}, \tag{2.1.21}$$

and

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{p}_j^{k,H^+}(x)dx = (\overline{\mathbf{h}}_+)_{jk}^n. \tag{2.1.22}$$

(2.1.21) is easy to understand as $(\mathbf{h}_+)_{j+1/2,k}^{-,n}$ was obtained by evaluating the polynomials from WENO reconstruction at the interface $x_{j+1/2}$ shown in (1.1.42). (2.1.22) is the principle behind WENO reconstruction (1.1.31). $\mathbf{p}_j^{k,H^+}(x)$ is a 3-th order accurate approximation to the function $(\mathbf{h}_+)^{-,n}(x)$ on $[x_{j-1/2}, x_{j+1/2})$ and 4th order at $x = x_{j+1/2}$, if $\mathbf{w}^n$ is smooth. The exactness of an N-point Gauss-Lobatto quadrature rule [ZS10], [ZS12] gives.

$$(\overline{\mathbf{h}}_+)_{jk}^n = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{p}_j^{k,H^+}(x)dx = \sum_{\alpha=1}^{N} \hat{\omega}_\alpha \mathbf{p}_j^{k,H^+}(\hat{x}_j^\alpha). \tag{2.1.23}$$

Here an $N = 4$ point Gauss-Lobatto quadrature rule is applied required by a 5-cell WENO reconstruction [ZS12], [ZS10]. $\hat{x}_j^\alpha$ are the nodes and $\hat{\omega}_\alpha$ are the corresponding weights

Denoting

$$\mathbf{p}_j^{*,k,H^+} = \frac{1}{1-\hat{\omega}_N}\sum_{\alpha=1}^{N-1}\hat{\omega}_\alpha \mathbf{p}_j^{k,H^+}(\hat{x}_j^\alpha), \tag{2.1.24}$$

we rewrite (2.1.17) as follows

$$\mathbf{H}^+ = \frac{1}{2}(1-\hat{\omega}_N)\mathbf{p}_j^{*,k,H^+} + (\frac{1}{2}\hat{\omega}_N - \alpha_x\lambda_x)(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k} + \alpha_x\lambda_x(\mathbf{h}_+)^{-,n}_{j-\frac{1}{2},k}. \tag{2.1.25}$$

Therefore, if $\mathbf{p}_j^{*,k,H^+}$, $(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}$ and $(\mathbf{h}_+)^{-,n}_{j-\frac{1}{2},k}$ are inside the admissible set G, then $\mathbf{H}^+ \in G$ under the CFL condition

$$\alpha_x\lambda_x \le \frac{1}{2}\hat{\omega}_N. \tag{2.1.26}$$

Similarly, for (2.1.18), there exists a vector of polynomials of degree 2, $\mathbf{p}_j^{k,H^-}(x)$ such that

$$\mathbf{p}_j^{k,H^-}(x_{j-\frac{1}{2}}) = (\mathbf{h}_-)^{+,n}_{j-\frac{1}{2},k}, \tag{2.1.27}$$

and

$$\frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{p}_j^{k,H^-}(x)dx = (\overline{\mathbf{h}}_-)^n_{jk}. \tag{2.1.28}$$

$\mathbf{p}_j^{k,H^-}$ is a 3th order accurate approximation to the function $(\mathbf{h}_-)^{+,n}(x)$ on $(x_{j-1/2}, x_{j+1/2}]$ and 4th order at $x = x_{j-1/2}$, if $\mathbf{w}^n$ is smooth. Let

$$\mathbf{p}_j^{*,k,H^-} = \frac{1}{1-\hat{\omega}_1}\sum_{\alpha=2}^{N}\hat{\omega}_\alpha \mathbf{p}_j^{k,H^-}(\hat{x}_j^\alpha), \tag{2.1.29}$$

where $\hat{x}_j^\alpha$ and $\hat{\omega}_\alpha$ are defined in Table 2.1. Then (2.1.18) becomes

$$\mathbf{H}^- = \frac{1}{2}(1-\hat{\omega}_1)\mathbf{p}_j^{*,k,H^-} + (\frac{1}{2}\hat{\omega}_1 - \alpha_x\lambda_x)(\mathbf{h}_-)^{+,n}_{j-\frac{1}{2},k} + \alpha_x\lambda_x(\mathbf{h}_-)^{+,n}_{j+\frac{1}{2},k}. \tag{2.1.30}$$

Table 2.2: 4-point Gauss-Lobatto quadrature rule on $[y_{k-1/2}, y_{k+1/2}]$

α	$\hat{y}_k^\alpha$	$\hat{\omega}_\alpha$
1	$y_{k-1/2}$	$\frac{1}{6}$
2	$y_{k-1/2} + \left(1 - \frac{1}{\sqrt{5}}\right)\frac{\Delta y}{2}$	$\frac{5}{6}$
3	$y_{k+1/2} - \left(1 - \frac{1}{\sqrt{5}}\right)\frac{\Delta y}{2}$	$\frac{5}{6}$
4	$y_{k+1/2}$	$\frac{1}{6}$

Therefore, if $\mathbf{p}_j^{*,k,H^-}$, $(\mathbf{h}_-)^{+,n}_{j+\frac{1}{2},k}$ and $(\mathbf{h}_-)^{+,n}_{j-\frac{1}{2},k}$ are inside the admissible set G, then $\mathbf{H}^- \in G$ under the CFL condition

$$\alpha_x \lambda_x \leq \frac{1}{2}\hat{\omega}_1. \tag{2.1.31}$$

Note that $\hat{\omega}_1 = \hat{\omega}_N$ for Gaussian quadratures, (2.1.26) and (2.1.31) are the same CFL condition in the x direction.

Next, we apply the same strategy in the y direction. For each j, the WENO reconstruction in the y direction gives a vector of quadratic functions $\mathbf{p}_k^{j,Q^+}$ defined inside $[y_{k-1/2}, y_{k+1/2}]$ satisfying

$$\mathbf{p}_k^{j,Q^+}(y_{k-1/2}) = (\mathbf{q}_+)^{-,n}_{j,k+1/2}, \tag{2.1.32}$$

and

$$\frac{1}{\Delta y}\int_{y_{k-1/2}}^{y_{k+1/2}} \mathbf{p}_k^{j,Q^+}(y)\,dy = (\overline{\mathbf{q}}_+)^n_{jk}. \tag{2.1.33}$$

$\mathbf{p}_k^{j,Q^+}$ is a 3rd order approximation to $(\mathbf{q}_+)^{-,n}(y)$ on $[y_{k-1/2}, y_{k+1/2})$ and 4th order at $y = y_{k+1/2}$. Let's denote

$$\mathbf{p}_k^{*,j,Q^+} = \frac{1}{1-\hat{\omega}_N}\sum_{\alpha=1}^{N-1} \hat{\omega}_\alpha \mathbf{p}_k^{j,Q^+}(\hat{y}_k^\alpha), \tag{2.1.34}$$

where $\hat{y}_k^\alpha$ and $\hat{\omega}_\alpha$, given in Table 2.2 are the integration points and the corresponding weights, respectively. (2.1.19) thus becomes

$$\mathbf{Q}^+ = \frac{1}{2}(1-\hat{\omega}_N)\mathbf{p}_k^{*,j,Q^+} + (\frac{1}{2}\hat{\omega}_N - \alpha_y\lambda_y)(\mathbf{q}_+)^{-,n}_{j,k+\frac{1}{2}} + \alpha_y\lambda_y(\mathbf{q}_+)^{-,n}_{j,k-\frac{1}{2}}. \tag{2.1.35}$$

If $\mathbf{p}_k^{*,j,Q^+}$, $(\mathbf{q}_+)^{-,n}_{j,k+\frac{1}{2}}$ and $(\mathbf{q}_+)^{-,n}_{j,k-\frac{1}{2}}$ are inside G, then $\mathbf{Q}^+ \in G$ if

$$\alpha_y\lambda_y \le \frac{1}{2}\hat{\omega}_N. \tag{2.1.36}$$

It leaves (2.1.20) to finish this section. For every j, there exists a vector of quadratic functions $\mathbf{p}_k^{j,Q^-}(y)$ on $[y_{k-1/2}, y_{k+1/2}]$ satisfying

$$\mathbf{p}_k^{j,Q^-}(y_{k-1/2}) = (\mathbf{q}_-)^{+,n}_{j,k-1/2}, \tag{2.1.37}$$

and

$$\frac{1}{\Delta y}\int_{y_{k-1/2}}^{y_{k+1/2}} \mathbf{p}_k^{j,Q^-}(y)dy = (\overline{\mathbf{q}}_-)^n_{jk}. \tag{2.1.38}$$

$\mathbf{p}_k^{j,Q^-}(y)$ is a 3rd order approximation of $(\mathbf{q}_-)^{+,n}(y)$ on $(y_{k-1/2}, y_{k+1/2}]$, and 4th order at $y = y_{k-1/2}$. Let's denote

$$\mathbf{p}_k^{*,j,Q^-} = \frac{1}{1-\hat{\omega}_1}\sum_{\alpha=2}^{N}\hat{\omega}_\alpha\mathbf{p}_k^{j,Q^-}(\hat{y}_k^\alpha), \tag{2.1.39}$$

where $\hat{y}_k^\alpha$ and $\hat{\omega}_\alpha$ are given in Table 2.2. (2.1.20) becomes

$$\mathbf{Q}^- = \frac{1}{2}(1-\hat{\omega}_1)\mathbf{p}_k^{*,j,Q^-} + (\frac{1}{2}\hat{\omega}_1 - \alpha_y\lambda_y)(\mathbf{q}_-)^{+,n}_{j,k+\frac{1}{2}} + \alpha_y\lambda_y(\mathbf{q}_-)_{j,k-\frac{1}{2}}. \tag{2.1.40}$$

If $\mathbf{p}_k^{*,j,Q^-}$, $(\mathbf{q}_-)^{+,n}_{j,k+\frac{1}{2}}$ and $(\mathbf{q}_-)_{j,k-\frac{1}{2}}$ are all inside the admissible set G, then $\mathbf{Q}^- \in G$ if

$$\alpha_y\lambda_y \le \frac{1}{2}\hat{\omega}_1. \tag{2.1.41}$$

Note that because $\hat{\omega}_1 = \hat{\omega}_N$, (2.1.36) and (2.1.41) give the same CFL condition in the y direction. Combining the CFL conditions in the x (2.1.26) and y (2.1.36) directions, the finalized CFL condition is as follows

$$\alpha_x \lambda_x \leq \frac{1}{12}, \text{ and } \alpha_y \lambda_y \leq \frac{1}{12}, \tag{2.1.42}$$

since $\hat{\omega}_N = \frac{1}{6}$ in a four-point Gauss-Lobatto quadrature rule.

We have shown in Section 2.1.1 that if $\mathbf{w}_{jk}^n \in G$, then $\mathbf{f}^{\pm}(\mathbf{w}_{jk}^n) \in G$ and $\mathbf{g}^{\pm}(\mathbf{w}_{jk}^n) \in G$. In Section 2.1.2, we have also concluded that if the flux functions evaluated at the interfaces $(\mathbf{h}_+)_{j\pm1/2,k}^{-,n}$, $(\mathbf{h}_-)_{j\pm1/2,k}^{+,n}$, $(\mathbf{q}_+)_{j,k\pm1/2}^{-,n}$, and $(\mathbf{q}_-)_{j,k\pm1/2}^{+,n}$, and the weighted average values inside the cell I_{jk} except one end point, $\mathbf{p}_j^{*,k,H^+}$, $\mathbf{p}_j^{*,k,H^-}$, $\mathbf{q}_k^{*,j,Q^+}$, and $\mathbf{q}_k^{*,j,Q^-}$, are all inside the admissible G, then $\mathbf{w}_{jk}^{n+1}$ will be inside G, given that $\mathbf{w}_{jk}^n \in G$. We only need to prove that if $\mathbf{f}^{\pm}(\mathbf{w}_{jk}^n)$ and $\mathbf{g}^{\pm}(\mathbf{w}_{jk}^n)$ are both inside G, then the flux functions evaluated at the interfaces $(\mathbf{h}_+)_{j\pm1/2,k}^{-,n}$, $(\mathbf{h}_-)_{j\pm1/2,k}^{+,n}$, $(\mathbf{q}_+)_{j,k\pm1/2}^{-,n}$, and $(\mathbf{q}_-)_{j,k\pm1/2}^{+,n}$, and the weighted average values inside the cell I_{jk} except one end point, $\mathbf{p}_j^{*,k,H^+}$, $\mathbf{p}_j^{*,k,H^-}$, $\mathbf{q}_k^{*,j,Q^+}$, and $\mathbf{q}_k^{*,j,Q^-}$, are all inside the admissible G. By then, we will obtain that if $\mathbf{w}_{jk}^n \in G$, then $\mathbf{w}_{jk}^{n+1} \in G$. A limiter is introduced in the following section to fill the last gap.

2.1.3 Limiter

Let's first consider $(\overline{\mathbf{h}}_+)_{jk}^n$, $(\mathbf{h}_+)_{j+1/2,k}^{-,n}$ and $\mathbf{p}_j^{*,k,H^+}$ as a group. $(\overline{\mathbf{h}}_+)_{jk}^n \in G$ because

$$(\overline{\mathbf{h}}_+)_{jk}^n = \mathbf{f}^+(\mathbf{w}_{jk}^n) \in G. \tag{2.1.43}$$

Recall that

$$(\overline{\mathbf{h}}_+)_{jk}^n = (1-\hat{\omega}_N)\mathbf{p}_j^{*,k,H^+} + \hat{\omega}_N(\mathbf{h}_+)_{j+1/2,k}^{-,n}. \tag{2.1.44}$$

Let's start modifying the density function. Our limiter is modified from [ZS11b], [ZS10] and [YWS13].

Density:

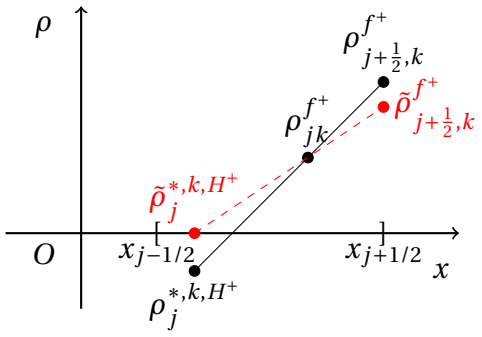
- Denote $\rho^{f^+}_{jk}$, $\rho^{f^+}_{j+\frac{1}{2},k}$, and ρ^{*,k,H^+}_{j} to be the first components of $\bar{\mathbf{h}}^{n,+}_{jk}$, $(\mathbf{h}_+)^{-,n}_{j+1/2,k}$, and $\mathbf{p}^{*,k,H^+}_{j}$, respectively.

- For each cell I_{jk}, modify the density flux in the x-direction by setting:

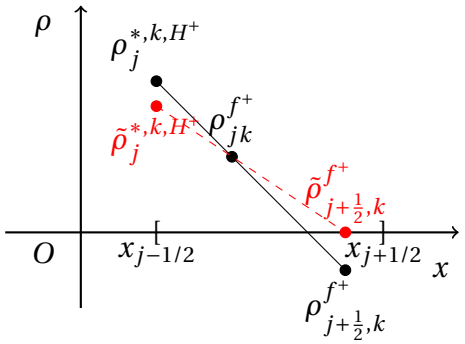
$$\tilde{\rho}^{f^+}_{j+\frac{1}{2},k} = \theta_\rho\left(\rho^{f^+}_{j+\frac{1}{2},k} - \rho^{f^+}_{jk}\right) + \rho^{f^+}_{jk},\ \theta_\rho = \min\left\{\frac{\rho^{f^+}_{jk}}{\rho^{f^+}_{jk} - \rho^{1}_{\min}}, 1\right\}, \tag{2.1.45}$$

where $\rho^1_{\min} = \min\left\{\rho^{f^+}_{j+\frac{1}{2},k}, \rho^{*,k,H^+}_{j}\right\}$.

Notice that $\frac{\rho^{f^+}_{jk}}{\rho^{f^+}_{jk} - \rho^1_{\min}}$ is nonnegative, and hence, $0 \le \theta_\rho \le 1$. Note that θ_ρ can take 0 since $\rho^{f^+}_{jk}$ can be zero. The following two diagrams illustrate the cases when one of $\rho^{f^+}_{j+\frac{1}{2},k}$ and ρ^{*,k,H^+}_{j} is negative, we can adjust the behaviors by introducing a parameter θ_ρ. $\rho^{f^+}_{j+\frac{1}{2},k}$ and ρ^{*,k,H^+}_{j} cannot be both negative, as their weighed average $\rho^{f^+}_{jk} \ge 0$.
If $\rho^{*,k,H^+}_{j} < 0$:



If $\rho^{f^+}_{j+\frac{1}{2},k} < 0$:



The above limiter for the density function is modified from the positivity preserving limiter in [ZS11b], [ZS10] and [YWS13]. Their limiter was applied on the cell average values of the density function. They thus introduced a small parameter $\varepsilon = 10^{-13}$ and defined θ_ρ as follows

$$\min\left\{\frac{\rho^{f^+}_{jk} - \varepsilon}{\rho^{f^+}_{jk} - \rho^1_{\min}}, 1\right\}. \tag{2.1.46}$$

to make sure that the density values are not less than 10^{-13}. Their corresponding admissible set was as follows

$$G^\varepsilon = \left\{\mathbf{w} = (\rho,\ m,\ n)^T : \rho \geq \varepsilon, m^2 + n^2 \leq (S+\varepsilon)^2\rho^2\right\} \tag{2.1.47}$$

where

$$S > 0, \text{ and } S^2 = \max_{x,y}\left\{u^2(x,y,0) + v^2(x,y,0)\right\}.$$

Since our limiter is applied on the split fluxes, where zero flux is acceptable, we deleted the ε part in the θ_ρ and G^ε definitions. Moreover, the split fluxes $\mathbf{f}^\pm(\mathbf{w})^n_{jk}$ and $\mathbf{g}^\pm(\mathbf{w})^n_{jk}$ fail to be inside G^ε when the split fluxes are zero.

Velocity:

Now let's modify the velocity functions [YWS13] so that the maximum principle is satisfied.

- If $(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}$ with the updated $\tilde{\rho}^{f^+}_{j+\frac{1}{2},k}$ is inside the admissible set G, then $\theta_1 = 1$. Otherwise, let

$$\theta_1 = \frac{||(\overline{\mathbf{h}}_+)^n_{jk} - \mathbf{s}^1_{jk}||}{||(\overline{\mathbf{h}}_+)^{-,n}_{jk} - (\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}||} \tag{2.1.48}$$

  where $||\cdot||$ is the Euclidean norm, and $\mathbf{s}^1_{jk}$ is the intersection point of the straight line

$$\mathbf{s}^1(t) = (1-t)(\overline{\mathbf{h}}_+)^n_{jk} + t(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}, \quad 0 \le t \le 1 \tag{2.1.49}$$

  and the surface of the admissible set G

$$\partial G = \left\{ \mathbf{w} = (\rho,\ m,\ n)^T : \rho = 0, m^2 + n^2 \le S^2\rho^2 \right\}. \tag{2.1.50}$$

- Similarly, if $\mathbf{p}^{*,k,H^+}_j$ with the updated $\tilde{\rho}^{*,k,H^+}_j$ is inside G, then $\theta_2 = 1$. Otherwise, we can define

$$\theta_2 = \frac{||(\overline{\mathbf{h}}_+)^n_{jk} - \mathbf{s}^2_{jk}||}{||(\overline{\mathbf{h}}_+)^{-,n}_{jk} - \mathbf{p}^{*,k,H^+}_j||} \tag{2.1.51}$$

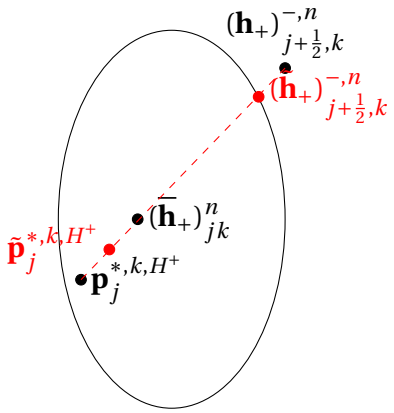
  where $\mathbf{s}^2_{ij}$ is the intersection point of the straight line

$$\mathbf{s}^2(t) = (1-t)\overline{\mathbf{h}}^{n,+}_{ij} + t\mathbf{p}^{*,j,H^+}_i, \quad 0 \le t \le 1 \tag{2.1.52}$$

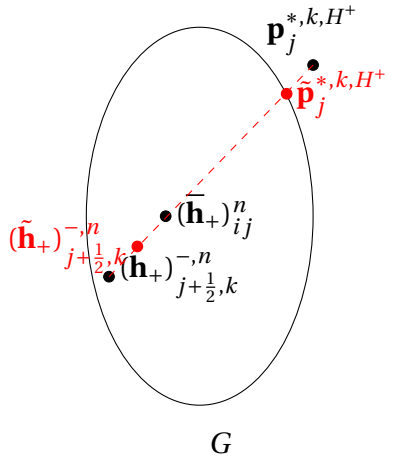
  and the surface ∂G. Modify $(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}$ as follows

$$(\tilde{\mathbf{h}}_+)^{-,n}_{j+\frac{1}{2},k} = \theta\left((\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k} - (\overline{\mathbf{h}}_+)^n_{jk}\right) + (\overline{\mathbf{h}}_+)^n_{jk}, \quad \theta = \min_{1,2}\{\theta_1, \theta_2\} \tag{2.1.53}$$

The condition $m^2 + n^2 \leq S^2\rho^2$ is not linear, so we use an ellipse to denote the convex set G. The following two diagrams illustrate the two cases. When $(\mathbf{h}_+)^{-,n}_{j+\frac{1}{2},k}$ with the updated $\tilde{\rho}^{f^+}_{j+\frac{1}{2},k}$ is not inside G:



When $\mathbf{p}^{*,k,H^+}_j$ with the updated $\tilde{\rho}^{*,k,H^+}_j$ is not inside G:



We apply the same process to $(\mathbf{h}_-)^{+,n}_{j+\frac{1}{2},k}$, $(\mathbf{q}_+)^{-,n}_{j+\frac{1}{2},k}$ and $(\mathbf{q}_-)^{+,n}_{j+\frac{1}{2},k}$ individually. Finally, we obtain

$$\hat{\mathbf{f}}_{j+\frac{1}{2},k} = \alpha_x\left((\tilde{\mathbf{h}}_+)^{-,n}_{j+\frac{1}{2},k} - (\tilde{\mathbf{h}}_-)^{+,n}_{j+\frac{1}{2},k}\right) \tag{2.1.54}$$

$$\hat{\mathbf{g}}_{j,k+\frac{1}{2}} = \alpha_y\left((\tilde{\mathbf{q}}_+)^{-,n}_{j,k+\frac{1}{2}} - (\tilde{\mathbf{q}}_-)^{+,n}_{j,k+\frac{1}{2}}\right) \tag{2.1.55}$$

The limiter for the velocity functions is modified from the limiter introduced in [YWS13]. We omitted the ϵ term in the convex set G.

We thus come to the semi-discrete scheme

$$\frac{d}{dt}\mathbf{w}_{jk}(t) = -\frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{j+\frac{1}{2},k} - \hat{\mathbf{f}}_{j-\frac{1}{2},k}\right) - \frac{1}{\Delta y}\left(\hat{\mathbf{g}}_{j,k+\frac{1}{2}} - \hat{\mathbf{g}}_{j,k-\frac{1}{2}}\right). \tag{2.1.56}$$

2.1.4 TVD-RK3 Time Integration

Note that the limiter from the previous section is developed over a forward Euler method

$$\mathbf{w}_j^{n+1} = \mathbf{w}_j^n + \Delta t \mathbf{L}(\mathbf{w}^n). \tag{2.1.57}$$

Now we will show that the TVD RK3 time integration gives a convex combination of forward Euler methods, and thus by using the limiter mentioned in Section 2.1.3, the numerical solution obtained from the full scheme is also in G. The TVD RK3 [SO88], [SO89], [LOC94], [ZS12], [LXX13], is given as follows

$$\begin{aligned}
\mathbf{w}^{(0)} =& \mathbf{w}^n, \\
\mathbf{w}^{(1)} =& \mathbf{w}^{(0)} + \Delta t \mathbf{L}(\mathbf{w}^{(0)}), \\
\mathbf{w}^{(2)} =& \frac{3}{4}\mathbf{w}^{(0)} + \frac{1}{4}\left(\mathbf{w}^{(1)} + \Delta t \mathbf{L}(\mathbf{w}^{(1)})\right), \\
\mathbf{w}^{n+1} =& \frac{1}{3}\mathbf{w}^{(0)} + \frac{2}{3}\left(\mathbf{w}^{(2)} + \Delta t \mathbf{L}(\mathbf{w}^{(2)})\right)
\end{aligned} \tag{2.1.58}$$

It is natural that $\mathbf{w}^{(0)} \in G$. $\mathbf{w}^{(1)}$ is obtained as a forward Euler method with respect to $\mathbf{w}^{(0)}$, so $\mathbf{w}^{(1)} \in G$. As for $\mathbf{w}^{(2)}$, $\mathbf{w}^{(1)} + \Delta t \mathbf{L}(\mathbf{w}^{(1)} \in G$ as it is another forward Euler expression with respect to $\mathbf{w}^{(1)}$. So $\mathbf{w}^{(2)} \in G$. $\mathbf{w}^{n+1} \in G$ since $\mathbf{w}^{(2)} + \Delta t \mathbf{L}(\mathbf{w}^{(2)}$ is another forward Euler method with respect to $\mathbf{w}^{(2)}$. At each time level $t = t_n$, we find the maximal character-

istic speeds in x and y directions

$$\alpha_x = \max_{j,k} |u_{jk}^n|, \alpha_y = \max_{j,k} |v_{jk}^n|. \tag{2.1.59}$$

Then we find the proceeding time step size

$$\Delta t = CFL \times \min\left\{\frac{\Delta x}{\alpha_x}, \frac{\Delta y}{\alpha_y}\right\} \tag{2.1.60}$$

where

$$CFL \leq \frac{1}{2}\hat{\omega}_N = \frac{1}{2}\frac{1}{6} = \frac{1}{12}. \tag{2.1.61}$$

Thus, the semi-discrete scheme (2.1.56) obtained in the previous section together with the TVD-RK3 time integration method given in Section 2.1.4 will give a high order scheme. In practice, we will see 5th order of accuracy in Section 2.2.1.

2.2 Numerical Results

2.2.1 Accuracy Test

We make use of the Example 7. in [YWS13] to test the order of convergence of our finite difference method. The two dimensional pressureless system has the following initial condition

$$\begin{aligned} \rho(x,y,0) =& \rho_0(x+y) = \exp\big(\sin(x+y)\big), \\ u(x,y,0) =& u_0(x+y) = \frac{1}{3}\big(\cos(x+y)+2\big), \\ v(x,y,0) =& v_0(x+y) = \frac{1}{3}\big(\sin(x+y)+2\big). \end{aligned} \tag{2.2.1}$$

This specific example has the analytical solution in the following implicit form

$$\begin{aligned} u(x,y,t) &= u_0(z), \\ v(x,y,t) &= v_0(z), \\ \rho(x,y,t) &= \frac{\rho(z)}{1+t\big(u_0'(z)+v_0'(z)\big)}, \end{aligned} \tag{2.2.2}$$

where

$$z + t\big(u_0(z) + v_0(z)\big) = x + y. \tag{2.2.3}$$

Note that [YWS13] had a typo for missing t in the denominator. It is straightforward to plug (2.2.2) back in the pressureless system (2.0.1) and conclude there is a t term in the density solution. We can use Newton's method [New11] to solve for z in the above equation. For each pair (x, y) at time t, we define

$$f(z) = z + t\big(u_0(z) + v_0(z)\big) - x - y. \tag{2.2.4}$$

Note that

$$f'(z) = 1 + t\big(u_0'(z) + v_0'(z)\big) = 1 + \frac{1}{3}t\big(\cos(z) - \sin(z)\big), \tag{2.2.5}$$

and

$$f''(z) = t\big(u_0''(z) + v_0''(z)\big) = -\frac{1}{3}t\big(\sin(z) + \cos(z)\big). \tag{2.2.6}$$

[YWS13] chose the final time $t = 0.1$, so

$$f'(z) = 1 + \frac{1}{3} \times 0.1\big(\cos(z) - \sin(z)\big) \in [1 - \frac{\sqrt{2}}{30}, 1 + \frac{\sqrt{2}}{30}] \neq 0. \tag{2.2.7}$$

$f''(z)$ is continuous everywhere. With those two conditions satisfied, we simply choose $x+y$ as the initial guess and start the iteration of the Newton's method as follows

$$\begin{aligned} z_0 &= x+y, \\ z_{n+1} &= z_n - \frac{f(z_n)}{f'(z_n)},\ n \in \mathbb{N}. \end{aligned} \tag{2.2.8}$$

The iteration stops when

$$|z_{n+1} - z_n| \leq 10^{-13}, \tag{2.2.9}$$

and z_{n+1} is the final approximation to z at the given location (x, y). We compare the analytical solution ρ^a and the numerical solution ρ^u to the density function in L_1, L_2 and L_∞ norms. We obtain Table 2.3 for the orders of convergence of our finite difference WENO method on the pressureless system of Euler equations (2.0.1).

2.2.2 Improvements

We didn't achieve the good result of at least 5th order of convergence like what we got from the one dimensional problem. The errors and orders of accuracy are listed in Table 2.3. Let's study how it happened. Recall the convex set defined in Section 2.1.1

$$G = \left\{ \mathbf{w} = (\rho,\ \ m,\ \ n)^T : \rho \geq 0, m^2 + m^2 \leq S^2\rho^2 \right\} \tag{2.2.10}$$

where

$$S > 0, \text{ and } S^2 = \max_{x,y} \left\{ u^2(x,y,0) + v^2(x,y,0) \right\}.$$

In particular, the second condition $u^2 + v^2 \leq S^2$ is not a linear condition. The nature of WENO reconstruction is to find the appropriate polynomials from neighboring cells

Table 2.3: Without the global condition that the limiter is applied when the first components of the split fluxes are less than 0.1.
L_1, L_2 and L_∞ Errors and Orders of accuracy

$N \times N$	L_1 Error	L_1 Order	L_2 Error	L_2 Order	L_∞ Error	L_∞ Order
Density ρ						
10×10	$5.4251e-01$	-	$1.1599e-01$	-	$3.6920e-02$	-
20×20	$4.0899e-01$	0.4076	$1.2896e-01$	-0.1529	$5.9678e-02$	-0.6928
40×40	$3.5863e-02$	3.5115	$1.4419e-02$	3.1609	$1.0325e-02$	2.5311
80×80	$1.7141e-02$	1.0650	$1.0252e-02$	0.4921	$1.0858e-02$	-0.0726
160×160	$5.4042e-03$	1.6653	$4.0584e-03$	1.3369	$5.5673e-03$	0.9637
320×320	$1.1339e-03$	2.2528	$1.2209e-03$	1.7329	$1.8726e-03$	1.5719
640×640	$9.9335e-05$	3.5129	$1.3250e-04$	3.2039	$2.9798e-04$	2.6518
Momentum ρu						
10×10	$4.8896e-01$	-	$1.0444e-01$	-	$3.0476e-02$	-
20×20	$2.5032e-01$	0.9659	$7.6433e-02$	0.4504	$3.4919e-02$	-0.1963
40×40	$2.6463e-02$	3.2417	$1.0641e-02$	2.8446	$7.7500e-03$	2.1717
80×80	$1.3584e-02$	0.9621	$7.9598e-03$	0.4188	$8.3993e-03$	-0.1161
160×160	$3.8176e-03$	1.8311	$2.8440e-03$	1.4848	$3.9752e-03$	1.0793
320×320	$7.7737e-04$	2.2960	$8.0697e-04$	1.8173	$1.3190e-03$	1.5916
640×640	$6.4194e-05$	3.5981	$8.5545e-05$	3.2378	$1.9114e-04$	2.7867
Momentum ρv						
10×10	$6.8892e-01$	-	$1.3974e-01$	-	$4.2746e-02$	-
20×20	$4.8949e-01$	0.4931	$1.5469e-01$	-0.1466	$7.0364e-02$	-0.7190
40×40	$4.2323e-02$	3.5318	$1.7160e-02$	3.1723	$1.2274e-02$	2.5192
80×80	$1.9265e-02$	1.1354	$1.1641e-02$	0.5598	$1.2145e-02$	0.0153
160×160	$6.2059e-03$	1.6343	$4.6980e-03$	1.3091	$6.3640e-03$	0.9323
320×320	$1.3241e-03$	2.2286	$1.4302e-03$	1.7158	$2.1827e-03$	1.5438
640×640	$1.1608e-04$	3.5118	$1.5487e-04$	3.2072	$3.4748e-04$	2.6511

and evaluate the polynomials at certain locations. Even though the points $(\rho, m, n)^T$ we used to reconstruct the polynomials are inside the convex set G, it is not guaranteed that the reconstructed point values are inside G, let alone a convex combination of those point values. That being said, there are many points where the second condition $u^2 + v^2 \leq S^2$ is violated, but not enough to cause damage to the algorithm. More specifically, the velocity u or v gets a little too large to be inside G, but not large enough to dramatically decrease the size of Δt. Therefore, the program can run smoothly without the limiter. With the strong conditions of the limiter, we manually changed plenty of point values inside the domain, which leads to worse order of convergence.

Considering the two special behaviors of the analytical solutions to the pressureless system: Dirac delta shock solutions and vacuum solutions. We need a limiter to maintain the boundedness of the velocity functions in the vacuum cases, but not the delta shock cases, as only density close to vacuum leads to the velocities functions to blow up. Here we manually add a global condition. If the density components of the split fluxes are greater than or equal to 0.1, we will skip the limiter. Otherwise, the limiter will be applied to maintain the algorithm. By comparing the results from without any limiter and with the limiter along with the global condition, they are identical. This tells us that the limiter was not used in this smooth case. The errors and orders of accuracy are listed in Table 2.4.

2.2.3 Numerical Applications

We applied our finite difference method along with the limiter with the global condition on the 23 cases introduced in [Pan19]. All of the cases start with the following

Table 2.4: With the global condition that the limiter is applied when the first components in the split fluxes are less than 0.1.
L_1, L_2, and L_∞ Errors and Orders of Accuracy

$N \times N$	L_1 Error	L_1 Order	L_2 Error	L_2 Order	L_∞ Error	L_∞ Order
Density ρ						
10×10	$3.0033e-01$	-	$5.5593e-02$	-	$1.6167e-02$	-
20×20	$4.1681e-02$	2.8491	$8.2732e-03$	2.7484	$2.3343e-03$	2.7920
40×40	$3.9207e-03$	3.4102	$1.1284e-03$	2.8742	$6.8668e-04$	1.7653
80×80	$3.7834e-04$	3.3734	$1.2701e-04$	3.1512	$8.8550e-05$	2.9551
160×160	$5.1602e-05$	2.8742	$2.0164e-05$	2.6550	$1.5513e-05$	2.5130
320×320	$2.6078e-06$	4.3065	$1.2266e-06$	4.0391	$1.1204e-06$	3.7914
640×640	$3.8272e-08$	6.0904	$1.8286e-08$	6.0678	$2.1365e-08$	5.7126
1280×1280	$8.9436e-10$	5.4193	$2.2734e-10$	6.3297	$1.4132e-10$	7.2401
Momentum ρu						
10×10	$2.8231e-01$	-	$5.1805e-02$	-	$1.5490e-02$	-
20×20	$3.5682e-02$	2.9840	$6.7232e-03$	2.9459	$2.0866e-03$	2.8921
40×40	$4.8350e-03$	2.8836	$1.3227e-03$	2.3457	$8.5828e-04$	1.2816
80×80	$3.4039e-04$	3.8283	$1.1494e-04$	3.5246	$7.6195e-05$	3.4937
160×160	$4.8047e-05$	2.8247	$2.1134e-05$	2.4432	$2.0090e-05$	1.9232
320×320	$2.0604e-06$	4.5435	$1.0432e-06$	4.3404	$1.2043e-06$	4.0602
640×640	$2.9893e-08$	6.1070	$1.3835e-08$	6.2366	$1.3845e-08$	6.4426
1280×1280	$7.6022e-10$	5.2972	$1.9779e-10$	6.1282	$1.3448e-10$	6.6859
Momentum ρv						
10×10	$4.1589e-01$	-	$7.6165e-02$	-	$2.1588e-02$	-
20×20	$6.6756e-02$	2.6392	$1.3065e-02$	2.5489	$4.1403e-03$	2.3824
40×40	$5.9495e-03$	3.4881	$1.5649e-03$	3.0616	$9.1108e-04$	2.1841
80×80	$5.2276e-04$	3.5085	$1.6931e-04$	3.2084	$1.1208e-04$	3.0230
160×160	$5.9453e-05$	3.1363	$2.3425e-05$	2.8536	$1.6397e-05$	2.7731
320×320	$3.1153e-06$	4.2543	$1.4793e-06$	3.9851	$1.5089e-06$	3.4418
640×640	$4.7153e-08$	6.0459	$2.2965e-08$	6.0093	$2.6549e-08$	5.8287
1280×1280	$1.0119e-09$	5.5421	$2.7145e-10$	6.4026	$1.7778e-10$	7.2224

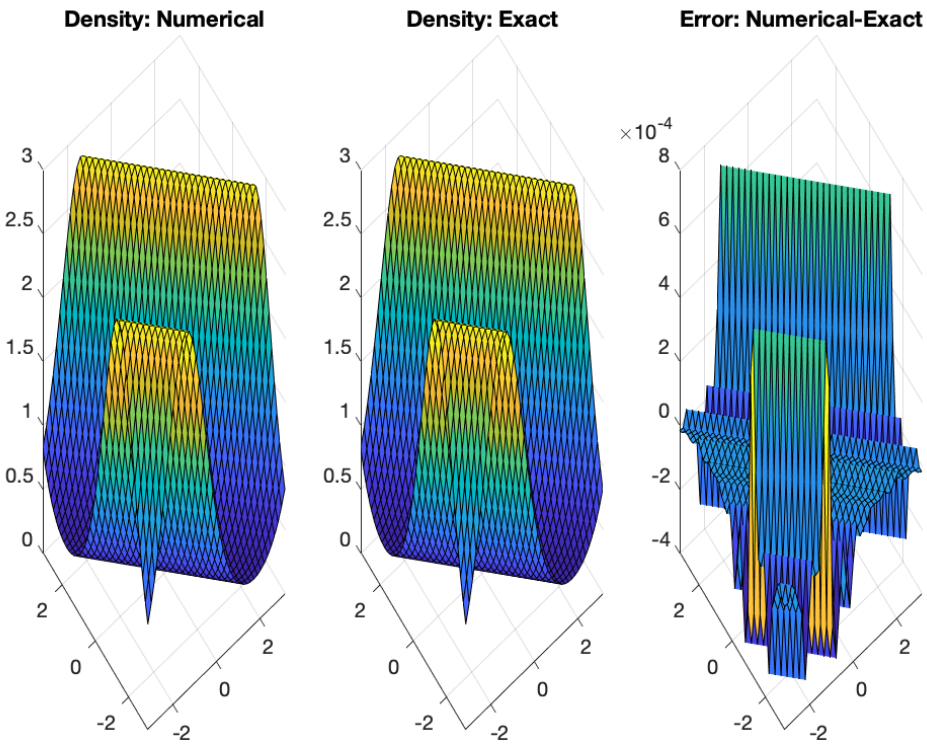


Figure 2.1: Smooth case with the global condition: $N \times N = 40 \times 40$, $T = 0.1$.

initial condition

$$(\rho, u, v)(0, x, y) = (\rho_i, u_i, v_i), \text{ in the } i\text{th quadratic}, \ i = 1, 2, 3, 4 \tag{2.2.11}$$

where $\rho_i > 0$, u_i and v_i are constants. We plot all of the figures in (x, y) plane at the final time T.

Case 1.

Case 1. is the only solution with the contact discontinuities and the vacuum behaviors satisfying

$$u_3 = u_4 < u_1 = u_2, v_1 = v_4 < v_2 = v_3. \tag{2.2.12}$$

We took

$$u_1 = u_2 = 3,\ u_3 = u_4 = -3,\ v_1 = v_4 = -2,\ v2 = v_3 = 2, \tag{2.2.13}$$

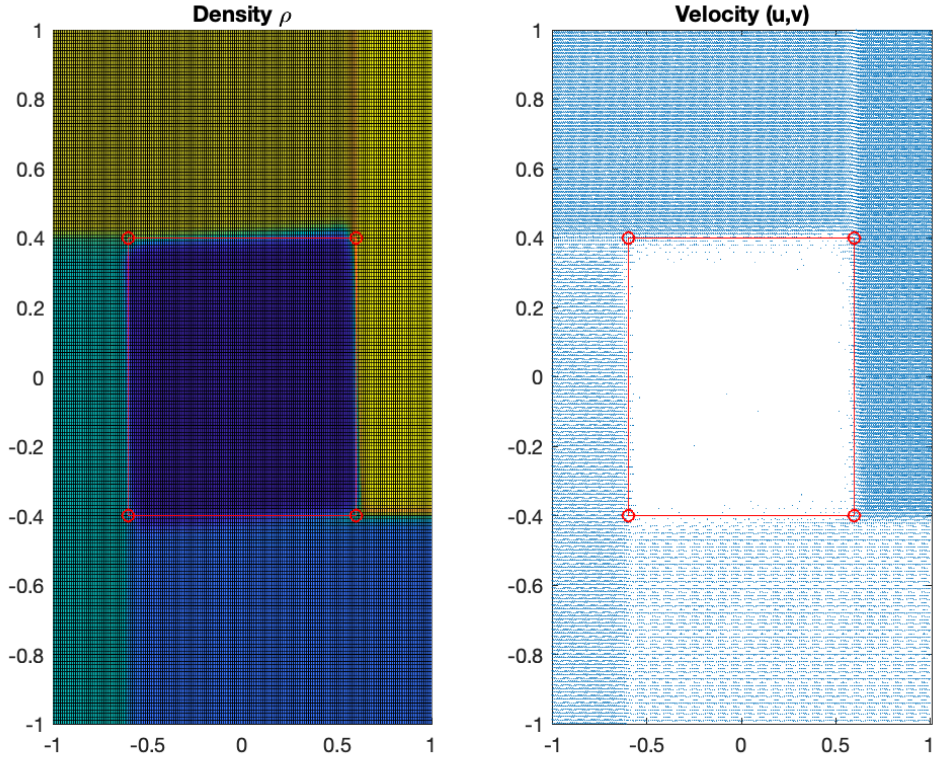
as the initial conditions for the velocity field. The choice for the density is arbitrary. Considering that we want to be able to tell the locations of the contact discontinuities, we need to make the two adjacent density values are distinct, i.e., ρ_1 is not equal to ρ_2 or ρ_4, ρ_2 is not equal to ρ_1 or ρ_3, etc. In our example, we took

$$\rho_1 = 4, \rho_2 = 3, \rho_3 = 2, \rho_4 = 1 \tag{2.2.14}$$

as the initial density conditions. The four red circles are corresponding to $\Xi_i, i = 1,2,3,4$ in [Pan19]. Each Ξ_i has the Cartesian coordinates

$$\Xi_i = (u_i T, v_i T), \tag{2.2.15}$$

where T is the final time of the numerical simulation. Figure 2.2a gave the numerical simulation of Case 1. In Figure 2.2b, ξu and ηv denote the x and y directions, respectively. Note that the two quantities are not multiplied. Pang in [Pan19] defined $\xi = x/t$ and $\eta = y/t$. For numerical simulations at any time, t is fixed. u and v denote the velocities in x and y directions.



(a) Case 1. $N \times N = 200 \times 200$, $T = 0.2$.

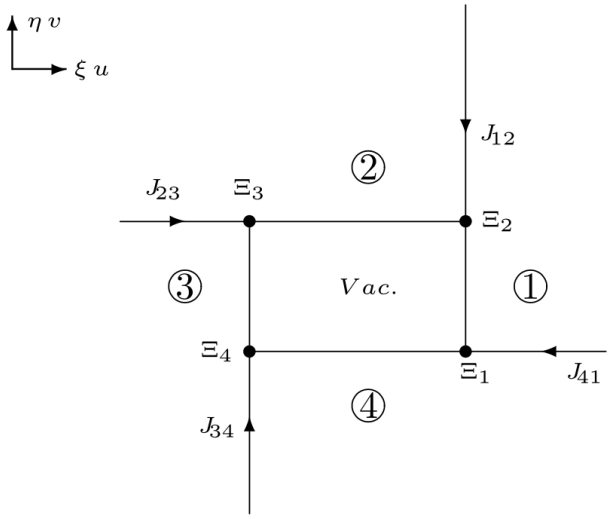


Fig. 4.1. The solution for Case 4.1(i).

(b) Case 1. Screenshot from [Pan19]

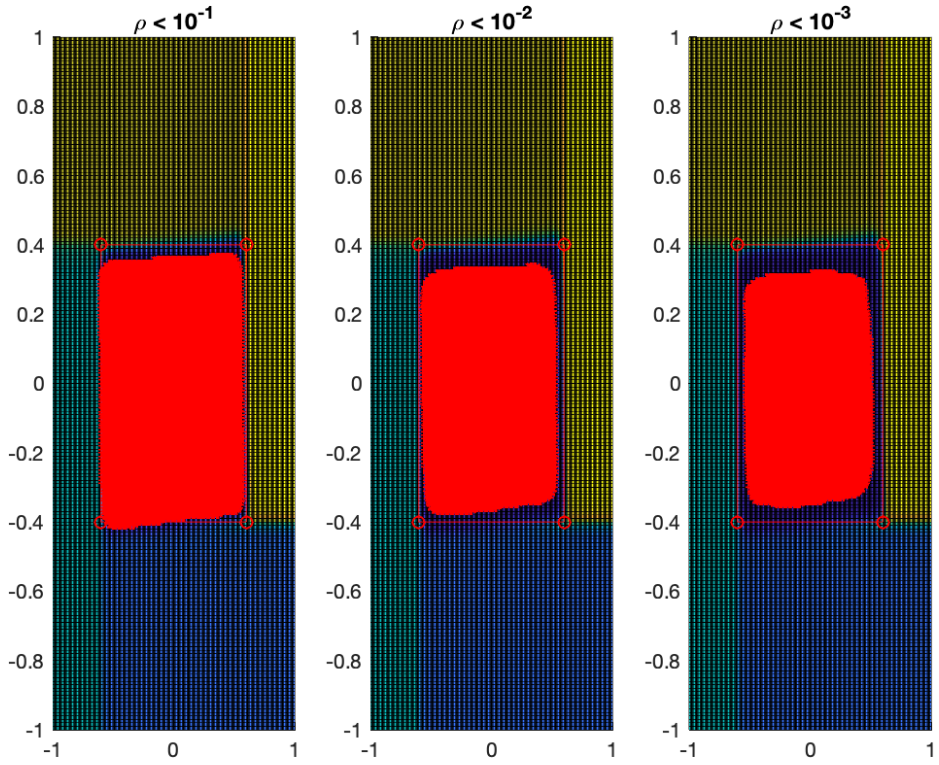


Figure 2.2: Case 1. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

Figure 2.2 introduced more details about the vacuum area. The red portions in the three subfigures denote the areas where density is less than $10^{-1}, 10^{-2}$ and 10^{-3}, respectively. The following table gives more detailed numbers of points with density lower than certain small numbers. The minimum density value of this simulation is

$$\rho_{min} = 1.6136 \times 10^{-8}. \tag{2.2.16}$$

Table 2.5: Case 1. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	8,186	6,912	6,100	3,817

Case 2.

The initial conditions for Case 2. requires

$$u_1 = u_2 < u_3 = u_4, v_1 = v_4 < v_2 = v_3. \tag{2.2.17}$$

There was no condition for the density mentioned in [Pan19], but the directions of the two shocks implied that the initial density value in the first quadrant is notably bigger than the values from other quadrants. Therefore, we used the following set of initial data for Case 2.

$$u_1 = u_2 = -1, u_3 = u_4 = 3, v_1 = v_4 = -1, v_2 = v_3 = 1, \tag{2.2.18}$$

and

$$\rho_1 = 10, \rho_2 = 1, \rho_3 = 3, \rho_4 = 2. \tag{2.2.19}$$

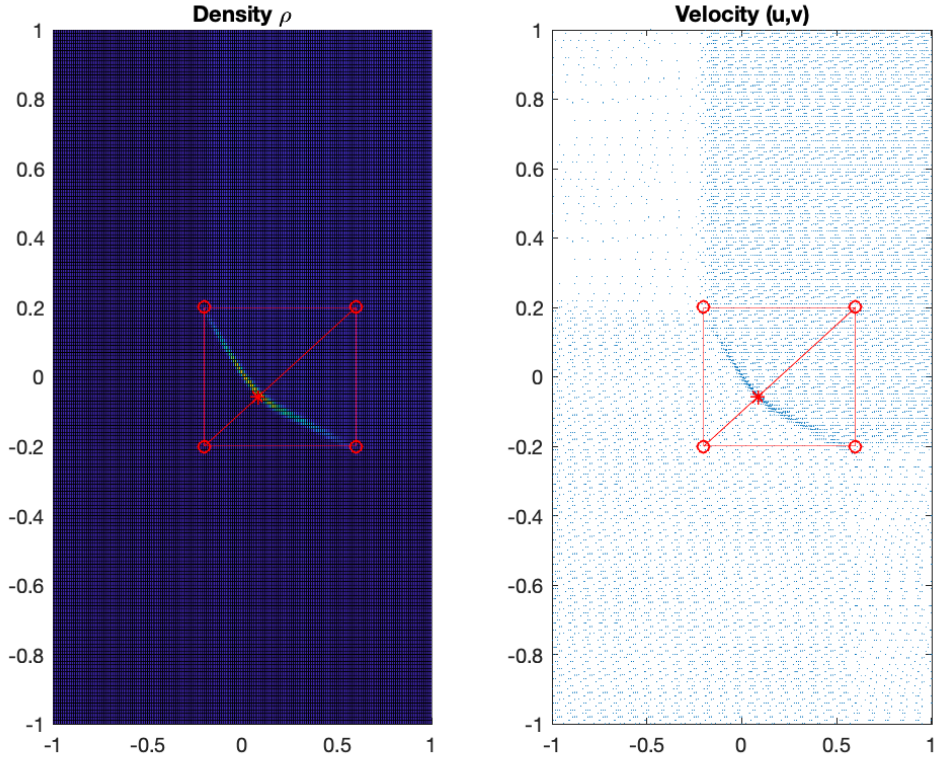
In Case 2. the two contact discontinuities J_{12} and J_{23} meet at the point Ξ_2 and form a new delta shock. The two contact discontinuities J_{34} and J_{41} meet at the point Ξ_4 and form another delta shock. The two newly developed delta shocks follow the bold paths, $\Xi_2 U_\delta^{13}$ and $\Xi_4 U_\delta^{13}$, and will collide at the point U_δ^{13}. The intersection point U_δ^{13} has the Cartesian coordinates

$$U_\delta^{13} = \Big(\frac{\sqrt{\rho_1} u_1 + \sqrt{\rho_3} u_3}{\sqrt{\rho_1} + \sqrt{\rho_3}} T, \frac{\sqrt{\rho_1} v_1 + \sqrt{\rho_3} v_3}{\sqrt{\rho_1} + \sqrt{\rho_3}} T \Big). \tag{2.2.20}$$

Case 3.

Case 3. requires

$$u_4 < u_3 < u_1 < u_2, v_1 < v_4 < v_2 < v_3. \tag{2.2.21}$$



(a) Case 2. $N \times N = 200 \times 200$, $T = 0.2$.

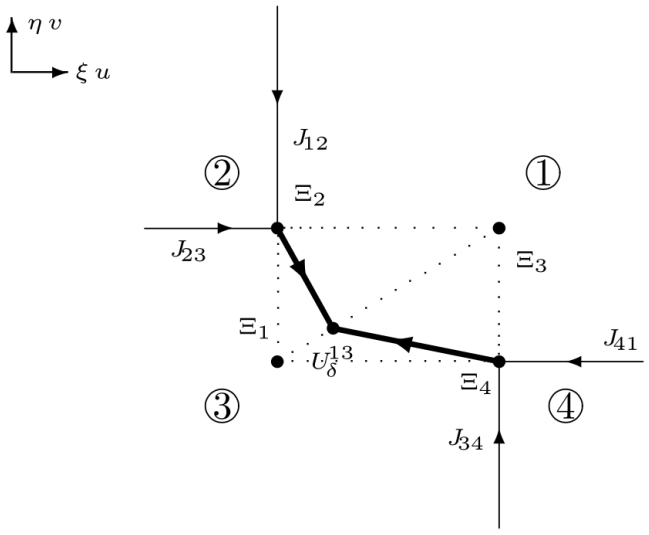


Fig. 4.2. The solution for Case 4.1(ii).

(b) Case 2. Screenshot from [Pan19]

Again, there is no condition for the density choice. Especially there is no contact discontinuity in this figure to cause confusion when comes to the location, the initial density choice can be arbitrary. We took

$$\rho_1 = 1, \rho_2 = 1, \rho_3 = 1, \rho_4 = 1 \tag{2.2.22}$$

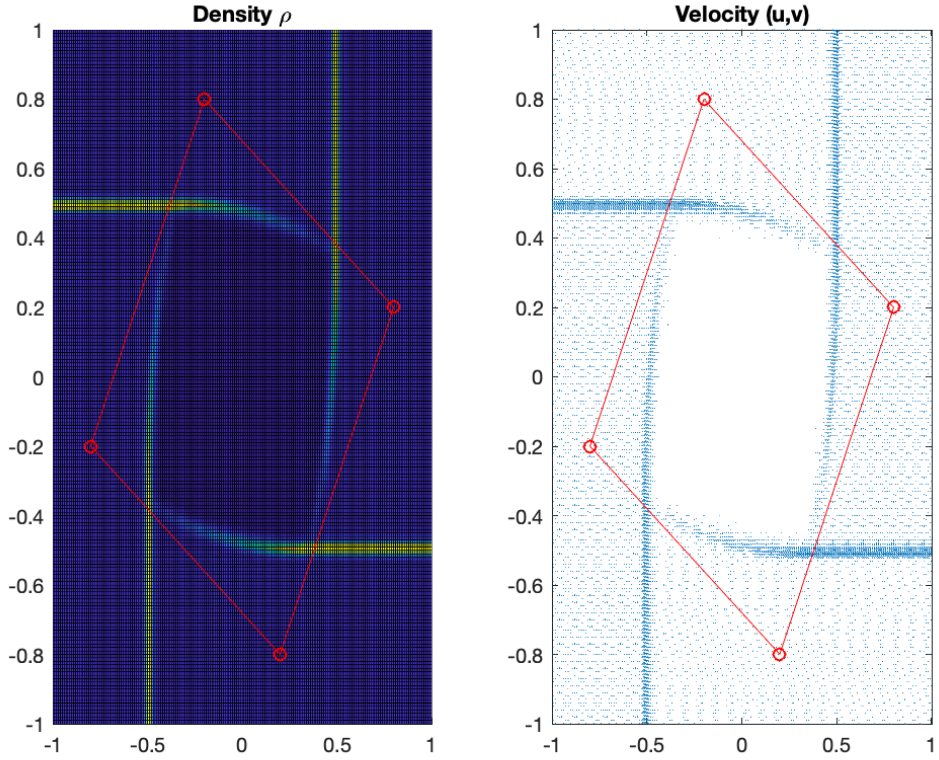
as the initial density condition.

$$u_1 = 1, u_2 = 4, u_3 = -1, u_4 = -4, v_1 = -4, v_2 = 1, v_3 = 4, v_4 = -1 \tag{2.2.23}$$

as the initial velocity conditions. In particular, the delta shocks δ_{12}, δ_{23}, δ_{34}, and δ_{41} are located on

$$\begin{aligned}
\delta_{12} : x &= \frac{\sqrt{\rho_1}u_1 + \sqrt{\rho_2}u_2}{\sqrt{\rho_1} + \sqrt{\rho_2}} T, \\
\delta_{23} : y &= \frac{\sqrt{\rho_2}v_2 + \sqrt{\rho_3}v_3}{\sqrt{\rho_2} + \sqrt{\rho_3}} T, \\
\delta_{34} : x &= \frac{\sqrt{\rho_3}u_3 + \sqrt{\rho_4}u_4}{\sqrt{\rho_3} + \sqrt{\rho_4}} T, \\
\delta_{41} : y &= \frac{\sqrt{\rho_4}v_4 + \sqrt{\rho_1}v_1}{\sqrt{\rho_4} + \sqrt{\rho_1}} T,
\end{aligned} \tag{2.2.24}$$

respectively. Case 3. combines the delta shock behaviors and the vacuum behavior. All of the bold lines and curves represent the delta shocks. The area inside the shocks is a vacuum region.



(a) Case 3. $N \times N = 200 \times 200$ and $T = 0.2$.

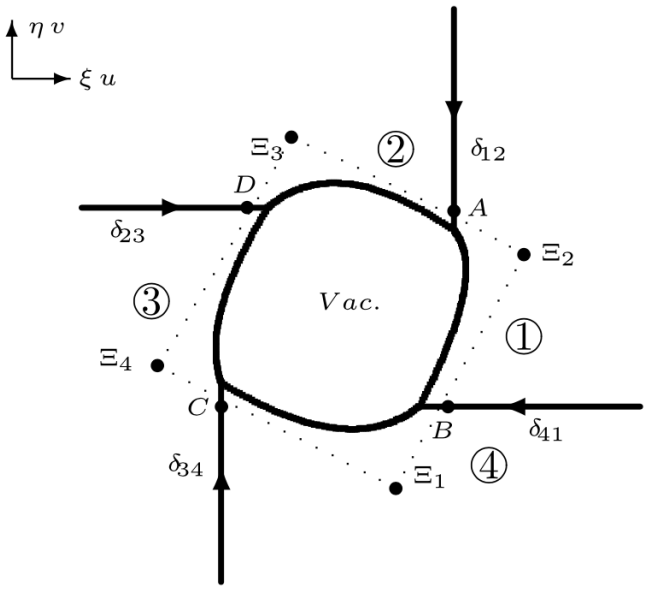


Fig. 4.3. The solution for Case 4.2(i).

(b) Case 3. Screenshot from [Pan19]

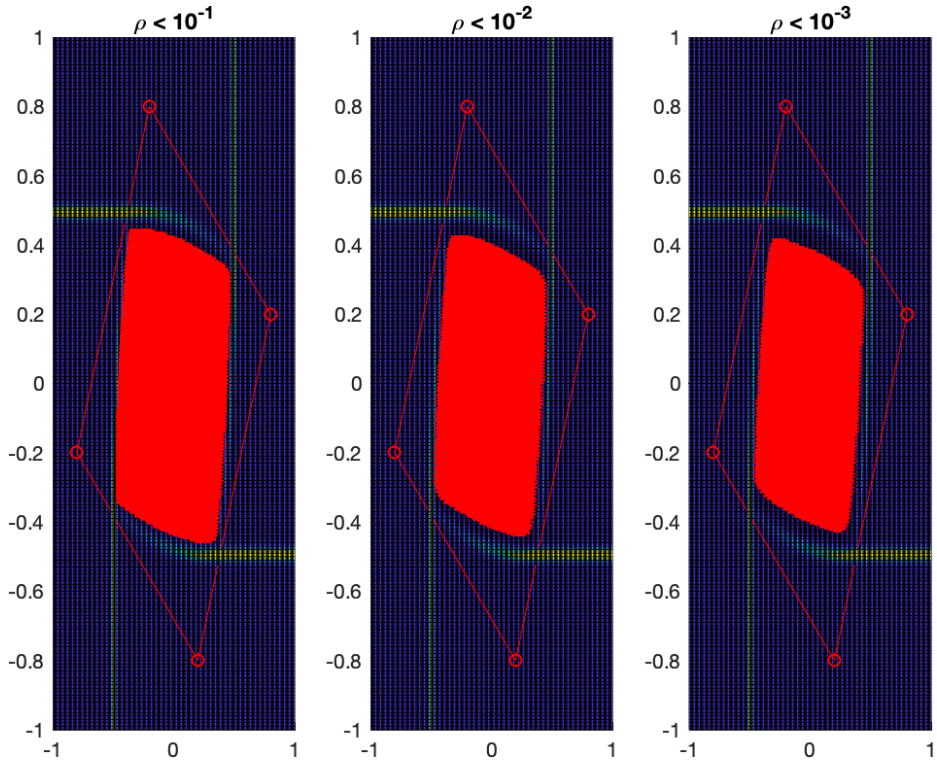


Figure 2.3: Case 3. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 3. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 9.8315 \times 10^{-11}. \tag{2.2.25}$$

More details about the vacuum area are listed in the following table.

Table 2.6: Case 3. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	6,488	5,824	5,384	4,912

Case 4.

For Case 4., Pang gave a generic initial condition

$$u_1 < u_2, u_4 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.26}$$

and

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_2 & v_2 & 1 \end{pmatrix} \geq 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} \leq 0. \tag{2.2.27}$$

After studying the diagram attentively, we further restrict the initial condition to be

$$u_1 < u_2 < u_4 < u_3, v_1 < v_4 < v_2 < v_3, \tag{2.2.28}$$

and

$$u_2 \leq \frac{\sqrt{\rho_1}u_1 + \sqrt{\rho_3}u_3}{\sqrt{\rho_1} + \sqrt{\rho_3}} \leq u_4, v_4 < \frac{\sqrt{\rho_1}v_1 + \sqrt{\rho_3}v_3}{\sqrt{\rho_1} + \sqrt{\rho_3}} < v_2 \tag{2.2.29}$$

together with the determinant conditions in (2.2.27). Specifically, we took

$$\rho_1 = 20, \rho_2 = 1, \rho_3 = 2, \rho_4 = 1, \tag{2.2.30}$$

$$u_1 = -2, u_2 = -1.5, u_3 = 2, u_4 = 1, v_1 = -2, v_2 = 1, v_3 = 2, v_4 = -1. \tag{2.2.31}$$

Case 4. Delta shocks δ_{12} and δ_{23} intersect at Ξ_2 and form a new delta shock. Delta shocks δ_{34} and δ_{41} intersect at Ξ_4 and form another delta shock. The two newly developed delta shocks along the curved paths, $\Xi_2 U_\delta^{13}$ and $\Xi_4 U_\delta^{13}$, will collide at point U_δ^{13}

on the segment $\Xi_1\Xi_3$, where U_δ^{13} is located at

$$U_\delta^{13} = \Big(\frac{\sqrt{\rho_1}u_1 + \sqrt{\rho_3}u_3}{\sqrt{\rho_1}+\sqrt{\rho_3}}T, \frac{\sqrt{\rho_1}v_1 + \sqrt{\rho_3}v_3}{\sqrt{\rho_1}+\sqrt{\rho_3}}T\Big). \tag{2.2.32}$$

The sad thing is that we have no idea what additional conditions are need to obtain the curvatures of the two delta shocks. In [Pan19], Pang consistently used straight lines to illustrate delta shocks from the intersection of two contact discontinuities, and curvy lines to illustrate delta shocks from the intersection of two behaviors with at least one delta shock. Pang did not introduce details about what kind of conditions led to different shapes of the solution behaviors.

Case 5.

Case 5. is a complicated case where the boundary of the vacuum area is delta shocks. [Pan19] gave the initial condition as follows

$$u_4 < u_1 < u_2 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.33}$$

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_2 & v_2 & 1 \end{pmatrix} \geq 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} \geq 0, \tag{2.2.34}$$

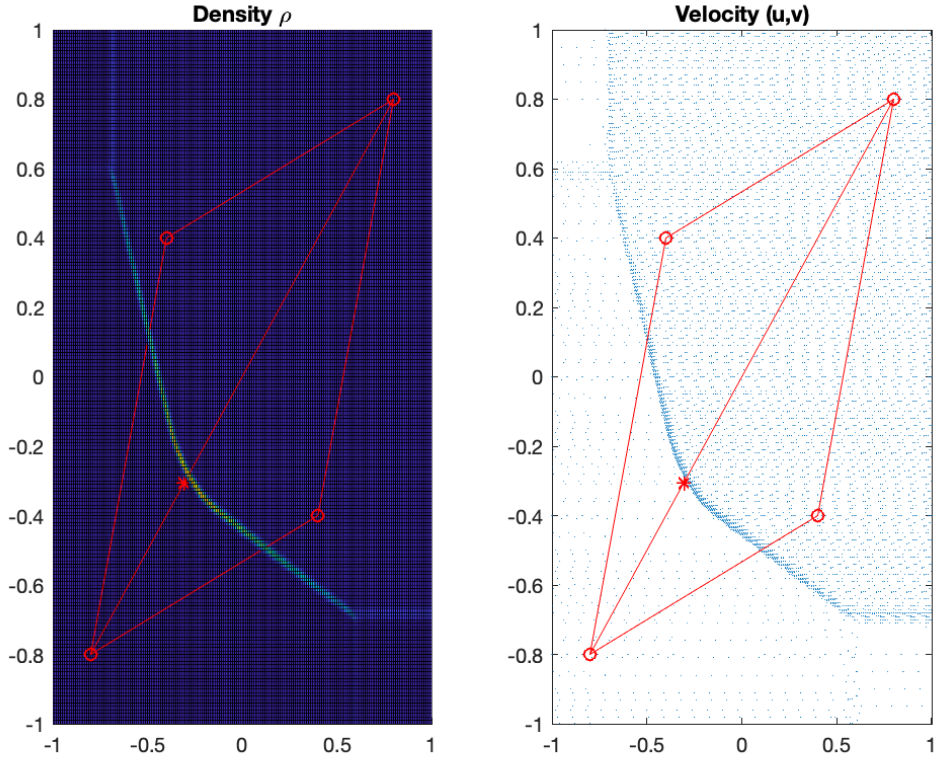
and

$$v_\delta^{41} < v_\delta^{13}, u_\delta^{34} < u_1. \tag{2.2.35}$$

From the diagram in [Pan19], we further need the condition that

$$v_4 < v_2. \tag{2.2.36}$$

Because of the conditions in (2.2.42), delta shocks δ_{12} and δ_{23} intersect and we see



(a) Case 4. $N \times N = 200 \times 200$ and $T = 0.2$.

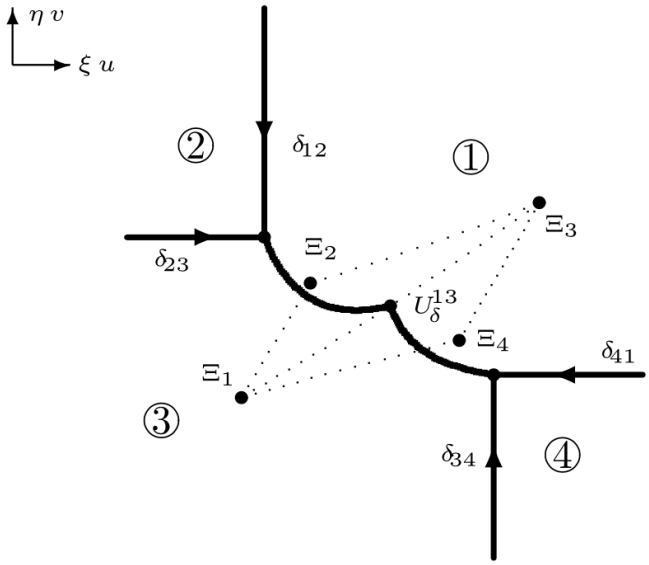


Fig. 4.4. The solution for Case 4.2(ii).

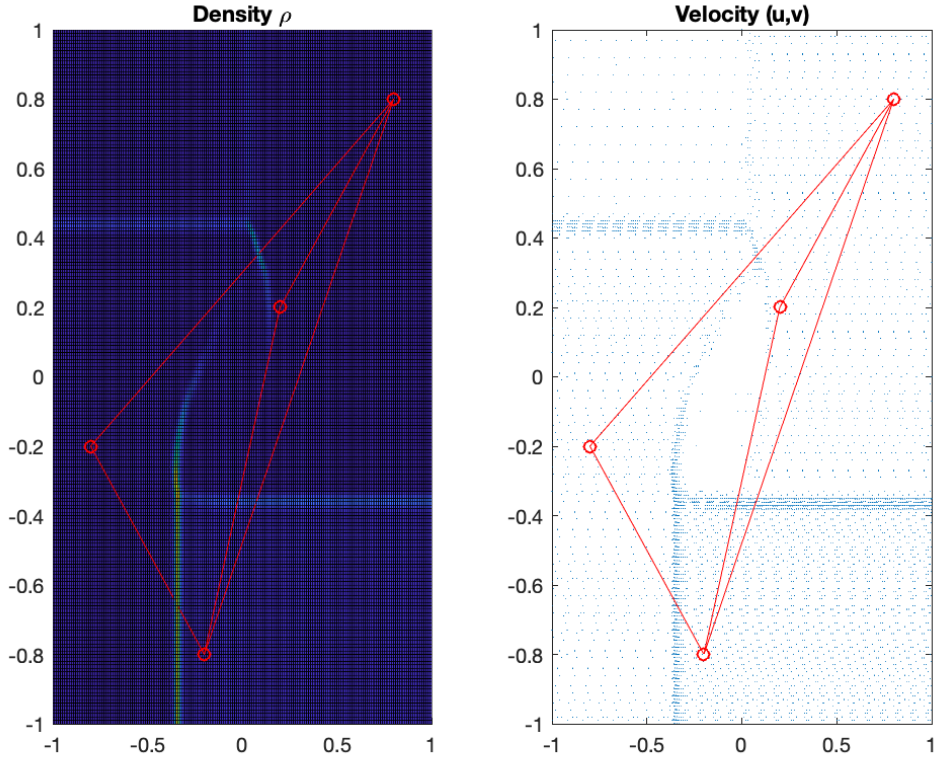
(b) Case 4. Screenshot from [Pan19]

another delta shock from the intersection. Delta shocks δ_{34} and δ_{41} do not intersect. In our numerical simulation, the delta shock above δ_{41} is much lower than the delta shock δ_{41}, so it is too weak to change the direction of δ_{41}. We couldn't obtain the curvature below δ_{41}. There was no explicit explanation regarding the choice of initial conditions affecting the shape of the solution. For Case 5., we took

$$\rho_1 = 1, \rho_2 = 2, \rho_3 = 1, \rho_4 = 6 \tag{2.2.37}$$

and

$$u_1 = -1, u_2 = 1, u_3 = 4, u_4 = -4, v_1 = -4, v_2 = 1, v_3 = 4, v_4 = -1. \tag{2.2.38}$$



(a) Case 5. $N \times N = 200 \times 200$, and $T = 0.2$.

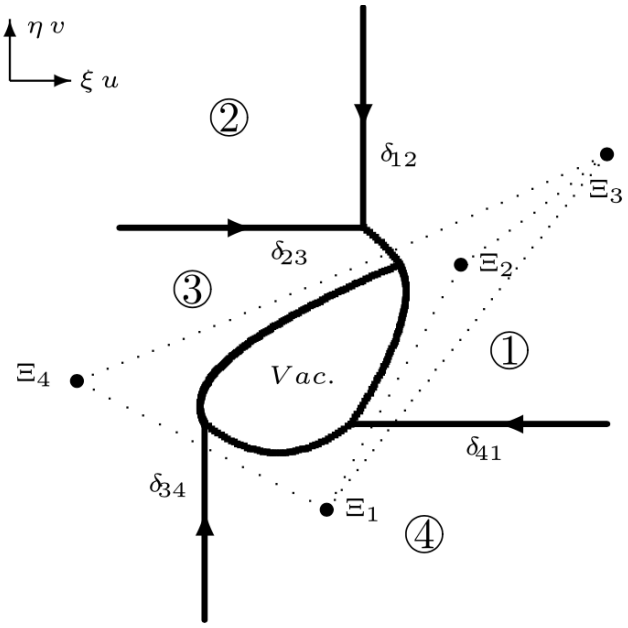


Fig. 4.5. The solution for Case 4.2(iii)a.

(b) Case 5. Screenshot from [Pan19]

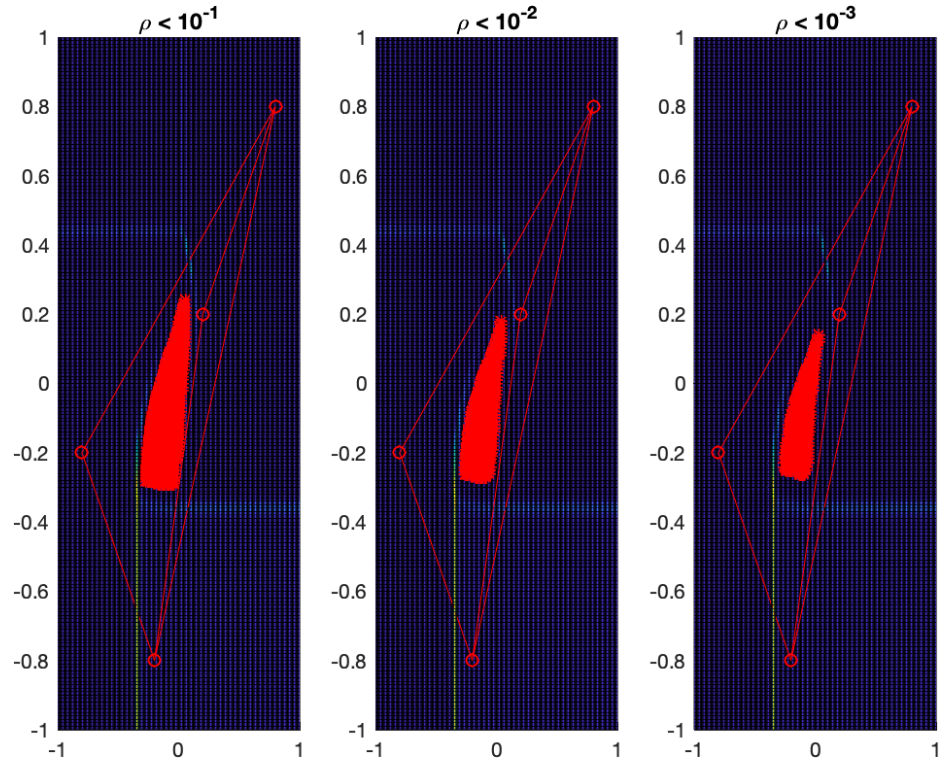


Figure 2.4: Case 5. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 5. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 8.5615 \times 10^{-7}. \tag{2.2.39}$$

More details about the vacuum area are listed in the following table.

Table 2.7: Case 5. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	973	664	487	373

Case 6.

[Pan19] gave the initial condition as follows

$$u_4 < u_1 < u_2 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.40}$$

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_2 & v_2 & 1 \end{pmatrix} \geq 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} \geq 0, \tag{2.2.41}$$

and

$$v_\delta^{41} < v_\delta^{13}, u_\delta^{34} > u_1. \tag{2.2.42}$$

Like Case 5. we still need the additional condition

$$v_4 < v_2. \tag{2.2.43}$$

So we took the following initial conditions

$$\rho_1 = \rho_2 = \rho_3 = \rho_4 = 1 \tag{2.2.44}$$

and

$$u_1 = -1, u_2 = 1, u_3 = 4, u_4 = -4, v_1 = -4, v_2 = 1, v_3 = 4, v_4 = -1. \tag{2.2.45}$$

Because the vacuum area is small, we set the numerical domain to be $(-40, 40)$ and the final time $T = 8$. Case 6. is the first case that our finite difference method behaves not well. Delta shocks δ_{12} and δ_{23} intersect and develop a new delta shock δ_{13}^A. Delta shocks δ_{34} and δ_{41} intersect and develop a new delta shock δ_{13}^B. Delta shocks δ_{13}^A and δ_{13}^B form a vacuum area. We will see that Case 14. and Case 17 are also cases of this

type where two newly developed delta shocks form a vacuum area in between. All of the three numerical results are not satisfying as there are only a few points with close-to-zero density values.

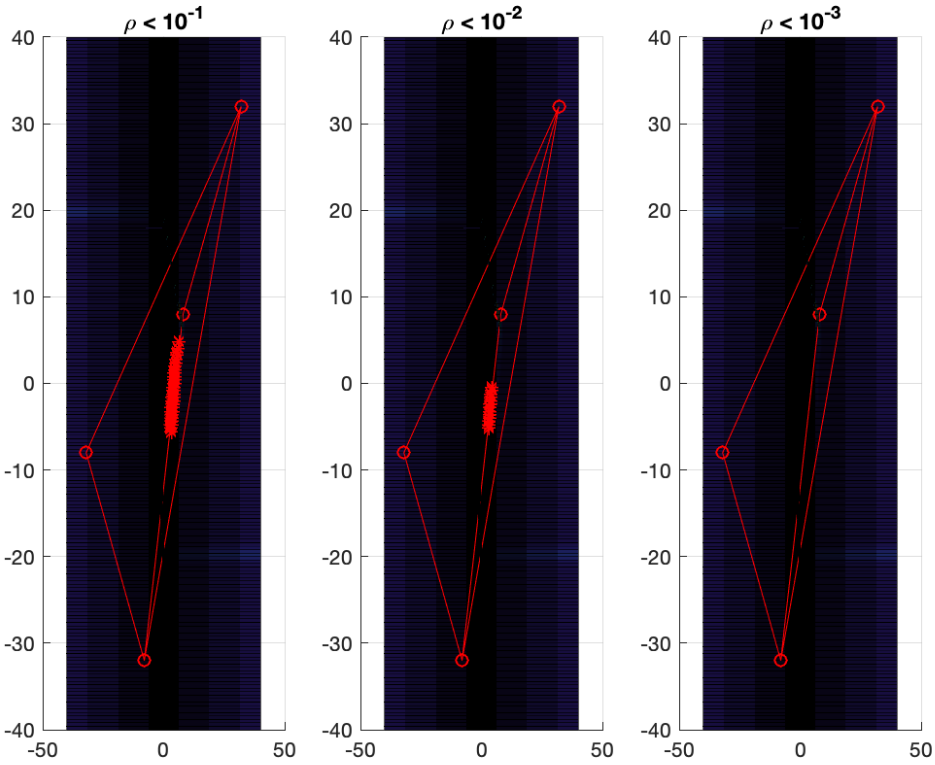
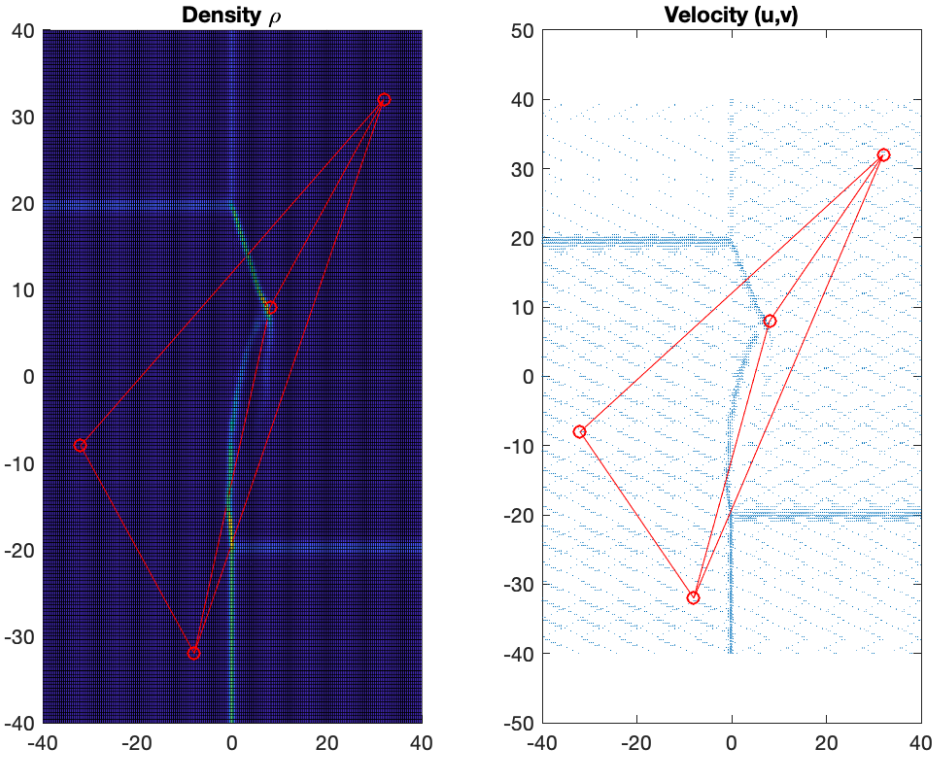


Figure 2.5: Case 6. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 6. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 4.1 \times 10^{-3}. \tag{2.2.46}$$

More details about the vacuum area are listed in the following table.



(a) Case 6. $N \times N = 200 \times 200$, and $T = 8$.

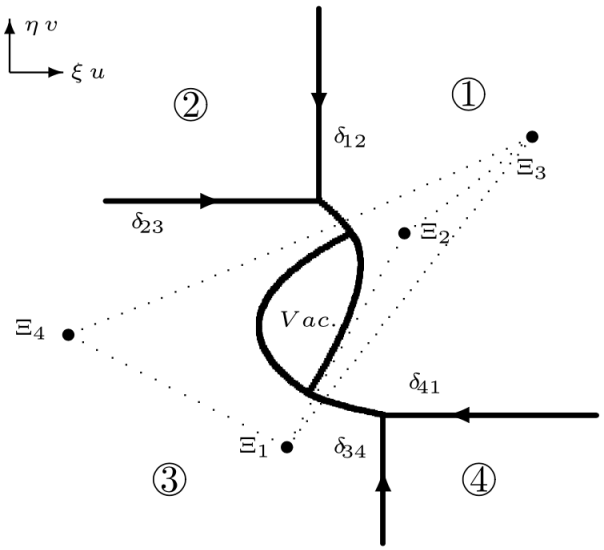


Fig. 4.6. The solution for Case 4.2(iii)b.

(b) Case 6. Screenshot from [Pan19]

Table 2.8: Case 6. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	48	14	0	0

Case 7.

[Pan19] gave the following restrictions

$$u_1 = u_2, u_4 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.47}$$

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_2 & v_2 & 1 \end{pmatrix} > 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} > 0, \tag{2.2.48}$$

and

$$v_1 < v_\delta^{41} < v_2. \tag{2.2.49}$$

To obtain the shape in the given diagram for Case 7., we further updated the conditions

$$u_4 < u_3 < u_1 = u_2, v_1 < v_4 < v_2 < v_3, \tag{2.2.50}$$

and

$$u_\delta^{34} > u_\delta^{24} \tag{2.2.51}$$

together with (2.2.48). The inequality (2.2.51) comes from the fact that the delta shock δ_{34} travels fast in the x-direction. It doesn't intersect with the delta shock δ_{23}, instead, intersects with the shock on the right of the delta shock δ_{23}. We omit the condition (2.2.58) because $v_1 < v_\delta^{41} < v_4$, as a weight average of v_1 and v_4, and therefore $v_\delta^{41} < v_2$ by transitivity. In addition, the location in the original diagram in [Pan19] is inaccurate.

It is supposed to be in between the y-values of Ξ_2 and Ξ_3, as

$$v_\delta^{23} = \frac{\sqrt{\rho_2}v_2 + \sqrt{\rho_3}v_3}{\sqrt{\rho_2} + \sqrt{\rho_3}}. \tag{2.2.52}$$

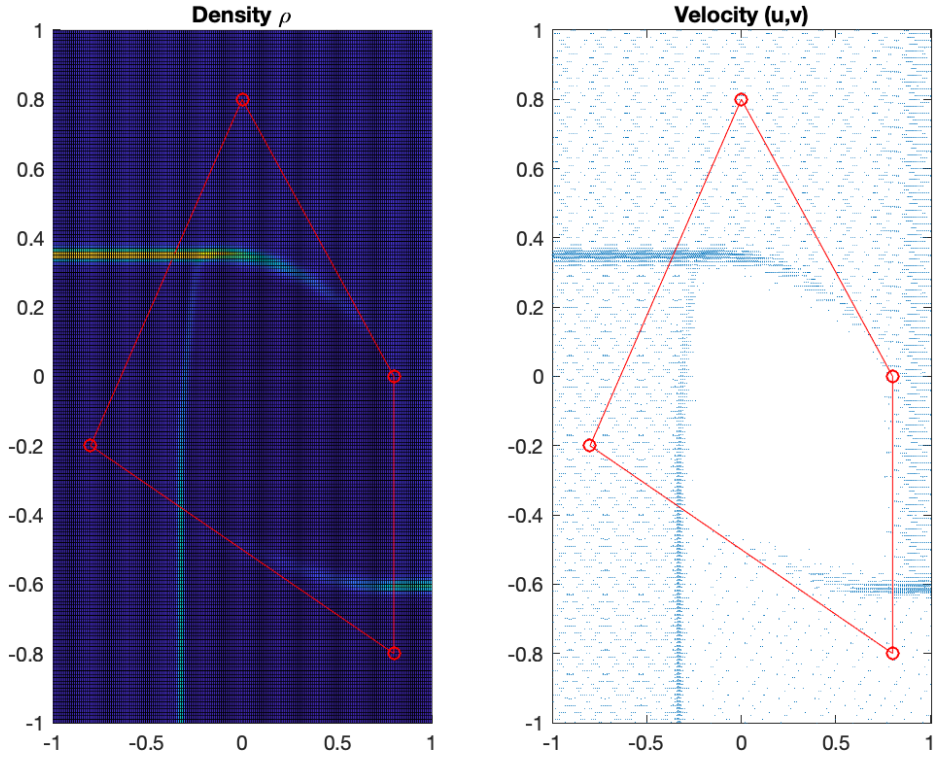
Specifically, we took the following set of initial conditions for Case 7.

$$\rho_1 = 4, \rho_2 = 3, \rho_3 = 2, \rho_4 = 1, \tag{2.2.53}$$

and

$$u_1 = 4, u_2 = 4, u_3 = 0, u_4 = -4, v_1 = -4, v_2 = 0, v_3 = 4, v_4 = -1. \tag{2.2.54}$$

Case 7. consists of three delta shocks, δ_{23}, δ_{34} and δ_{41}, and one contact discontinuity, J_{12}.



(a) Case 7. $N \times N = 200 \times 200$, and $T = 0.2$.

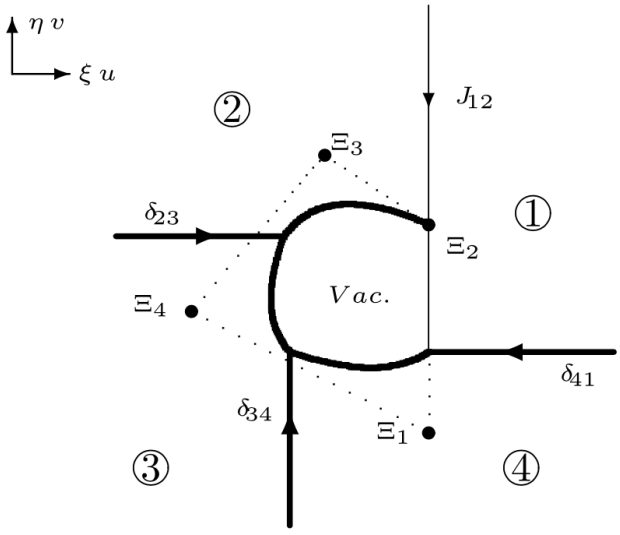


Fig. 4.7. The solution for Case 4.3(i)a.

(b) Case 7. Screenshot from [Pan19]

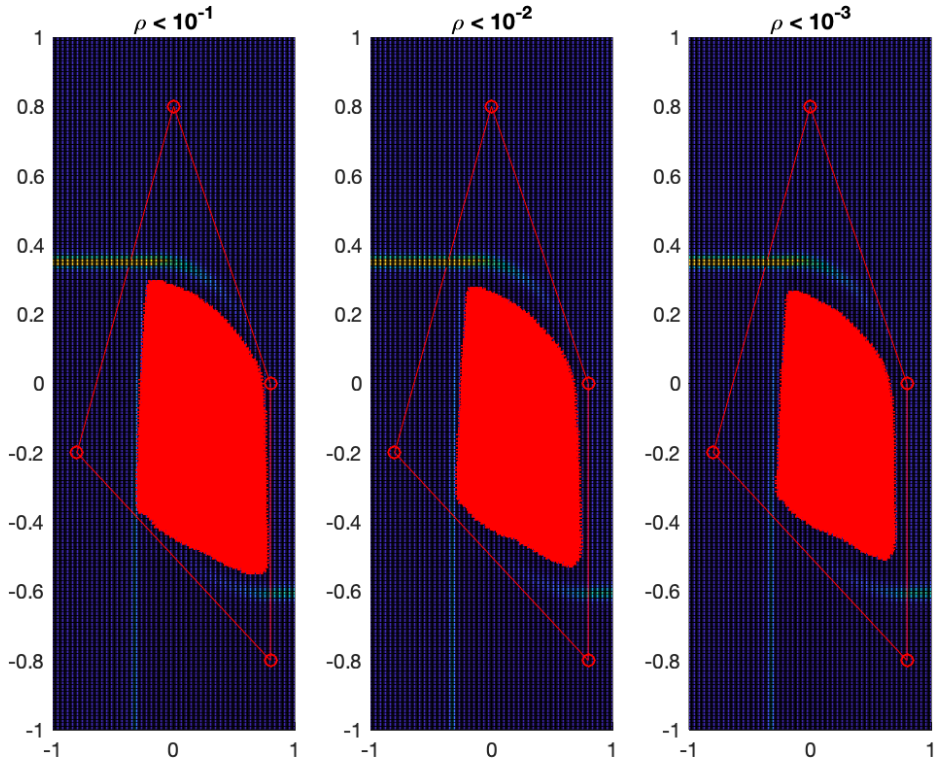


Figure 2.6: Case 7. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 7. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 1.4110 \times 10^{-12}. \tag{2.2.55}$$

More details about the vacuum area are listed in the following table.

Table 2.9: Case 7. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	6,406	5,588	5,102	4,630

Case 8.

[Pan19] gave the following restrictions

$$u_1 = u_2, u_4 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.56}$$

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_2 & v_2 & 1 \end{pmatrix} > 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} > 0, \tag{2.2.57}$$

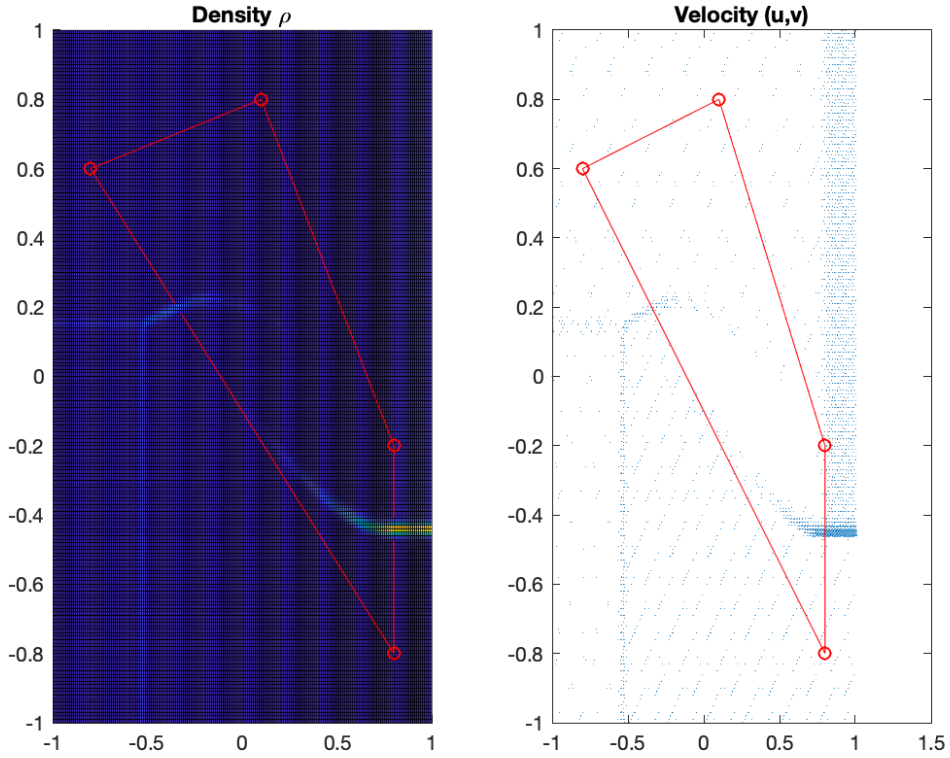
and

$$v_1 < v_\delta^{41} < v_2. \tag{2.2.58}$$

The above conditions are the same as the ones given for Case 7. We need to further study the diagram and come up with more detailed restrictions. As the only difference between Case 8. and Case 7 is the location of the delta shock δ_{34}, we only need to update the condition (2.2.51) for Case 7. to be

$$u_\delta^{34} < u_\delta^{24}. \tag{2.2.59}$$

Thus, we made sure that the delta shock δ_{34} would intersect with the delta shock δ_{23} first, and form a new delta shock. Case 8. is very much similar to Case 7. except that the delta shock δ_{34} is slower in the x-direction, so it intersects with the delta shock δ_{23} first and form a new delta shock. The rest of the restrictions are the same as those of Case 7.



(a) Case 8. $N \times N = 200 \times 200$, and $T = 0.2$.

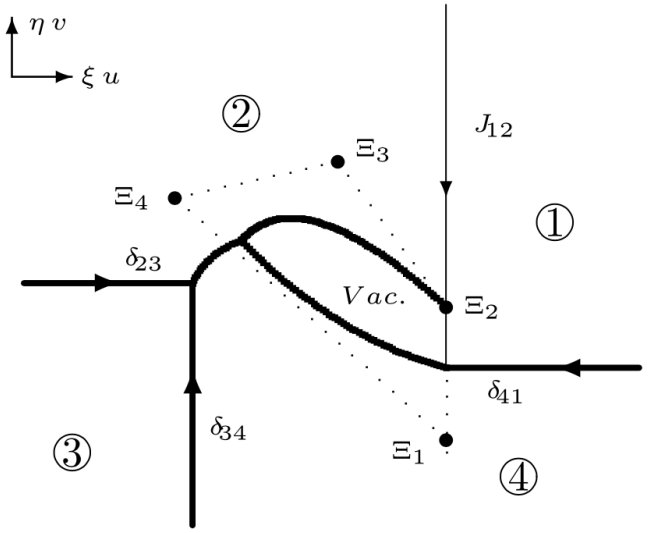


Fig. 4.8. The solution for Case 4.3(i)b.

(b) Case 8. Screenshot from [Pan19]

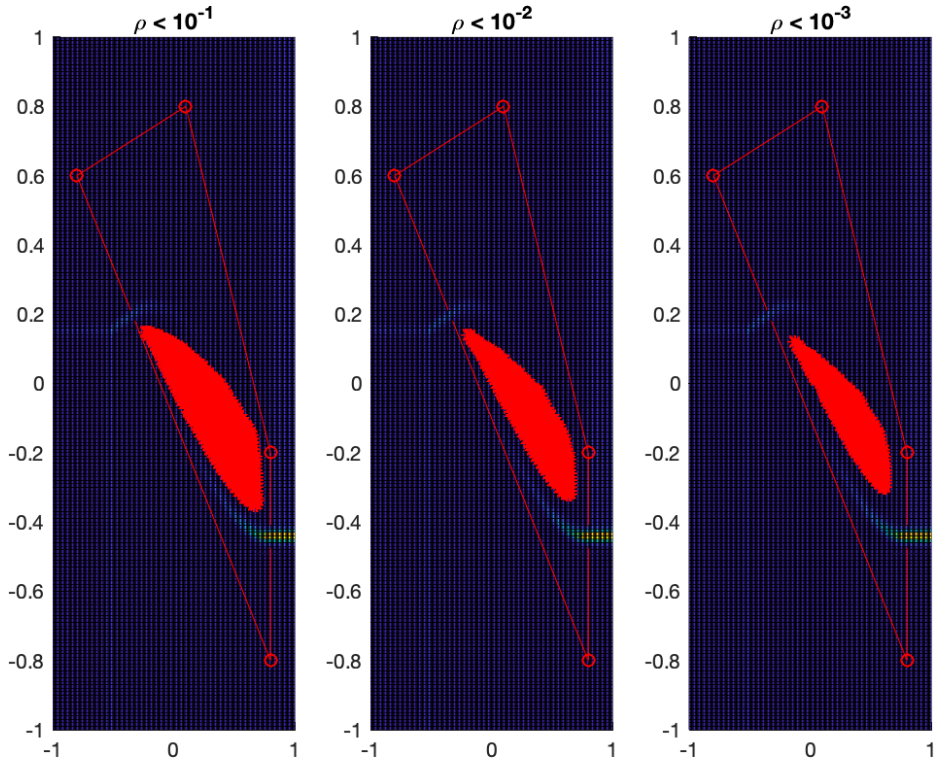


Figure 2.7: Case 8. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 8. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 7.1027 \times 10^{-8}. \tag{2.2.60}$$

More details about the vacuum area are listed in the following table.

Table 2.10: Case 8. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	1,566	1,115	846	609

Case 9.

Case 9. is closely related to Case 7 as well. From the diagrams, the difference mainly lies in the location of the delta shock δ_{41}. [Pan19] gave the following restrictions

$$u_1 = u_2, u_4 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.61}$$

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{pmatrix} > 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} > 0, \tag{2.2.62}$$

and

$$v_1 < v_2 < v_\delta^{41}. \tag{2.2.63}$$

When we looked at the diagram more closely, we updated the restrictions to

$$u_4 < u_3 < u_1 = u_2, v_1 < v_2 < v_4 < v_3. \tag{2.2.64}$$

and

$$u_\delta^{34} > u_\delta^{24}, \tag{2.2.65}$$

together with the condition (2.2.62). The condition (2.2.63) made sure that the delta shock is fast enough in the y direction, so the delta shock on the left of the contact discontinuity J_{12} intersects with the newly developed delta shock from the the delta shock δ_{41} and the contact discontinuity J_{12}, instead of J_{12} itself. As for the condition (2.2.73), it led to the location of the delta shock δ_{34} intersecting with the delta shock on the right of the delta shock δ_{23}. In the end, we chose the following set of the initial conditions for Case 9.

$$\rho_1 = 10, \rho_2 = 3, \rho_3 = 60, \rho_4 = 10, \tag{2.2.66}$$

and

$$u_1 = 4, u_2 = 4, u_3 = 0.5, u_4 = -4, v_1 = -4, v_2 = -1, v_3 = 4, v_4 = 3. \tag{2.2.67}$$

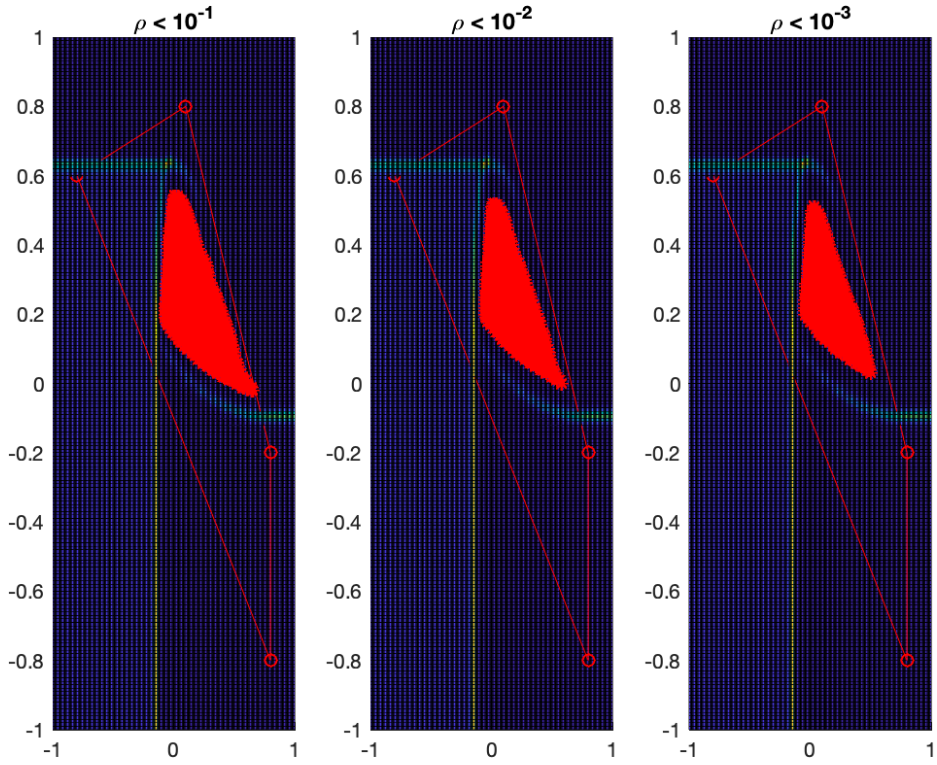
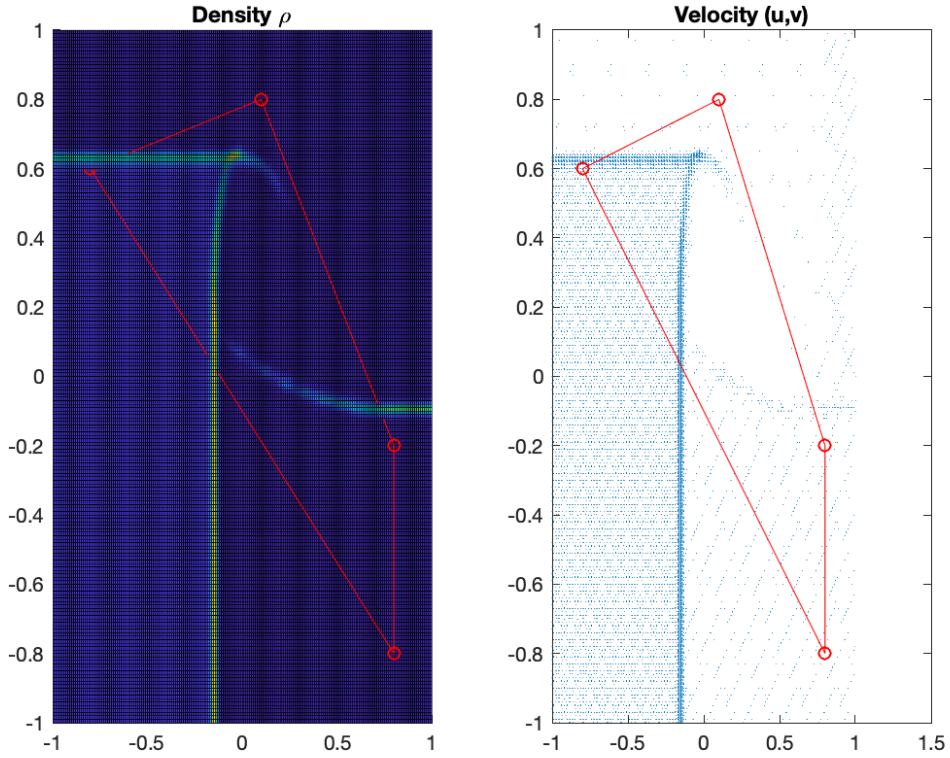


Figure 2.8: Case 9. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 9. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 6.3272 \times 10^{-11}. \tag{2.2.68}$$

More details about the vacuum area are listed in the following table.



(a) Case 9. $N \times N = 200 \times 200$, and $T = 0.2$.

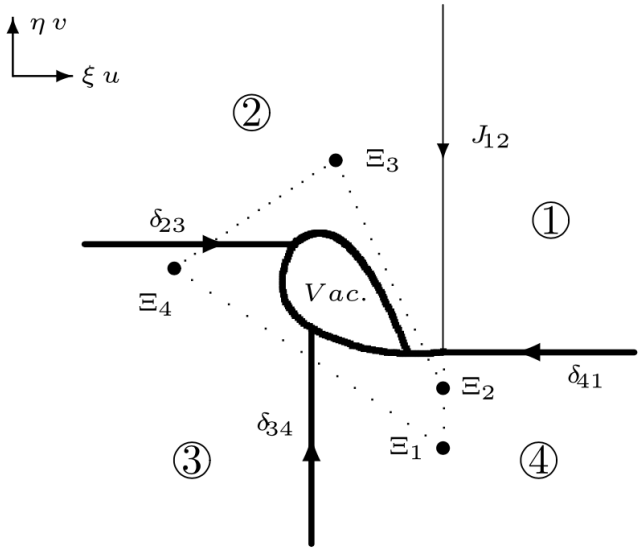


Fig. 4.9. The solution for Case 4.3(ii)a.

(b) Case 9. Screenshot from [Pan19]

Table 2.11: Case 9. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	1,727	1,365	1,154	942

Case 10.

[Pan19] gave the following restrictions

$$u_1 = u_2, u_4 < u_3, v_1 < v_4, v_2 < v_3, \tag{2.2.69}$$

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{pmatrix} > 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} > 0, \tag{2.2.70}$$

and

$$v_1 < v_2 < v_\delta^{41}. \tag{2.2.71}$$

These conditions are the same as those given for Case 9. But we further add more details to be

$$u_4 < u_3 < u_1 = u_2, v_1 < v_2 < v_4 < v_3. \tag{2.2.72}$$

and

$$u_\delta^{34} < u_\delta^{24}, \tag{2.2.73}$$

together with conditions in (2.2.70) and (2.2.71). So we chose the following initial conditions

$$\rho_1 = 10, \rho_2 = 3, \rho_3 = 1, \rho_4 = 10, \tag{2.2.74}$$

and

$$u_1 = 4, u_2 = 4, u_3 = 0.5, u_4 = 4, v_1 = -4, v_2 = -1, v_3 = 4, v_4 = 3. \tag{2.2.75}$$

Case 10. is a pair with Case 9. The only difference lies in the location of the delta shock δ_{34}. In Case 10, the delta shock δ_{34} is slower in the x-direction, so it intersects with the delta shock δ_{23} first, and form a new delta shock, just like what happened in Case 8. We kept the same set for velocities, and changed the relative velocity of the delta shock δ_{34} by modifying the choice of the densities in the four quadrants.

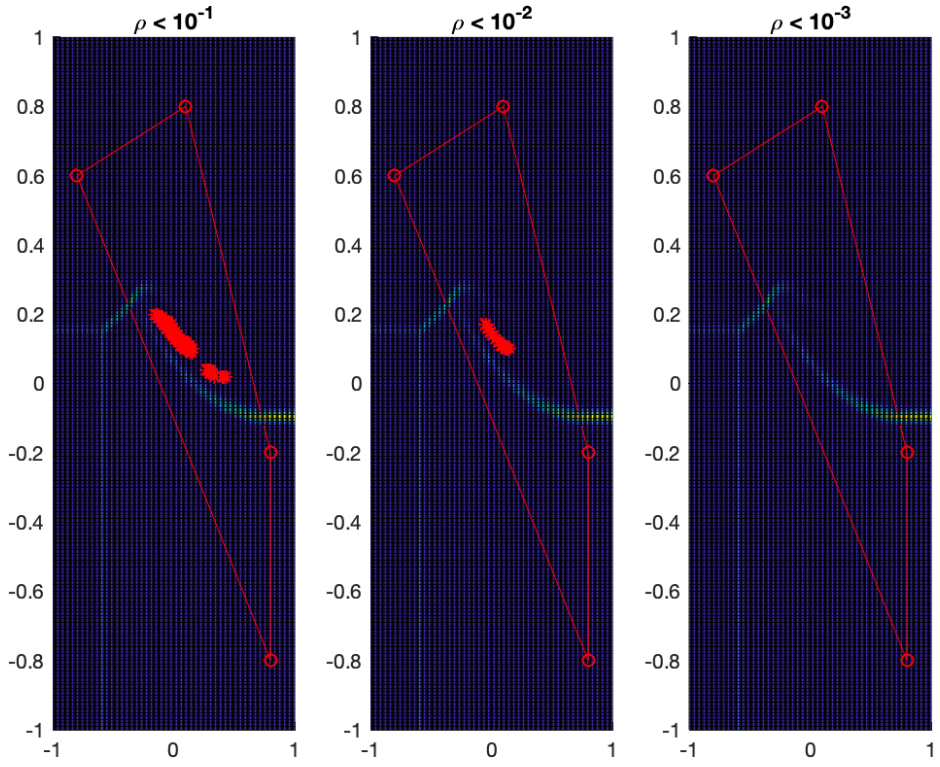
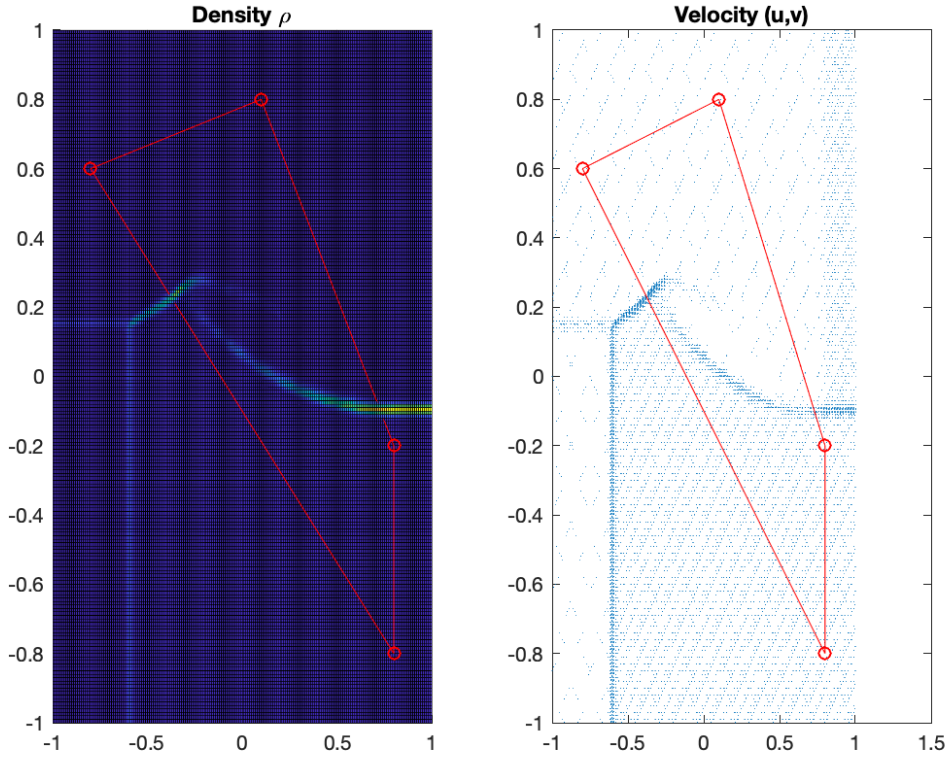


Figure 2.9: Case 10. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 6. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 3 \times 10^{-3}. \tag{2.2.76}$$

More details about the vacuum area are listed in the following table.



(a) Case 10. $N \times N = 200 \times 200$, and $T = 0.2$.

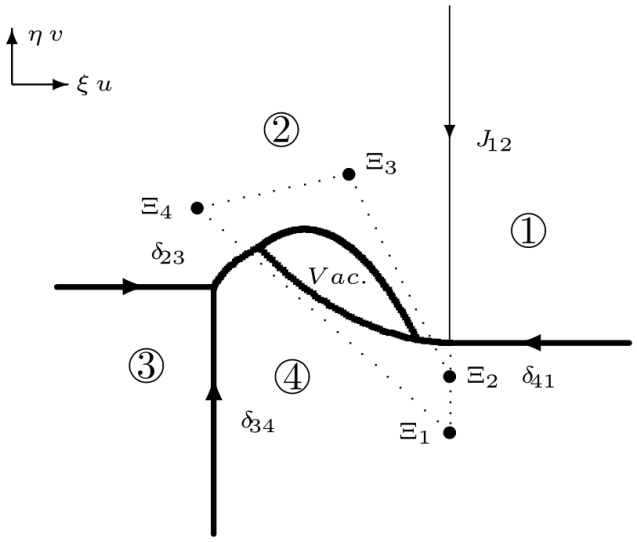


Fig. 4.10. The solution for Case 4.3(ii)b.

(b) Case 10. Screenshot from [Pan19]

Table 2.12: Case 10. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	88	20	0	0

Case 11.

[Pan19] gave the following restrictions

$$u_1 = u_2, u_4 < u_3, v_1 < v_4, v_2 < v_3, v_1 < v_2, \tag{2.2.77}$$

and

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_3 & v_3 & 1 \end{pmatrix} < 0, \text{ and } det\begin{pmatrix} u_1 & v_1 & 1 \\ u_2 & v_2 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} < 0. \tag{2.2.78}$$

We added more details in the restrictions

$$u_1 = u_2 < u_4 < u_3, v_1 < v_4 < v_2 < v_3. \tag{2.2.79}$$

So we took the following set of initial conditions

$$\rho_1 = 4, \rho_2 = 1, \rho_3 = 10, \rho_4 = 2, \tag{2.2.80}$$

and

$$u_1 = 2, u_2 = 2, u_3 = 4, u_4 = 3, v_1 = -2, v_2 = 2, v_3 = 4, v_4 = 0, \tag{2.2.81}$$

Case 11. consists of there delta shocks along the axis directions and one contact discontinuity. The contact discontinuity J_{12} meets with the delta shock δ_{23} and form a new delta shock δ_{13}^{A}. The two delta shocks δ_{34} and δ_{41} intersect and form a new delta

shock δ_{13}^{B}. The two newly developed delta shocks δ_{13}^{A} and δ_{13}^{B} meet at U_{δ}^{13}. U_{δ}^{13} has the Cartesian coordinates

$$U_{\delta}^{13} = \Big(\frac{\sqrt{\rho_1} u_1 + \sqrt{\rho_3} u_3}{\sqrt{\rho_1} + \sqrt{\rho_3}} T, \frac{\sqrt{\rho_1} v_1 + \sqrt{\rho_3} v_3}{\sqrt{\rho_1} + \sqrt{\rho_3}} T \Big). \tag{2.2.82}$$

Note that we didn't obtain the curvature in the diagram. [Pan19] did not explicitly state the conditions to achieve the curvature.

Case 12.

Case 12. consists of two delta shocks and two contact discontinuities along the axis directions. [Pan19] required

$$u_4 < u_3 < u_1 = u_2, v_1 < v_4, v_1 < v_{\delta}^{41} < v_2 = v_3. \tag{2.2.83}$$

In addition, the diagram required that

$$v_1 < v_4 < v_2 = v_3, \tag{2.2.84}$$

and

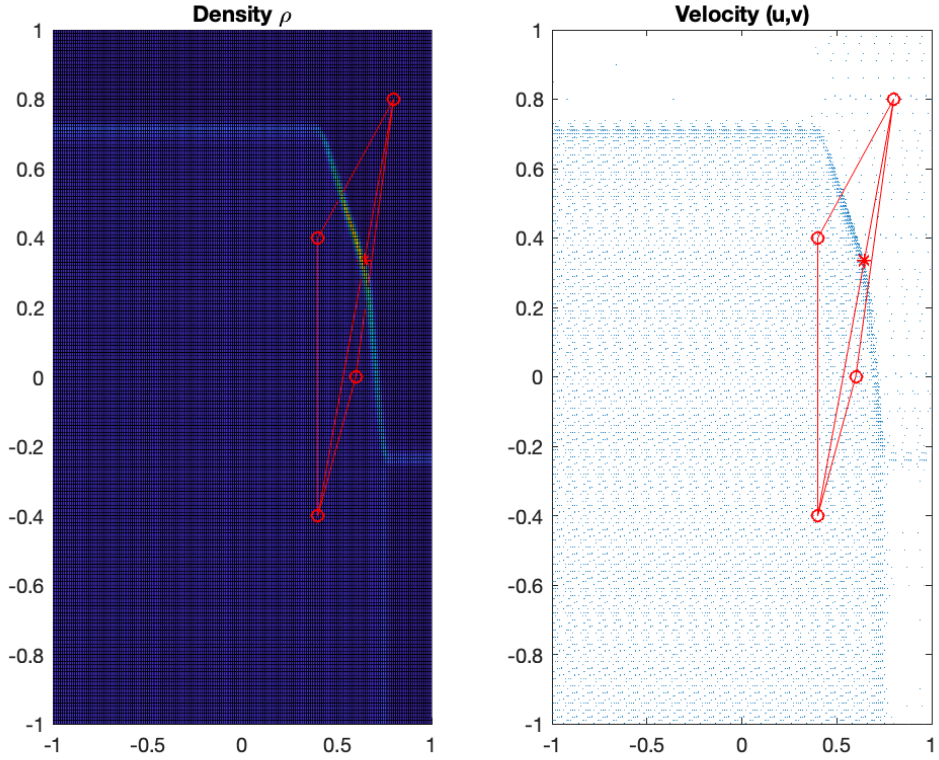
$$u_{\delta}^{34} > u_{\delta}^{24}. \tag{2.2.85}$$

We took the following set of initial conditions

$$\rho_1 = 1, \rho_2 = 2, \rho_3 = 1, \rho_4 = 1, \tag{2.2.86}$$

and

$$u_1 = 4, u_2 = 4, u_3 = -1, u_4 = -4, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = 1. \tag{2.2.87}$$



(a) Case 11. $N \times N = 200 \times 200$, and $T = 0.2$.

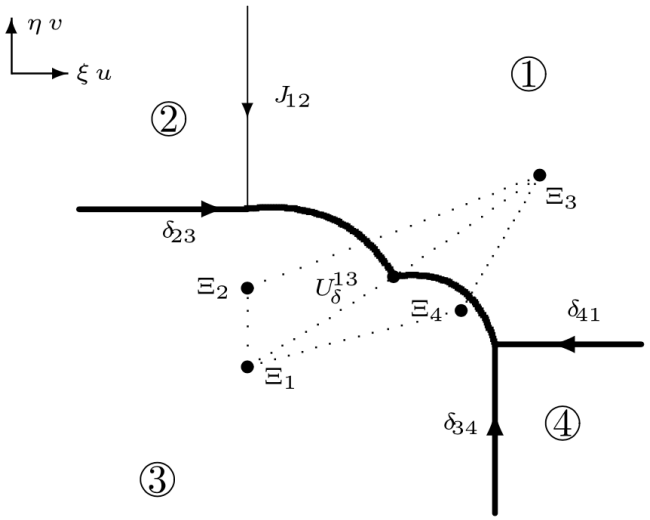
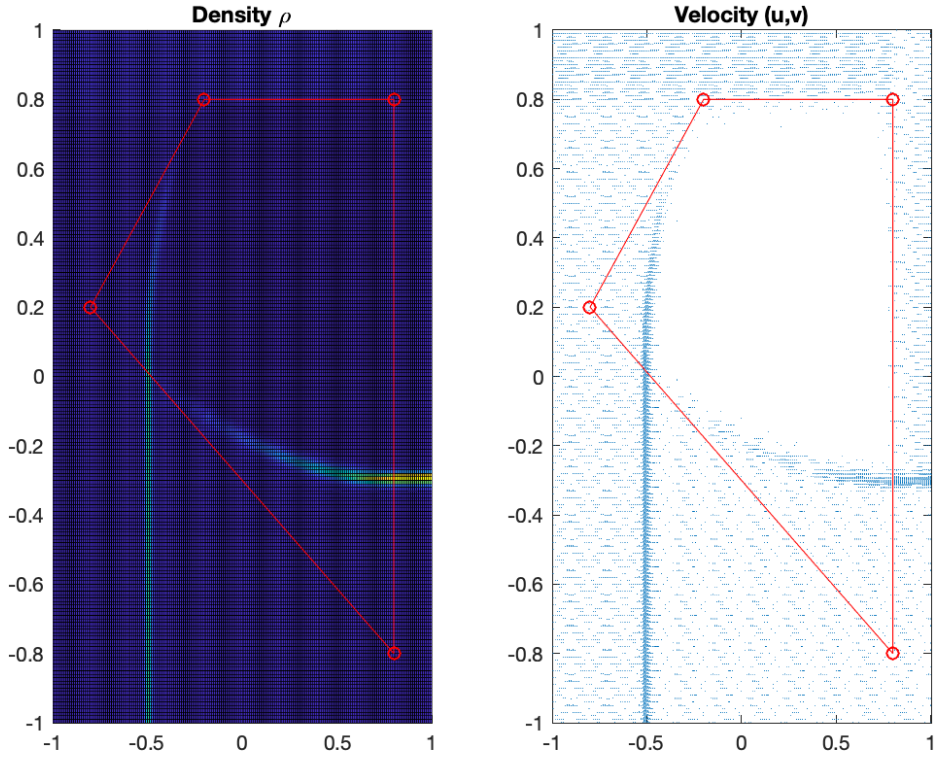


Fig. 4.11. The solution for Case 4.3(iii).

(b) Case 11. Screenshot from [Pan19]



(a) Case 12. $N \times N = 200 \times 200$, and $T = 0.2$.

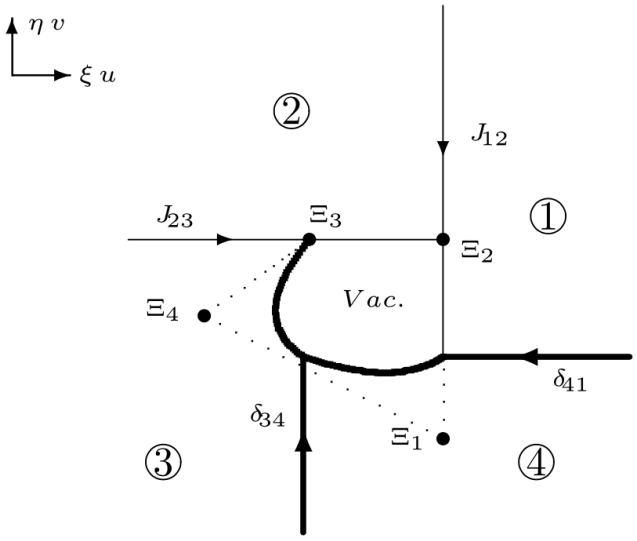


Fig. 4.12. The solution for Case 4.4(1)(i)a.

(b) Case 12. Screenshot from [Pan19]

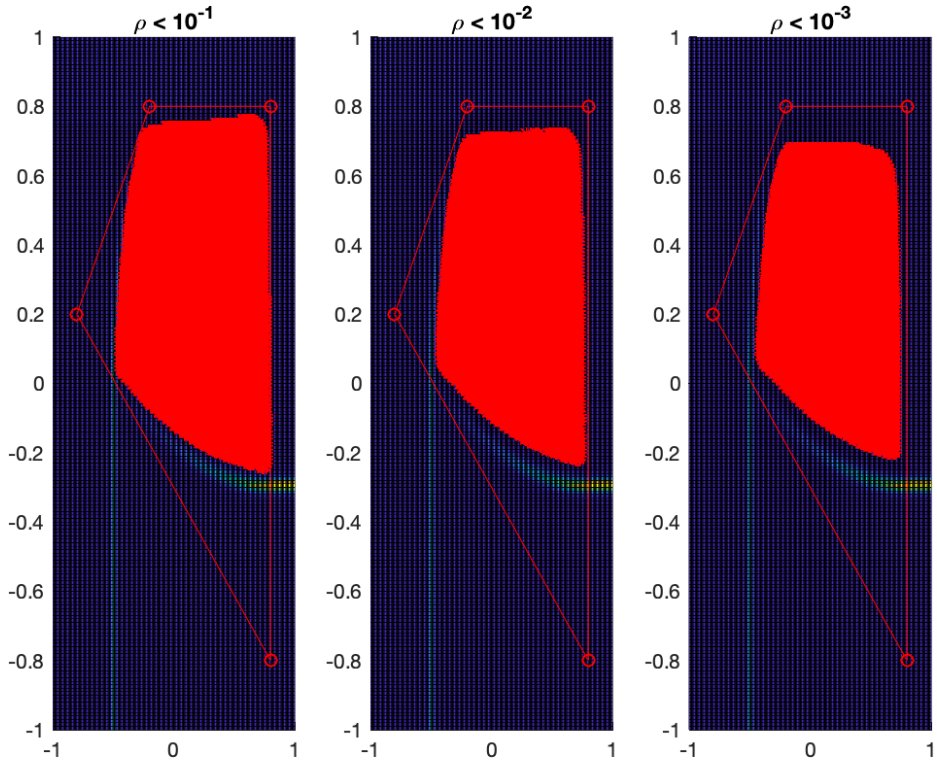


Figure 2.10: Case 12. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 12. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 2.1221 \times 10^{-12}. \tag{2.2.88}$$

More details about the vacuum area are listed in the following table.

Table 2.13: Case 12. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	10,108	9,011	8,182	7,426

Case 13.

[Pan19] gave the same initial conditions as Case 12. Note that the delta shock δ_{41} is below the point Xi_2 since $v_\delta^{41} < v_2$. We further studied the diagram in Case 13. and concluded the following additional conditions

$$v_1 < v_\delta^{41} < v_2 = v_3 < v_4, \tag{2.2.89}$$

and

$$u_\delta^{34} < u_\delta^{24}. \tag{2.2.90}$$

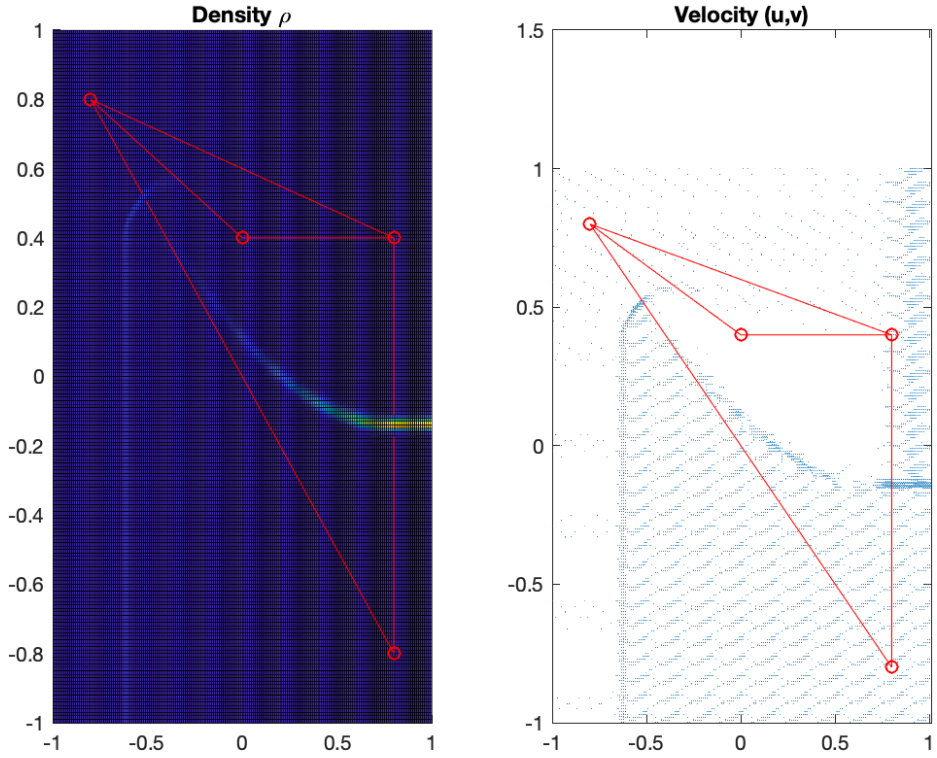
So we had the following set of initial conditions

$$\rho_1 = 20, \rho_2 = 2, \rho_3 = 1, \rho_4 = 10, \tag{2.2.91}$$

and

$$u_1 = 4, u_2 = 4, u_3 = 0, u_4 = -4, v_1 = -4, v_2 = 2, v_3 = 2, v_4 = 4. \tag{2.2.92}$$

Case 13. also consists of two delta shocks and two contact discontinuities along the axis directions. The contact discontinuity J_{23} meet with the delta shock δ_{34} and form a new delta shock. The contact discontinuity J_{12} meet with the delta shock δ_{41} and form a new delta shock.



(a) Case 13. $N \times N = 200 \times 200$, and $T = 0.2$.

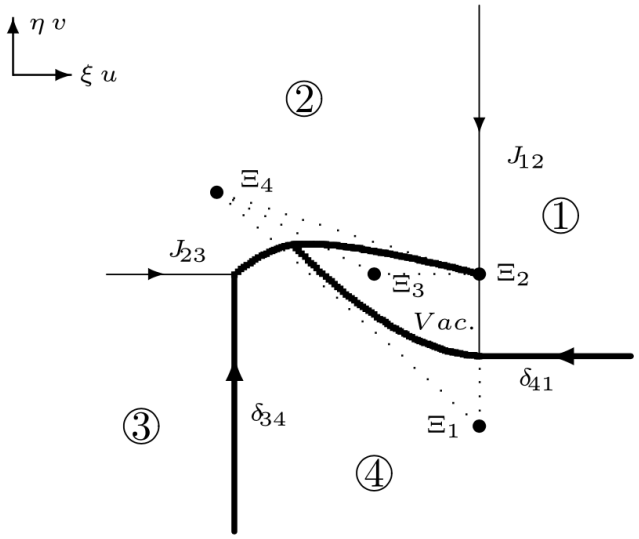


Fig. 4.13. The solution for Case 4.4(1)(i)b.

(b) Case 13. Screenshot from [Pan19]

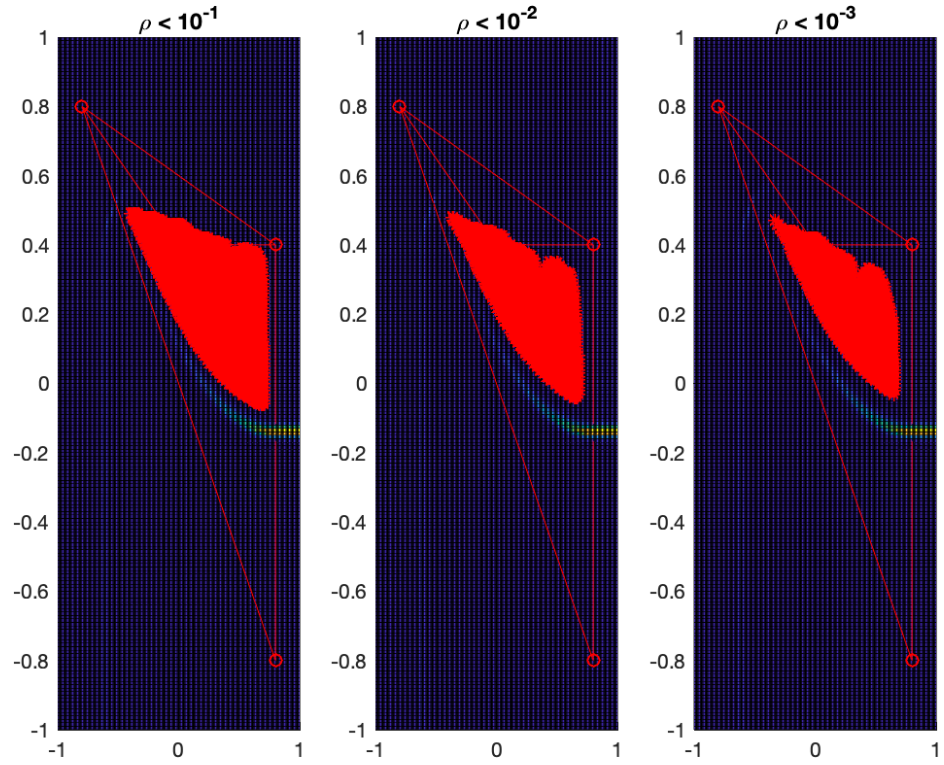


Figure 2.11: Case 13. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 13. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 2.6446 \times 10^{-9}. \tag{2.2.93}$$

More details about the vacuum area are listed in the following table.

Table 2.14: Case 13. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	3,169	2,379	1,975	1,643

Case 14.

[Pan19] stated the initial conditions to be

$$u_4 < u_3 < u_1 = u_2, v_1 < v_4, v_2 = v_3, v_1 < v_2 < v_\delta^{41}. \tag{2.2.94}$$

We simplified the above conditions to be

$$u_4 < u_3 < u_1 = u_2, v_1 < v_2 = v_3 < v_\delta^{41} < v_4. \tag{2.2.95}$$

Together with the additional condition regarding the delta shock δ_{34} meeting the contact discontinuity J_{23}

$$u_\delta^{34} < u_\delta^{24}, \tag{2.2.96}$$

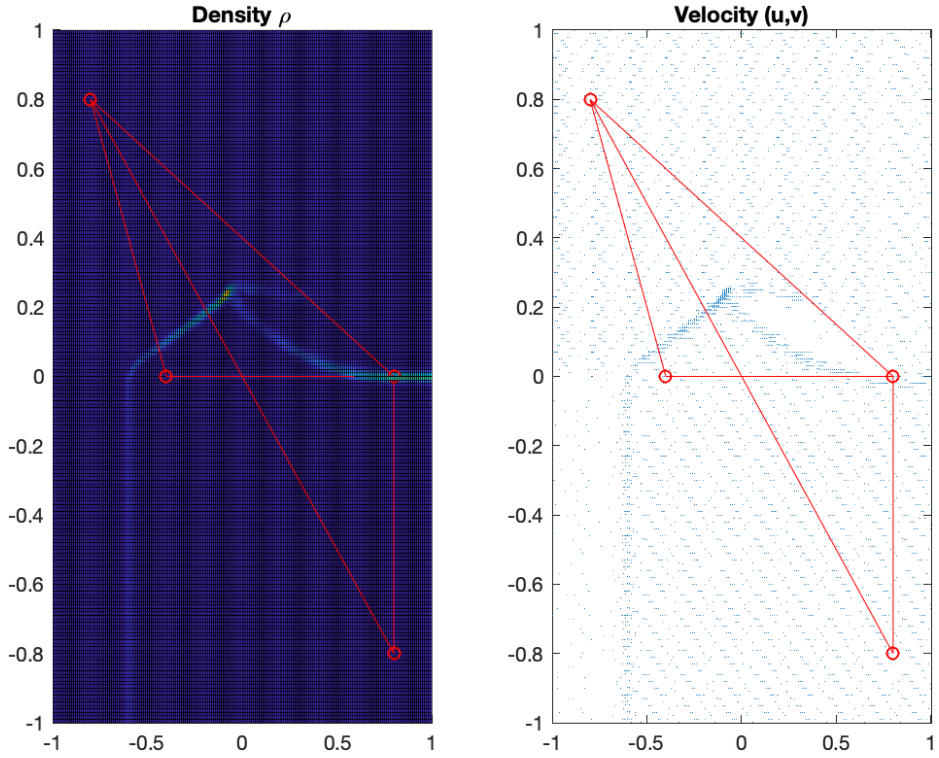
we come to the following set of initial conditions for Case 14.

$$\rho_1 = 1, \rho_2 = 2, \rho_3 = 1, \rho_4 = 1, \tag{2.2.97}$$

and

$$u_1 = 4, u_2 = 4, u_3 = -2, u_4 = -4, v_1 = -4, v_2 = 0, v_3 = 0, v_4 = 4. \tag{2.2.98}$$

Case 14. also consists of two delta shocks and two contact discontinuities along the axis directions. The contact discontinuity J_{23} meet with the delta shock δ_{34} and form a new delta shock δ_{24}^A. The contact discontinuity J_{12} meet with the delta shock δ_{41} and form a new delta shock δ_{24}^B. The two newly developed delta shocks δ_{24}^A and δ_{24}^B form a vacuum area in between.



(a) Case 14. $N \times N = 200 \times 200$, and $T = 0.2$.

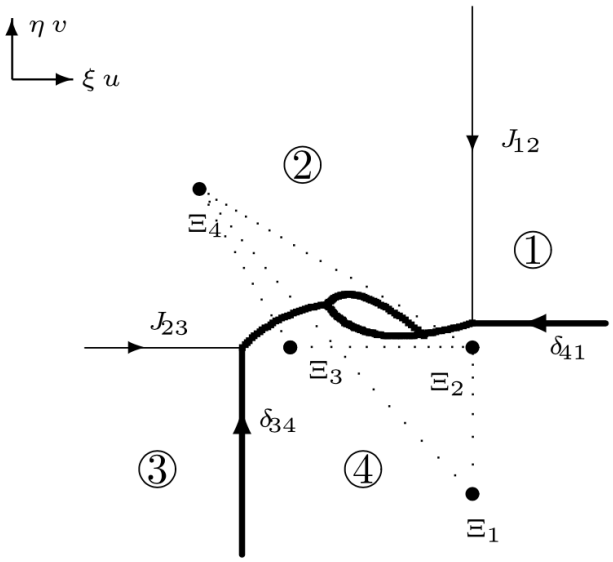


Fig. 4.14. The solution for Case 4.4(1)(ii).

(b) Case 14. Screenshot from [Pan19]

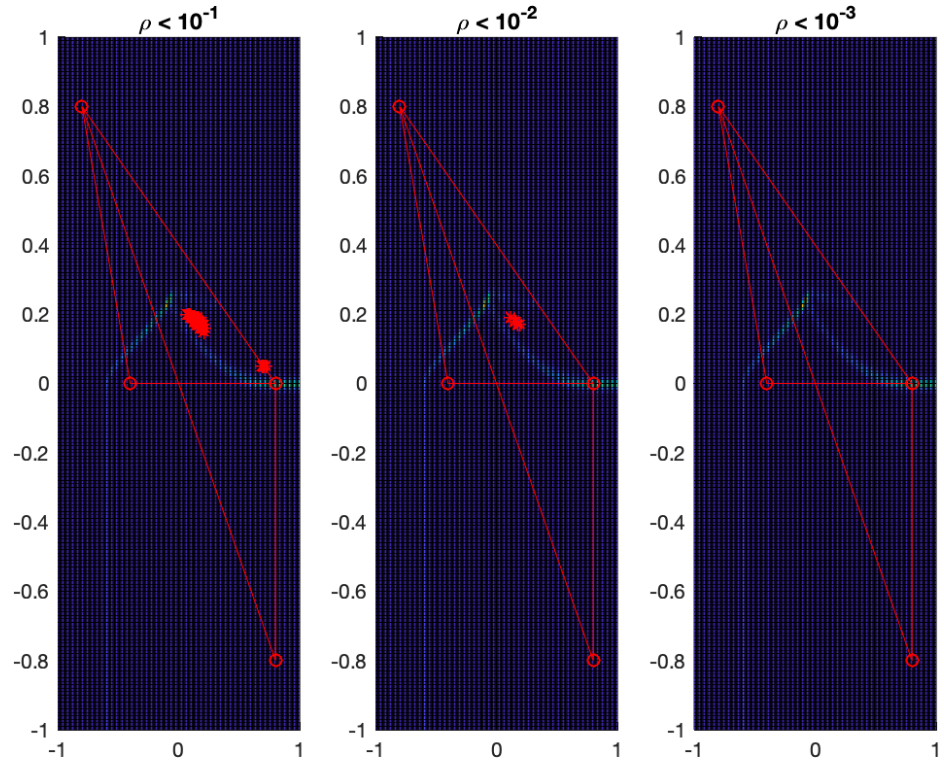


Figure 2.12: Case 14. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 14. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 4.4 \times 10^{-3}. \tag{2.2.99}$$

More details about the vacuum area are listed in the following table.

Table 2.15: Case 14. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	35	5	0	0

Case 15.

[Pan19] stated the initial conditions to be

$$u_4 < u_3 < u_1 = u_2, v_2 = v_3 < v_1 < v_4. \tag{2.2.100}$$

There is no explicit condition for densities. So we took the following initial conditions

$$\rho_1 = 3, \rho_2 = 4, \rho_3 = 3, \rho_4 = 4, \tag{2.2.101}$$

and

$$u_1 = 4, u_2 = 4, u_3 = -2, u_4 = -4, v_1 = -2, v_2 = -4, v_3 = -4, v_4 = 4. \tag{2.2.102}$$

Case 15. does not have vacuum area in the solution. The contact discontinuity J_{23} meet with the delta shock δ_{34} and form a new delta shock δ_{24}^A. The contact discontinuity J_{12} meet with the delta shock δ_{41} and form a new delta shock δ_{24}^B. The two newly developed delta shocks δ_{24}^A and δ_{24}^B intersect at U_δ^{24}. U_δ^{24} has the Cartesian coordinates

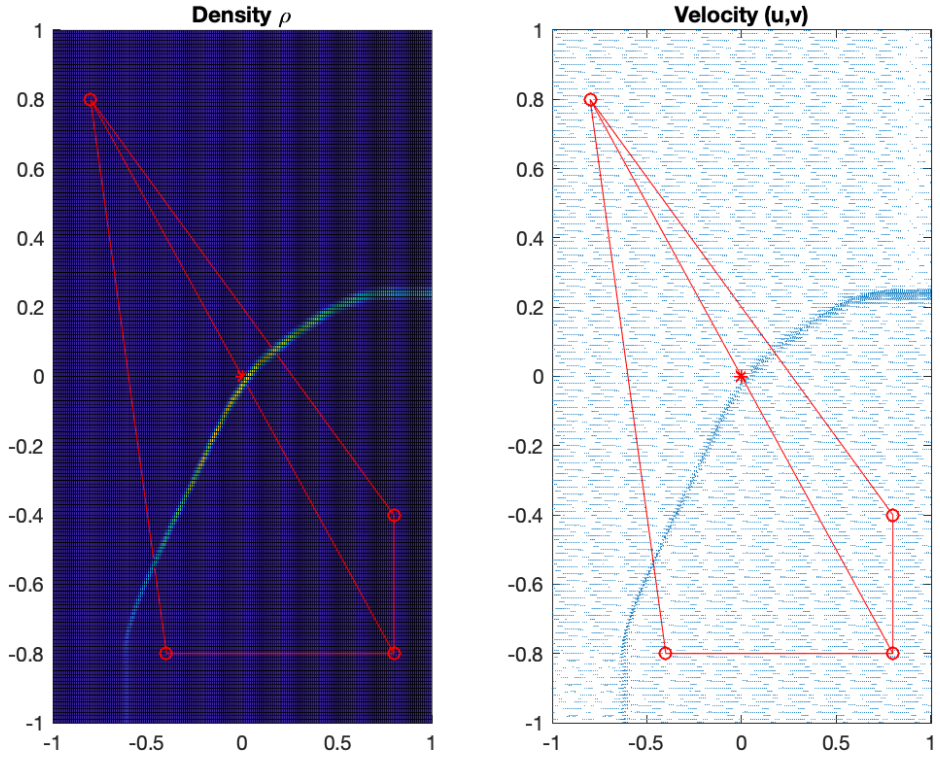
$$U_\delta^{24} = \Big(\frac{\sqrt{\rho_2}u_2 + \sqrt{\rho_4}u_4}{\sqrt{\rho_2} + \sqrt{\rho_4}} T, \frac{\sqrt{\rho_2}v_2 + \sqrt{\rho_4}v_4}{\sqrt{\rho_2} + \sqrt{\rho_4}} T \Big). \tag{2.2.103}$$

We still could not obtain the curvature showed in the diagram. We guess the choice of the four initial density values is in charge of the shape of the solution, but we do not have any analytical proof for this.

Case 16.

[Pan19] stated the initial conditions to be

$$u_4 < u_1 = u_2 < u_3, v_1 < v_2 = v_3, v_1 < v_4. \tag{2.2.104}$$



(a) Case 15. $N \times N = 200 \times 200$, and $T = 0.2$.

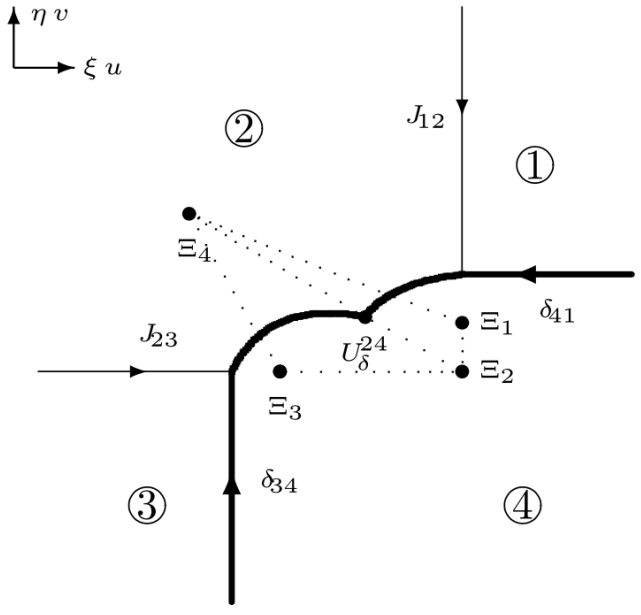


Fig. 4.15. The solution for Case 4.4(1)(iii).

(b) Case 15. Screenshot from [Pan19]

The diagram for Case 16. required explicitly that

$$v_1 < v_4 < v_2 = v_3. \tag{2.2.105}$$

We further added the following condition to guarantee the vacuum solution

$$u_\delta^{34} < u_1, v_\delta^{41} < v_\delta^{13}. \tag{2.2.106}$$

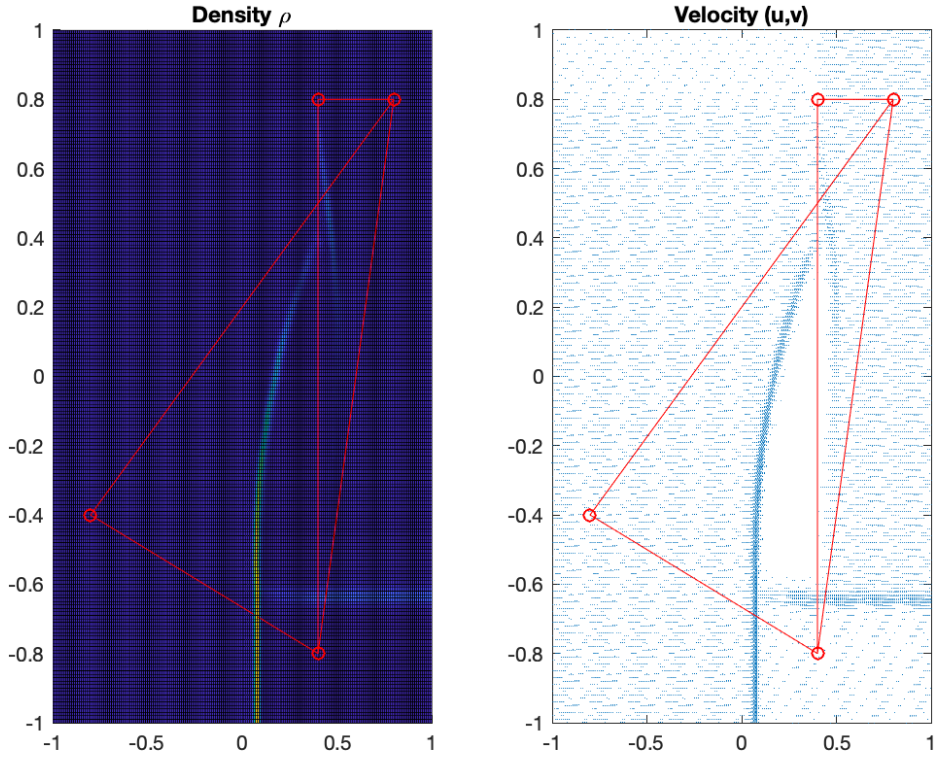
We obtained these conditions inspired by Case 5. and Case 6. So we took the following set of initial conditions

$$\rho_1 = 4, \rho_2 = 1, \rho_3 = 3, \rho_4 = 2, \tag{2.2.107}$$

and

$$u_1 = 2, u_2 = 2, u_3 = 4, u_4 = -4, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = -2. \tag{2.2.108}$$

Case 16. also consists of two delta shocks and two contact discontinuities along the axis directions. But this time, the two contact discontinuities meet at the point Xi_2 and form a new delta shock. The newly developed delta shock surround the vacuum area together with the other two delta shocks δ_{34} and δ_{41}.



(a) Case 16. $N \times N = 200 \times 200$, and $T = 0.2$.

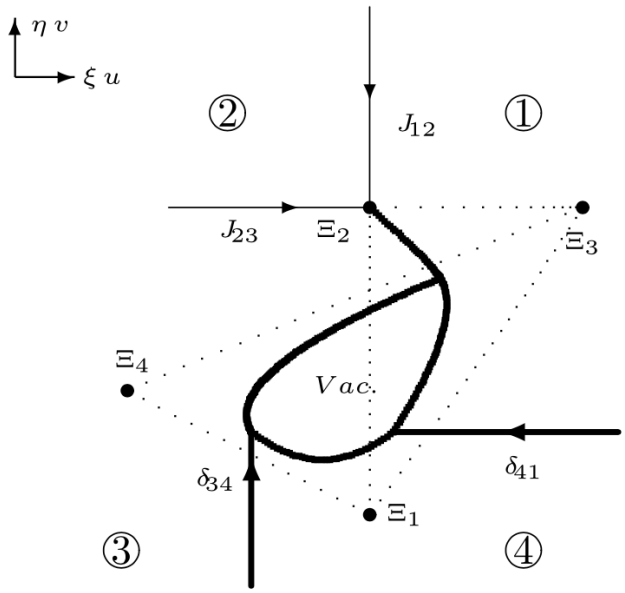


Fig. 4.16. The solution for Case 4.4(1)(iv)a.

(b) Case 16. Screenshot from [Pan19]

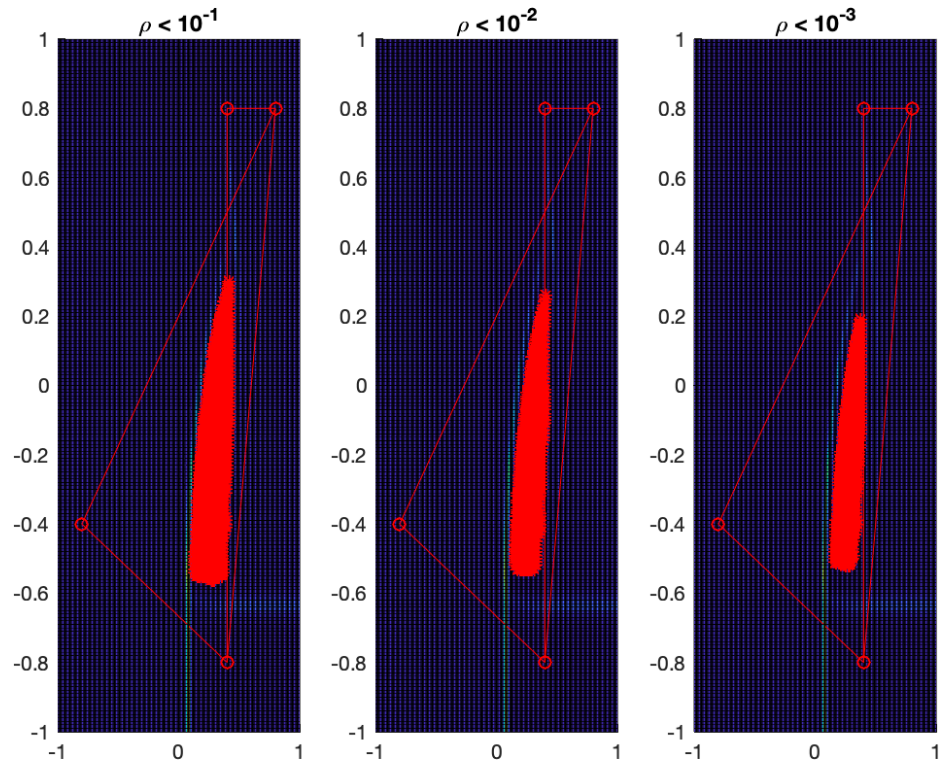


Figure 2.13: Case 16. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 16. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 5.1140 \times 10^{-6}. \tag{2.2.109}$$

More details about the vacuum area are listed in the following table.

Table 2.16: Case 16. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	1,217	764	508	152

Case 17.

[Pan19] gave the same conditions as those for Case 16. Clearly we need additional restrictions.

$$u_4 < u_1 = u_2 < u_3, v_1 < v_4 < v_2 = v_3, \tag{2.2.110}$$

and

$$u_\delta^{34} > u_1, v_\delta^{41} < v_\delta^{13}. \tag{2.2.111}$$

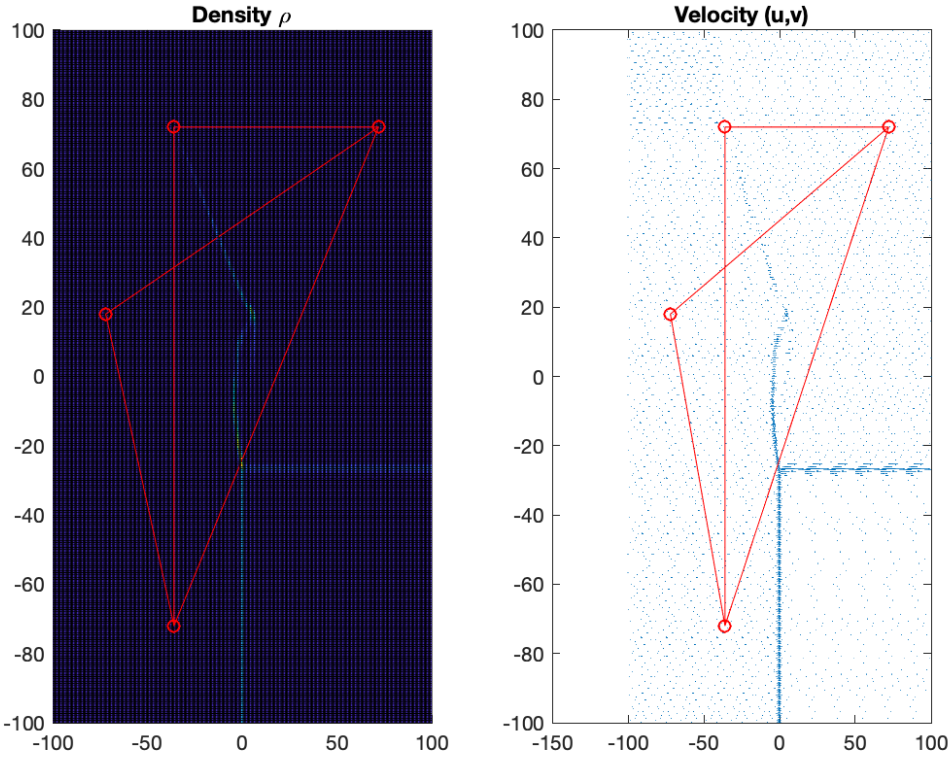
We eventually took the following set of initial conditions

$$\rho_1 = 1, \rho_2 = 2, \rho_3 = 1, \rho_4 = 1, \tag{2.2.112}$$

and

$$u_1 = -2, u_2 = -2, u_3 = 4, u_4 = -4, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = 1. \tag{2.2.113}$$

Case 17. is another case that our finite difference does not give a very good numerical solution. Note that the only difference between Case 17. and Case 16. is the location of the delta shock δ_{34}. The delta shock δ_{34} travels faster in the x-direction so it meets with the delta shock δ_{41} instead of the delta shock on the left. Because the vacuum area of Case 17. is small and has fewer points with density close to zero. We set the numerical domain to be $(-100, 100)$ and the final time $T = 18$. Moreover, we have 400×400 grid points for this case.



(a) Case 17. $N \times N = 300 \times 300$, and $T = 18$.

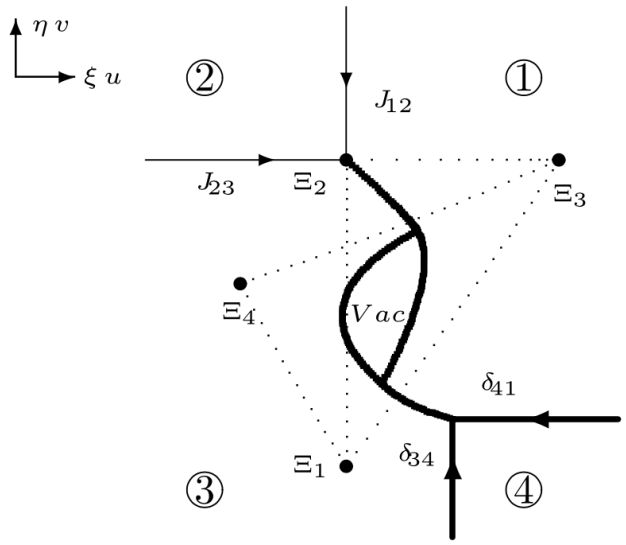


Fig. 4.17. The solution for Case 4.4(1)(iv)b.

(b) Case 17. Screenshot from [Pan19]

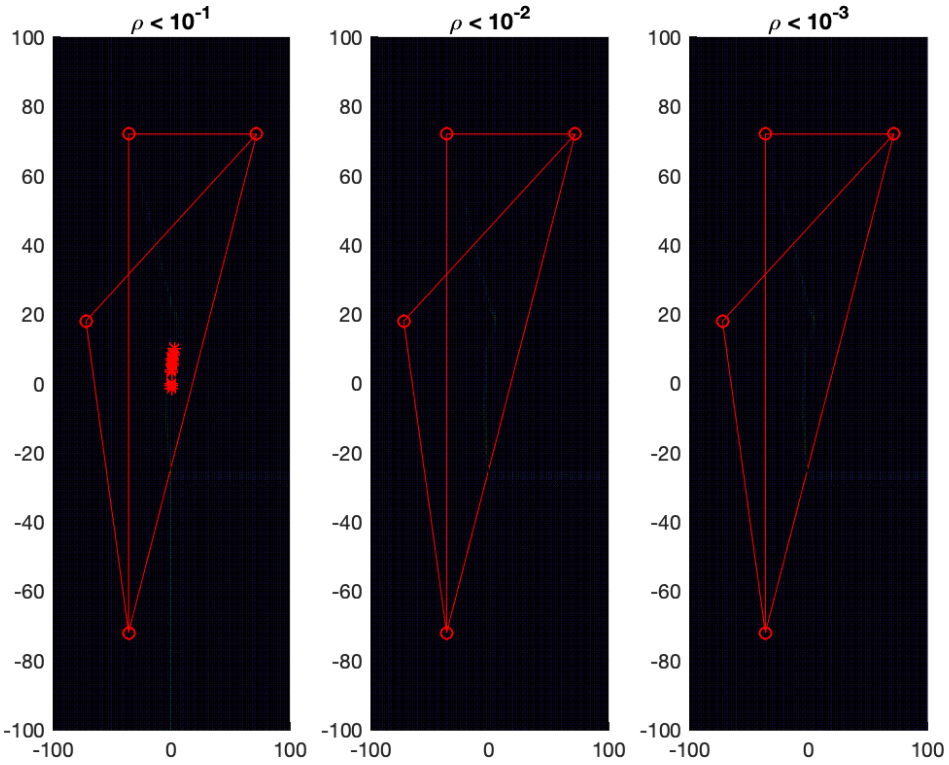


Figure 2.14: Case 17. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 17. there are $301 \times 301 = 90,601$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 3.41 \times 10^{-2}. \tag{2.2.114}$$

More details about the vacuum area are listed in the following table.

Table 2.17: Case 17. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	90,601	12	0	0	0

Case 18.

[Pan19] required that the initial conditions satisfied

$$u_1 = u_2 < u_4 < u_3, v_1 < v_2 = v_3, v_1 < v_4, \tag{2.2.115}$$

and

$$det\begin{pmatrix} u_1 & v_1 & 1 \\ u_3 & v_3 & 1 \\ u_4 & v_4 & 1 \end{pmatrix} < 0. \tag{2.2.116}$$

After studying the diagram for Case 17. we needed one more condition

$$v_1 < v_4 < v_\delta^{13} < v_2 = v_3. \tag{2.2.117}$$

So we took the initial conditions as

$$\rho_1 = 4, \rho_2 = 2, \rho_3 = 1, \rho_4 = 2, \tag{2.2.118}$$

and

$$u_1 = -4, u_2 = -4, u_3 = 4, u_4 = 2, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = -2. \tag{2.2.119}$$

Case 18. does not have vacuum area in the solution. The contact discontinuities J_{12} and J_{23} intersect at the point Ξ_2 and develop a new delta shock $\delta_{13}^{\Xi_2}$. The two delta shocks δ_{34} and δ_{41} meet and form a new shock heading towards U_δ^{13}. The two newly developed delta shocks meet at the point U_δ^{13}. U_δ^{13} has the Cartesian coordinates

$$U_\delta^{13} = \Big(\frac{\sqrt{\rho_1}u_1 + \sqrt{\rho_3}u_3}{\sqrt{\rho_1} + \sqrt{\rho_3}}T, \frac{\sqrt{\rho_1}v_1 + \sqrt{\rho_3}v_3}{\sqrt{\rho_1} + \sqrt{\rho_3}}T\Big). \tag{2.2.120}$$

Case 19.

[Pan19] gave the following initial conditions

$$u_3 = u_4 < u_1 = u_2, v_1 < v_\delta^{41} < v_2 < v_3, v_1 < v_4 < v_\delta^{23} < v_3. \tag{2.2.121}$$

After studying the diagram for Case 19. we simplified the second conditions in (2.2.121) to be

$$v_1 < v_4 < v_2 < v_3. \tag{2.2.122}$$

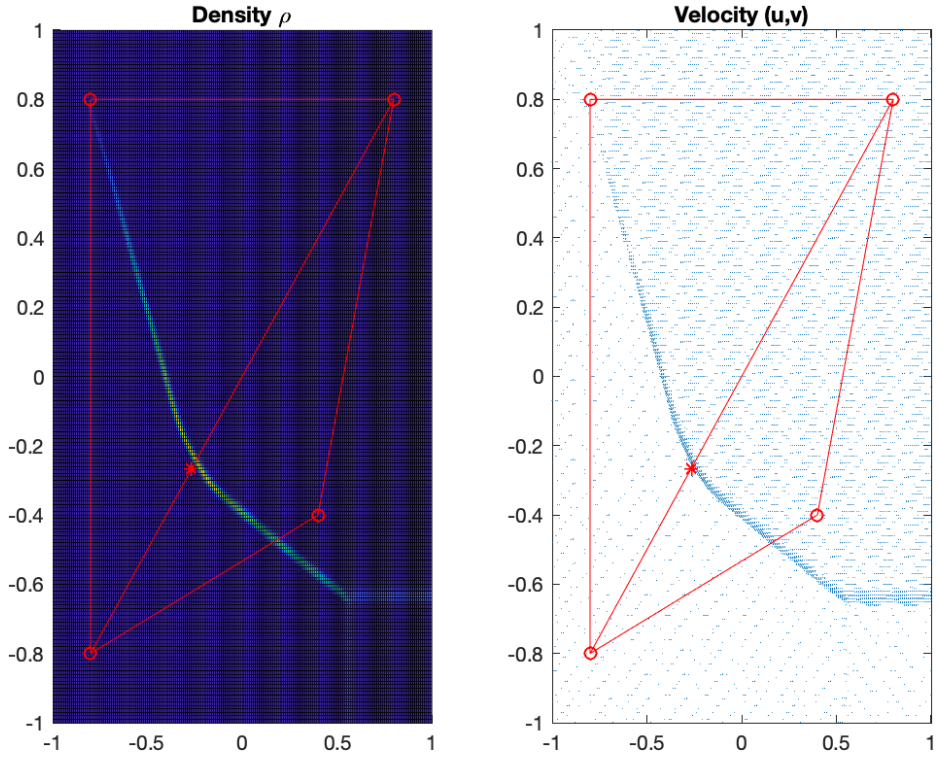
So we took

$$\rho_1 = 4, \rho_2 = 3, \rho_3 = 2, \rho_4 = 1, \tag{2.2.123}$$

and

$$u_1 = 3, u_2 = 3, u_3 = -3, u_4 = -3, v_1 = -4, v_2 = 2, v_3 = 4, v_4 = -2. \tag{2.2.124}$$

Case 19. consists of delta shocks, contact discontinuities and a vacuum area in the solution. The delta shock δ_{23} meets with the contact discontinuity J_{34} before J_{34} arrives at the point Ξ_3. The newly developed delta shock stops at the point Ξ_2. On the right hand side, the delta shock δ_{41} meets with the contact discontinuity J_{12} before J_{12} arrives at the point Ξ_1. The newly developed delta shock stops at the point Ξ_4. The vacuum area in the center is surrounded with two delta shocks and two contact discontinuities.



(a) Case 18. $N \times N = 200 \times 200$, and $T = 0.2$.

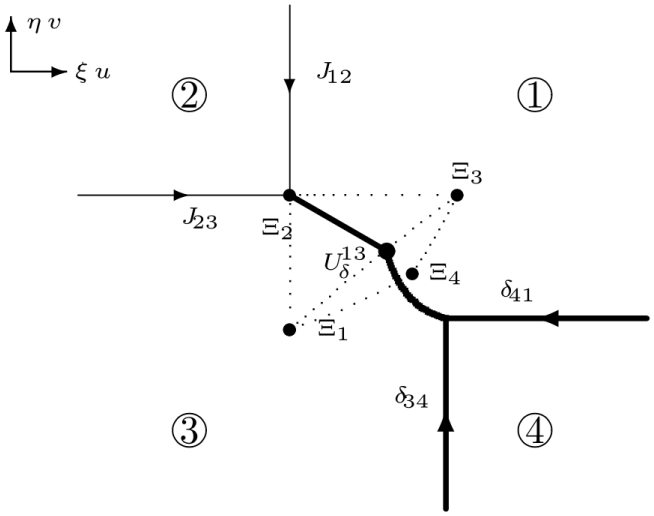
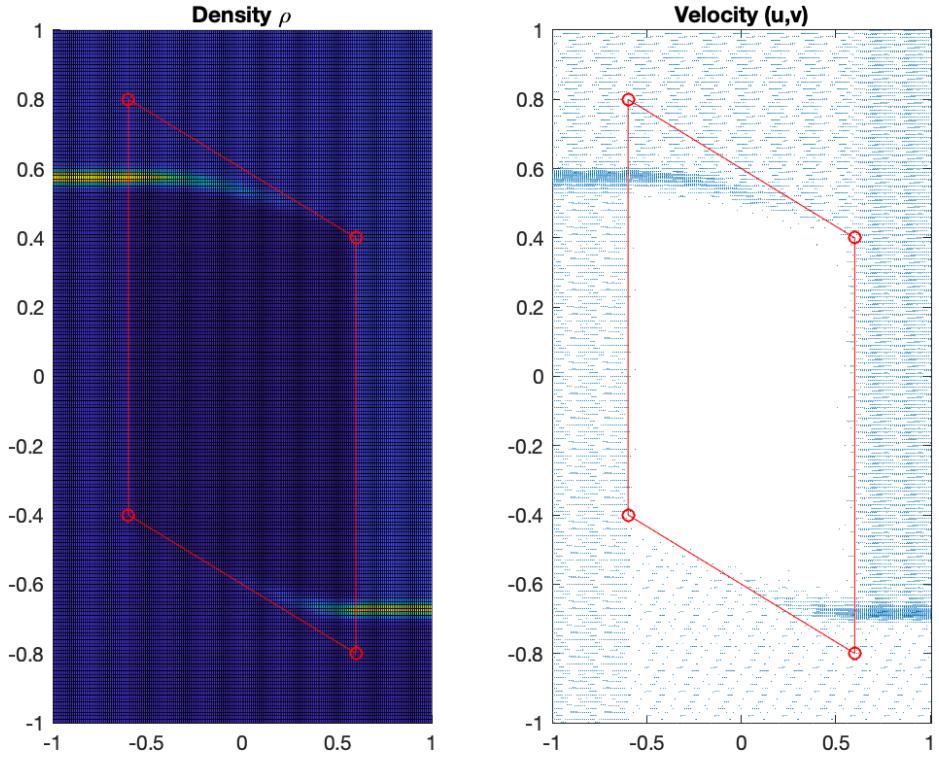


Fig. 4.18. The solution for Case 4.4(1)(v).

(b) Case 18. Screenshot from [Pan19]



(a) Case 19. $N \times N = 200 \times 200$, and $T = 0.2$.

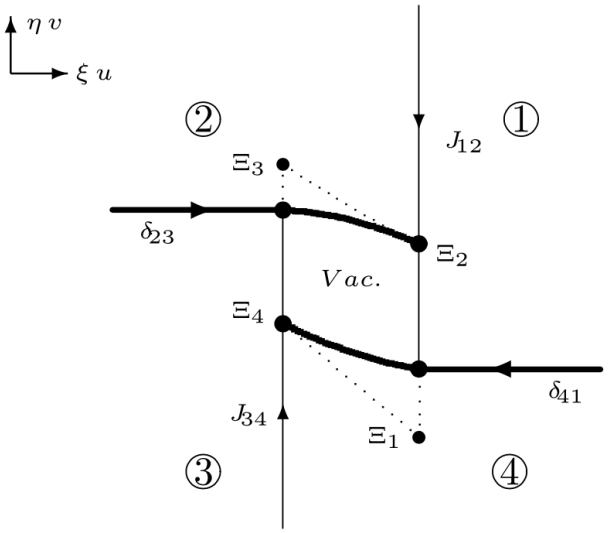


Fig. 4.19. The solution for Case 4.4(2).

(b) Case 19. Screenshot from [Pan19]

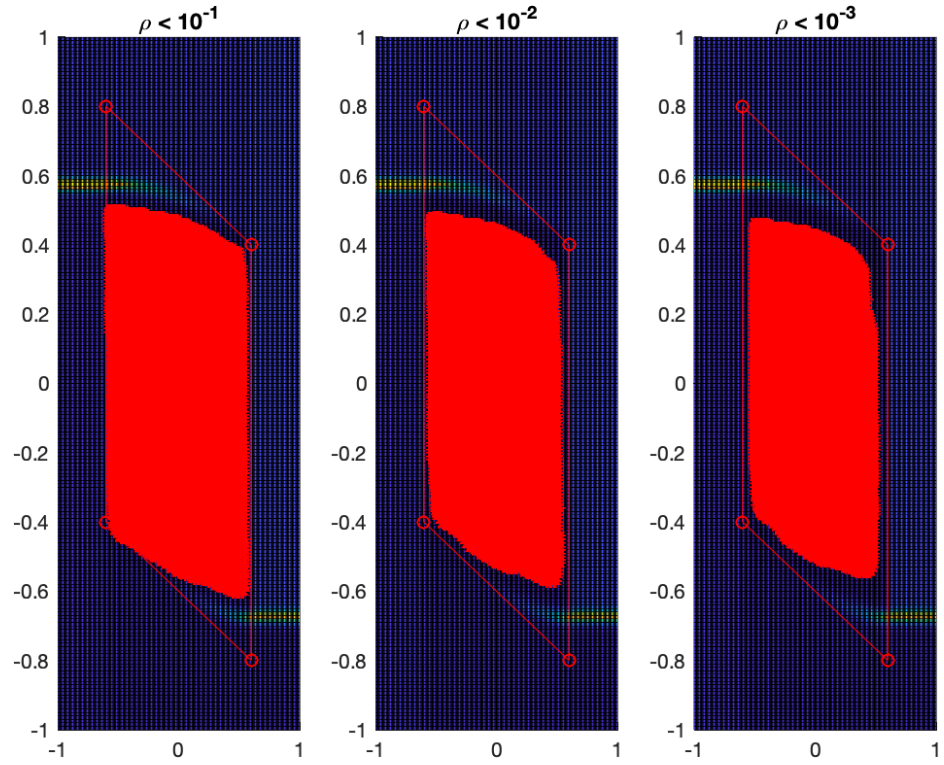


Figure 2.15: Case 19.The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 19. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 7.2903 \times 10^{-12}. \tag{2.2.125}$$

More details about the vacuum area are listed in the following table.

Table 2.18: Case 19. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	10,828	9,589	8,802	7,991

Case 20.

[Pan19] required that

$$u_3 = u_4 < u_1 < u_2, v_4 = v_1 < v_2 = v_3. \tag{2.2.126}$$

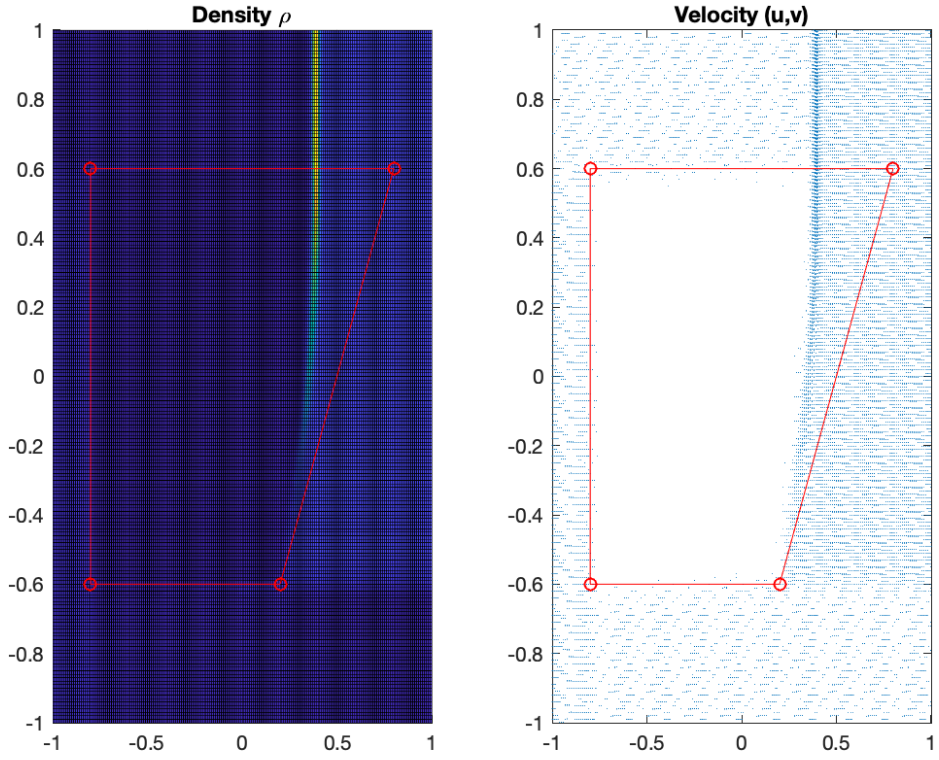
We do not need to add any additional conditions for Case 20. So we took the following initial conditions

$$\rho_1 = 4, \rho_2 = 1, \rho_3 = 2, \rho_4 = 1, \tag{2.2.127}$$

and

$$u_1 = 1, u_2 = 4, u_3 = -4, u_4 = -4, v_1 = -3, v_2 = 3, v_3 = 3, v_4 = -3. \tag{2.2.128}$$

Case 20. also consists delta shocks, contact discontinuities and a vacuum area in the solution. The contact discontinuities J_{34} and J_{41} meet at the point Ξ_4. The latter stops at the point Ξ_4 while the former continues to stop at the point Ξ_3. The contact discontinuity J_{23} meets with the contact discontinuity J_{34} and continues to meet with the delta shock δ_{12} before reaching the point Ξ_2. And thus a new delta shock is developed and meets with the contact discontinuity J_{41}. The newly developed delta shock eventually stops at the point Ξ_1.



(a) Case 20. $N \times N = 200 \times 200$, and $T = 0.2$.

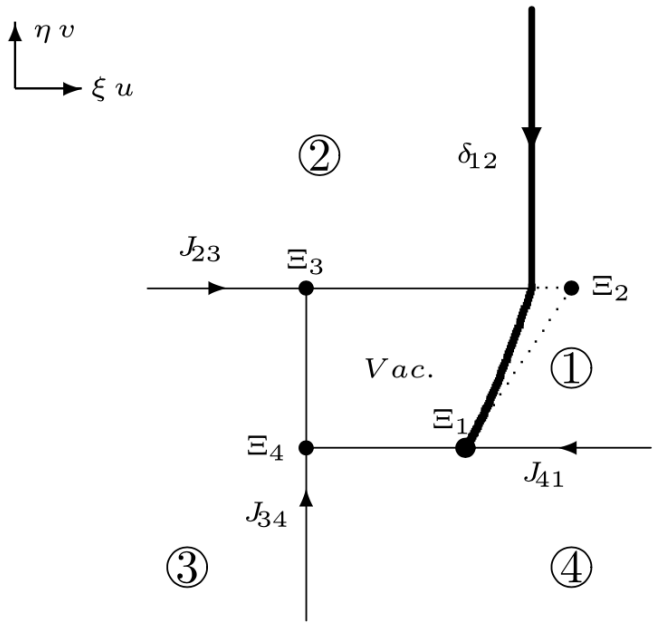


Fig. 4.20. The solution for Case 4.5(i).

(b) Case 20. Screenshot from [Pan19]

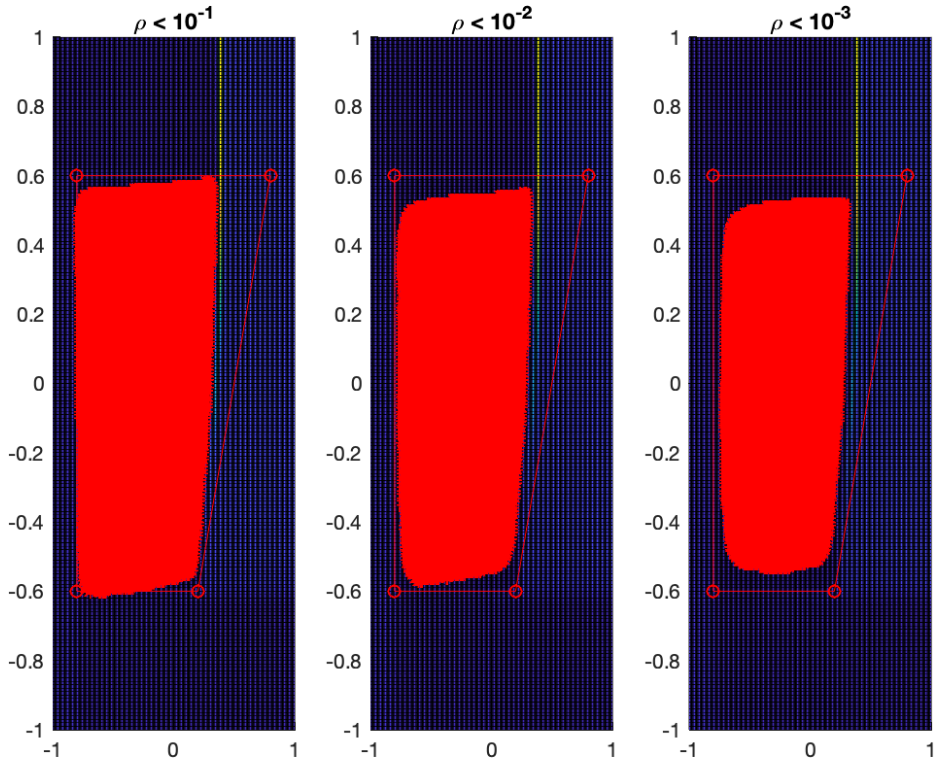


Figure 2.16: Case 20.The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 20. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 1.5430 \times 10^{-12}. \tag{2.2.129}$$

More details about the vacuum area are listed in the following table.

Table 2.19: Case 20. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	11,728	10,443	9,473	8,610

Case 21.

Case 21. has an incorrect diagram in [Pan19]. Pang explicitly stated that Case 21. required

$$u_1 < u_\delta^{12} < u_3 = u_4 < u_2, v_1 = v_4 < v_2 = v_3. \tag{2.2.130}$$

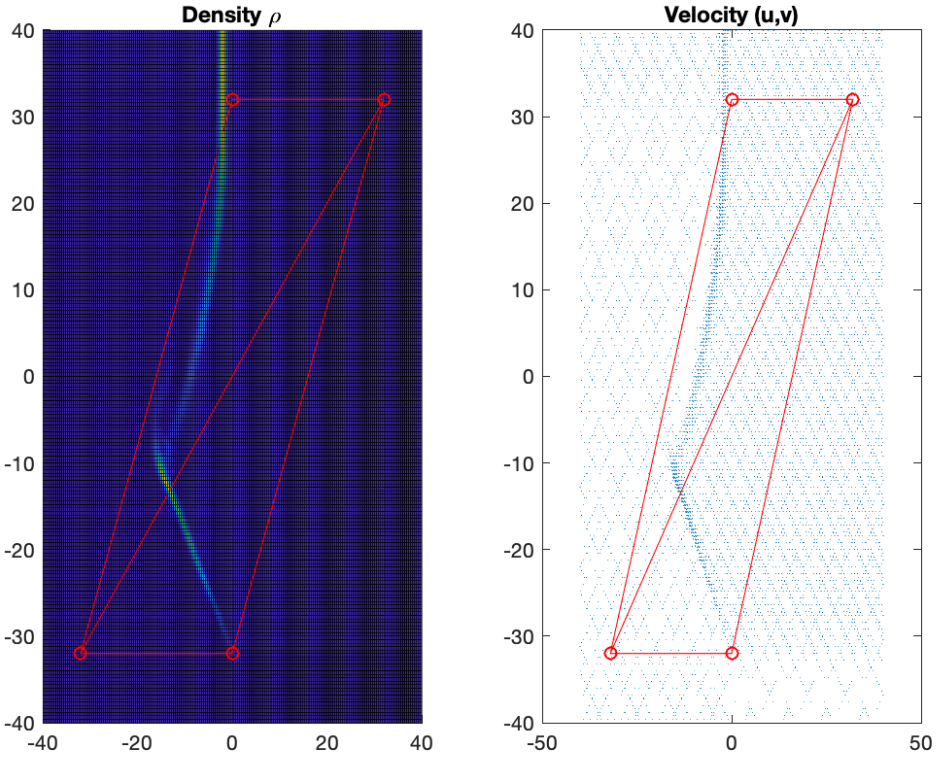
In particular, $u_3 = u_4$ leads to the two points Ξ_3 and Ξ_4 are supposed to be on the same vertical line. The diagram is accurate so we have no idea what the shape of the vacuum area should be like. We took the following initial conditions based on (2.2.130)

$$\rho_1 = 4, \rho_2 = 3, \rho_3 = 2, \rho_4 = 1, \tag{2.2.131}$$

and

$$u_1 = -4, u_2 = 4, u_3 = 0, u_4 = 0, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = -4. \tag{2.2.132}$$

We further extended the numerical domain to be $[-40, 40]$ and the final time $T = 8$ to obtain a better shaped result.



(a) Case 21. $N \times N = 200 \times 200$, and $T = 8$.

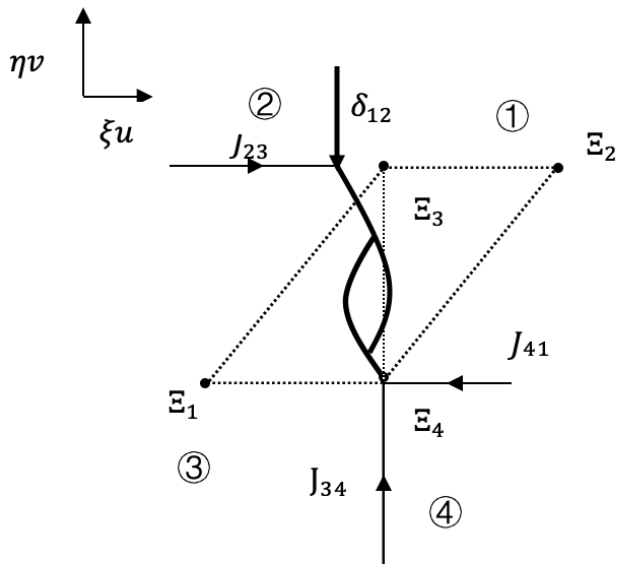


Fig. 4.21. The solution for Case 4.5(ii)a.

(b) Case 21. Diagram from [Pan19] is inaccurate. We draw this based on the given initial condition.

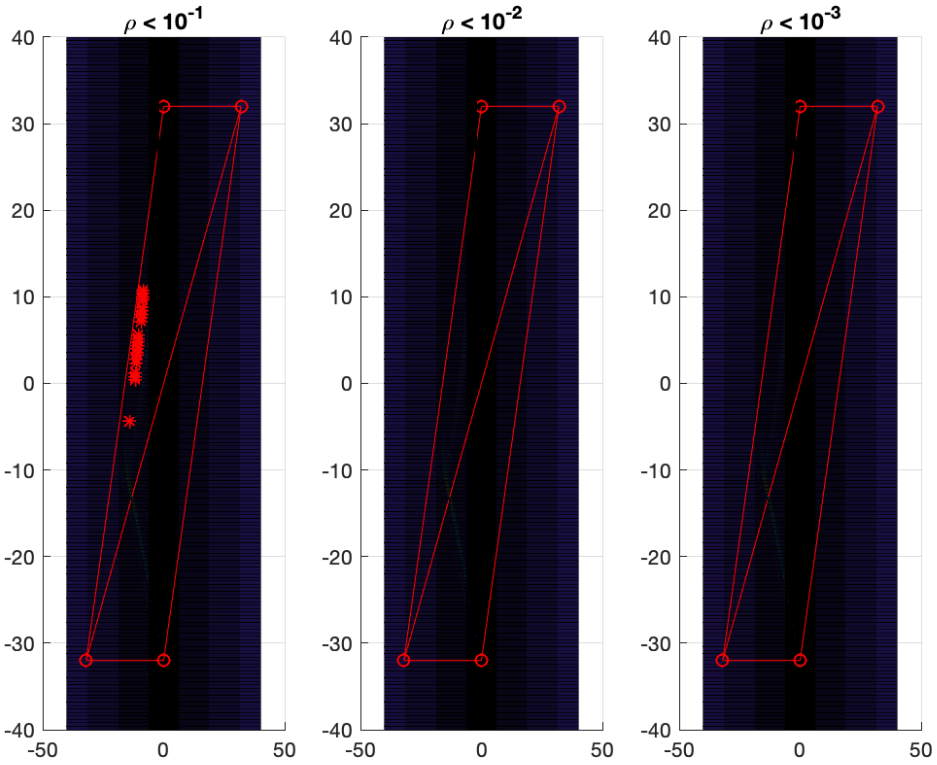


Figure 2.17: Case 21. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 21. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 4.34 \times 10^{-2}. \tag{2.2.133}$$

More details about the vacuum area are listed in the following table.

Table 2.20: Case 21. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	22	0	0	0

Case 22.

[Pan19] gave the initial conditions

$$u_1 < u_3 = u_4 < u_\delta^{12} < u_2, v_1 = v_4 < v_2 = v_3. \tag{2.2.134}$$

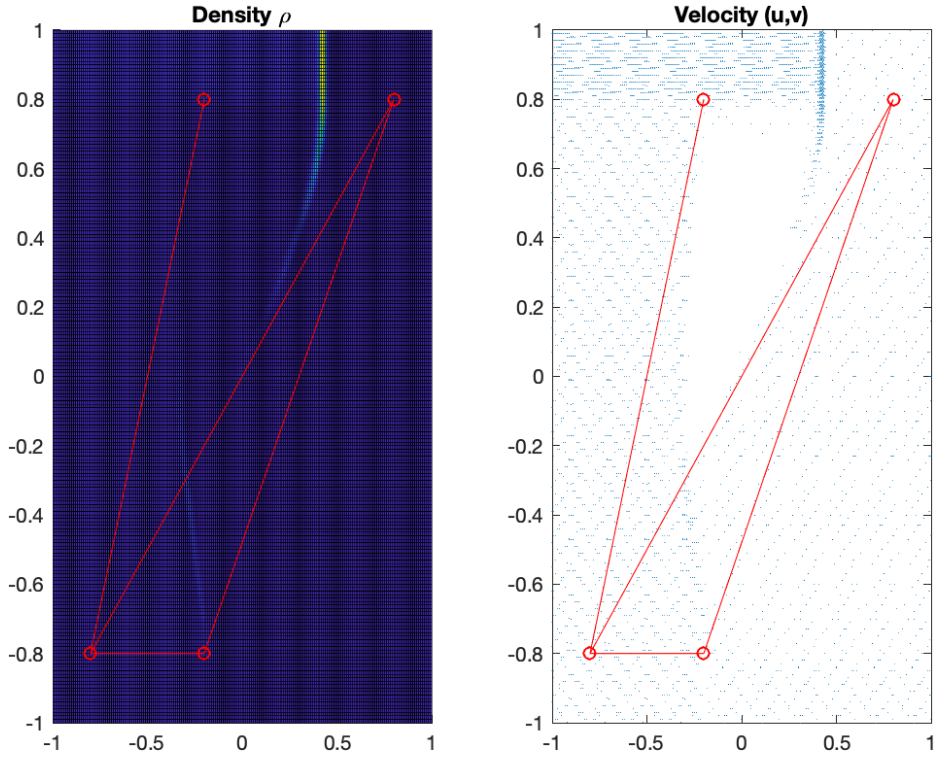
There is no additional condition needed for Case 22. We took

$$\rho_1 = 1, \rho_2 = 10, \rho_3 = 3, \rho_4 = 2, \tag{2.2.135}$$

and

$$u_1 = -4, u_2 = 4, u_3 = -1, u_4 = -1, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = -4. \tag{2.2.136}$$

Case 22. The two contact discontinuities J_{34} and J_{41} meet at the point Ξ_4 and develop a new delta shock $\delta_{13}^{\Xi_4}$. The contact discontinuity J_{23} and the delta shock δ_{12} meet and form a new delta shock. This delta shock meets with the other newly developed delta shock $\delta_{13}^{\Xi_4}$ and form a new delta shock that ends at the point Ξ_3.



(a) Case 22. $N \times N = 200 \times 200$, and $T = 0.2$.

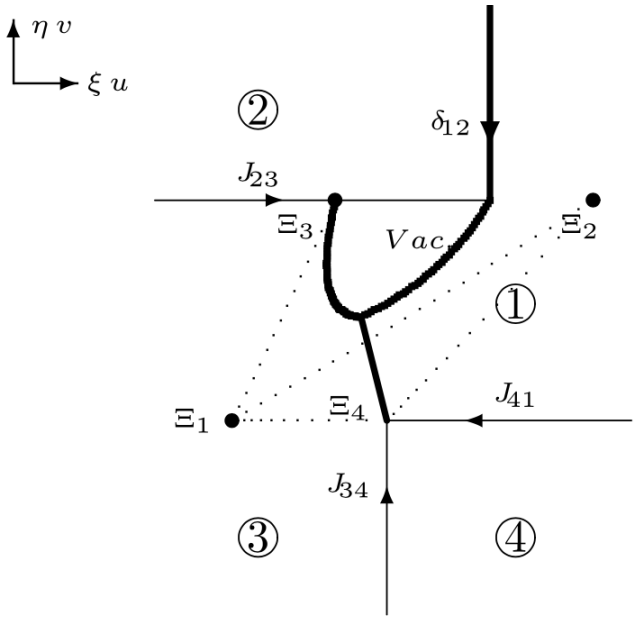


Fig. 4.22. The solution for Case 4.5(ii)b.

(b) Case 22. Screenshot from [Pan19]

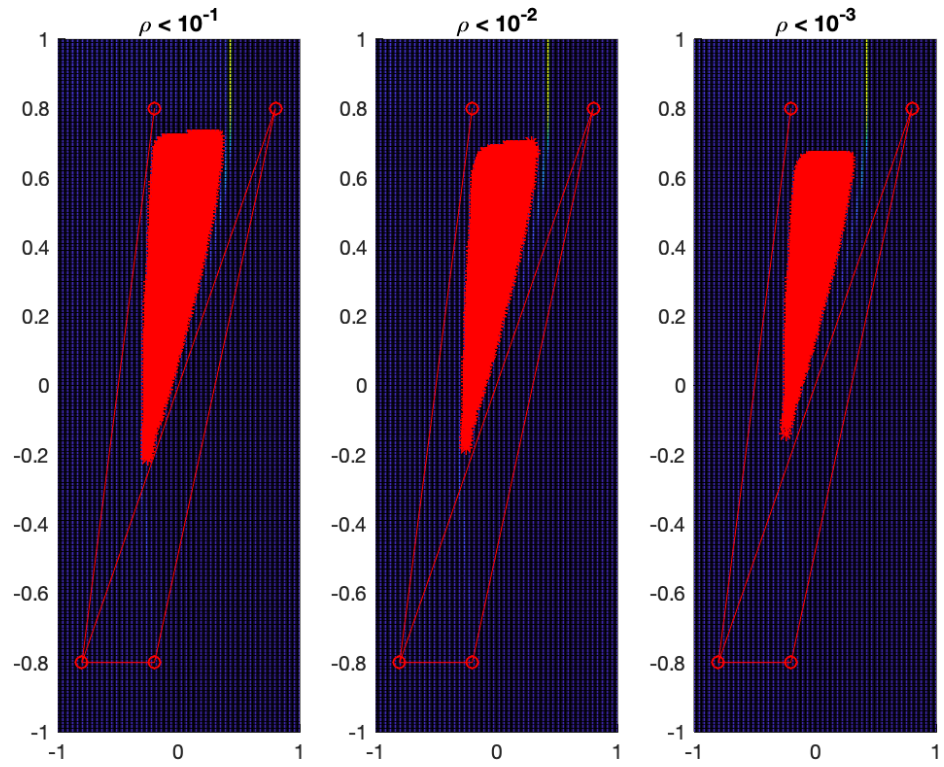


Figure 2.18: Case 22. The red areas denote points where the density ρ is less than 10^{-1}, 10^{-2} and 10^{-3}, respectively.

In Case 22. there are $201 \times 201 = 40,401$ points in total. The minimum density value of this simulation is

$$\rho_{min} = 1.8713 \times 10^{-9}. \tag{2.2.137}$$

More details about the vacuum area are listed in the following table.

Table 2.21: Case 22. Number of points with different density conditions

Density Condition	Total	$\rho < 10^{-1}$	$\rho < 10^{-2}$	$\rho < 10^{-3}$	$\rho < 10^{-4}$
Number of Points	40,401	2,896	2,303	1,928	1,587

Case 23.

[Pan19] gave the initial conditions that

$$u_1 < u_2 < u_3 = u_4, v_4 = v_1 < v_2 = v_3. \tag{2.2.138}$$

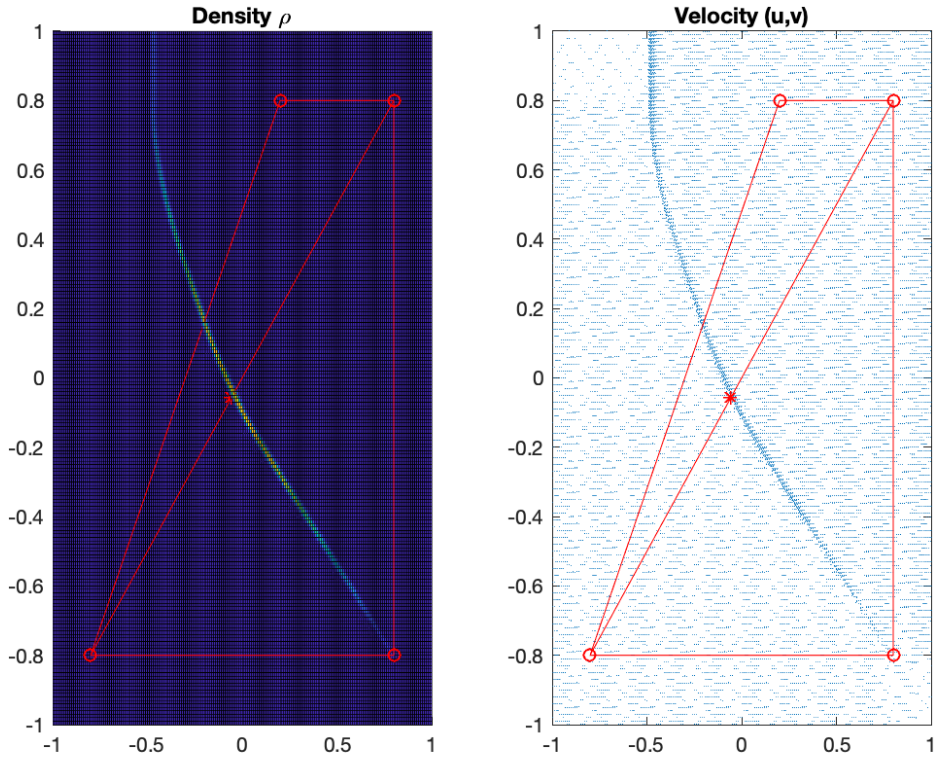
We took the following set of initial conditions

$$\rho_1 = 4, \rho_2 = 1, \rho_3 = 3, \rho_4 = 2, \tag{2.2.139}$$

and

$$u_1 = -4, u_2 = 1, u_3 = 4, u_4 = 4, v_1 = -4, v_2 = 4, v_3 = 4, v_4 = -4. \tag{2.2.140}$$

Case 23. The two contact discontinuities J_{34} and J_{41} meet at the point Ξ_4 and develop a new delta shock $\delta_{13}^{\Xi_4}$. The contact discontinuity J_{23} and the delta shock δ_{12} meet and form a new delta shock. The two newly developed delta shocks meet at the point U_{δ}^{13}.



(a) Case 23. $N \times N = 200 \times 200$, and $T = 0.2$.

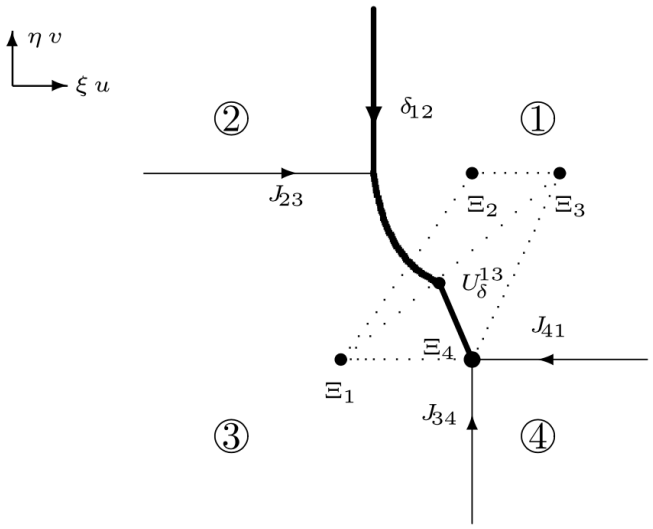


Fig. 4.23. The solution for Case 4.5(ii)c.

(b) Case 23. Screenshot from [Pan19]

Bibliography

[AH94] RK Agarwal and DW Halt. A modified cusp scheme in wave/particle split form for unstructured grid euler flows. *Frontiers of Computational Fluid Dynamics*, 1995, 1994.

[BBS02] MC Bento, Orfeu Bertolami, and Anjan A Sen. Generalized chaplygin gas, accelerated expansion, and dark-energy-matter unification. *Physical Review D*, 66(4):043507, 2002.

[BBT06] Christophe Berthon, Michael Breuß, and Marc-Olivier Titeux. A relaxation scheme for the approximation of the pressureless Euler equations. *Numer. Methods Partial Differential Equations*, 22(2):484–505, 2006.

[BG98] Yann Brenier and Emmanuel Grenier. Sticky particles and scalar conservation laws. *SIAM J. Numer. Anal.*, 35(6):2317–2328, 1998.

[BJ99] François Bouchut and François James. Duality solutions for pressureless gases, monotone scalar conservation laws, and uniqueness. *Comm. Partial Differential Equations*, 24(11-12):2173–2189, 1999.

[BJL03] François Bouchut, Shi Jin, and Xiantao Li. Numerical approximations of pressureless and isothermal gas dynamics. *SIAM J. Numer. Anal.*, 41(1):135–158, 2003.

[BM12] Laurent Boudin and Julien Mathiaud. A numerical scheme for the one-dimensional pressureless gases system. *Numer. Methods Partial Differential Equations*, 28(6):1729–1746, 2012.

[Bou94] F. Bouchut. On zero pressure gas dynamics. In *Advances in kinetic theory and computing*, volume 22 of *Ser. Adv. Math. Appl. Sci.*, pages 171–190. World Sci. Publ., River Edge, NJ, 1994.

[BTV02] Neven Bilic, Gary B Tupper, and Raoul D Viollier. Dark matter, dark energy and the chaplygin gas. *arXiv preprint astro-ph/0207423*, 2002.

[Cha02] SA Chaplygin. On gas jets [o gazovyh struyah], 1902.

[CKR07] Alina Chertock, Alexander Kurganov, and Yurii Rykov. A new sticky particle method for pressureless gas dynamics. *SIAM J. Numer. Anal.*, 45(6):2408—2441, 2007.

[CL03] Gui-Qiang Chen and Hailiang Liu. Formation of δ-shocks and vacuum states in the vanishing pressure limit of solutions to the Euler equations for isentropic fluids. *SIAM J. Math. Anal.*, 34(4):925–938, 2003.

[CM14] Demetrios Christodoulou and Shuang Miao. *Compressible flow and Euler's equations*, volume 9 of *Surveys of Modern Mathematics*. International Press, Somerville, MA; Higher Education Press, Beijing, 2014.

[CSA94] Jay Casper, Chi-Wang Shu, and H Atkins. Comparison of two formulations for high-order accurate essentially nonoscillatory schemes. *AIAA journal*, 32(10):1970–1977, 1994.

[DLY18] Haoyu Dong, Changna Lu, and Hongwei Yang. The finite volume weno with lax–wendroff scheme for nonlinear system of euler equations. *Mathematics*, 6(10):211, 2018.

[ERS96] Weinan E, Yu. G. Rykov, and Ya. G. Sinai. Generalized variational principles, global weak solutions and behavior with random initial data for systems of conservation laws arising in adhesion particle dynamics. *Comm. Math. Phys.*, 177(2):349–380, 1996.

[Eul57] Leonhard Euler. Principes généraux du mouvement des fluides. *Mémoires de l'académie des sciences de Berlin*, pages 274–315, 1757.

[FA18] VMC Ferreira and PP Avelino. Extended family of generalized chaplygin gas models. *Physical Review D*, 98(4):043515, 2018.

[GKMP05] Vittorio Gorini, Alexander Kamenshchik, Ugo Moschella, and Vincent Pasquier. The chaplygin gas as a model for dark energy. In *The Tenth Marcel Grossmann Meeting: On Recent Developments in Theoretical and Experimental General Relativity, Gravitation and Relativistic Field Theories (In 3 Volumes)*, pages 840–859. World Scientific, 2005.

[Hu00] Jiaxin Hu. Two-dimensional Riemann problem for pressureless gas dynamics equations with functional solutions. *Quart. Appl. Math.*, 58(2):251–264, 2000.

[HW07] Jan S Hesthaven and Tim Warburton. *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*. Springer Science & Business Media, 2007.

[Kri17] A. J. Kriel. Error analysis of flux limiter schemes at extrema. *J. Comput. Phys.*, 328:371–386, 2017.

[Lax54] Peter D Lax. Weak solutions of nonlinear hyperbolic equations and their numerical computation. *Communications on pure and applied mathematics*, 7(1):159–193, 1954.

[Lev04] Randall J. Leveque. The dynamics of pressureless dust clouds and delta waves. *J. Hyperbolic Differ. Equ.*, 1(2):315–327, 2004.

[LOC94] Xu-Dong Liu, Stanley Osher, and Tony Chan. Weighted essentially non-oscillatory schemes. *J. Comput. Phys.*, 115(1):200–212, 1994.

[LPR99] Doron Levy, Gabriella Puppo, and Giovanni Russo. Central weno schemes for hyperbolic systems of conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, 33(3):547–571, 1999.

[LPR02] Doron Levy, Gabriella Puppo, and Giovanni Russo. A fourth-order central WENO scheme for multidimensional hyperbolic systems of conservation laws. *SIAM J. Sci. Comput.*, 24(2):480–506, 2002.

[LXX13] Jun Luo, Lijun Xuan, and Kun Xu. Comparison of fifth-order weno scheme and finite volume weno-gas-kinetic scheme for inviscid and viscous flow simulation. *Communications in Computational Physics*, 14(3):599–620, 2013.

[MSON+21] GA Monerat, CGM Santos, G Oliveira-Neto, EV Corrêa Silva, and LG Ferreira Filho. The dynamics of the early universe in a model with radiation and a generalized chaplygin gas-early universe with radiation and a

generalized chaplygin gas. *The European Physical Journal Plus*, 136(1):34, 2021.

[New11] Isaac Newton. *De analysi per aequationes numero terminorum infinitas*. 1711.

[Pan17] Yicheng Pang. Delta shock wave in the compressible Euler equations for a Chaplygin gas. *J. Math. Anal. Appl.*, 448(1):245–261, 2017.

[Pan19] Yicheng Pang. The Riemann problem for the two-dimensional zero-pressure Euler equations. *J. Math. Anal. Appl.*, 472(2):2034–2074, 2019.

[PHW18] Yicheng Pang, Min Hu, and Jinhuan Wang. Riemann problem for a compressible perfect fluid with a constant external force for the Chaplygin gas. *Bound. Value Probl.*, pages Paper No. 89, 17, 2018.

[Roe81] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43(2):357–372, 1981.

[Set07] MR Setare. Holographic chaplygin gas model. *Physics Letters B*, 648(5-6):329–332, 2007.

[Shu90] Chi-Wang Shu. Numerical experiments on the accuracy of eno and modified eno schemes. *Journal of Scientific Computing*, 5(2):127–149, 1990.

[Shu98] Chi-Wang Shu. Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws. In *Advanced numerical approximation of nonlinear hyperbolic equations (Cetraro, 1997)*, volume 1697 of *Lecture Notes in Math.*, pages 325–432. Springer, Berlin, 1998.

[Shu99] Chi-Wang Shu. High order eno and weno schemes for computational fluid dynamics. In *High-order methods for computational physics*, pages 439–582. Springer, 1999.

[SO88] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially nonoscillatory shock-capturing schemes. *J. Comput. Phys.*, 77(2):439–471, 1988.

[SO89] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes, ii. In *Upwind and High-Resolution Schemes*, pages 328–374. Springer, 1989.

[Sod78] Gary A Sod. A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws. *Journal of computational physics*, 27(1):1–31, 1978.

[SZ89] Sergei F Shandarin and Ya B Zeldovich. The large-scale structure of the universe: Turbulence, intermittency, structures in a self-gravitating medium. *Reviews of Modern Physics*, 61(2):185, 1989.

[SZE+92] Chi-Wang Shu, Thomas A Zang, Gordon Erlebacher, David Whitaker, and Stanley Osher. High-order eno schemes applied to two-and three-dimensional compressible flow. *Applied Numerical Mathematics*, 9(1):45–71, 1992.

[Tsi39] Hsue-Shen Tsien. Two-dimensional subsonic flow of compressible fluids. *J. Aeronaut. Sci.*, 6:399–407, 1939.

[Van91] Bram Vanleer. Flux-vector splitting for the 1990s. In *NASA, Lewis Research Center, Computational Fluid Dynamics Symposium on Aeropropulsion*, 1991.

[vK48] Theodore von Kármán. Effect of compressibility in aerodynamics. *Ciencia y Técnica*, 110:251–278, 305–320, 1948.

[WC84] Paul Woodward and Phillip Colella. The numerical simulation of two-dimensional fluid flow with strong shocks. *Journal of computational physics*, 54(1):115–173, 1984.

[YWS13] Yang Yang, Dongming Wei, and Chi-Wang Shu. Discontinuous Galerkin method for Krause's consensus models and pressureless Euler equations. *J. Comput. Phys.*, 252:109–127, 2013.

[ZS10] Xiangxiong Zhang and Chi-Wang Shu. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *J. Comput. Phys.*, 229(23):8918–8934, 2010.

[ZS11a] Xiangxiong Zhang and Chi-Wang Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 467(2134):2752–2776, 2011.

[ZS11b] Xiangxiong Zhang and Chi-Wang Shu. Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms. *J. Comput. Phys.*, 230(4):1238–1248, 2011.

[ZS12] Xiangxiong Zhang and Chi-Wang Shu. Positivity-preserving high order finite difference WENO schemes for compressible Euler equations. *J. Comput. Phys.*, 231(5):2245–2258, 2012.