High order numerical methods for hyperbolic equations: superconvergence, and applications to δ-singularities and cosmology

by

Yang Yang

B.S., University of Science and Technology of China, Hefei, China 2009

M.S., Brown university, Providence, RI 2011

A Dissertation Submitted in Partial Fulfillment of the

Requirements for the Degree of Doctor of Philosophy

in the Division of Applied Mathematics at Brown University

Providence, Rhode Island

May 2013

This dissertation by Yang Yang is accepted in its present form
by the Division of Applied Mathematics as satisfying the
dissertation requirement for the degree of Doctor of Philosophy.

Date________________ ____
Chi-Wang Shu, Ph.D., Advisor

Recommended to the Graduate Council

Date________________ ____
Jan Hesthaven, Ph.D., Reader

Date________________ ____
Johnny Guzmán, Ph.D., Reader

Approved by the Graduate Council

Date________________ ____
Peter Weber, Dean of the Graduate School

The Vita of Yang Yang

Yang Yang was born in Tianjin, China, on 03 March 1987, the son of Min Liu and Zhiyong Yang. After completing his work at Tianjin Nankai High School, he went on to the University of Science and Technology of China where he studied mathematics and received his Bachelor of Science in July 2009. After that, he entered the Division of Applied Mathematics at Brown University.

Publications:

- Y. Yang, I. Roy, C.-W. Shu and L.-Z. Fang, *Effect of dust on Lyα photon transfer in optically thick halo* , The Astrophysical Journal, 739 (2011), 91(11).
- Y. Yang and C.-W. Shu, *Analysis of optimal superconvergence of discontinuous Galerkin method for linear hyperbolic equations*, SIAM Journal on Numerical Analysis, 50 (2012), 3110-3133.
- Y. Yang and C.-W. Shu, *Discontinuous Galerkin method for hyperbolic equations involving δ-singularities: negative-order norm error estimates and applications*, Numerische Mathematik, to appear.
- Y. Yang, I. Roy, C.-W. Shu and L.-Z. Fang, *Angular distribution of Lyα resonant photons emergent from optically thick medium*, submitted.
- Y. Yang, D. Wei and C.-W. Shu, *Discontinuous Galerkin method for Krause's consensus models and pressureless Euler equations*, submitted.

Address: (Updated 22 March 2013)

182 George Street
Providence, Rhode Island 02912
(401) 523-8384

This thesis was typed by the author.

Acknowledgments

First of all, I would like to thank my thesis advisor, Professor Chi-Wang Shu, who assigned such an interesting topic to me. Professor Shu is really patient, supportive, available and productive. I can discuss with him almost anytime I wish. Moreover, he also introduced other areas in applied mathematics and computation techniques to me, such as inverse problem and parallel programming, which benefited me during my on-site interview.

Second, I want to thank Professor Li-Zhi Fang, who was a professor in physics department at the University of Arizona. Professor Fang gave me plenty of guidance on cosmology. However, very sad news came to me in April 6th 2012, that Professor Fang passed away. I was so astonished, since we finished our last project only one week earlier and started a new one. His enthusiasm in science and research is really impressive to me. Moreover, I also finished this project with Doctor Ishani Roy. With the help of Ishani, I got familiar with the project quickly.

Next, I would like to thank my family, who supported me a lot. I understand that they missed me very much during my stay in US, even though they never said so. Moreover, I appreciate my girl friend, who never showed up these years. Therefore, I had nothing to do but concentrated myself on the projects everyday, including holidays, such as Christmas and Chinese new years.

I also want to thank my group members, since the whole group formed a big family. Thank Xiong Meng and Doctor Qiang Zhang, who gave me suggestions on the project of superconvergence. Thank Doctor Xiangxiong Zhang, who shared his experience about the positivity-preserving limiter with me. Thank all the other members, who introduced other research areas in the group seminars.

Finally, I would like to thank all of the people who have helped me to complete my thesis.

Contents

List of Tables

List of Figures

Part I

Discontinuous Galerkin methods for hyperbolic equations

Chapter 1

Introduction

We consider discontinuous Galerkin (DG) methods for solving hyperbolic equations

$$\begin{aligned} &u_t + f(u)_x = g(x,t), && (x,t) \in R \times (0,T], \\ &u(x,0) = u_0(x), && x \in R. \end{aligned} \tag{1.1}$$

The DG method is a class of finite element methods using completely discontinuous piecewise polynomial space to represent as the numerical solutions and as the test functions. The method, first introduced in 1973 by Reed and Hill [86], was generalized by Johnson and Pitkäranta to solve scalar linear hyperbolic equations with L^p-norm error estimates [64]. Subsequently, Cockburn et al. studied Runge-Kutta discontinuous Galerkin (RKDG) methods for hyperbolic conservation laws in a series of papers [35, 32, 33, 36]. As mentioned in [26], for linear hyperbolic equations, by using piecewise polynomial of degree k, the DG approximation is $(k+3/2)$-th order superconvergent towards a particular projection of the exact solution. However, numerical experiments demonstrate that a rate of convergence of $k+2$. We will use a dual argument to prove this property.

One application of the DG methods is to solve hyperbolic equations involving δ-functions. It is well known that generic solutions of hyperbolic equations are not smooth. Discontinuities or even δ-singularities may appear in the solutions. The

DG methods have been shown to be L^2-stable for nonlinear hyperbolic equations with L^2-solutions which may contain discontinuities [60, 55]. In our work, we assume that the initial condition u_0, or the source term $g(x,t)$, or the solution $u(x,t)$ to (1.1) contains δ-singularities. Such problems appear often in applications, such as pressureless Euler equatons, and are difficult to approximate numerically. Many numerical techniques rely on modifications with smooth kernels which may smear such singularities, leading to large errors in the approximation. On the other hand, the DG methods are based on weak formulations and can be designed directly solve such problems without modifications, leading to very accurate results. We will provide numerical examples to demonstrate this advantage. Moreover, we will give rigorous error estimates for the DG methods on some linear problems involving δ-singularities. As demonstrate that the DG approximations are high order accurate under suitable negative-order norms. This makes possible extraction of superconvergence solution by convolving the numerical solutions with suitable kernels. For nonlinear models equations, we will apply boundary-preserving (BP) limiters and prove the L^1-stability of the schemes.

Chapter 2

Preliminaries

In this chapter, we consider the hyperbolic equation (1.1) on the interval $[0, 2\pi]$.

2.1 The DG scheme

First, we divide the computational domain $\Omega = [0, 2\pi]$ into N cells

$$0 = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N+\frac{1}{2}} = 2\pi,$$

and denote

$$I_j = \left(x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}\right)$$

as the cells. $h_j = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$ denotes the length of each cell. We also define $h = h_{\max} = \max_j h_j$ and $h_{\min} = \min_j h_j$ as the lengths of the largest and smallest cells, respectively. We consider regular meshes, that is $h_{\max} \leq \Lambda h_{\min}$ where $\Lambda \geq 1$ is a constant during mesh refinement. Clearly, if $\Lambda = 1$, the mesh is uniform.

Next, we define

$$V_h = \{v : v|_{I_j} \in \mathcal{P}^k(I_j), j = 1, \cdots, N\}$$

as the finite element space, where $\mathcal{P}^k(I_j)$ denotes the space of polynomials in I_j of degree at most k. We also define

$$H_h^1 = \{\phi : \phi|_{I_j} \in H^1(I_j),\ \forall j\}.$$

The DG scheme we consider is the following: find $u_h \in V_h$, such that for any $v_h \in V_h$

$$((u_h)_t, v_h)_j = (f(u_h), (v_h)_x)_j - \hat{f}_{j+\frac{1}{2}} v_h^-|_{j+\frac{1}{2}} + \hat{f}_{j-\frac{1}{2}} v_h^+|_{j-\frac{1}{2}} + (g(x,t), v_h)_j, \qquad (2.1)$$

where $(w, v)_j = \int_{I_j} wvdx$, and $v_h^-|_{j+\frac{1}{2}} = v_h(x_{j+\frac{1}{2}}^-)$ denotes the left limit of the function v_h at $x_{j+\frac{1}{2}}$. Likewise for v_h^+. Moreover, the numerical flux $\hat{f}$ is a single valued function defined at the cell interfaces and in general depends on the values of the numerical solution u_h from both sides of the interfaces

$$\hat{f}_{j+\frac{1}{2}} = \hat{f}(u_h(x_{j+\frac{1}{2}}^-), u_h(x_{j+\frac{1}{2}}^+)).$$

In general, we use monotone fluxes.

For the linear case $f(u) = u$, we consider the upwind fluxes $\hat{f} = u_h^-$. Then the numerical scheme (2.1) can be written as

$$\begin{aligned} ((u_h)_t, v_h)_j &= (u_h, (v_h)_x)_j - u_h^- v_h^-|_{j+\frac{1}{2}} + u_h^- v_h^+|_{j-\frac{1}{2}} + (g(x,t), v_h)_j \qquad (2.2)\\ &= -((u_h)_x, v_h)_j - [u_h] v_h^+|_{j-\frac{1}{2}} + (g(x,t), v_h)_j, \qquad (2.3) \end{aligned}$$

where $[u_h]_{j-\frac{1}{2}} = u_h(x_{j-\frac{1}{2}}^+) - u_h(x_{j-\frac{1}{2}}^-)$ is the jump of u_h across $x_{j-\frac{1}{2}}$. We use (2.2) and (2.3) in Chapters 3 and 4 for the error estimates. Define

$$\mathcal{H}_j(u_h, v_h) = (u_h, (v_h)_x)_j - u_h^- v_h^-|_{j+\frac{1}{2}} + u_h^- v_h^+|_{j-\frac{1}{2}}, \qquad (2.4)$$

such that the DG scheme can be written as

$$((u_h)_t, v_h)_j = \mathcal{H}_j(u_h, v_h) + (g(x,t), v_h)_j.$$

If we do not consider the source term (i.e. $g(x,t) = 0$), the scheme becomes

$$((u_h)_t, v_h)_j = \mathcal{H}_j(u_h, v_h). \tag{2.5}$$

2.2 Norms

We now define some norms that we use throughout the thesis.

Denote $\|u\|_{0,I_j}$ as the standard L^2-norm of u on cell I_j. For any non-negative natural number ℓ, we also define the norm and seminorm of the Sobolev space $H^\ell(I_j)$ as

$$\|u\|_{\ell,I_j} = \left\{ \sum_{0\le\alpha\le\ell} \left\| \frac{d^\alpha u}{dx^\alpha} \right\|_{0,I_j}^2 \right\}^{1/2}, \quad |u|_{\ell,I_j} = \left\| \frac{d^\ell u}{dx^\ell} \right\|_{0,I_j}.$$

For convenience, if $\ell = 0$, the corresponding index will be omitted.

We also define the L^∞-norm and seminorm by

$$\|u\|_{\ell,\infty,I_j} = \max_{0\le\alpha\le\ell} \left\| \frac{d^\alpha u}{dx^\alpha} \right\|_{\infty,I_j}, \quad |u|_{\ell,\infty,I_j} = \left\| \frac{d^\ell u}{dx^\ell} \right\|_{\infty,I_j},$$

where $\|u\|_{\infty,I_j}$ is the standard L^∞-norm of u on cell I_j. Clearly, the L^∞-norm is stronger than the L^2-norm, and in one cell I_j, we have

$$\|u\|_{I_j} \le h_j^{1/2} \|u\|_{\infty,I_j}. \tag{2.6}$$

Moreover, we define the norms on $D = \cup_{j\in\Gamma} I_j$ for some index set Γ as follows:

$$\|u\|_{\ell,D} = \left(\sum_{j\in\Gamma} \|u\|_{\ell,I_j}^2 \right)^{1/2}, \quad \|u\|_{\ell,\infty,D} = \max_{j\in\Gamma} \|u\|_{\ell,\infty,I_j}.$$

For convenience, if $D = \Omega = [0, 2\pi]$, the corresponding index will be omitted.

Finally, the negative order Sobolev norm can be defined as

$$\|u\|_{-\ell,D} = \sup_{\phi \in C_0^\infty(D)} \frac{\int_D u(x)\phi(x)dx}{\|\phi\|_{\ell,D}}.$$

2.3 Properties of the finite element space

In this section, we study the basic properties of the finite element space. Let us start with the classical inverse properties.

Lemma 2.3.1. *Assume $u \in V_h$, then there exists a constant $C > 0$ independent of h and u such that*

$$\left\| \frac{d^\alpha u}{dx^\alpha} \right\|_D \le Ch^{-\alpha}\|u\|_D, \qquad \alpha \ge 1, \tag{2.7}$$

$$\sum_{I_j \in D} \left(\left| u^-_{j+\frac{1}{2}} \right| + \left| u^+_{j-\frac{1}{2}} \right| \right) \le Ch^{-1/2}\|u\|_D, \tag{2.8}$$

where D can be the single cell I_j or the whole computational domain Ω.

Proof: The proof is trivial. We just use the fact that the norms in finite dimensional spaces are equivalent. So we skip it here.

We define $\mathbb{P}_\ell(p)$ as the ℓ-th order L^2-projection of p into V_h, such that

$$(\mathbb{P}_\ell(p), u)_j = (p, u)_j \ , \quad \forall u \in \mathcal{P}^\ell(I_j). \tag{2.9}$$

In addition, if $\ell \ge 1$, we can also define two Gauss-Radau projections $\mathbb{P}_+$ and $\mathbb{P}_-$ as:

$$(\mathbb{P}_+(p), u)_j = (p, u)_j \ , \quad \forall u \in \mathcal{P}^{\ell-1}(I_j) \ , \quad \text{and} \quad \mathbb{P}_+(p)(x^+_{j-1/2}) = p(x^+_{j-1/2}), \tag{2.10}$$

$$(\mathbb{P}_-(p), u)_j = (p, u)_j \ , \quad \forall u \in \mathcal{P}^{\ell-1}(I_j) \ , \quad \text{and} \quad \mathbb{P}_-(p)(x^-_{j+1/2}) = p(x^-_{j+1/2}). \tag{2.11}$$

The projections $\mathbb{P}_+$ and $\mathbb{P}_-$ are different from the exact collocation at different end points of each cell.

For the projection $\mathbb{P}_h$, which is either $\mathbb{P}_k$, $\mathbb{P}_+$ or $\mathbb{P}_-$, we denote the error operator by $\mathbb{P}_h^\perp = \mathbb{I} - \mathbb{P}_h$, where $\mathbb{I}$ is the identity operator. By a scaling argument, we obtain the following property [28].

Lemma 2.3.2. *Suppose the function $u(x) \in C^{k+1}(I_j)$, then there exists a positive constant C independent of h and u, such that*

$$\|\mathbb{P}_h^\perp u\|_{I_j} \le C h_j^{k+1} |u|_{k+1,I_j} \quad \textit{and} \quad \|\mathbb{P}_h^\perp u\|_{\infty,I_j} \le C h_j^{k+1} |u|_{\infty,k+1,I_j}. \tag{2.12}$$

Moreover, one can also prove the following superconvergence property [5].

Lemma 2.3.3. *Suppose $u(x) \in C^{k+2}(I_j)$, and x_j is one of the downwind-biased Radau points in the cell I_j, then*

$$|(u - \mathbb{P}_- u)(x_j)| \le C h_j^{k+2} |u|_{k+2,\infty,I_j}. \tag{2.13}$$

However, if u is highly oscillatory or discontinuous, we cannot obtain any useful estimate of $\|\mathbb{P}_h^\perp u\|$ by the two lemmas above. Therefore, we consider the following estimate.

Lemma 2.3.4. *Suppose $u(x)$ is a bounded function, then*

$$\|\mathbb{P}_h u\|_{\infty,I_j} \le C \|u\|_{\infty,I_j}, \quad \textit{and} \quad \|\mathbb{P}_h^\perp u\|_{\infty,I_j} \le C \|u\|_{\infty,I_j}. \tag{2.14}$$

Proof: For the simplicity of the presentation, we will only prove it for the $\mathbb{P}_-$ projection. We consider the projection on the reference cell $T = [-1, 1]$ and define a special norm in $\mathcal{P}^k(T)$ as

$$|||v||| = \max \left\{ |v(1)|, \left| \int_{-1}^{1} v(s) s^p ds \right| : 0 \le p \le k-1 \right\}.$$

It is not difficult to show this is indeed a norm. Since all norms in $\mathcal{P}^k$ are equivalent, we have $\|v\|_{\infty,T} \leq C|||v|||$ for any $v \in \mathcal{P}^k(T)$. Therefore, for any bounded function u,

$$\|\mathbb{P}_- u\|_{\infty,T} \leq C|||\mathbb{P}_- u||| = C|||u||| \leq C\|u\|_{\infty,T}.$$

This proves the assertion on the reference cell. The general case follows from a standard scaling argument.

By using (2.6) and Lemma 2.3.4, we obtain

$$\|\mathbb{P}_h^{\perp} u\|_{I_j} \leq h_j^{1/2} \|\mathbb{P}_h^{\perp} u\|_{\infty,I_j} \leq C h_j^{1/2} \|u\|_{\infty,I_j}.$$

Now, we consider the projection of functions depending not only on the spatial variable x but also on the time variable t. Suppose $u(x,t)$ is a function differentiable and integrable with respect to t and assume t_1 and t_2 are two real values such that $t_1 < t_2$. Then we have

$$\mathbb{P}_h\left(u_t(x,t)\right) = \left(\mathbb{P}_h u(x,t)\right)_t, \quad \text{and} \quad \mathbb{P}_h\left(\int_{t_1}^{t_2} u(x,t)dt\right) = \int_{t_1}^{t_2} (\mathbb{P}_h u(x,t))dt. \quad (2.15)$$

As a result, we do not need to distinguish $\mathbb{P}_h(u_t(x,t))$ and $(\mathbb{P}_h u(x,t))_t$, and can simply denote them as $\mathbb{P}_h u_t$.

2.4 Properties of the DG spatial discretization

In this subsection, we present some basic properties about the bilinear form $\mathcal{H}_j$ and the L^2-stability condition [30]. We consider the linear case, namely (1.1) with $f(u) = u$. The following lemma is given by Cockburn [30].

Lemma 2.4.1. *Suppose u_h is the DG numerical solution which satisfies* (2.5) *in each cell. By using the upwind flux, we have*

$$\|u_h(T)\|^2 + \int_0^T \sum_{1\leq j\leq N} [u_h(t)]_{j+1/2}^2 \, dt \leq \|u_h(0)\|^2. \quad (2.16)$$

Lemma 2.4.2. *Suppose $v_h \in V_h$ and $q(x) \in H_h^1$, then the two Gauss-Radau projections satisfy*

$$\mathcal{H}_j(\mathbb{P}_-^\perp q(x), v_h) = 0, \quad \textit{and} \quad \mathcal{H}_j(v_h, \mathbb{P}_+^\perp q(x)) = 0. \tag{2.17}$$

Proof: The proof is straight forward. So we skip it here.
If we define $(u_h, v_h) = \sum_j (u_h, v_h)_j$ and $\mathcal{H}(p,q) = \sum_j \mathcal{H}_j(p,q)$, then

Lemma 2.4.3. *Suppose $p(x) \in H_h^1$ and $v_h \in V_h$, there holds*

$$\mathcal{H}(\mathbb{P}_-^\perp p(x), v_h) = 0, \quad \textit{and} \quad \mathcal{H}(v_h, \mathbb{P}_+^\perp p(x)) = 0. \tag{2.18}$$

Proof: The proof follows from Lemma 2.4.2 and the definition of $\mathcal{H}$. So we skip it.

2.5 The error equation

In this subsection, we also consider linear equation (i.e. $f(u) = u$ in (1.1)) and proceed to construct the error equations. From (2.2) and definition (2.4), we have for any $v_h \in V_h$

$$((u_h)_t, v_h)_j = \mathcal{H}_j(u_h, v_h) + (g(x,t), v_h)_j.$$

Clearly, the exact solution u satisfies a similar equation

$$(u_t, v_h)_j = \mathcal{H}_j(u, v_h) + (g(x,t), v_h)_j.$$

Denote the error between the exact solution and the DG numerical solution to be $e(t) = u(t) - u_h(t)$. Then we have

$$(e_t, v_h)_j = \mathcal{H}_j(e, v_h), \text{ for any } v_h \in V_h.$$

Following the usual treatment in finite element analysis, we divide the error into the form $e(t) = \eta(t) - \xi(t)$, where

$$\eta(t) = u(t) - \mathbb{P}_{-}u(t), \quad \text{and} \quad \xi(t) = u_h(t) - \mathbb{P}_{-}u(t).$$

From Lemma 2.4.2, we obtain the error equations of the DG scheme. Suppose $v_h \in V_h$ then

$$\begin{aligned}(e_t, v_h)_j &= -\mathcal{H}_j(\xi, v_h)\\ &= -(\xi, v_{hx})_j + \xi^- v_h^-|_{j+\frac{1}{2}} - \xi^- v_h^+|_{j-\frac{1}{2}} \qquad (2.19)\\ &= (\xi_x, v_h)_j + [\xi] v_h^+|_{j-\frac{1}{2}} \qquad (2.20)\end{aligned}$$

because of upwinding. Equations (2.19) and (2.20) are fundamental in our analysis later.

Let us finish this section by proving the following lemma.

Lemma 2.5.1. *Suppose $\bar{\xi}$ is the cell average of ξ, that is $\bar{\xi} = \bar{\xi}_j = \frac{1}{h_j}\int_{I_j} \xi dx$ in cell I_j, for any $j = 1, \cdots, N$. Then we have*

$$\|\xi - \bar{\xi}\|_{I_j} \le Ch_j \|\xi_x\|_{I_j} \le Ch_j \|\mathbb{P}_k e_t\|_{I_j} \le Ch_j \|e_t\|_{I_j}. \qquad (2.21)$$

Proof: The right inequality is trivial and the left one follows from the Poincaré inequality. So we only need to prove the middle one.

Suppose Q is the Legendre polynomial of degree k in [-1,1] and define $P = (-1)^k Q$. Then P satisfies the following three properties:
(1) P is uniformly bounded: $\|P\|_{\infty,[-1,1]} \le 1$;
(2) P evaluated at the left boundary is 1: $P(-1) = 1$;
(3) P is orthogonal to any polynomials with degree no greater than $k-1$: $\int_{-1}^{1} PRdx = 0$ for any $R(x) \in \mathcal{P}^{k-1}([-1, 1])$.
Define $P_j(x) = P(\frac{2(x-x_j)}{h_j})$, then P_j also satisfies the corresponding three properties

in the cell I_j. In (2.20), we take $v_h = \xi_x - aP_j$, where $a = \xi_x^+|_{j-1/2}$ is a real number, to obtain

$$\begin{aligned}
\|\xi_x\|_{I_j}^2 &= (\mathbb{P}_k e_t, \xi_x - aP_j)_j \\
&\le \|\mathbb{P}_k e_t\|_{I_j} \left(\|\xi_x\|_{I_j} + |a|\|P_j\|_{I_j}\right) \\
&\le \|\mathbb{P}_k e_t\|_{I_j} \left(\|\xi_x\|_{I_j} + Ch_j^{-1/2}\|\xi_x\|_{I_j} h_j^{1/2}\right) \\
&\le C\|\mathbb{P}_k e_t\|_{I_j}\|\xi_x\|_{I_j}, \qquad (2.22)
\end{aligned}$$

where the constant C does not depend on j, h or u. Here, for the second step we use the Cauchy-Schwarz inequality, for the third one we use (2.6) and (2.8), and the last step is trivial. We finish the proof by dividing both sides of the above equation by $\|\xi_x\|_{I_j}$.

Chapter 3

Analysis of optimal superconvergence for linear hyperbolic equations

In this chapter, we study one-dimensional linear hyperbolic conservation laws

$$\begin{aligned} &u_t + u_x = 0, && (x,t) \in R \times (0,T], \\ &u(x,0) = u_0(x), && x \in R, \end{aligned} \tag{3.1}$$

where the initial datum u_0 is sufficiently smooth. We will consider both the periodic boundary condition $u(0,t) = u(2\pi,t)$ and the initial-boundary value problem with $u(0,t) = g(t)$. We use piecewise k-th degree polynomials to approximate the solution in each cell and prove that, under suitable initial discretization, the rate of convergence for the error between the DG solution and the exact solution is of order $(k+2)$-th at the downwind-biased Radau points. Moreover, we also prove order $(k+2)$-th superconvergence of the cell averages as well as the error between the DG solution and a particular type of projection of the exact solution.

In [116], Zhang and Shu gave explicitly formulae for the DG solution in the case of piecewise linear functions for the linear convection equation on uniform meshes.

The leading error term is shown to be of a constant magnitude independent of time t. This motivates the division of the numerical error into two parts, one being the leading term and the other one being a superconvergent term.

In [5, 6], Adjerid et al. proved the $(k+2)$-th order superconvergence of the DG solutions at the downwind-biased Radau points for ordinary differential equations. Later, Adjerid and Weihart [7, 8] investigated the local DG error for multi-dimensional first-order linear symmetric and symmetrizable hyperbolic systems. The authors showed that the projection of the local DG error is also $(k+2)$-th order superconvergent at the downwind-biased Radau points by performing a local error analysis on Cartesian meshes. The global superconvergence is given by numerical experiments. In [7, 8], only initial-boundary value problems are considered, and the local DG error estimate is valid for t sufficiently large. Subsequently, Adjerid and Baccouch [4] investigated the global convergence of the implicit residual-based *a posteriori* error estimates, and proved that these estimates at a fixed time t converge to the true spatial error in the L^2 norm under mesh refinement. In [25], Cheng and Shu proved $(k+\frac{3}{2})$-th order superconvergence of the DG solution towards a particular projection of the exact solution. The authors considered the case of piecewise linear polynomials ($k=1$) on uniform meshes with periodic boundary conditions for the linear conservation laws. Later Cheng and Shu also proved the same $(k+\frac{3}{2})$-th order superconvergence when using piecewise k-th degree polynomials with arbitrary k on arbitrary regular meshes in [26]. In [26] the authors considered both periodic boundary conditions and initial-boundary value problems. However, the convergence rate obtained in [26] is not optimal. Numerical tests showed that the error of the DG solution towards the particular projection of the exact solution is $(k+2)$-th order accurate, even on highly non-uniform meshes, when a special initial discretization is used. Recently, in [123] Zhong and Shu revisited the same problem and showed that the error between the DG numerical solution and the exact solution is $(k+2)$-th order superconvergent at the downwind-biased Radau points and $(2k+1)$-th order superconvergent at the downwind point in each cell on uniform meshes with periodic boundary conditions for $k = 1$, 2 and 3. The proofs in [25, 123] use Fourier analysis and work only for uniform meshes and periodic boundary conditions. Moreover, Fourier analysis is difficult to perform for higher polynomial degree k since it relies explicitly on the structure of the algorithm matrices. In [26], a different framework to prove the superconvergence results that does not rely on Fourier analysis is offered and the results are valid for both periodic boundary conditions and initial-boundary value problems. In this chapter, we improve upon the result in [26]. A new technique is adopted to obtain the optimal rate of superconvergence. The proof works for arbitrary regular meshes and schemes of any order. Even though the proof is given for the simple scalar equation (3.1), the same superconvergent results can be obtained for one-dimensional linear systems using similar points.

3.1 Statement of the main result

Before proceeding to the main theorem, we first introduce a special initial discretization. We wish to require

$$(u_h)_t = \mathbb{P}_-(u_t) \quad \text{and} \quad \|\mathbb{P}_- u - u_h\|_\Omega = \mathcal{O}(h^{k+2}). \tag{3.2}$$

Notice that the special projection $\mathbb{P}_-$ is used in the error estimates of the DG methods to derive optimal L^2-error bounds in the literature, e.g., in [118]. As in [26], we will prove that the numerical solution is closer to this special projection of the exact solution than to the exact solution itself. The exact way to discretize the initial data to achieve the property (3.2) will be given in Section 3.2.1. We can now state our main theorem.

Theorem 3.1.1. *Let $u(x,t) \in C^{k+4}$ be the exact solution of the linear hyperbolic equation* (3.1) *and u_h be the numerical solution of the DG scheme* (2.5). *The finite element space V_h is made up of piecewise polynomials of degree $k \geq 1$ on regular* *meshes, i.e. the ratio of the length of the largest cell to that of the smallest one is bounded during mesh refinement. At time T there holds the following estimate*

$$\left(\frac{1}{N}\sum_{j=1}^{N}|(u-u_h)(x_j)|^2\right)^{\frac{1}{2}} \leq C(1+T^2)h^{k+2}\|u\|_{k+4,\infty,\Omega}, \tag{3.3}$$

where Ω is the computational domain, and x_j is any one of the downwind-biased Radau points in the cell I_j. The constant C does not depend on h, T or u.

Remark 3.1.1. *Theorem 3.1.1 is valid for both periodic boundary condition and initial-boundary value problems.*

Corollary 3.1.1. *Suppose the conditions in the above theorem are satisfied, then we have*

$$\|\overline{u-u_h}\|_{L^2(\Omega)} \leq C(1+T^2)h^{k+2}\|u\|_{k+4,\Omega}, \tag{3.4}$$

$$\|\mathbb{P}_{-}u-u_h\|_{L^2(\Omega)} \leq C(1+T^2)h^{k+2}\|u\|_{k+4,\Omega}, \tag{3.5}$$

where $\overline{u-u_h}$ denotes the cell average of $u-u_h$, and the constant C does not depend on h, T or u.

3.2 Proof of the main result

In this section, we first discuss how to discretize the initial datum, then prove the main result, Theorem 3.1.1, and finally briefly discuss the application of the superconvergence results. The proof can be divided into several steps. Briefly, by using the triangle inequality, we separate $|(u-u_h)(x_j)|$ into two parts, $|(u-\mathbb{P}_{-}u)(x_j)|$ and $|\xi(x_j)|$. The superconvergence of the first term is given by Lemma 2.3.3 while the second one is more difficult to deal with and we separate this process into two steps. In the first step, we consider the estimates of e_t and e_{tt}. In the second step, we use a quadrature formula and consider the dual problem of (3.1). Besides the main theorem, we also prove Corollary 3.1.1 in this section.

3.2.1 The initial discretization

In this subsection we consider the suitable discretization of the initial datum. As mentioned in Section 3.1, we would like to have the initial solution satisfy $\xi_t = 0$ and $\|\xi\|_\Omega \le Ch^{k+2}$, see (3.2). We start from the requirement $\xi_t = 0$ and check whether a special numerical initial solution can be constructed which also satisfies the second requirement $\|\xi\|_\Omega \le Ch^{k+2}$. Let us start from the following lemma. Taking $v_h = 1$ in (2.19), we have

Lemma 3.2.1. $\int_{I_j} e_t dx = 0,\ \forall\ 1 \le j \le N$ *if and only if* $\xi^-_{j+\frac{1}{2}}$ *is a constant which does not depend on* j.

Denote the constant mentioned in the previous lemma as S. Clearly, such a constant gives us freedom to control $\|\xi\|_{I_j}$, as is shown in the following lemma.

Lemma 3.2.2. *Suppose* $\|e_t\|_{I_j} \le Ch_j^{k+3/2}$*, then* $S \le Ch_j^{k+2}$ *if and only if* $\|\xi\|_{I_j} \le Ch_j^{k+5/2}$.

Proof: Suppose $\|\xi\|_{I_j} \le Ch_j^{k+5/2}$, then by Lemma 2.3.1 we have $S \le Ch_j^{k+2}$. On the other hand, suppose $S \le Ch_j^{k+2}$, then by Lemma 2.5.1

$$\begin{aligned} \bar{\xi}_j &= \xi^-_{j+\frac{1}{2}} - (\xi - \bar{\xi}_j)^-_{j+\frac{1}{2}} \\ &\le S + Ch_j^{-1/2}\|\xi - \bar{\xi}_j\|_{I_j} \\ &\le S + Ch_j^{1/2}\|e_t\|_{I_j} \\ &\le Ch_j^{k+2}. \end{aligned}$$

Therefore,

$$\|\xi\|_{I_j} \le \|\bar{\xi}_j\|_{I_j} + \|\xi - \bar{\xi}_j\|_{I_j} \le Ch_j^{k+5/2}.$$

Remark 3.2.1. *The condition* $\|e_t\|_{I_j} \le Ch_j^{k+3/2}$ *in Lemma 3.2.2 is true because we require* $\xi_t = 0$. *Actually, we can show* $\|e_t\|_{I_j} \le Ch_j^{k+1}|u|_{k+2,I_j}$. *We will also use this estimate of* e_t *later.*

There is a straightforward corollary of the above lemma.

Corollary 3.2.1. *Suppose the initial solution satisfies $\xi_t = 0$ and $S \leq Ch^{k+2}$, then $\|\xi\|_\Omega \leq Ch^{k+2}$.*

Now let us proceed to construct the initial solution u_h from $\xi_t = 0$.

Lemma 3.2.3. *Suppose $\int_{I_j} e_t = 0$, then ξ_x is uniquely determined by $\mathbb{P}_k e_t$ in the cell I_j.*

Proof: Let $v_h^+|_{j-\frac{1}{2}} = 0$ in equation (2.20), then we have

$$(\mathbb{P}_k e_t, v_h)_j = (\xi_x, v_h)_j. \tag{3.6}$$

Since the equation is linear, we only need to prove uniqueness. That is, suppose $(\mathbb{P}_k e_t, v_h)_j = 0$, $\forall\ v_h \in V_h$ and $v_h^+|_{j-\frac{1}{2}} = 0$, then $\xi_x = 0$. To prove this, let $p(x)$ be an arbitrary polynomial of degree no more than k and $v_h = p - p^+_{j-\frac{1}{2}}$. Then

$$(\mathbb{P}_k e_t, p)_j = (\mathbb{P}_k e_t, p - p^+_{j-\frac{1}{2}})_j = 0.$$

This implies that $\mathbb{P}_k e_t = 0$. By Lemma 2.5.1, we obtain $\xi_x = 0$.

Remark 3.2.2. *The expression of u_t can be obtained by the partial differential equation. Therefore it is not difficult to obtain $\mathbb{P}_k e_t$ from $\xi_t = 0$.*

Now, the only thing left is to determine the value of the constant $S = \xi^-_{j-\frac{1}{2}}$. By Corollary 3.2.1 we can simply take $S = 0$. However, such S does not satisfy the conservation of mass. If we consider periodic boundary condition we can select a special S such that $\int_\Omega \xi = 0$. We first prove that such as S satisfies the property $S \leq Ch^{k+2}$. Actually,

$$\int_\Omega \xi dx = \sum_{j=1}^{N} \bar{\xi}_j h_j = \sum_{j=1}^{N} \left(S - (\xi - \bar{\xi})^-_{j+\frac{1}{2}} \right) h_j,$$

which yields

$$S|\Omega| = \sum_{j=1}^{N} (\xi - \bar{\xi})^-_{j+\frac{1}{2}} h_j. \tag{3.7}$$

Then we obtain

$$S \le \frac{C}{|\Omega|} \sum_{j=1}^{N} \|e_t\|_{I_j} h_j^{3/2} \le \frac{C}{|\Omega|} \left(\sum_{j=1}^{N} h_j^{2k+5} \right)^{1/2} |u|_{k+2,\Omega} \le \frac{C}{\sqrt{|\Omega|}} h^{k+2} |u|_{k+2,\Omega}. \tag{3.8}$$

In the first inequality in (3.8) we use Lemma 2.5.1 and (2.8). For the second inequality we use the Cauchy-Schwartz inequality and the estimate $\|e_t\|_{I_j} \le C h_j^{k+1} |u|_{k+2,I_j}$ obtained in Remark 3.2.1. The last inequality follows from the fact that $\sum h_j = |\Omega|$ and $h_j \le h$.

Now we summarize how to implement the initial discretization. We divide the process into the following steps:

(1) Let $\xi_t = 0$, then compute the value of e_t using the PDE.

(2) Use Lemma 3.2.3 to find ξ_x.

(3) Compute $\xi - \bar{\xi}$ in each cell from ξ_x and the fact that $\int_{I_j} (\xi - \bar{\xi}) dx = 0$.

(4) Express S by using (3.7) or simply by taking $S = 0$.

(5) Calculate ξ from the expressions of S and ξ_x.

(6) Recover $u_h = \xi + \mathbb{P}_- u$.

From the process mentioned above, we observe that the initial solution is uniquely determined by the requirements $\xi_t = 0$ and $\int_\Omega \xi dx = 0$ or $\xi^-_{j-\frac{1}{2}} = 0$.

3.2.2 Step 1

Now, we proceed to prove Theorem 3.1.1. The estimates of $\|e_t\|_\Omega$ and $\|e_{tt}\|_\Omega$ follow from Lemma 2.3 in [26] with some minor changes. We skip the proofs and state the results in the following two equations:

$$\|e_{tt}(t)\|_{\Omega} \leq Ch^{k+1}|u|_{k+3,\Omega} + Cth^{k+1}|u|_{k+4,\Omega}, \tag{3.9}$$

$$\|e_{t}(t)\|_{\Omega} \leq Ch^{k+1}|u|_{k+2,\Omega} + Cth^{k+1}|u|_{k+3,\Omega}. \tag{3.10}$$

Therefore, by Lemma 2.5.1 we have

$$\|\xi - \bar{\xi}\| \leq C(1+t)h^{k+2}\|u\|_{k+3,\Omega}. \tag{3.11}$$

Before proceeding to the optimal error estimates of $\|\xi\|_{\Omega}$, we use a superconvergence result to prove the optimal error estimate of $\|e\|_{\infty,\Omega}$.

Following [26], we can easily prove

$$\|\xi(t)\|_{\Omega} \leq C(1+t)h^{k+3/2}\|u\|_{k+3,\Omega}.$$

Since ξ is a polynomial of degree at most k in each cell, we have

$$\|\xi(t)\|_{\infty,I_j} \leq Ch^{-1/2}\|\xi(t)\|_{I_j} \leq Ch^{-1/2}\|\xi(t)\|_{\Omega} \leq C(1+t)h^{k+1}\|u\|_{k+3,\Omega}.$$

Notice that the right hand side of the above equation does not depend on j. We can therefore take the maximum on both sides to obtain

$$\|\xi(t)\|_{\infty,\Omega} \leq C(1+t)h^{k+1}\|u\|_{k+3,\Omega}.$$

Finally, by Lemma 2.3.2, we obtain

$$\|e(t)\|_{\infty,\Omega} \leq C(1+t)h^{k+1}\|u\|_{\infty,k+3,\Omega}. \tag{3.12}$$

3.2.3 Step 2

Now, we proceed to estimate $e(x_j)$. By Lemma 2.3.3, only $\xi(x_j)$ is considered. Denote the downwind-biased Radau points of the cell I_j as x_j^i, $0 \le i \le k$. Also denote ψ_j^i to be a polynomial of degree k in cell I_j, such that

$$\psi_j^i(x_\ell) = \begin{cases} 1 & x_\ell = x_j^i \\ 0 & x_\ell \ne x_j^i \end{cases} . \tag{3.13}$$

By the Gauss-Radau quadrature, we have $\xi(x_j^i) = \frac{2}{w_i h_j}(\xi, \psi_j^i)$, where the constant w_i is the weight of the quadrature at the i^{th} downwind-biased Radau point on the reference interval $[-1, 1]$. Therefore, we only need to estimate (ξ, ψ_j^i) for any $0 \le i \le k$. Clearly, $\|\psi_j^i\|_\infty \le C$, where the positive constant C does not depend on i, j or h. Motivated by [34], we consider the dual problem of (3.1). For convenience, we denote by C a generic positive constant that does not depend on h, T or u, but may depend on Λ. Recall that Λ is the ratio of the length of the largest cell to that of the smallest one.

We begin by considering the solution to the dual problem:

(1) For the periodic boundary condition, find a function ϕ_j^i such that $\phi_j^i(\cdot, t)$ satisfies

$$\begin{aligned} &{\phi_j^i}_t + {\phi_j^i}_x = 0, && (x,t) \in R \times (0,T], \\ &\phi_j^i(x,T) = \psi_j^i(x), && x \in R, \\ &\phi_j^i(0,t) = \phi_j^i(2\pi, t), && t \in [0,T]. \end{aligned} \tag{3.14}$$

(2) For the initial boundary value problem, find a function ϕ_j^i such that $\phi_j^i(\cdot, t)$ satisfies

$$\begin{aligned} &{\phi_j^i}_t + {\phi_j^i}_x = 0, && (x,t) \in R \times (0,T], \\ &\phi_j^i(x,T) = \psi_j^i(x), && x \in R, \\ &\phi_j^i(2\pi,t) = 0, && t \in [0,T]. \end{aligned} \tag{3.15}$$

For convenience, we drop the subscript j as well as the superscript i and denote ψ

as ψ_j^i and ϕ as ϕ_j^i. Following [34]

$$\begin{aligned}(e(T),\psi) &= (e,\phi)(0) + \int_0^T (e,\phi)_t dt \\ &= (e,\phi)(0) + \int_0^T [(e_t,\phi) + (e,\phi_t)]dt. \end{aligned} \tag{3.16}$$

We apply $\mathbb{P}_+$ to deal with the term $(e_t,\phi) + (e,\phi_t)$, with the definition of the projection given in (2.10). Recalling that $e = \eta - \xi$ where the notations of ξ and η can be found in Section 2.5, we have

$$\begin{aligned}(e_t,\phi) + (e,\phi_t) &= (e_t,\mathbb{P}_+^\perp\phi) + (e_t,\mathbb{P}_+\phi) - (e,\phi_x) \\ &= (e_t,\mathbb{P}_+^\perp\phi) + \mathcal{H}(e,\mathbb{P}_+\phi) - (\eta,\phi_x) + (\xi,\phi_x) \\ &= (e_t,\mathbb{P}_+^\perp\phi) - \mathcal{H}(\xi,\mathbb{P}_+\phi) - (\eta,\phi_x) + \mathcal{H}(\xi,\phi) \\ &\quad - \sum_{j=2}^N \xi^-[\phi]_{j-\frac{1}{2}} + \xi^-\phi^-|_{N+\frac{1}{2}} - \xi^-\phi^+|_{\frac{1}{2}} \\ &= (e_t,\mathbb{P}_+^\perp\phi) - (\eta,\phi_x) - \sum_{j=2}^N \xi^-[\phi]_{j-\frac{1}{2}} + \xi^-\phi^-|_{N+\frac{1}{2}} - \xi^-\phi^+|_{\frac{1}{2}}, \end{aligned} \tag{3.17}$$

where for the last equality we use Lemma 2.4.3.

For the periodic boundary condition, the above turns out to be

$$(e_t,\phi) + (e,\phi_t) = (e_t,\mathbb{P}_+^\perp\phi) - (\eta,\phi_x) - \sum_{j=1}^N \xi^-[\phi]_{j-\frac{1}{2}}. \tag{3.18}$$

Integrating in t and noticing the fact that

$$\int_0^T \sum_{j=1}^N \xi^-[\phi]_{j-\frac{1}{2}} = 0,$$

since $[\phi(t)]_{i-\frac{1}{2}} = 0$ except for at most finitely many t, we have

$$\int_0^T [(e_t, \phi) + (e, \phi_t)]dt = \int_0^T (e_t, \mathbb{P}_+^\perp \phi)dt + \int_0^T (\eta, \phi_t)dt. \tag{3.19}$$

For the initial boundary value problem, keeping in mind the fact that $\xi_{\frac{1}{2}}^- = 0$, (3.17) becomes

$$(e_t, \phi) + (e, \phi_t) = (e_t, \mathbb{P}_+^\perp \phi) - (\eta, \phi_x) - \sum_{j=2}^{N} \xi^-[\phi]_{j-\frac{1}{2}} + \xi^-\phi^-|_{N+\frac{1}{2}}. \tag{3.20}$$

Integrating the above equation in t, and noticing the fact that

$$\int_0^T \sum_{j=1}^{N} \xi^-[\phi]_{j-\frac{1}{2}} = 0, \quad \text{and} \quad \int_0^T \xi^-\phi^-|_{N+\frac{1}{2}} = 0,$$

since $[\phi(t)]_{i-\frac{1}{2}} = 0$ except for at most finitely many t, and $\phi^-(t)|_{N+\frac{1}{2}} = 0$ when $t < T$, we again obtain (3.19).

We use integration by parts on the second term of the right hand side of (3.19),

$$\int_0^T (\eta, \phi_t)dt = (\eta, \psi)(T) - (\eta, \phi)(0) - \int_0^T (\eta_t, \phi)dt. \tag{3.21}$$

Plugging (3.21) into the second term on the right hand side of (3.19), then plugging (3.19) into the right hand side of (3.16), we obtain

$$(e(T), \psi) = (e, \phi(0)) + \int_0^T (e_t, \mathbb{P}_+^\perp \phi)dt + (\eta, \psi)(T) - (\eta, \phi)(0) - \int_0^T (\eta_t, \phi)dt.$$

Noticing that $\mathbb{P}_- u - u_h = e - \eta$, we have

$$(\mathbb{P}_- u - u_h, \psi)(T) = \Pi_1^j + \Pi_2^j + \Pi_3^j,$$

where

$$\Pi_1^j = (\mathbb{P}_- u - u_h, \phi)(0),$$

$$\Pi_2^j = -\int_0^T (\mathbb{P}_-^\perp u_t, \phi)dt,$$

$$\Pi_3^j = \int_0^T (e_t, \mathbb{P}_+^\perp \phi)dt.$$

For the first term, notice the fact that at $t = 0$, the support of ϕ_j is of length at least $h_{\min}$. Therefore, each cell contains at most $\lceil\Lambda\rceil + 1$ such ϕ_j, where $\lceil\Lambda\rceil$ denotes the smallest integer no smaller than Λ. In Section 3.2.1 we obtained the estimate $\|\xi(0)\|_\Omega \le Ch^{k+2}|u|_{k+2,\Omega}$, such that

$$\begin{aligned}\sum_{j=1}^N (\Pi_1^j)^2 &= \sum_{j=1}^N (\mathbb{P}_- u - u_h, \phi_j)^2(0) \\ &\le Ch(\lceil\Lambda\rceil + 1)\|\xi\|_\Omega^2 \\ &\le Ch^{2k+5}|u|_{k+2,\Omega}^2. \end{aligned} \tag{3.22}$$

The estimate of Π_2^j

We proceed to estimate $\Pi_2^j = -\int_0^T (\mathbb{P}_-^\perp u_t, \phi)$. For simplicity, only a periodic boundary condition is considered. However the estimate of the initial-boundary value problem can be obtained in exactly the same way. We extend our meshes onto the whole real line periodically, so the domain under consideration in this and the next subsections is $R \times [0, T]$. Clearly, the characteristic line which passes through $(x_{j-\frac{1}{2}}, T)$, denoted by l_j, is $t = x + T - x_{j-\frac{1}{2}},\ 0 < t < T$. We also assume that l_j and the cell boundary $x_{i-\frac{1}{2}} \times [0, T]$ intersect at $t = t_i^j$. Denote the support of ϕ in $R \times [0, T]$ as Ω_j, then the boundaries of the cells separate Ω_j into several pieces, as shown in Figure 3.1. Denote $\Omega_i^j = \Omega_j \cap I_i \times [0, T]$ and $k_j = \min\{i : \Omega_i^j \text{ is not empty}\}$. Also define $\Delta_j = \{k_j, k_j + 1, \cdots, j\}$ to be the index set which contains the subscripts of all the nonempty pieces. Then we can easily realize the following properties:

(1) $\Omega_j = \cup_{i\in\Delta_j}\Omega_i^j$ and $|\Delta_j| = j - k_j + 1 \leq \lceil \frac{T\Lambda}{h} \rceil + 1$.

(2) Denote $\widetilde{\Delta}_j = \{i \in \Delta_j | \Omega_i^j$ is not a parallelogram$\}$, then $|\widetilde{\Delta}_j| \leq \lceil \Lambda \rceil + 2$ and $j \in \widetilde{\Delta}_j$.

(3) Among those which are not parallelograms, Ω_j^j is a triangle which lies in the region $R \times [T - h, T]$, and by denoting $\widetilde{\Omega}_j = \cup_{i\in\widetilde{\Delta}_j\backslash j}\Omega_i^j$, we have $\widetilde{\Omega}_j \in R \times [0, 2h]$.

(4) Suppose Ω_i^j is a parallelogram then the vertices are $(x_{i-\frac{1}{2}}, t_i^j)$, $(x_{i+\frac{1}{2}}, t_{i+1}^j)$, $(x_{i+\frac{1}{2}}, t_{i+1}^{j+1})$, and $(x_{i-\frac{1}{2}}, t_i^{j+1})$.

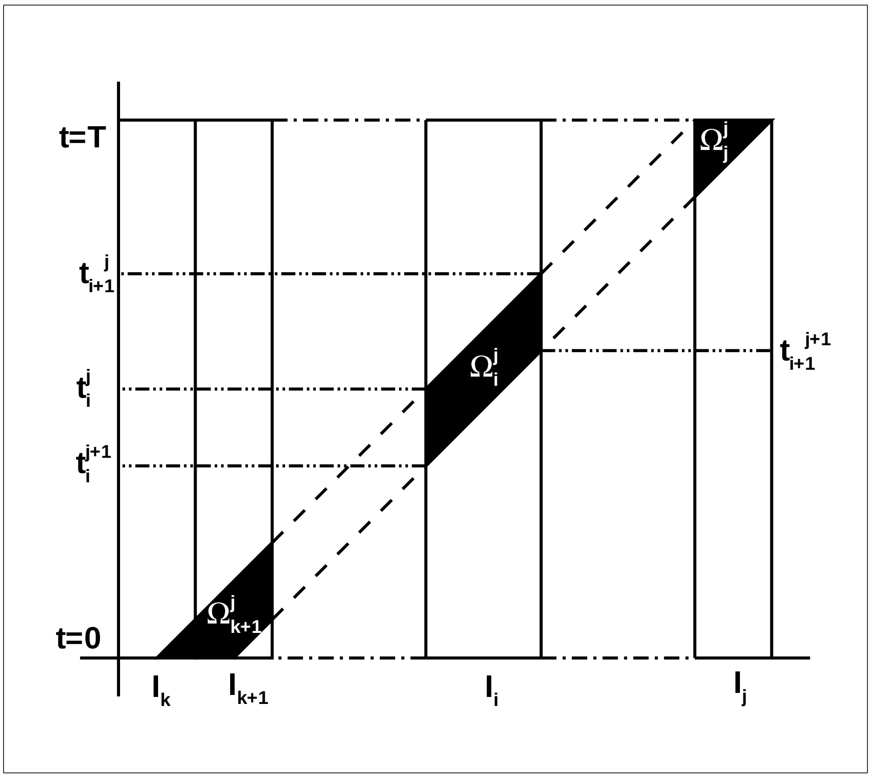

Figure 3.1: The support of ϕ: black polygons along the dashed line.

Now we can proceed to obtain the estimate

$$\int_0^T (\mathbb{P}_-^\perp u_t, \phi) = \sum_{i\in\Delta_j} \int_{\Omega_i^j} \mathbb{P}_-^\perp u_t \ \phi \ dxdt.$$

Consider the parallelogram Ω_i^j. And noticing the fact

$$\begin{aligned}\int_{t_i^{j+1}}^{t_{i+1}^j} \left(\mathbb{P}_-^\perp u_t(t_i^{j+1}), \phi\right) dt &= \left(\mathbb{P}_-^\perp u_t(t_i^{j+1}), \int_{t_i^{j+1}}^{t_{i+1}^j} \phi \; dt\right) \\ &= \left(\mathbb{P}_-^\perp u_t(t_i^{j+1}), \int_{I_j} \psi \; dx\right) \\ &= 0. \end{aligned} \tag{3.23}$$

Then we have

$$\begin{aligned}\int_{\Omega_i^j} \mathbb{P}_-^\perp u_t \; \phi \; dxdy &= \int_{t_i^{j+1}}^{t_{i+1}^j} (\mathbb{P}_-^\perp u_t, \phi) dt \\ &= \int_{t_i^{j+1}}^{t_{i+1}^j} (\mathbb{P}_-^\perp u_t(t) - \mathbb{P}_-^\perp u_t(t_i^{j+1}), \phi) dt \\ &= \int_{t_i^{j+1}}^{t_{i+1}^j} \left(\mathbb{P}_-^\perp \left(\int_{t_i^{j+1}}^{t} u_{tt}(\tau) d\tau\right), \phi\right) dt \\ &= \int_{t_i^{j+1}}^{t_{i+1}^j} \int_{t_i^{j+1}}^{t} \left(\mathbb{P}_-^\perp u_{tt}(\tau), \phi\right) d\tau dt \\ &\leq Ch^{k+4} |u|_{k+3,\infty,\Omega}. \end{aligned} \tag{3.24}$$

Now, we consider Ω_j^j and $\widetilde{\Omega}_j$. By using the third property of the partition of the support of Ω_j, we have

$$\int_{\Omega_j^j} \mathbb{P}_-^\perp u_t \; \phi \; dxdt \leq Ch^{k+3} |u|_{k+2,\infty,\Omega},$$

and

$$\int_{\widetilde{\Omega}_j} \mathbb{P}_-^\perp u_t \; \phi \; dxdt \leq Ch^{k+3} |u|_{k+2,\infty,\Omega}.$$

Combining the above, we obtain

$$\int_0^T (\mathbb{P}_-^\perp u_t, \phi) \leq Ch^{k+3} |u|_{k+2,\infty,\Omega} + CTh^{k+3} |u|_{k+3,\infty,\Omega}.$$

The estimate of Π_3^j

We still consider periodic boundary condition and follow the procedure in the previous section. However, there are two differences:

(1) The support of $\mathbb{P}_+^{\perp}\phi$, denoted as T_j, is different from Ω_j.

(2) We do not have the local estimates of $\|e_{tt}\|_{I_j}$ or $\|e_t\|_{I_j}$.

To deal with the first one, we define $T_i^j = T_j \cap I_i \times (0,T)$. Clearly T_i^j is a rectangle covering Ω_i^j and can be written as $T_i^j = I_i \times (t_0, t_1)$, where $t_0 = \inf\{t : \exists x \in I_i \ s.t. \ (x,t) \in \Omega_i^j\}$ and $t_1 = \sup\{t : \exists x \in I_i \ s.t. \ (x,t) \in \Omega_i^j\}$. If Ω_i^j is a parallelogram, then $t_0 = t_i^{j+1}$ and $t_1 = t_{i+1}^j$ (see Figure 3.2). We also denote $\widetilde{T}_j = \cup_{i\in\tilde{\Delta}_j\setminus j} T_i^j$, then $T_j = \cup_{i\in\Delta_j} T_i^j$. Moreover, it is not difficult to obtain $T_j^j \subset I_j \times (T-h, T)$ and $\widetilde{T}_j \subset R \times (0, 2h)$. Consider T_i^j such that Ω_i^j is a parallelogram. And notice that

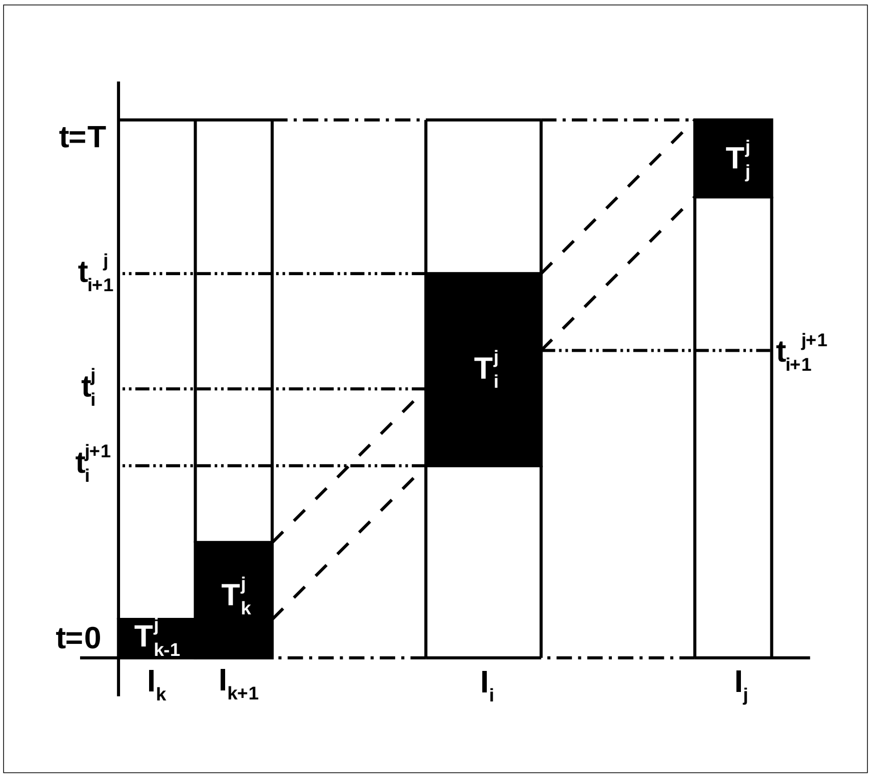

Figure 3.2: The support of $\mathbb{P}_+\phi$: the black boxes along the dashed line.

$$\begin{aligned}\int_{t_i^{j+1}}^{t_{i+1}^j}(e_t(t_i^{j+1}),\mathbb{P}_+^\perp\phi)dt &= \left(e_t(t_i^{j+1}),\mathbb{P}_+^\perp\int_{t_i^{j+1}}^{t_{i+1}^j}\phi\ dt\right)\\ &= \left(e_t(t_i^{j+1}),\mathbb{P}_+^\perp\int_{I_j}\psi\ dx\right)\\ &= 0. \qquad (3.25)\end{aligned}$$

Then we have

$$\begin{aligned}\int_{T_i^j} e_t\ \mathbb{P}_+^\perp\phi\ dxdt &= \int_{t_i^{j+1}}^{t_{i+1}^j}(e_t,\mathbb{P}_+^\perp\phi)dt\\ &= \int_{t_i^{j+1}}^{t_{i+1}^j}\left(e_t(t)-e_t(t_i^{j+1}),\mathbb{P}_+^\perp\phi\right)dt\\ &= \int_{t_i^{j+1}}^{t_{i+1}^j}\left(\int_{t_i^{j+1}}^{t} e_{tt}(\tau)d\tau,\mathbb{P}_+^\perp\phi\right)dt\\ &\le Ch^{3/2}\int_{t_i^{j+1}}^{t_{i+1}^j}\|e_{tt}\|_{I_i}dt. \qquad (3.26)\end{aligned}$$

Observe that $\sup\{t:(x,t)\in\widetilde{T}_j\}\le 2h$ and $\inf\{t:(x,t)\in T_j^j\}\ge T-h$. Therefore

$$\int_{T_j^j} e_t\ \mathbb{P}_+^\perp\phi\ dxdt \le Ch_j^{1/2}\int_{T-h}^{T}\|e_t\|_{I_j}dt,$$

and

$$\begin{aligned}\int_{\widetilde{T}_j} e_t\ \mathbb{P}_+^\perp\phi\ dxdt &\le C\int_0^{2h}\sum_{i\in\widetilde{\Delta}_j\backslash j}\|e_t\|_{I_i}h_i^{1/2}dt\\ &\le Ch^{1/2}\int_0^{2h}\left(\sum_{i\in\widetilde{\Delta}_j\backslash j}\|e_t\|_{I_i}^2\right)^{1/2}dt. \qquad (3.27)\end{aligned}$$

Combining the above, we obtain the estimate

$$\Pi_3^j\le C\Gamma_1^j+C\Gamma_2^j+C\Gamma_3^j,$$

where

$$\Gamma_1^j = h^{3/2} \sum_{i \in \Delta_j \backslash \widetilde{\Delta}_j} \int_{t_i^{j+1}}^{t_{i+1}^j} \|e_{tt}\|_{I_i} dt,$$

$$\Gamma_2^j = h_j^{1/2} \int_{T-h}^{T} \|e_t\|_{I_j} dt,$$

$$\Gamma_3^j = h^{1/2} \int_0^{2h} \left(\sum_{i \in \widetilde{\Delta}_j \backslash j} \|e_t\|_{I_i}^2 \right)^{1/2} dt.$$

As mentioned at the beginning, we do not have the local estimates of $\|e_t\|$ or $\|e_{tt}\|$, so we need to consider the summation with respect to j.

First, we consider Γ_1^j. Keeping in mind the fact that, for any $t \in (0, T)$ and $1 \le i \le N$, the information of $\|e_{tt}(t)\|_{I_i}$ is contained in at most $\lceil \Lambda \rceil + 1$ instances of Γ_1^j, we have

$$\begin{aligned}
\sum_{j=1}^N |\Gamma_1^j|^2 &\le \sum_{j=1}^N h^3 \sum_{i \in \Delta_j \backslash \widetilde{\Delta}_j} \lceil \frac{T\Lambda}{h} \rceil \left(\int_{t_i^{j+1}}^{t_{i+1}^j} \|e_{tt}\|_{I_i} dt \right)^2 \\
&\le CTh^3 \sum_{j=1}^N \sum_{i \in \Delta_j \backslash \widetilde{\Delta}_j} \int_{t_i^{j+1}}^{t_{i+1}^j} \|e_{tt}\|_{I_i}^2 dt \\
&\le CT(\lceil \Lambda \rceil + 1) h^3 \int_0^T \|e_{tt}\|_\Omega^2 dt \\
&\le CTh^{2k+5} \int_0^T (|u|_{k+3,\Omega} + t|u|_{k+4,\Omega})^2 dt \\
&\le Ch^{2k+5} \left(T^2 |u|_{k+3,\Omega}^2 + T^4 |u|_{k+4,\Omega}^2 \right). \qquad (3.28)
\end{aligned}$$

The second term is easy to deal with since

$$\begin{aligned}\sum_{j=1}^{N}|\Gamma_2^j|^2 &\leq \sum_{j=1}^{N} h^2\int_{T-h}^{T}\|e_t\|_{I_j}^2 dt\\ &= h^2\int_{T-h}^{T}\|e_t\|_{\Omega}^2 dt\\ &\leq Ch^{2k+4}\int_{T-h}^{T}(|u|_{k+2,\Omega}+t|u|_{k+3,\Omega})^2 dt\\ &\leq Ch^{2k+5}\left(|u|_{k+2,\Omega}^2+T^2|u|_{k+3,\Omega}^2\right)dt. \qquad (3.29)\end{aligned}$$

The third term is also not difficult. Notice that for fixed i, $\int_0^{2h}\|e_t\|_{I_i}$ is contained in at most $\lceil\Lambda\rceil+1$ instances of Γ_3^j, we have

$$\begin{aligned}\sum_{j=1}^{N}|\Gamma_3^j|^2 &\leq \sum_{j=1}^{N} h^2\int_{0}^{2h}\sum_{i\in\tilde{\Delta}_j\setminus j}\|e_t\|_{I_j}^2 dt\\ &\leq (\lceil\Lambda\rceil+1)h^2\int_{0}^{2h}\|e_t\|_{\Omega}^2 dt\\ &\leq Ch^{2k+4}\int_{0}^{2h}(|u|_{k+2,\Omega}+t|u|_{k+3,\Omega})^2 dt\\ &\leq Ch^{2k+5}\left(|u|_{k+2,\Omega}^2+h^2|u|_{k+3,\Omega}^2\right)dt. \qquad (3.30)\end{aligned}$$

Combining the above, we obtain

$$\sum_{j=1}^{N}|\Pi_3^j|^2 \leq C\sum_{j=1}^{N}\left(|\Gamma_1^j|^2+|\Gamma_2^j|^2+|\Gamma_3^j|^2\right) \leq C(1+T^4)h^{2k+5}\|u\|_{k+4,\Omega}^2.$$

Remark 3.2.3. *By the same method mentioned in this subsection, we can also derive that*

$$\sum_{j=1}^{N}|\Pi_2^j|^2 \leq C(1+T^2)h^{2k+5}\|u\|_{k+4,\Omega}^2.$$

Here the upper bound is of T^2 instead of T^4 since $\|\eta_t\|_\Omega$ and $\|\eta_{tt}\|_\Omega$ do not grow in time.

3.2.4 Final estimate

Now we proceed to the final estimate of $|\xi(x_j)|$. We simply sum up all the previous estimates and obtain

$$\begin{aligned}\sum_{j=1}^{N}|(\xi,\psi_j)|^2 &\leq 3\sum_{j=1}^{N}\left(|\Pi_1^j|^2+|\Pi_2^j|^2+|\Pi_3^j|^2\right)\\ &\leq Ch^{2k+5}\left(|u|_{k+2,\Omega}^2+(1+T^2)\|u\|_{k+4,\Omega}^2+(1+T^4)\|u\|_{k+4,\Omega}^2\right)\\ &\leq Ch^{2k+5}(1+T^4)\|u\|_{k+4,\Omega}^2. \end{aligned} \qquad (3.31)$$

Therefore,

$$\begin{aligned}\frac{1}{N}\sum_{j=1}^{N}|\xi(x_j)|^2 &= \frac{1}{N}\sum_{j=1}^{N}\left|\frac{2}{h_j}(\xi,\psi_j)\right|^2\\ &\leq Ch^{2k+4}(1+T^4)\|u\|_{k+4,\Omega}^2.\end{aligned}$$

By Lemma 2.3.3,

$$\frac{1}{N}\sum_{j=1}^{N}|(u-u_h)(x_j)|^2 \leq Ch^{2k+4}(1+T^4)\|u\|_{k+4,\infty,\Omega}^2. \qquad (3.32)$$

This completes the proof of the main theorem.

Notice that, throughout the proof, we have not used any special property of ψ. Hence we can take ψ to be the indicator function of cell I_j, which yields the estimate of (3.4). Finally, (3.5) follows from (3.11) and (3.4).

3.2.5 Applications

In Section 3.2.2, we proved optimal error estimates in L^∞ by using the superconvergence of $\xi = u_h - \mathbb{P}_- u$. This can be considered as an application of the superconvergence result. Let us also briefly discuss some other applications. A notice is to use the superconvergence of the cell averages to construct a new *a posterior* error

indicator, in the same spirit as in [67] where the jump sizes at cell interfaces are used as an *a posterior* error indicator. To expose this, we denote v as the numerical solution instead of u_h in this section and consider the cell I_j only. Following the superconvergence result, the cell average of the numerical solution $\overline{v_j}$ is superclose to that of the exact solution $\overline{u_j}$. If we construct another numerical cell average $\widetilde{v_j}$ which is not superconvergent, then the difference $\widetilde{v_j} - \overline{v_j}$ is a good *a posterior* error indicator of the local error. We can construct $\widetilde{v_j}$ in the following steps:
(1) Extend the polynomial numerical solution from the two neighboring cells, denoted by v_{j-1} and v_{j+1}, to the cell I_j.
(2) Compute the cell averages of v_{j-1} and v_{j+1} in the cell I_j, and denote them by $\widetilde{v_{j-1}}$ and $\widetilde{v_{j+1}}$ respectively.
(3) Define

$$\widetilde{v_j} = \theta\widetilde{v_{j-1}} + (1-\theta)\widetilde{v_{j+1}},$$

where $0 \le \theta \le 1$. In general, the new cell average $\widetilde{v_j}$ is only $(k+1)$-th order accurate.
(4) The *a posterior* computable quantity $\widetilde{v_j} - \overline{v_j}$ is asymptotically equal to the error $\widetilde{v_j} - \overline{u_j}$ and is therefore a good indicator of the local error.
Numerical evidence will be given in Section 3.3.

Based on the superconvergence of the downwind-biased Radau points, we can construct another *a posterior* error indicator. We use v_j and u_j as the numerical and exact solutions in cell I_j, respectively. Denote x_j^i, $0 \le i \le k$ as the downwind-biased Radau points and $\hat{S} = \{x_{j-\frac{1}{2}}, x_j^0, \cdots, x_j^k\}$. Then we construct another numerical approximation $w_j \in \mathcal{P}^{k+1}(I_j)$ which interpolates v_j at $x \in \hat{S}$. For convenience, if $\hat{x} \in \hat{S}$ is located at the cell interface, $v(\hat{x})$ is denoted as the left limiter of the numerical approximation. Therefore, we have $w_j(x^+_{j-\frac{1}{2}}) = v_{j-1}(x^-_{j-\frac{1}{2}})$ and $w_j(x^-_{j+\frac{1}{2}}) = v_j(x^-_{j+\frac{1}{2}})$. We define $w(x)$ to be such a function such that for any j, $w(x)$ agrees with w_j in cell I_j. Obviously, $w(x)$ is a continuous function. In (3.12), we mentioned that the error between the numerical solution and the exact solution is not superconvergent. If we can show that the new numerical approximation w_j is superconvergent in the

L^∞-norm, the difference $v_j - w_j$ is a good *a posterior* error indicator of the local error. Let $\widetilde{w}_j \in \mathcal{P}^{k+1}(I_j)$ be a polynomial thus interpolates u_j at $x \in \hat{S}$. Clearly,

$$\|u_j - \widetilde{w}_j\|_{\infty, I_j} \leq Ch^{k+2}. \tag{3.33}$$

On the other hand, for any $\hat{x} \in \hat{S}$, Theorem 3.1.1 yields $|v_j(\hat{x}) - u_j(\hat{x})| \leq Ch^{k+3/2}$. Then we have $|w_j(\hat{x}) - \widetilde{w}_j(\hat{x})| \leq Ch^{k+3/2}$. Define a special norm in $\mathcal{P}^{k+1}(I_j)$ as

$$||v||_{\hat{S}} = \max_{x \in \hat{S}} \{|v(x)|\}.$$

It is not difficult to show this is indeed a norm. Since all norms in $\mathcal{P}^{k+1}$ are equivalent, we have $\|v\|_{\infty, I_j} \leq C||v||_{\hat{S}}$ for any $v \in \mathcal{P}^{k+1}(I_j)$. Therefore,

$$\|w_j - \widetilde{w}_j\|_{\infty, I_j} \leq C||w_j - \widetilde{w}_j||_{\hat{S}} \leq Ch^{k+3/2}.$$

By using (3.33), we have

$$\|u_j - w_j\|_{\infty, I_j} \leq Ch^{k+3/2},$$

which further yields

$$\|u_j - w_j\|_\infty \leq Ch^{k+3/2}.$$

3.3 Numerical tests

The purpose of this section is to verify our main result, Theorem 3.1.1 as well as Corollary 3.1.1, and to present numerical evidence suggesting that the proved rate of superconvergence is optimal. In most cases, we consider random meshes (that is, each cell boundary point is randomly and independently perturbed from a uniform mesh up to a given percentage) and use Λ to denote the ratio of the length of the largest cell to that of the smallest one.

Example 3.1. *We solve the following equation*

$$\begin{cases} u_t + u_x = 0 \\ u(x,0) = e^{\sin(x)} \\ u(0,t) = u(2\pi, t) \end{cases} . \tag{3.34}$$

The exact solution to this problem is

$$u(x,t) = e^{\sin(x-t)}.$$

We use ninth order SSP Runge-Kutta discretization in time [46] and take $\Delta t = 0.05h_{\min}$ to reduce the time error. Non-uniform meshes are obtained by randomly and independently perturbing each node in a uniform mesh by up to 40%, and the example is tested with both $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials. The error in Theorem 3.1.1 at different downwind-biased Radau points at $t = 1$ on random meshes of N cells are computed. In Table 3.1, we observe $(2k+1)$-th order superconvergence at the downwind point and $(k+2)$-th order superconvergence at other Radau points. The initial solution is obtained by exactly the same way as mentioned in Section 3.2.1. The downwind-biased Radau points on the interval [-1,1] are $-\frac{1}{3}$ and 1 for $\mathcal{P}^1$ polynomials, and are $\frac{-1-\sqrt{6}}{5}$, $\frac{-1+\sqrt{6}}{5}$ and 1 for $\mathcal{P}^2$ ones.

Table 3.2 shows the rate of convergence of the error ξ. We observe that the order is $k+2$, indicating that the estimate in (3.5) is sharp.

Moreover, we also test the superconvergence for the cell average. Table 3.3 shows the result for Example 3.1 by using the method mentioned in Section 3.2.1 as well as with L^2- and $\mathbb{P}_-$-projection for the initial discretization. From the table, we find the convergent rates to be of order $2k+1$, $k+\frac{3}{2}$ and at least $k+2$, respectively, for the three different ways of numerical initial discretization.

Now we follow the steps in Section 3.2.5 and construct the new numerical cell average. Following the same notations, θ is taken to be $\theta = 1$, i.e. we consider the

Table 3.1: The error e at the Radau points for (3.34) when using $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials.

				1^{st} Radau point		2^{nd} Radau point		downwind point	
Polynomial	N	h_{max}	Λ	error	order	error	order	error	order
$\mathcal{P}^1$	50	0.202	6.056	1.86E-04	-			1.34E-04	-
	100	0.111	6.169	3.76E-05	2.64			2.87E-05	2.55
	200	5.408e-02	7.339	3.89E-06	3.17			2.92E-06	3.19
	400	2.781e-02	6.956	4.90E-07	3.12			3.80E-07	3.07
$\mathcal{P}^2$	50	0.202	6.056	1.11E-06	-	1.09E-06	-	2.13E-07	-
	100	0.111	6.169	1.70E-07	3.09	1.42E-07	3.37	1.78E-08	4.10
	200	5.408e-02	7.339	1.03E-08	3.93	7.94E-09	4.04	4.28E-10	5.21
	400	2.781e-02	6.956	7.74E-10	3.89	5.81E-10	3.93	1.37E-11	5.18

Table 3.2: The error ξ for equation (3.34) when using $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials.

L^2 norm of ξ			$\mathcal{P}^1$ Polynomial		$\mathcal{P}^2$ Polynomial	
N	h_{max}	Λ	L^2 error	order	L^2 error	order
50	0.202	6.056	4.09E-04	-	1.85E-06	-
100	0.111	6.169	8.67E-05	2.56	2.93E-07	3.05
200	5.408e-02	7.339	8.84E-06	3.19	1.70E-08	3.99
400	2.781e-02	6.956	1.11E-06	3.12	1.29E-19	3.88

Table 3.3: The cell average of the error e for equation (3.34) when using $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials.

L^2-norm of the cell average of e				$\mathcal{P}^1$ polynomial		$\mathcal{P}^2$ polynomial	
initial discretization	N	h_{max}	Λ	L^2 error	order	L^2 error	order
$u_{ht} = \mathbb{P}_- u_t$	50	0.202	6.056	3.84E-04	-	5.18E-07	-
$\int_\Omega (u_h - u)dx = 0$	100	0.111	6.169	8.23E-05	2.54	4.70E-08	3.96
	200	5.408e-02	7.339	8.42E-06	3.19	1.08E-09	5.29
	400	2.781e-02	6.956	1.06E-06	3.11	3.33E-11	5.23
L^2 projection	50	0.202	6.056	3.87E-04	-	5.56E-06	-
	100	0.111	6.169	1.19E-04	1.94	1.18E-06	2.56
	200	5.408e-02	7.339	1.38E-05	3.02	1.11E-07	3.31
	400	2.781e-02	6.956	2.02E-06	2.89	7.65E-09	4.02
$\mathbb{P}_-$ projection	50	0.202	6.056	3.35E-04	-	1.07E-06	-
	100	0.111	6.169	7.38E-05	2.50	9.30E-08	4.03
	200	5.408e-02	7.339	7.72E-06	3.16	3.47E-09	4.60
	400	2.781e-02	6.956	9.81E-07	3.10	1.72E-10	4.52

extension from the downwind cell to the right. We use $\mathcal{P}^2$ polynomials on a uniform mesh with $N = 100$. Define piecewise constants $s(x)$ such that in cell I_j

$$s(x) = s_j = \frac{\widetilde{v_j} - \overline{v_j}}{\widetilde{v_j} - \overline{u_j}} - 1,$$

where $\overline{u_j}$ denotes the cell average of the exact solution u in cell I_j. We compute and observe that $\|s\|_{\infty,\Omega} = 7.90 \times 10^{-4}$, indicating that the computable quantity $\widetilde{v_j} - \overline{v_j}$ is a good estimate of the local error $\widetilde{v_j} - \overline{u_j}$. From Figure 3.3, we observe that the computable quantity $|\bar{v} - \tilde{v}|$ captures the profile of the local error $|u - v|$ (computed at the middle point in each cell) well.

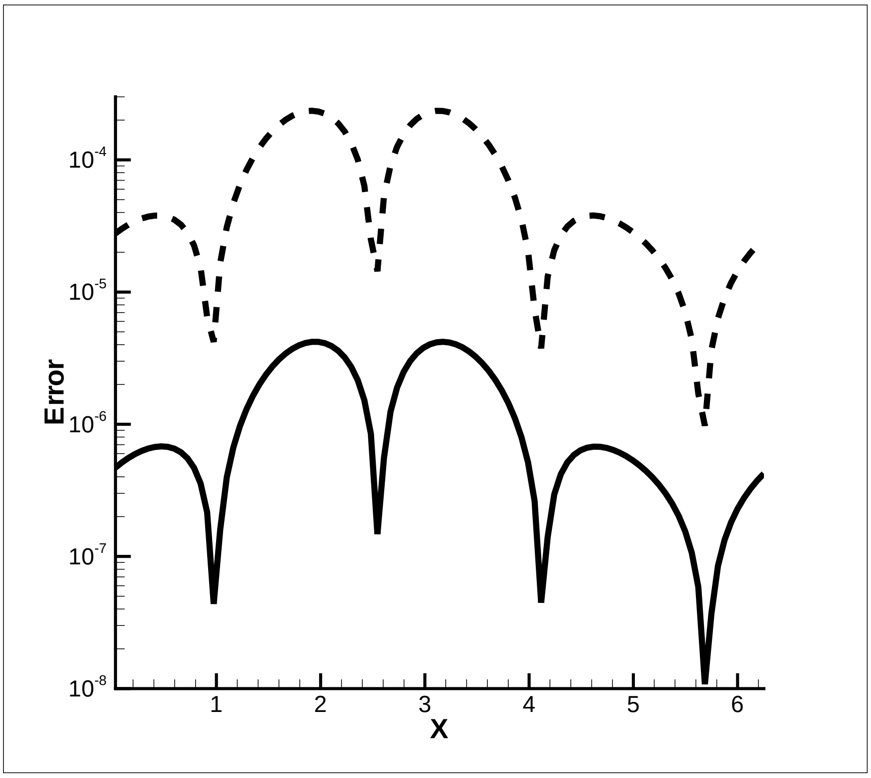

Figure 3.3: Comparison between $|u - v|$ (solid line) and $|\bar{v} - \tilde{v}|$ (dashed line).

Example 3.2. *We solve the following initial boundary value problem*

$$\begin{cases} u_t + u_x = 0 \\ u(x,0) = \sin(x) \\ u(0,t) = \sin(-t) \end{cases} . \tag{3.35}$$

The exact solution to this problem is

$$u(x,t) = \sin(x-t).$$

We use third order SSP Runge-Kutta discretization in time, take $\Delta t = 0.1h_{\min}^2$ to reduce the time error, and test the example with both $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials. The same quantities as in Example 3.1 on the same kind of random meshes of N cells are computed. The initial solution is obtained as given in Section 3.2.1. In Table 3.4 we observe that the error between the DG solution and the exact solution is $(2k+1)$-th order superconvergent at the downwind point and $(k+2)$-th order superconvergent at the other Radau points.

Table 3.4: The error e at the Radau points for (3.35) when using $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials.

				1^{st} Radau point		2^{nd} Radau point		downwind point	
Polynomial	N	h_{max}	Λ	error	order	error	order	error	order
$\mathcal{P}^1$	50	0.202	6.056	6.06E-05	-			2.88E-05	-
	100	0.111	6.169	8.92E-06	3.16			4.30E-06	3.14
	200	5.408e-02	7.339	1.02E-06	3.04			4.73E-07	3.09
	400	2.781e-02	6.956	1.25E-07	3.15			5.82E-08	3.15
$\mathcal{P}^2$	50	0.202	6.056	4.10E-07	-	3.11E-07	-	1.30E-08	-
	100	0.111	6.169	3.20E-08	4.21	2.38E-08	4.24	5.51E-10	5.22
	200	5.408e-02	7.339	1.84E-09	4.00	1.38E-09	3.99	1.45E-11	5.06
	400	2.781e-02	6.956	9.36E-11	4.48	1.05E-10	3.87	4.31E-13	5.28

Table 3.5 shows the $(k+2)$-th order superconvergence of the error ξ in the L^2-

norm, demonstrating that the estimate in (3.5) is sharp.

Table 3.5: The error ξ for (3.35) when using $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials.

L^2-norm of ξ			$\mathcal{P}^1$ polynomial		$\mathcal{P}^2$ polynomial	
N	h_{max}	Λ	L^2 error	order	L^2 error	order
50	0.202	6.056	1.24E-04	-	6.63E-07	-
100	0.111	6.169	1.87E-05	3.13	5.32E-08	4.17
200	5.408e-02	7.339	2.10E-06	3.06	3.03E-09	4.01
400	2.781e-02	6.956	2.58E-07	3.15	2.30E-10	3.88

As in Example 3.1, we also test the superconvergence for the cell average. Table 3.6 shows the result for Example 3.2 using the method in Section 3.2.1 as well as the L^2- and $\mathbb{P}_-$-projections for the initial discretization. We observe results similar to those in the periodic case.

Example 3.3. *We solve the following two-dimensional problem*

$$\begin{cases} u_t + u_x + u_y = 0 \\ u(x, y, 0) = \sin(x + y) \end{cases}, \tag{3.36}$$

with periodic boundary condition on the domain $[0, 2\pi]^2$.

The exact solution is

$$u(x, y, t) = \sin(x + y - 2t).$$

We use a random rectangular mesh defined as

$$0 = x_{\frac{1}{2}} < \cdots < x_{N_x+\frac{1}{2}} = 2\pi,\ 0 = y_{\frac{1}{2}} < \cdots < y_{N_y+\frac{1}{2}} = 2\pi$$

and

$$I_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}].$$

Table 3.6: The cell average of the error e for (3.35) when using $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials.

L^2 norm of the cell average of e				$\mathcal{P}^1$ polynomial		$\mathcal{P}^2$ polynomial	
initial discretization	N	h_{max}	Λ	L^2 error	order	L^2 error	order
$u_{ht} = \mathbb{P}_- u_t$	50	0.202	6.056	1.11E-04	-	3.75E-08	-
$u_{h_{j+\frac{1}{2}}}^- = \mathbb{P}_- u_{j+\frac{1}{2}}^-$	100	0.111	6.169	1.67E-05	3.12	1.63E-09	5.18
	200	5.408e-02	7.339	1.90E-06	3.04	4.07E-11	5.16
	400	2.781e-02	6.956	2.35E-07	3.14	1.14E-12	5.37
L^2 projection	50	0.202	6.056	1.86E-04	-	2.59E-06	-
	100	0.111	6.169	5.48E-05	2.02	3.59E-07	3.26
	200	5.408e-02	7.339	6.86E-06	2.91	3.11E-08	3.43
	400	2.781e-02	6.956	1.02E-06	2.87	2.86E-09	3.59
$\mathbb{P}_-$ projection	50	0.202	6.056	1.01E-04	-	2.49E-07	-
	100	0.111	6.169	1.53E-05	3.12	1.48E-08	4.66
	200	5.408e-02	7.339	1.72E-06	3.06	5.73E-10	4.55
	400	2.781e-02	6.956	2.06E-07	3.20	2.48E-11	4.72

We define the approximation space as

$$V_h^k = \left\{u_h : u_h|_{I_{i,j}} \in \mathcal{Q}^k(I_{i,j}),\ 1 \le i \le N_x, 1 \le j \le N_y\right\},$$

where $\mathcal{Q}^k(I_{i,j})$ denotes all the tensor product polynomials of degree at most k in x and in y on $I_{i,j}$. The Gauss-Radau projection $\mathbb{P}_-$ is defined as follows:

$$\int_{I_{i,j}} (\mathbb{P}_-u - u)v_h dxdy = 0$$

for any $v_h \in V_h^{k-1}$,

$$\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} (\mathbb{P}_-u(x_{i+\frac{1}{2}}, y) - u(x_{i+\frac{1}{2}}, y))w_h(y)dy = 0$$

for any $w_h \in \mathcal{P}^{k-1}$,

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} (\mathbb{P}_-u(x, y_{j+\frac{1}{2}}) - u(x, y_{j+\frac{1}{2}}))z_h(x)dx = 0$$

for any $z_h \in \mathcal{P}^{k-1}$, and

$$\mathbb{P}_-u(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}) = u(x_{i+\frac{1}{2}}, y_{j+\frac{1}{2}}).$$

We also use an upwind flux and ninth order SSP Runge-Kutta discretization in time with $\Delta t = 0.1h_{\min}$. We test the example with both $\mathcal{Q}^1$ and $\mathcal{Q}^2$ polynomials, and compute $\xi = u_h - \mathbb{P}_-u$, the error of the cell average, as well as the error on the downwind-biased Radau points and the downwind point. For simplicity, we do not consider the Radau points one by one, but select the one which gives the largest error in each cell. The initial solution is given by the $\mathbb{P}_-$ projection.

It appears that similar superconvergence results are valid also in two-dimensions. However, the technique of the proof in this chapter, in particular the part related to the special projections, does not seem to be easily extendable to two-dimensions.

Table 3.7: Superconvergence results for equation (3.36) when using $\mathcal{Q}^1$ and $\mathcal{Q}^2$ polynomials.

				$\mathcal{Q}^1$ polynomial		$\mathcal{Q}^2$ polynomial	
Error	$N_x \times N_y$	h_{max}	Λ	L^2 error	order	L^2 error	order
$\|\mathbb{P}_- u - u_h\|$	10×10	0.997	3.722	2.65E-02	-	8.40E-04	-
	20×20	0.552	4.809	4.41E-03	3.04	6.97E-05	4.22
	40×40	0.270	7.339	5.12E-04	3.02	3.64E-06	4.13
	80×80	0.133	6.326	6.33E-05	2.96	2.33E-07	3.89
$\|\overline{u - u_h}\|$	10×10	0.997	3.722	4.19E-02	-	6.47E-04	-
	20×20	0.552	4.809	7.65E-03	2.88	5.26E-05	4.25
	40×40	0.270	7.339	9.12E-04	2.98	2.20E-06	4.44
	80×80	0.133	6.326	1.14E-04	2.94	1.16E-07	4.17
$\max_{i,j} \|(u - u_h)(x,y)\|$	10×10	0.997	3.722	2.28E-02	-	1.56E-03	-
(x,y) is the Downwind-	20×20	0.552	4.809	5.65E-03	2.37	1.35E-04	4.14
biased Radau points in $I_{i,j}$	40×40	0.270	7.339	6.39E-04	3.05	8.88E-06	3.81
	80×80	0.133	6.326	1.01E-04	2.62	7.15E-07	3.57
$\max_{i,j} \|(u - u_h)(x,y)\|$	10×10	0.997	3.722	1.99E-02	-	3.72E-04	-
(x,y) is the downwind	20×20	0.552	4.809	2.79E-03	3.33	4.09E-05	3.74
point in $I_{i,j}$	40×40	0.270	7.339	3.93E-04	2.74	1.49E-07	4.64
	80×80	0.133	6.326	4.65E-05	3.02	1.34E-07	3.42

3.4 Concluding remarks

We have studied the behavior of the error between the upwinded DG solution and the exact solution for sufficiently smooth solutions of linear conservation laws. We prove that under suitable initial discretization, the error between the DG solution and the exact solution is $(k+2)$-th order superconvergent at the downwind-biased Radau points. Moreover, numerical experiments show that the convergent rate is $(2k+1)$-th order superconvergent at the downwind point as well as in the L^2-norm of the cell average. We also prove that the DG solution is superconvergent with the rate $k+2$ towards a particular projection of the exact solution.

Chapter 4

Negative-order norm error estimates for linear hyperbolic equations involving δ-functions

In this chapter, we develop and analyze DG methods for solving hyperbolic conservation laws involving δ-functions. δ-function has many equivalent definitions, and one of them is the following weak form:

$$\int_R \delta(x)v(x)dx = v(0),\ \ v(x) \in C(R).$$

As mentioned in Chapter 1, the DG methods are based on a weak form, and can be designed to approximate δ-functions directly. However, in DG methods, the test functions are completely discontinuous across the cell interfaces. If the δ-function is placed at the cell interface, then the numerical scheme is not well-defined. In [68], the author introduced the distribution theory for discontinuous test function, which extended the definition of δ-function as

$$\int_R \delta(x)v(x)dx = \frac{v(0+)+v(0-)}{2},\ \ v(x) \text{ is piecewise continuous},$$

which further completes the definition of the DG schemes.

We consider negative-order norm error estimates. Such norms can be used to detect the oscillations of a function. In [34], Cockburn et al. proved high order superconvergence error estimates of DG methods including their divided differences for hyperbolic equations with smooth solutions in negative-order norms. They also demonstrated that the application of the post-processing techniques of Bramble and Schatz [17] can yield superconvergence in the strong L^2-norm. Other related works include [92, 103, 111, 91], where one-sided filter and local derivative post-processing were considered. We will extend the work in [34], and consider the general case where the exact solutions are not smooth.

The first example of non-smooth solutions for hyperbolic equations is the following problem

$$\begin{aligned} &u_t + u_x = 0, && (x,t) \in R \times (0,T], \\ &u(x,0) = u_0(x), && x \in R, \end{aligned} \tag{4.1}$$

where the initial solution $u_0(x)$ has compact support, with a discontinuity at $x = 0$, but is otherwise smooth. Clearly, the exact solution of (4.1) is discontinuous along the characteristic line $x = t$ and the numerical DG solution has spurious oscillations around this discontinuity line, which we refer to as the pollution region. There are not too many works in the literature studying error estimates of DG methods for problems with discontinuous solutions. The first work in this direction seems to be that of Johnson et al. [63, 64, 65] for DG methods in both space and time. They have shown that, with linear space-time elements, the width of the pollution region is of the size at most $\mathcal{O}(h^{1/2} \log 1/h)$. More recently, Cockburn and Guzmán [31] and Zhang and Shu [117] revisited this problem with the RKDG methods and obtained similar results. Especially, in [31], the left boundary of the pollution region is shown to be at most $\mathcal{O}(h^{2/3} \log(1/h))$ from the singularity for piecewise linear DG method with second order Runge-Kutta time discretization on uniform meshes. The first problem we consider in this chapter is (4.1) with the initial condition $u_0(x)$ having a δ-singularity at $x = 0$. We consider a semi-discrete DG method and use the result in [117] to prove superconvergence results estimated in negative-order norms outside the pollution region. By convolving the DG solution with a suitable kernel, the post-processed approximation is $(2k+1)$-th order accurate in a region slightly smaller than the one above. The rate of convergence agrees with that in [34], in which the initial datum $u_0(x)$ was assumed to be sufficiently smooth.

Hyperbolic conservation laws with source terms have been analyzed by several authors [21, 47, 66, 99, 72]. In particular, in [47], the authors studied the following problem

$$\begin{aligned} &u_t + f(u)_x = g(x,t), && (x,t) \in R \times (0,T], \\ &u(x,0) = u_0(x), && x \in R, \end{aligned} \tag{4.2}$$

where f is a smooth convex function ($f''(u) > 0$ for all u) and $g(x,t) = G_x(x,t)$ with G being a bounded, piecewise smooth function, and constructed L^∞-stable Godunov-type difference schemes. In [98], Santos and de Oliveira studied hyperbolic conservation laws whose source terms contain δ-singularities, and investigated the convergence of numerical discretization by using a finite volume scheme. Later, they considered a class of high resolution methods in [79]. In [78], Noussair studied the wave behavior of (4.2), where the source term also depends on u but not on the time variable t. We note that all these previous works did not provide any error estimates in the smooth region away from the singularities. In this chapter, we investigate a simpler case by assuming $f(u) = u$ and $g(x,t) = G'(x)$ in the source term in (4.2), where $G(x)$ is a step function which does not depend on the time variable t. We show that by convolving the DG solution with a suitable kernel, the post-processed approximation turns out to be $(2k + 1)$-th order superconvergent in the smooth region.

4.1 Post-processing

4.1.1 Convolution kernel

Now, we proceed to describe the type of post-processing to be considered, following Bramble and Schatz [17]. Let χ be the indicator function of the interval $(-\frac{1}{2}, \frac{1}{2})$, We define recursively the functions $\psi^{(l)}$ as

$$\psi^{(1)} = \chi, \qquad \psi^{(n+1)} = \psi^{(n)} * \psi^{(1)}, \text{ for } n \geq 1.$$

The numerical solution is post-processed by convolving it with a kernel $K^{(\nu,l)}(x)$ which satisfies the following properties:
(1) It has a compact support.
(2) It reproduces polynomials p of degree $\nu - 1$ by convolution: $K^{(\nu,l)} * p = p$.
(3) It is a linear combination of B-splines and is of the form

$$K^{(\nu,l)}(x) = \sum_{\gamma \in \mathbb{Z}} k_\gamma^{\nu,l} \psi^{(l)}(x - \gamma).$$

The weights $k_\gamma^{\nu,l} \in R$ are chosen so that (2) is satisfied. See [17, 34] for more details. We also define $K_H^{(\nu,l)}(x) = K^{(\nu,l)}(x/H)/H$ and $\psi_H^{(l)}(x) = \psi^{(l)}(x/H)/H$ and it is not difficult to verify that

$$D^\alpha(\psi^{(\beta)}) * v = \psi_H^{(\beta-\alpha)} * \partial_H^\alpha v,$$

where $\partial_H v(x) = \frac{1}{H}(v(x+\frac{1}{2}H) - v(x-\frac{1}{2}H))$. In general, we take $H = nh,\ n = 1, 2, \cdots$. This property is important as it allows us to express derivatives of the convolution with the kernel in terms of simple difference quotients.

4.1.2 An approximation result

Let us investigate the relationship between $u - K_h^{2k+2,k+1} * u_h$ and the negative-order norm estimates of divided differences of the error $u - u_h$.

Theorem 4.1.1 (Bramble and Schatz [17]). *Suppose the kernel $K_h^{\nu,l}$ satisfies the properties listed in Section 4.1.1. Let v be a function in $L^2(\Omega_1)$, where Ω_1 is an open set in Ω, and u be a function in $H^\nu(\Omega_1)$. Further assume Ω_0 to be an open set in Ω_1 such that $\Omega_0 + 2supp(K_h^{\nu,l}) \subset\subset \Omega_1$. Then we have*

$$\|u - K_h^{\nu,l} * v\|_{\Omega_0} \leq \frac{h^\nu}{\nu!} C_1 |u|_{\nu,\Omega_1} + C_1 C_2 \sum_{0 \leq \alpha \leq l} \|\partial_h^\alpha (u - v)\|_{-l,\Omega_1},$$

where $C_1 = \sum_{\gamma \in \mathbb{Z}} |k_\gamma^{\nu,l}|$ and C_2 only depends on $\Omega_0, \Omega_1, \nu,$ and l.

There is a straightforward corollary.

Corollary 4.1.1. *Suppose the conditions in Theorem 4.1.1 are satisfied. Further assume that $\|\partial_h^\alpha (u - v)\|_{-l,\Omega_1} \leq Ch^\mu$ is valid for all $\alpha \leq l$ and $\nu \geq \mu$. Then we have*

$$\|u - K_h^{\nu,l} * v\|_{\Omega_0} \leq Ch^\mu,$$

where C only depends on $\Omega_0, \Omega_1, \nu,$ and l.

4.2 Singular initial condition

Let us consider problem (4.1) and use upwind fluxes. We first state the main results in Theorem 4.2.1 and then give the proofs. We provide the negative-order norm error estimates in the whole space as well as in the region away from the singularities.

4.2.1 Main results

The following lemma is the semi-discrete version of the result in Zhang and Shu [117]. For completeness we will give its proof in Section 4.5.

Lemma 4.2.1. *Let u be the exact solution of the initial value problem* (4.1)*, where the initial condition $u_0(x) \in C^{k+2}$ except for one singularity at $x = 0$. Let u_h be the solution of the DG method* (2.5) *at time T, where the finite element space V_h is made up of the piecewise polynomials of degree $k \geq 1$. Suppose h is the maximum cell length. Then there holds the following error estimate*

$$\|u(T) - u_h(T)\|_{\Omega \backslash \mathcal{R}_T} \leq Ch^{k+1}, \tag{4.3}$$

where $\mathcal{R}_T = (T - Ch^{1/2}\log(1/h), T + Ch^{1/2}\log(1/h))$, and the bounding constant $C > 0$ does not depend on h.

We will use Lemma 4.2.1 to prove the following theorem.

Theorem 4.2.1. *Suppose $u \in C^{2k+2}$ and the conditions of Lemma 4.2.1 are satisfied. Then by taking $\Omega_0 + 2supp(K_h^{2k+2,k+1}) \subset\subset \Omega_1 \subset\subset \Omega\backslash\mathcal{R}_T$, we have*

$$\|u(T) - u_h(T)\|_{-(k+1)} \leq Ch^k, \tag{4.4}$$

$$\|u(T) - u_h(T)\|_{-(k+2)} \leq Ch^{k+1/2}, \tag{4.5}$$

$$\|u(T) - u_h(T)\|_{-(k+1),\Omega_1} \leq Ch^{2k+1}, \tag{4.6}$$

$$\|u(T) - K_h^{2k+2,k+1} * u_h(T)\|_{\Omega_0} \leq Ch^{2k+1}, \tag{4.7}$$

where the positive constant C does not depend on h. Here the mesh is assumed to be uniform for (4.7) *but can be regular and non-uniform for the other three inequalities.*

Remark 4.2.1. *To obtain* (4.7)*, we have to assume the mesh is uniformly distributed, that is $h_j = h$, $\forall\ j$. This is a result of the negative-order norm estimates* *of the divided differences. Actually, we denote* $w = \partial_h u$ *and* $w_h = \partial_h u_h$. *Clearly,* w *satisfies* (4.1) *with initial condition* $w(x,0) = \partial_h u(x,0)$. *If we shift the mesh by* $\frac{h}{2}$, *then* w_h *satisfies numerical scheme* (2.5). *By the same analysis for the proof of* (4.6), *we obtain*

$$\|\partial_h(u - u_h)\|_{-(k+1),\Omega_1} = \|w - w_h\|_{-(k+1),\Omega_1} \le Ch^{2k+1}.$$

The estimates for higher order divided differences can be obtained by exactly the same line in this remark. Therefore, (4.7) *follows directly from Corollary 4.1.1.*

Remark 4.2.2. *The error estimates in the* $-(k+1)$*-th order norm are used for problems with singular initial conditions while the estimates in the* $-(k+2)$*-th order norm are used for problems with singular source terms.*

4.2.2 A proof of Theorem 4.2.1

In this subsection, we give the discretization and prove the first three estimates in Theorem 4.2.1.

Initial discretization

From now on, we assume the δ-singularity of the initial datum is contained in cell I_i. For simplicity, we also assume the singularity is concentrated at 0, denoted as $\delta(x)$. We apply the L^2-projection $\mathbb{P}_k$ to discretize the initial condition to obtain $\|u_h(0)\| \le Ch^{-1/2}$. At $t = 0$, for any function $\phi \in C_0^\infty(\Omega)$, we have, for the cell I_i which contains the δ-singularity,

$$\begin{aligned}
(u - u_h, \phi)_i &= (u - u_h, \phi - \mathbb{P}_k\phi)_i \\
&= (u, \phi - \mathbb{P}_k\phi)_i \\
&\le \|\phi - \mathbb{P}_k\phi\|_{\infty,I_i} \\
&\le Ch^{k+\frac{1}{2}}|\phi|_{k+1,I_i}.
\end{aligned}$$

In other cells, following the same analysis as above, we have

$$(u-u_h,\phi)_j = (u-\mathbb{P}_k u, \phi - \mathbb{P}_k\phi)_j \le Ch^{2k+2}|u_0|_{k+1,I_j}|\phi|_{k+1,I_j}. \tag{4.8}$$

The $-(k+1)$-th order error estimate on Ω

In this subsection, we proceed to prove (4.4). The proof mostly follows [34]. We begin by considering the solution to the dual problem: Find a function ϕ such that $\phi(\cdot,t)$ satisfies

$$\begin{aligned} &\phi_t + \phi_x = 0, && (x,t)\in\Omega\times(0,T),\\ &\phi(x,T) = \Phi(x), && x\in\Omega. \end{aligned} \tag{4.9}$$

Assuming Φ is an arbitrary function in $C_0^\infty(\Omega)$, we have, following [34],

$$\begin{aligned} (u(T)-u_h(T),\Phi) &= (u-\mathbb{P}_k u,\phi)(0) - \int_0^T [((u_h)_t,\phi)+(u_h,\phi_t)]dt && (4.10)\\ &= (u-\mathbb{P}_k u,\phi)(0) - \int_0^T \sum_{j=1}^N [u_h](\phi-\mathbb{P}_k\phi)^+|_{j-\frac{1}{2}} && (4.11)\\ &\le Ch^{k+1/2}|\Phi|_{k+1} + Ch^{k+1/2}|\Phi|_{k+1}\int_0^T \left(\sum_{j=1}^N [u_h]_{j-\frac{1}{2}}^2\right)^{1/2} dt. \end{aligned}$$

Using the Cauchy-Schwartz inequality and Lemma 2.4.1, we have

$$\begin{aligned} \int_0^T \left(\sum_{j=1}^N [u_h]_{j-\frac{1}{2}}^2\right)^{1/2} dt &\le T^{1/2}\left(\int_0^T \sum_{j=1}^N [u_h]_{j-\frac{1}{2}}^2 dt\right)^{1/2}\\ &\le T^{1/2}\|u_h(0)\|\\ &\le CT^{1/2}h^{-1/2}. \end{aligned}$$

Combining the above, we can see

$$\begin{aligned}\|u(T)-u_h(T)\|_{-(k+1)} &= \sup_{\Phi\in C_0^\infty(\Omega)} \frac{(u(T)-u_h(T),\Phi)}{\|\Phi\|_{k+1}}\\ &\le \sup_{\Phi\in C_0^\infty(\Omega)} \frac{Ch^{k+1/2}|\Phi|_{k+1}+CT^{1/2}h^k|\Phi|_{k+1}}{\|\Phi\|_{k+1}}\\ &\le CT^{1/2}h^k + Ch^{k+1/2}.\end{aligned}$$

Now, we consider the extension to higher dimensions. The proof of the following corollary is straightforward and is similar to the one-dimensional case, and is thus omitted.

Corollary 4.2.1. *Let Ω be an open set in $\mathbb{R}^d$, and u be the exact solution of the following initial value problem*

$$\begin{aligned}&u_t+\textstyle\sum_{j=1}^d u_{x_j}=0, &&(x,t)\in\Omega\times(0,T],\\ &u(x,0)=\delta(f(x)), &&x\in\Omega,\end{aligned}$$

where $f(x):\mathbb{R}^d\to\mathbb{R}$ is a smooth function. Denote $\Gamma_h=\{K\}$ as a regular triangulation of $\mathbb{R}^d$, whose elements K are open and have diameter h_K less than or equal to h. In each K, denote ∂K_- and ∂K_+ as the inflow and outflow edges respectively. Let u_h be the DG approximation which satisfies

$$(u_{ht},v_h)_K=\sum_{i=1}^d(u_h,(v_h)_{x_i})_K+\sum_{i=1}^d(u_h^-,v_h^+)_{\partial K_-}-\sum_{i=1}^d(u_h^-,v_h^-)_{\partial K_+},\quad v_h\in V_h,$$

where the finite element space V_h is made up of the piecewise polynomials of degree $k\ge 1$. Suppose the total measure of the cells which contain δ-singularities initially is mh, then there holds the following estimate

$$\|u(T)-u_h(T)\|_{-k-1}\le C\sqrt{m}T^{1/2}h^k+Ch^{k+d/2}, \tag{4.12}$$

where the bounded constant $C > 0$ does not depend on h or T.

The $-(k+2)$-th order error estimate on Ω

We shall prove (4.5). To do so, we apply $\mathbb{P}_+$ to estimate the term $(u_{ht}, \phi) + (u_h, \phi_t)$. By using (2.4) and Lemma 2.4.3, we obtain

$$\begin{aligned}(u_{ht}, \phi) + (u_h, \phi_t) &= (u_{ht}, \mathbb{P}_+^\perp \phi) + (u_{ht}, \mathbb{P}_+ \phi) - (u_h, \phi_x) \\ &= (u_{ht}, \mathbb{P}_+^\perp \phi) + \mathcal{H}(u_h, \mathbb{P}_+ \phi) - \mathcal{H}(u_h, \phi) \\ &= (u_{ht}, \mathbb{P}_+^\perp \phi). \qquad (4.13)\end{aligned}$$

Integrating in t, we obtain

$$\int_0^T (u_{ht}, \phi) + (u_h, \phi_t) dt = (u_h, \mathbb{P}_+^\perp \phi)(T) - (u_h, \mathbb{P}_+^\perp \phi)(0) - \int_0^T (u_h, \mathbb{P}_+^\perp \phi_t) dt. \qquad (4.14)$$

Applying Lemma 2.4.1, we have

$$\begin{aligned}\int_0^T (u_{ht}, \phi) + (u_h, \phi_t) dt &\le \|u_h(0)\| \, \left(\|(\mathbb{P}_+^\perp \phi)(0)\| + \|(\mathbb{P}_+^\perp \phi)(T)\|\right) + \int_0^T \|u_h(0)\| \, \|\mathbb{P}_+^\perp \phi_t(t)\| dt \\ &\le C\|u_h(0)\| \left(h^{k+1}|\Phi|_{k+1} + \int_0^T h^{k+1}|\Phi|_{k+2} dt\right) \\ &\le C(1+T)h^{k+1}\|u_h(0)\|\|\Phi\|_{k+2}.\end{aligned}$$

From the above we observe

$$\begin{aligned}\|u(T) - u_h(T)\|_{-(k+2)} &= \sup_{\Phi \in C_0^\infty(\Omega)} \frac{(u(T) - u_h(T), \Phi)}{\|\Phi\|_{k+2}} \\ &\le \sup_{\Phi \in C_0^\infty(\Omega)} \frac{Ch^{k+\frac{1}{2}}|\Phi|_{k+1} + C(1+T)h^{k+\frac{1}{2}}\|\Phi\|_{k+2}}{\|\Phi\|_{k+2}} \\ &\le C(1+T)h^{k+\frac{1}{2}}.\end{aligned}$$

4.2.3 The negative-order error estimate on $\Omega_1 \subset\subset \Omega\backslash\mathcal{R}_T$

We proceed to prove (4.6). To estimate the negative-order norm of $u - u_h$ at time T on Ω_1, we need to assume $\Phi \in C_0^\infty(\Omega_1)$ instead of $C_0^\infty(\Omega)$. Moreover, we also assume the exact solution $u \in C(\Omega)$, this is because we are allowed to modify the exact solution in the cell which contains the δ-singularity, keeping the numerical solution u_h untouched. More details of this assumption can be found in [117, 31] or Section 4.5.2. Therefore, in (4.11), we have

$$\begin{aligned}
\sum_{j=1}^{N}[u_h](\phi - \mathbb{P}_k\phi)^+|_{j-\frac{1}{2}} &= \sum_{j=1}^{N}[u_h - \mathbb{P}_k u + \mathbb{P}_k u - u](\phi - \mathbb{P}_k\phi)^+|_{j-\frac{1}{2}} \\
&\leq Ch^k \left(\|u - \mathbb{P}_k u\|_{\Omega_1} + \|\mathbb{P}_k u - u_h\|_{\Omega_1}\right)|\phi|_{k+1} \\
&\leq Ch^k \left(\|u - \mathbb{P}_k u\|_{\Omega_1} + \|u - u_h\|_{\Omega_1}\right)|\phi|_{k+1} \\
&\leq Ch^{2k+1}|\phi|_{k+1},
\end{aligned}$$

where we use Lemmas 2.3.1 and 2.3.2 in the second inequality and Lemma 4.2.1 in the last one. Inserting this into (4.11), we have

$$(u(T) - u_h(T), \Phi) \leq (u - \mathbb{P}_k u, \phi)(0) + Ch^{2k+1}|\phi|_{k+1}. \tag{4.15}$$

By using (4.8), we obtain the estimate we want

$$\begin{aligned}
\|u(T) - u_h(T)\|_{-(k+1),\Omega_1} &= \sup_{\Phi \in C_0^\infty(\Omega_1)} \frac{(u(T) - u_h(T), \Phi)}{\|\Phi\|_{k+1}} \\
&\leq \sup_{\Phi \in C_0^\infty(\Omega_1)} \frac{Ch^{2k+1}\|\Phi\|_{k+1} + Ch^{2k+2}\|\Phi\|_{k+1}}{\|\Phi\|_{k+1}} \\
&\leq Ch^{2k+1},
\end{aligned}$$

where the constant $C > 0$ is independent of h.

4.3 Singular source term

In this section, we briefly discuss a linear inhomogeneous evolution equation of a function

$$u(x,t) : \Omega \times (0,\infty) \to R$$

of the form

$$\begin{cases} u_t(x,t) + Lu(x,t) = f(x,t), & (x,t) \in \Omega \times (0,\infty), \\ u(x,0) = 0, & x \in \Omega, \end{cases} \tag{4.16}$$

with L being a linear differential operator that does not involve time derivatives. If we multiply the above equation by a smooth function $\phi(x,t)$, then integrate over space and time, we obtain

$$\int_0^\infty \int_\Omega [\phi u_t + \phi Lu]\, dxdt = \int_0^\infty \int_\Omega f(x,t)\phi(x,t) dxdt.$$

Integrating by parts and assuming zero boundary condition, we have

$$\int_0^\infty \int_\Omega [u\phi_t + uL^*\phi]\, dxdt + \int_0^\infty \int_\Omega f(x,t)\phi(x,t) dxdt = 0, \tag{4.17}$$

where L^* is the dual operator of L.

Definition 4.3.1. *The function $u(x,t)$ is called a weak solution of the equation* (4.16)*, if* (4.17) *holds for all functions $\phi \in C_0^1(\Omega \times R^+)$.*

4.3.1 Duhamel's principle

Now, we consider linear hyperbolic conservation laws with source terms. To deal with such problems we apply Duhamel's principle, which is applicable to linear parabolic and hyperbolic PDE and yields an integral representation in terms of the solutions of more tractable PDEs.

Lemma 4.3.1 (Duhamel's principle). *The solution to* (4.16) *is*

$$u(x,t) = \int_0^t (P^s f)(x,t)ds,$$

where $P^s f$ is the solution of the problem

$$\begin{cases} P_t(x,t) + LP(x,t) = 0, \ (x,t) \in \Omega \times (s,\infty), \\ P(x,s) = f(x,s), \qquad\qquad x \in \Omega. \end{cases} \tag{4.18}$$

Notice that $P^s f$ is the solution to the homogeneous PDE with the source term f serving as the initial condition at time $t = s$. To prove the lemma, we can simply check that the expression of u satisfies (4.16). More details and a proof can be found in [62], in which the PDE is a second order wave equation. The above lemma requires suitable regularity of u. However, Duhamel's principle is also valid in the following weak sense.

Lemma 4.3.2. *Suppose $u(x,t)$ is the weak solution of equation* (4.16)*, then*

$$u(x,t) = \int_0^t (P^s f)(x,t)ds$$

in the sense of distribution, where $P^s f$ is the weak solution of equation (4.18)*.*

The proof directly follows from the definition of the weak solution and the proof of Duhamel's principle, so we omit it here.

Finally, we extend Duhamel's principle to the DG schemes. For simplicity, we consider the following equation

$$\begin{cases} u_t(x,t) + u_x(x,t) = \delta(x), \ (x,t) \in \Omega \times (0,T), \\ u(x,0) = u_0(x), \qquad\qquad x \in \Omega, \end{cases} \tag{4.19}$$

with $u_0 = 0$. For general smooth $u_0(x)$, the same result can be obtained by superposition. We define the finite element approximation $u_h : [0,T] \to V_h$ as the solution to

$$\begin{aligned}&(u_{ht}, \chi)_j = \mathcal{H}_j(u_h, \chi) + (\delta(x), \chi)_j, \ \forall \chi \in V_h,\\ &u_h(0) = 0,\end{aligned} \tag{4.20}$$

where $\mathcal{H}_j(\cdot, \cdot)$ is the DG bilinear form defined in (2.4). Then the semi-discrete version of Duhamel's principle is given in Lemma 4.3.3.

Lemma 4.3.3. *The solution of* (4.20) *can be written in the form* $u_h = \int_0^t p^s(x,t)ds$ *where* $p^s(x,t)$ *is the solution of the following scheme: find* $p \in V_h$ *such that*

$$\begin{aligned}&(p_t, \chi)_j = \mathcal{H}_j(p, \chi), \ \forall \chi \in V_h,\\ &p(s) = \mathbb{P}_k \delta(x).\end{aligned} \tag{4.21}$$

The proof is straightforward, since u_h in (4.20) and $\int_0^t p^s(x,t)ds$ share the same initial condition and the same system of ODEs, noticing the fact that $(\mathbb{P}_k\delta, \chi) = (\delta, \chi)$.

In what follows, we would like to rewrite the inhomogeneous equations (4.19) and (4.20) into homogeneous ones (4.26) and (4.21) by using Lemma 4.3.2 and Lemma 4.3.3 respectively. Then we apply the estimates of $P^s f - p^s$, which have been given in Theorem 4.2.1, to prove the main result, Theorem 4.3.1.

4.3.2 Error estimates

We first state the main result Theorem 4.3.1 and then give the proof.

Theorem 4.3.1. *Suppose* u *is the exact solution of equation* (4.19)*, and* u_h *is the numerical solution which satisfies* (4.20)*. Denote* $\mathcal{R}_T = I_i \cup (T - C\log(1/h)h^{1/2}, T + C\log(1/h)h^{1/2})$*, where* I_i *is the cell which contains the concentration of the* δ*-singularity* *on the source term. Then we have the following estimates*

$$\|u(T) - u_h(T)\|_{-(k+1)} \leq Ch^k, \tag{4.22}$$

$$\|u(T) - u_h(T)\|_{-(k+2)} \leq Ch^{k+1/2}, \tag{4.23}$$

$$\|u - u_h\|_{-(k+1),\Omega_1} \leq Ch^{2k+1}, \tag{4.24}$$

$$\|u(T) - K_h^{2k+2,k+1} * u_h(T)\|_{\Omega_0} \leq Ch^{2k+1}, \tag{4.25}$$

where $\Omega_0 + 2supp(K_h^{2k+2,k+1}) \subset\subset \Omega_1 \subset\subset R\backslash\mathcal{R}_T$. Here the mesh is assumed to be uniform for (4.25) but can be regular and non-uniform for the other three inequalities.

Remark 4.3.1. *As mentioned in Remark 4.2.1,* (4.25)*, which requires uniform meshes, follows from* (4.24)*. Moreover, we also skip the proofs of* (4.22) *and* (4.23)*, since they follow easily from* (4.4) *and* (4.5) *in Theorem 4.2.1.*

Now we proceed to prove (4.24). Denote v^s as the exact solution of the following equation

$$\begin{aligned} &u_t + u_x = 0, && (x,t) \in \Omega \times (s,T], \\ &u(x,s) = \delta(x), && x \in \Omega, \end{aligned} \tag{4.26}$$

and v_h^s as the solution of the numerical scheme (4.21). For convenience, if $s = 0$, the superscript will be omitted. We consider the dual problem defined the same way as (4.9). By Lemma 4.3.3 and Lemma 4.3.2, we have

$$(u - u_h, \Phi)(T) = \int_0^T (v^s - v_h^s, \Phi)(T)ds. \tag{4.27}$$

By using (4.10) and (4.13), and the fact that v_h is the L^2-projection of v at $t = 0$, we obtain

$$(v^s - v_h^s, \Phi)(T) = ((v - v_h)(0), \mathbb{P}_+^{\perp}\phi(s)) - \int_s^T (v_{ht}(t-s), \mathbb{P}_+^{\perp}\phi(t))dt,$$

which further yields

$$(u-u_h,\Phi)(T)=\Pi_1-\Pi_2,$$

where $\Pi_1=\int_0^T((v-v_h)(0),\mathbb{P}_+^\perp\phi(s))ds$, and $\Pi_2=\int_0^T\int_s^T(v_{ht}(t-s),\mathbb{P}_+^\perp\phi(t))dtds$. First consider the second term,

$$\begin{aligned}\Pi_2&=-\int_0^T\int_0^t((v_h)_s(t-s),\mathbb{P}_+^\perp\phi(t))dsdt\\&=\int_0^T(v_h(t)-v_h(0),\mathbb{P}_+^\perp\phi(t))dt.\end{aligned}$$

Therefore,

$$\begin{aligned}(u-u_h,\Phi)(T)&=\int_0^T(v(0),\mathbb{P}_+^\perp\phi(s))ds-\int_0^T(v_h(t),\mathbb{P}_+^\perp\phi(t))dt\\&=G_1-G_2.\end{aligned}$$

We claim $G_1=0$. Actually, for any $\tau\in I_i$, $\int_\tau^{T+\tau}\Phi(x)dx$ does not depend on τ, since $\Phi(x)$ vanishes in the neighborhood of $x=0$ and $x=T$. Therefore,

$$G_1=\left(\delta(x),\mathbb{P}_+^\perp\int_0^T\phi(x,s)ds\right)=\left(\delta(x),\mathbb{P}_+^\perp\int_x^{x+T}\Phi(y)dy\right)_i=0.$$

Now, we only need to estimate G_2. Since

$$(v_h(t),\mathbb{P}_+^\perp\phi(t))=(v_h(t)-v(t)+v(t)-\mathbb{P}_{k-1}v,\mathbb{P}_+^\perp\phi(t)),$$

by Lemma 4.2.1 and Lemma 2.3.2, we have $G_2\le Ch^{2k+1}|\phi|_{k+1}$. Finally, we obtain

$$\|u-u_h\|_{-(k+1),\Omega_1}\le Ch^{2k+1}$$

and complete the proof of Theorem 4.3.1.

4.4 Numerical examples

In this section, we provide numerical experiments to demonstrate our theoretical results for the post-processor and to illustrate the performance of the DG schemes. We denote by d the distance between the singularities and the region under consideration. In all the figures, if not otherwise stated, the numerical solutions are plotted using six Gaussian points in each cell.

4.4.1 Singular initial condition

Example 4.1. *We solve the following problem*

$$\begin{aligned} &u_t + u_x = 0, && (x,t) \in [0,\pi] \times (0,1], \\ &u(x,0) = \sin(2x) + \delta(x-0.5), && x \in [0,\pi], \end{aligned} \tag{4.28}$$

with periodic boundary condition $u(0,t) = u(\pi,t)$.

Clearly, the exact solution is

$$u(x,t) = \sin(2x-2t) + \delta(x-t-0.5).$$

We use a ninth order SSP Runge-Kutta discretization in time [46] and take the time step $\Delta t = 0.1h$. We test the example by using $\mathcal{P}^k$ polynomials with $k = 1, 2, 3$ on uniform meshes, and compute the L^2-norm of the error after post-precessing in the region away from the singularity at $t = 0.5$. By taking $d = 0.2$, the region under consideration is $[0, 0.8] \cup [1.2, \pi]$. In Table 4.1, we can observe at least $(2k+1)$-th order convergence. Moreover, we observe that the rate of convergence settles to the asymptotic value when the total number of cells is around $\frac{dN}{\pi} = \frac{0.2\times 500}{\pi} \approx 30$, no matter which degree of polynomials we use. The initial discretization is obtained by taking the L^2 projection.

Figure 4.1 shows the numerical solution with and without post-processing. We

Table 4.1: L^2-norm of the error between the numerical solution and the exact solution for (4.28) after post-processing in the region away from the singularity.

		$\mathcal{P}^1$ polynomial		$\mathcal{P}^2$ polynomial		$\mathcal{P}^3$ polynomial	
N	d	error	order	error	order	error	order
200	0.2	6.88E-05	-	8.40e-07	-	1.48E-09	-
300	0.2	1.41E-05	3.92	3.56e-10	19.2	3.98E-13	20.3
400	0.2	5.89E-06	3.02	1.98e-11	10.1	4.42E-16	23.7
500	0.2	3.01E-06	3.01	6.13e-12	5.25	7.49E-17	7.95
600	0.2	1.74E-06	3.00	2.37e-12	5.21	1.76E-17	7.94

use $\mathcal{P}^2$ polynomials and take $h = 0.01$. From the figure we observe some localized oscillations near the discontinuity and that the post-processor does not significantly smear the singularity.

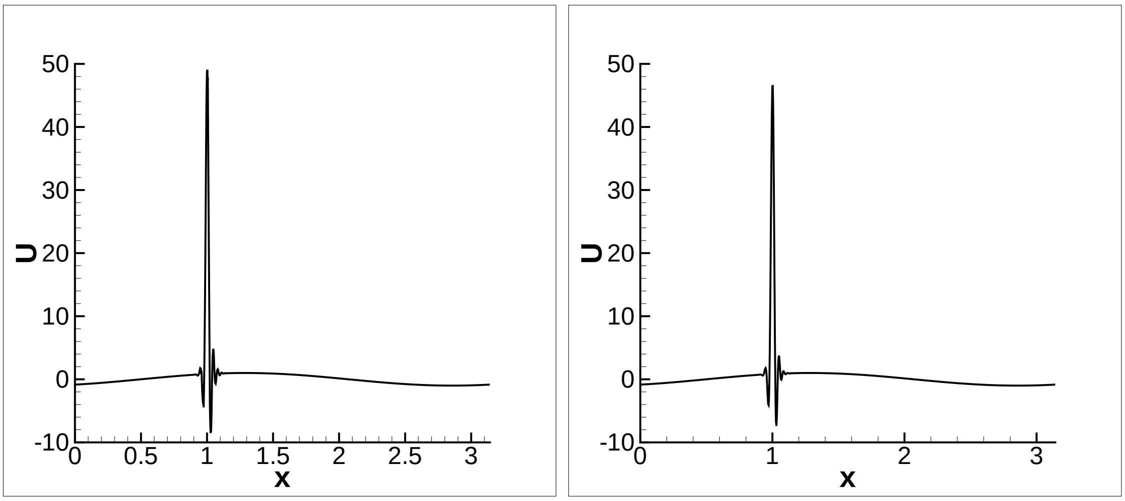

Figure 4.1: Numerical solution for (4.28) at $t = 0.5$ with (right) and without (left) post-processing.

Example 4.2. *We consider the following two dimensional problem*

$$\begin{aligned} &u_t + u_x + u_y = 0, && (x,y,t) \in [0,2\pi] \times [0,2\pi] \times (0,1], \\ &u(x,0) = \sin(x+y) + \delta(x+y-2\pi), && (x,y) \in [0,2\pi] \times [0,2\pi], \end{aligned} \tag{4.29}$$

with periodic boundary condition.

Clearly, the exact solution is

$$u(x, t) = \sin(x + y - 2t) + \delta(x + y - 2t) + \delta(x + y - 2t - 2\pi).$$

We use $\mathcal{Q}^k$ polynomial approximation spaces with $k = 1$ and 2, where $\mathcal{Q}^k$ is the space of tensor product polynomials of degree at most $k \geq 0$. We also apply the same time discretization as in Example 4.1 and compute the L^2-norm of the error after post-precessing in the region away from the singularity at $t = 0.5$. Moreover, we take $d = 0.4$. In Table 4.2, we can observe $(2k + 1)$-th order convergence.

Table 4.2: L^2-norm of the error between the numerical solution and the exact solution for (4.29) after post-processing in the region away from the singularity.

		$\mathcal{Q}^1$ polynomial		$\mathcal{Q}^2$ polynomial	
N	d	error	order	error	order
400	0.4	2.60E-05	-	3.23e-08	-
500	0.4	1.24E-05	3.32	2.47e-10	20.0
600	0.4	7.16E-06	3.01	1.19e-11	16.6
700	0.4	4.50E-06	3.01	5.11e-12	5.47
800	0.4	3.01E-06	3.02	2.53e-12	5.29

It appears that similar results are valid in two dimensions. However, the technique of proof in this chapter, in particular the part related to the special projections in (2.10) and (2.11), does not seem to be easily extendable to two dimensions.

Moreover, Figure 4.2 shows the numerical solution by plotting the numerical cell averages. We use $\mathcal{Q}^2$ polynomials and take $N = 100$. From the figure we can observe two lines of δ-singularities.

Even though the theory in this chapter is given only for scalar linear equations for simplicity, it generalizes to linear systems in a straightforward way.

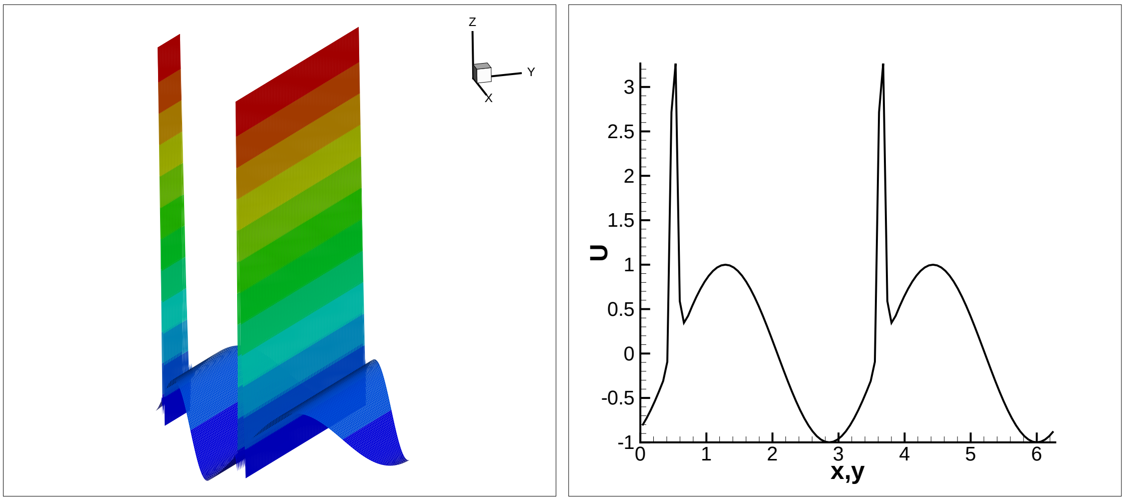

Figure 4.2: Numerical solution (left) and the cut plot along $x = y$ (right) for (4.29) at $t = 0.5$.

Example 4.3. *We solve the following linear system*

$$\begin{aligned} & u_t - v_x = 0, && (x,t) \in [0,2] \times (0,0.4], \\ & v_t - u_x = 0, && (x,t) \in [0,2] \times (0,0.4], \\ & u(x,0) = \delta(x-1), v(x,0) = 0, && x \in [0,2]. \end{aligned} \tag{4.30}$$

Clearly, the exact solution (the Green's function) is

$$u(x,t) = \frac{1}{2}\delta(x-1-t) + \frac{1}{2}\delta(x-1+t), \quad v(x,t) = \frac{1}{2}\delta(x-1+t) - \frac{1}{2}\delta(x-1-t).$$

We use a third order SSP Runge-Kutta discretization in time [46] and take the time step $\Delta t = 0.1h$. Figure 4.3 shows the numerical solutions at $t = 0.4$ with $\mathcal{P}^3$ polynomials and $h = 0.01$. We observe that the numerical solutions capture the profiles of the exact solutions quite well. Since we have not used any limiter, there are some localized oscillations near the singularities.

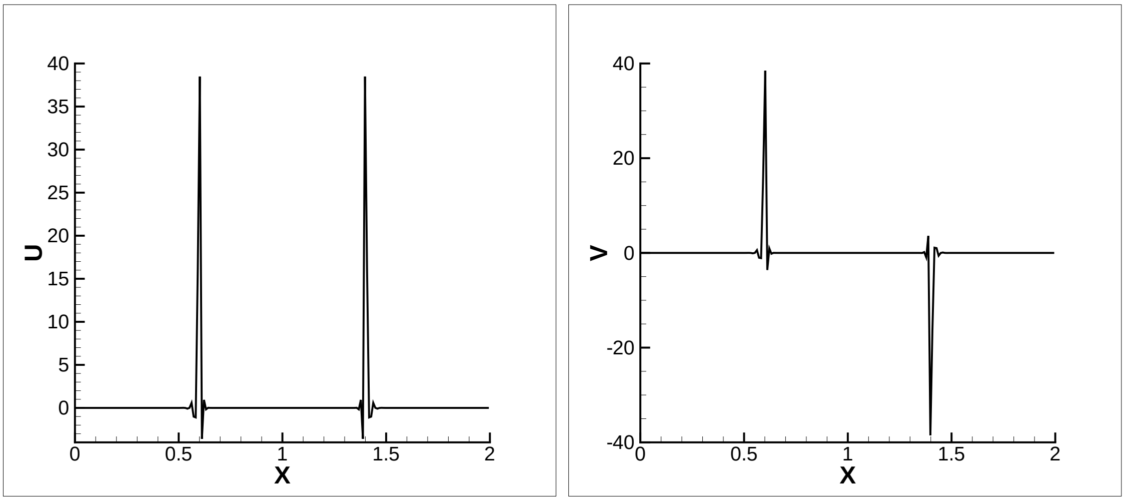

Figure 4.3: Solutions of u (left) and v (right) for (4.30) at $t = 0.4$.

4.4.2 Singular source term

Example 4.4. *We solve the following problem*

$$\begin{aligned} &u_t + u_x = \delta(x-\pi), && (x,t) \in [0,2\pi] \times (0,1], \\ &u(x,0) = \sin(x), && x \in [0,2\pi], \\ &u(0,t) = 0, && t \in (0,1]. \end{aligned} \tag{4.31}$$

Clearly, the exact solution is

$$u(x,t) = \sin(x-t) + \chi_{[\pi,\pi+t]},$$

where $\chi_{[a,b]}$ denotes the indicator function of the interval $[a,b]$. We use the same time discretization as in the previous example, and use both $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials to approximate the exact solution on uniform meshes. We compute the L^2-norm of the error after post-precessing in the region away from the singularities at $t = 0.5$. In this example, we also take $d = 0.2$, and the region under consideration is $[0,\pi-0.2]\cup[\pi+0.2,\pi+0.3]\cup[\pi+0.7,2\pi]$. In Table 4.3, we observe $(2k+1)$-th order convergence. The initial discretization is again obtained by taking the L^2 projection.

Table 4.3: L^2-norm of the error between the numerical solution and the exact solution for (4.31) after post-processing in the region away from the singularity.

		$\mathcal{P}^1$ polynomial		$\mathcal{P}^2$ polynomial	
N	d	error	order	error	order
401	0.2	1.74E-06	-	4.29E-08	-
801	0.2	5.92E-09	8.22	6.80E-13	15.9
1601	0.2	7.36E-10	3.03	1.34E-17	12.3
3201	0.2	9.19E-11	3.01	3.86E-18	5.13
6401	0.2	1.15E-11	3.01	1.16E-19	5.07

Moreover, Figure 4.4 shows the numerical solutions with and without post-processing. We use $\mathcal{P}^2$ polynomials and take $h = 0.01$. We observe that the post-processor does not smear the singularity and that it effectively damps out the oscillations near the left singularity.

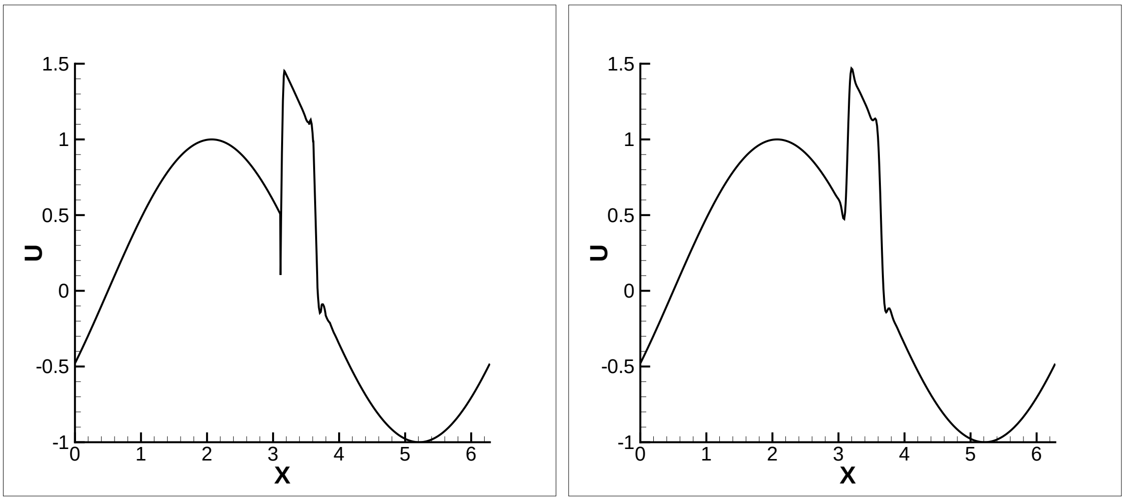

Figure 4.4: Numerical solutions for (4.31) at $t = 0.5$ with (right) and without (left) post-processing.

Example 4.5. *We solve the following problem*

$$\begin{aligned} &u_t + ((x+1)u)_x = \delta(x-c), && (x,t) \in [0,1.5] \times (0,1], \\ &u(x,0) = 0, && x \in [0,1.5], \\ &u(0,t) = 0, && t \in (0,1]. \end{aligned} \tag{4.32}$$

The exact solution is

$$u(x,t) = \frac{1}{1+x}[H(x-c) - H(x+1-(c+1)e^t)],$$

where $H(x)$ is the Heaviside function defined as

$$H(x) = \begin{cases} 0, \; x < 0, \\ 1, \; x \geq 0. \end{cases}$$

We take $c = \frac{\pi}{20}$ and compute the solution at $t = 0.5$ with $\mathcal{P}^1$ and $\mathcal{P}^2$ polynomials. Since $(x+1)$ is always positive, by using upwind fluxes, we always consider u_h^- to be the numerical flux at the cell interfaces. For time discretization, the classical fourth order Runge-Kutta method is used with $\Delta t = h^2$. In this example, we take $d = 0.1$, and the region under consideration is $[0, \frac{\pi}{20} - 0.1] \cup [\frac{\pi}{20} + 0.1, (\frac{\pi}{20} + 1)\sqrt{e} - 1.1] \cup [(\frac{\pi}{20} + 1)\sqrt{e} - 0.9, 1.5]$. In Table 4.4, we can observe $(k+1)$-th and $(2k+1)$-th order convergence before and after post-processing respectively.

Moreover, Figure 4.5 shows the numerical solution with $\mathcal{P}^2$ polynomials and $h = 0.01$. We use the cell averages to plot the left panel of the figure. We observe that the numerical solution agrees well with the exact solution away from the singularities. Since we have not used any limiter, there are some localized oscillations near the singularity on the right. It is interesting to observe that there are very few numerical oscillations near the left singularity. In the middle panel of Figure 4.5, we use six Gaussian points to plot, and the detailed zoom for the left singularity is given in the right panel. Clearly, the numerical solution only oscillates in the cell [0.15,0.16]. No

Table 4.4: L^2-norm of the error between the numerical solution with and the exact solution for (4.32) before and after post-processing in the region away from the singularity.

			$\mathcal{P}^1$ polynomial		$\mathcal{P}^2$ polynomial	
	N	d	error	order	error	order
before post-processing	400	0.1	7.08E-07	-	1.10E-08	-
	800	0.1	1.21E-07	2.56	4.37E-11	7.98
	1600	0.1	3.02E-08	2.00	5.46E-12	3.00
	3200	0.1	7.55E-09	2.00	6.83E-13	3.00
	6400	0.1	1.89E-09	2.00	8.53E-14	3.00
after post-processing	400	0.1	4.92E-07	-	8.65E-09	-
	800	0.1	7.49E-11	12.7	4.54E-14	17.5
	1600	0.1	7.43E-12	3.33	5.13E-19	16.4
	3200	0.1	9.31E-13	3.00	1.76E-20	4.86
	6400	0.1	1.16E-13	3.00	5.75E-22	4.94

oscillation is observed in the left figure for cell averages, and only one undershoot can be observed in the middle and right panels for which six Gaussian points are plotted. This can be explained by the size of the pollution region. In Theorem 4.3.1, we have proved that, for such singularities, $\mathcal{R}_T$ contains only one cell. This implies that the numerical solution will oscillate within that cell, which clearly agrees with our observation.

4.5 Proof of Lemma 4.2.1

In this section we prove Lemma 4.2.1. The main line of proof is based on ideas in [31, 117]. For simplicity, we only consider a δ-singularity in (4.1), hence $u_0(x) = \delta(x) + f(x)$, where $f(x)$ is sufficiently smooth and has compact support on the computational domain Ω.

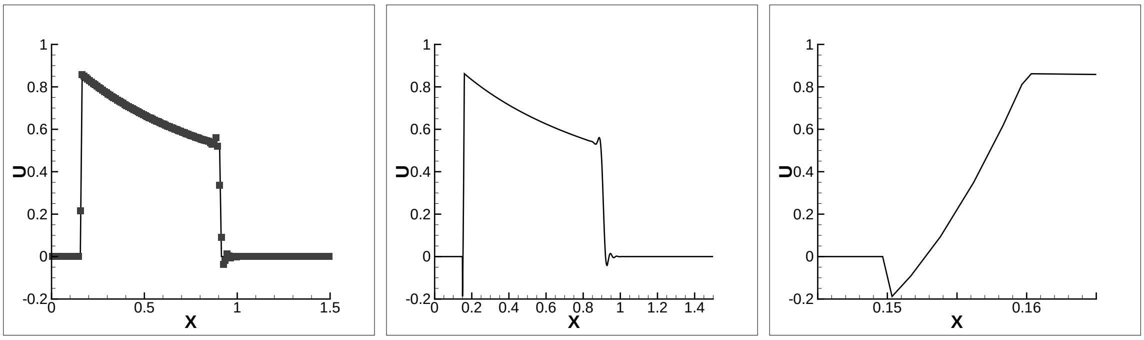

Figure 4.5: Numerical solutions for (4.32) at $t = 0.5$ plotted for the cell averages (left), six Gaussian points (middle) and the detailed zoom (right). In the left panel, the solid line is the exact solution and the symbols are the cell averages of the numerical solution.

4.5.1 The weight function

Let $\varphi(x)$ be a positive bounded function, which can be taken as a weight function. For any function $q \in H_h^1$, we define the weighted L^2-norm as

$$\|q\|_{\varphi,D} = \left(\int_D q^2 \varphi dx \right)^{\frac{1}{2}}$$

in the domain D. If $\varphi = 1$ or $D = \Omega$, the corresponding subscript will be omitted.

We will consider two weight functions $\varphi^1(x,t)$ and $\varphi^{-1}(x,t)$, respectively, to determine the left-hand and right-hand boundary of the region $\mathcal{R}_T$ such that, outside this region, we can resume the $(k+1)$-th order accuracy in the L^2-norm. Both weight functions are related to the cut-off of the exponent function $\phi(r) \in C^1 : \Omega \to \mathbb{R}$,

$$\phi(r) = \begin{cases} 2 - e^r, \; r < 0, \\ e^{-r}, \qquad r > 0, \end{cases}$$

and they are defined as the solutions of the linear hyperbolic problem,

$$\varphi_t^a + \varphi_x^a = 0, \tag{4.33}$$

$$\varphi^a(x,0) = \phi\left(\frac{a(x-x_c)}{\gamma h^\sigma}\right), \tag{4.34}$$

where $\gamma > 0$, $0 < \sigma < 1$ and x_c are three parameters which will be chosen later. We always assume $\gamma h^{\sigma-1} \geq 1$ in this section.

In [117], the authors have listed several properties about the two weight functions. Here, we state those that will be used.

Proposition 4.5.1. *For each weight function $\varphi^a(x,t)$, the following properties hold*

$$1 \leq \varphi^a(x,t) \leq 2, \;\; a(x - x_c - t) \leq 0, \tag{4.35}$$

$$0 < \varphi^a(x,t) < h^s, \; a(x - x_c - t) > s\,\log(1/h)\gamma h^\sigma. \tag{4.36}$$

Lemma 4.5.1. *Let $\mathbb{V}$ be a Gauss-Radau projection, either $\mathbb{P}_-$ or $\mathbb{P}_+$. For any sufficiently smooth function $p(x)$, there exists a positive constant C independent of h and p, such that*

$$\|\mathbb{V}^\perp p\|_{\varphi,D} \leq Ch^{k+1}\|\partial_x^{k+1}p\|_{\varphi,D}, \tag{4.37}$$

$$\|\mathbb{V}^\perp(\varphi v_h)\|_{\varphi^{-1},D} \leq C\gamma^{-1}h^{1-\sigma}\|v_h\|_{\varphi,D}, \tag{4.38}$$

$$\|\mathbb{V}(\varphi v_h)\|_{\varphi^{-1},D} \leq C\|v_h\|_{\varphi,D}. \tag{4.39}$$

where D is either the single cell I_j or the whole computational domain Ω.

Lemma 4.5.2. *For any function $v \in V_h$ there holds the following identity*

$$\mathcal{H}(v, \varphi v) = -\frac{1}{2}\sum_j \varphi_{j+\frac{1}{2}}[v]^2_{j+\frac{1}{2}} + \frac{1}{2}(v, \varphi_x v). \tag{4.40}$$

4.5.2 The smooth solution

We consider the following problem

$$v_t + v_x = 0, \tag{4.41}$$

$$v(x,0) = v_0(x), \tag{4.42}$$

where the initial condition $v_0(x)$, is a sufficiently smooth function modified from the original initial condition $u_0(x) = \delta(x) + f(x)$ such that it agrees with $u_0(x)$ for all $x \in \Omega \backslash I_i$, and satisfies

$$|\partial_x^\alpha v_0(x)| \leq C h^{-\alpha-1}, \quad x \in I_i,$$

where I_i is the cell containing $x = 0$.

4.5.3 Error representation and error equations

Denote the error by $e = v - u_h$, where u_h approximates the solution to (4.1) or (4.41). Clearly, e also satisfies (2.5). We divide the error into the form $e = \eta - \xi$, where

$$\eta = v - \mathbb{P}_- v = \mathbb{P}_-^\perp v, \quad \text{and} \quad \xi = u_h - \mathbb{P}_- v.$$

Following [117], we obtain

$$\begin{aligned}
\frac{d\|\xi\|_\varphi^2}{dt} &= 2\left(\xi_t, \mathbb{P}_+^\perp(\varphi\xi)\right) + 2\left(\xi_t, \mathbb{P}_+(\varphi\xi)\right) - (\xi, \varphi_x\xi) \\
&= 2\left(\xi_t, \mathbb{P}_+^\perp(\varphi\xi)\right) + 2\left(\eta_t, \mathbb{P}_+(\varphi\xi)\right) - 2\left(e_t, \mathbb{P}_+(\varphi\xi)\right) - (\xi, \varphi_x\xi) \\
&= 2\left(\xi_t, \mathbb{P}_+^\perp(\varphi\xi)\right) + 2\left(\eta_t, \mathbb{P}_+(\varphi\xi)\right) - 2\mathcal{H}(e, \mathbb{P}_+(\varphi\xi)) - (\xi, \varphi_x\xi) \\
&= 2\left(\xi_t, \mathbb{P}_+^\perp(\varphi\xi)\right) + 2\left(\eta_t, \mathbb{P}_+(\varphi\xi)\right) + 2\mathcal{H}(\xi, \mathbb{P}_+(\varphi\xi)) - (\xi, \varphi_x\xi) \\
&= 2\left(\xi_t, \mathbb{P}_+^\perp(\varphi\xi)\right) + 2\left(\eta_t, \mathbb{P}_+(\varphi\xi)\right) + 2\mathcal{H}(\xi, \varphi\xi) - (\xi, \varphi_x\xi) \\
&= 2\Pi_1 + 2\Pi_2 - \Pi_3,
\end{aligned}$$

where

$$\Pi_1 = \left(\xi_t, \mathbb{P}_+^{\perp}(\varphi\xi)\right), \quad \Pi_2 = (\eta_t, \mathbb{P}_+(\varphi\xi)), \quad \Pi_3 = \sum_j \varphi_{j+\frac{1}{2}}[\xi]^2_{j+\frac{1}{2}}.$$

First we estimate Π_1. Denote $w = \xi_t - \mathbb{P}_{k-1}\xi_t$. From the scheme (2.20), we have

$$(\xi_t, w)_j = (\eta_t, w)_j - (e_t, w)_j = (\eta_t, w)_j - [\xi]_{j-\frac{1}{2}} w^+_{j-\frac{1}{2}}.$$

Inserting this into Π_1 and defining $\psi = \sqrt{\varphi}$, we obtain

$$\begin{aligned}
(\xi_t, \mathbb{P}_+^{\perp}(\varphi\xi))_j &= \left(\frac{(\xi_t, w)_j}{\|w\|^2_{I_j}} w, \mathbb{P}_+^{\perp}(\varphi\xi)\right)_j \\
&= \left(\left((\eta_t, w)_j - [\xi]_{j-1/2} w^+_{j-\frac{1}{2}}\right) \frac{w}{\|w\|^2_{I_j}}, \mathbb{P}_+^{\perp}(\varphi\xi)\right)_j \\
&\leq \frac{C}{\|w\|_{I_j}} \left(|(\psi\eta_t, w)_j| + \left|[\psi\xi]_{j-1/2} w^+_{j-\frac{1}{2}}\right|\right) \left\|\psi^{-1}\mathbb{P}_+^{\perp}(\varphi\xi)\right\|_{I_j} \\
&\leq \frac{Ch^{1-\sigma}}{\gamma}\left(\|\eta_t\|^2_{\varphi,I_j} + \|\xi\|^2_{\varphi,I_j}\right) + \frac{Ch^{1/2-\sigma}}{\gamma}\left(\varphi_{j-1/2}[\xi]^2_{j-1/2} + \|\xi\|^2_{\varphi,I_j}\right).
\end{aligned}$$

Summing up with respect to j, we obtain

$$(\xi_t, \mathbb{P}_+^{\perp}(\varphi\xi)) \leq \frac{Ch^{1-\sigma}}{\gamma}\left(\|\eta_t\|^2_{\varphi} + \|\xi\|^2_{\varphi}\right) + \frac{Ch^{1/2-\sigma}}{\gamma}\left(\sum_j \varphi_{j-1/2}[\xi]^2_{j-1/2} + \|\xi\|^2_{\varphi}\right).$$

For Π_2, it is not difficult to see that

$$\Pi_2 \leq C\|\eta_t\|_{\varphi}\ \|\xi\|_{\varphi} \leq C(\|\eta_t\|^2_{\varphi} + \|\xi\|^2_{\varphi}).$$

Then if γ is large enough and $\sigma = \frac{1}{2}$, we have

$$2\Pi_1 + 2\Pi_2 - \Pi_3 \leq C\left(\|\eta_t\|^2_{\varphi} + \|\xi\|^2_{\varphi}\right).$$

By Gronwall's inequality,

$$\|\xi(T)\|_\varphi^2 \le C\int_0^T \|\eta_t\|_\varphi^2 dt + C\|\xi(0)\|_\varphi^2. \tag{4.43}$$

4.5.4 The final estimate

This part is almost the same as in [117]. We will only discuss the left-hand boundary of $\mathcal{R}_T$ since the discussion for the right one is similar. Denote $x_L(t) = t + x_c$ with

$$x_c = -2s\log(1/h)\gamma h^\sigma,$$

where s and γ are sufficiently large and $\sigma = 1/2$. As mentioned before, the δ-singularity in the initial datum is assumed to be contained in cell I_i. By proposition 4.5.1, we obtain $0 < \phi(x) < h^s$ for any $x \in I_i$. We choose v_0 to satisfy $\mathbb{P}_k v_0 = \mathbb{P}_k u_0 = u_h(0)$. Then

$$\|\xi(0)\|_\varphi \le \|\xi(0)\|_{\varphi,L^2(R\backslash I_i)} + \|\xi(0)\|_{\varphi,L^2(I_i)} \le Ch^{k+1}\|f\|_{k+2} + Ch^{s-1/2}.$$

If s is large enough, then $\|\xi(0)\|_\varphi \le Ch^{k+1}$.

Define the domain $\mathcal{R}_T^+ = (x_L(T), \infty)$, then

$$\|u_h - v\|_{R\backslash\mathcal{R}_T^+} \le \|u_h - v\|_{\varphi,R\backslash\mathcal{R}_T^+} \le \|\eta\|_{\varphi,R\backslash\mathcal{R}_T^+} + \|\xi\|_\varphi \le Ch^{k+1}\|f\|_{k+1} + \|\xi\|_\varphi.$$

To estimate the second term on the right hand side, we use (4.43). Denote

$$w(t) = \max\{x_{j+\frac{1}{2}} : x_{j-\frac{1}{2}} < t + \frac{1}{2}x_c, \forall j\},$$

and $\mathcal{R}_1(t) = (-\infty, w(t))$, $\mathcal{R}_2(t) = R\backslash\mathcal{R}_1(t) = (w(t), \infty)$. If $\gamma h^{\sigma-1}$ is large enough, $\mathcal{R}_1(t)$ stays away from the bad interval $[t-h, t+h]$ where $v(x,t) \ne u(x,t)$, then we have

$$\|\eta_t\|_{\varphi,\mathcal{R}_1(t)} \le Ch^{k+1}\|f\|_{k+2}.$$

Now we proceed to estimate $\|\eta_t\|_{\varphi,\mathcal{R}_2(t)}$. Since $\mathcal{R}_2$ contains the entire bad region, we will use the property of the weight function. By (4.36) we have $\varphi \le h^s$ in this zone. Then we obtain

$$\|\eta_t\|_{\varphi,\mathcal{R}_2(t)} \le Ch^{s/2}\|\eta_t\|_{\mathcal{R}_2(t)} \le Ch^{s/2+k+1}\|\partial_x^{k+2}v\|_{\mathcal{R}_2(t)} \le Ch^{(s-3)/2}+Ch^{s/2+k+1}\|f\|_{k+2,\mathcal{R}_2(t)}.$$

Similarly, we can estimate the right-hand side of the non-smooth region. If we take s large enough, we have

$$\|u_h - u(x,T)\|_{\mathbb{R}\backslash\mathcal{R}_T^+} = \|u_h - v(x,T)\|_{\mathbb{R}\backslash\mathcal{R}_T^+} \le Ch^{k+1}\|f\|_{k+2} + Ch^{(s-3)/2} \le Ch^{k+1}.$$

4.6 Concluding remarks

In this chapter, we use a DG method to solve linear hyperbolic conservation laws involving δ-singularities. We investigate the negative-order norm error estimates for the accuracy of the DG approximations to linear hyperbolic conservation laws with singular initial data or singular source terms, and obtain error estimates in the L^2-norm after post-processing in one space dimension. Numerical experiments demonstrate that the estimates are optimal. The results in this chapter offers evidence that the DG method is a good algorithm for problems involving δ-singularities in their solutions.

Chapter 5

Applications to Krause consensus models and pressureless Euler equations

In this chapter, we apply DG methods to solve hyperbolic conservation law

$$\begin{aligned} &\mathbf{u}_t + \mathbf{f}(\mathbf{u})_x = 0, && (x,t) \in R \times (0,T], \\ &\mathbf{u}(x,0) = \mathbf{u}_0(x), && x \in R, \end{aligned} \tag{5.1}$$

and its two dimensional version, where the exact solution $\mathbf{u}(x,t)$ contains δ-singularities. We extend the work in Chapter 4 and consider applications to two model equations: Krause's consensus models and the pressureless Euler equations. These two models are used to describe the collision of particles, and the distributions can be identified as density functions. If the particles are places at a single point, the density function is a δ-function and is difficult to approximate numerically. Recently, in [119], genuinely maximum-principle-satisfying high order DG schemes for scalar equations and two-dimensional incompressible flows in vorticity-streamfunction formulation have been constructed. Subsequently, positivity-preserving high order DG schemes for compressible Euler equations were given in [120]. We will extend the ideas in

[119, 120] to construct bound-preserving high order DG schemes for the Krause's consensus models and pressureless Euler equations.

For the first example, we discuss the following Krause's consensus model equation

$$\begin{aligned} &\rho_t + F_x = 0, \qquad x \in [0,1], t > 0, \\ &\rho(x,0) = \rho_0(x),\ t > 0, \end{aligned} \tag{5.2}$$

where ρ is the density function, which is always positive. The flux F is given by

$$F(x,t) = v(x,t)\rho(x,t),$$

and the velocity v is defined by

$$v(x,t) = \int_0^1 (y-x)\xi(y-x)\rho(y,t)dy,$$

where $0 \leq \xi(x) \leq 1$ is supported on a ball centered at zero with radius R. In [19], Canuto et al. investigated the discretized version of the PDE and proved that when the time t tends to infinity, the density function ρ will converge to some isolated δ-singularities, and the distance between any two of these δ-singularities cannot be less than R. Some computational results are shown in [19] based on a first order finite volume method. For two dimensions, if the initial density is rotationally invariant, the limit density should also be rotationally invariant, and hence can only be a single δ located at the center. However, direct computations on rectanglular meshes yield more than one δ singularity for sufficiently small R as a resultof the meshes not being invariant under rotation. In this chapter, following ideas in [23, 24], we construct a special mesh to obtain symmetry and convergence to physically relevant solutions. Computational results are given to demonstrate the advantages of high order DG schemes.

For the second example, we discuss the pressureless Euler equation

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x = 0, \quad t > 0,\ x \in \mathbb{R}, \tag{5.3}$$

$$\mathbf{w} = \begin{pmatrix} \rho \\ m \end{pmatrix}, \quad \mathbf{f}(\mathbf{w}) = \begin{pmatrix} m \\ \rho u^2 \end{pmatrix},$$

with $m = \rho u$, where ρ is the density function and u is the velocity. Pressureless Euler equations in one space dimension have been analyzed at the theoretical level intensively, e.g. [13, 14, 18, 22, 41]. Some numerical methods have also been studied by several authors [15, 10, 27, 16]. In [15], only first and second order numerical schemes were considered. Except for those in [15], no other methods seem to have been designed to solve (5.3) directly. In [16], the authors added an artificial viscosity and built a diffusive scheme. In [27], the authors applied the sticky particle methods to the equation and showed that the approximation satisfies the original system within a certain residual. In [10], the authors introduced a new variable and added one more equation to the system, leading to more computational cost. In this chapter, we consider high order DG scheme and approximate the equation without modification. Physically, the density ρ is positive and the velocity u satisfies a maximum principle. We extend the idea in [120] and construct suitable limiters to fulfill these two requirements while maintaining high order accuracy. Moreover, numerical evidences demonstrate that the scheme is good for approximations in the presence of vacuum. Finally, our scheme works well in two dimensions. To the authors' knowledge, not too much works in the literature focus on the two dimensional equations, and some theoretical results can be found in [96, 97]. However, complete existence and uniqueness results are not available. Therefore, our scheme offers a good tool to study two-dimensional pressureless Euler equations and other similar equations.

5.1 Preliminaries

5.1.1 Limiters

In this subsection, we use forward Euler for time discretization and briefly discuss the construction of bound-preserving limiters [121]. We denote $\mathbf{u}_j^n$ and $\overline{\mathbf{u}}_j^n$ to be the numerical solution and its cell average at time level n in cell I_j. If we consider generic numerical solution on the whole computational domain Ω, then the subscript j will be omitted. Suppose the exact solution of (5.1) is in some convex set G and we are interested in constructing numerical solutions which are also in G. The whole precess can be divided into three steps.

In the first step, we consider a first order scheme

$$\begin{aligned}\mathbf{u}_j^{n+1} &= \mathbf{u}_j^n + \lambda\left(\hat{\mathbf{f}}\left(\mathbf{u}_{j-1}^n, \mathbf{u}_j^n\right) - \hat{\mathbf{f}}\left(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n\right)\right)\\ &= \frac{1}{2}\left(\mathbf{u}_j^n + 2\lambda\hat{\mathbf{f}}\left(\mathbf{u}_{j-1}^n, \mathbf{u}_j^n\right)\right) + \frac{1}{2}\left(\mathbf{u}_j^n - 2\lambda\hat{\mathbf{f}}\left(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n\right)\right)\\ &= \frac{1}{2}\mathbf{H}_1\left(\mathbf{u}_{j-1}^n, \mathbf{u}_j^n, 2\lambda\right) + \frac{1}{2}\mathbf{H}_2\left(\mathbf{u}_j^n, \mathbf{u}_{j+1}^n, 2\lambda\right), \qquad (5.4)\end{aligned}$$

where

$$\mathbf{H}_1\left(\mathbf{u}, \mathbf{v}, c\right) = \mathbf{v} + c\,\hat{\mathbf{f}}(\mathbf{u}, \mathbf{v}), \quad \mathbf{H}_2\left(\mathbf{u}, \mathbf{v}, c\right) = \mathbf{u} - c\,\hat{\mathbf{f}}(\mathbf{u}, \mathbf{v}). \qquad (5.5)$$

Here $\mathbf{u}_j^n = \overline{\mathbf{u}}_j^n$ is a constant in each cell I_j, and $\lambda = \frac{\Delta t}{\Delta x}$ is the ratio of time and space mesh sizes. For many two-point first order numerical fluxes, we can prove the following property.

Property 5.1.1. *Suppose G is a convex set and $\mathbf{u}, \mathbf{v} \in G$, then there exists a positive constant $C_\star$, such that, for any $0 < c < C_\star$, we have $\mathbf{H}_1\left(\mathbf{u}, \mathbf{v}, c\right), \mathbf{H}_2\left(\mathbf{u}, \mathbf{v}, c\right) \in G$.*

Based on the above property, we can easily obtain that, under the CFL condition $\lambda < \frac{C_\star}{2}$, $\mathbf{u}^n \in G$ implies $\mathbf{u}^{n+1} \in G$.

Next, we study high order schemes and assume $\mathbf{u}^n \in G$. Consider the equation satisfied by the numerical cell averages

$$\bar{\mathbf{u}}_j^{n+1} = \bar{\mathbf{u}}_j^n + \lambda \left(\hat{\mathbf{f}}(\mathbf{u}_{j-\frac{1}{2}}^-, \mathbf{u}_{j-\frac{1}{2}}^+) - \hat{\mathbf{f}}(\mathbf{u}_{j+\frac{1}{2}}^-, \mathbf{u}_{j+\frac{1}{2}}^+) \right). \tag{5.6}$$

Let α_i be the Legendre Gauss-Lobatto quadrature weights for the interval $[-\frac{1}{2}, \frac{1}{2}]$ such that $\sum_{i=0}^M \alpha_i = 1$, with $2M - 3 \geq k$, and denote the corresponding Gauss-Lobatto points in cell I_j as $\{\check{x}_i^j\}$. Then the Gauss-Lobatto quadrature yields

$$\bar{\mathbf{u}}_j^n = \sum_{i=0}^M \alpha_i \mathbf{u}_j^n(\check{x}_i^j).$$

Clearly, $\mathbf{u}_j^n(\check{x}_0^j) = \mathbf{u}_{j-\frac{1}{2}}^+$ and $\mathbf{u}_j^n(\check{x}_M^j) = \mathbf{u}_{j+\frac{1}{2}}^-$. Therefore,

$$\begin{aligned} \bar{\mathbf{u}}_j^{n+1} &= \sum_{i=0}^M \alpha_i \mathbf{u}_j^n(\check{x}_i^j) + \lambda \left(\hat{\mathbf{f}}(\mathbf{u}_{j-\frac{1}{2}}^-, \mathbf{u}_{j-\frac{1}{2}}^+) - \hat{\mathbf{f}}(\mathbf{u}_{j+\frac{1}{2}}^-, \mathbf{u}_{j+\frac{1}{2}}^+) \right) \\ &= \sum_{i=1}^{M-1} \alpha_i \mathbf{u}_j^n(\check{x}_i^j) + \alpha_0 \mathbf{H}_1 \left(\mathbf{u}_{j-\frac{1}{2}}^-, \mathbf{u}_{j-\frac{1}{2}}^+, \frac{\lambda}{\alpha_0} \right) + \alpha_M \mathbf{H}_2 \left(\mathbf{u}_{j+\frac{1}{2}}^-, \mathbf{u}_{j+\frac{1}{2}}^+, \frac{\lambda}{\alpha_M} \right). \end{aligned}$$

If the numerical flux satisfies Property 5.1.1, we have $\mathbf{H}_1 \left(\mathbf{u}_{j-\frac{1}{2}}^-, \mathbf{u}_{j-\frac{1}{2}}^+, \frac{\lambda}{\alpha_0} \right) \in G$ and $\mathbf{H}_2 \left(\mathbf{u}_{j+\frac{1}{2}}^-, \mathbf{u}_{j+\frac{1}{2}}^+, \frac{\lambda}{\alpha_M} \right) \in G$, provided the suitable CFL condition $\lambda < \alpha_0 C_\star$ is satisfied. Here, we use the fact that $\alpha_0 = \alpha_M$. Since $\mathbf{u}_j^n(\check{x}_i^j) \in G$ and G is a convex set, we have $\bar{\mathbf{u}}_j^{n+1} \in G$.

Finally, we can modify the numerical solution through the simple scaling limiter $\tilde{\mathbf{u}}_j^{n+1} = \bar{\mathbf{u}}_j^{n+1} + \theta \left(\mathbf{u}_j^{n+1} - \bar{\mathbf{u}}_j^{n+1} \right)$. By taking suitable $\theta \in [0, 1]$, we have $\tilde{\mathbf{u}}_j^{n+1} \in G$, and $\tilde{\mathbf{u}}_j^{n+1}$ is used as the numerical solution at time level $n + 1$. In many situations we can prove that this modification does not affect the high order accuracy of the original solution $\mathbf{u}_j^{n+1}$ [121].

5.1.2 High order time discretizations

We will use strong stability preserving (SSP) high order time discretizations to solve the ODE system $\mathbf{u}_t = \mathbf{L}\mathbf{u}$. More details of these time discretizations can be found in [102, 101, 46]. In this chapter, we use the third order SSP Runge-Kutta method [102]

$$\begin{aligned} \mathbf{u}^{(1)} &= \mathbf{u}^n + \Delta t \mathbf{L}(\mathbf{u}^n), \\ \mathbf{u}^{(2)} &= \frac{3}{4}\mathbf{u}^n + \frac{1}{4}\left(\mathbf{u}^{(1)} + \Delta t \mathbf{L}(\mathbf{u}^{(1)})\right), \\ \mathbf{u}^{n+1} &= \frac{1}{3}\mathbf{u}^n + \frac{2}{3}\left(\mathbf{u}^{(2)} + \Delta t \mathbf{L}(\mathbf{u}^{(2)})\right), \end{aligned} \tag{5.7}$$

and the third order SSP multi-step method [101]

$$\mathbf{u}^{n+1} = \frac{16}{27}\left(\mathbf{u}^n + 3\Delta t \mathbf{L}(\mathbf{u}^n)\right) + \frac{11}{27}\left(\mathbf{u}^{n-3} + \frac{12}{11}\Delta t \mathbf{L}(\mathbf{u}^{n-3})\right). \tag{5.8}$$

Since a SSP time discretization is a convex combination of the forward Euler, by using the limiter mentioned in Section 5.1.1, the numerical solution obtained from the full scheme is also in G.

5.2 Krause's consensus models

In this section we apply DG methods to Krause's consensus models, extending the results in [19].

5.2.1 Positivity-preserving high order schemes

We consider (5.2) in more detail. For this model, we define $G = \{\rho : \rho > 0\}$. Clearly, G is a convex set. We start with the following first order scheme

$$\rho_j^{n+1} = \rho_j^n + \lambda\left(v_{j-\frac{1}{2}} h(\rho_{j-1}^n, \rho_j^n) - v_{j+\frac{1}{2}} h(\rho_j^n, \rho_{j+1}^n)\right), \tag{5.9}$$

where $h(\cdot,\cdot)$ is a numerical flux, and $\rho_j^n = \overline{\rho}_j^n$ is the numerical approximation to the exact solution in cell I_j at time level n, with $\overline{\rho}_j^n$ being its cell average. Moreover, $v_{j-\frac{1}{2}}$ is the numerical velocity at the interface $x_{j-\frac{1}{2}}$, given by

$$v_{j-\frac{1}{2}} = \sum_{i=1}^{N} \int_{I_i} (y - x_{j-\frac{1}{2}})\xi(y - x_{j-\frac{1}{2}})\rho_i^n(y)dy.$$

For this model, we use an upwind flux, i.e.

$$vh(u,w) = \begin{cases} vu \;\; v \geq 0, \\ vw \;\; v < 0, \end{cases}$$

and define H_1 and H_2 as

$$H_1(u,w,c) = w + cvh(u,w), \quad H_2(u,w,c) = u - cvh(u,w).$$

Then the scheme (5.9) can be written as

$$\rho_j^{n+1} = \frac{1}{2}H_1(\rho_{j-1}^n, \rho_j^n, 2\lambda) + \frac{1}{2}H_2(\rho_j^n, \rho_{j+1}^n, 2\lambda).$$

Based on the above notations, we can prove the following lemma.

Lemma 5.2.1. *Suppose $\rho^n > 0$, then under the CFL condition*

$$\max_j |v_{j-\frac{1}{2}}|\lambda < \frac{1}{2},$$

we have $\rho^{n+1} > 0$.

Proof: We consider $H_1(\rho_{j-1}^n, \rho_j^n, 2\lambda)$ first. If $v_{j-\frac{1}{2}} < 0$, then

$$H_1(\rho_{j-1}^n, \rho_j^n, 2\lambda) = \rho_j^n + 2\lambda v_{j-\frac{1}{2}} h(\rho_{j-1}^n, \rho_j^n) = (1 + 2\lambda v_{j-\frac{1}{2}})\rho_j^n > 0.$$

On the other hand, if $v_{j-\frac{1}{2}} \geq 0$, then

$$H_1(\rho_{j-1}^n, \rho_j^n, 2\lambda) = \rho_j^n + 2\lambda v_{j-\frac{1}{2}} h(\rho_{j-1}^n, \rho_j^n) = \rho_j^n + 2\lambda v_{j-\frac{1}{2}} \rho_{j-1}^n > 0.$$

Similarly, we have

$$H_2(\rho_j^n, \rho_{j+1}^n, 2\lambda) = \rho_j^n - 2\lambda v_{j+\frac{1}{2}} h(\rho_j^n, \rho_{j+1}^n) > 0.$$

Therefore,

$$\rho_j^{n+1} = \frac{1}{2} H_1(\rho_{j-1}^n, \rho_j^n, 2\lambda) + \frac{1}{2} H_2(\rho_j^n, \rho_{j+1}^n, 2\lambda) > 0.$$

This completes the proof.

Now, we consider high order schemes. The analysis in Section 5.1.1 implies the following theorem.

Theorem 5.2.1. *Suppose the DG solution $\rho^n > 0$, then under the CFL condition*

$$\max_j |v_{j-\frac{1}{2}}| \lambda < \alpha_0,$$

we have $\bar{\rho}^{n+1} > 0$.

Based on the above theorem, we can modify the density ρ_j^n in the following steps.

- Set up a small number $\varepsilon = \min\left\{10^{-13}, \overline{\rho}_j^n\right\}$.
- Compute $m_j = \min_i \rho_j^n(\check{x}_i^j)$, where $\{\check{x}_i^j\}$ are the Gauss-Lobatto points in cell I_j.
- If $m_j < \varepsilon$, then we take

  $$\theta = \frac{\overline{\rho}_j^n - \varepsilon}{\overline{\rho}_j^n - m_j},$$

  and use

  $$\widetilde{\rho}_j^n = \overline{\rho}_j^n + \theta(\rho_j^n - \overline{\rho}_j^n) \tag{5.10}$$

as the DG approximation in cell I_j at time level n.

Based on the above steps, the numerical density is always positive. Therefore

$$\|\rho^n\|_{L^1(\Omega)} = \int_\Omega \rho^n(x)dx = \int_\Omega \rho^0(x)dx = \|\rho_0\|_{L^1(\Omega)}, \tag{5.11}$$

where $\|u\|_{L^1(\Omega)}$ is the standard L^1-norm of u on Ω. Clearly, (5.11) implies the L^1 stability for the DG scheme. Moreover, we can also derive a sufficient CFL condition which does not depend on the numerical velocity in Theorem 5.2.1. Actually, Note that

$$v_{j-\frac{1}{2}} = \sum_{i=1}^{N} \int_{I_i} (y - x_{j-\frac{1}{2}})\xi(y - x_{j-\frac{1}{2}})\rho_i^n(y)dy \leq R\|\rho_0\|_{L^1(\Omega)},$$

such that the sufficient CFL condition is

$$\lambda \leq \frac{\alpha_0}{R\|\rho_0\|_{L^1(\Omega)}}. \tag{5.12}$$

To summarize, we have the following theorem.

Theorem 5.2.2. *Under the CFL condition* (5.12)*, the DG scheme with the positivity-preserving limiter for equation* (5.2) *is* L^1 *stable and the density function is always positive.*

5.2.2 Numerical experiments

In this subsection, some numerical examples will be given to demonstrate the good performance of the DG scheme.

Example 5.1. *We consider the following problem*

$$\begin{aligned} &\rho_t + (v\rho)_x = 0, \quad x \in [0,1], t > 0, \\ &\rho(x,0) = \rho_0(x), \quad t > 0, \end{aligned} \tag{5.13}$$

where the velocity v is defined by

$$v(x,t) = \int_{x-R}^{x+R} (y-x)\rho(y,t)dy.$$

We apply the positivity-preserving limiter and use $\mathcal{P}^0$ and $\mathcal{P}^1$ polynomials. Moreover, we use the third order SSP Runge-Kutta discretization in time [102] with $\Delta t = 0.1\Delta x$. Figure 5.1 shows the numerical approximations of $\rho(x)$ at $t = 1000$, with $N = 400$, $\rho_0 = 1$, and $R = 0.02$. We can observe 22 δ-singularities in each panel, and the distance between any two adjacent singularities is greater than R. The algorithm is quite stable in this simulation. Moreover, the $\mathcal{P}^1$ solution in the right panel is more accurate than the $\mathcal{P}^0$ one in the left panel, since the heights of the δ-singularities are almost doubled.

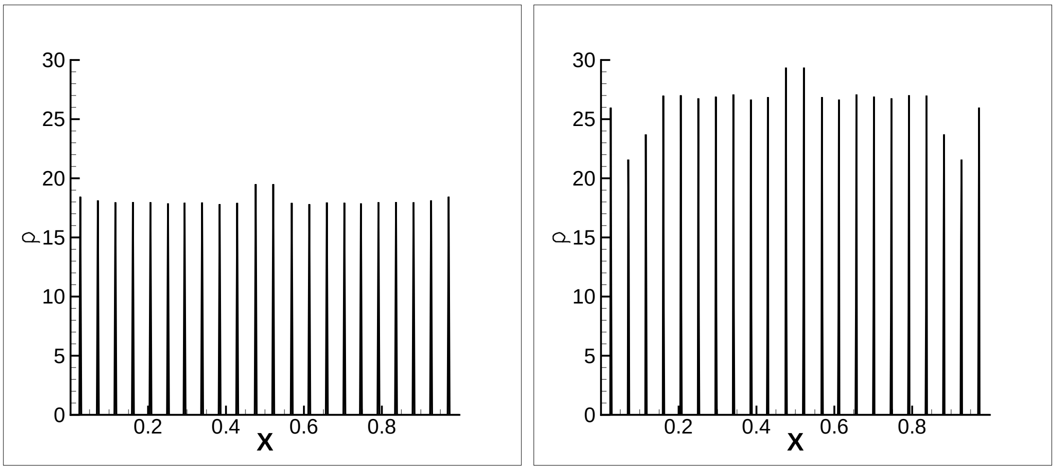

Figure 5.1: Numerical density for (5.13) at $t = 1000$ with $N = 400$ when using $\mathcal{P}^0$ (left) and $\mathcal{P}^1$ (right) polynomials.

Example 5.2. *We consider the model problem in two dimensions.*

$$\begin{aligned} &\rho_t + div(\mathbf{v}\rho) = 0, && \mathbf{x} \in [-1,1]^2, t > 0, \\ &\rho(\mathbf{x},0) = \rho_0(\mathbf{x}), && t > 0, \end{aligned} \tag{5.14}$$

where the velocity $\mathbf{v}$ *is defined by*

$$\mathbf{v}(\mathbf{x},t) = \int_{B_R(\mathbf{x})} (\mathbf{y}-\mathbf{x})\rho(\mathbf{y},t)d\mathbf{y}.$$

In this example, we take $R = 0.1$ and

$$\rho_0(\mathbf{x}) = \begin{cases} 1 \; r < 0.5, \\ 0 \; r > 0.5, \end{cases}$$

where $r = \|\mathbf{x}\|$ is the Euclidean norm of $\mathbf{x}$. In [19], the authors demonstrated that the exact solution should be a single delta placed at the origin. However, by using rectangle meshes, we observe more than one delta singularity for R sufficiently small as a consequence of the meshes not being invariant under rotation. To address this problem, we follow [23, 24], and construct a special equal-angle-zoned mesh. The structure of the mesh is given in Figure 5.2. By using this mesh, the limit density given in Figure 5.3 is a single delta placed at the origin.

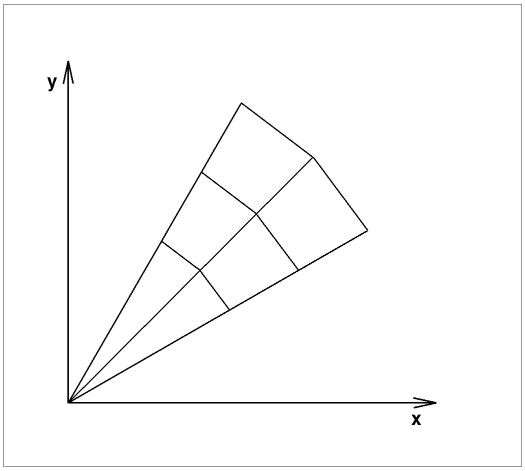

Figure 5.2: Equal-angle-zoned mesh.

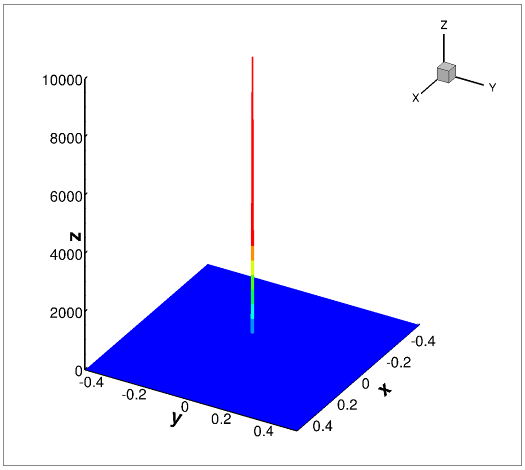

Figure 5.3: Numerical density ρ for (5.14) at $t = 2000$ with $N = 200$ when using $\mathcal{P}^0$ polynomials.

5.3 Pressureless Euler equations

In this section, we apply DG methods to the pressureless Euler equations.

5.3.1 Numerical schemes in one dimension

We study (5.3) in more detail. Physically, the density is positive and the velocity satisfies the maximum principle. Therefore, we define

$$G = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \end{pmatrix} : \rho > 0, a\rho \leq m \leq b\rho \right\},$$

where

$$a = \min u_0(x), \qquad b = \max u_0(x), \tag{5.15}$$

with u_0 being the initial velocity. Clearly, G is a convex set. As mentioned in Section 5.1.1, we start with the following first order scheme,

$$\mathbf{w}_j^{n+1} = \mathbf{w}_j^n + \lambda \left(\mathbf{h}(\mathbf{w}_{j-1}^n, \mathbf{w}_j^n) - \mathbf{h}(\mathbf{w}_j^n, \mathbf{w}_{j+1}^n) \right), \tag{5.16}$$

where $\mathbf{h}(\cdot,\cdot)$ is a numerical flux and $\mathbf{w}_j^n = \left(\rho_j^n, m_j^n\right)^T$ is the numerical approximation to the exact solution in cell I_j at time level n. Moreover, we define $\overline{\mathbf{w}}_j^n = \left(\overline{\rho}_j^n, \overline{m}_j^n\right)^T$ as its cell average. Clearly, for a first order scheme, $\mathbf{w}_j^n = \overline{\mathbf{w}}_j^n$ in (5.16). For simplicity, we use u_j^n for $\frac{m_j^n}{\rho_j^n}$ as the numerical velocity throughout this section. In this problem, we consider the Godunov flux [15]. Suppose that at the cell interface $x = x_{j-\frac{1}{2}}$ we have two numerical approximations $\mathbf{w}_\ell = (\rho_\ell, m_\ell)^T$ and $\mathbf{w}_r = (\rho_r, m_r)^T$ from the left and right respectively. Then the Godunov flux is given as

$$\left(\mathbf{h}\left(\mathbf{w}_\ell, \mathbf{w}_r\right)\right)^T = (\widehat{\rho u}_{j-\frac{1}{2}}, \widehat{\rho u^2}_{j-\frac{1}{2}}) = \begin{cases} (m_\ell, \rho_\ell u_\ell^2) & u_\ell > 0, u_r > 0, \\ (0,0) & u_\ell \le 0, u_r > 0, \\ (m_r, \rho_r u_r^2) & u_\ell \le 0, u_r \le 0, \\ (m_\ell, \rho_\ell u_\ell^2) & u_\ell > 0, u_r \le 0, v > 0, \\ (m_r, \rho_r u_r^2) & u_\ell > 0, u_r \le 0, v < 0, \\ (\frac{m_\ell + m_r}{2}, \rho_\ell u_\ell^2 = \rho_r u_r^2) & u_\ell > 0, u_r \le 0, v = 0, \end{cases} \tag{5.17}$$

where

$$u_\ell = \frac{m_\ell}{\rho_\ell}, \quad u_r = \frac{m_r}{\rho_r}, \text{ and } v = \frac{\sqrt{\rho_\ell} u_\ell + \sqrt{\rho_r} u_r}{\sqrt{\rho_\ell} + \sqrt{\rho_r}}.$$

For this problem, $\mathbf{H}_1$ and $\mathbf{H_2}$ are taken to be

$$\mathbf{H}_1\left(\mathbf{u}, \mathbf{v}, c\right) = \mathbf{v} + c\,\mathbf{h}\left(\mathbf{u}, \mathbf{v}\right), \quad \mathbf{H_2}\left(\mathbf{u}, \mathbf{v}, c\right) = \mathbf{u} - c\,\mathbf{h}\left(\mathbf{u}, \mathbf{v}\right). \tag{5.18}$$

Clearly, (5.16) can be written as

$$\mathbf{w}_j^{n+1} = \frac{1}{2}\mathbf{H}_1\left(\mathbf{w}_{j-1}^n, \mathbf{w}_j^n, 2\lambda\right) + \frac{1}{2}\mathbf{H}_2\left(\mathbf{w}_j^n, \mathbf{w}_{j+1}^n, 2\lambda\right).$$

Before proceeding to the theoretical results for the scheme, we would like to introduce the following lemma. The proof is trivial and is omitted.

Lemma 5.3.1. *Suppose* $\{\check{x}_i\}$ *are positive real numbers, and* $a \leq \check{y}_i \leq b, \forall i,$ *then*

$$a \leq \frac{\sum_{i=1}^n \check{x}_i \check{y}_i}{\sum_{i=1}^n \check{x}_i} \leq b.$$

We will use Lemma 5.3.1 to prove the following lemma.

Lemma 5.3.2. *Suppose* $\mathbf{w}^n \in G$*, then under the CFL condition*

$$\lambda < \frac{1}{2\max(|a|, |b|)},$$

where a *and* b *are defined in (5.15), we have* $\mathbf{w}^{n+1} \in G$.

Proof: We will only prove $\mathbf{H}_1\left(\mathbf{w}_{j-1}^n, \mathbf{w}_j^n, 2\lambda\right) \in G$. The proof for $\mathbf{H}_2\left(\mathbf{w}_j^n, \mathbf{w}_{j+1}^n, 2\lambda\right) \in G$ follows the same lines. Define $\mathbf{H}_1\left(\mathbf{w}_{j-1}^n, \mathbf{w}_j^n, 2\lambda\right) = (\check{\rho}, \check{m})^T$, then the velocity derived from $\mathbf{H}_1$ is given as

$$\check{u} = \frac{\check{m}}{\check{\rho}} = \frac{m_j^n + 2\lambda\widehat{\rho u^2}_{j-\frac{1}{2}}}{\rho_j^n + 2\lambda\widehat{\rho u}_{j-\frac{1}{2}}}. \tag{5.19}$$

We will prove $\check{\rho} > 0$ and $a \leq \check{u} \leq b$. To do so, we have to determine what $\{\check{x}_i\}$ and $\{\check{y}_i\}$ should be in Lemma 5.3.1, by testing the different choices for the numerical flux in (5.17). For simplicity, we define $\widehat{u}_{j-\frac{1}{2}} = \frac{\widehat{\rho u^2}_{j-\frac{1}{2}}}{\widehat{\rho u}_{j-\frac{1}{2}}}$, if $\widehat{\rho u}_{j-\frac{1}{2}} \neq 0$.

- If $\widehat{\rho u}_{j-\frac{1}{2}} = m_{j-1}^n$, then $\widehat{u}_{j-\frac{1}{2}} = u_{j-1}^n > 0$. We take $\check{x}_1 = \rho_j^n$, $\check{y}_1 = u_j^n$ and $\check{x}_2 = 2\lambda m_{j-1}^n$, $\check{y}_2 = u_{j-1}^n$.

- If $\widehat{\rho u}_{j-\frac{1}{2}} = m_j^n$, then $\widehat{u}_{j-\frac{1}{2}} = u_j^n \leq 0$. We take $\check{x}_1 = \rho_j^n + 2\lambda m_j^n$, $\check{y}_1 = u_j^n$.

- If $\widehat{\rho u}_{j-\frac{1}{2}} = (m_{j-1}^n + m_j^n)/2$, then $\widehat{\rho u^2}_{j-\frac{1}{2}} = (m_{j-1}^n u_{j-1}^n + m_j^n u_j^n)/2$. We combine the two situations above, and take $\check{x}_1 = \rho_j^n + \lambda m_j^n$, $\check{y}_1 = u_j^n$ and $\check{x}_2 = \lambda m_{j-1}^n$, $\check{y}_2 = u_{j-1}^n$.

- If $\widehat{\rho u}_{j-\frac{1}{2}} = 0$, then $\widehat{\rho u^2}_{j-\frac{1}{2}} = 0$. We take $\check{x}_1 = \rho_j^n$, $\check{y}_1 = u_j^n$.

Clearly, in each case, $\check{x}_i > 0$, therefore $\check{\rho} > 0$. Moreover, we observe $a \leq \check{y}_i \leq b$, so that by Lemma 5.3.1, we have $a \leq \check{u} \leq b$.

Now, we consider high order schemes. By the same analysis as in Section 5.1.1, we have the following result.

Theorem 5.3.1. *Suppose* $\mathbf{w}^n \in G$, *then under the CFL condition*

$$\lambda < \frac{\alpha_0}{\max(|a|,|b|)},$$

we have $\overline{\mathbf{w}}^{n+1} \in G$.

Based on this, we can modify the numerical solution $\mathbf{w}_j^n$ while keeping the cell average untouched. Due to rounding error, we define

$$G^\varepsilon = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \end{pmatrix} : \rho \geq \varepsilon, a - \varepsilon \leq \frac{m}{\rho} \leq b + \varepsilon \right\},$$

$$\partial G^\varepsilon = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \end{pmatrix} : \rho \geq \varepsilon, \frac{m}{\rho} = a - \varepsilon \text{ or } b + \varepsilon \right\}.$$

Then the modification of $\mathbf{w}_j^n$ is given in the following steps.

- Set up a small number $\varepsilon = 10^{-13}$.

- If $\overline{\rho}_j^n > \varepsilon$, then proceed to the following steps. Otherwise, ρ_j^n is identified as the approximation to vacuum, and the velocity is undefined. Therefore, we take $\widetilde{\mathbf{w}}_j^n = \overline{\mathbf{w}}_j^n$ as the numerical solution and skip the following steps.

- Modify the density first: Compute $m_j = \min_i \rho_j^n(\check{x}_i^j)$, where $\{\check{x}_i^j\}$ are the Gauss-Lobatto points in cell I_j, and get $\widetilde{\rho}_j^n$ by (5.10). Then use $\widetilde{\rho}_j^n$ as the new numerical density ρ_j^n.

- Modify the velocity: Define $\mathbf{q}_i^j = \mathbf{w}_j^n(\breve{x}_i^j)$ in cell I_j. If $\mathbf{q}_i^j \in G^\varepsilon$, then take $\theta_i^j = 1$. Otherwise, take
$$\theta_i^j = \frac{\left\|\overline{\mathbf{w}}_j^n - \mathbf{s}_i^j\right\|}{\left\|\overline{\mathbf{w}}_j^n - \mathbf{q}_i^j\right\|},$$
where $\|\cdot\|$ is the Euclidean norm, and $\mathbf{s}_i^j$ is the intersection point of the straight line
$$\mathbf{s}(t) = (1-t)\overline{\mathbf{w}}_j^n + t\mathbf{q}_i^j, \quad 0 \le t \le 1,$$
and the surface ∂G^ε. Define $\theta_j = \min_{i=0,\cdots,m} \theta_i^j$, and use
$$\widetilde{\mathbf{w}}_j^n = \overline{\mathbf{w}}_j^n + \theta_j(\mathbf{w}_j^n - \overline{\mathbf{w}}_j^n),$$
as the DG approximation in cell I_j.

5.3.2 Numerical schemes in two dimensions

We extend our work to two dimensions and consider the following equation

$$\mathbf{w}_t + \mathbf{f}(\mathbf{w})_x + \mathbf{g}(\mathbf{w})_y = 0, \quad t > 0,\ (x,y) \in \mathbb{R}^2, \tag{5.20}$$

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

with

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

where ρ is the density function and (u, v) is the velocity field. We define

$$G = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \\ n \end{pmatrix} : \rho > 0, m^2 + n^2 \le S^2 \rho^2 \right\},$$

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. Clearly, G is a convex set.

For simplicity, we use uniform rectangular meshes. The cell is defined as $I_{ij} = \left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}} \right] \times \left[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}} \right]$, and the mesh sizes in x and y directions are denoted as Δx and Δy respectively. At time level n, we approximate the exact solution with a vector of polynomials of degree k, $\mathbf{w}_{ij}^n = (\rho_{ij}^n, m_{ij}^n, n_{ij}^n)^T$, and define the cell average $\overline{\mathbf{w}}_{ij}^n = (\overline{\rho}_{ij}^n, \overline{m}_{ij}^n, \overline{n}_{ij}^n)^T$. Moreover, we denote $\mathbf{w}^+_{i-\frac{1}{2},j}(y), \mathbf{w}^-_{i+\frac{1}{2},j}(y), \mathbf{w}^+_{i,j-\frac{1}{2}}(x), \mathbf{w}^-_{i,j+\frac{1}{2}}(x)$ as the traces of $\mathbf{w}$ on the four edges of cell I_{ij} respectively. More details can be found in [120]. We use (u_{ij}^n, v_{ij}^n) for $(\frac{m_{ij}^n}{\rho_{ij}^n}, \frac{n_{ij}^n}{\rho_{ij}^n})$ as the numerical velocity field in cell I_{ij} at time level n, and define $a_1 = \max_{ij} |u_{ij}^n|$ and $a_2 = \max_{ij} |v_{ij}^n|$. For simplicity, if we consider a generic numerical solution on the whole computational domain at time level n, the subscript ij will be omitted.

We only consider high order schemes, and the one satisfied by the cell averages can be written as

$$\begin{aligned}
\overline{\mathbf{w}}_{ij}^{n+1} = \overline{\mathbf{w}}_{ij}^n &+ \frac{\Delta t}{\Delta x \Delta y} \int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} \mathbf{h}_1 \left(\mathbf{w}^-_{i-\frac{1}{2},j}(y), \mathbf{w}^+_{i-\frac{1}{2},j}(y) \right) - \mathbf{h}_1 \left(\mathbf{w}^-_{i+\frac{1}{2},j}(y), \mathbf{w}^+_{i+\frac{1}{2},j}(y) \right) dy \\
&+ \frac{\Delta t}{\Delta x \Delta y} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{h}_2 \left(\mathbf{w}^-_{i,j-\frac{1}{2}}(x), \mathbf{w}^+_{i,j-\frac{1}{2}}(x) \right) - \mathbf{h}_2 \left(\mathbf{w}^-_{i,j+\frac{1}{2}}(x), \mathbf{w}^+_{i,j+\frac{1}{2}}(x) \right) dx, \quad (5.21)
\end{aligned}$$

where $\mathbf{h}_1(\cdot,\cdot)$ and $\mathbf{h}_2(\cdot,\cdot)$ are one-dimensional numerical fluxes. For this problem, we also use the Godunov flux. Suppose $(x,y) = (x_{i-\frac{1}{2}}, y_0)$ is a point on the vertical cell interface, at which we have two numerical approximations $\mathbf{w}_\ell = (\rho_\ell, m_\ell, n_\ell)^T$ and $\mathbf{w}_r = (\rho_r, m_r, n_r)^T$ from left and right respectively. Then the Godunov flux

$(\mathbf{h}_1(\mathbf{w}_\ell, \mathbf{w}_r))^T$ can be written as

$$\left(\widehat{\rho u}, \widehat{\rho u^2}, \widehat{\rho uv}\right) = \begin{cases} (m_\ell, \rho_\ell u_\ell^2, \rho_\ell u_\ell v_\ell) & u_\ell > 0, u_r > 0, \\ (0,0,0) & u_\ell \leq 0, u_r > 0, \\ (m_r, \rho_r u_r^2, \rho_r u_r v_r) & u_\ell \leq 0, u_r \leq 0, \\ (m_\ell, \rho_\ell u_\ell^2, \rho_\ell u_\ell v_\ell) & u_\ell > 0, u_r \leq 0, v > 0, \\ (m_r, \rho_r u_r^2, \rho_r u_r v_r) & u_\ell > 0, u_r \leq 0, v < 0, \\ \frac{1}{2}(m_\ell + m_r, \rho_\ell u_\ell^2 + \rho_r u_r^2, m_\ell v_\ell + m_r v_r) & u_\ell > 0, u_r \leq 0, v = 0, \end{cases}$$

where

$$(u_\ell, v_\ell) = \left(\frac{m_\ell}{\rho_\ell}, \frac{n_\ell}{\rho_\ell}\right), \ (u_r, v_r) = \left(\frac{m_r}{\rho_r}, \frac{n_\ell}{\rho_\ell}\right), \text{ and } v = \frac{\sqrt{\rho_\ell} u_\ell + \sqrt{\rho_r} u_r}{\sqrt{\rho_\ell} + \sqrt{\rho_r}}.$$

The numerical flux $\mathbf{h}_2 = \left(\widehat{\rho v}, \widehat{\rho uv}, \widehat{\rho v^2}\right)^T$ can be defined in a similar way on the horizontal cell interfaces.

For accuracy, we use L-point Gauss quadratures with $L \geq k+1$ to approximate the integrals in (5.21). More details of this requirement can be found in [32]. The Gauss quadrature points on $\left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$ and $\left[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}\right]$ are denoted by

$$\hat{p}_i^x = \left\{\hat{x}_i^\beta : \beta = 1, \cdots, L\right\} \text{ and } \hat{p}_j^y = \left\{\hat{y}_j^\beta : \beta = 1, \cdots, L\right\},$$

respectively. Also, we denote $\hat{w}_\beta$ as the corresponding weights on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$. Following the same notation as in previous sections, we use

$$\check{p}_i^x = \{\check{x}_i^\alpha : \alpha = 0, \cdots, M\} \text{ and } \check{p}_j^y = \left\{\check{y}_j^\alpha : \alpha = 0, \cdots, M\right\}$$

as the Gauss-Lobatto points on $\left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$ and $\left[y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}\right]$ respectively. Also, we denote $\check{w}_\alpha$ as the corresponding weights on the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$.

Let $\lambda_1 = \frac{\Delta t}{\Delta x}$ and $\lambda_2 = \frac{\Delta t}{\Delta y}$, then the numerical scheme (5.21) becomes

$$\begin{aligned}\overline{\mathbf{w}}_{ij}^{n+1} &= \overline{\mathbf{w}}_{ij}^{n} + \lambda_1 \sum_{\beta=1}^{L} \hat{w}_\beta \left[\mathbf{h}_1\left(\mathbf{w}^-_{i-\frac{1}{2},\beta}, \mathbf{w}^+_{i-\frac{1}{2},\beta}\right) - \mathbf{h}_1\left(\mathbf{w}^-_{i+\frac{1}{2},\beta}, \mathbf{w}^+_{i+\frac{1}{2},\beta}\right)\right] \\ &+ \lambda_2 \sum_{\beta=1}^{L} \hat{w}_\beta \left[\mathbf{h}_2\left(\mathbf{w}^-_{\beta,j-\frac{1}{2}}, \mathbf{w}^+_{\beta,j-\frac{1}{2}}\right) - \mathbf{h}_2\left(\mathbf{w}^-_{\beta,j+\frac{1}{2}}, \mathbf{w}^+_{\beta,j+\frac{1}{2}}\right)\right], \qquad (5.22)\end{aligned}$$

where $\mathbf{w}^-_{i-\frac{1}{2},\beta} = \mathbf{w}^-_{i-\frac{1}{2},j}(\hat{y}_j^\beta)$ is a point value in the Gauss quadrature. Likewise for the other point values. As the general treatment, we rewrite the cell average on the right hand side as

$$\overline{\mathbf{w}}_{ij}^{n} = \sum_{\alpha=0}^{M}\sum_{\beta=1}^{L} \breve{w}_\alpha \hat{w}_\beta \mathbf{w}^1_{\alpha\beta} = \sum_{\alpha=0}^{M}\sum_{\beta=1}^{L} \breve{w}_\alpha \hat{w}_\beta \mathbf{w}^2_{\beta\alpha},$$

where $\mathbf{w}^1_{\alpha\beta}$ and $\mathbf{w}^2_{\beta\alpha}$ denote $\mathbf{w}^n_{ij}(\breve{x}_i^\alpha, \hat{y}_j^\beta)$ and $\mathbf{w}^n_{ij}(\hat{x}_i^\beta, \breve{y}_j^\alpha)$ respectively. We extend the definitions of $\mathbf{H}_1$ and $\mathbf{H}_2$ in (5.18) to two-dimensional problems and define

$$\begin{aligned}\mathbf{H}^1_1(\mathbf{u},\mathbf{v},c) &= \mathbf{v} + c\,\mathbf{h}_1(\mathbf{u},\mathbf{v}), \quad \mathbf{H}^1_2(\mathbf{u},\mathbf{v},c) = \mathbf{u} - c\,\mathbf{h}_1(\mathbf{u},\mathbf{v}), \\ \mathbf{H}^2_1(\mathbf{u},\mathbf{v},c) &= \mathbf{v} + c\,\mathbf{h}_2(\mathbf{u},\mathbf{v}), \quad \mathbf{H}^2_2(\mathbf{u},\mathbf{v},c) = \mathbf{u} - c\,\mathbf{h}_2(\mathbf{u},\mathbf{v}).\end{aligned}$$

Let $\mu = a_1\lambda_1 + a_2\lambda_2$, then scheme (5.22) can be written as

$$\begin{aligned}\overline{\mathbf{w}}_{ij}^{n+1} &= C_1 \sum_{\beta=1}^{L} \hat{w}_\beta \left(\sum_{\alpha=1}^{M-1} \breve{w}_\alpha \mathbf{w}^1_{\alpha\beta} + \breve{w}_0 \mathbf{H}^1_1\left(\mathbf{w}^-_{i-\frac{1}{2},\beta}, \mathbf{w}^+_{i-\frac{1}{2},\beta}, \mu_1\right) + \breve{w}_M \mathbf{H}^1_2\left(\mathbf{w}^-_{i+\frac{1}{2},\beta}, \mathbf{w}^+_{i+\frac{1}{2},\beta}, \mu_1\right)\right) \\ &+ C_2 \sum_{\beta=1}^{L} \hat{w}_\beta \left(\sum_{\alpha=1}^{M-1} \breve{w}_\alpha \mathbf{w}^2_{\beta\alpha} + \breve{w}_0 \mathbf{H}^2_1\left(\mathbf{w}^-_{\beta,j-\frac{1}{2}}, \mathbf{w}^+_{\beta,j-\frac{1}{2}}, \mu_2\right) + \breve{w}_M \mathbf{H}^2_2\left(\mathbf{w}^-_{\beta,i+\frac{1}{2}}, \mathbf{w}^+_{\beta,j+\frac{1}{2}}, \mu_2\right)\right),\end{aligned}$$

where

$$C_1 = \frac{a_1\lambda_1}{\mu},\ C_2 = \frac{a_2\lambda_2}{\mu},\ \mu_1 = \frac{\mu}{a_1\breve{w}_0},\ \mu_2 = \frac{\mu}{a_2\breve{w}_0}.$$

Now, we can state the main theorem.

Theorem 5.3.2. *Suppose* $\mathbf{w}^n \in G$*, then under the CFL condition*

$$\frac{\Delta t}{\Delta x}a_1 + \frac{\Delta t}{\Delta y}a_2 \leq \hat{w}_0,$$

we have $\overline{\mathbf{w}}^{n+1} \in G$*.*

Proof: For simplicity, we only prove $\mathbf{H}_1^1\left(\mathbf{w}^-_{i-\frac{1}{2},\beta}, \mathbf{w}^+_{i-\frac{1}{2},\beta}, \mu_1\right) \in G, \forall\beta$, and define

$$\mathbf{H}_1^1\left(\mathbf{w}^-_{i-\frac{1}{2},\beta}, \mathbf{w}^+_{i-\frac{1}{2},\beta}, \mu_1\right) = \left(\check{\rho}, \check{m}, \check{n}\right)^T, \check{u} = \frac{\check{m}}{\check{\rho}}, \check{v} = \frac{\check{n}}{\check{\rho}}.$$

Following the same analysis as in Lemma 5.3.2, we have $\check{\rho} > 0$. Therefore, we need only prove $\check{u}^2 + \check{v}^2 \leq S^2$. By the assumption, we have

$$\mathbf{w}^-_{i-\frac{1}{2},\beta} = \left(\rho^-_{i-\frac{1}{2},\beta}, m^-_{i-\frac{1}{2},\beta}, n^-_{i-\frac{1}{2},\beta}\right)^T \in G \text{ and } \mathbf{w}^+_{i-\frac{1}{2},\beta} = \left(\rho^+_{i-\frac{1}{2},\beta}, m^+_{i-\frac{1}{2},\beta}, n^+_{i-\frac{1}{2},\beta}\right)^T \in G.$$

Denote $\mathbf{h}_1\left(\mathbf{w}^-_{i-\frac{1}{2},\beta}, \mathbf{w}^+_{i-\frac{1}{2},\beta}\right) = \left(\widehat{\rho u}_{i-\frac{1}{2},\beta}, \widehat{\rho u^2}_{i-\frac{1}{2},\beta}, \widehat{\rho uv}_{i-\frac{1}{2},\beta}\right)^T$ as the corresponding numerical flux, and for any unit vector $\mathbf{n} = (n_1, n_2)^T$, define $\check{w} = \check{u}n_1 + \check{v}n_2$. Then

$$\begin{aligned}\check{w} &= \frac{m^+_{i-\frac{1}{2},\beta}n_1 + n^+_{i-\frac{1}{2},\beta}n_2 + \mu_1\left(\widehat{\rho u^2}_{i-\frac{1}{2},\beta}n_1 + \widehat{\rho uv}_{i-\frac{1}{2},\beta}n_2\right)}{\rho^+_{i-\frac{1}{2},\beta} + \mu_1\widehat{\rho u}_{i-\frac{1}{2},\beta}} \\ &= \frac{\rho^+_{i-\frac{1}{2},\beta}w^+_{i-\frac{1}{2},\beta} + \mu_1\widehat{\rho u}_{i-\frac{1}{2},\beta}\widehat{w}_{i-\frac{1}{2},\beta}}{\rho^+_{i-\frac{1}{2},\beta} + \mu_1\widehat{\rho u}_{i-\frac{1}{2},\beta}},\end{aligned}$$

where

$$w^+_{i-\frac{1}{2},\beta} = \frac{m^+_{i-\frac{1}{2},\beta}n_1 + n^+_{i-\frac{1}{2},\beta}n_2}{\rho^+_{i-\frac{1}{2},\beta}}, \quad \widehat{w}_{i-\frac{1}{2},\beta} = \frac{\widehat{\rho u^2}_{i-\frac{1}{2},\beta}n_1 + \widehat{\rho uv}_{i-\frac{1}{2},\beta}n_2}{\widehat{\rho u}_{i-\frac{1}{2},\beta}}.$$

We can easily show that $|w^+_{i-\frac{1}{2},\beta}| \leq S$ and $|\widehat{w}_{i-\frac{1}{2},\beta}| \leq S$. Following the same lines as is the proof of Lemma 5.3.2, we have $|\check{w}| \leq S$. Choosing $\mathbf{n}$ to be parallel with $(\check{u}, \check{v})$, we have $\check{u}^2 + \check{v}^2 \leq S^2$, completing the proof.

Remark 5.3.1. *Since* $a_1 \leq S$ *and* $a_2 \leq S$*, another sufficient CFL condition in Theorem 5.3.2 is* $\frac{\Delta t}{\Delta x} + \frac{\Delta t}{\Delta y} \leq \frac{\hat{w}_0}{S}$.

Based on the above theorem, we can modify the numerical solution $\mathbf{w}_{ij}^n$ keeping the cell average untouched. Due to the rounding error, we define

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

$$\partial G^\varepsilon = \left\{ \mathbf{w} = \begin{pmatrix} \rho \\ m \\ n \end{pmatrix} : \rho \geq \varepsilon, m^2 + n^2 = (S+\varepsilon)^2 \rho^2 \right\}.$$

Then the modification of $\mathbf{w}_{ij}^n$ is given in the following steps.

- Set up a small number $\varepsilon = 10^{-13}$.
- If $\overline{\rho}_{ij}^n > \varepsilon$, then proceed to the following steps. Otherwise, ρ_{ij}^n is identified as the approximation to vacuum, and the velocity is undefined. Therefore, we take $\widetilde{\mathbf{w}}_{ij}^n = \overline{\mathbf{w}}_{ij}^n$ as the numerical solution and skip the following steps.
- Modify the density first: Compute $m_{ij} = \min_{\alpha\beta} \left\{ \rho_{ij}^n(\check{x}_i^\alpha, \hat{y}_j^\beta), \rho_{ij}^n(\hat{x}_i^\beta, \check{y}_j^\alpha) \right\}$. If $m_{ij} < \varepsilon$, then take $\widetilde{\rho}_{ij}^n$ as

  $$\widetilde{\rho}_{ij}^n = \overline{\rho}_{ij}^n + \theta_{ij}\left(\rho_{ij}^n - \overline{\rho}_{ij}^n\right),$$

  with
  $$\theta_{ij} = \frac{\overline{\rho}_{ij}^n - \varepsilon}{\overline{\rho}_{ij}^n - m_{ij}},$$
  and use $\widetilde{\rho}_{ij}^n$ as the new numerical density ρ_{ij}^n.
- Modify the velocity: Consider $\mathbf{w}_{\alpha\beta}^1$ and $\mathbf{w}_{\beta\alpha}^2$ in cell I_{ij}. If $\mathbf{w}_{\alpha\beta}^1 \in G^\varepsilon$, then take

$\theta^1_{\alpha\beta} = 1$. Otherwise, take

$$\theta^1_{\alpha\beta} = \frac{\left\|\overline{\mathbf{w}}^n_{ij} - \mathbf{s}^1_{\alpha\beta}\right\|}{\left\|\overline{\mathbf{w}}^n_{ij} - \mathbf{w}^1_{\alpha\beta}\right\|},$$

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

$$\mathbf{s}^1(t) = (1-t)\overline{\mathbf{w}}^n_{ij} + t\mathbf{w}^1_{\alpha\beta}, \quad 0 \le t \le 1,$$

and the surface ∂G^ε. Similarly, we can define $\theta^2_{\beta\alpha}$ in the same way for $\mathbf{w}^2_{\beta\alpha}$. Finally, we use

$$\widetilde{\mathbf{w}}^n_{ij} = \overline{\mathbf{w}}^n_{ij} + \theta(\mathbf{w}^n_{ij} - \overline{\mathbf{w}}^n_{ij}), \quad \theta = \min_{\alpha,\beta}\left\{\theta^1_{\alpha\beta}, \theta^2_{\beta\alpha}\right\},$$

as the DG approximation in cell I_{ij}.

5.3.3 Numerical experiments

Let us provide numerical experiments to demonstrate the good performance of the DG scheme for solving pressureless Euler equations. In all numerical simulations, if not otherwise stated, we use third order schemes and take $N = 100$.

One space dimension

We consider the problem in one space dimension and solve (5.3) with different initial conditions.

Example 5.3. *We consider the following initial data*

$$\rho_0(x) = \sin(x) + 2, \;\; u_0(x) = \sin(x) + 2, \tag{5.23}$$

with periodic boundary condition.

Clearly, the exact solution is

$$u(x,t) = u_0(x_0), \quad \rho(x,t) = \frac{\rho_0(x_0)}{1+u_0'(x_0)},$$

where x_0 is given implicitly by

$$x_0 + tu_0(x_0) = x.$$

We use the third order SSP multi-step method in time [101] with $\Delta t = 0.01\Delta x^2$, and test the example by using $\mathcal{P}^k$ polynomials with $k = 1, 2, 3$ on uniform meshes. Table 5.1 shows the L^2-norm of the error at $t = 0.1$. We observe $(k + 0.5)$-th order convergence.

Table 5.1: L^2-norm of the error between the numerical density and the exact density for initial condition (5.23).

	k=1		k=2		k=3	
N	error	order	error	order	error	order
20	1.41E-02	-	6.84E-04	-	3.40e-5	-
40	4.18E-03	1.76	1.04E-04	2.72	2.82e-6	3.59
80	1.30E-03	1.68	1.55E-05	2.74	2.26e-7	3.64
160	4.24E-04	1.62	2.41E-06	2.69	1.83e-8	3.62
320	1.51E-04	1.49	3.80E-07	2.67	1.49e-9	3.63

Example 5.4. *We consider the following initial condition*

$$\rho_0(x) = \begin{cases} 1 \quad\quad x < 0, \\ 0.25 \; x > 0, \end{cases} \qquad u_0(x) = \begin{cases} 1 \; x < 0, \\ 0 \; x > 0. \end{cases} \tag{5.24}$$

Clearly, the exact solution is

$$(\rho(x,t), u(x,t)) = \begin{cases} (1,1) \quad\quad x < 2t/3, \\ (0.25, 0) \; x > 2t/3, \end{cases}$$

and at $x = \frac{2t}{3}$, the density should be a δ-function. Figure 5.4 shows the numerical density and velocity at $t = 0.5$ using $\mathcal{P}^1$ polynomials. From the figure, we observe the numerical solution capture the profile of the exact solution quite well.

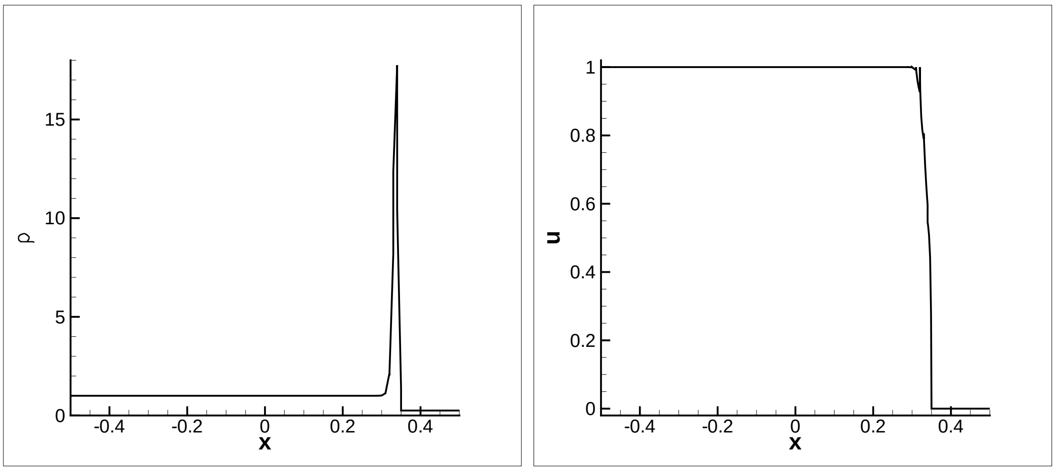

Figure 5.4: Numerical density (left) and velocity (right) at $t = 0.5$ with $\mathcal{P}^1$ polynomials for initial condition (5.24).

Example 5.5. *We consider the following initial condition*

$$\rho_0(x) = 0.5, \quad u_0(x) = \begin{cases} -0.5 & x < -0.5, \\ 0.4 & -0.5 < x < 0, \\ 0.4 - x & 0 < x < 0.8, \\ -0.4 & x > 0.8. \end{cases} \tag{5.25}$$

The exact solution for $t < 1$ is

$$(\rho(x,t), u(x,t)) = \begin{cases} (0.5, -0.5) & x < -0.5 - 0.5t, \\ (0, \text{undefined}) & -0.5 - 0.5t < x < -0.5 + 0.4t, \\ (0.5, 0.4) & -0.5 + 0.4t < x < 0.4t, \\ (\frac{0.5}{1-t}, \frac{0.4-x}{1-t}) & 0.4t < x < 0.8 - 0.4t, \\ (0.5, -0.4) & x > 0.8 - 0.4t. \end{cases}$$

Figure 5.5 shows the numerical density and velocity at $t = 0.5$. From the figure, we can observe some local oscillations near the singularities. This is not surprising as we have not used any limiters other than the bound-preserving ones for the DG scheme.

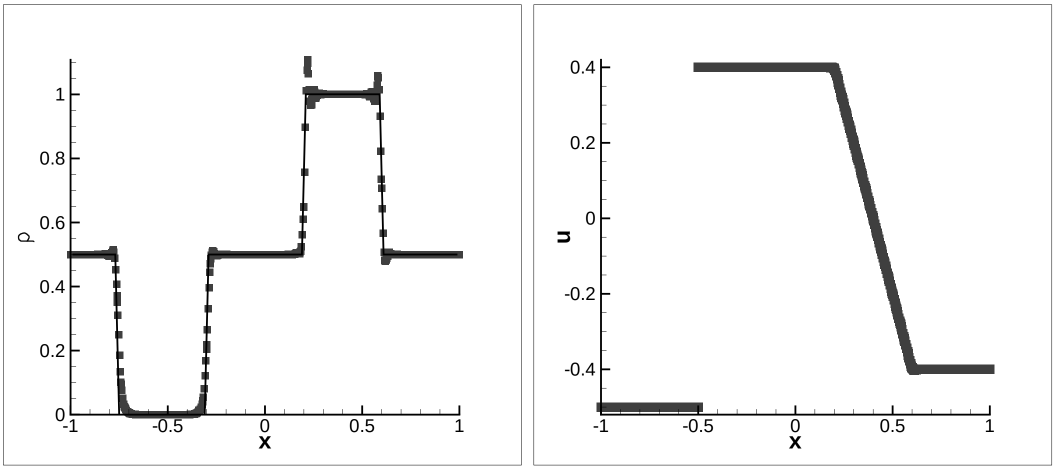

Figure 5.5: Numerical density (left) and velocity (right) at $t = 0.5$ for initial condition (5.25). The solid line shows the exact solution while the symbols show the numerical solution.

Example 5.6. *We consider the following initial condition*

$$\rho_0(x) = 0.5, \quad u_0(x) = \begin{cases} -0.5 \; x < 0, \\ 0.4 \quad x > 0. \end{cases} \tag{5.26}$$

The exact solution is

$$(\rho(x,t), u(x,t)) = \begin{cases} (0.5, -0.5) & x < -0.5t, \\ (0, \text{undefined}) & -0.5t < x < 0.4t, \\ (0.5, 0.4) & x > 0.4t. \end{cases}$$

Figure 5.6 shows the numerical density and velocity at $t = 0.5$. From the figure, we can observe some local oscillations near the singularities.

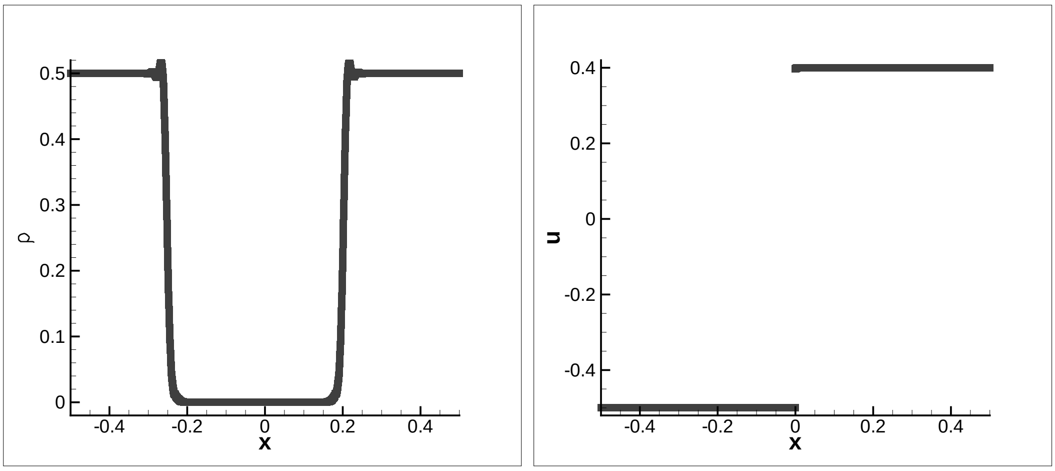

Figure 5.6: Numerical density (left) and velocity (right) at $t = 0.5$ for initial condition (5.26).

Two dimensions

We consider the problem in two dimensions and solve (5.20) with different initial conditions.

Example 5.7. *We consider the following initial condition*

$$\begin{aligned} \rho(x,y,0) &= \rho_0(x+y) = \exp(\sin(x+y)), \\ u(x,y,0) &= u_0(x+y) = \tfrac{1}{3}(\cos(x+y)+2), \\ v(x,y,0) &= v_0(x+y) = \tfrac{1}{3}(\sin(x+y)+2). \end{aligned} \tag{5.27}$$

The exact solution is

$$u(x,y,t) = u_0(z_0), \quad v(x,y,t) = v_0(z_0), \quad \rho(x,y,t) = \frac{\rho_0(z_0)}{1+u_0'(z_0)+v_0'(z_0)},$$

where z_0 is given implicitly by

$$z_0 + t(u_0(z_0) + v_0(z_0)) = x + y.$$

We use the third order SSP multi-step method in time [101] with $\Delta t = 0.01\Delta x^{3/2}$, and test the example by using $\mathcal{P}^k$ polynomials with $k = 1, 2, 3$. Table 5.2 shows the L^2-norm of the error at $t = 0.1$. From the table, we again observe about $(k+0.5)$-th order convergence.

Table 5.2: L^2-norm of the error between the numerical density and the exact density for initial condition (5.27).

	k=1		k=2		k=3	
N	error	order	error	order	error	order
10	0.512	-	0.107	-	3.42E-02	-
20	0.176	1.54	3.12E-02	1.78	3.57E-03	3.26
40	6.48E-02	1.44	8.52E-03	1.87	4.86E-04	2.88
80	2.32E-02	1.48	1.39E-03	2.62	3.97E-05	3.61
160	9.08E-03	1.35	1.92E-04	2.86	3.65E-06	3.45

Example 5.8. *We consider the following initial condition*

$$\rho(x, y, 0) = \frac{1}{100}, \quad (u, v)(x, y, 0) = (-\frac{1}{10}\cos\theta, -\frac{1}{10}\sin\theta), \tag{5.28}$$

where θ is the polar angle.

Since all the particles are moving towards the origin, the density function at $t > 0$ should be a single delta at the origin. Different from Example 5.2 in Section 5.2.2, we can observe only one delta located at the origin by using rectangle mesh as shown in Figure 5.7.

Example 5.9. *We consider the following initial condition*

$$\rho(x, y, 0) = \frac{1}{10}, \quad (u, v)(x, y, 0) = \begin{cases} (-0.25, -0.25) & x > 0, y > 0, \\ (0.25, -0.25) & x < 0, y > 0, \\ (0.25, 0.25) & x < 0, y < 0, \\ (-0.25, 0.25) & x > 0, y < 0. \end{cases} \tag{5.29}$$

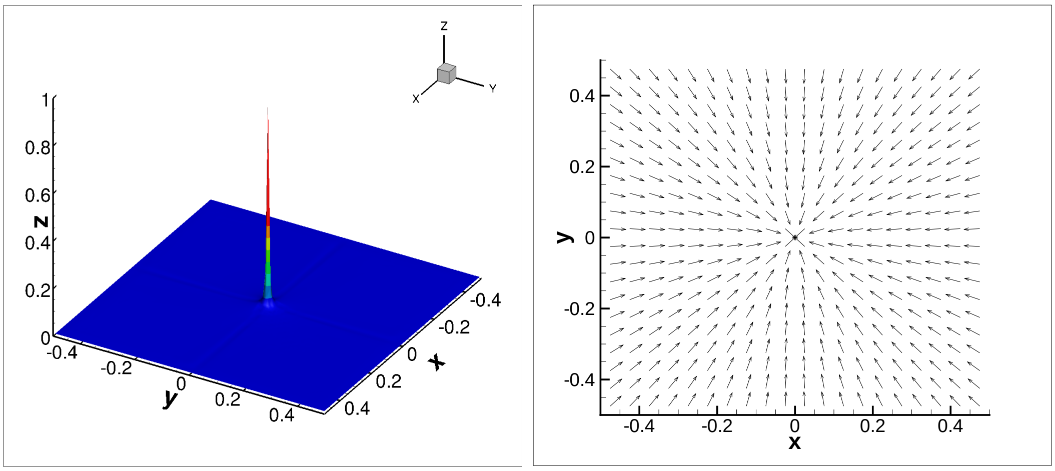

Figure 5.7: Numerical density (left) and velocity field (right) at $t = 0.5$ for initial condition (5.28).

Figure 5.8 shows the numerical density and velocity field at $t = 0.5$. From the figure, we can observe δ-singularities located at the origin and two axes.

Example 5.10. *We consider the following initial condition*

$$\rho(x,y,0) = \frac{1}{100}, \quad (u,v)(x,y,0) = \begin{cases} (\cos\theta, \sin\theta) & r < 0.3, \\ (-\frac{1}{2}\cos\theta, -\frac{1}{2}\sin\theta) & r > 0.3, \end{cases} \tag{5.30}$$

where $r = \sqrt{x^2+y^2}$ and θ is the polar angle.

Figure 5.9 shows the numerical density (contour plot) and velocity field at $t = 0.5$. From the figure, we can observe δ-shocks located on a circle and vacuum inside.

Example 5.11. *We consider the following initial condition*

$$\rho(x,y,0) = 0.5, \quad (u,v)(x,y,0) = \begin{cases} (0.3, 0.4) & x > 0, y > 0, \\ (-0.4, 0.3) & x < 0, y > 0, \\ (-0.3, -0.4) & x < 0, y < 0, \\ (0.4, -0.3) & x > 0, y < 0. \end{cases} \tag{5.31}$$

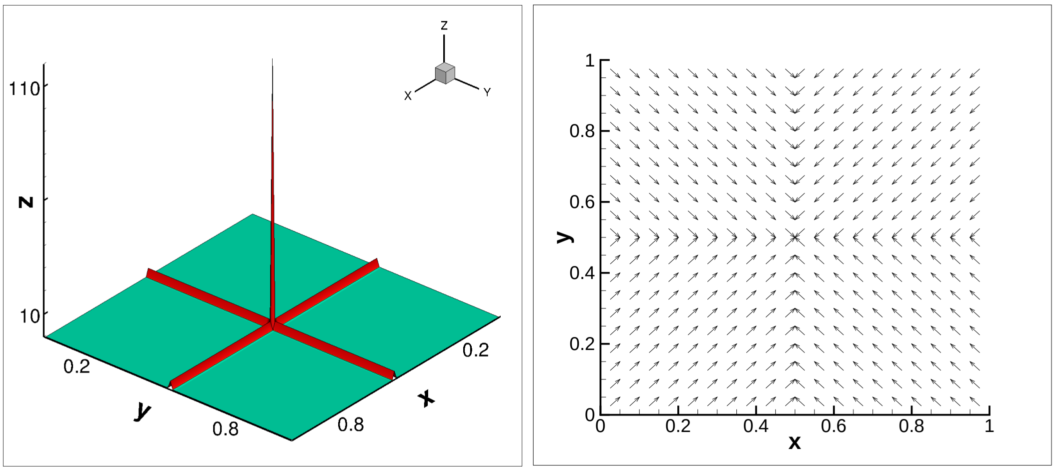

Figure 5.8: Numerical density (left) and velocity field (right) at $t = 0.5$ for initial condition (5.29).

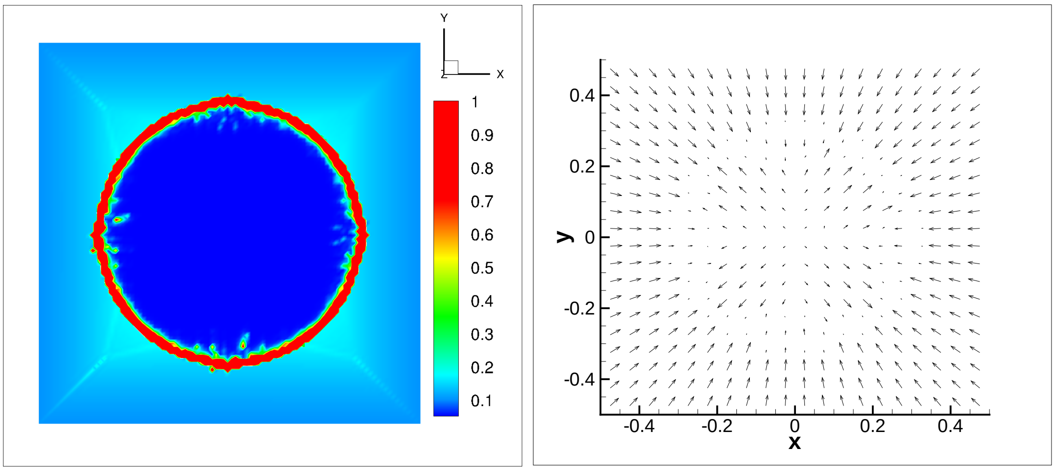

Figure 5.9: Numerical density (left) and velocity field (right) at $t = 0.5$ for initial condition (5.30).

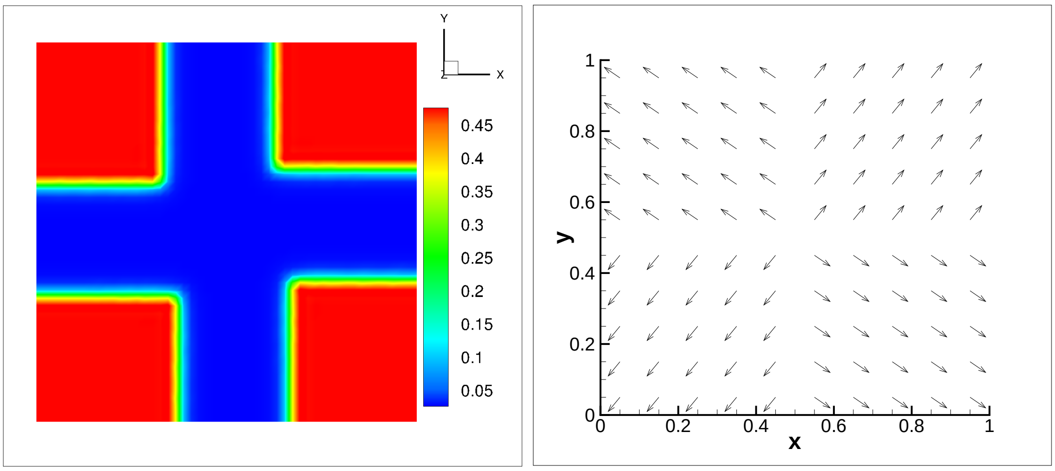

Figure 5.10: Numerical density (left) and velocity field (right) at $t = 0.4$ with $N = 50$ for initial condition (5.31).

Figure 5.10 shows the numerical density (contour plot) and velocity field with $N = 50$ at $t = 0.4$. From the figure, we can observe that the numerical solution approximates the vacuum quite well.

5.4 Concluding remarks

In this chapter, we developed DG methods to solve hyperbolic conservation laws involving δ-singularities. We study Krause's consensus models and pressureless Euler equations to demonstrate the stability and high resolution of the DG approximations. Moreover, numerical experiments show that the scheme is also good for approximations in the presence of vacuum. In future work we will extend DG methods to other equations involving δ-singularities to wider areas of applications.

Part II

Numerical cosmology

In this second part, we will study the Lyα photons transfer in an optically thick medium. Lyα photons have been widely applied to study the physics of luminous objects at various epochs of the universe, such as Lyα emitters, Lyα blob, damped Lyα system, Lyα forest, fluorescent Lyα emission, star-forming galaxies, quasars at high redshifts as well as optical afterglow of gamma ray bursts [49, 43, 38, 69]. The resonant scattering of Lyα photons with neutral hydrogen atoms has a profound effect on the time, space and frequency dependencies of Lyα photons transfer in an optically thick medium. Lyα photons emerging from an optically thick medium would carry rich information of the photon sources and halo surrounding the source of the Lyα photon. The profiles of the emission and absorption of the Lyα radiation are powerful tools to constrain the mass density, velocity, temperature and the fraction of neutral hydrogen of the optically thick medium. Radiation transfer of Lyα photons in an optically thick medium is fundamentally important.

The radiative transfer of Lyα photons in a medium consisting of neutral hydrogen atoms has been extensively studied either analytically or numerically for more than half a century. In [3], the focus was on the numerical approximation of the redistribution function of resonant scattering, and no solution of the integro-differential equation of the radiative transfer has been found. Before that, Field [44] gave the first analytical solution of the integro-differential equation for the case of both medium and source uniformly distributed in the whole space. Analytical solutions of the frequency profile of photons emergent from optically thick halo are found based on the Fokker-Planck (F-P) approximation of the integro-differential equation [51, 76, 37]. Besides the F-P approximation, Monte Carlo (MC) simulations are also popular in solving the transfer of resonant photons (e.g. [75, 122, 107, 110, 70, 82, 114, 115].

However, many important topics cannot be seen with the above-mentioned solutions. Besides the Field's analytical solution, all others are time-independent, and therefore miss the detailed balance relationship of resonant scattering [93] and cannot be used to describe the formation and evolution of the Wouthuysen-Field (W-F) local thermalization of the Lyα photon frequency distribution [113, 44, 45], which is important for the emission and absorption of the hydrogen 21 cm line (e.g. [42]. The rich features of the Lyα photon transfer referring to the W-F local thermalization are fully missed. The F-P equation is based on the Eddington approximation, which assumes that the radiation intensity is a linear function of angular (direction) variable. Therefore, the solutions of the F-P equation do not provide the information of the evolution of the angular distribution of Lyα photons.

Recently, a state-of-the-art numerical method has been introduced to solve the integro-differential equation of the radiative transfer with resonant scattering [83, 84, 85, 87]. The solver is based on the weighted essentially non-oscillatory (WENO) scheme [61]. With the WENO solver, many physical features of the transfer of Lyα photons in an optically thick medium [88, 89, 90], missed in the F-P equation approximations, have been revealed. For instance, the WENO solution shows that the time scale of the formation of the W-F local thermal equilibrium is only about a few hundred times that of the resonant scattering. It also shows that the double peaked frequency profile of the Lyα photon, emerging from an optically thick medium, does not follow the time-independent solutions of the F-P equation [42, 87, 88, 89, 90]. These results indicate the needs of re-visiting problems which have been studied only via the F-P time-independent approximation. In this part, we will use a WENO solver to study the effect of dust and angular distribution of Lyα resonant photons transfer in an optically thick halo.

Chapter 6

WENO solver of transfer equations of resonant photons

6.1 Basic theory

6.1.1 Radiative transfer equation of a dusty halo

The radiative transfer equation of Lyα photons in a spherical halo with dust is given by

$$\begin{aligned}
&\frac{\partial I}{\partial \eta} + \mu\frac{\partial I}{\partial r} + \frac{(1-\mu^2)}{r}\frac{\partial I}{\partial \mu} - \gamma\frac{\partial I}{\partial x} = \\
&\quad -\phi(x;a)I + \int \mathcal{R}(x,\mu,x',\mu';a)I(\eta,r,x',\mu')dx'd\mu'/2 \\
&\quad -\kappa(x)I + A\kappa(x)\int \mathcal{R}^d(x,x';\mu,\mu';a)I(\eta,r,x',\mu')dx'd\mu' + S,
\end{aligned} \tag{6.1}$$

where S is the source and $I(t, r_p, x, \mu)$ is the specific intensity, which is a function of time t, radial coordinate r_p, frequency x and the direction angle, $\mu = \cos\theta$, with respect to the radial vector $\mathbf{r}$.

In (6.1), we use the dimensionless time η defined as $\eta = cn_{\rm HI}\sigma_0 t$ and the dimensionless radial coordinate r defined as $r = n_{\rm HI}\sigma_0 r_p$, where $n_{\rm HI}$ is the numer density of HI, and $\sigma_0/\pi^{1/2}$ is the cross section of HI resonent scatering of Lyα photons at resonant frequency $\nu_0 = 2.46\times10^{15}s^{-1}$. That is, η and r are, respectively, in the units of mean free flight-time and mean free path of photon ν_0 with respect to the resonant scattering without dust scattering and absorption. Without resonant scattering, a signal propagates in the radial direction with the speed of light and the orbit of the signal is $r = \eta + \text{const}$.

$\phi(x, a)$ is the normalized Voigt profile [57] given as

$$\phi(x, a) = \frac{a}{\pi^{3/2}} \int_{-\infty}^{\infty} dy \frac{e^{-y^2}}{(x - y)^2 + a^2}. \tag{6.2}$$

As usual, the photon frequency ν in (6.1) is described by the dimensionless frequency $x \equiv (\nu - \nu_0)/\Delta\nu_D$, and $\Delta\nu_D = \nu_0(v_T/c) = 1.06 \times 10^{11}(T/10^4)^{1/2}$ Hz is the Doppler broadening by the thermal motion $v_T = \sqrt{2k_BT/m}$, T being the gas temperature of the halo. The parameter a in (6.2) is the ratio of the natural to the Doppler broadening. For the Lyα line, $a = 4.7 \times 10^{-4}(T/10^4)^{-1/2}$. The optical depth of Lyα photons with respect to HI resonant scattering is $\tau_s(x) = n_{\rm HI}\sigma(x)dr_p$, where $\sigma(x) = \sigma_0\phi(x, a)$ is the cross section of scattering at ν, and therefore, the dimensionless size of the halo R is equal to the optical depth $\tau_0 = n_{\rm HI}\sigma_0R$.

The re-distribution function $\mathcal{R}(x, \mu, x', \mu'; a)$ of (6.2), the derivation of which is given in Section 6.1.4, gives the probability of a photon absorbed at the frequency x' direction μ', and re-emitted at the frequency x direction μ. It depends on the details of the scattering [54, 56, 58]. If we consider coherent scattering without recoil, the re-distribution function with the Voigt profile is

$$\mathcal{R}(x, \mu, x', \mu'; a) = \int_0^{2\pi} \frac{a}{4\pi^3\beta} \int_{-\infty}^{\infty} e^{-u^2} \left[a^2 + \left(\frac{x + x'}{2} - \alpha u\right)^2\right]^{-1} \exp\left(-\frac{(x - x')^2}{4\beta^2}\right) du d\phi, \tag{6.3}$$

where $H = \sqrt{1 - \mu^2}\sqrt{1 - \mu'^2}\cos\phi + \mu\mu'$, $\alpha = \sqrt{\frac{1+H}{2}}$, and $\beta = \sqrt{\frac{1-H}{2}}$. In the case of $a = 0$, i.e., considering only the Doppler broadening, the re-distribution function is

$$\mathcal{R}(x,\mu,x',\mu') = \int_0^{2\pi} \frac{1}{2\pi^2\sqrt{1-H^2}} \exp\left[-\frac{2^2 - 2xx'H + x'^2}{1-H^2}\right] d\phi, \qquad (6.4)$$

where H is exactly the same as in (6.3). The redistribution function of (6.4) is normalized as

$$\frac{1}{2}\int_{-1}^{1}\int_{-\infty}^{\infty} \mathcal{R}(x,\mu,x',\mu')dx'd\mu' = \phi(x,0) = \pi^{-1/2}e^{-x^2}.$$

With this normalization, the total number of photons is conserved in the evolution described by (6.1). That is, the destruction processes of Lyα photons, such as the two-photon process [105, 80], are ignored in (6.3). The recoil of atoms is not considered in (6.3) or (6.4).

The absorption and scattering of dust are described by the term $\kappa(x)I$ in (6.1), where $\kappa(x) = \sigma_d/\sigma_0$, which is of the order of $10^{-8}(T/10^4)^{1/2}$ [40, 39]. The term with A in (6.1) describes albedo, i.e. $A \equiv \sigma_s/\sigma_d$, where σ_s is the cross section of dust scattering and σ_d is the effective cross-section per hydrogen atom, which describes the absorption and scattering of dust. Generally, A lies approximately between 0.3 and 0.4 [81, 112]. With dust, the optical depth is given by

$$\tau(x) = \tau_0\phi(x,a) + \tau_d(x) \qquad (6.5)$$

where the dust optical depth $\tau_d(x) = n_{\rm HI}\sigma_d(x)R$. This is equal to assuming that dust is uniformly distributed in IGM. The effects of inhomogeneous density distributions of dust [77, 48] will not be studied.

Since dust generally is much heavier than a single atoms, the recoil of dust particles can be neglected when colliding with a photon. Under this “heavy dust” approximation, photons do not change their frequency during the collision with dust. The redistribution function of dust $\mathcal{R}^d$ is independent of x and x', and is simply given by a phase function as

$$\mathcal{R}^d(\mu,\mu') = \frac{1}{4\pi}\int_0^{2\pi} d\phi' \frac{1-g^2}{(1+g^2-2g\bar{\mu})^{3/2}} = \sum_{l=0}^{\infty} \frac{(2l+1)}{2} g^l P_l(\mu) P_l(\mu'), \tag{6.6}$$

where $\bar{\mu} = \mu\mu' + \sqrt{(1-\mu^2)(1-\mu'^2)}\cos\phi'$ and P_l is the Legendre function. The factor g in (6.6) is the asymmetry parameter. For isotropic scattering, $g = 0$. The cases of $g = +1$ and -1 correspond to complete forward and backward scattering, respectively. Generally, the factor g is a function of the wavelength. For the Lyα photon, we will take $g = 0.73$ for realistic dust scattering [73]. The integral of (6.6) is performed in Section 6.1.5.

In (6.1) we neglect the effect of collision transition from $H(2p)$ state to $H(2s)$ state, which can significantly affect the escape of Lyα photons when the HI column density is higher than 10^{21} cm^{-2} and dust absorption is very small [76]. This generally is out of the parameter range used below. We are also not considering the effects of bulk motion of the medium of halos (e.g. [104, 114]).

In (6.1), the term with the parameter γ is due to the expansion of the universe. If $n_{\rm H}$ is equal to the mean of the number density of cosmic hydrogen, we have $\gamma = \tau_{GP}^{-1}$, and τ_{GP} is the Gunn-Peterson optical depth. Since the Gunn-Peterson optical depth is of the order of 10^6 at high redshift (e.g. [89]), the parameter γ is of the order of $10^{-5} - 10^{-6}$. Therefore, if the optical depth of halos is less than or equal to 10^6, the term with γ in (6.1) can be ignored.

6.1.2 Integrated redistribution function

In general, it is difficult to solve the transfer equation for noncoherent scattering. Therefore, the scattering is assumed to be isotropic and we need to integrate $\mathcal{R}(x,\mu,x',\mu';a)$ over the angular direction. Denote $\mathcal{R}(x,x';a)$ as the angular averaged re-distribution function. It gives the probability of a photon absorbed at the frequency x', and re-emitted at the frequency x. If we consider coherent scattering without recoil, the angular-averaged re-distribution function with the Voigt profile can be written as,

$$\mathcal{R}(x,x';a) = \tag{6.7}$$
$$\frac{1}{\pi^{3/2}}\int_{|x-x'|/2}^{\infty} e^{-u^2}\left[\tan^{-1}\left(\frac{x_{\rm min}+u}{a}\right)-\tan^{-1}\left(\frac{x_{\rm max}-u}{a}\right)\right]du$$

where $x_{\rm min}=\min(x,x')$ and $x_{\rm max}=\max(x,x')$. In the case of $a=0$, i.e. including only the Doppler broadening, the re-distribution function is

$$\mathcal{R}(x,x') = \frac{1}{2}{\rm erfc}[\max(|x|,|x'|)]. \tag{6.8}$$

The angular-averaged re-distribution function of (6.8) is normalized as $\int_{-\infty}^{\infty}\mathcal{R}(x,x')dx' = \phi(x,0) = \pi^{-1/2}e^{-x^2}$. With this normalization, the total number of photons is also conserved in the evolution described by (6.1).

6.1.3 Eddington approximation

Equation (6.6) indicates that the transfer equation (6.1) can be solved with the Legendre expansion $I(\eta,r,x,\mu)=\sum_l I_l(\eta,r,x)P_l(\mu)$. If we take only the first two terms, $l=0$ and 1, it is the Eddington approximation as

$$I(\eta,r,x,\mu)\simeq J(\eta,r,x)+3\mu F(\eta,r,x) \tag{6.9}$$

where

$$J(\eta,r,x)=\frac{1}{2}\int_{-1}^{+1}I(\eta,r,x,\mu)d\mu, \quad F(\eta,r,x)=\frac{1}{2}\int_{-1}^{+1}\mu I(\eta,r,x,\mu)d\mu. \tag{6.10}$$

They are, respectively, the angularly averaged specific intensity and flux. Defining $j = r^2 J$ and $f = r^2 F$, (6.1) yields the equations of j and f as

$$\frac{\partial j}{\partial \eta} + \frac{\partial f}{\partial r} = -(1-A)\kappa j - \phi(x;a)j + \int \mathcal{R}(x,x';a)jdx' + \gamma\frac{\partial j}{\partial x} + r^2 S, \quad (6.11)$$

$$\frac{\partial f}{\partial \eta} + \frac{1}{3}\frac{\partial j}{\partial r} = -(1-Ag)\kappa f + \gamma\frac{\partial f}{\partial x} - \phi(x;a)f + \frac{2}{3}\frac{j}{r}. \quad (6.12)$$

The mean intensity $j(\eta, r, x)$ describes the x photons trapped in the position r at time η by the resonant scattering, while the flux $f(\eta, r, x)$ describes the photons in transit.

For spherical halo with a central source, the term S of (6.1) can be replaced by a boundary condition of $I(\eta, r, x, \mu)$ at $r = 0$. If the angular distribution of photons is independent of photon's frequency, we have

$$r^2 I(\eta, r, x, \mu)|_{r\to 0} = S_0 T(\eta)\Theta(\mu)\phi(x). \quad (6.13)$$

where the functions $T(\eta)$, $\Theta(\mu)$, and $\phi(x)$ describe, respectively, the time-dependence, angular- and frequency-distributions of photons of the source. In this case, the source of (6.13) can be replaced by a boundary condition at $r = 0$ as

$$f(\eta, 0, x) = S_0 T(\eta)\phi(x)\frac{1}{2}\int \mu\Theta(\mu)d\mu = S_0\phi_s(x). \quad (6.14)$$

For example, if we take the boundary condition at $r = 0$ to be

$$r^2 I(\eta, r, x, \mu)|_{r\to 0} = \begin{cases} 6\mu\pi^{-1/2}e^{-x^2}, & \mu > 0, \\ 0, & \mu < 0. \end{cases} \quad (6.15)$$

With (6.15) one can find I from (6.1), and then find j and f via (6.10). Therefore, the corresponded boundary condition of (6.14) is

$$f(\eta, r = 0, x) = \pi^{-1/2}e^{-x^2}. \quad (6.16)$$

We will use these two boundary conditions in Chapter 8.

In (6.13) and (6.14), S_0 is the intensity. Since (6.1), (6.11) and (6.12) are linear, the solutions of $j(x)$, $f(x)$ and $I(x)$ for given $S_0 = S$ are equal to $Sj_1(x)$, $Sf_1(x)$ and $SI_1(x)$, where $j_1(x)$, $f_1(x)$ and $I_1(x)$ are the solutions of $S_0 = 1$.

On the outside of the halo, $r > R$, no photons propagate in the direction $\mu < 0$. The boundary condition at $r = R$ of (6.1) should be

$$I(\eta, R, x, \mu) = 0, \quad \mu < 0. \tag{6.17}$$

For (6.11), we have $\int_0^{-1} \mu I(\eta, R, x, \mu) d\mu = 0$ [108], and the boundary condition is then

$$j(\eta, R, x) = 2f(\eta, R, x). \tag{6.18}$$

If the source becomes to emit photon at $t = 0$, the initial condition should be

$$I(0, r, x, \mu) = 0, \tag{6.19}$$

for (6.1), and

$$j(0, r, x) = f(0, r, x) = 0, \tag{6.20}$$

for (6.11).

6.1.4 Re-distribution function

In this section, we proceed to the re-distribution function $\mathcal{R}(x, \mu, x', \mu'; a)$ in (6.1). We consider isotropic scattering and the case of $a = 0$. The re-distribution function

$$\mathcal{R}(x, \mathbf{n}, x', \mathbf{n}') = \frac{1}{\pi \sin \alpha} \exp\left[-\frac{x^2 - 2xx' \cos\alpha + x'^2}{\sin^2 \alpha} \right]$$

gives the probability that a photon with frequency x' and direction $\mathbf{n}'$ within an element of solid angle $d\omega'$ is absorbed and re-emitted with frequency x and direction

$\mathbf{n}$ in $d\omega$ [56], where α is the angle between $\mathbf{n}$ and $\mathbf{n}'$. Choose x-axis such that $\mathbf{n}'$ lies in the xz-plane, then

$$\mathbf{n}' = (\sin\theta', 0, \cos\theta')$$

and

$$\mathbf{n} = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta),$$

where ϕ is the azimuthal angle, $\mu = \cos\theta$ and $\mu' = \cos\theta'$. With the above notation, $d\omega = \frac{1}{4\pi} d\phi d\mu$ and

$$\cos\alpha = \mathbf{n}' \cdot \mathbf{n} = \sin\theta\sin\theta'\cos\phi + \cos\theta\cos\theta'.$$

Integrating over ϕ, we obtain

$$\mathcal{R}(x,\mu,x',\mu') = \int_0^{2\pi} \frac{1}{2\pi^2\sqrt{1-H^2}} \exp\left[-\frac{2^2 - 2xx'H + x'^2}{1-H^2}\right] d\phi,$$

where $H = \sqrt{1-\mu^2}\sqrt{1-\mu'^2}\cos\phi + \mu\mu'$. If we consider $a \neq 0$, and follow the same line above, we obtain

$$\mathcal{R}(x,\mu,x',\mu';a) = \int_0^{2\pi} \frac{a}{4\pi^3\beta} \int_{-\infty}^{\infty} e^{-u^2} \left[a^2 + \left(\frac{x+x'}{2} - \alpha u\right)^2\right]^{-1} \exp\left(-\frac{(x-x')^2}{4\beta^2}\right) du d\phi,$$

where $H = \sqrt{1-\mu^2}\sqrt{1-\mu'^2}\cos\phi + \mu\mu'$, $\alpha = \sqrt{\frac{1+H}{2}}$, and $\beta = \sqrt{\frac{1-H}{2}}$. We can verify numerically that the angular averaged re-distribution function is exactly the same as the one obtained by Hummer [56], i.e.

$$\frac{1}{2}\int_{-1}^{1} \mathcal{R}(x,\mu,x',\mu';a) d\mu' = \frac{1}{2}\int_{-1}^{1} \mathcal{R}(x,\mu,x',\mu';a) d\mu = \mathcal{R}(x,x';a).$$

6.1.5 Integral of the phase function

Equation (6.6) can be rewritten as

$$\mathcal{R}^d(\mu,\mu') = \frac{1}{4\pi}\int_0^{2\pi} d\phi' \frac{1-g^2}{|\mathbf{I}-g\mathbf{I}'|^{\frac{3}{2}}} \tag{6.21}$$

where $\mathbf{I}$ and $\mathbf{I}'$ are unit vector on the direction of polar angle θ and θ', and azimuth angle ϕ and ϕ', respectively. That is $\mathbf{I}\cdot\mathbf{I} = \mathbf{I}'\cdot\mathbf{I}' = 1$ and $\mathbf{I}\cdot\mathbf{I}' = \cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi')$, and $\mu = \cos\theta$, $\mu' = \cos\theta$. We have

$$\frac{d}{dg}\frac{1}{|\mathbf{I}-g\mathbf{I}'|^{1/2}} = \frac{1-g^2}{2g|\mathbf{I}-g\mathbf{I}'|^{3/2}} - \frac{1}{2g|\mathbf{I}-g\mathbf{I}'|^{1/2}},$$

and therefore,

$$\frac{1-g^2}{|\mathbf{I}-g\mathbf{I}'|^{3/2}} = 2g\frac{d}{dg}\frac{1}{|\mathbf{I}-g\mathbf{I}'|^{1/2}} + \frac{1}{|\mathbf{I}-g\mathbf{I}'|^{1/2}}. \tag{6.22}$$

The expansion with Legendre functions $P_l(\cos\gamma)$ gives

$$\frac{1}{|\mathbf{I}-g\mathbf{I}'|^{1/2}} = \sum_{l=0}^{\infty} g^l P_l(\cos\gamma), \tag{6.23}$$

and then

$$\frac{1-g^2}{|\mathbf{I}-g\mathbf{I}'|^{3/2}} = \sum_{l=1}^{\infty} 2lg^l P_l(\cos\gamma) + \sum_{l=0}^{\infty} g^l P_l(\cos\gamma). \tag{6.24}$$

Since $\cos\gamma = \cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi')$, we have the following identity for the Legendre function $P_l(\cos\gamma)$ as

$$P_l(\cos\gamma) = P_l(\cos\theta)P_l(\cos\theta') + 2\sum_{m=1}^{m=l}\frac{(l-m)!}{(l+m)!}P_l^m(\cos\theta)P_l^m(\cos\theta')\cos[m(\phi-\phi')]. \tag{6.25}$$

The integral of ϕ' in (6.21) kills the second term of (6.25), we have

$$\begin{aligned}\mathcal{R}^d(\mu,\mu') &= \frac{1}{4\pi}2\pi\left[\sum_{l=1}^{\infty}2lg^lP_l(\cos\theta)P_l(\cos\theta')+\sum_{l=0}^{\infty}g^lP_l(\cos\theta)P_l(\cos\theta')\right] \quad (6.26)\\ &= \frac{1}{2}\left[\sum_{l=1}^{\infty}2lg^lP_l(\mu)P_l(\mu')+\sum_{l=0}^{\infty}g^lP_l(\mu)P_l(\mu')\right].\end{aligned}$$

Using the orthogonal relation $\int_{-1}^{1}P_l(\mu)P_{l'}(\mu)d\mu=\frac{2}{2l+1}\delta_{l,l'}$, we have

$$R_0(g)=\frac{1}{2}\int_{-1}^{1}d\mu\int_{-1}^{1}d\mu' R^d(\mu,\mu')=1, \quad (6.27)$$

for which only the term $l=0$ in (6.26) has contribution. Similarly,

$$R_1(g)=\frac{1}{2}\int_{-1}^{1}d\mu\int_{-1}^{1}d\mu'\mu R^d(\mu,\mu')=\frac{1}{2}\int_{-1}^{1}d\mu\int_{-1}^{1}d\mu'\mu' R^d(\mu,\mu')=0, \quad (6.28)$$

$$R_2(g)=\frac{1}{2}\int_{-1}^{1}d\mu\int_{-1}^{1}d\mu'\mu\mu' R^d(\mu,\mu')=\frac{g}{3}. \quad (6.29)$$

These results are used in deriving (6.11) and (6.12).

6.2 Numerical algorithm for equation (6.1)

We solve (6.1) with initial and boundary conditions (6.13), (6.17) and (6.19). For simplicity, we ignore the effect of dust (i.e. $\kappa(x)=0$). Our computational domain is $(r,x,\mu)\in[0,r_{\max}]\times[x_{\text{left}},x_{\text{right}}]\times[-1,1]$, where $r_{\max},x_{\text{left}}$ and x_{right} are chosen such that the solution vanishes to zero outside the boundaries. We choose mesh sizes with grid refinement tests to ensure proper numerical resolution. In the following, we describe numerical techniques involved in our algorithm, including approximations to spatial derivatives, numerical integration, numerical boundary condition and time evolution.

6.2.1 Conservation law

To use the WENO algorithm, we first rewrite (6.1) into the form of a conservation law. Noticing the boundary condition (6.13), we define $I' = r^2 I$, so(6.1) becomes

$$\begin{aligned}&\frac{\partial I'}{\partial \eta} + \mu \frac{\partial I'}{\partial r} + \frac{1}{r}\frac{\partial (1-\mu^2) I'}{\partial \mu} - \gamma \frac{\partial I'}{\partial x} = \\ &\quad -\phi(x;a) I' + \int \mathcal{R}(x,\mu,x',\mu';a) I'(\eta,r,x',\mu') dx' d\mu'/2 + r^2 S. \qquad (6.30)\end{aligned}$$

For simplicity, we drop the prime, and use $I(\eta,r,x,\mu)$ for $I'(\eta,r,x,\mu)$ below.

6.2.2 The WENO algorithm: approximations to the spatial derivatives

The spatial derivative terms in (6.30) are approximated by a fifth-order finite difference WENO scheme.

We first give the WENO reconstruction procedure for approximating $\frac{\partial I}{\partial x}$,

$$\frac{\partial I(\eta^n, r_i, x_j, \mu_k)}{\partial x} \approx \frac{1}{\Delta x}(\hat{h}_{j+1/2} - \hat{h}_{j-1/2})$$

with fixed $\eta = \eta^n$, $r = r_i$ and $\mu = \mu_k$. The numerical flux $\hat{h}_{j+1/2}$ is obtained by the fifth-order WENO approximation in an upwind fashion, because the wind direction is fixed (negative). Denote

$$h_j = I(\eta^n, r_i, x_j, \mu_k), \qquad j = -2, -1, \cdots, N_x + 3$$

with fixed n, i and k. The numerical flux from the WENO procedure is obtained by

$$\hat{h}_{j+1/2} = \omega_1 \hat{h}^{(1)}_{j+1/2} + \omega_2 \hat{h}^{(2)}_{j+1/2} + \omega_3 \hat{h}^{(3)}_{j+1/2}, \qquad (6.31)$$

where $\hat{h}^{(m)}_{j+1/2}$ are the three third-order fluxes on three different stencils given by

$$\begin{aligned}
\hat{h}^{(1)}_{j+1/2} &= -\frac{1}{6}h_{j-1} + \frac{5}{6}h_j + \frac{1}{3}h_{j+1}, \\
\hat{h}^{(2)}_{j+1/2} &= \frac{1}{3}h_j + \frac{5}{6}h_{j+1} - \frac{1}{6}h_{j+2}, \\
\hat{h}^{(3)}_{j+1/2} &= \frac{11}{6}h_{j+1} - \frac{7}{6}h_{j+2} + \frac{1}{3}h_{j+3},
\end{aligned}$$

and the nonlinear weights ω_m are given by

$$\omega_m = \frac{\breve{\omega}_m}{\sum_{l=1}^{3}\breve{\omega}_l}, \quad \breve{\omega}_l = \frac{\gamma_l}{(\epsilon + \beta_l)^2},$$

where ϵ is a parameter to avoid the denominator to become zero and is taken as $\epsilon = 10^{-8}$. The linear weights γ_l are given by

$$\gamma_1 = \frac{3}{10}, \quad \gamma_2 = \frac{3}{5}, \quad \gamma_3 = \frac{1}{10},$$

and the smoothness indicators β_l are given by

$$\begin{aligned}
\beta_1 &= \frac{13}{12}(h_{j-1} - 2h_j + h_{j+1})^2 + \frac{1}{4}(h_{j-1} - 4h_j + 3h_{j+1})^2, \\
\beta_2 &= \frac{13}{12}(h_j - 2h_{j+1} + h_{j+2})^2 + \frac{1}{4}(h_j - h_{j+2})^2, \\
\beta_3 &= \frac{13}{12}(h_{j+1} - 2h_{j+2} + h_{j+3})^2 + \frac{1}{4}(3h_{j+1} - 4h_{j+2} + h_{j+3})^2.
\end{aligned}$$

Similarly, we give the WENO procedure in approximating $\frac{\partial(1-\mu^2)I}{\partial\mu}$,

$$\frac{\partial(1-\mu_j^2)I(\eta^n, r_i, x_k, \mu_j)}{\partial\mu} \approx \frac{1}{\Delta\mu}(\hat{h}_{j+1/2} - \hat{h}_{j-1/2})$$

with fixed $\eta = \eta^n$, $r = r_i$ and $x = x_k$. The numerical flux $\hat{h}_{j+1/2}$ is also obtained by the fifth-order WENO approximation in an upwind fashion, however the wind direction here is positive, opposite from that of $\frac{\partial I}{\partial x}$. Denote

$$h_j = (1-\mu_j^2) I(\eta^n, r_i, x_k, \mu_j), \qquad j = -3, -2, \cdots, N_\mu + 2$$

with fixed n, i and k. The numerical flux from the WENO procedure is obtained by

$$\hat{h}_{j+1/2} = \omega_1 \hat{h}^{(1)}_{j+1/2} + \omega_2 \hat{h}^{(2)}_{j+1/2} + \omega_3 \hat{h}^{(3)}_{j+1/2}, \tag{6.32}$$

where $\hat{h}^{(m)}_{j+1/2}$ are the three third-order fluxes on three different stencils given by

$$\begin{aligned}
\hat{h}^{(1)}_{j+1/2} &= -\frac{1}{6}h_{j+2} + \frac{5}{6}h_{j+1} + \frac{1}{3}h_j, \\
\hat{h}^{(2)}_{j+1/2} &= \frac{1}{3}h_{j+1} + \frac{5}{6}h_j - \frac{1}{6}h_{j-1}, \\
\hat{h}^{(3)}_{j+1/2} &= \frac{11}{6}h_j - \frac{7}{6}h_{j-1} + \frac{1}{3}h_{j-2},
\end{aligned}$$

and the nonlinear weights ω_m are given as

$$\omega_m = \frac{\breve{\omega}_m}{\sum_{l=1}^3 \breve{\omega}_l}, \quad \breve{\omega}_l = \frac{\gamma_l}{(\epsilon + \beta_l)^2},$$

where ϵ is taken as $\epsilon = 10^{-8}$. The linear weights γ_l are also given by

$$\gamma_1 = \frac{3}{10}, \quad \gamma_2 = \frac{3}{5}, \quad \gamma_3 = \frac{1}{10},$$

and the smoothness indicators β_l are given by

$$\begin{aligned}
\beta_1 &= \frac{13}{12}(h_{j+2} - 2h_{j+1} + h_j)^2 + \frac{1}{4}(h_{j+2} - 4h_{j+1} + 3h_j)^2, \\
\beta_2 &= \frac{13}{12}(h_{j+1} - 2h_j + h_{j-1})^2 + \frac{1}{4}(h_{j+1} - h_{j-1})^2, \\
\beta_3 &= \frac{13}{12}(h_j - 2h_{j-1} + h_{j-2})^2 + \frac{1}{4}(3h_j - 4h_{j-1} + h_{j-2})^2.
\end{aligned}$$

In the end, we approximate the r-derivative in (6.30), following the reconstruction procedures mentioned above. However, we need to check the wind direction at the r-boundary of each cell. When $\mu > 0$, the wind direction is positive, and we use (6.32) to approximate the numerical flux, while when $\mu < 0$, we use equation (6.31).

6.2.3 High order numerical integration

The integration of the resonance scattering term is calculated by a fifth order quadrature [100]

$$\int_{\mu_{left}}^{\mu_{right}} f(\mu)d\mu = \Delta\mu \sum_{k=1}^{N_\mu} \omega_k f(\mu_k) + O(\Delta\mu^5),$$

where $\mu_k = \mu_{left} + (k - \frac{1}{2})d\mu$ and the weights are defined as,

$$\omega_1 = \frac{6463}{5760}, \quad \omega_2 = \frac{1457}{1920}, \quad \omega_3 = \frac{741}{640}, \quad \omega_4 = \frac{5537}{5760},$$
$$\omega_{N_\mu-3} = \frac{5537}{5760}, \quad \omega_{N_\mu-2} = \frac{741}{640}, \quad \omega_{N_\mu-1} = \frac{1457}{1920}, \quad \omega_{N_\mu} = \frac{6463}{5760},$$

and $\omega_k = 1$ otherwise.

6.2.4 Numerical boundary condition

Following Carrillo et al. [20], at $\mu = -1$ and $\mu = 1$, we take the boundary conditions as, for $\mu > 0$,

$$\begin{aligned} I(\eta, r, x, -1-\mu) &= I(\eta, r, x, -1+\mu), \\ I(\eta, r, x, 1+\mu) &= I(\eta, r, x, 1-\mu), \end{aligned}$$

motivated by the physical meaning of μ as the cosine of the angle to the $z-$axis. We also explicitly impose $\hat{h}_{\frac{1}{2}} = \hat{h}_{N_\mu+\frac{1}{2}} = 0$ for the first and last numerical fluxes in order to enforce conservation of mass.

6.2.5 Time Evolution

To evolve in time, we use the third-order TVD Runge-Kutta time discretization [102]. For system of ODEs $u_t = L(u)$, the third order Runge-Kutta method is

$$\begin{aligned} u^{(1)} &= u^n + \Delta\tau L(u^n, \tau^n), \\ u^{(2)} &= \frac{3}{4}u^n + \frac{1}{4}(u^{(1)} + \Delta\tau L(u^{(1)}, \tau^n + \Delta\tau)), \\ u^{n+1} &= \frac{1}{3}u^n + \frac{2}{3}(u^{(2)} + \Delta\tau L(u^{(2)}, \tau^n + \frac{1}{2}\Delta\tau)). \end{aligned}$$

6.3 Numerical algorithm for equations (6.11) and (6.12)

We will solve (6.11) and (6.12) with boundary and initial conditions (6.14), (6.18) and (6.20) by using the WENO solver.

To solve (6.11) and (6.12) as a system, our computational domain is $(r, x) \in [0, r_{\max}] \times [x_{\text{left}}, x_{\text{right}}]$. As mentioned in the Section 6.2, $r_{\max}, x_{\text{left}}$ and x_{right} are chosen such that the solution vanishes outside the boundaries. In the following, we describe numerical techniques involved in our algorithm, including the characteristic decomposition, and numerical boundary condition. All other algorithms can be found in Section 6.2.

6.3.1 Characteristic decomposition

We consider the WENO reconstruction procedure for approximating the r-derivatives only. We need to perform the WENO procedure based on a characteristic decomposition. To accomplish this, we write the left-hand side of (6.11) and (6.12) as

$$\mathbf{u}_t + A\mathbf{u}_r,$$

where $\mathbf{u} = (j, f)^T$ and

$$A = \begin{pmatrix} 0 & 1 \\ \frac{1}{3} & 0 \end{pmatrix}$$

is a constant matrix. To perform the characteristic decomposition, we first compute the eigenvalues, the right eigenvectors and the left eigenvectors of A and denote them by Λ, R and R^{-1}. We then project $\mathbf{u}$ to the local characteristic fields $\mathbf{v}$ with $\mathbf{v} = R^{-1}\mathbf{u}$. Now $\mathbf{u}_t + A\mathbf{u}_r$ of the original system is decoupled as two independent equations as $\mathbf{v}_t + \Lambda\mathbf{v}_r$. We approximate the derivative $\mathbf{v}_r$ component by component, each with the correct upwind direction, with the WENO reconstruction procedure similar to the procedure described in Section 6.2. In the end, we transform $\mathbf{v}_r$ back to the physical space by $\mathbf{u}_r = R\mathbf{v}_r$. We refer the readers to [29] for more implementation details.

6.3.2 Numerical Boundary Condition

To implement the boundary condition (6.18), we also need to perform a characteristic decomposition as discussed above. Using the same notation as before, we project $\mathbf{u}$ to the local characteristic fields $\mathbf{v}$ with $\mathbf{v} = R^{-1}\mathbf{u}$. Denote $\mathbf{v} = (v_1, v_2)^T$, now $\mathbf{u}_t + A\mathbf{u}_r$ of the original system is decoupled to two independent scalar operators given by

$$\frac{\partial v_1}{\partial t} + \lambda_1 \frac{\partial v_1}{\partial r}; \qquad \frac{\partial v_2}{\partial t} + \lambda_2 \frac{\partial v_2}{\partial r}$$

where $\lambda_1 = \frac{\sqrt{3}}{3}$ and $\lambda_2 = -\frac{\sqrt{3}}{3}$. The characteristic line starting from the boundary $r = r_{\max}$ for the first equation is pointing outside the computational domain while the one for the second equation is pointing inside. For well-posedness of our system, we need to impose the boundary condition there as

$$v_2 = \alpha v_1 + \beta$$

with constants α and β. We can calculate the values of α and β based on equation (6.18) and the left and right eigenvectors of A. For example, if we take

$$R = \begin{pmatrix} \frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix},$$

we recover that $\alpha = 7 - 4\sqrt{3}$ and $\beta = 0$. We use extrapolation to obtain the value of v_1 and then compute the value v_2. In the end, we transfer $\mathbf{v}$ back to the physical space by $\mathbf{u} = R\mathbf{v}$.

Chapter 7

Effect of dust on Lyα photon transfer in optically thick halo

We will investigate, in this chapter, the effects of the dust on the Lyα photons transfer in an optically thick medium. Dust can be produced at epochs of low and moderate redshifts, and even at redshift as high as 6 [106]. Absorption and scattering of dust have been used to explain the observations on Lyα emission and absorption [59], such as the escaping fraction of Lyα photons [52, 53, 11]; the redshift-dependence of the ratio between Lyα emitters and Lyman Break galaxies [109]; and the "evolution" of the double-peaked profile [71].

However, it is still unclear whether the time scale of a photon escaping from optically thick halo will be increasing (or decreasing) when the halo is dusty. It is also unclear whether the effects of dust absorption can be estimated by the random walk picture [50]. As for the dust effect on the double-peaked profile, the current results given by different studies seem to be contradictory: some claims that the dust absorption leads to the narrowing of the double-peaked profile [71], while others conclude that the width between the two peaks apparently should be increasing due to the dust absorption [110]. We will focus on these basic problems, and examine them with the solution of the integro-differential equation of radiative transfer.

7.1 Basic theory

We study the radiative transfer (6.11) and (6.12) with boundary and initial condition (6.14), (6.18) and (6.20).

7.1.1 Dust models

We consider three models of the dust as follows:

I. pure scattering, $A = 1,\ g = 0.73$: dust causes only anisotropic scattering, but no absorption;

II. scattering and absorption, $A = 0.32,\ g = 0.73$: dust causes both absorption and anisotropic scattering.

III. pure absorption, $A = 0$: dust causes only absorption, but no scattering;

Models I and III do not occur in reality. They are, however, helpful to reveal the effects of pure scattering and absorption on the radiative transfer.

Since $\kappa(x)$ is on the order of 10^{-8}, its effect will be significant only for halos with optical depth $\tau_0 \geq 10^6$, and ignorable for $\tau_0 \leq 10^5$. To illustrate the dust effect, we use halos of $R = \tau_0 \leq 10^4$, and take larger κ to be $\simeq 10^{-4} - 10^{-2}$. We consider below only the case of grey dust, i.e. κ is independent of frequency x. This certainly is not realistic dust. Yet, the frequency range given in the solution below mostly are in the range $|x| < 4$. Therefore, the approximation of grey dust would be proper if the cross section of dust is not strongly frequency dependent in the range $|x| < 4$.

7.1.2 Numerical example: Wouthuysen-Field thermalization

As the first example of numerical solutions, we show the Wouthuysen-Field (W-F) effect, which requires that the distribution of Lyα photons in the frequency space should be thermalized near the resonant frequency ν_0. The W-F effect illustrates the difference between the analytical solutions of the Fokker-Planck approximation and that of equations (6.11) and (6.12). The former can not show the local thermalization [76], while the latter can [88]. All problems related to the W-F local thermal equilibrium should be studied with the integro-differential equation (6.1).

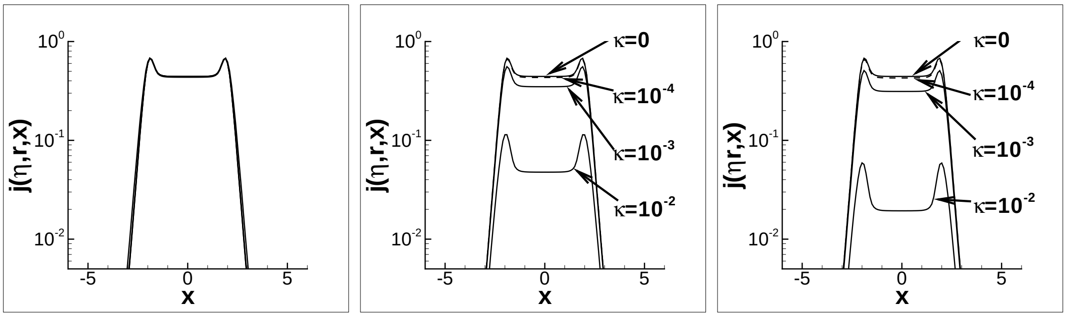

Figure 7.1: The mean intensity $j(\eta, r, x)$ at $\eta = 500$ and $r = 100$ for dust models I (left panel), II (middle panel) and III (right panel). The source is $S_0 = 1$ and $\phi_s(x) = (1/\sqrt{\pi})e^{-x^2}$. The parameter $a = 10^{-3}$. In each panel, κ is taken to be 0, 10^{-4}, 10^{-3} and 10^{-2}.

Figure 7.1 presents a solution of the mean intensity $j(\eta, r, x)$ at time radial $\eta =$ 500 coordinate $r = 10^2$ for halo with size $R \gg r = 10^2$. The three panels correspond to dust models I (left panel), II (middle panel) and III (right panel). The source is taken to have a Gaussian profile $\phi_s(x) = (1/\sqrt{\pi})e^{-x^2}$ and unit intensity $S_0 = 1$. The solutions of Figure 7.1 actually are independent of R, if $R \gg 10^2$. The intensity of j is decreasing from left to right in Figure 7.1, because the absorption is increasing with the models from I to III.

A remarkable feature shown in Figure 7.1 is that all $j(\eta, r, x)$ have a flat plateau in the range $|x| \leq 2$. This gives the frequency range of the W-F local thermalization [88, 89]. The range of the flat plateau $|x| \leq 2$ is almost dust-independent, either for model I or for models II and III. This is expected, as neither the absorption nor scattering given by the κ term of (6.1) changes the frequency distribution of photons. The redistribution function (6.6) also does not change the frequency distribution of photons. This point can also be seen from (6.11) and (6.12), in which the κ terms are frequency-independent. The evolution of the frequency distribution of photons is due only to the resonant scattering.

Since thermalization will erase all frequency features within the range $|x| \leq 2$, the double-peaked structure does not retain information of the photon frequency distribution within $|x| < 2$ at the source. That is, the results in Figure 7.1 will hold for any source $S_0\phi_s(x)$ with arbitrary $\phi_s(x)$ which is non-zero within $|x| < 2$ [88, 89]. This property can also be used as a test of the simulation code. It requires that simulation result in a flat plateau, regardless of the whether source is monochromatic or with a finite width around ν_0.

7.2 Dust effects on photon escape

7.2.1 Model I: scattering of dust

To study the effects of dust scattering on the Lyα photon escape, we show in Figure 7.2 the flux $f(\eta, r, x)$ of Lyα photons emerging from halos at the boundary $r = R = 10^2$ for Model I. The three panels of Figure 7.2 correspond to $\kappa = 10^{-4}$, 10^{-3}, and 10^{-2} from left to right, respectively. The source starts to emit photons at $\eta = 0$ with a stable luminosity $S_0 = 1$, and with a Gaussian profile $\phi_s(x) = (1/\sqrt{\pi})e^{-x^2}$.

Figure 7.2 clearly shows that the time-evolution of $f(\eta, r, x)$ is κ-independent. Although the cross section of dust scattering increases about 100 times from $\kappa = 10^{-4}$ to $\kappa = 10^{-2}$, the curves of the left and right panels in Figure 7.2 are almost identical.

According to the scenario of "single longest excursion", photon escape is not a process of Brownian random walk in the spatial space, but a transfer in the frequency space [80, 9, 1, 2, 51, 12]. A photon will escape, once its frequency is transferred from $|x| < 2$ to $|x| > 2$, on which the medium is transparent. On the other hand, dust scattering given by the redistribution function equation (6.6) does not change photon frequency. Dust scattering has no effect on the transfer in the frequency space.

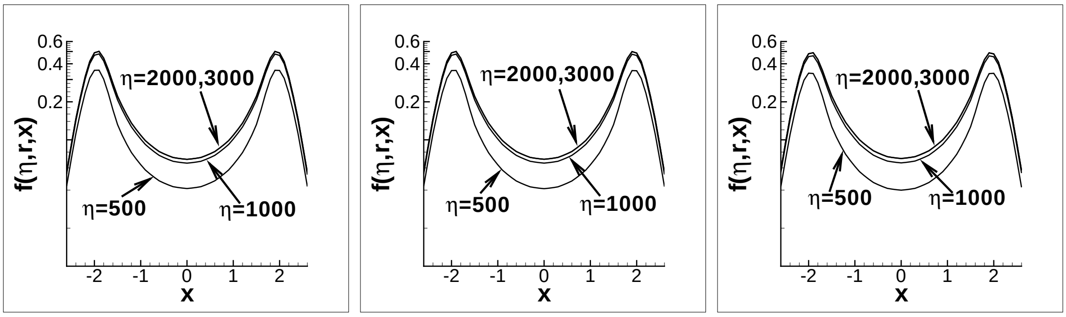

Figure 7.2: Flux $f(\eta, r, x)$ of Lyα photons emergent from halos at the boundary $R = 10^2$, and for the dust model I $A = 1,\ g = 0.73$. The parameter κ is taken to be 10^{-4} (left), 10^{-3}(middle) and 10^{-2} (right). The source is $S_0 = 1$ and $\phi_s(x) = (1/\sqrt{\pi})e^{-x^2}$. The parameter $a = 10^{-3}$.

Moreover, photons with frequency $|x| < 2$ are quickly thermalized after a few hundred resonant scattering. In the local thermal equilibrium state, the angular distribution of photons is isotropic. Thus, even if the dust scattering is anisotropic $g \neq 0$ with respect to the direction of the incident particle, the angular distribution will keep isotropic undergoing a $g \neq 0$ scattering. Hence, dust scattering also has no effect on the angular distribution.

7.2.2 Model III: absorption of dust

Similar to Figure 7.2, we present in Figure 7.3 the flux of Model III, i.e. dust causes only absorption without scattering. All other parameters of Figure 7.3 are the same as in Figure 7.2. In the left panel of Figure 7.3, the curves at the time $\eta = 2000$ and 3000 are the same. It means the flux $f(\eta, R, x)$ at the boundary R is already stable, or saturated at the time $\eta \geq 2000$. The small difference between the curves of $\eta = 1000$ and $\eta \geq 2000$ of the left panel indicates that the flux is still not yet completely saturated at the time $\eta = 1000$. However, comparing the middle and right panels of Figure 7.3, we see that for $\kappa = 10^{-3}$, the flux has already saturated at $\eta = 1600$, while it has saturated at $\eta = 800$ for $\kappa = 10^{-2}$. That is, the stronger the dust absorption, the shorter the saturation time scale. The time scales of escape or saturation do not increase by dust absorption, and even decrease with respect to the medium without dust. Stronger absorption leads to shorter time scale of saturation.

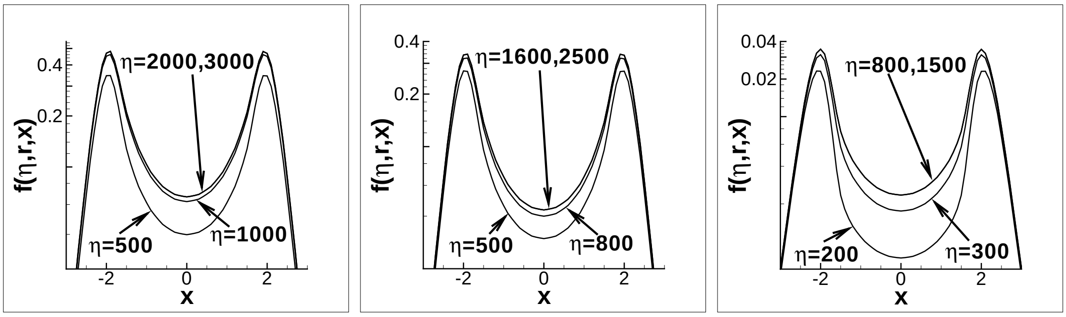

Figure 7.3: Flux $f(\eta, r, x)$ of Lyα photons emergent from halos at the boundary $r = R = 10^2$. The parameters of the dust are $A = 0$ and $\kappa = 10^{-4}$ (left), 10^{-3} (middle) and 10^{-2} (right). Other parameters are the same as in Figure 7.2.

Obviously, dust absorption does not help in producing photons for the "single longest excursion". Therefore, dust absorption cannot make the time scale of producing photons for "single longest excursion" smaller. However, dust absorptions are effective in reducing the number of photons trapped in the state of local thermalized equilibrium $|x| < 2$ (see also Section 7.3.2). This indicates that the higher the value of κ, shorter the time scale of saturation.

7.2.3 Effective absorption optical depth

Since Lyα photons underwent a large number of resonant scattering before escaping from the halo with optical depth $\tau_0 \gg 1$, it is generally believed that a small absorption of dust will lead to a significant decrease of the flux. However, it is still unclear what the exact relationship between the dust absorption and the resonant scattering is. This problem should be measured by the effective optical depth of dust absorption of Lyα photons in $R = \tau_0 \gg 1$ halos.

To calculate the effective optical depth, we first give the total flux of Lyα photons emerging from a halo of radius R, which is defined as $F(\eta) = \int f(\eta, R, x)dx$. Figure 7.4 plots $F(\eta)$ as a function of time η for halo with sizes $R = \tau_0 = 10^2$ and 10^4. The curves typically are growing and then saturating. The three panels correspond to the dust models I, II and III from left to right. The upper panels are of $R = 10^2$, and lower panels for $R = 10^4$. In each panel of $R = 10^2$, we have three curves corresponding to $\kappa = 10^{-4}$, 10^{-3} and 10^{-2}, respectively. In cases of $R = 10^4$, we take $\kappa = 10^{-4}$ and 10^{-3}.

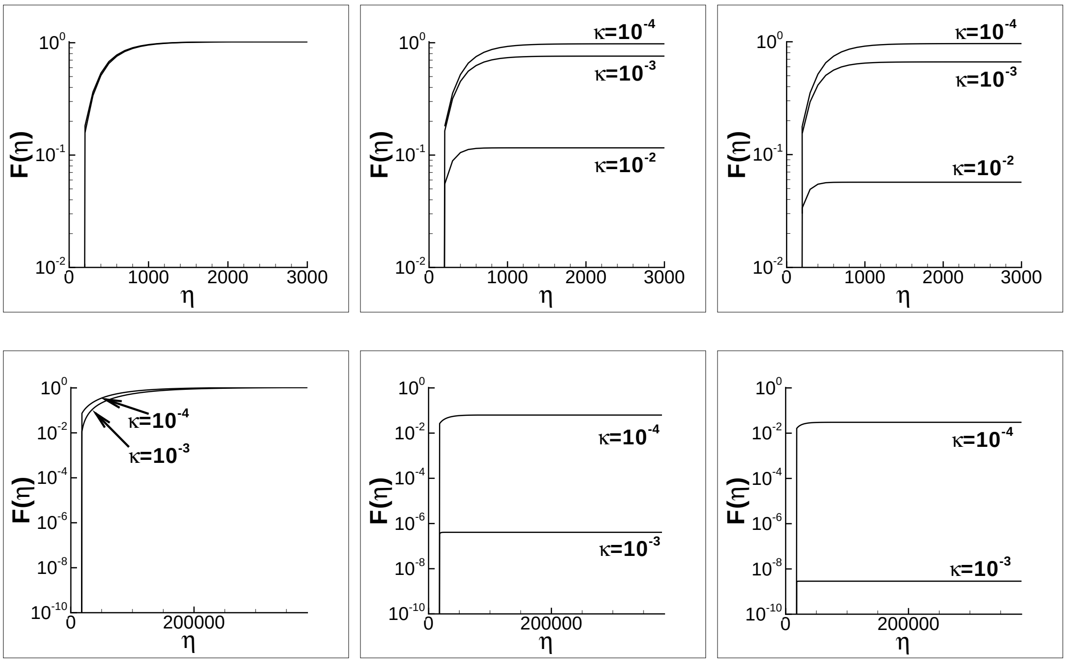

Figure 7.4: The time evolution of the total flux $F(\eta)$ at the boundary of halos with $R = \tau_0 = 10^2$ (upper panels), and $R = \tau_0 = 10^4$ (lower Panels). The source of $S_0 = 1$ and $\phi_s(x) = (1/\sqrt{\pi})e^{-x^2}$ starts to emit photons at time $\eta = 0$. The parameters of dust are $(A = 1, g = 0.73)$ (left); $(A = 0.32,\ g = 0.73)$ (middle) and $A = 0$ (right). In each panel of $R = 10^2$, κ is taken to be 10^{-4}, 10^{-3} and 10^{-2}. In the cases of $R = 10^4$, κ is taken to be 10^{-4}, 10^{-3}.

The left panel of Figure 7.4 shows that the three curves of $\kappa = 10^{-4}$, 10^{-3} and 10^{-2} are almost the same. This is consistent with Figure 7.2 that for Model I, the time-evolution of f are κ-independent for the pure scattering dust. For the pure absorption dust (the right panel of Figure 7.4), the saturated flux is smaller for larger κ. We can also see from Figure 7.4 that the time scale of approaching saturation is smaller for larger κ. The result of model II is in between that for models I and III.

With the saturated flux of Figure 7.4, one can define the effective absorption optical depth by $\tau_{\rm effect} \equiv -(1/\kappa)\ln f_S$. The results are shown in Table 7.1, in which τ_a is the dust absorption depth. It is interesting to see that the effective absorption optical depth is always equal to a few times of the optical depth of resonant scattering τ_0, regardless of whether τ_a is less than 1. Namely, the effective absorption depth $\tau_{\rm effect}$ of dust is roughly proportional to τ_0.

Table 7.1: Effective absorption optical depth $\tau_{\rm effect}$

		Model II			Model III		
$R = \tau_0$	κ	τ_a	f_S	$\tau_{\rm effect}$	τ_a	f_S	$\tau_{\rm effect}$
10^2	10^{-4}	0.0068	0.978	2.2×10^2	0.01	0.963	3.8×10^2
10^2	10^{-3}	0.068	0.760	2.7×10^2	0.10	0.670	4.0×10^2
10^2	10^{-2}	0.68	0.116	2.2×10^2	1.00	0.057	2.9×10^2
10^4	10^{-4}	0.68	6.28×10^{-2}	2.8×10^4	1.00	3.02×10^{-2}	3.5×10^4
10^4	10^{-3}	6.8	4.07×10^{-7}	1.5×10^4	10.0	2.87×10^{-9}	1.97×10^4

According to the random walk scenario, if a medium has optical depths of absorption τ_a and scattering τ_s, the effective absorption optical depth should be equal to $\tau_{\rm effect} = \sqrt{\tau_a(\tau_a + \tau_s)}$ [95]. However, the results of the last line of Table 7.1 show that the random walk scenario does not work for the dust effect on resonant photon transfer. This result is consistent with Figures 7.2 and 7.3. When the optical depth of dust is lower than the optical depth of resonant scattering τ_0, the time scale of photon escaping is not affected by the dust, but is proportional to τ_0, and therefore, the absorption is also proportional to τ_0.

7.2.4 Escape coefficient

With the total flux, we can define the escaping coefficient of Lyα photon as $f_{\rm esc}(\eta, \tau_0) \equiv F(\eta)/F_0$, where F_0 is the flux of the center source. Figure 7.5 shows $f_{\rm esc}(\eta, \tau_0)$ at three times $\eta = 5 \times 10^3$, 10^4 and $3.2{\times}10^4$ for Model II and $\kappa = 10^{-3}$. At $\eta = 5 \times 10^3$, the flux of halos with $\tau_0 \leq 10^3$ is saturated. At $\eta = 10^4$, halos with $\tau_0 \leq 3 \times 10^3$ are saturated, and all halos of $\tau_0 \leq 10^4$ are saturated at $\eta = 3.2 \times 10^4$.

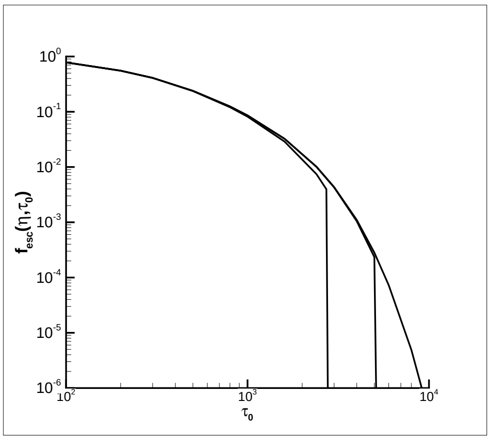

Figure 7.5: Escaping coefficient $f_{\rm esc}(\eta)$ as a function of the optical depth τ_0 of halo at time $\eta = 5 \times 10^3$, 10^4, and $3.2{\times}10^4$ from bottom to up. Dust is modeled by II, $A = 0.32$, $g = 0.73$, and $\kappa = 10^{-3}$.

7.3 Dust effects on double-peaked profile

7.3.1 Dust and the frequency of double peaks

A remarkable feature of Lyα photon emerging from an optically thick medium is the double-peaked profile. Figures 7.1, 7.2 and 7.3 have shown that the double peak frequencies $x_+ = |x_-|$ are almost independent of either the scattering or the absorption of dust. In this section, we consider halos of size R or τ_0 larger than 10^2. Figure 7.6 presents the double peak frequency $|x_\pm|$ as a function of $a\tau_0$, where the parameter a is taken to be 10^{-2} (left) and 5×10^{-3} (right). Comparing the curves with dust and without dust in Figure 7.6 we conclude that the dust effect on $|x_\pm|$ is very small woth $a\tau_0 = aR = 10^2$.

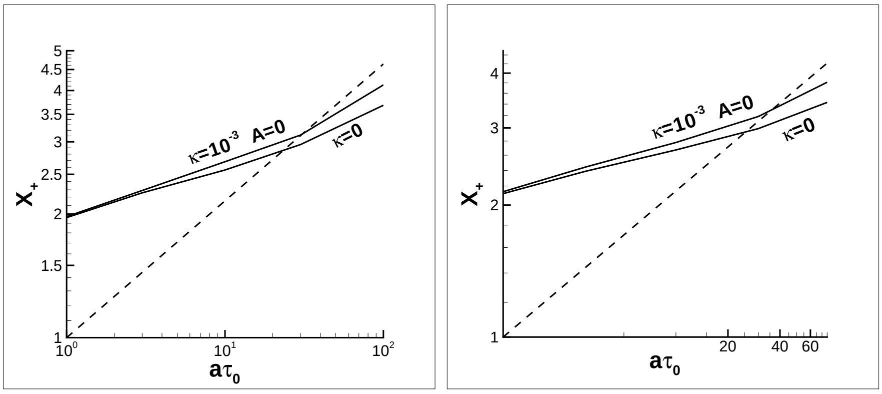

Figure 7.6: The two-peak frequencies $x_+ = |x_-|$ as a function of $a\tau_0$. The parameter a is taken to be 10^{-2} (left) and 5×10^{-3} (right). Dust model III (pure absorption) is used, and κ is taken to be 10^{-3}. The dashed straight line gives $\log x_\pm$-$\log a\tau$ with slope 1/3, which is to show the $(a\tau)^{1/3}$-law of $x_\pm$.

In the range $a\tau_0 < 20$, the $|x_\pm|$-τ_0 relation is almost flat with $|x_\pm| \simeq 2$. It is because the double-peaked profile is given by the frequency range of the locally thermal equilibrium. The positions of the two peaks, x_+ and x_-, essentially are at the maximum and minimum frequencies of the local thermalization. The frequency range of the local thermal equilibrium state is determined mainly by the Doppler broadening, and weakly dependent on τ_0. Thus, we always have $x_\pm \simeq \pm 2$. When the optical depth is larger, $a\tau_0 \sim 10^2$, more and more photons of the flux are attributed to the resonant scattering by the Lorentzian wing of the Voigt profile. In this phase, $|x_\pm|$ will increase with τ_0.

Figure 7.6 also shows a line $x_\pm = \pm(a\tau_0)^{1/3}$, which is given by the analytical solution of the Fokker-Planck approximation, in which the Doppler broadening core in the Voigt profile being ignored [51, 76, 37]. The numerical solutions of (6.1) or (6.11) and (6.12) deviate from the $(a\tau_0)^{1/3}$-law at all parameter range of Figure

7.6. The deviation at $a\tau_0 < 20$ is caused by the Doppler broadening core in the Voigt profile is ignored in the Fokker-Planck approximation, so no locally thermal equilibrium can be reached. Therefore, in the range $a\tau_0 < 20$, $|x_\pm|$ of the WENO solution is larger than the $(a\tau_0)^{1/3}$-law. In the range of $a\tau_0 > 20$, the Fokker-Planck approximation yields a faster diffusion of photons in the frequency space. This point can be seen in the comparison between a Fokker-Planck solution with Field's analytical solution (Figure 1 in [94]). In this range, the numerical results of $|x_\pm|$ is less than the $(a\tau_0)^{1/3}$-law.

7.3.2 No narrowing and no widening

The dust effect has been used to explain the narrowing of the width between the two peaks [71]. conversely, it is also used to explain the widening of the width between the two peaks [110]. However, Figures 7.1, 7.2, 7.3 and 7.6 show that the width between the two peaks of the profile is very weakly dependent on dust scattering and absorption. This result supports, at least in the parameter range considered in Figures 7.1, 7.2, 7.3, neither the narrowing nor the widening of the two peaks.

If dust absorption can cause narrowing, the absorption should be weaker at $|x| \sim 0$, and stronger at $|x| \geq 2$. Similarly, if dust absorption can cause widening, the absorption should be weaker at $|x| \sim 2$, and stronger at $|x| \sim 0$. To test these assumptions, Figure 7.7 plots $\ln[f(\eta, r, x, \kappa = 0)/f(\eta, r, x, \kappa)]$ as a function of x. It measures the x(frequency)-dependence of the flux ratio with and without dust absorption. We take large η, and then the fluxes in Figure 7.7 are saturated. Figure 7.7 shows that the absorption in the range $|x| < 2$ is much stronger than for $|x| > 2$, and therefore, the assumption of the narrowing is ruled out. Figure 7.7 shows also that the curves of $\ln[f(\eta, r, x, \kappa = 0)/f(\eta, r, x, \kappa = 10^{-3})]$ are almost flat in the range $|x| < 2$. Therefore, the assumption of widening of the two peaks can also be ruled out.

Since the cross sections of dust absorption and scattering are assumed to be

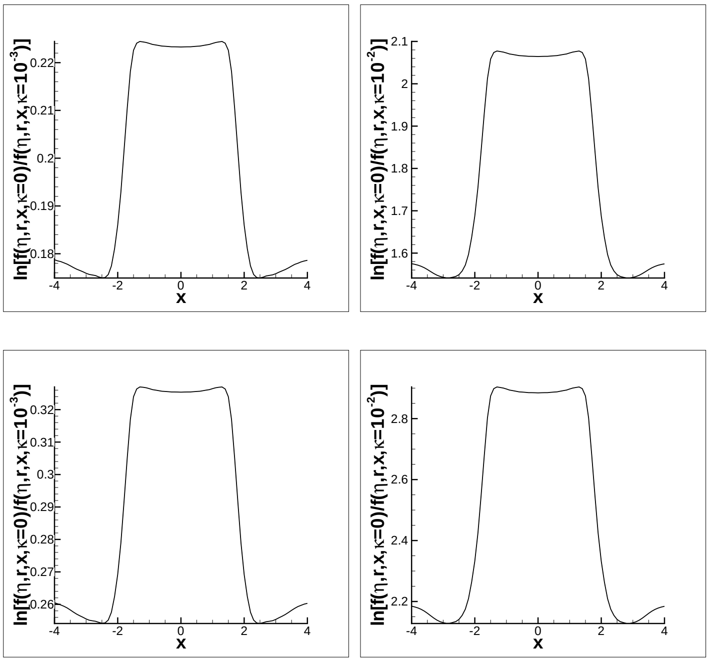

Figure 7.7: $\ln[f(\eta, r, x, \kappa = 0)/f(\eta, r, x, \kappa)]$ as function of x for model II (up), and III (bottom), and $\kappa = 10^{-3}$ (left) and 10^{-2} (right). Other parameters are the same as in Figure 7.2.

frequency-independent. (6.11) and (6.12) do not contain any frequency scales other than that from resonant scattering. However, either narrowing or widening would require to have frequency scales different from that of resonant scattering. This is not possible if the dust is gray.

7.3.3 Profile of absorption spectrum

If the radiation from the sources has a continuous spectrum, the effect of a neutral hydrogen halos is to produce an absorption line at $\nu = \nu_0$. The profile of the absorption line can also be found by solving (6.11) and (6.12), but replacing the boundary equation (6.14) by

$$f(\eta, 0, x) = S_0. \tag{7.1}$$

That is, we assume that the original spectrum is flat in the frequency space. The spectrum of the flux emerging from the halo of $R = 10^2$ and 10^4 with central source for (7.1) for dust models I, II and III are shown in Figure 7.8. All curves are for large η, i.e. they are saturated.

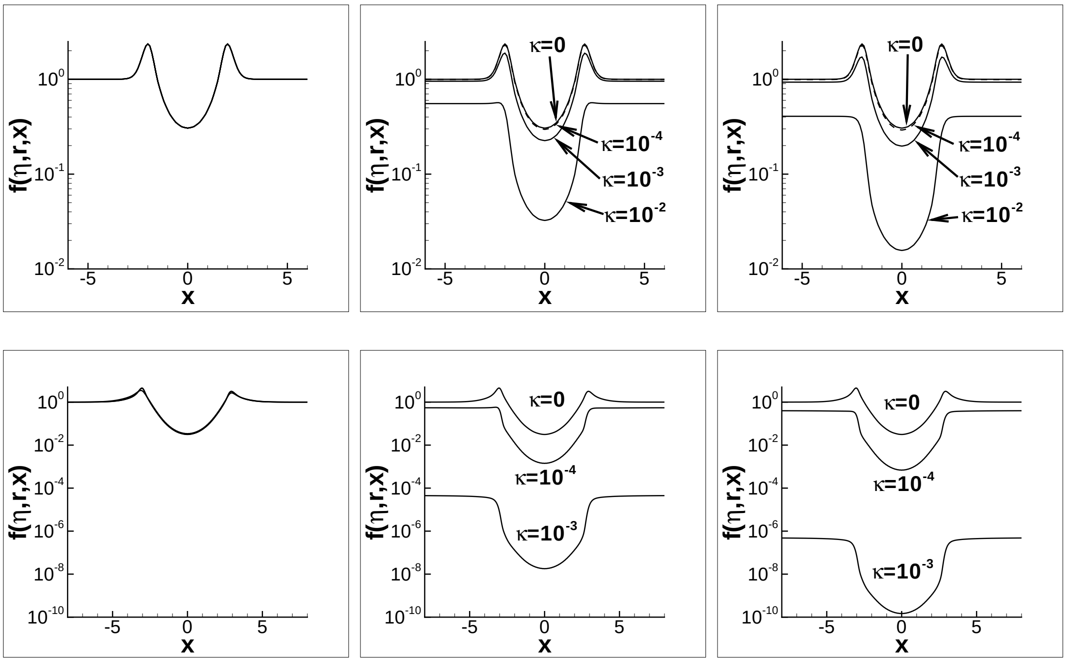

Figure 7.8: The spectrum of the flux emergent from halo of $R = 10^2$ (upper panels) and 10^4 (lower panels) with central source of equation (7.1) for the dust model I (left), II (middle) and III (right). Other parameters are the same as in Figure 7.2.

The optical depths at the frequency $|x| > 4$ are small, and therefore, the Eddington approximation might no longer be valid. However, those photons do not strongly involve the resonant scattering, and hence they do not significantly affect the solution around $x = 0$. The solutions of Figure 7.8 is still useful to study the profiles of f around $x = 0$.

The flux profile of Figure 7.8 are absorption lines with the width given by the double peaks similar to the double peaked structure of the emission line. The flux at the double peaks is even higher than for the flat wing. It is because more photons are stored in the frequency range $|x| < 2$. According to the redistribution function equation (6.7), the probability of transferring a x' photon to a $|x| < |x'|$ photon is larger than that from $|x'|$ to $|x| > |x'|$. Therefore, if the original spectrum is flat, the net effect of resonant scattering is to bring photons with frequency $|x| > 2$ to $|x| < 2$. Photons stored $|x| < 2$ are thermalized, and therefore, in the range $|x| < 2$, the profile will be the same as the emission line, and the double peaks can be higher than the wing. It makes the shoulder at $|x| \sim 2$.

As expected, for model I (left panels of Figure 7.8), the double profile is completely κ-independent. Dusty scattering does not change the flux and its profile. For models II and III, the higher the κ, the lower the flux of the wing, because the dust absorption is assumed to be frequency-independent. The positions of the double peaks, in the absorption spectrum are also κ-independent. This once again shows that dust absorption and scattering causes neither narrowing nor widening of the double-peaked profile. However, for higher κ the flux of the peaks is lower. When the absorption is very strong, the double-peaked structure might disappear, but will not be narrowed or widened.

7.4 Discussions and conclusions

The study of dust effects on radiative transfer has had a long history related to extinction. However, dust effects on radiative transfer of resonant photons actually have not been carefully investigated. Existing works are mostly based on the solutions of the Fokker-Planck approximation, or Monte Carlo simulation. These results are important. We revisited these problems with the WENO solver of the integro-differential equation of the resonant radiative transfer, and have found some features which have not been addressed in previous works. These features are summarized as follows.

First, the random walk picture in the physical space will no longer be available for estimating the effective optical depth of dust absorption. For a medium with the optical depth of absorption and resonant scattering to be $\tau_a \gg 1$, $\tau(\nu_0) \gg 1$ and $\tau_s(\nu_0) \gg \tau_a$, the effective absorption optical depth is found to be almost independent of τ_a, and to be equal to about a few times of $\tau_s(\nu_0)$.

Second, dust absorption will, of course, yield the decrease of the flux of Lyα photons emergent from optical thick medium. However, if the absorption cross-section of dust is frequency independent, the double-peaked structure of the frequency profile is basically dust-independent. The double-peaked structure does not narrow or widen by the absorption and scattering of dust.

Third, the time scales of Lyα photon transfer basically are independent of dust scattering and absorption. Since these time scales are mainly determined by the kinetics in the frequency space. However, dust does not affect the behavior of the transfer in the frequency space if the cross section of the dust is wavelength-independent. The local thermal equilibrium makes the anisotropic scattering ineffective on the angular distribution of photons. Dust absorption and scattering do not lead to the increase or decrease of the time of storing Lyα photons in the halos.

The differences between the time-independent solutions of the Fokker-Planck approximation, or Monte Carlo simulation and the WENO solution of (6.1) is mainly related to the W-F effect. Therefore, all above-mentioned features can already be clearly seen with halos of $\tau_0 \sim 10^2$, in which the W-F local thermal equilibrium has been well established.

In this context, most calculation in this chapter is on holes with $\tau_0 < 10^5$. This range of τ_0 certainly is unable to describe halos with column number density of HI larger than 10^{17} cm^{-2} (e.g. [90]). Nevertheless, the result of $\tau_0 < 10^5$ would already be useful for studying the 21 cm region around high-redshift sources, of which the optical depth typically is [74, 89].

$$\tau_0 = 3.9 \times 10^5 f_{\rm HI} \left(\frac{T}{10^4{\rm K}}\right)^{-1/2} \left(\frac{1+z}{10}\right)^3 \left(\frac{\Omega_b h^2}{0.022}\right) \left(\frac{R_{\rm ph}}{10{\rm kpc}}\right), \qquad (7.2)$$

where f_{HI} is the fraction of HI. All other parameters in (7.2) is taken from the concordance ΛCDM mode. For these objects the relation between dimensionless η and physical time t is given by

$$t = 5.4 \times 10^{-2} f_{\mathrm{HI}}^{-1} \left(\frac{T}{10^4\mathrm{K}}\right)^{1/2} \left(\frac{1+z}{10}\right)^{-3} \left(\frac{\Omega_b h^2}{0.022}\right)^{-1} \eta, \quad \mathrm{yr}. \tag{7.3}$$

The 21 cm emission rely on the W-F effect. On the other hand, the time-scale of the evolution of the 21 region is short. The effect of dust on the time-scales of Lyα evolution should be considered.

We have not considered the Lyα photons produced by the recombination in the ionized halo. If the halo is optical thick, photons from the recombination will also be thermalized. The information of where the photon comes from will be forgotten during the thermalization. Therefore, photons from recombination should not show any difference from those emitted from central sources. Only the photons formed very close to the boundary of the halo will not be thermalized, and may yield different behavior.

Chapter 8

Angular distribution of Lyα resonant photons emergent from optically thick medium

This chapter will study the angular distribution of Lyα photon transferring in an optically thick medium. Previous methods are based on the Eddington approximation and the evolution of the angular distribution is completely ignored. However, the evolution of angular distribution actually is significant. In a thermalized or statistical equilibrium state, the angular distribution of photons should be isotropic, regardless of the initial angular distribution. Therefore, one can expect that the angular distribution of Lyα photons with resonant frequency ν_0 should be isotropic. On the other hand, the angular distribution of photons with frequency different from ν_0 might be anisotropic, as those photons are not involved in the evolution of thermalization or statistical equilibrium. Consequently, the angular distributions of Lyα resonant photons from optically thick medium should be frequency-dependent. It definitely cannot be described by the Eddington approximation. The evolution of the angular distribution of resonant photons is not trivial. We still use the WENO solver, and solve the photon transfer in both frequency and angular spaces.

8.1 Transfer equations of resonant photons without dust

We solve the radiative transfer equation of Lyα resonant photon in a spherically symmetric medium containing neutral HI. For simplicity, we ignore the effect of dust (i.e. $\kappa(x) = 0$), and (6.1) is

$$\frac{\partial I}{\partial \eta} + \mu \frac{\partial I}{\partial r} + \frac{(1-\mu^2)}{r}\frac{\partial I}{\partial \mu} - \gamma \frac{\partial I}{\partial x} = \\ -\phi(x;a)I + \int \mathcal{R}(x,\mu,x',\mu';a)I(\eta,r,x',\mu')dx'd\mu'/2 + S. \tag{8.1}$$

The boundary and initial conditions are still given in (6.13), (6.17) and (6.19).

8.1.1 Test with Field's analytical solution

We first test the WENO solver with analytical solutions. Assuming that the specific intensity and source S are homogeneous in the r and μ space, i.e. $I(\eta, r, x, \mu)$ is independent of variables r and μ. (8.1) becomes

$$\frac{\partial J}{\partial \eta} - \gamma \frac{\partial J}{\partial x} = -\phi(x)J + \int R(x,x')J(\eta,x')dx' + S, \tag{8.2}$$

where

$$J(\eta, x) = \frac{1}{2}\int I(\eta, r, x, \mu)d\mu. \tag{8.3}$$

Take $\gamma = 0$, Voigt parameter $a = 0$, the source $S = \phi(x) = \pi^{-1/2}e^{-x^2}$, and the initial radiative field $I(x, \eta = 0) = 0$. The time-dependent solution of(8.2) is [44, 94].

$$J(x,\eta) = \pi^{-1/2}[1-\exp(-\eta e^{-x^2})] \\ + \int_x^\infty e^{w^2}[1-(1+\eta e^{-w^2})\exp(-\eta e^{-w^2})]\mathrm{erf}(w)dw. \tag{8.4}$$

Our solver seeks a solution I from (8.1). One can then give J via (8.3). It is interesting to see whether the solution to (8.4) can be reproduced, if we also assume that the source S in (8.1) is spatially homogeneous, but μ-dependent, i.e. $S = \Theta(\mu)\phi(x) = \Theta(\mu)\pi^{-1/2}e^{-x^2}$ where $\Theta(\mu)$ describes the angular distribution of photons from the source. We consider both an isotropic source

$$S = \pi^{-1/2}e^{-x^2}, \quad -1 \leq \mu \leq 1, \tag{8.5}$$

and an anisotropic source as follows,

$$S = \begin{cases} 2(n+1)\mu^n\pi^{-1/2}e^{-x^2}, & 0 < \mu \leq 1, \\ 0, & -1 \leq \mu \leq 0, \end{cases} \tag{8.6}$$

where n is taken to be a positive integer.Obviously, the larger the n, the stronger the emission in the direction $\mu = 1$. The factor $2(n+1)$ is for normalization: $\frac{1}{2}\int_0^1 2(n+1)\mu^n d\mu = 1$.

The numerical results with sources (8.5) and (8.6) with $n = 4$ and 6 are shown in Figure 8.1. It is expected that the numerical solution with source (8.5) (the left

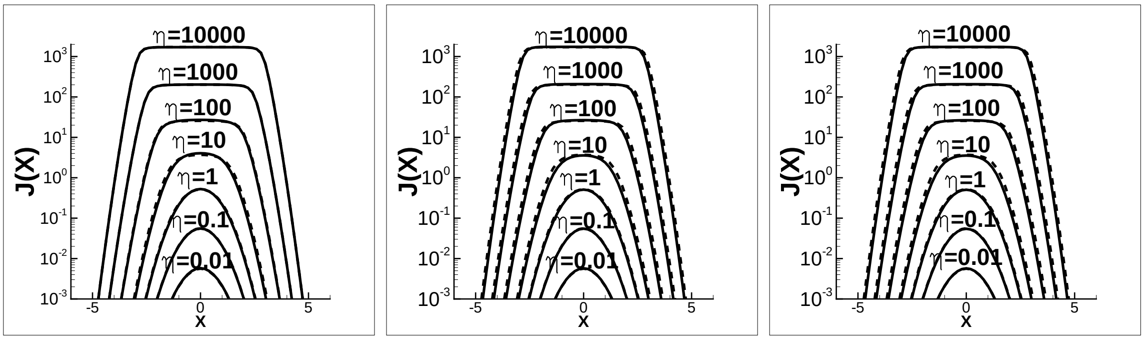

Figure 8.1: The WENO numerical solutions (solid lines) of equation (8.1) assuming the sources S is a) $S = \pi^{-1/2}e^{-x^2}$ for all μ (left); b) $S = 10\mu^4\pi^{-1/2}e^{-x^2}$ (middle); and c) $S = 14\mu^6\pi^{-1/2}e^{-x^2}$ (right) for $\mu > 0$ and S=0 for $\mu < 0$. The Field's analytical solution is shown with dot lines.

panel of Figure 8.1) should follow the analytical solution (8.4) well, as the isotropic source is the same as that used to find the analytical solution.

It is interesting to see that the WENO solutions of $n = 4$ (middle panel) and $n = 6$ (right panel) also follow the analytical solution to (8.4) well. It seems to indicate that the evolution of the frequency space is independent of the μ-space.

8.1.2 Time scale of the statistical equilibrium of the angular distribution

A remarkable feature of the solutions of Figure 8.1 is a flat plateau in the range $|x| < 2$ at time $\eta > 100$. The flat plateau is caused by the Wouthuysen-Field local thermalization of frequency distribution of resonant photon [113, 44, 45]. The flat plateau actually is the Boltzmann statistical equilibrium distribution around $x = 0$ when the atomic mass is infinite. If the mass is finite, i.e. considering the recoil in the re-distribution functions (6.3) or (6.4), the flat plateau will become e^{-2bx}, where $b = h\nu_0/mv_Tc$, which is the local Boltzmann distribution required by the Wouthuysen-Field effect [88]. The resonant scattering between photons and HI atoms leads to the Boltzmann distribution of the photon frequency distribution around $x = 0$ with the temperature equal to that of HI atoms.

When resonant photons undergo the local thermalization in the frequency space, the angular distribution should be approaching statistical equilibrium. The anisotropic μ-distributions have to evolve to be isotropic (statistical equilibrium). We calculate all the μ-distributions at the time η corresponding to the three panels of Figure 8.1. The result is plotted in Figure 8.2. The μ-distribution of left panel is always isotropic. This is expected as the source is isotropic, which is already in the state of statistical equilibrium.

The middle and right panels of Figure 8.2 show the evolution of an anisotropic μ-distribution equation (8.6). The time scale for approaching an isotropic distribution seems to be independent of the anisotropy of sources. It is always equal to about

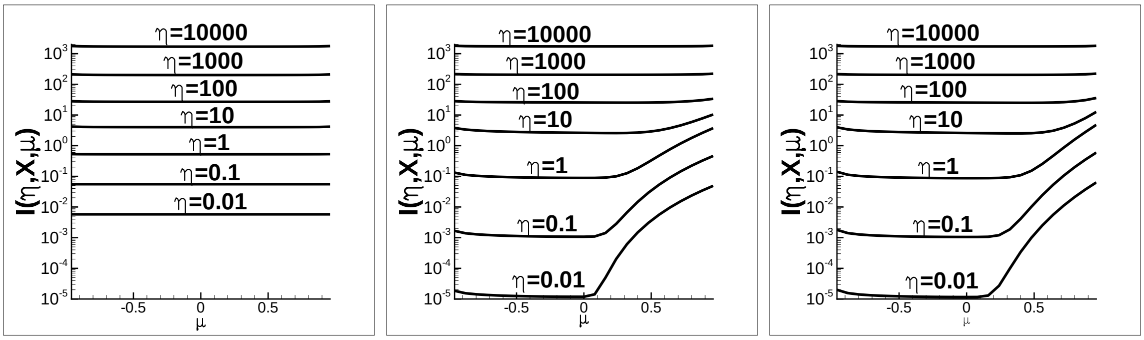

Figure 8.2: The WENO numerical solutions of angular distributions from equation (8.1) at $x = 0$, assuming the sources S are a) $S = \pi^{-1/2}e^{-x^2}$ for all μ (left) and b) $S = 10\mu^4\pi^{-1/2}e^{-x^2}$ (middle); and c) $S = 14\mu^6\pi^{-1/2}e^{-x^2}$ (right) for $\mu > 0$ and S=0 for $\mu < 0$.

$\eta \sim 100$ for both $n = 4$ and $n = 6$, i.e. the μ-distribution will become isotropic after 100 times of resonant scattering. This time scale is about the same as that of the W-F thermalization (Figure 8.1). Therefore, the thermalization in the frequency space and the isotropic distribution in the μ-space are realized at about the same time.

8.2 Precision of the Eddington approximation

8.2.1 Equations of the Eddington approximation

Following the same analysis as in Section 6.1.3, (8.1) yields the equations of j and f as

$$\frac{\partial j}{\partial \eta} + \frac{\partial f}{\partial r} = -\phi(x;a)j + \int \mathcal{R}(x,x';a)j dx' + \gamma\frac{\partial j}{\partial x} + r^2 S, \tag{8.7}$$

$$\frac{\partial f}{\partial \eta} + \frac{1}{3}\frac{\partial j}{\partial r} - \frac{2}{3}\frac{j}{r} = -\phi(x;a)f + \gamma\frac{\partial f}{\partial x}, \tag{8.8}$$

with initial and boundary conditions (6.14), (6.18) and (6.20).

8.2.2 Profiles of j and f

Figure 8.3 does show small differences between the solutions with and without the Eddington approximation, even though both solutions are given by the same source. The difference comes from the contribution of the terms of $l > 2$ in the Legendre expansion. The difference between the profiles with and without the Eddington approximation becomes smaller when the time η is larger. It is because larger η corresponds to larger optical depth. The Eddington approximation generally is good for optically thick medium.

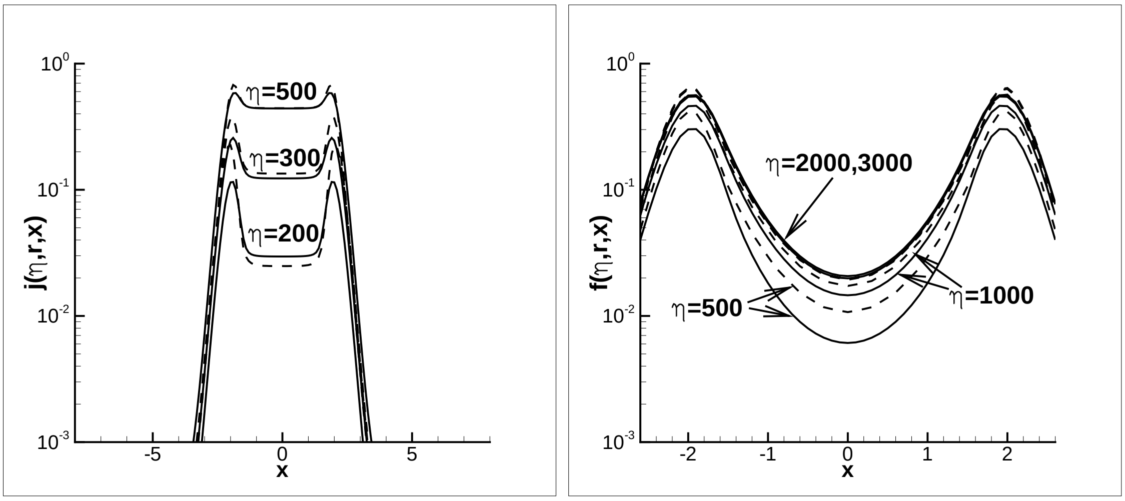

Figure 8.3: A comparison of the solutions j and f with the Eddington approximation (dashed curves) and the solutions of f without the Eddington approximation (solid curves). Relevant parameters are $r = R = 10^2$, and $a = 10^{-3}$.

Now we consider different sources. We re-do the solutions of j and f with equation (8.1) by taking $S = \delta(\mu - \frac{1}{2})$ and $S = \delta(\mu - 1)$. We use polynomials of degree 6 to approximate the delta sources. The results are given in Figure 8.4, which shows the same shape of the profiles. That is, the profiles of j and f are not affected by the angular distribution of photons from the source. It is probably because the μ-distribution quickly evolves into the statistical equilibrium state, the initial anisotropy of the μ distribution is forgotten.

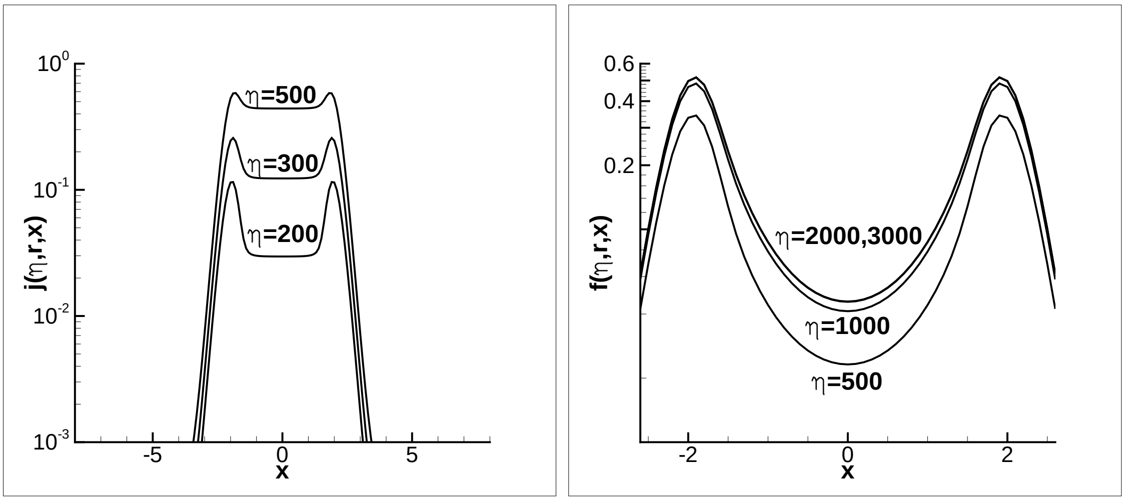

Figure 8.4: A comparison of the solution s j and f with S=$\delta(\mu-1/2)$ and $\delta(\mu-1)$. Relative parameters are $r = R = 100$, $a = 10^{-3}$.

8.3 Angular distributions

8.3.1 Frequency dependence

Although the Eddington approximation is acceptable when calculating the profile of Lyα photons in the frequency space, it fails in the μ-space. The result in Figure 8.2 shows that the μ-distribution is isotropic at frequency ν_0. On the other hand, the μ-distribution will no longer be isotropic at frequency $|x| \geq 2$, because photons of $|x| \geq 2$ have not undergone scattering. Consequently, the angular distribution of photons emerging from optically thick halo should be frequency(energy)-dependent.

We calculate the μ-distribution of photons from a halo with $R = 500$ with the central source given by (6.15), i.e. photons from the source can be described by the Eddington approximation (6.9). The result is shown in Figure 8.5. The μ distributions at frequencies $x = 0$ and 0.8 are basically straight lines in the whole

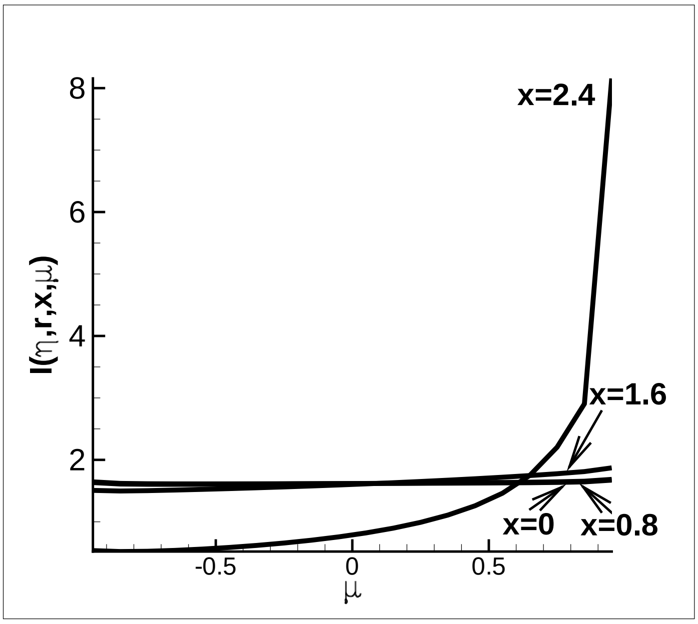

Figure 8.5: The μ distribution of photons emergent from a halo with radius $R = 500$. The frequencies are $x = 0.0, 0.8, 1.6$ and 2.4. The relevant parameters of the calculation are $\eta = 1.2 \times 10^4$, $\gamma = 0$, and $a = 10^{-3}$.

range $-1 \leq \mu \leq 1$. That is, I can be described by the Eddington approximation equation (6.9).

At $x = 1.6$, the μ-distribution begins to deviate from a straight line, i.e. deviating from an Eddington approximation. At $x = 2.4$, the μ-distribution shows a very sharp spike at $\mu = 1$. That is, the angular distribution of photons with frequency at the two peaks (Figure 8.3) is significantly different from isotropic, but is dominated by photons of $\mu = 1$. This result is consistent with the "single shot picture" [2, 12], in which photons with frequency $|x| < 2$ mainly undergo a diffusion in the frequency space; once a photon diffuses to $|x| \geq 2$, it will take "single longest excursion" to leave for outside of the halo. Therefore, the two peaks of the flux f at frequency $x_{\pm} \simeq \pm(2 - 3)$ are dominated by photons from "single longest excursion" photons, of which $\mu \sim 1$.

8.3.2 Dependence of the initial anisotropy

The source of Figure 8.5 given by (6.15) has $\Theta(\mu) = 6\mu$ $(\mu > 0)$, which is linear in μ. We now consider sources with higher anisotropy with $\Theta(\mu)$ given by

$$\Theta(\mu) = \begin{cases} 2(n+2)\mu^n, \; 0 < \mu \leq 1, \\ 0, \qquad\qquad\quad \mu < 0. \end{cases} \tag{8.9}$$

When the integer n is large, $\Theta(\mu)$ is similar to a δ function $\delta(\mu - 1)$, i.e. most photons are in the direction $\mu = 1$.

We repeat the calculation of Figure 8.5, but using the source equation (8.9) with $n = 1$, 2, 4, 6 and 8. The result is plotted in Figure 8.6. It is interesting to see that the μ-distributions are independent of n, but depend on x. It is easy to explain the n-independence of the two top panels of Figure 8.6, both of which have frequency $|x| \leq 2$. In this frequency range, the evolution of the specific intensity I is governed by the local thermalization of x-space and entropy increasing of μ-space. These processes lead to the Boltzmann distribution in the energy space and

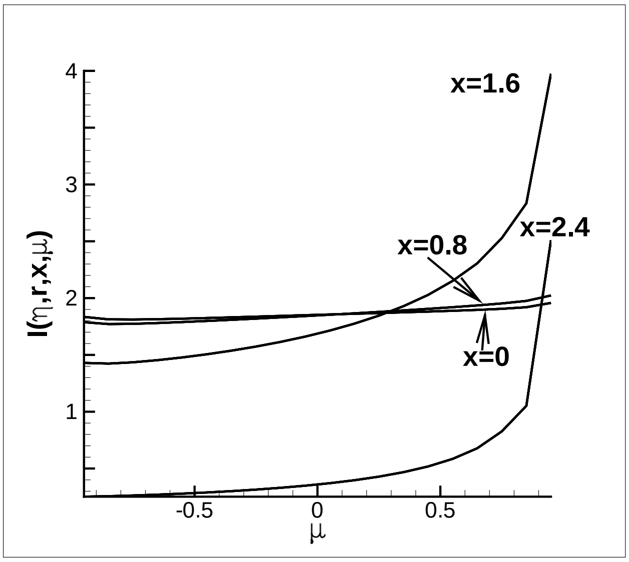

Figure 8.6: The μ-distribution of halo with radius R =100 and source equation (8.9) with n=1, 2, 4, 6, and 8. For each x, the curves for different n overlap with each other. The frequencies are taken to be x = 0, 0.8, 1.6, 2.4. The parameters of the halos are $\eta = 3500$, $\gamma = 0$, and $a = 10^{-3}$.

isotropic distribution in the angular space, regardless of the initial distributions in either frequency- or angular spaces. In other words, the initial distribution is forgotten during the local thermalization and approaching statistical equilibrium.

However, the mechanism of the local thermalization and approaching isotropic distribution seems to be unable to explain why the two curves at x=1.6 and x=2.4 of Figure 8.6 also show n-independence. The μ-distribution of these two curves of Figure 8.6 are highly anisotropic. Therefore, they do not have to be the result of the local thermalization and approaching statistical equilibrium. Why do they also show the behavior of forgetting the initial angular distributions? The reason is as follows. In the first phase of resonant photon evolution, Lyα photons are trapped in the range of $|x| \leq 2$ within the time scales of a few tens or hundred scattering [89]. The trapped photons have already forgotten their initial state. On the other hand, photons with $|x| \geq 2$ mostly come from the diffusion of trapped photons from

$|x| \leq 2$ to $|x| \geq 2$ [88]. Thus, all photons of $|x| \geq 2$ emerging from the optically thick halo essentially have the same initial condition, given by the $|x|$ space diffusion of trapped photons. Therefore, the initial distributions before they are trapped have been forgotten. This property can also be seen in Figure 8.2, in which, although the sources of the middle and right panels are different from each other, the behaviors of the time-evolution of the μ-distribution are about the same. This result also implies that it is impossible to recover the information of the distribution of photons emitted by the central source.

8.3.3 Collimation of photons of the double peaks

A common feature of Figure 8.6 is to show a very sharp spike at $\mu \sim 1$ when $|x| = 1.6$ and $|x| = 2.4$, corresponding to the double peaks of Figures 8.3 and 8.4. Therefore, the spiky distribution of μ indicates that the photons with frequency at the double peaks have formed a forward beam.

In order to measure the angular size of the $\mu = 1$ spikes, we fit the μ-distributions of Figure 8.5 at $x = 1.6$ and $x = 2.4$ with polynomials of μ. We find that both curves can be well fitted with polynomials of μ having leading terms $A\mu^{16} + B\mu^{15} + ...$, A, B being fitting coefficients. The terms of either μ^{16} or μ^{15} are much sharper than the central source equation (8.9) μ^n with $n \leq 6$. Therefore, the radiative transfer at the double peaks of frequency space plays the role of forward collimator. It made the photons form forward beams.

If we define the spread angle β of the forward beam as the angle of half intensity, this number can be estimated by $\cos^{16}\beta = 1/2$, and therefore, $\beta \sim 0.29$rad. This result is again consistent with the "single shot picture". The double peaks mainly consist of photons from a single shot, which moves in the forward direction.

8.3.4 Large halo

We calculate the μ-distribution of Lyα photons in a halo with large radius $R = 1000$, and the central source is given by equation (8.9) and $n = 6$. The results are given in Figure 8.7, which shows the dependence of the μ distribution on the radial variable r in the halo.

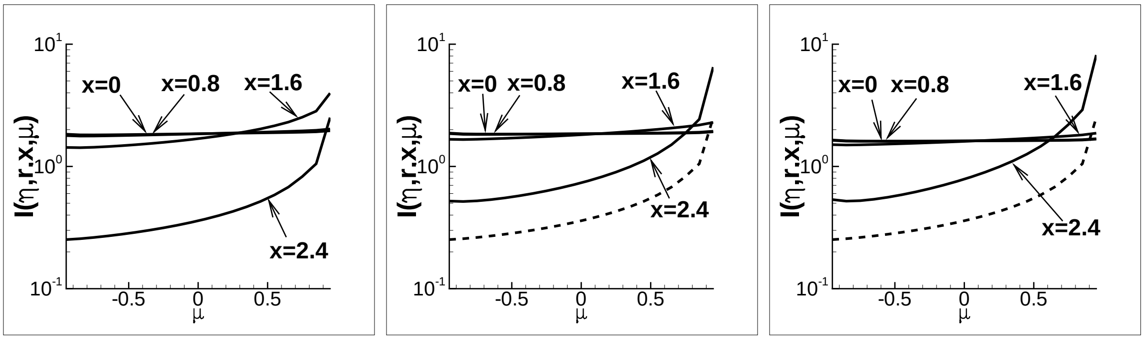

Figure 8.7: The μ-distributions at radial positions $r = 100$ (left), 300 (middle), 500 (right) of a halo with radius $R = 1000$. The source is given by equation (8.9) and $n = 6$. The frequencies are taken to be $x = 0$, 0.8, 1.6, 2.4. The dotted curves in the middle and right panels are μ-distributions at $r = 100$ $x = 2.4$. Other relevant parameters are $\eta = 1.2 \times 10^4$, $\gamma = 0$ and $a = 10^{-3}$.

Although the photons from the source of $n = 6$ are highly anisotropic, all the μ-distributions of $x = 0.0$ and 0.8 at $r = 100$, 300, 500 are straight lines. That is, the specific intensity I can be well approximated by the Eddington approximation equation (6.9). This result is consistent with Section 8.3.2. The r-dependence of the μ-distribution of $|x| = 2.4$ photons is also consistent with the result of section 8.3.3: the larger the r, the sharper the μ-distribution. The r transfer leads to the collimation.

The behavior of r-dependence of the μ-distribution at $x = 1.6$ is very different from that of $x = 0.0$, 0.8, and 2.4. The μ distribution at $r = 100$ is about the same as Figure 8.6, i.e. it has undergone an evolution of forward collimation, having a sharp spike at $\mu = 1$. However, the μ distribution will no longer show a spike at

$r = 300$ and 500. When r is equal to or less than about 200, the r-dependence of the μ distribution is similar to the $|x| = 2.4$ photons. However, when $r \geq 200$, the r-dependence of the μ distribution is similar to the $|x| = 0$ and 0.8 photons. This is because the optical depth at $|x| = 1.6$ is larger than $|x| = 2.4$, the "single shot picture" is working well at $r \sim 100$ for both $|x| = 1.6$ and 2.4. However, at $r \geq 200$, the single shot picture can still work well for $|x| = 2.4$, but not so well for $|x| = 1.6$.

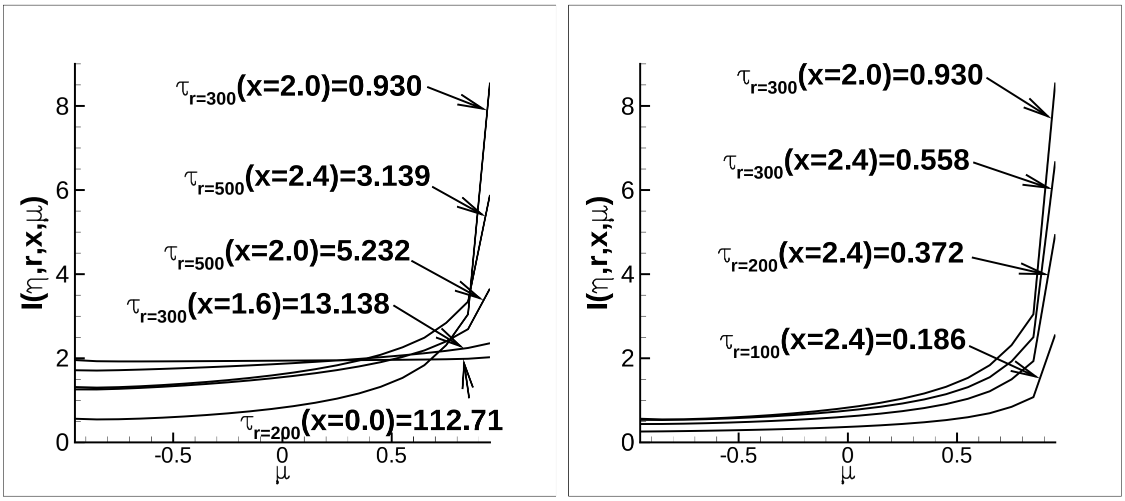

Figure 8.8: The μ-distributions with respect to the effective optical depth. Relevant parameters are $\eta = 1.2 \times 10^4$, $\gamma = 0$ and $a = 10^{-3}$.

We consider the evolution of the angular distribution with respect to effective radial optical depth $\tau_r(x) = \tau_0\phi(x, a)$. The result is plotted in Figure 8.8. The μ-distribution is isotropic if the effective radial optical depth is large and it will no longer be isotropic when the depth becomes small. If we define the transition between isotropic and anisotropic μ-distribution occurs when $I(r, x, 1) = 2I(r, x, -1)$, then at the transition, the critical optical depth is $\tau_{crit} \approx 5$, as can be found from table 8.1.

Table 8.1: critical effective optical depth $\tau_{\rm crit}$

r	100	200	300	400	500
x	1.6	1.8	1.8	2.0	2.0
$\tau_r(x)$	4.41	4.46	6.69	4.17	5.21

8.3.5 Effect of anisotropic scattering

All calculations in the previous sections are based on the re-distribution function equation (6.3), which consider only isotropic scattering. If we consider dipole scattering, it seems to introduce a new factor leading to anisotropy and yield new anisotropic behavior. However, HI atoms are in thermal equilibrium, and their distribution is isotropic. The dipole scattering, as average, does not contain any parameter of specific direction. It will not add any anisotropic behavior. Therefore, all conclusions in the previous sections should still hold.

8.4 Conclusion

The transfer of Lyα resonant photons from a central source in a halo consisting of HI generally is considered as a problem of radiative transfer in an optically thick medium. However, the "optically thick medium" assumption is true only when the frequency of Lyα photons lies in a narrow range $|x| \leq 2$. The cross section of resonant scattering is very sensitive to the photon frequency. It quickly becomes small when the frequency of Lyα photons has only a small deviation from the range $|x| \leq 2$. For those photons, the halo is optically moderate thick, or even thin. Therefore, in order to understand the transfer of Lyα photons with frequency around the resonant peak, we need to find the solutions of the integro-differential equation (8.1) in optically thick as well as moderate thick and even thin medium. That is, although the halo is optically thick for resonant photons, one should not treat (8.1) by using the condition of optical thick.

To find solution of (8.1) with desired precision in frequency ranges of optically thick as well as moderately thick, the algorithm must handle the extremely flat distribution ($|x| < 2$) and its sharp boundary ($|x| \sim 2$) of I. These features can be properly captured by the state-of-the-art numerical method, WENO scheme, as it has high order of accuracy and good convergence in capturing discontinuities as well as being superior to piecewise smooth solutions containing discontinuities. The WENO solver has been shown to be powerful to solve the integro-differential equation of radiative transfer of resonant photons Lyα. In this chapter, we develop the WENO algorithm to be able to solve the integral-differential equation (8.1) in frequency and angular space simultaneously.

We have first shown that the Eddington approximation can yield reasonable results of the frequency profile of photons emergent from optically thick halos. Since the Eddington approximation assumes I is linearly depends on μ, all the physics of the angular distribution of Lyα photons are missing. A cost of the Fokker-Planck equations is also to ignore all the effects of the evolutions of angular distribution.

The physics of the evolution of the angular distribution is rich. As has been known, resonant scattering couples the transfers of resonant photons in the physical space and the frequency space. We show, in this chapter, that the resonant scattering leads to the coupling between the evolutions of resonant photons in the frequency space and the angular space as well. The evolution of the resonant photon distribution in the μ space is strongly dependent on the frequency. Photons with frequency $|x| \leq 2$ undergo the procedure of approaching statistical equilibrium, and their angular distribution is isotropic after a few tens or hundred scattering, regardless of whether the initial angular distribution is isotropic. On the other hand, the angular distribution of photons with $|x| \geq 2$ is strongly anisotropic, even if the initial angular distribution is isotropic.

An interesting feature is that the anisotropic angular distributions at frequency

$|x| \sim 2$ are independent of the initial angular distributions. Different initial angular distributions yield the same anisotropic angular distributions after a few tens or hundred scattering. This is because photons at frequency $|x| \sim 2$ are not directly from the source, but come from the trapped photons within $|x| \leq 2$, for which the initial distributions have been forgotten. Therefore, it seems to be impossible to recover the property of the source with the observed μ-distribution of Lyα photons either in the range of $|x| \leq 2$ or in $|x| \geq 2$.

Another interesting feature of an optically thick halo is the collimation of photons with frequencies of the double peaks. This is also because photons trapped in $|x| \leq 2$ are thermal. When the trapped photons diffuse to $|x| \geq 2$, they have two possible fates. One is to get out of the holes by a single shot if photons move forward. If a photon has not taken a single shot, the resonant scattering will lead it back to the region of $|x| \leq 2$. Therefore, photon transfer in optically thick medium is a collimator. Although photons stored in an optically thick halo are thermal, the μ-distribution is isotropic and the double peak only picks up photons of a single shot, i.e. moving forward.

Bibliography

[1] T. Adams, *The escape of resonance-line radiation from extremely opaque media*, The Astrophysical Journal, 174 (1972), 439-448.

[2] T. Adams, *The mean photon path length in extremely opaque media*, The Astrophysical Journal, 201 (1975), 350-351.

[3] T. Adams, D.G. Hummer, G.B. Rybicki *Numerical evaluation of the redistribution function $R_{II-A}(x, xy)$ and of the associated scattering integral*, Journal of Quantitative Spectroscopy and Radiative Transfer, 11 (1971), 1365-1376.

[4] S. Adjerid and M. Baccouch, *Asymptotically exact a posteriori error estimates for a one-dimensional linear hyperbolic problem*, Applied Numerical Mathematics 60 (2010), 903-914.

[5] S. Adjerid, K. Devine, J. Flaherty and L. Krivodonova, *A posteriori error estimation for discontinuous Galerkin solutions of hyperbolic problems*, Computational Methods in Applied Mechanics and Engineering 191 (2002), 1097-1112.

[6] S. Adjerid and T. Massey, *Superconvergence of discontinuous Galerkin solutions for a nonlinear scalar hyperbolic problem*, Computer Methods in Applied Mechanics and Engineering 195 (2006), 3331-3346.

[7] S. Adjerid and T. Weinhart, *Discontinuous Galerkin error estimation for linear symmetric hyperbolic systems*, Computer Methods in Applied Mechanics and Engineering 198 (2009), 3113-3129.

[8] S. Adjerid and T. Weinhart, *Discontinuous Galerkin error estimation for linear symmetrizable hyperbolic systems*, Mathematics of Computations 80 (2011), 1335-1367.

[9] L.W. Avery and L. House, *An investigation of resonance-line scattering by the Monte Carlo technique*, The Astrophysical Journal, 152 (1968), 493-508.

[10] C. Berthon, M. Breuss and M.-O. Titeux, *A relaxation scheme for the approximation of the pressureless Euler equations*, Numerical Methods for Partial Differential Equations, 22, (2006), 484-505.

[11] G. Blanc et. al., *The HETDEX pilot survey. II. the evolution of the Lyα escape fraction form the ultraviolet slope and luminosity function of 1.9¡z¡3.8 LAEs*, The Astrophysical Journal, 736 (2010), 31(21).

[12] J.R.M. Bonilha et. al., *Monte Carlo calculation for tesonance scattering with absorption or differential expansion*, The Astrophysical Journal, 233 (1979), 649-660.

[13] F. Bouchut, *On zero pressure gas dynamics*, Advances in Kinetic Theorey and Computing, Would Scientific, River Edge, NJ, (1994), 171-190.

[14] F. Bouchut and F. James *Duality solutions for pressureless gases, monotone scalar conservation laws, and uniqueness*, Communications in Partial Differential Equations, 24, (1999), 2173-2189.

[15] F. Bouchut, S. Jin and X. Li, *Numerical approximations of pressureless and isothermal gas dynamics*, SIAM Journal on Numerical Analysis, 41, (2003), 135-158.

[16] L. Boudin and J. Mathiaud, *A numerical scheme for the one-dimensional pressureless gases system*, Numerical Methods for Partial Differential Equations, 28, (2006), 1729-1746.

[17] J.H. Bramble and A.H. Schatz, *High order local accuracy by averaging in the finite element method*, Mathematics of Computation, 31 (1977), 94-111.

[18] Y. Brenier and E. Grenier, *Sticky particles and scalar conservation laws*, SIAM Journal on Numerical Analysis, 35, (1998), 2317-2328.

[19] C. Canuto, F. Fagnani and P. Tilli, *An Eulerian approach to the analysis of Krause's consensus models*, SIAM Journal on Control and Optimization, 50, (2012), 243-265.

[20] J. Carrillo, *2D semiconductor device simulations by WENO-Boltzman schemes:Efficiency, boundary conditions and comparison to Monte Carlo methods*, Journal of Computational Physics, 214 (2006), 55-80.

[21] A. Chalaby, *On convergence of numerical schemes for hyperbolic conservation laws with stiff source terms*, Mathematics of Computation, 66 (1997), 527-545.

[22] G.-Q. Chen and H. Liu, *Formation of δ-shocks and vacuum states in the vanishing pressure limit of solutions to the Euler equations for isentropic fluids*, SIAM Journal on Mathematical Analysis, 34, (2003), 925-938.

[23] J. Cheng and C.-W. Shu, *A cell-centered Lagrangian scheme with the preservation of symmetry and conservation properties for compressible fluid flows in two-dimensional cylindrical geometry*, Journal of Computational Physics, 229, (2010), 7191-7206.

[24] J. Cheng and C.-W. Shu, *Improvement on spherical symmetry in two-dimensional cylindrical coordinates for a class of control volume Lagrangian schemes*, Communications in Computational Physics, 11, (2012), 1144-1168.

[25] Y. Cheng and C.-W. Shu, *Superconvergence and time evolution of discontinuous Galerkin finite element solutions*, Journal of Computational Physics, 227 (2008), 9612-9627.

[26] Y. Cheng and C.-W. Shu, *Superconvergence of discontinuous Galerkin and local discontinuous Galerkin schemes for linear hyperbolic and convection-diffusion equations in one space dimension*, SIAM Journal on Numerical Analysis, 47 (2010), 4044-4072.

[27] A. Chertock, A. Kurganov and Y. Rykov, *A new sticky particle method for pressureless gas dynamics*, SIAM Journal on Numerical Analysis, 45, (2007), 2408-2441.

[28] P.G. Ciarlet. *Finite Element Method for Elliptic Problems*, North-Holland, Amsterdam, 1978.

[29] B. Cockburn et. al., *In advanced numerical approximation of nonlinear hyperbolic equations*, ed. A. Quarteroni (Lecture Notes in Mathematics, Vol. 1697 Berlin: Springer), 450

[30] B. Cockburn, *A Introduction to the discontinuous Galerkin methods for convection dominated problems*, High-Order Method for Computational Physics (T. Barth and H. Deconink, eds.), Lecture Notes in Computational Science and Engineering, vol. 9, Springer-Verlag Berlin Heidelberg, 1999, 69-224.

[31] B. Cockburn and J. Guzmán, *Error estimate for the Runge-Kutta discontinuous Galerkin method for the transport equation with discontinuous initial data*, SIAM Journal on Numerical Analysis, 46 (2008), 1364-1398.

[32] B. Cockburn, S. Hou and C.-W. Shu, *The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws, IV: the multidimensional case*, Mathematics of Computation, 54 (1990), 545-581.

[33] B. Cockburn, S.-Y. Lin and C.-W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems*, Journal of Computational Physics, 84 (1989), 90-113.

[34] B. Cockburn, M. Luskin, C.-W. Shu and E. Süli, *Enhanced accuracy by post-processing for finite element methods for hyperbolic equations*, Mathematics of Computation, 72 (2003), 577-606.

[35] B. Cockburn and C.-W. Shu, *TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws, II: general framework*, Mathematics of Computation, 52 (1989), 411-435.

[36] B. Cockburn and C.-W. Shu, *The Runge-Kutta discontinuous Galerkin method for conservation laws, V: multidimensional system*, Journal of Computational Physics, 35 (1998), 199-224.

[37] M. Dijkstra, Z. Haiman, M. Spaans, *Lyα radiation from collapsing protogalaxies*, The Astrophysical Journal, 649 (2006), 14-36.

[38] M. Dijkstra, A. Loeb, *Lyα blobs as an observational signature of cold accretion streams into galaxies*, Monthly Notices of the Royal Astronomical Society, 400 (2009), 1109-1120

[39] B.T. Draine, *Scattering by interstellar dust grains. I. optical and ultraviolet*, The Astrophysical Journal, 598 (2003), 1017-1025.

[40] B.T. Draine and H.M. Lee *Optical properties of interstellat graphite and silicate grains*, The Astrophysical Journal, 285 (1984), 89-108.

[41] W. 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*, Communications in Mathematical Physics, 177 (1996), 349-380.

[42] L.Z. Fang *The zeroth law of thermodynamics of the photon-hydrogen system and 21cm cosmology*, International Journal of Modern Physics D, 18 (2009), 1943-1954.

[43] M. Fardal *Cooling radiation and the Lyα luminosity of forming galaxies*, The Astrophysical Journal, 562 (2001), 605-617.

[44] G.B. Field *Excitation of the hydrogen 21-cm line*, Proceedings of the IRE, 46 (1958), 240-250.

[45] G.B. Field *The time relaxation fo a resonance-line profile*, The Astrophysical Journal, 129 (1959), 551-564.

[46] S. Gottlieb, C.-W. Shu and E. Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM Review, 43 (2001), 89-112.

[47] J.M. Greenberg, A.Y. Leroux, R. Baraille and A. Noussair, *Analysis and approximation of conservation laws with source terms*, SIAM Journal on Numerical Analysis, 34 (1997), 1980-2007.

[48] Z. Haiman and M. Spaans, *Models for dusty Lyα emitters at high redshift*, The Astrophysical Journal, 518 (1999), 138-144.

[49] Z. Haiman, M. Spaans and E. Quataert, *Lyα cooling radiation from high-redshift halos*, The Astrophysical Journal, 537 (2000), L5-L8.

[50] M. Hansen and S.P. Oh, *Lyα radiative transfer in a multiphase medium*, Monthly Notices of the Royal Astronomical Society, 367 (2006), 979-1002.

[51] J.P. Harrington, *The scattering of resonance-line radiation in the limit of large optical depth*, Monthly Notices of the Royal Astronomical Society, 162 (1973), 43-52.

[52] M. Hayes, *Escape of about five per cent of Lyα photons froom high-redshift star-forming galaxies*, Nature, 464 (2010), 562-565.

[53] M. Hayes, *On the redshift evolution of the Lyα escape fraction and the dust content of galaxies*, The Astrophysical Journal, 730 (2011), 8(13).

[54] L.G. Henyey and J.L. Greenstein, *Diffuse radiation in the galaxy*, The Astrophysical Journal, 93 (1941), 70-83.

[55] S. Hou and X.-D. Liu, *Solutions of multi-dimensional hyperbolic systems of conservation laws by square entropy condition satisfying discontinuous Galerkin method*, Journal of Scientific Computing, 31 (2007), 127-151.

[56] D.G. Hummer, *Non-coherent scattering: I. The redistribution function with Doppler broadening*, Monthly Notices of the Royal Astronomical Society, 125 (1962), 21-37.

[57] D.G. Hummer, *The Voigt function: An eight-significant-figure table and generating procedure*, Memoirs of the Royal Astronomical Society, 70 (1965), 1-32.

[58] D.G. Hummer, *Non-coherent scattering: VI. Solution of the transfer problem with a frequency-dependent source function*, Monthly Notices of the Royal Astronomical Society, 145 (1969), 95-120.

[59] D.G. Hummer and P.B. Kunasz, *Energy loss by resonance line photons in an absorbing medium*, The Astrophysical Jornal, 236 (1980), 609-618.

[60] G.-S. Jiang and C.-W. Shu, *On a cell entropy inequality for discontinuous Galerkin methods*, Mathematics of Computation, 62 (1994), 531-538.

[61] G.-S. Jiang and C.-W. Shu, *Efficient implementation of weighted ENO schemes*, Journal of Computational Physics, 126 (1996), 202-228.

[62] F. John, *Partial different equations*, Springer-Verlag, New York, 1971

[63] C. Johnson, U. Nävert and J. Pitkäranta, *Finite element methods for linear hyperbolic problems*, Computer Methods in Applied Mechanics and Engineering, 45 (1984), 285-312.

[64] C. Johnson and J. Pitkäranta, *An analysis of the discontinuous Galerkin method for a scalar hyperbolic equation*, Mathematics of Computation, 46 (1986), 1-26.

[65] C. Johnson, A. Schatz and L. Walhbin, *Crosswind smear and pointwise errors in streamline diffusion finite element methods*, Mathematics of Computation, 49 (1987), 25-38.

[66] B. Koren, *A robust upwind discretization method for advection, diffusion and source terms*, Notes on Numerical Fluid Mechanics, 45, Vieweg, Braunschweig (1993), 117-138.

[67] L. Krivodonova, J. Xin, J.-F. Remacle, N. Chevaugeon and J.E. Flaherty, *Shock detection and limiting with discontinuous Galerkin methods for hyperbolic conservation laws*, Applied Numerical Mathematics, 48 (2004), 323-338.

[68] P. Kurasov, *Distribution theory for discontinuous test functions and differential operators with generalized coefficients*, Journal of Mathematical Analysis and Applications, 201 (1996), 297-323.

[69] M. Latif, et. al., *Lyα emission from the first galaxies: signatures of accretion and infall in the presence of line trapping*, Monthly Notices of the Royal Astronomical Society: Letters, 413 (2011), L33-L37.

[70] P. Laursen and J. Sommer-Larsen, *Lyα resonsant scattering in young galaxies: Predictions from cosmological simulations*, The Astrophysical Journal, 657 (2007), L69-L72.

[71] P. Laursen, J. Sommer-Larsen and A. Andersen, *Lyα radiative transfer with dust: escape fractions from simulated high-redshift galaxies*, The Astrophysical Jornal, 704 (2009), 1640-1656.

[72] R.J. LeVeque and H.C. Yee, *A study of numerical methods for hyperbolic conservation laws with stiff source terms*, Journal of Computational Physics, 86 (1990), 187-210.

[73] A. Li, B.T. Draine, *On ultrasmall silicate grains in the diffuse interstellar medium*, The Astrophysical Jornal, 550 (2001), L213-L217.

[74] J. Liu, et. al., *21cm signals from early ionizing sources*, The Astrophysical Jornal, 663 (2007), 1-9.

[75] A. Loeb and G.B. Rybicki, *Scattered Lyα radiation around sources before cosmological reionization*, The Astrophysical Journal, 524 (1999), 527-535.

[76] D. Neufeld, *The transfer of resonance-line radiation in static astrophysical media*, The Astrophysical Jornal, 350 (1990), 216-241.

[77] D. Neufeld, *The escape of Lyα radiation from a multiphase interstellar medium*, The Astrophysical Jornal, 370 (1991), L85-L88.

[78] A. Noussair, *Analysis of nonlinear resonance in conservation laws with point sources and well-balanced scheme*, Studies in Applied Mathematics, 104 (2000), 313-352.

[79] P. de Oliveira and J. Santos, *On a class of high resolution methods for solving hyperbolic conservation laws with source terms*, Applied Nonlinear Analysis, 432-445, (Adélia Sequeira, Hugo Beirão da Veiga, and Juha Hans Videman, eds.), Kluwer academic publishers, New York, Boston, Dordrecht, London, Moscow, 1999.

[80] D.E. Osterbrock, *The escape of resonance-line radiation from an optically thick nebula*, The Astrophysical Jornal, 135 (1962), 195-216.

[81] Y. Pei, *Interstellar dust from the milky way to the magellanic clouds*, The Astrophysical Jornal, 395 (1992), 130-139.

[82] M. Pierleoni, A. Maselli and B. Ciardi, *Crashα: coupling continuum and line radiative transfer*, Monthly Notice of the Royal Astronomical Society, 393 (2009), 872-884.

[83] J. Qiu, et. al., *A WENO algorithm for the radiative transfer and ionized sphere at reionization*, New Astronomy, 12 (2006), 1-10.

[84] J. Qiu, et. al., *A WENO algorithm of the temperature and ionization profiles around a point source*, New Astronomy, 12 (2007), 398-409.

[85] J. Qiu, et. al., *A WENO algorithm for the growth of ionized regions at the reionization epoch*, New Astronomy, 13 (2008), 1-11.

[86] W.H. Reed and T.R. Hill, *Triangular mesh for the Neutron transport equation*, Los Alamos Scientific Laboratory Report LA-UR-73-479, Los Alamos, NM, 1973.

[87] I. Roy, et. al., *A WENO algorithm for radiative transfer with resonant scattering and the Wouthuysen-Field coupling*, New Astronomy, 14 (2009), 513-520.

[88] I. Roy, et. al., *Time evolution of Wouthuysen-Field coupling*, The Astrophysical Journal, 694 (2009), 1121-1130.

[89] I. Roy, et. al., *Wouthuysen-Field coupling in the 21 cm region around high-redshift sources*, The Astrophysical Journal, 704 (2009), 1992-2003.

[90] I. Roy, C.-W. Shu and L.-Z. Fang, *Resonant scattering and Lyα radiation emergent from nertral hydrogen halos*, The Astrophysical Journal, 716 (2010), 604-614.

[91] J.K. Ryan and B. Cockburn, *Local derivative post-processing for the discontinuous Galerkin method*, Journal of Computational Physics 228 (2009), 8642-8664.

[92] J.K. Ryan and C.-W. Shu, *On a one-sided post-processing technique for the discontinuous Galerkin methods*, Methods and Applications of Analysis, 10 (2003), 295-308.

[93] G.B. Rybicki, *Improved Fokker-Planck equation for resonance-line scattering*, The Astrophysical Journal, 647 (2006), 709-718.

[94] G.B. Rybicki and I. dell'Antonio, *The time development of a resonance line in the expanding universe*, The Astrophysical Journal, 427 (1994), 603-617.

[95] G.B. Rybicki and A.P. Lightman, Radiative Processes in Astrophysics, (New York: Wiley).

[96] Yu. G. Rykov, *Propagation of shock wave type singularities in equations of two-dimensional zero-pressure gas dynamics*, Mathematical notes, 66 (1999), 628-635.

[97] Yu. G. Rykov, *On the nonhamiltonian character of shocks in 2-D pressureless gas*, Bollettino della Unione Matematica Italiana. Serie VIII. Sezione B. Articoli di Ricerca Matematica, 5 (2002), 55-78.

[98] J. Santos and P. de Oliveira, *A converging finite volume scheme for hyperbolic conservation laws with source terms*, Journal of Computational and Applied Mathematics, 111 (1999), 239-251.

[99] H.J. Schroll and R. Winther, *Finite-difference Schemes for scalar conservation laws with source terms*, IMA Journal of Numerical Analysis, 16 (1996), 201-215.

[100] J. Shen, C.-W. Shu, M. Zhang, *A high order WENO scheme for a hierarchical size-structured population*, Journal of Scientific Computing, 33 (2007), 279-291.

[101] C.-W. Shu, *Total-variation-diminishing time discretizations*, SIAM Journal on Scientific and Statistical Computing, 9 (1988), 1073-1084.

[102] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock-capturing schemes*, Journal of Computational Physics, 77 (1988), 439-471.

[103] P.V. Slingerland and J.K. Ryan and C. Vuik *Position-dependent smooth-increasing accuracy-conserving (SIAC) filtering for improving discontinuous Galerkin solutions*, SIAM Journal on Scientific Computing, 33 (2011), 802-825.

[104] M. Spaans and J. Silk, *Pregalactic black hole formation with an atomic hydrogen equation of state*, The Astrophysical Journal, 652 (2006), 902-906.

[105] L. Spitzer and J. Greenstein, *Continuous emission from planetary nebulae*, The Astrophysical Journal, 114 (1951), 407-420.

[106] G. Stratta, et. al., *Dust properties at z=6.3 in the host galaxy of GRB 050904*, The Astrophysical Journal, 661 (2007), L9-L12.

[107] A. Tasitsiomi, *Lyα radiative transfer in cosmological simulations and application to a z 8 Lyα emitter*, The Astrophysical Journal, 645 (2006), 792-813.

[108] W. Unno, *Theoretical line contour of the Lyα radiation of ionized helium and the excitation of Bowen lines in planetary nebulae*, Publications of the astronomical society of Japan, 7 (1955), 81-103.

[109] A. Verhamme, *3D Lyα radiation transfer. III. constraints on gas and stellar properties of z 3 Lyα break galaxies (LBG) and implications for high-z LBGs and Lyα emmitters*, The Astrophysical Journal, 491 (2008), 89-111.

[110] A. Verhamme, D. Schaerer and A. Maselli, *3D Lyα radiation transfer. I. Understanding Lyα line profile morphologies*, The Astrophysical Journal, 460 (2006), 397-413.

[111] D. Walfisch, J.K. Ryan, R.M. Kirby and R. Haimes, *One-sided smoothness-increasing accuracy-conserving filtering for enhanced streamline integration*

through discontinuous fields, Journal of Scientific Computing, 38 (2009), 164-184.

[112] J.C. Weingartner and B.T. Draine, *Dust grain-size distribution and extinction in the milky way, large magellanic cloud, and small magellaanic cloud*, The Astrophysical Journal, 548 (2001), 296-309.

[113] S.A. Wouthuysen, *On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line*, The Astrophysical Journal, 57 (1952), 31-32.

[114] W. Xu and X.-P. Wu, *On the formation of Lyα emission from resonantly scattered continnum photons of gamma-ray burst's afterglow*, The Astrophysical Journal, 710 (2010), 1432-1443.

[115] W. Xu and X.-P. Wu and L.-Z. Fang, *Time-dependent behaviour of Lyα photon transfer in a high-redshift optically thick medium*, Monthly Notice of the Royal Astronomical Society, 418 (2011), 853-862.

[116] M. Zhang and C.-W. Shu, *An analysis of and a comparison between the discontinuous Galerkin and the spectral finite volume methods*, Computers and Fluids, 34 (2005), 581-592.

[117] Q. Zhang and C.-W. Shu, *Error estimates for the third order explicit Runge-Kutta discontinuous Galerkin method for linear hyperbolic equation in one-dimension with discontinuous initial data*, submitted to Numerische Mathematik.

[118] Q. Zhang and C.-W. Shu, *Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws*, SIAM Journal on Numerical Analysis, 42 (2004), 641-666.

[119] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, Journal of Computational Physics, 229 (2010), 3091-3120.

[120] X. Zhang and C.-W. Shu, *On positivity preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes*, Journal of Computational Physics, 229 (2010), 8918-8934.

[121] X. Zhang and C.-W. Shu, *Maximum-principle-satisfying and positivity-preserving high order schemes for conservation laws: Survey and new developments*, Proceedings of the Royal Society A, 467 (2011), 2752-2776.

[122] Z. Zheng and J. Miralda-Escude, *Monte Carlo simulation of Lyα scattering and application to damped Lyα systems*, The Astrophysical Journal, 578 (2002), 33-42.

[123] X. Zhong and C.-W. Shu, *Numerical resolution of discontinuous Galerkin methods for time dependent wave equations*, Computer Methods in Applied Mechanics and Engineering, 200 (2011), 2814-2827.