Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational Physics 214 (2006) 672–697
JOURNAL OF COMPUTATIONAL PHYSICS
www.elsevier.com/locate/jcp

Hamiltonian-preserving schemes for the Liouville equation of geometrical optics with discontinuous local wave speeds

Shi Jin [a,b,*], Xin Wen [b,1]

[a] Department of Mathematics, University of Wisconsin – Madison, Van Vleck Hall, 480 Lincoln Drive, Madison, WI 53706-1388, USA
[b] Department of Mathematical Sciences, Tsinghua University, Beijing 100084, PR China

Received 8 June 2005; received in revised form 11 October 2005; accepted 12 October 2005
Available online 6 December 2005

Abstract

In this paper, we construct two classes of Hamiltonian-preserving numerical schemes for a Liouville equation with discontinuous local wave speed. This equation arises in the phase space description of geometrical optics, and has been the foundation of the recently developed level set methods for multivalued solution in geometrical optics. We extend our previous work in [S. Jin, X. Wen, Hamiltonian-preserving schemes for the Liouville equation with discontinuous potentials, Commun. Math. Sci. 3 (2005) 285–315] for the semiclassical limit of the Schrödinger equation into this system. The designing principle of the Hamiltonian preservation by building in the particle behavior at the interface into the numerical flux is used here, and as a consequence we obtain two classes of schemes that allow a hyperbolic stability condition. When a plane wave hits a flat interface, the Hamiltonian preservation is shown to be equivalent to Snell's law of refraction in the case when the ratio of wave length over the width of the interface goes to zero, when both length scales go to zero. Positivity, and stabilities in both l^1 and l^∞ norms, are established for both schemes. The approach also provides a selection criterion for a unique solution of the underlying linear hyperbolic equation with singular (discontinuous and measure-valued) coefficients. Benchmark numerical examples are given, with analytic solution constructed, to study the numerical accuracy of these schemes.

Keywords: Geometrical optics; Liouville equation; Level set method; Hamiltonian-preserving schemes; Snell's law of refraction

1. Introduction

In this paper, we construct and study numerical schemes for the Liouville equation in d-dimension:

$$f_t + \nabla_{\mathbf{v}} H \cdot \nabla_{\mathbf{x}} f - \nabla_{\mathbf{x}} H \cdot \nabla_{\mathbf{v}} f = 0, \quad t > 0, \ \ \mathbf{x}, \mathbf{v} \in R^d, \tag{1.1}$$

* Corresponding author. Tel.: +1 608 263 3302; fax: +1 608 263 8891.
E-mail addresses: jin@math.wisc.edu (S. Jin), wenxin@amss.ac.cn (X. Wen).
[1] Present address: Institute of Computational Mathematics and Scientific and Engineering Computing, Chinese Academy of Science, Beijing 100080, China.

0021-9991/$ - see front matter
doi:10.1016/j.jcp.2005.10.012

where the Hamiltonian H is given by

$$H(\mathbf{x},\mathbf{v},t) = c(\mathbf{x})|\mathbf{v}| = c(\mathbf{x})\sqrt{v_1^2 + v_2^2 + \cdots + v_d^2} \tag{1.2}$$

with $c(\mathbf{x}) > 0$ being the local wave speed. $f(t,\mathbf{x},\mathbf{v})$ is the density distribution of particles depending on position $\mathbf{x}$, time t and the slowness vector $\mathbf{v}$. In this paper, we are interested in the case when $c(\mathbf{x})$ contains *discontinuities* corresponding to different indices of refraction at different media. This discontinuity will generate an *interface* at the point of discontinuity of $c(\mathbf{x})$, and as a consequence waves crossing this interface will undergo transmissions and reflections. The incident and transmitted waves obey Snell's law of refraction.

The bicharacteristics of the Liouville equation (1.1) satisfy the Hamiltonian system:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = c(\mathbf{x})\frac{\mathbf{v}}{|\mathbf{v}|}, \quad \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = -|\mathbf{v}|\nabla_{\mathbf{x}} c. \tag{1.3}$$

In classical mechanics the Hamiltonian (1.2) of a particle remains a constant along particle trajectory, even across an interface.

This Liouville equation arises in the phase space description of geometrical optics. It is the high frequency limit of the wave equation

$$u_{tt} - c(\mathbf{x})^2\Delta u = 0, \quad t > 0, \ \ \mathbf{x} \in R^d. \tag{1.4}$$

Recently, several phase space based level set methods are based on this equation, see [13,16,22,31]. High frequency limit of wave equations with transmissions and reflections at the interfaces was studied in [1,30,39]. A Liouville equation based level set method for the wave front, but with only reflection, was introduced in [7]. It was also suggested to smooth out the local wave speed in [31].

The Liouville equation (1.1) is a linear wave equation, with the characteristic speed determined by bicharacteristic (1.3). If $c(\mathbf{x})$ is smooth, then the standard numerical methods (for example, the upwind scheme and its higher order extensions) for linear wave equations give satisfactory results. However, if $c(\mathbf{x})$ is discontinuous, then the conventional numerical schemes suffer from two problems. Firstly, the characteristic speed $c_{\mathbf{x}}$ of the Liouville equation is *infinity* at the discontinuous point of wave speed. When numerically approximating $c_{\mathbf{x}}$ crossing the interface (for example by smoothing out $c(\mathbf{x})$ [31]), the numerical derivative of c is of $\mathrm{O}(1/\Delta x)$, with Δx the mesh size in the physical space. Thus an explicit scheme needs time step $\Delta t = \mathrm{O}(\Delta x \Delta v)$ with Δv the mesh size in particle slowness space. This is very expensive. Moreover, a conventional numerical scheme in general does not preserve a *constant Hamiltonian* across the interface, usually leading to poor or even incorrect numerical resolutions by ignoring the discontinuities of $c(\mathbf{x})$. Theoretically, there is a uniqueness issue for weak solutions to these linear hyperbolic equations with singular wave speeds [6,9,19,34,35]. It is not clear which weak solution a standard numerical discretization that ignores the discontinuity of $c(\mathbf{x})$ will select.

We also remark that Hamiltonian or sympletic schemes have been introduced for Hamiltonian ODEs and PDEs in order to preserve the Hamiltonian or sympletic structures, see for example [15,28]. To our knowledge, no such schemes have been constructed for Hamiltonian systems with discontinuous Hamiltonians.

In this paper, we construct a class of numerical schemes that are suitable for the Liouville equation (1.1) with a discontinuous local wave speed $c(\mathbf{x})$. An important feature of our schemes is that they are consistent with the constant Hamiltonian across the interface. This gives a selection criterion for a unique solution to the governing equation. As done in [24] for the Liouville equation for the semiclassical limit of the linear Schrödinger equation, we call such schemes *Hamiltonian-preserving schemes*. A key idea of these schemes is to build the behavior of a particle at the interface – either cross over with a changed velocity or be reflected with a negative velocity – into the numerical flux. This idea was formerly used by Perthame and Semioni in their work [33] to construct a well-balanced kinetic scheme for the shallow water equations with a (discontinuous) bottom topography which can capture the steady state solutions – corresponding to a constant energy – of the shallow water equations when the water velocity is zero. As a consequence, these new schemes allow a typical hyperbolic stability condition $\Delta t = \mathrm{O}(\Delta x, \Delta v)$.

We extend both classes of the Hamiltonian-preserving schemes developed in [24] here. One (called *Scheme I*) is based on a finite difference approach, and involves interpolation in the slowness space. The second (called *Scheme II*) uses a finite volume approach, and numerical quadrature rule in the slowness space is needed. These new schemes allow a typical hyperbolic stability condition $\Delta t = \mathrm{O}(\Delta x, \Delta v)$. We will also establish the

positivity and stability theory for both schemes. It is proved that Scheme I is positive, l^∞ contracting, and l^1 stable under a hyperbolic stability condition, while Scheme II is positive, l^∞ stable and l^1 contracting under the same stability condition.

By building in the wave behavior at the interface, we have also provided a selection principle to pick up a unique solution to this linear hyperbolic equation with singular coefficients. For a plane wave hitting a flat interface, we show that it selects the solution that describes the interface condition in geometrical optics governed by *Snell's law of refraction* when the wave length is much shorter than the width of the interface while both lengths go to zero.

In geometrical optics applications, one has to solve the Liouville equation like (1.1) with *measure-valued* initial data

$$f(\mathbf{x},\mathbf{v},0)=\rho_0(\mathbf{x})\delta(\mathbf{v}-\mathbf{u}_0(\mathbf{x})), \tag{1.5}$$

see for example [12,22,38]. The solution at later time remains measure-valued (with finite or even infinite number of concentrations – corresponding to *multivalued* solutions in the physical space). Computation of multivalued solutions in geometrical optics and more generally in nonlinear PDEs has been a very active area of recent research, see [2–5,8,10,11,13,14,16–18,20,23,31,37,41].

Numerical methods for the Liouville equation with measure-valued initial data (1.5) could easily suffer from poor resolution due to the numerical approximation of the initial data as well as numerical dissipation. The level set method proposed in [21,22] decomposes f into ϕ and ψ_i $(i=1,\ldots,d)$ where ϕ and ψ_i solve the same Liouville equation with initial data

$$\phi(\mathbf{x},\mathbf{v},0)=\rho_0(\mathbf{x}),\quad \psi_i(\mathbf{x},\mathbf{v},0)=v_i-u_{i0}(\mathbf{x}), \tag{1.6}$$

respectively. (We remark here that the common *zeroes* of ψ_i give the multivalued slowness, see [8,23,21,22].) This allows the numerical computations for bounded rather than measure-valued solution of the Liouville equation, which greatly enhances the numerical resolution (see [22]). The moments can be recovered through

$$\rho(\mathbf{x},t)=\int f(\mathbf{x},\mathbf{v},t)\,\mathrm{d}\mathbf{v}=\int \phi(\mathbf{x},\mathbf{v},t)\Pi_{i=1}^{d}\delta(\psi_i)\,\mathrm{d}\mathbf{v}, \tag{1.7}$$

$$\mathbf{u}(\mathbf{x},t)=\frac{1}{\rho(\mathbf{x},t)}\int f(\mathbf{x},\mathbf{v},t)\mathbf{v}\,\mathrm{d}\mathbf{v}=\int \phi(\mathbf{x},\mathbf{v},t)\mathbf{v}\Pi_{i=1}^{d}\delta(\psi_i)\,\mathrm{d}\mathbf{v}/\rho(\mathbf{x},t). \tag{1.8}$$

Thus one only involves numerically the delta-function at the output time!

Numerical computations of multivalued solution for smooth $c(\mathbf{x})$ using this technique were given in [22]. In this paper, we will also give numerical examples using this technique with a discontinuous $c(\mathbf{x})$.

The more general case with partial transmissions and reflections will be studied in a forthcoming paper [26].

This paper is organized as follows. In Section 2, we first show that the usual finite difference scheme to solve the Liouville equation with a discontinuous wave speed suffers from the severe stability constraint. We then present the design principle of our Hamiltonian-preserving scheme by describing the behavior of waves at an interface. We present Scheme I in one space dimension in Section 3 and study its positivity and stability in both l^∞ and l^1 norms. Scheme II in one space dimension is presented and studied in Section 4. We extend these schemes to two space dimension in Section 5 in the simple case of interface aligning with the grids and a plane wave. Numerical examples, with analytical solutions constructed, are given in Section 6 to verify the accuracy of the schemes. For comparison, we also present numerical solutions by methods ignoring or smearing the discontinuity of c. We make some concluding remarks in Section 7.

2. The design principle of the Hamiltonian-preserving scheme

2.1. Deficiency of the usual finite difference schemes

Consider the numerical solution of the Liouville equation in one physical space dimension

$$f_t+c(x)\,\mathrm{sign}(\xi)f_x-c_x|\xi|f_\xi=0 \tag{2.1}$$

with a discontinuous wave speed $c(x)$.

We employ a uniform mesh with grid points at $x_{i+\frac{1}{2}}$, $i = 0, \ldots, N$, in the x-direction and $\xi_{j+\frac{1}{2}}$, $j = 0, \ldots, M$ in the ξ-direction. The cells are centered at (x_i, ξ_j), $i = 1, \ldots, N$, $j = 1, \ldots, M$ with $x_i = \frac{1}{2}(x_{i+\frac{1}{2}} + x_{i-\frac{1}{2}})$ and $\xi_j = \frac{1}{2}(\xi_{j+\frac{1}{2}} + \xi_{j-\frac{1}{2}})$. The mesh size is denoted by $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$, $\Delta\xi = \xi_{j+\frac{1}{2}} - \xi_{j-\frac{1}{2}}$. We also assume a uniform time step Δt and the discrete time is given by $0 = t_0 < t_1 < \cdots < t_L = T$. We introduce mesh ratios $\lambda_x^t = \frac{\Delta t}{\Delta x}$, $\lambda_\xi^t = \frac{\Delta t}{\Delta \xi}$, assumed to be fixed. We define the cell average of f as

$$f_{ij} = \frac{1}{\Delta x\, \Delta\xi} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \int_{\xi_{j-\frac{1}{2}}}^{\xi_{j+\frac{1}{2}}} f(x, \xi, t)\, \mathrm{d}\xi\, \mathrm{d}x.$$

A typical semi-discrete finite difference method for this equation is

$$\partial_t f_{ij} + c_i \operatorname{sign}(\xi_j)\, \frac{f_{i+\frac{1}{2},j} - f_{i-\frac{1}{2},j}}{\Delta x} - Dc_i\, |\xi_j|\, \frac{f_{i,j+\frac{1}{2}} - f_{i,j-\frac{1}{2}}}{\Delta\xi} = 0, \tag{2.2}$$

where the numerical fluxes $f_{i+\frac{1}{2},j}$, $f_{i,j+\frac{1}{2}}$ are defined by the upwind scheme, and Dc_i is some numerical approximation of c_x at $x = x_i$.

Such a discretization suffers from two problems:

- If an explicit time discretization is used, the CFL condition for this scheme requires the time step to satisfy
$$\Delta t \max_i \left[\frac{c_i}{\Delta x} + \frac{|Dc_i| \max_j |\xi_j|}{\Delta\xi} \right] \leqslant 1. \tag{2.3}$$
Since the wave speed $c(x)$ is discontinuous at some points, $\max_i |Dc_i| = \mathrm{O}(1/\Delta x)$, so the CFL condition (2.3) requires $\Delta t = \mathrm{O}(\Delta x \Delta\xi)$, which is too expensive for a hyperbolic problem.
- The above discretization in general does not preserve a constant Hamiltonian $H = c|\xi|$ across the discontinuities of c, thus may not produce numerical solutions consistent with, for example, Snell's law of refraction.

2.2. Behavior of waves at the interface

When a wave moves with its density distribution governed by the Liouville equation (1.1), *the Hamiltonian $H = c|\mathbf{v}|$ should be preserved across the interface:*

$$c^+|\mathbf{v}^+| = c^-|\mathbf{v}^-|, \tag{2.4}$$

where the superscripts $\pm$ indicate the right and left limits of the quantity at the interface.

We will discuss the wave behavior in one and two space dimensions, respectively.

- One space dimension. The 1D case is simple. Consider the case when, at an interface, the characteristic on the left of the interface is given by $\xi^- > 0$. Then the particle definitely crosses the interface and $\xi^+ = \frac{c^-}{c^+}\xi^-$.
- Two space dimension, when an incident plane wave hits an interface that aligns with the grid. In the 2D case, $\mathbf{x} = (x, y)$, $\mathbf{v} = (\xi, \eta)$. Consider the case that the interface is a line parallel to the y-axis. The incident wave has slowness (ξ^-, η^-) to the left side of the interface, with $\xi^- > 0$. Since the interface is vertical, (1.3) implies that η will not change across the interface, while ξ has three possibilities:
  (1) $c^- > c^+$. In this case, the local wave speed decreases, so the wave will cross the interface and increase its ξ value in order to maintain a constant Hamiltonian. (2.4) implies
  $$\xi^+ = \sqrt{\left(\frac{c^-}{c^+}\right)^2 (\xi^-)^2 + \left[\left(\frac{c^-}{c^+}\right)^2 - 1\right](\eta^-)^2}.$$
  (2) $c^- < c^+$ and $(\frac{c^-}{c^+})^2(\xi^-)^2 + [(\frac{c^-}{c^+})^2 - 1](\eta^-)^2 > 0$. In this case the wave can also cross the interface with a reduced ξ value. (2.4) still gives

$$\xi^+ = \sqrt{\left(\frac{c^-}{c^+}\right)^2 (\xi^-)^2 + \left[\left(\frac{c^-}{c^+}\right)^2 - 1\right](\eta^-)^2}.$$

(3) $c^- < c^+$ and $(\frac{c^-}{c^+})^2(\xi^-)^2 + [(\frac{c^-}{c^+})^2 - 1](\eta^-)^2 < 0$. In this case, there is no possibility for the wave to cross the interface, so the wave will be reflected with slowness $(-\xi^-, \eta^-)$.

If $\xi^- < 0$, similar behavior can also be analyzed using the constant Hamiltonian condition (2.4).

Remark 2.1. In general, one cannot define a unique weak solution to a linear hyperbolic equation with singular (discontinuous or measure-valued) coefficients. By using the wave behavior described above, we give a *selection criterion* for a unique solution. This solution is the one when the wave length of the incident wave is much smaller than the width of the interface, both of which go to zero. It is equivalent to Snell's law of refraction:

$$\frac{\sin\theta_i}{c^-} = \frac{\sin\theta_t}{c^+}, \tag{2.5}$$

where θ_i and θ_t stand for angles of incident and transmitted waves, respectively. This is to say:

$$\frac{\eta^-}{c^-\sqrt{(\xi^-)^2 + (\eta^-)^2}} = \frac{\eta^+}{c^+\sqrt{(\xi^+)^2 + (\eta^+)^2}}. \tag{2.6}$$

If $c = c(x)$, then (1.3) implies that

$$\eta^+ = \eta^-. \tag{2.7}$$

Clearly (2.6) and (2.7) imply (2.4).

Of course this is not the only physically relevant way to choose a solution. In particular, this principle excludes the more general case that allows partial reflections and transmissions. It applies to the case when the wave length of the incident wave is much shorter than the width of the interface as both lengths go to zero. The more general case of partial transmissions and reflections is a topic of a forthcoming paper [26].

The main ingredient in the *well-balanced* kinetic scheme by Perthame and Semioni [33] for the shallow water equations with topography was to build in the Hamiltonian-preserving mechanism into the numerical flux in order to preserve the steady state solution of the shallow water equations when the water velocity is zero. This is achieved using the fact that the density distribution f remains unchanged along the characteristic, thus

$$f(t, x^+, \xi^+) = f(t, x^-, \xi^-) \tag{2.8}$$

at a discontinuous point x of $c(x)$, where for example, ξ^+ is defined through the constant Hamiltonian condition (2.4).

In this paper, we use this mechanism for the numerical approximation to the Liouville equation (1.1) with a discontinuous wave speed. This approximation, by its design, maintains a constant Hamiltonian modulus the numerical approximation error across the interface. In [24] we introduced two Hamiltonian-preserving schemes for the Liouville equation arising from the semiclassical limit of the linear Schrödinger equation by incorporating this particle behavior into the numerical flux.

3. Scheme I: a finite difference approach

3.1. A Hamiltonian-preserving numerical flux

We now describe our first finite difference scheme (called *Scheme I*) for the Liouville equation with a discontinuous local wave speed.

Assume that the discontinuous points of wave speed c are located at the grid points. Let the left and right limits of $c(x)$ at point $x_{i+1/2}$ be $c^+_{i+\frac{1}{2}}$ and $c^-_{i+\frac{1}{2}}$, respectively. Note that if c is continuous at $x_{i+1/2}$, then $c^+_{i+\frac{1}{2}} = c^-_{i+\frac{1}{2}}$. We approximate c by a piecewise linear function

$$c(x) \approx c^+_{i-1/2} + \frac{c^-_{i+1/2} - c^+_{i-1/2}}{\Delta x}(x - x_{i-1/2}).$$

We also define the averaged wave speed as $c_i = \frac{c^+_{i-\frac{1}{2}} + c^-_{i+\frac{1}{2}}}{2}$. We will adopt the flux splitting technique used in [33]. The semi-discrete scheme (with continuous time) reads

$$(f_{ij})_t + \frac{c_i \operatorname{sign}(\xi_j)}{\Delta x}\left(f^-_{i+\frac{1}{2},j} - f^+_{i-\frac{1}{2},j}\right) - \frac{c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}}{\Delta x \Delta \xi}|\xi_j|\left(f_{i,j+\frac{1}{2}} - f_{i,j-\frac{1}{2}}\right) = 0, \tag{3.1}$$

where the numerical fluxes $f_{i,j+\frac{1}{2}}$ is defined using the upwind discretization. Since the characteristics of the Liouville equation maybe different on the two sides of the interface, the corresponding numerical fluxes should also be different. The essential part of our algorithm is to define the split numerical fluxes $f^-_{i+\frac{1}{2},j}, f^+_{i-\frac{1}{2},j}$ at each cell interface. We will use (2.8) to define these fluxes.

Assume c is discontinuous at $x_{i+1/2}$. Consider the case $\xi_j > 0$. Using upwind scheme, $f^-_{i+\frac{1}{2},j} = f_{ij}$. However,

$$f^+_{i+1/2,j} = f(x^+_{i+1/2}, \xi^+_j) = f(x^-_{i+1/2}, \xi^-_j)$$

while ξ^-_j is obtained from $\xi^+_j = \xi_j$ from (2.4). Since ξ^-_j may not be a grid point, we have to define it approximately. The first approach is to locate the two cell centers that bound ξ^-_j, then use a linear interpolation to evaluate the needed numerical flux at ξ^-. The case of $\xi_j < 0$ is treated similarly. The detailed algorithm to generate the numerical flux is given below.

Algorithm I

- if $\xi_j > 0$

  $f^-_{i+\frac{1}{2},j} = f_{ij}$,

  $\xi^- = \frac{c^+_{i+\frac{1}{2}}}{c^-_{i+\frac{1}{2}}}\xi_j$

  if $\xi_k \leqslant \xi^- < \xi_{k+1}$ for some k

  then $f^+_{i+\frac{1}{2},j} = \frac{\xi_{k+1}-\xi^-}{\Delta\xi} f_{i,k} + \frac{\xi^- - \xi_k}{\Delta\xi} f_{i,k+1}$,
- if $\xi_j < 0$

  $f^+_{i+\frac{1}{2},j} = f_{i+1,j}$,

  $\xi^+ = \frac{c^-_{i+\frac{1}{2}}}{c^+_{i+\frac{1}{2}}}\xi_j$

  if $\xi_k \leqslant \xi^+ < \xi_{k+1}$ for some k

  then $f^-_{i+\frac{1}{2},j} = \frac{\xi_{k+1}-\xi^+}{\Delta\xi} f_{i+1,k} + \frac{\xi^+ - \xi_k}{\Delta\xi} f_{i+1,k+1}$.

The above algorithm for evaluating numerical fluxes is of first order. One can obtain a second order flux by incorporating the slope limiter, such as van Leer or minmod slope limiter [29], into the above algorithm. This can be achieved by replacing f_{ik} with $f_{ik} + \frac{\Delta x}{2}s_{ik}$, and replacing $f_{i+1,k}$ with $f_{i+1,k} - \frac{\Delta x}{2}s_{i+1,k}$ in the above algorithm for all the possible index k, where s_{ik} is the slope limiter in the x-direction.

After the spatial discretization is specified, one can use any time discretization for the time derivative.

3.2. Positivity and l^∞ contraction

Since the exact solution of the Liouville equation is positive when the initial profile is, it is important that the numerical solution inherits this property.

We only consider the scheme using the first order numerical flux, and the forward Euler method in time. Without loss of generality, we consider the case $\xi_j > 0$ and $c^-_{i+\frac{1}{2}} < c^+_{i-\frac{1}{2}}$ for all i (the other cases can be treated similarly with the same conclusion). The scheme reads

$$\frac{f^{n+1}_{ij} - f^n_{ij}}{\Delta t} + c_i \frac{f_{ij} - (d_1 f_{i-1,k} + d_2 f_{i-1,k+1})}{\Delta x} - \frac{c^-_{i+\frac{1}{2}} - c_{i-\frac{1}{2}+}}{\Delta x}\xi_j \frac{f_{ij} - f_{i,j-1}}{\Delta \xi} = 0,$$

where d_1, d_2 are non-negative and $d_1 + d_2 = 1$. We omit the superscript n of f. The above scheme can be rewritten as

$$f_{ij}^{n+1} = \left(1 - c_i\lambda_x^t - \frac{\left|c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}\right|}{\Delta x}|\xi_j|\lambda_\xi^t\right) f_{ij} + c_i\lambda_x^t(d_1 f_{i-1,k} + d_2 f_{i-1,k+1}) + \frac{\left|c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}\right|}{\Delta x}|\xi_j|\lambda_\xi^t f_{i,j-1}. \tag{3.2}$$

Now we investigate the positivity of scheme (3.2). This is to prove that if $f_{ij}^n \geqslant 0$ for all (i,j), then this is also true for f^{n+1}. Clearly, one just needs to show that all coefficients for f^n are non-negative. A sufficient condition for this is clearly

$$1 - c_i\lambda_x^t - \frac{\left|c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}\right|}{\Delta x}|\xi_j|\lambda_\xi^t \geqslant 0,$$

or

$$\Delta t \max_{i,j}\left[\frac{c_i}{\Delta x} + \frac{\frac{\left|c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}\right|}{\Delta x}}{|\xi_j|}\Delta\xi\right] \leqslant 1. \tag{3.3}$$

This CFL condition is similar to the CFL condition (2.3) of the usual finite difference scheme *except* that the quantity $\frac{|c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}|}{\Delta x}$ now represents the wave speed gradient at its *smooth* point, which has a *finite* upper bound. Thus our scheme allows a time step $\Delta t = O(\Delta x, \Delta\xi)$, a significant improvement over a standard discretization.

According to the study in [32], our second order scheme, which incorporates slope limiter into the first order scheme, is positive under the half CFL condition, namely, the constant on the right-hand side of (3.3) is 1/2.

The above conclusion is analyzed based on forward Euler time discretization. One can draw the same conclusion for the second order TVD Runge–Kutta time discretization [40].

The l^∞-contracting property of this scheme follows easily, because the coefficients in (3.2) are positive and the sum of them is 1.

3.3. The l^1-stability of Scheme I

In this section, we prove the l^1-stability of Scheme I (with the first order numerical flux and the forward Euler method in time). The proof is similar to that in [25] with difference in details due to different particle behaviors at the interface.

For simplicity, we consider the case when the wave speed has only one discontinuity at grid point $x_{m+\frac{1}{2}}$ with $c^-_{m+\frac{1}{2}} > c^+_{m+\frac{1}{2}}$, and $c'(x) > 0$ at smooth points. The other cases, namely, when $c'(x) \leqslant 0$, or when the wave speed has several discontinuous points with increased or decreased jumps, can be discussed similarly. Denote $\lambda_c \equiv c^+_{m+\frac{1}{2}}/c^-_{m+\frac{1}{2}} < 1$.

We consider the general case that $\xi_1 < 0$, $\xi_M > 0$. For this case, the study in [22] suggests that the computational domain should exclude a set $O_\xi = \{(x,\xi) \in \mathbb{R}^2 | \xi = 0\}$ which causes singularity in the velocity field. For example, we can exclude the following index set

$$D_o = \left\{(i,j)\middle| |\xi_j| < \frac{\Delta\xi}{2}\right\},$$

from the computational domain.

Since $c(x)$ has a discontinuity, we also define an index set

$$D_l^4 = \{(i,j) | x_i \leqslant x_m, \xi_j < \lambda_c\xi_1\}.$$

Due to the slowness change across the wave speed jump at $x_{m+\frac{1}{2}}$, D_l^4 represents the area where waves come from outside of the domain $[x_1, x_N]\times[\xi_1, \xi_M]$. In order to implement our scheme conveniently, this index set is also excluded from the computational domain. Thus the computational domain is chosen as

$$E_d = \{(i,j)|i = 1,\ldots,N, j = 1,\ldots,M\} \setminus \{D_o \cup D_l^4\}. \tag{3.4}$$

A sketch of E_d and D_l^4 is shown in Fig. 1 in Section 4.2.

As a result of excluding the index set D_o from the computational domain, the computational domain is split into two independent parts

$$E_d = \{(i,j) \in E_d|\xi_j > 0\} \cup \{(i,j) \in E_d|\xi_j < 0\} \equiv E_d^+ \cup E_d^-.$$

The l^1-stability study of Scheme I can be carried out in these two domains, respectively. In the following we prove the l^1-stability of Scheme I in the domain E_d^-. The study in the domain E_d^+ can be made similarly.

We define the l^1-norm of a numerical solution $u_{i\,j}$ in the set E_d^- to be

$$|f|_1 = \frac{1}{N_d^-} \sum_{(i,j)\in E_d^-} |f_{ij}|$$

with N_d^- being the number of elements in E_d^-. Given the initial data $f_{ij}^0, (i,j) \in E_d^-$. Denote the numerical solution at time T to be f_{ij}^L, $(i,j) \in E_d^-$. To prove the l^1-stability, we need to show that $|f^L|_1 \leqslant C\,|f^0|_1$.

Due to the linearity of the scheme, the equation for the error between the analytical and the numerical solution is the same as (3.2), so in this section, f_{ij} will denote the error. We assume there is no error at the boundary, thus $f_{ij}^n = 0$ at the boundary. If the l^1-norm of the error introduced at each time step in incoming boundary cells is ensured to be o(1) part of $|f^n|_1$, our following analysis still applies.

Now denote

$$A_i = \frac{1}{\Delta x}\left|c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}}\right|. \tag{3.5}$$

Assume an upper bound for the wave speed slope is A_u, $A_i < A_u\ \forall i$. These notations will be used below as well as in the stability proof of Scheme II. One also has $\frac{1}{\Delta x}|c_i - c_{i-1}| < A_u\ \forall i$. Assume the wave speed has a lower bound C_m, $c_i > C_m > 0\ \forall i$.

When $\xi_j < 0$, Scheme I is given by

(1) if $i = m$,

$$f_{ij}^{n+1} = (1 - A_i|\xi_j|\lambda_\xi^t - c_i\lambda_x^t)f_{ij} + A_i|\xi_j|\lambda_\xi^t f_{i,j+1} + c_i\lambda_x^t f_{i+1,j}, \tag{3.6}$$

(2) $$f_{mj}^{n+1} = (1 - A_m|\xi_j|\lambda_\xi^t - c_m\lambda_x^t)f_{mj} + A_m|\xi_j|\lambda_\xi^t f_{m,j+1} + c_m\lambda_x^t(d_{jk}f_{m+1,k} + d_{j,k+1}f_{m+1,k+1}), \tag{3.7}$$

where $0 \leqslant d_{jk} \leqslant 1$ and $d_{jk} + d_{j,k+1} = 1$. In (3.7) k is determined by $\xi_k \leqslant \frac{\xi_j}{\lambda_c} < \xi_{k+1}$.

When summing up all absolute values of f_{ij}^{n+1} in (3.6) and (3.7), one typically gets the following inequality:

$$|f^{n+1}|_1 \leqslant \frac{1}{N_d^-} \sum_{(i,j)\in E_d^-} \alpha_{ij}|f_{ij}^n|, \tag{3.8}$$

where the coefficients $\alpha_{i\,j}$ are positive. One can check that, under the CFL condition (3.3), $\alpha_{i\,j} \leqslant 1 + 2A_u\Delta t$ except for possibly $(i,j) \in D_{m+1}^-$ defined as

$$D_{m+1}^- = \{(i,j) \in E_d^-|i = m+1\}.$$

We next derive the bound for M^- defined as

$$M^- = \max_{(m+1,j)\in D_{m+1}^-} \alpha_{m+1,j}.$$

Define the set

$$S_j^{m+1} = \left\{j'|\xi_{j'} < 0, \left|\frac{\xi_{j'}}{\lambda_c} - \xi_j\right| < \Delta\xi\right\} \quad \text{for} \quad (m+1,j) \in D_{m+1}^-.$$

Let the number of elements in S_j^{m+1} be N_j^{m+1}. One can check that $N_j^{m+1} \leqslant 2\lambda_c + 1$ because every two elements $j_1', j_2' \in S_j^m$ satisfy $|\frac{\xi_{j_1'}}{\lambda_c} - \frac{\xi_{j_2'}}{\lambda_c}| \geqslant \frac{\Delta\xi}{\lambda_c}$.

On the other hand, one can easily check from (3.6) and (3.7), for $(m+1, j) \in D_{m+1}^-$,

$$\alpha_{m+1,j} < 1 - c_{m+1}\lambda_x^t + c_m\lambda_x^t(2\lambda_c + 1) = 1 + (c_m + c_{m+1})\lambda_x^t + \mathrm{O}(\Delta x),$$

so for sufficiently small Δx, M^- can be bounded by

$$M^- < 1 + 2(c_m + c_{m+1})\lambda_x^t.$$

Denote $M' = 2(c_m + c_{m+1})\lambda_x^t$. From (3.8),

$$|f^{n+1}|_1 < (1 + 2A_u\Delta t)|f^n|_1 + \frac{M'}{N_d^-}\sum_{(m+1,j)\in D_{m+1}^-}|f_{m+1,j}^n|. \tag{3.9}$$

We now establish the following theorem:

Theorem 3.1. *Under the CFL condition (3.3), the scheme (3.6), (3.7) is l^1-stable*

$$|f^L|_1 < C|f^0|_1.$$

Proof. From (3.9),

$$|f^L|_1 < (1 + 2A_u\Delta t)^L\left\{|f^0|_1 + \frac{M'}{N_d^-}\sum_{n=0}^{L-1}\left[\sum_{(m+1,j)\in D_{m+1}^-}|f_{m+1,j}^n|\right]\right\}. \tag{3.10}$$

It remains to estimate

$$S = \sum_{n=0}^{L-1}\left\{\sum_{(m+1,j)\in D_{m+1}^-}|f_{m+1,j}^n|\right\}. \tag{3.11}$$

Define the set

$$S_r = \{(i,j)|x_i > x_{m+\frac{1}{2}}, (m+1,j) \in D_{m+1}^-\}$$

$\forall (i,j) \in S_r$, due to the zero boundary condition and the upwind nature of the scheme, one has

$$f_{ij}^n = \sum_{(p,q)\in S_r, p\geqslant i}\beta_{pq}^{ijn0}f_{pq}^0, \quad (i,j) \in S_r \tag{3.12}$$

with $\beta_{pq}^{ijn0} \geqslant 0$.

Notice $D_{m+1}^- \subset S_r$,

$$S \leqslant \sum_{(p,q)\in S_r}\left(\sum_{n=0}^{L-1}\sum_{(m+1,j)\in D_{m+1}^-}\beta_{pq}^{m+1,jn0}\right)|f_{pq}^0| \equiv \sum_{(p,q)\in S_r}F(p,q)|f_{pq}^0|, \tag{3.13}$$

where we have defined

$$F(p,q) = \sum_{n=0}^{L-1}\sum_{(m+1,j)\in D_{m+1}^-}\beta_{pq}^{m+1,jn0}, \quad (p,q) \in S_r. \tag{3.14}$$

The next step is to estimate these coefficients. Define

$$\beta_{pq}^{ij0} = \sum_{n=0}^{\infty}\beta_{pq}^{ijn0}, \quad (i,j),(p,q) \in S_r, p \geqslant i,$$

then (3.14) gives

$$F(p,q) = \sum_{(m+1,j)\in D_{m+1}^-}\sum_{n=0}^{L-1}\beta_{pq}^{m+1,jn0} \leqslant \sum_{(m+1,j)\in D_{m+1}^-}\beta_{pq}^{m+1,j0}, \quad (p,q) \in S_r.$$

We first evaluate $\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ij0}_{pq}$ when $i=p$. Denote $c^{ij}_1=1-A_i|\xi_j|\lambda^t_\xi-c_i\lambda^t_x$, $c^{ij}_2=A_i|\xi_j|\lambda^t_\xi$, $c^i_3=c_i\lambda^t_x$. Under the CFL condition (3.3), c^{ij}_1, c^{ij}_2, c^i_3 are positive. One also has $c^{ij}_1+c^{ij}_2\leqslant 1-C_m\lambda^t_x$. From scheme (3.6), it can be directly computed

$$\sum_{(m+1,j)\in D^-_{m+1}}\beta^{pj0}_{pq}<\sum_{n=0}^{\infty}(1-C_m\lambda^t_x)^n=\frac{1}{C_m\lambda^t_x}. \tag{3.15}$$

We now study the relation between $\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ij0}_{pq}$ and $\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,j0}_{pq}$ when $i<p$. From scheme (3.6),

$$\beta^{ij,n+1,0}_{pq}=c^{ij}_1\beta^{ijn0}_{pq}+c^{ij}_2\beta^{i,j+1,n0}_{pq}+c^i_3\beta^{i+1,jn0}_{pq}. \tag{3.16}$$

Summing up j in (3.16) gives

$$\begin{aligned}\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ij,n+1,0}_{pq}&=\sum_{(m+1,j)\in D^-_{m+1}}(c^{ij}_1+c^{i,j-1}_2)\beta^{ijn0}_{pq}+c^i_3\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,jn0}_{pq}\\&<(1-c^i_3+A_u\lambda^t_\xi\Delta\xi)\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ijn0}_{pq}+c^i_3\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,jn0}_{pq},\end{aligned} \tag{3.17}$$

then a sum for n from 0 to ∞ in (3.17) gives

$$(c^i_3-A_u\lambda^t_\xi\Delta\xi)\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ij0}_{pq}<c^i_3\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,j0}_{pq},$$

so

$$\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ij0}_{pq}<\frac{c^i_3}{c^i_3-A_u\lambda^t_\xi\Delta\xi}\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,j0}_{pq}<\left(1+\frac{A_u}{C_m}\Delta x+o(\Delta x)\right)\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,j0}_{pq}.$$

Thus for sufficiently small Δx, one has

$$\sum_{(m+1,j)\in D^-_{m+1}}\beta^{ij0}_{pq}<\left(1+\frac{2A_u}{C_m}\Delta x\right)\sum_{(m+1,j)\in D^-_{m+1}}\beta^{i+1,j0}_{pq},\quad i<p. \tag{3.18}$$

We now can evaluate $F(p,q)$ for $(p,q)\in S_r$. From the definition of S_r, when $(p,q)\in S_r$, one has $p\geqslant m+1$.

$$\begin{aligned}F(p,q)&\leqslant\sum_{(m+1,j)\in D^-_{m+1}}\beta^{m+1j0}_{pq}<\left(1+\frac{2A_u}{C_m}\Delta x\right)\sum_{(m+1,j)\in D^-_{m+1}}\beta^{m+2j0}_{pq}<\cdots<\left(1+\frac{2A_u}{C_m}\Delta x\right)^{p-m-1}\sum_{(m+1,j)\in D^-_{m+1}}\beta^{pj0}_{pq}\\&<\exp\left(\frac{2A_u}{C_m}(x_N-x_1)\right)\sum_{(m+1,j)\in D^-_{m+1}}\beta^{pj0}_{pq}<\exp\left(\frac{2A_u}{C_m}(x_N-x_1)\right)\frac{1}{C_m\lambda^t_x}\equiv C_T.\end{aligned} \tag{3.19}$$

Therefore, from (3.13) one gets

$$S\leqslant\sum_{(p,q)\in S_r}F(p,q)|f^0_{pq}|<C_T\sum_{(p,q)\in S_r}|f^0_{pq}|\leqslant C_T\sum_{(p,q)\in E^-_d}|f^0_{pq}|=C_TN^-_d|f^0|_1. \tag{3.20}$$

Combing (3.10) and (3.20),

$$|f^L|_1<(1+2A_u\Delta t)^L\{|f^0|_1+C_TM'|f^0|_1\}<\exp(2A_uT)[1+C_TM']|f^0|_1\equiv C|f^0|_1,$$

where $C\equiv\exp(2A_{u\ T})[1+C'_{TM}]$. Thus Theorem 3.1 is proved. □

One can prove the similar conclusion for index set E^+_d.

Remark 3.1. Theorem 3.1 holds for *any* l^1 initial data. The corresponding result in the case of semiclassical limit of Schrödinger equation [25] excludes the case of measure-valued initial data.

4. Scheme II: a finite volume approach

4.1. A Hamiltonian-preserving numerical flux

In this section, we derive another flux based on the finite volume approach which results in an l^1-contracting scheme. We call this scheme as *Scheme II*.

Assuming the mesh grid is such that $\xi_{j-\frac{1}{2}}$ and $\xi_{j+\frac{1}{2}}$ do not have opposite sign. By integrating the Liouville equation (2.1) over the cell $[x_{i-1/2},\ x_{i+1/2}]\times[\xi_{j-1/2},\xi_{j+1/2}]$, one gets the following semi-discrete flux splitting scheme:

$$(f_{ij})_t+\frac{\operatorname{sign}(\xi_j)}{\Delta x}\left(c^-_{i+\frac{1}{2}}f^-_{i+\frac{1}{2},j}-c^+_{i-\frac{1}{2}}f^+_{i-\frac{1}{2},j}\right)-\frac{c^-_{i+\frac{1}{2}}-c_{i-\frac{1}{2}^+}}{\Delta x\,\Delta\xi}\left(|\xi_{j+\frac{1}{2}}|f_{i,j+\frac{1}{2}}-|\xi_{j-\frac{1}{2}}|f_{i,j-\frac{1}{2}}\right)=0. \tag{4.1}$$

In the finite volume approach, the numerical fluxes are defined as integrals of solution along the cell interface which depend on the sign of ξ_j and $\frac{c^-_{i+\frac{1}{2}}-c^+_{i-\frac{1}{2}}}{\Delta x}$. To illustrate the basic idea, we assume $\xi_j>0$, $\frac{c^-_{i+\frac{1}{2}}-c^+_{i-\frac{1}{2}}}{\Delta x}<0$. In this case

$$f^-_{i+\frac{1}{2},j}=\frac{1}{\Delta\xi}\int_{\xi_{j-\frac{1}{2}}}^{\xi_{j+\frac{1}{2}}}f\left(x^-_{i+\frac{1}{2}},\xi,t\right)\mathrm{d}\xi,$$

$$f_{i,j+\frac{1}{2}}=\frac{1}{c^-_{i+\frac{1}{2}}-c^+_{i-\frac{1}{2}}}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}c_xf\left(x,\xi^-_{j+\frac{1}{2}},t\right)\mathrm{d}x.$$

Note that $f(x,\xi,t)$ may be discontinuous at the grid point $x=x_{i+\frac{1}{2}}$ and $\xi=\xi_{j+\frac{1}{2}}$.

By using condition (2.8):

$$f^+_{i+\frac{1}{2},j}=\frac{1}{\Delta\xi}\int_{\xi_{j-\frac{1}{2}}}^{\xi_{j+\frac{1}{2}}}f\left(x^+_{i+\frac{1}{2}},\xi,t\right)\mathrm{d}\xi=\frac{1}{\Delta\xi}\int_{\xi_{j-\frac{1}{2}}}^{\xi_{j+\frac{1}{2}}}\overline{f}\left(x^-_{i+\frac{1}{2}},\xi,t\right)\mathrm{d}\xi, \tag{4.2}$$

where $\overline{f}$ is defined as

$$\overline{f}\left(x^-_{i+\frac{1}{2}},\xi,t\right)=f\left(x^-_{i+\frac{1}{2}},\frac{c^+_{i+\frac{1}{2}}}{c^-_{i+\frac{1}{2}}}\xi,t\right).$$

Using change of variable on (4.2) leads to

$$f^+_{i+\frac{1}{2},j}=\frac{1}{\Delta\xi}\int_{\xi_{j-\frac{1}{2}}}^{\xi_{j+\frac{1}{2}}}f\left(x^-_{i+\frac{1}{2}},\frac{c^+_{i+\frac{1}{2}}\xi}{c^-_{i+\frac{1}{2}}},t\right)\mathrm{d}\xi=\frac{c^-_{i+\frac{1}{2}}}{c^+_{i+\frac{1}{2}}}\frac{1}{\Delta\xi}\int_{c^+_{i+\frac{1}{2}}\xi_{j-\frac{1}{2}}/c^-_{i+\frac{1}{2}}}^{c^+_{i+\frac{1}{2}}\xi_{j+\frac{1}{2}}/c^-_{i+\frac{1}{2}}}f\left(x^-_{i+\frac{1}{2}},\xi,t\right)\mathrm{d}\xi. \tag{4.3}$$

The integral in (4.3) will be approximated by a quadrature rule. Since the end point $c^+_{i+\frac{1}{2}}\xi_{j+\frac{1}{2}}/c^-_{i+\frac{1}{2}}$ in (4.3) may not be a grid point in the ξ-direction, special care needs to be taken at both ends of the interval

$$\left[c^+_{i+\frac{1}{2}}\xi_{j-\frac{1}{2}}\Big/c^-_{i+\frac{1}{2}},c^+_{i+\frac{1}{2}}\xi_{j+\frac{1}{2}}\Big/c^-_{i+\frac{1}{2}}\right]. \tag{4.4}$$

We propose the following evaluation of the split fluxes $f^{\pm}_{i+\frac{1}{2},j}$ in (4.1).

Algorithm II

- if $\xi_j>0$
  $f^-_{i+\frac{1}{2},j}=f_{ij}$,
  $\xi^-_1=\frac{c^+_{i+\frac{1}{2}}}{c^-_{i+\frac{1}{2}}}\xi_{j-\frac{1}{2}}$, $\quad\xi^-_2=\frac{c^+_{i+\frac{1}{2}}}{c^-_{i+\frac{1}{2}}}\xi_{j+\frac{1}{2}}$
  - if $\xi_{k-\frac{1}{2}}\leqslant\xi^-_1<\xi^-_2\leqslant\xi_{k+\frac{1}{2}}$ for some k
    $f^+_{i+\frac{1}{2},j}=f_{ik}$
  - else $\xi_{k-\frac{1}{2}}\leqslant\xi^-_1<\xi_{k+\frac{1}{2}}<\cdots<\xi_{k+s-\frac{1}{2}}<\xi^-_2\leqslant\xi_{k+s+\frac{1}{2}}$ for some k, s
    $f^+_{i+\frac{1}{2},j}=\frac{c^-_{i+\frac{1}{2}}}{c^+_{i+\frac{1}{2}}}\left\{\frac{\xi_{k+\frac{1}{2}}-\xi^-_1}{\Delta\xi}f_{ik}+f_{i,k+1}+\cdots+f_{i,k+s-1}+\frac{\xi^-_2-\xi_{k+s-\frac{1}{2}}}{\Delta\xi}f_{i,k+s}\right\}$
  - end

- if $\xi_j < 0$

  $f^+_{i+\frac{1}{2},j} = f_{i+1,j}$,

  $\xi^+_1 = \frac{c^-_{i+\frac{1}{2}}}{c^+_{i+\frac{1}{2}}}\xi_{j-\frac{1}{2}}, \quad \xi^+_2 = \frac{c^-_{i+\frac{1}{2}}}{c^+_{i+\frac{1}{2}}}\xi_{j+\frac{1}{2}}$

  – if $\xi_{k-\frac{1}{2}} \leqslant \xi^+_1 < \xi^+_2 \leqslant \xi_{k+\frac{1}{2}}$ for some k

  $f^-_{i+\frac{1}{2},j} = f_{i+1,k}$

  – else $\xi_{k-\frac{1}{2}} \leqslant \xi^+_1 < \xi_{k+\frac{1}{2}} < \cdots < \xi_{k+s-\frac{1}{2}} < \xi^+_2 \leqslant \xi_{k+s+\frac{1}{2}}$ for some k, s

  $$f^-_{i+\frac{1}{2},j} = \frac{c^+_{i+\frac{1}{2}}}{c^-_{i+\frac{1}{2}}}\left\{\frac{\xi_{k+\frac{1}{2}}-\xi^+_1}{\Delta\xi}f_{i+1,k} + f_{i+1,k+1} + \cdots + f_{i+1,k+s-1} + \frac{\xi^+_2-\xi_{k+s-\frac{1}{2}}}{\Delta\xi}f_{i+1,k+s}\right\}$$

  – end

- end

Remark 4.1. The above algorithm uses a first order quadrature rule at the ends of the interval (4.4), thus it is of first order even if the slope limiters in x-direction are incorporated into the algorithm. One can also use a second order quadrature rule at the ends of intervals (4.4). But the resulting second order scheme is no longer l^1-contracting, which is the property of Scheme II, as will be proved in the next subsection. One can still prove that this scheme is l^1-stable, similar to the property of Scheme I. Compared with Scheme I, this scheme is second order accurate and l^1-stable, but more complex to implement. We will not present the detail of this numerical scheme in this paper.

4.2. The l^1-contraction, l^∞-stabilities and positivity of Scheme II

In this subsection, we study the l^1 and l^∞ stability of Scheme II. Its positivity is obvious under the CFL condition (3.3).

Theorem 4.1. *If the forward Euler time discretization is used, then the flux given by Algorithm I yields the scheme (4.1) which is l^1-contracting and l^∞-stable.*

Proof. In this proof we only discuss the case when the wave speed has one discontinuity at grid point $x_{m+\frac{1}{2}}$ with $c^-_{m+\frac{1}{2}} > c^+_{m+\frac{1}{2}}$, and $c'(x) > 0$ at smooth points. The other situations can be discussed similarly.

We consider the general case that $\xi_1 < 0$, $\xi_M > 0$. We assume the mesh is such that 0 is a grid point in ξ-direction. In this case, the index set

$$D_o = \left\{(i,j)\middle||\xi_j| < \frac{\Delta\xi}{2}\right\}$$

that needs to be excluded from the computational domain is null. As such, the cell interface $\{(x,\xi)|\xi = 0\}$ is actually the computational domain boundary where appropriate boundary conditions should be imposed [22]. As discussed in Section 3.3, the computational domain is chosen as

$$E_d = \{(i,j)|i = 1,\ldots,N, j = 1,\ldots,M\} \setminus D^4_l,$$

where

$$D^4_l = \left\{(i,j)\middle|x_i \leqslant x_m, \xi_{j-\frac{1}{2}} < \frac{c^+_{m+\frac{1}{2}}}{c^-_{m+\frac{1}{2}}}\xi_{\frac{1}{2}}\right\}.$$

Define some subsets of E_d

$$D^+_m = \left\{(m,j)|\xi_j \geqslant \frac{\Delta\xi}{2}\right\},$$

$$D^+_{m+1} = \left\{(m+1,j)\middle|\xi_j \geqslant \frac{\Delta\xi}{2}\right\},$$

$$D_m^- = \left\{ (m,j) \left| \frac{c^+_{m+\frac{1}{2}} \xi_{\frac{1}{2}}}{c_{m+\frac{1}{2}}} \leqslant \xi_{j-\frac{1}{2}} \leqslant -\Delta\xi \right. \right\},$$

$$D_{m+1}^- = \left\{ (m+1,j) \left| \xi_j \leqslant -\frac{\Delta\xi}{2} \right. \right\}.$$

These domains are shown in Fig. 1.

Recall the definition of A_i in (3.5). Our scheme (4.1) with Algorithm II can be made precise as

(1) if $\xi_j > 0$, $i \neq m+1$,

$$f_{ij}^{n+1} = \left(1 - A_i \xi_{j-\frac{1}{2}} \lambda^t_\xi - c^-_{i+\frac{1}{2}} \lambda^t_x\right) f_{ij} + A_i \xi_{j+\frac{1}{2}} \lambda^t_\xi f_{i,j+1} + c^+_{i-\frac{1}{2}} \lambda^t_x f_{i-1,j}, \tag{4.5}$$

(2) if $\xi_j < 0$, $i \neq m$,

$$f_{ij}^{n+1} = \left(1 - A_i |\xi_{j-\frac{1}{2}}| \lambda^t_\xi - c^+_{i-\frac{1}{2}} \lambda^t_x\right) f_{ij} + A_i |\xi_{j+\frac{1}{2}}| \lambda^t_\xi f_{i,j+1} + c^-_{i+\frac{1}{2}} \lambda^t_x f_{i+1,j}, \tag{4.6}$$

(3) if $\xi_j > 0$,

$$f_{m+1,j}^{n+1} = \left(1 - A_{m+1} \xi_{j-\frac{1}{2}} \lambda^t_\xi - c^-_{m+\frac{3}{2}} \lambda^t_x\right) f_{m+1,j} + A_{m+1} \xi_{j+\frac{1}{2}} \lambda^t_\xi f_{m+1,j+1} + c^+_{m+\frac{1}{2}} \lambda^t_x f^+_{m+\frac{1}{2},j}, \tag{4.7}$$

(4) if $\xi_j < 0$,

$$f_{mj}^{n+1} = \left(1 - A_m |\xi_{j-\frac{1}{2}}| \lambda^t_\xi - c^+_{m-\frac{1}{2}} \lambda^t_x\right) f_{mj} + A_m |\xi_{j+\frac{1}{2}}| \lambda^t_\xi f_{m,j+1} + c^-_{m+\frac{1}{2}} \lambda^t_x f^-_{m+\frac{1}{2},j}, \tag{4.8}$$

where we omit the superscript n on the right-hand side.

By summing up (4.5)–(4.8) for $(i,j) \in E_d$, one typically gets the following expression:

$$\sum_{(i,j)\in E_d} |f_{ij}^{n+1}| \leqslant \sum_{(i,j)\in E_d} \alpha_{ij} |f_{ij}| + \sum_{(m+1,j)\in D^+_{m+1}} c^+_{m+\frac{1}{2}} \lambda^t_x |f^+_{m+\frac{1}{2},j}| + \sum_{(m,j)\in D^-_m} c^-_{m+\frac{1}{2}} \lambda^t_x |f^-_{m+\frac{1}{2},j}| \equiv I_1 + I_2 + I_3. \tag{4.9}$$

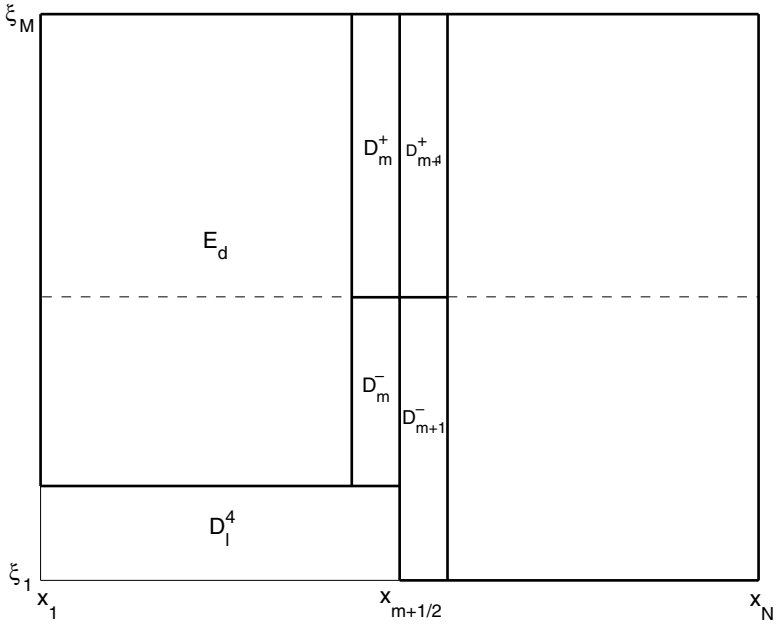

Fig. 1. Sketch of the index sets D_m^+, D_{m+1}^+, D_m^-, D_{m+1}^-, D_l^4.

As in the proof of stability of Scheme I, we assume that f satisfies the zero boundary condition. In this situation, the coefficients α_{ij} in (4.9) satisfy

$$\alpha_{ij} \leqslant 1, \quad (i,j) \in E_d \setminus \{D_m^+ \cup D_{m+1}^-\}, \tag{4.10}$$

$$\alpha_{ij} \leqslant 1 - c^-_{m+\frac{1}{2}}\lambda_x^t, \quad (i,j) \in D_m^+, \tag{4.11}$$

$$\alpha_{ij} \leqslant 1 - c^+_{m+\frac{1}{2}}\lambda_x^t, \quad (i,j) \in D_{m+1}^-. \tag{4.12}$$

We now study the relation between I_2 and $\sum_{(m,j)\in D_m^+}|c^-_{m+\frac{1}{2}}\lambda_x^t f_{mj}|$. Let

$$p_{M+1} = \frac{c^+_{m+\frac{1}{2}}\xi_{M+\frac{1}{2}}}{c^-_{m+\frac{1}{2}}},$$

and assume

$$\xi_{k-\frac{1}{2}} < p_{M+1} \leqslant \xi_{k+\frac{1}{2}} \leqslant \xi_{M+\frac{1}{2}}.$$

Assume $\xi_{J_2-\frac{1}{2}} = 0$ for some J_2, since

$$\frac{1}{\lambda_x^t c^-_{m+\frac{1}{2}}} I_2 \leqslant \sum_{j=J_2}^{k-1}|f_{mj}| + \frac{p_{N+1}-\xi_k}{\Delta\xi}|f_{mk}| \leqslant \sum_{(m,j)\in D_m^+}|f_{mj}|,$$

thus

$$I_2 \leqslant \sum_{(m,j)\in D_m^+}\left|c^-_{m+\frac{1}{2}}\lambda_x^t f_{mj}\right|. \tag{4.13}$$

Similarly, one gets

$$I_3 \leqslant \sum_{(m+1,j)\in D_{m+1}^-}\left|c^+_{m+\frac{1}{2}}\lambda_x^t f_{m+1,j}\right|. \tag{4.14}$$

Combining (4.9)–(4.14) gives

$$\sum_{(i,j)\in E_d}|f_{ij}^{n+1}| \leqslant \sum_{(i,j)\in E_d}|f_{ij}^n|. \tag{4.15}$$

This is the l^1-contracting property of Scheme II.

Next we prove the l^∞-stability. Observing that the coefficients on the right-hand side of (4.5)–(4.8) are positive, it remains to estimate the *sum of these coefficients* (SC). In (4.5), the SC is

$$\mathrm{SC}_1 = 1 + (c^+_{i-\frac{1}{2}} - c^-_{i+\frac{1}{2}})\lambda_x^t + A_i\Delta\xi\lambda_\xi^t < 1 + 2A_u\Delta t. \tag{4.16}$$

In (4.6), the SC is

$$\mathrm{SC}_2 = 1 + (c^-_{i+\frac{1}{2}} - c^+_{i-\frac{1}{2}})\lambda_x^t - A_i\Delta\xi\lambda_\xi^t < 1 + 2A_u\Delta t. \tag{4.17}$$

Now we derive the SC in (4.8). Denote

$$\xi_1' = \frac{c^-_{m+\frac{1}{2}}}{c^+_{m+\frac{1}{2}}}\xi_{j-\frac{1}{2}}, \quad \xi_2' = \frac{c^-_{m+\frac{1}{2}}}{c^+_{m+\frac{1}{2}}}\xi_{j+\frac{1}{2}}. \tag{4.18}$$

The condition $c^+_{m+\frac{1}{2}} < c^-_{m+\frac{1}{2}}$ gives $\xi_2' - \xi_1' > \Delta\xi$. Therefore, it is impossible that $\xi_{k-\frac{1}{2}} \leqslant \xi_1' < \xi_2' \leqslant \xi_{k+\frac{1}{2}}$ for any k. Assume $\xi_{k-\frac{1}{2}} \leqslant \xi_1' < \xi_{k+\frac{1}{2}} < \cdots < \xi_{k+s-\frac{1}{2}} < \xi_2' \leqslant \xi_{k+s+\frac{1}{2}}$ with $s \geqslant 1$. In this case

$$f^-_{m+\frac{1}{2},j} = \frac{c^+_{m+\frac{1}{2}}}{c^-_{m+\frac{1}{2}}}\left\{\frac{\xi_{k+\frac{1}{2}}-\xi_1'}{\Delta\xi}f_{m+1,k} + f_{m+1,k+1} + \cdots + f_{m+1,k+s-1} + \frac{\xi_2'-\xi_{k+s-\frac{1}{2}}}{\Delta\xi}f_{m+1,k+s}\right\}. \tag{4.19}$$

Substituting (4.19) into (4.8) yields the evaluation

$$\begin{aligned}\mathrm{SC}_3 &= 1 - A_m\Delta\xi\lambda^t_\xi - c^+_{m-\frac{1}{2}}\lambda^t_x + c^-_{m+\frac{1}{2}}\lambda^t_x\left[\frac{c^+_{m+\frac{1}{2}}}{c^-_{m+\frac{1}{2}}}\left(\frac{\xi_{k+\frac{1}{2}}-\xi'_1}{\Delta\xi}+\frac{\xi_{k+\frac{3}{2}}-\xi_{k+\frac{1}{2}}}{\Delta\xi}+\cdots+\frac{\xi'_2-\xi_{k+s-\frac{1}{2}}}{\Delta\xi}\right)\right]\\ &= 1 - A_m\Delta\xi\lambda^t_\xi - c^+_{m-\frac{1}{2}}\lambda^t_x + c^-_{m+\frac{1}{2}}\lambda^t_x < 1 + 2A_u\Delta t. \end{aligned} \tag{4.20}$$

Now we consider case (4.7). Denote

$$\xi'_1 = \frac{c^+_{m+\frac{1}{2}}}{c^-_{m+\frac{1}{2}}}\xi_{j-\frac{1}{2}}, \quad \xi'_2 = \frac{c^+_{m+\frac{1}{2}}}{c^-_{m+\frac{1}{2}}}\xi_{j+\frac{1}{2}}. \tag{4.21}$$

In this case, we know $\xi'_2-\xi'_1<\Delta\xi$. So there are two cases $\xi_{k-\frac{1}{2}}\leqslant\xi'_1<\xi'_2\leqslant\xi_{k+\frac{1}{2}}$ or $\xi_{k-\frac{1}{2}}\leqslant\xi'_1<\xi_{k+\frac{1}{2}}<\xi'_2\leqslant\xi_{k+\frac{3}{2}}$ corresponding, respectively, to

$$f^+_{m+\frac{1}{2},j} = f_{mk} \tag{4.22}$$

or

$$f^+_{m+\frac{1}{2},j} = \frac{c^-_{m+\frac{1}{2}}}{c^+_{m+\frac{1}{2}}}\left\{\frac{\xi_{k+\frac{1}{2}}-\xi'_1}{\Delta\xi}f_{mk}+\frac{\xi'_2-\xi_{k+\frac{1}{2}}}{\Delta\xi}f_{m,k+1}\right\}. \tag{4.23}$$

Similar to the deduction of (4.20), one can check, for both cases, that

$$\mathrm{SC}_4 = 1 + A_{m+1}\Delta\xi\lambda^t_\xi - c^-_{m+\frac{3}{2}}\lambda^t_x + c^+_{m+\frac{1}{2}}\lambda^t_x < 1 + 2A_u\Delta t. \tag{4.24}$$

Combining (4.16), (4.17), (4.20) and (4.24), one gets

$$|f^{n+1}|_\infty < (1+2A_u\Delta t)|f^n|_\infty,$$

thus

$$|f^L|_\infty < (1+2A_u\Delta t)^L|f^0|_\infty < e^{2A_uT}|f^0|_\infty. \tag{4.25}$$

This is the l^∞-stability property of Scheme II. □

5. The schemes in two space dimension

Consider the Liouville equation in two space dimension:

$$f_t + \frac{c(x,y)\xi}{\sqrt{\xi^2+\eta^2}}f_x + \frac{c(x,y)\eta}{\sqrt{\xi^2+\eta^2}}f_y - c_x\sqrt{\xi^2+\eta^2}f_\xi - c_y\sqrt{\xi^2+\eta^2}f_\eta = 0. \tag{5.1}$$

We employ a uniform mesh with grid points at $x_{i+\frac{1}{2}}$, $y_{j+\frac{1}{2}}$, $\xi_{k+\frac{1}{2}}$, $\eta_{l+\frac{1}{2}}$ in each direction. The cells are centered at (x_i,y_j,ξ_k,η_l) with $x_i=\frac{1}{2}(x_{i+\frac{1}{2}}+x_{i-\frac{1}{2}})$, $y_j=\frac{1}{2}(y_{j+\frac{1}{2}}+y_{j-\frac{1}{2}})$, $\xi_k=\frac{1}{2}(\xi_{k+\frac{1}{2}}+\xi_{k-\frac{1}{2}})$, $\eta_l=\frac{1}{2}(\eta_{l+\frac{1}{2}}+\eta_{l-\frac{1}{2}})$. The mesh size is denoted by $\Delta x=x_{i+\frac{1}{2}}-x_{i-\frac{1}{2}}$, $\Delta y=y_{j+\frac{1}{2}}-y_{j-\frac{1}{2}}$, $\Delta\xi=\xi_{k+\frac{1}{2}}-\xi_{k-\frac{1}{2}}$, $\Delta\eta=\eta_{l+\frac{1}{2}}-\eta_{l-\frac{1}{2}}$. We define the cell average of f as

$$f_{ijkl} = \frac{1}{\Delta x\,\Delta y\,\Delta\xi\,\Delta\eta}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}}\int_{\xi_{k-\frac{1}{2}}}^{\xi_{k+\frac{1}{2}}}\int_{\eta_{l-\frac{1}{2}}}^{\eta_{l+\frac{1}{2}}} f(x,y,\xi,\eta,t)\,\mathrm{d}\eta\,\mathrm{d}\xi\,\mathrm{d}y\,\mathrm{d}x.$$

Similar to the 1D case, we approximate $c(x,y)$ by a piecewise bilinear function, and for convenience, we always provide two interface values of c at each cell interface. When c is smooth at a cell interface, the two interface values are identical. We also define the averaged wave speed in a cell by averaging the four cell interface values

$$c_{ij} = \frac{c^+_{i-\frac{1}{2},j}+c^-_{i+\frac{1}{2}}+c^+_{i,j-\frac{1}{2}}c^-_{i,j+\frac{1}{2}}}{4}.$$

The 2D Liouville equation (5.1) can be semi-discretized as

$$(f_{ijkl})_t + \frac{c_{ij}\xi_k}{\Delta x\sqrt{\xi_k^2+\eta_l^2}}\left(f^-_{i+\frac{1}{2},jkl} - f^+_{i-\frac{1}{2},jkl}\right) + \frac{c_{ij}\eta_l}{\Delta y\sqrt{\xi_k^2+\eta_l^2}}\left(f^-_{i,j+\frac{1}{2},kl} - f^+_{i,j-\frac{1}{2},kl}\right)$$
$$- \frac{c^-_{i+\frac{1}{2},j} - c^+_{i-\frac{1}{2},j}}{\Delta x\,\Delta\xi}\sqrt{\xi_k^2+\eta_l^2}\left(f_{ij,k+\frac{1}{2},l} - f_{ij,k-\frac{1}{2},l}\right) - \frac{c^-_{i,j+\frac{1}{2}} - c^+_{i,j-\frac{1}{2}}}{\Delta y\,\Delta\eta}\sqrt{\xi_k^2+\eta_l^2}\left(f_{ijk,l+\frac{1}{2}} - f_{ijk,l-\frac{1}{2}}\right) = 0,$$

where the numerical fluxes $f_{ij,k+\frac{1}{2},l}$, $f_{ijk,l+\frac{1}{2}}$ are provided by the upwind approximation, and the split fluxes values $f^-_{i+\frac{1}{2},jkl}$, $f^+_{i-\frac{1}{2},jkl}$, $f^-_{i,j+\frac{1}{2},kl}$, $f^+_{i,j-\frac{1}{2},kl}$ should be obtained using similar but slightly different algorithm for the 1D case, since the particle behavior at the interface is different in 2d from that in 1d, as described in Section 2. For example, to evaluate $f^\pm_{i+\frac{1}{2},jkl}$ we can extend Algorithm I as

Algorithm I in 2D

- if $\xi_k > 0$
  $f^-_{i+\frac{1}{2},jkl} = f_{ijkl}$,
  - if $\left(\frac{C^+_{i+\frac{1}{2},j}}{C^-_{i+\frac{1}{2},j}}\right)^2\xi_k^2 + \left[\left(\frac{C^+_{i+\frac{1}{2},j}}{C^-_{i+\frac{1}{2},j}}\right)^2 - 1\right]\eta_l^2 > 0$

    $\xi^- = \sqrt{\left(\frac{C^+_{i+\frac{1}{2},j}}{C^-_{i+\frac{1}{2},j}}\right)^2\xi_k^2 + \left[\left(\frac{C^+_{i+\frac{1}{2},j}}{C^-_{i+\frac{1}{2},j}}\right)^2 - 1\right]\eta_l^2}$

    if $\xi_{k'} \leqslant \xi^- < \xi_{k'+1}$ for some k'

    then $f^+_{i+\frac{1}{2},jkl} = \frac{\xi_{k'+1}-\xi^-}{\Delta\xi}f_{ij,k',l} + \frac{\xi^- - \xi_{k'}}{\Delta\xi}f_{ij,k'+1,l}$
  - else
    $f^+_{i+\frac{1}{2},jkl} = f_{i+1,j,k',l}$ where $\xi_{k'} = -\xi_k$
  - end

- if $\xi_k < 0$
  $f^+_{i+\frac{1}{2},jkl} = f_{i+1,jkl}$,
  - if $\left(\frac{C^-_{i+\frac{1}{2},j}}{C^+_{i+\frac{1}{2},j}}\right)^2\xi_k^2 + \left[\left(\frac{C^-_{i+\frac{1}{2},j}}{C^+_{i+\frac{1}{2},j}}\right)^2 - 1\right]\eta_l^2 > 0$

    $\xi^+ = -\sqrt{\left(\frac{C^-_{i+\frac{1}{2},j}}{C^+_{i+\frac{1}{2},j}}\right)^2\xi_k^2 + \left[\left(\frac{C^-_{i+\frac{1}{2},j}}{C^+_{i+\frac{1}{2},j}}\right)^2 - 1\right]\eta_l^2}$

    if $\xi_{k'} \leqslant \xi^+ < \xi_{k'+1}$ for some k'

    then $f^-_{i+\frac{1}{2},jkl} = \frac{\xi_{k'+1}-\xi^+}{\Delta\xi}f_{i+1,j,k',l} + \frac{\xi^+ - \xi_{k'}}{\Delta\xi}f_{i+1,j,k'+1,l}$
  - else
    $f^-_{i+\frac{1}{2},jkl} = f_{i,j,k',l}$ where $\xi_{k'} = -\xi_k$
  - end

The flux $f^{\pm}_{i,j+\frac{1}{2},kl}$ can be constructed similarly.

The 2d version of Scheme II can be constructed similarly.

As introduced in Section 2.2, the essential difference between 1D and 2D flux definition is that in 2D case, the phenomenon that a wave is reflected at the interface does occur. While in 1D, a wave is always transmitted across an interface with a change of slowness.

Since the gradient of the wave speed at its smooth points are bounded by an upper bound, this scheme, similar to the 1D scheme, is also subject to a hyperbolic CFL condition under which the scheme is positive, and Hamiltonian preserving.

6. Numerical examples

In this section, we present numerical examples to demonstrate the validity of the proposed schemes and to study their accuracy. In the numerical computations the second order TVD Runge–Kutta time discretization [40] is used. In Example 6.2, we compare the results of Schemes I and II. In other examples, we present the numerical results using Scheme I.

Example 6.1. An 1D problem with exact L^{∞}-solution. Consider the 1D Liouville equation

$$f_t + c(x)\,\mathrm{sign}(\xi) f_x - c_x|\xi| f_\xi = 0 \tag{6.1}$$

with a discontinuous wave speed given by

$$c(x) = \begin{cases} 0.6, & x < 0, \\ 0.5, & x > 0. \end{cases}$$

The initial data are given by

$$f(x,\xi,0) = \begin{cases} 1, & x < 0,\ \xi > 0,\ \sqrt{x^2+\xi^2} < 1, \\ 1, & x > 0,\ \xi < 0,\ \sqrt{x^2+\xi^2} < 1, \\ 0, & \text{otherwise}, \end{cases} \tag{6.2}$$

as shown in the upper part in Fig. 2 which depicts the non-zero part of $f(x,\xi,0)$.

The exact solution at $t = 1$ is given by

$$f(x,\xi,1) = \begin{cases} 1, & 0 < x < 0.5,\ 0 < \xi < 1.2\sqrt{1-(1.2x-0.6)^2}, \\ 1, & 0 < x < 0.5,\ -\sqrt{1-(x+0.5)^2} < \xi < 0, \\ 1, & -0.4 < x < 0,\ 0 < \xi < \sqrt{1-(x-0.6)^2}, \\ 1, & -0.6 < x < 0,\ -\frac{1}{1.2}\sqrt{1-(\frac{x}{1.2}+0.5)^2} < \xi < 0, \\ 0, & \text{otherwise}, \end{cases} \tag{6.3}$$

as shown in the lower left part in Fig. 2.

The numerical solution computed with a 100×101 cell on the domain $[-1.5, 1.5] \times [-1.5, 1.5]$ using Scheme I is shown in the lower right part in Fig. 2. The time step is chosen as $\Delta t = \frac{1}{2}\Delta\xi$. It shows a good agreement with the exact solution.

Table 1 compares the l^1-error of the numerical solutions computed by Scheme I using 50×51, 100×101 and 200×201 cells, respectively. This comparison shows that the convergence rate of the numerical solution in

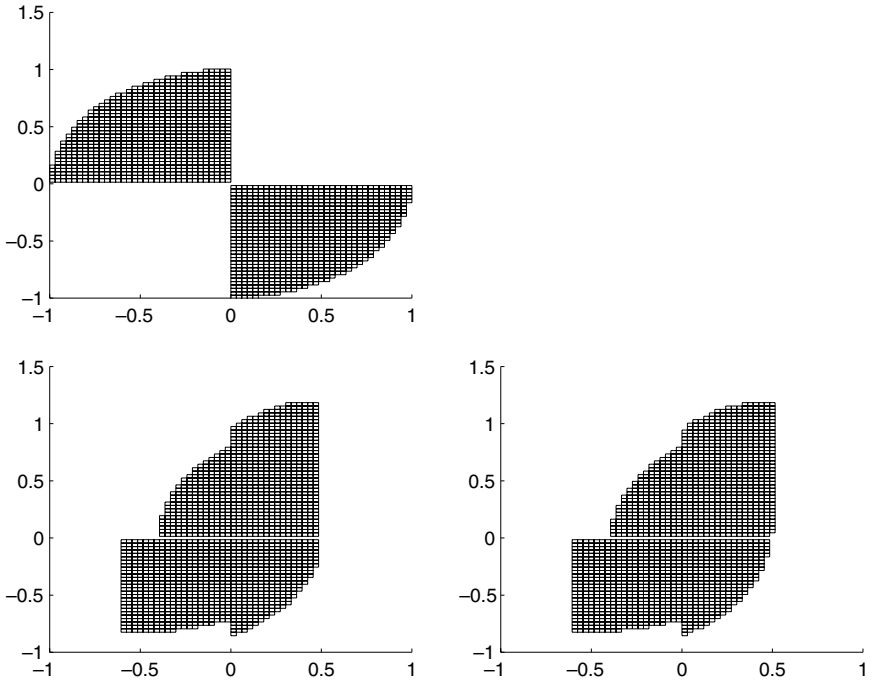

Fig. 2. Example 6.1, solution in the phase space. Upper: non-zero part of the initial data; lower left: non-zero part of exact solution $f(x,\xi,1)$; lower right: the part of numerical solution $f(x,\xi,1) > 0.5$ computed by the 100×101 mesh. The horizontal axis is the position, the vertical axis is the slowness.

Table 1
l^1 error of numerical solutions for f on different meshes

Meshes	50×51	100×101	200×201
	0.269575	0.171837	0.102073

l^1-norm is about 0.68. This agrees with the well established theory [27,42], that the l^1-error by finite difference scheme for a discontinuous solution of a linear hyperbolic equation is at most halfth order.

Example 6.2. Computing the physical observables of an 1D problem with measure-valued solution. Consider the 1D Liouville equation (6.1), where the wave speed is

$$c(x) = \begin{cases} \frac{1}{e-1}, & x \leqslant -1, \\ \frac{1}{e-1} + 1 + x, & -1 < x < 0, \\ \frac{1}{e-1} + 0.5 - x, & 0 < x < 1, \\ \frac{1}{e-1} - 0.5, & x \geqslant 1, \end{cases}$$

the initial data are given by

$$f(x,\xi,0) = \delta(\xi - w(x)) \tag{6.4}$$

with

$$w(x) = \begin{cases} 0.8, & x \leqslant -1.5, \\ 0.8 - \frac{0.8}{(1.5)^2}(x+1.5)^2, & -1.5 < x \leqslant 0, \\ -0.8 + \frac{0.8}{(1.5)^2}(x-1.5)^2, & 0 < x < 1.5, \\ -0.8, & x \geqslant 1.5. \end{cases} \tag{6.5}$$

Fig. 3 plots $w(x)$ in a dashed line.

In this example we are interested in the approximation of the moments, such as the density

$$\rho(x,t) = \int f(x,\xi,t)\,\mathrm{d}\xi,$$

and the averaged slowness

$$u(x,t) = \frac{\int f(x,\xi,t)\xi\,\mathrm{d}\xi}{\int f(x,\xi,t)\,\mathrm{d}\xi}.$$

These quantities are computed by decomposition techniques described in Section 1. We first solve the level set function ψ and modified density function ϕ which satisfy the Liouville equation (6.1) with initial data $\xi - w(x)$ and 1, respectively. Then the desired physical observables ρ and u are computed from the numerical singular integrals (1.7), (1.8), which are computed by the technique described in [21].

The exact slowness and corresponding density at $t = 1$ are given in Appendix A. Fig. 3 shows the exact multivalued slowness in solid line.

Firstly, we give the results using the standard finite difference method (SFDM) by either ignoring or smoothing out the wave speed discontinuity. In the first case, denoted by SFDMI, one uses the same wave speed approximation as in the Hamiltonian-preserving schemes, except that a standard upwind flux is used to approximate the space derivative. Such a method has a hyperbolic CFL condition. In the second approach, denoted by SFDMS, one smoothes out the wave speed discontinuity throughout several grid points, then uses the standard upwind scheme for the space derivative. Fig. 4 presents the numerical densities given by these two approaches using 100×80 mesh. For SFDMS, we choose the transition zone width to be $5\Delta x$, and connect the discontinuous wave speed by a linear function through the transition zone. We take $\Delta t = \frac{1}{5}\Delta\xi$. One can observe that SFDMI gives a wrong solution, while the SFDMS gives correct but smeared solution across the interface. Compared with the results by Schemes I and II, as shown in Fig. 5, SFDMS gives poorer numerical resolution with a much smaller CFL number. In addition, for SFDMS, one has to choose the width of the transition zone properly in order to guarantee the correct solution. A too narrow transition zone leads to a more severe CFL condition and also may produce incorrect results as in SFDMI, while an appropriately wider transition zone relaxes the CFL condition and may allow a convergent solution, but leads to more smeared numerical solution across the interface.

We then present the results computed by our Hamiltonian-preserving schemes. In the computation, the time step is chosen as $\Delta t = \frac{1}{3}\Delta\xi$. We first present numerical results without treating $\xi = 0$ as the domain boundary and performing the delta function integrals (1.7), (1.8) without separating $\xi > 0$ and $\xi < 0$. This is feasible for this example because the zero points of the level set function ψ are away from $\xi = 0$. Fig. 5 shows

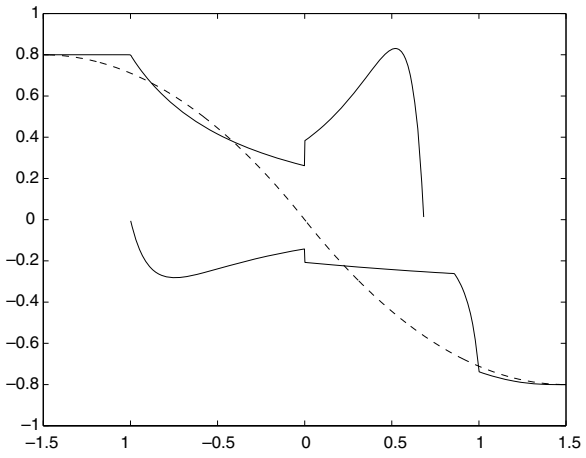

Fig. 3. Example 6.2, slowness. Dashed line: initial slowness $w(x)$; solid line: exact slowness at $t = 1$. The horizontal axis is position, the vertical axis is slowness quantity.

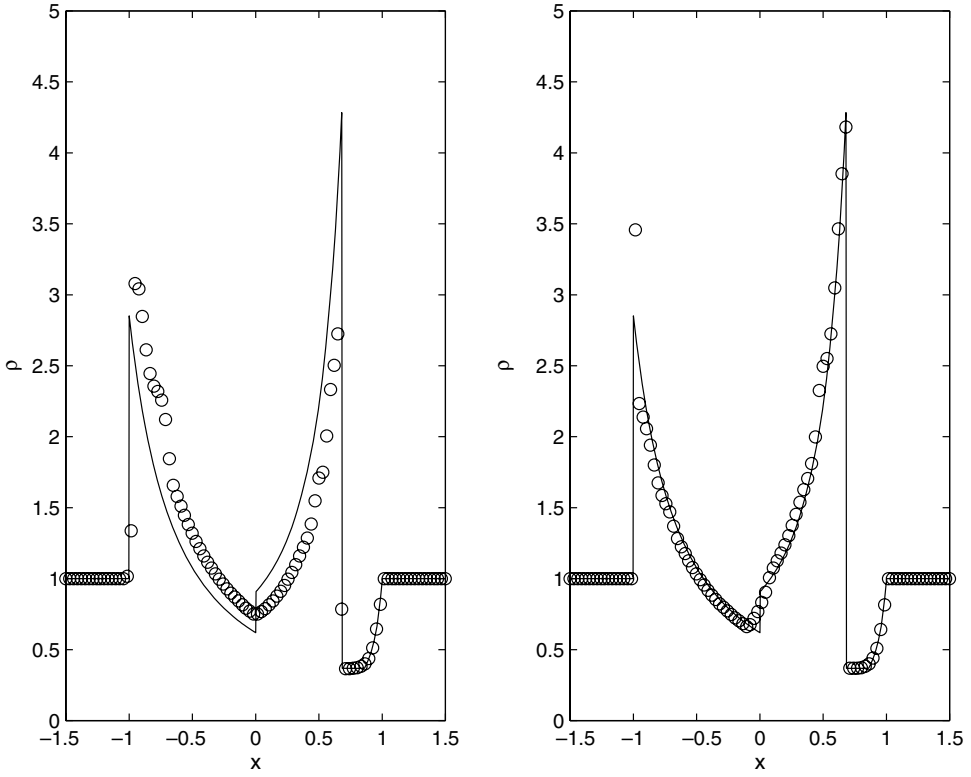

Fig. 4. Example 6.2, density $\rho(x,t)$ at $t=1$. Solid line: the exact solution; '$\circ$': the numerical solutions by the SFDM using 100×80 mesh. Left: SFDMI; Right: SFDMS.

the numerical density and averaged slowness on different meshes using Scheme I plotted against the exact solutions.

Table 2 presents the l^1-errors of numerical densities ρ computed with several different meshes on the domain $[-1.5,1.5]\times[-1.2,1.2]$. Table 3 presents the l^1-errors of numerical averaged slowness u. R_E in the tables denotes the estimated convergence rate. It can be observed that the l^1-convergence rate of the numerical solutions for ρ and u are about first order. In comparison, Scheme II generally has larger numerical errors than Scheme I in this test.

We next present numerical results by treating $\xi=0$ as the domain boundary. We use the meshes on the domain $[-1.5,1.5]\times[-1.2,1.2]$ such that 0 is the a grid point in ξ coordinate. We impose the outflow boundary condition at the domain boundary including the mesh interface $\xi=0$ as done in [22], and perform the delta function integrals (1.7), (1.8) on $\xi>0$ and $\xi<0$ separately.

Fig. 6 shows the calculated density using Schemes I and II together with the exact density on 400×320 mesh. It is observed that the Scheme I gives more accurate numerical solutions than Scheme II near the wave speed jump $x=0$. This is reasonable since Scheme I uses second order interpolation in constructing the Hamiltonian-preserving numerical fluxes while Scheme II only uses first order integration rule, thus Scheme I behaves better in preserving Hamiltonian across the wave speed discontinuity than Scheme II.

Table 4 presents the l^1-errors of numerical densities ρ computed with several different meshes. Table 5 presents the l^1-errors of numerical averaged slowness u. These results are more accurate than those given in Tables 2 and 3 due to the imposing boundary condition at $\xi=0$ as well as performing the delta function integrals on $\xi>0$ and $\xi<0$ separately.

It should be remarked here that the l^1-convergence rate of the numerical solutions reported in [24] is only halfth order due to the discontinuities that exist in the level set function ψ which influence the accuracy when evaluating the delta function integrals (1.7), (1.8) to obtain the moments. In the 1D Liouville equation of

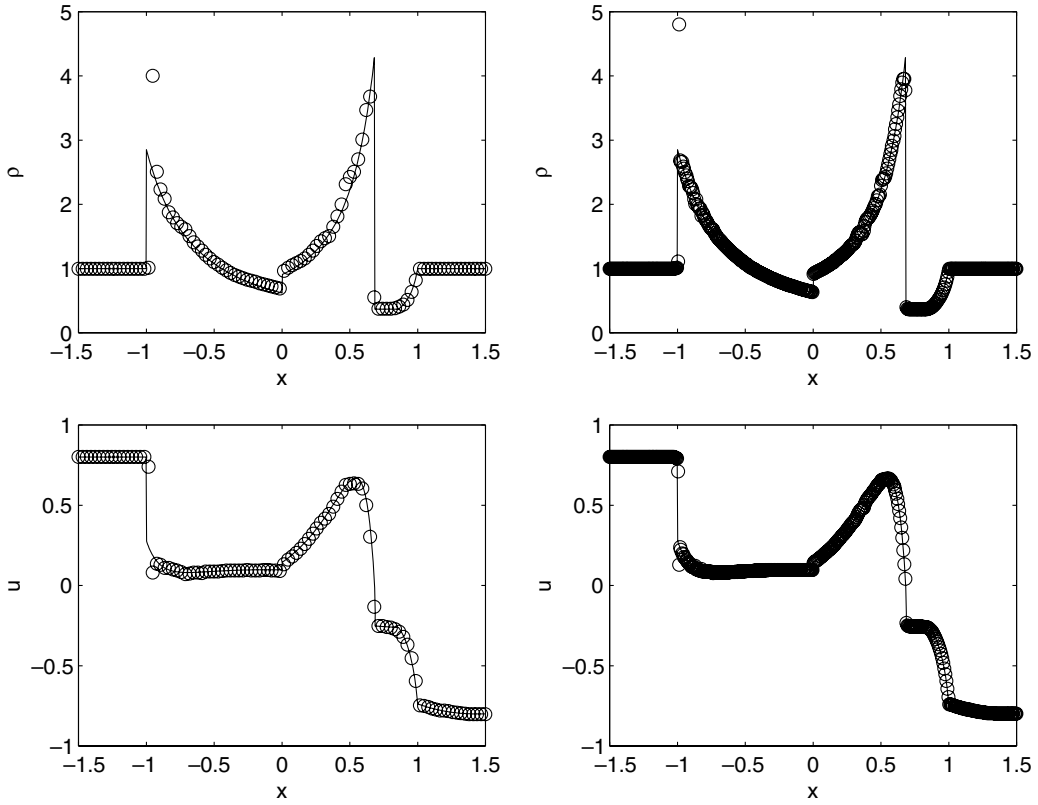

Fig. 5. Example 6.2, density $\rho(x,t)$ and averaged slowness $u(x,t)$ at $t=1$. Solid line: the exact solutions; '$\circ$': the numerical solutions without imposing the boundary condition at $\xi=0$. Upper: density; lower: averaged slowness. Left: 100×81 mesh; Right: 400×321 mesh.

Table 2
l^1 error of numerical densities ρ on different meshes, without imposing the boundary condition at $\xi=0$

Meshes	100×81	200×161	400×321	R_E
Scheme I	3.0174E − 1	1.1792E − 1	6.0890E − 2	1.15
Scheme II	3.0145E − 1	1.2041E − 1	6.2534E − 2	1.13

Table 3
l^1 error of numerical averaged slowness u on different meshes, without imposing the boundary condition at $\xi=0$

Meshes	100×81	200×161	400×321	R_E
Scheme I	4.1166E − 2	2.1229E − 2	8.9561E − 3	1.10
Scheme II	4.1249E − 2	2.2450E − 2	9.8773E − 3	1.03

geometrical optics, there are also such discontinuities for the level set function due to the change of particle slowness at the interface. But such discontinuities typically form the line parallel to the ξ-axis in the domain $\xi>0$ and $\xi<0$, respectively, since the velocity of particles are only dependent on the local wave speed and the sign of particle slowness, thus they do not influence the accuracy when evaluating the integrals (1.7), (1.8) of the delta function along ξ direction at most part of physical domain. Thus the accuracy loss in the moment evaluations does not occur here.

Example 6.3. Computing the physical observables of a 2D problem with a measure-valued solution. Consider the 2D Liouville equation (5.1) with a discontinuous wave speed given by

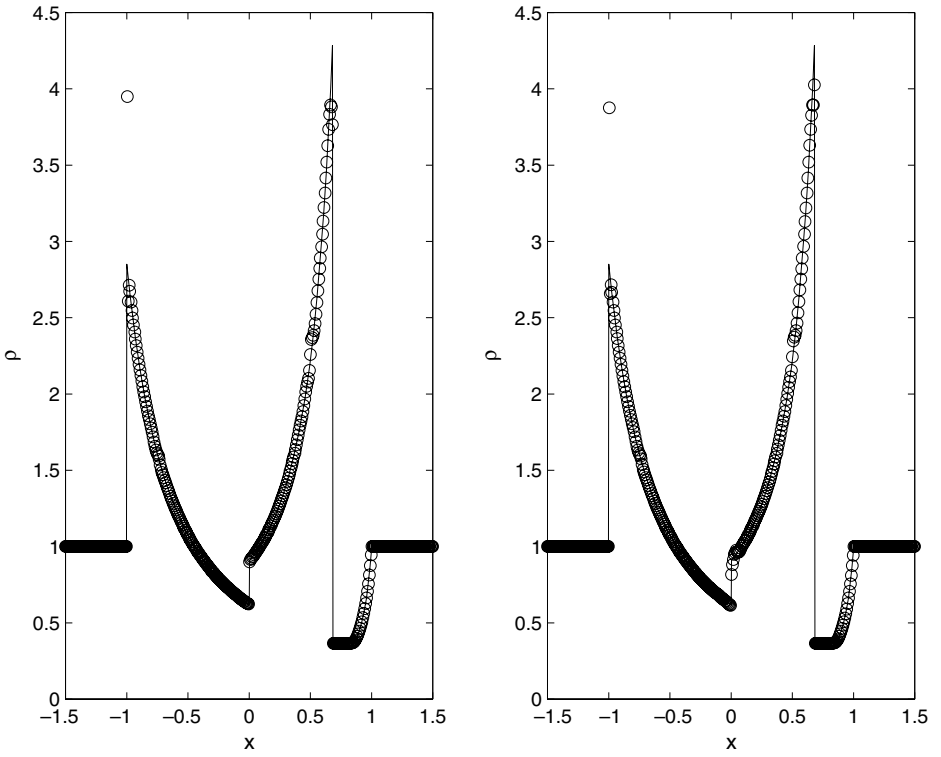

Fig. 6. Example 6.2, density $\rho(x,t)$ at $t=1$. Solid line: the exact solution; '$\circ$': the numerical solutions on 400×320 mesh by imposing the boundary condition at $\xi=0$. Left: Scheme I; Right: Scheme II.

Table 4
l^1 error of numerical densities ρ on different meshes, imposing the boundary condition at $\xi=0$

Meshes	100×80	200×160	400×320	R_E
Scheme I	1.3968E − 1	4.5126E − 2	2.3879E − 2	1.27
Scheme II	1.0579E − 1	5.4503E − 2	2.2259E − 2	1.12

Table 5
l^1 error of numerical averaged slowness u on different meshes, imposing the boundary condition at $\xi=0$

Meshes	100×80	200×160	400×320	R_E
Scheme I	1.8792E − 2	1.3353E − 2	4.2112E − 3	1.08
Scheme II	2.1149E − 2	1.4421E − 2	4.6247E − 3	1.10

$$c(x,y)=\begin{cases}\sqrt{0.8}, & x>0,y>0,\\ \sqrt{0.6}, & \text{else}\end{cases}$$

and a delta function initial data

$$f(x,y,\xi,\eta,0)=\rho(x,y,0)\delta(\xi-p(x,y))\delta(\eta-q(x,y)),$$

where

$$\rho(x,y,0)=\begin{cases}0 & x>-0.1,y>-0.1,\\ 1, & \text{else},\end{cases}$$
$$p(x,y)\equiv q(x,y)=0.6.$$

In this example we aim at computing the numerical density which is the first moment of this delta function solution

$$\rho(x,y,t) = \int\int f(x,y,\xi,\eta,t)\,\mathrm{d}\xi\,\mathrm{d}\eta.$$

The computational domain is chosen to be $[x,y,\xi,\eta] \in [-0.2,0.2]\times[-0.2,0.2]\times[0.3,0.9]\times[0.3,0.9]$.

Set $D_1 = \frac{0.4\sqrt{4}}{\sqrt{15}} - \frac{0.2\sqrt{2}}{3}$, $D_2 = \sqrt{2}$, $D_3 = \sqrt{\frac{9}{8}}$, the exact density at $t = 0.4$ is

$$\rho(x,y,0.4) = \begin{cases} 1, & x < 0 \text{ or } y < 0, \\ D_3, & 0 \leqslant x \leqslant D_1,\ y \geqslant D_2 x, \\ D_3, & 0 \leqslant y \leqslant D_1,\ y \leqslant \frac{x}{D_2}, \\ 0, & \text{otherwise} \end{cases},$$

as shown in the upper left part in Fig. 7 plotted on 50^2 space mesh.

In the computation of this example, the time step is chosen as $\Delta t = \frac{1}{2}\Delta x$. Fig. 7 shows, respectively, the numerical solutions of ρ with 14^4, 26^4 and 50^4 phase space meshes using Scheme I.

Table 6 presents the l^1 errors of ρ on $[0,0.2]\times[0,0.2]$ computed by Scheme I with several different meshes in phase space. The convergence order is about 1/2. In this example, since ρ is discontinuous initially, the modified density function ϕ is also discontinuous in the zero level set in phase space, contributing to the halfth order accuracy in l^1-convergence rate of ρ evaluated by formula (1.7).

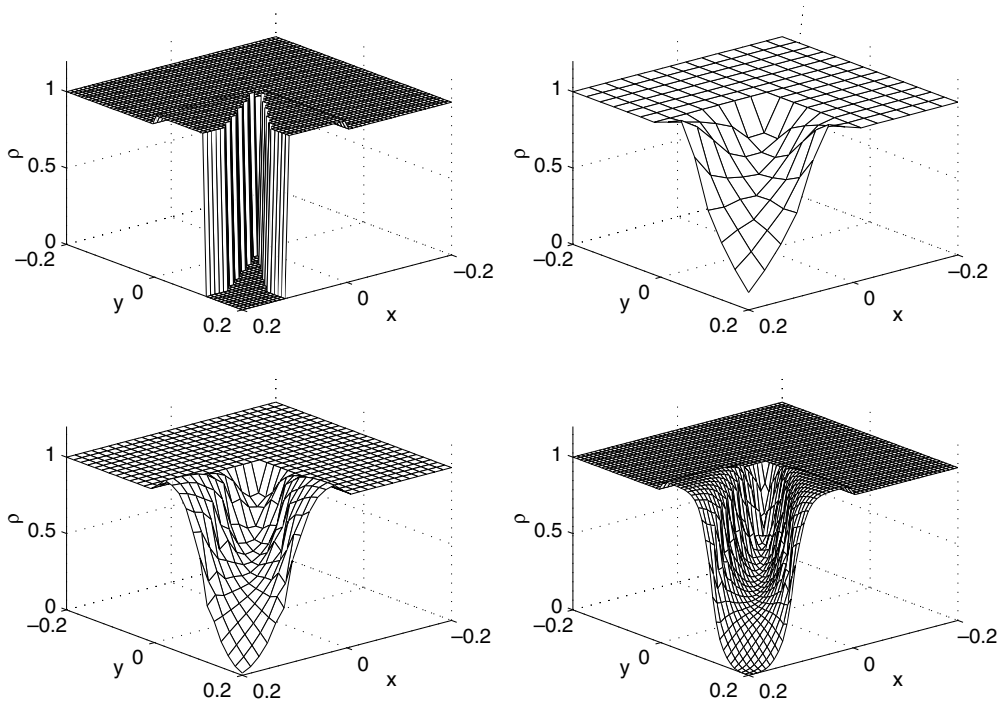

Fig. 7. Example 6.3, density at $t = 0.4$ in space. Upper left: the exact solution; upper right: the numerical solution using 14^4 mesh; lower left: the numerical solution using 26^4 mesh; lower right: the numerical solution using 50^4 mesh.

Table 6
l^1 error of numerical density ρ on $[0,0.2]\times[0,0.2]$ using different meshes

Grid points	14^4	26^4	50^4
	0.012411	0.010044	0.007741

7. Conclusion

In this paper, we constructed and studied two classes of Hamiltonian-preserving schemes for the Liouville equation arising in the phase space description of geometrical optics. These schemes are effective when the local wave speed is discontinuous, corresponding to different media. These schemes have a hyperbolic CFL condition, which is a significant improvement over a conventional discretization. The main idea is to build in the wave behavior at the interface – which conserves the Hamiltonian – into the numerical flux, as was previously done in [24,33]. This gives a selection criterion on the choice of a unique solution to this linear hyperbolic equation with singular coefficients. It allows the wave to be transmitted obeying Snell's law of refraction, or be reflected. We established the stability theory of these discretizations, and conducted numerical experiments to study the numerical accuracy.

In multidimension, we have presented the scheme only in the simple case when an incident plane wave hits the interface that aligns with the grids, and when the reflection and transmission of waves do not occur simultaneously. This idea was extended to the more general case with partial reflections and transmissions [26]. For a curved interface, the principle of Hamiltonian preserving can still be used, however, a different construction of numerical flux at the interface is needed. In addition, the same idea can also be extended to problems with external fields, such as the electrical or electromagnetic fields. There Vlasov–Poisson or Vlasov–Maxwell systems arise. It is also a worthwhile subject to extend it for anisotropic wave propagation [36] and the reduced Liouville equation which is obtained using the constant Hamiltonian. Currently, we are exploring the Hamiltonian-preserving schemes in these more general applications.

Acknowledgments

Research supported in part by NSF Grant No. DMS-0305080, Institute for Mathematics and its Applications (IMA) under a New Direction Visiting Professorship, NSFC under the Project 10228101 and the Basic Research Projects of Tsinghua University under the Project JC2002010.

Appendix A

This appendix gives the exact multivalued slowness and density at $t = 1$ for the problem in Example 6.2. Note that the multivalued slowness ω is the common *zeroes* of ψ_i defined in the introduction, while the averaged slowness u is given by (1.8).

- In the domain $-1.5 < x < -1$, $\omega(x)$ is single phased given by $\omega(x) = 0.8$ and the corresponding density is the constant 1.
- In the domain $-1 < x < 0$, ω has two phases. Set

$$x_1 = \frac{\ln((e-1)x + e) - 1}{e-1} - 1,$$

$$x_2 = \left(\frac{1}{2} + \frac{1}{e-1}\right)\left(1 - \frac{1}{(e-1)\left(\frac{e}{e-1} + x\right)}\right),$$

then

$$\omega_i(x) = \frac{w(x_i)c(x_i)}{c(x)}, \quad i = 1, 2.$$

The densities are given by

$$\rho_1(x) = \frac{1}{(e-1)x + e},$$

$$\rho_2(x) = \frac{\frac{1}{2} + \frac{1}{e-1}}{(e-1)\left(\frac{e}{e-1} + x\right)^2}.$$

- In the domain $0 < x < \frac{e+1}{2e}$, ω has two phases. Set

$$x_1 = \frac{1}{(e-1)\left(1-\frac{2x(e-1)}{e+1}\right)} - \frac{e}{e-1},$$
$$x_2 = \frac{e+1}{2e} + \frac{x}{e},$$

then

$$\omega_i(x) = \frac{w(x_i)c(x_i)}{c(x)}, \quad i = 1, 2.$$

The densities are given by

$$\rho_1(x) = \frac{2}{(e+1)\left(1-\frac{2x(e-1)}{e+1}\right)^2},$$
$$\rho_2(x) = \frac{1}{e}.$$

- In the domain $\frac{e+1}{2e} < x < 1$, ω is single phased. Set

$$x_1(x) = \begin{cases} \frac{e+1}{2e} + \frac{x}{e}, & x < \frac{1}{2}(e-1), \\ 1 + \frac{3-e}{2(e-1)}\left(1 - \ln\left(\frac{2(e-1)}{3-e}\left(\frac{e+1}{2(e-1)} - x\right)\right)\right), & x \geqslant \frac{1}{2}(e-1), \end{cases}$$

then

$$\omega(x) = \frac{w(x_1)c(x_1)}{c(x)}$$

with the corresponding density

$$\rho(x) = \begin{cases} \frac{1}{e}, & x < \frac{1}{2}(e-1), \\ \frac{3-e}{2(e-1)\left(\frac{e+1}{2(e-1)} - x\right)}, & x \geqslant \frac{1}{2}(e-1). \end{cases}$$

- In the domain $1 < x < 1.5$, ω is single phased given by

$$\omega(x) = \begin{cases} -0.8 + \frac{0.8}{(1.5)^2}\left(x - 2 + \frac{1}{e-1}\right)^2, & 1 < x < 2 - \frac{1}{e-1}, \\ -0.8, & 2 - \frac{1}{e-1} < x < 1.5, \end{cases}$$

the corresponding density is the constant 1.

References

[1] G. Bal, J.B. Keller, G. Papanicolaou, L. Ryzhik, Transport theory for acoustic waves with reflection and transmission at interfaces, Wave Motion 30 (1999) 303–327.
[2] J.-D. Benamou, Big ray tracing: multivalued travel time field computation using viscosity solutions of the Eikonal equation, J. Comput. Phys. 128 (1996) 463–474.
[3] J.-D. Benamou, Direct computation of multivalued phase space solutions for Hamilton–Jacobi equations, Commun. Pure Appl. Math. 52 (11) (1999) 1443–1475.
[4] J.-D. Benamou, An introduction to Eulerian geometrical optics (1992–2002), J. Sci. Comp. 19 (1-3) (2001) 63–93.
[5] J.-D. Benamou, I. Solliec, An Eulerian method for capturing caustics, J. Comput. Phys. 162 (1) (2000) 132–163.
[6] F. Bouchut, F. James, One-dimensional transport equations with discontinuous coefficients, Nonlinear Anal. Theory Meth. Appl. 32 (1998) 891–933.

[7] L.-T. Cheng, M. Kang, S. Osher, H. Shim, Y.-H. Tsai, Reflection in a level set framework for geometric optics, Comput. Model. Eng. Sci. 5 (2004) 347–360.
[8] L.-T. Cheng, H.-L. Liu, S. Osher, Computational high-frequency wave propagation using the level set method, with applications to the semi-classical limit of Schrödinger equations, Commun. Math. Sci. 1 (2003) 593–621.
[9] I. Capuzzo Dolcetta, B. Perthame, On some analogy between different approaches to first order PDEs with non-smooth coefficients, Adv. Math. Sci. Appl (1996).
[10] B. Engquist, O. Runborg, Multi-phase computations in geometrical optics, J. Comput. Appl. Math. 74 (1996) 175–192.
[11] B. Engquist, O. Runborg, Multiphase computations in geometrical optics, in: M. Fey, R. Jeltsch (Eds.), Hyperbolic Problems: Theory, Numerics, Applications, Internat. Ser. Numer. Math., vol. 129, Zrich, Switzerland, 1998. ETH Zentrum.
[12] B. Engquist, O. Runborg, Computational high frequency wave propagation, Acta Numer. 12 (2003) 181–266.
[13] B. Engquist, O. Runborg, A.-K. Tornberg, High-frequency wave propagation by the segment projection method, J. Comput. Phys. 178 (2) (2002) 373–390.
[14] E. Fatemi, B. Engquist, S. Osher, Numerical solution of the high frequency asymptotic expansion for the scalar wave equation, J. Comput. Phys. 120 (1995) 145–155.
[15] K. Feng, The Hamiltonian way for computing Hamiltonian dynamics, Applied and Industrial Mathematics, Venice 1989Math. Appl., vol. 56, Kluwer Academic Publishers, Dordrecht, 1991, pp. 17–35.
[16] S. Fomel, J.A. Sethian, Fast phase space computation of multiple arrivals, Proc. Natl. Acad. Sci. USA 99 (11) (2002) 7329–7334.
[17] L. Gosse, Using K-branch entropy solutions for multivalued geometric optics computations, J. Comput. Phys. 180 (2002) 155–182.
[18] L. Gosse, Multiphase semiclassical approximation of an electron in a one-dimensional crystalline lattice II. Impurities, confinement and Bloch oscillations, J. Comput. Phys. 201 (2004) 344–375.
[19] L. Gosse, F. James, Numerical approximations of one-dimensional linear conservation equations with discontinuous coefficients, Math. Comp. (2000) 987–1015.
[20] L. Gosse, S. Jin, X.T. Li, On two moment systems for computing multiphase semiclassical limits of the Schrodinger equation math, Model Methods Appl. Sci. 13 (2003) 1689–1723.
[21] S. Jin, H.L. Liu, S. Osher, R. Tsai, Computing multivalued physical observables for the semiclassical limit of the Schrodinger equation, J. Comput. Phys. 205 (2005) 222–241.
[22] S. Jin, H.L. Liu, S. Osher, R. Tsai, Computing multi-valued physical observables for high frequency limit of symmetric hyperbolic systems, J. Comput. Phys. 210 (2005) 497–518.
[23] S. Jin, S. Osher, A level set method for the computation of multi-valued solutions to quasi-linear hyperbolic PDE's and Hamilton–Jacobi equations, Commun. Math. Sci. 1 (3) (2003) 575–591.
[24] S. Jin, X. Wen, Hamiltonian-preserving schemes for the Liouville equation with discontinuous potentials, Commun. Math. Sci. 3 (2005) 285–315.
[25] S. Jin, X. Wen, The l^1-stability of a Hamiltonian-preserving scheme for the Liouville equation with discontinuous potentials, preprint.
[26] S. Jin, X. Wen, Hamiltonian-preserving schemes for the Liouville equation of geometrical optics with partial transmissions and reflections, SIAM J. Num. Anal., submitted.
[27] N.N. Kuznetsov, On stable methods for solving non-linear first order partial differential equations in the class of discontinuous functions, in: J.J.H. Miller (Ed.), Topics in Numerical Analysis III (Proc. Roy. Irish Acad. Conf.), Academic Press, London, 1977, pp. 183–197.
[28] B. Leimkuhler, S. Reich, Simulating Hamiltonian Dynamics, Cambridge University Press, Cambridge, 2005.
[29] R.J. LeVeque, Numerical Methods for Conservation Laws, Birkhäuser-Verlag, Basel, 1990.
[30] L. Miller, Refraction of high frequency waves density by sharp interfaces and semiclassical measures at the boundary, J. Math. Pures Appl. IX 79 (2000) 227–269.
[31] S. Osher, L.-T. Cheng, M. Kang, H. Shim, Y.-H. Tsai, Geometric optics in a phase-space-based level set and Eulerian framework, J. Comput. Phys. 179 (2) (2002) 622–648.
[32] B. Perthame, C.W. Shu, On positivity preserving finite volume schemes for Euler equations, Numer. Math. 73 (1996) 119–130.
[33] B. Perthame, C. Simeoni, A kinetic scheme for the Saint–Venant system with a source term, CALCOLO 38 (4) (2001) 201–231.
[34] G. Petrova, B. Popov, Linear transport equations with discontinuous coefficients, J. Math. Anal. Appl. 260 (2001) 307–324.
[35] F. Poupaud, M. Rascle, Measure solutions to the linear multidimensional transport equation with non-smooth coefficients, Commun. PDEs 22 (1997) 337–358.
[36] J. Qian, L.-T. Cheng, S. Osher, A level set based Eulerian approach for anisotropic wave propagations, Wave Motion 37 (2003) 365–379.
[37] O. Runborg, Some new results in multiphase geometrical optics, M2AN Math. Model. Numer. Anal. 34 (2000) 1203–1231.
[38] L. Ryzhik, G. Papanicolaou, J. Keller, Transport equations for elastic and other waves in random media, Wave Motion 24 (1996) 327–370.
[39] L. Ryzhik, G. Papanicolaou, J. Keller, Transport equations for waves in a half space, Commun. PDE's 22 (1997) 1869–1910.
[40] C.W. Shu, S. Osher, Efficient implementation of essentially non-oscillatory shock capturing scheme, J. Comput. Phys. 77 (1988) 439–471.
[41] W. Symes, J. Qian, A slowness matching Eulerian method for multivalued solutions of Eikonal equations, J. Sci. Comp. 19 (1–3) (2003) 501–526 (special issue in honor of the sixtieth birthday of Stanley Osher).
[42] T. Tang, Z.H. Teng, The sharpness of Kuznetsov's $O(\sqrt{\Delta x})$ L^1-error estimate for monotone difference schemes, Math. Comp. 64 (1995) 581–589.