Z. angew. Math. Phys. 52 (2001) 847–858
0010-2571/01/050858-12 $ 1.50+0.20/0

Zeitschrift für angewandte Mathematik und Physik ZAMP

Pseudoparticle Representation and positivity analysis of explicit and implicit Steger-Warming FVS schemes

Hua-Zhong Tang and Kun Xu

Abstract. This paper is about the pseudo-particle representation and the positivity analysis of an explicit and an implicit Steger and Warming's flux vector splitting (FVS) scheme for the compressible Euler equations. The positivity proof is based on the motion of pseudo-particles. For the explicit scheme, it shows that the density and the internal energy could keep non-negative values under the CFL-like condition for the Steger-Warming FVS scheme once the initial gas stays in a physically realizable state. For the implicit method, under a stronger CFL–like condition, the positivity property can also be preserved.

Mathematics Subject Classification (2000). 35L65, 65M06, 76M20, 76K05.

Keywords. Euler equations, flux vector splitting, positivity.

1. Introduction

The analysis of positivity for a numerical scheme has obtained much attention in the past years. Practically, it is very important for any scheme to avoid producing negative density or internal energy in the numerical simulation, especially in the high speed and lower density flow regions. Einfeldt et al. [3] first studied the behavior of Godunov–type methods near low densities. They showed that the Godunov scheme [5] is positively conservative while Roe's approximate Riemann solver [14] is not. They also modified Harten, Lax and van Leer's approximate Riemann solver [7] to become a positivity preserving scheme. Linde and Roe [9] discussed the conditions for a second–order multidimensional MUSCL–type scheme to remain positively conservative. Perthame [11] discussed the positivity property for the kinetic scheme (See also [8]). Perthame and Shu [12] discussed the positivity preserving finite volume methods for the compressible Euler equations in general. For the Kinetic Flux Vector Splitting (KFVS) scheme [13], due to the lack of dynamical coupling between the left and right moving particles across a cell interface, the flow updating process can be divided simply into a few subprocesses. Both the positivity and the entropy conditions can be proved by analyzing the same property in each subprocess [18, 10]. Recently, based on the similar consideration, Gressier et. al. [6] gave a general positivity analysis for the FVS schemes, which include the van Leer and the Steger-Warming's FVS methods. Estivalezes and Villedieu [4] constructed a general framework to transform a positive FVS scheme into a positive multidimensional higher–order accurate scheme with the implementation of anti–diffusive terms. The above numerical strategy cannot be used in the positivity proof for the Flux Difference Schemes (FDS), such as the Godunov method, due to their wave interactions. We refer readers to [2, 15, 16, 19, 20, 21] for details about the presentation and analysis of FVS and FDS schemes.

In terms of the physical originality, the idea of splitting fluxes into positive and negative parts can be traced back to the Beam scheme [15], where a few pseudo–particles are constructed from the macroscopic flow variables. In the current paper, we are going to obtain a pseudo-particle representation of the Steger-Warming method, from which the "equivalence" between the Steger-Warming method and the Beam scheme will be re-examined [21]. Based on the motion of individual particle, the positivity preserving property of the explicit and implicit FVS schemes can be conveniently analyzed. The results will show that the 1st-order explicit and implicit Steger-Warming FVS schemes could both preserve positive density and internal energy in its evolution process under a CFL–like condition.

2. Steger–Warming FVS scheme and its particle representation

Consider the one dimensional Euler equations of gas dynamics:

$$\frac{\partial U}{\partial t} + \frac{\partial \mathrm{F}(U)}{\partial x} = 0, \tag{2.1}$$

where

$$U = [\rho,\ m,\ \mathrm{E}]^{\mathrm{T}}, \quad \mathrm{F}(U) = [m,\ mu + p,\ u(\mathrm{E}+p)]^{\mathrm{T}}. \tag{2.2}$$

Here ρ is the density, u is the velocity, $m = \rho u$ is the momentum, $\mathrm{E} = \rho e + \frac{1}{2}\rho u^2$ is the energy density, e is the internal energy, and p is the pressure. For the ideal gas, the equation of state is $p = (\gamma - 1)\rho e$ and $1 < \gamma \le 3$.

The Jacobin matrix $\mathrm{A}(U)$ is given by

$$\mathrm{A}(U) \equiv \frac{\partial \mathrm{F}}{\partial U} = \begin{pmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{2}u^2 & (3-\gamma)u & \gamma - 1 \\ \frac{\gamma-2}{2}u^3 - \frac{a^2}{\gamma-1}u & \frac{3-2\gamma}{2}u^2 + \frac{a^2}{\gamma-1} & \gamma u \end{pmatrix},$$

and it has three real eigenvalues

$$\lambda_1 = u - a, \quad \lambda_2 = u, \quad \lambda_3 = u + a,$$

where a denotes the sound speed, $a = \sqrt{\gamma p/\rho}$. The matrix R of the corresponding right eigenvectors is

$$\mathrm{R} \equiv [\mathrm{R}^{(1)}, \mathrm{R}^{(2)}, \mathrm{R}^{(3)}] = \begin{pmatrix} 1 & 1 & 1 \\ u-a & u & u+a \\ H-ua & \frac{1}{2}u^2 & H+ua \end{pmatrix}.$$

Here H is the enthalpy defined by

$$H = \frac{\mathrm{E}+p}{\rho} = \frac{1}{2}u^2 + \frac{a^2}{\gamma-1}.$$

Thus we have

$$\mathrm{A} = \mathrm{R}\ \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)\ \mathrm{R}^{-1}.$$

With the definition of $\lambda_i^{\pm} = \frac{1}{2}(\lambda_i \pm |\lambda_i|)$, the Jacobin matrix A can be decomposed into a positive A^+ and a negative A^- component, such that

$$\mathrm{A} = \mathrm{A}^+ + \mathrm{A}^-,$$

where $\mathrm{A}^{\pm}$ is given by

$$\mathrm{A}^{\pm} = \mathrm{R}\ \mathrm{diag}(\lambda_1^{\pm}, \lambda_2^{\pm}, \lambda_3^{\pm})\ \mathrm{R}^{-1}.$$

Therefore, the matrix A^+ (A^-) has three non–negative (non–positive) real eigenvalues.

Based on the homogeneity property of the Euler equations (2.1),

$$\mathrm{F}(U) = \mathrm{A}(U)\ U,$$

and the above Steger and Warming's decomposition of the matrix $\mathrm{A}(U)$ into the positive and negative parts, the flux vector $\mathrm{F}(U)$ can be split as

$$\mathrm{F}(U) = \mathrm{F}^+(U) + \mathrm{F}^-(U) \equiv \mathrm{A}^+(U)\ U + \mathrm{A}^-(U)\ U. \tag{2.3}$$

In the following, we are going to present two numerical schemes based on the above flux vector splitting method and a fundamental theorem about its pseudo–particle representation. Let $x_j = j\Delta x$ ($j \in \mathcal{Z}$) be grid points in the x–direction, $t_n = n\Delta t$ ($n = 0, 1, 2, \cdots$) grid points in the t–direction, where Δx and Δt denote the corresponding grid sizes. If we define the cell averaged conservative variables by

$$U_j(t) = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} U(x,t)\ dx,$$

the initial data at each time level can be considered as a piecewise constant. The first order explicit and implicit Steger–Warming FVS schemes can be written in the following conservative form, respectively

$$U_j^{n+1} = U_j^n - \sigma(\mathrm{F}_{j+1/2}^n - \mathrm{F}_{j-1/2}^n), \qquad \text{—(Explicit)} \tag{2.4}$$

and

$$U_j^{n+1} = U_j^n - \sigma(\mathrm{F}_{j+1/2}^{n+1} - \mathrm{F}_{j-1/2}^{n+1}), \qquad \text{—(Implicit)} \tag{2.5}$$

where $\sigma = \Delta t/\Delta x$ and the numerical flux is

$$\mathrm{F}^k_{j+1/2} = \mathrm{F}^+(U^k_j) + \mathrm{F}^-(U^k_{j+1}), \quad k = n \text{ or } n+1. \tag{2.6}$$

For the Steger–Warming's splitting flux vector $\mathrm{F}^\pm(U)$ and flow variable U, we have

Theorem 2.1. *The macroscopic conservative variables U and the associated split flux component $\mathrm{F}^\pm$ in Eq.(2.3) can be written as follows*

$$U_j = \sum_{i=1}^{3} U_{i,j}, \quad \mathrm{F}^\pm(U_j) = \sum_{i=1}^{3} \lambda^\pm_{i,j} U_{i,j}, \tag{2.7}$$

where

$$U_{2,j} = c_2 \begin{pmatrix} \rho_j \\ \rho_j \lambda_{i,j} \\ \frac{1}{2}\rho_j \lambda^2_{i,j} \end{pmatrix}, \quad U_{i,j} = c_i \begin{pmatrix} \rho_j \\ \rho_j \lambda_{i,j} \\ \frac{1}{2}\rho_j \lambda^2_{i,j} + \frac{\gamma(3-\gamma)}{2}\rho_j e_j \end{pmatrix}, \quad i = 1 \text{ or } 3, \tag{2.8}$$

where $c_1 = c_3 = 1/2\gamma$ *and* $c_2 = (\gamma - 1)/\gamma$.

The proof of this theorem is not difficult, and will be omitted here. The above result means that the flow inside each cell j is considered as consisting of three particles and each particle is associated with its individual mass, momentum, and energy, i.e., $U_{i,j}$. Their speeds are $\lambda_{1,j} = u_j - a_j, \lambda_{2,j} = u_j$, and $\lambda_{3,j} = u_j + a_j$, where a_j is the sound speed for the fluid U_j inside cell j. The fluxes are equal to the particle variable $U_{i,j}$ multiplied by the corresponding particle velocity. The theoretical analysis in the next section will be based on the above particle representation.

At the early 80's, the Steger-Warming scheme was considered to be identical to the Beam scheme [21]. Actually, the Beam scheme is different from the above particle representation of the Steger-Warming method. For the Beam scheme, the individual velocities of three particles are

$$\lambda^{\mathrm{B}}_{1,j} = u_j - \sqrt{3p_j/\rho_j}, \quad \lambda^{\mathrm{B}}_{2,j} = u_j, \quad \lambda^{\mathrm{B}}_{3,j} = u_j + \sqrt{3p_j/\rho_j},$$

and its associated mass, momentum, and energy for i th particle in the cell j is

$$U^{\mathrm{B}}_{i,j} = c^{\mathrm{B}}_i \begin{pmatrix} \rho_j \\ \rho_j \lambda^{\mathrm{B}}_{i,j} \\ \frac{1}{2}\rho_j \left(\lambda^{\mathrm{B}}_{i,j}\right)^2 + \frac{3-\gamma}{2(\gamma-1)} p_j \end{pmatrix},$$

where $c^{\mathrm{B}}_1 = c^{\mathrm{B}}_3 = \frac{1}{6}$ and $c^{\mathrm{B}}_2 = \frac{2}{3}$. The split flux component for the Beam scheme is

$$\mathrm{F}^\pm(U_j) = \sum_{i=1}^{3} \lambda^{\mathrm{B},\pm}_{i,j} U^{\mathrm{B}}_{i,j}.$$

In the case of $\gamma = 3$, the Beam and Steger-Warming methods are identical. But, in the general case they have the following differences:

1. The particle speeds and the corresponding weights in terms of the total density are different.

2. The second particle in the Steger-Warming method has no internal energy, i.e., the total energy of this particle is equal to the kinetic energy. As shown in section 3, the absence of the internal energy will effect the positivity preserving property in the implicit Steger-Warming scheme (See **Theorem 3.2** in next section and **Theorem 1** in [17]).

3. Due to the point 2, there is no wonder that the Beam scheme has a better numerical behavior than the Steger-Warming method.

Numerically, the development of the Steger-Warming method is closely associated with the homogeneity of the Euler system, and the application of the Steger-Warming FVS method is limited to this kind of hyperbolic system, such as the isothermal and the Euler equations. However, for the Beam scheme, there is no such a limitation. The Beam scheme can be equally applied to the non-homogeneity system, such as the isentropic flow and the Euler equations with general equation of state. At end, we advocate that the Beam scheme deserves more serious discussion in modern CFD books.

3. Positivity analysis

In this section, we are going to analyze the positivity of the explicit and implicit Steger–Warming schemes (2.4) and (2.5). Before that, we first prove the following useful result:

Lemma 3.1. *Assume that* $\alpha_{i,j+k}$ *and* ρ_{j+k} $(k=-1,0,1;\forall j\in\mathcal{Z})$ *are some non–negative parameters, then we have*

$$\begin{aligned}\mathrm{B}_{il} \equiv &\left(\sum_{k=-1}^{1}\alpha_{i,j+k}\rho_{j+k}\lambda_{i,j+k}^{2}\right)\left(\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\right)\\ &+\left(\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\lambda_{l,j+k}^{2}\right)\left(\sum_{k=-1}^{1}\alpha_{i,j+k}\rho_{j+k}\right)\\ &-2\left(\sum_{k=-1}^{1}\alpha_{i,j+k}\rho_{j+k}\lambda_{i,j+k}\right)\left(\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\lambda_{l,j+k}\right)\geq 0,\end{aligned}$$

where $i,l=1,2,3$.

Proof. Expanding the expression B_{il} for each i and l , and factorizing them, respectively, gives

$$\begin{aligned}\mathrm{B}_{il} = &\alpha_{i,j-1}\rho_{j-1}\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}(\lambda_{i,j-1}-\lambda_{l,j+k})^2\\ &+\alpha_{i,j}\rho_j\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}(\lambda_{i,j}-\lambda_{l,j+k})^2\\ &+\alpha_{i,j+1}\rho_{j+1}\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}(\lambda_{i,j+1}-\lambda_{l,j+k})^2.\end{aligned}$$

Because $\alpha_{i,j}$ and ρ_j, $(i=1,2,3;\forall j\in\mathcal{Z})$ are non–negative, $\mathrm{B}_{il}\geq 0$ is satisfied. This completes the proof of **Lemma 3.1**.

3.1. The explicit Steger–Warming FVS scheme

For the explicit Steger–Warming FVS scheme (2.4), we have

Theorem 3.1. *Assume that* $1<\gamma\leq 3$ *. If* $\rho_j^n\geq 0$ *and* $e_j^n\geq 0$ $(\forall j\in\mathcal{Z}$ *), then under the CFL condition*

$$\sigma\max_{j\in\mathcal{Z}}\{|\lambda_{1,j}|,|\lambda_{2,j}|,|\lambda_{3,j}|\}\leq 1, \tag{3.1}$$

we have

$$\rho_j^{n+1}\geq 0,\qquad e_j^{n+1}\geq 0 \tag{3.2}$$

for all $j\in\mathcal{Z}$ *, where* ρ_j^{n+1} *,* m_j^{n+1} *, and* E_j^{n+1} *are computed by the explicit scheme (2.4).*

Proof. For the sake of convenience, let us introduce the notations

$$\alpha_{i,j-1}=\sigma\lambda_{i,j-1}^{+,n},\quad \alpha_{i,j}=1-\sigma|\lambda_{i,j}^n|,\quad \alpha_{i,j+1}=-\sigma\lambda_{i,j+1}^{-,n}.$$

With the results given in **Theorem 2.1**, we can rewrite Eq.(2.4) as

$$U_j^{n+1}=\sum_{i=1}^{3}\Big[\alpha_{i,j-1}U_{i,j-1}^n+\alpha_{i,j}U_{i,j}^n+\alpha_{i,j+1}U_{i,j+1}^n\Big]. \tag{3.3}$$

Specifically, the first equation in the above expression is

$$\rho_j^{n+1}=\sum_{i=1}^{3}c_i\Big[\alpha_{i,j-1}\rho_{j-1}^n+\alpha_{i,j}\rho_j^n+\alpha_{i,j+1}\rho_{j+1}^n\Big].$$

Under the CFL condition (3.1), all coefficients in front of $\rho_{j\pm1}^n$ and ρ_j^n are non–negative. Therefore, we have $\rho_j^{n+1}\geq 0$ for all $j\in\mathcal{Z}$.

Next, let us consider the proof of the second inequality of Eq.(3.2).

From Eq.(3.3), we have

$$
\begin{aligned}
2(\rho_j^{n+1})^2 e_j^{n+1} &\equiv 2\rho_j^{n+1}\mathrm{E}_j^{n+1} - (m_j^{n+1})^2 \\
&= \tfrac{\gamma(3-\gamma)}{2}\rho_j^{n+1}\sum_{i\neq 2} c_i \sum_{k=-1}^{1} \rho_{j+k}^n \alpha_{i,j+k} e_{j+k}^n \\
&+ \tfrac{1}{2}\sum_{i=1}^{3} c_i^2 \mathrm{B}_{ii} + c_1c_2\mathrm{B}_{12} + c_1c_3\mathrm{B}_{13} + c_2c_3\mathrm{B}_{23}.
\end{aligned}
$$

Using the conclusions of **Lemma 3.1**, we have $e_j^{n+1} \geq 0$ for all $j \in \mathcal{Z}$. Thus, the proof of **Theorem 3.1** is completed.

Remark. Using a different method, Gressier et. al. proved the same theorem [6]. At the same time, they gave another important theorem which states that *if a FVS scheme exactly preserves stationary contact discontinuities, then it cannot be positively conservative*. This is not surprising. As emphasized in [22], all FVS schemes have the intrinsic free wave or particle transport mechanism. In terms of the Boltzmann equation, the collision time is equal to the time step and the mean free path goes to the cell size. Due to the relation between the Boltzmann equation and the Navier-Stokes equations, it is not hard for us to understand that the FVS schemes are intrinsically solving the viscous governing equations. There is no corresponding contact discontinuity waves. The only way to reduce the dissipation in the FVS schemes is to implement particle collisions into the FVS schemes to modify the free transport mechanism.

3.2. The implicit Steger–Warming FVS scheme

For the implicit Steger-Warming scheme (2.5), we have

Theorem 3.2. *Assume that $2 \leq \gamma \leq 3$, and $\rho_j^n \geq 0$ and $e_j^n \geq 0$ for all $j \in \mathcal{Z}$, then (1) if there exits an integer j_1 such that $\alpha_\rho \equiv \rho_{j_1}^{n+1} \leq \rho_j^{n+1}$ for all $j \in \mathcal{Z}$, then $\alpha_\rho \geq 0$ (or $\rho_j^{n+1} \geq 0$ for all $j \in \mathcal{Z}$) under CFL–like condition*

$$\sigma \max_{j\in\mathcal{Z}}\{|\lambda_{1,j}|, |\lambda_{2,j}|, |\lambda_{3,j}|\} < \frac{1}{2}. \tag{3.4}$$

(2) If there is an integer j_2 such that $\alpha_e \equiv e_{j_2}^{n+1} \leq e_j^{n+1}$ for all $j \in \mathcal{Z}$, then $\alpha_e \geq 0$ (or $e_j^{n+1} \geq 0$ for all $j \in \mathcal{Z}$) under condition (3.4). Here ρ_j^{n+1} and e_j^{n+1} are computed by the implicit Steger–Warming scheme (2.5).

Proof. First, with the same technique in [1], let us introduce a small number s, which satisfies $0 < s \ll 1$. After that, we can rewrite Eq.(2.5) as

$$U_j^{n+1} = \frac{s}{1+s}U_j^n + \frac{1}{1+s}\widetilde{U}_j, \tag{3.5}$$

where

$$\widetilde{U}_j = U_j^{n+1} - s\sigma(\mathrm{F}_{j+1/2}^{n+1} - \mathrm{F}_{j-1/2}^{n+1}). \tag{3.6}$$

The above numerical flux $\mathrm{F}_{j+1/2}^{n+1}$ is defined in Eq.(2.6). Similarly, with the notations

$$\alpha_{i,j-1} = s\sigma\lambda_{i,j-1}^{+,n+1}, \quad \alpha_{i,j} = 1 - s\sigma|\lambda_{i,j}^{n+1}|, \quad \alpha_{i,j+1} = -s\sigma\lambda_{i,j+1}^{-,n+1},$$

Eq.(3.6) becomes

$$\widetilde{U}_j = \sum_{i=1}^{3} \widetilde{U}_{i,j} \equiv \sum_{i=1}^{3} \left[\alpha_{i,j-1}U_{i,j-1}^{n+1} + \alpha_{i,j}U_{i,j}^{n+1} + \alpha_{i,j+1}U_{i,j+1}^{n+1}\right]. \tag{3.7}$$

For example, the first component of $\widetilde{U}_j$ is

$$\widetilde{U}_j^{(1)} = \sum_{i=1}^{3} c_i\Big[\alpha_{i,j-1}\rho_{j-1}^{n+1} + \alpha_{i,j}\rho_j^{n+1} + \alpha_{i,j+1}\rho_{j+1}^{n+1}\Big].$$

With the CFL–like condition in (3.4) and $0 < s \ll 1$, all coefficients $\alpha_{i,j\pm k}$ ($k = 0, 1$) in the above expression are non–negative. Using the assumption of the theorem, we have

$$\widetilde{U}_j^{(1)} \geq \left[s\sigma \sum_{i=1}^{3} c_i(\lambda_{i,j-1}^{+} - \lambda_{i,j+1}^{-} - |\lambda_{i,j}|) + 1\right]\alpha_\rho,$$

for all $j \in \mathcal{Z}$. Combining the above inequality with Eq.(3.5), we get

$$\rho_j^{n+1} \geq \frac{1}{1+s}\left[s\sigma \sum_{i=1}^{3} c_i(\lambda_{i,j-1}^{+} - \lambda_{i,j+1}^{-} - |\lambda_{i,j}|) + 1\right]\alpha_\rho, \quad \forall j \in \mathcal{Z}. \tag{3.8}$$

Especially, if take $j = j_1$, then we have

$$\alpha_\rho \geq \frac{1}{1+s}\left[s\sigma \sum_{i=1}^{3} c_i(\lambda_{i,j_1-1}^{+} - \lambda_{i,j_1+1}^{-} - |\lambda_{i,j_1}|) + 1\right]\alpha_\rho. \tag{3.9}$$

On the other hand, since the CFL condition (3.4) implies

$$1 > \frac{1}{1+s}\left[s\sigma \sum_{i} c_i(\lambda_{i,j_1-1}^{+} - \lambda_{i,j_1+1}^{-} - |\lambda_{i,j_1}|) + 1\right],$$

we conclude $\alpha_\rho \geq 0$, i.e., $\rho_j^{n+1} \geq 0$ ($\forall j \in \mathcal{Z}$). Otherwise, multiplying both sides of the above inequality by α_ρ gives

$$\alpha_\rho < \frac{1}{1+s}\left[s\sigma \sum_{i} c_i(\lambda_{i,j_1-1}^{+} - \lambda_{i,j_1+1}^{-} - |\lambda_{i,j_1}|) + 1\right]\alpha_\rho.$$

Comparing it with inequality (3.9), we get

$$\alpha_\rho < \alpha_\rho.$$

which is inconsistent.

Next, let us prove that $\alpha_e \geq 0$. For the sake of convenience, we omit superscript $n+1$ in the following.

Substituting (2.5) with notations in Eqs.(3.5) and (3.6) into $2\rho_j \mathrm{E}_j - (m_j)^2$, we obtain

$$2\rho_j \mathrm{E}_j - (m_j)^2 = 2\Big[\tfrac{s}{1+s}\rho_j^n + \tfrac{1}{1+s}\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(1)}\Big]\Big[\tfrac{s}{1+s}\mathrm{E}_j^n + \tfrac{1}{1+s}\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(3)}\Big] - \Big[\tfrac{s}{1+s}m_j^n + \tfrac{1}{1+s}\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(2)}\Big]^2.$$

The right hand side of the above equation can be expanded as

$$\begin{aligned}2(\rho_j)^2 e_j = {} & \Big(\tfrac{s}{1+s}\Big)^2\Big[2\rho_j^n \mathrm{E}_j^n - (m_j^n)^2\Big] \\ & + \tfrac{2s}{(1+s)^2}\Big[\rho_j^n\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(3)} + \mathrm{E}_j^n\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(1)} - m_j^n\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(2)}\Big] \\ & + \Big(\tfrac{1}{1+s}\Big)^2\Big[2\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(1)}\cdot\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(3)} - \Big(\sum_{i=1}^{3}\widetilde{U}_{i,j}^{(2)}\Big)^2\Big].\end{aligned} \tag{3.10}$$

We denote the three terms on the right hand side of the above equation as I, II, and III, respectively. From the hypotheses of the current Theorem, $\mathrm{I} \geq 0$ is satisfied. Next, we estimate II and III terms. With the assumptions of the current Theorem and **Lemma 3.1**, we have

$$\begin{aligned}\mathrm{II} = {} & \tfrac{2s}{(1+s)^2}\sum_{i=1}^{3}\Big[\rho_j^n\widetilde{U}_{i,j}^{(3)} + \mathrm{E}_j^n\widetilde{U}_{i,j}^{(1)} - m_j^n\widetilde{U}_{i,j}^{(2)}\Big] = \tfrac{2s}{(1+s)^2}\sum_{i\neq 2} c_i\tfrac{\gamma(3-\gamma)}{2}\rho_j^n\rho_j e_j \\ & + \tfrac{2s}{(1+s)^2}\sum_{i=1}^{3} c_i\Big[\rho_j^n\rho_j e_j^n + \tfrac{1}{2}\rho_j^n\rho_j(\lambda_{i,j} - u_j^n)^2\Big] \\ & \geq \tfrac{s}{(1+s)^2}(3-\gamma)\alpha_e\rho_j\rho_j^n, \quad \forall j \in \mathcal{Z},\end{aligned} \tag{3.11}$$

and

$$\begin{aligned}\mathrm{III} = {} & \Big(\tfrac{1}{1+s}\Big)^2\Big[(3-\gamma)\widetilde{U}_j^{(1)}\Big(\gamma\sum_{l\neq 2}c_l\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}e_{j+k}\Big) \\ & + \tfrac{1}{2}\sum_{i=1}^{3}c_i^2\mathrm{B}_{ii} + c_1c_2\mathrm{B}_{12} + c_1c_3\mathrm{B}_{13} + c_2c_3\mathrm{B}_{23}\Big] \\ & \geq \Big(\tfrac{1}{1+s}\Big)^2(3-\gamma)\alpha_e\widetilde{U}_j^{(1)}\Big(\gamma\sum_{l\neq 2}c_l\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\Big), \quad \forall j \in \mathcal{Z}.\end{aligned} \tag{3.12}$$

Combining Eq.(3.10) with inequalities (3.11) and (3.12), we get

$$\begin{aligned}2\rho_j^2 e_j \geq {} & \tfrac{s}{(1+s)^2}(3-\gamma)\alpha_e\rho_j\rho_j^n + \Big(\tfrac{1}{1+s}\Big)^2(3-\gamma)\alpha_e\widetilde{U}_j^{(1)} \\ & \cdot\Big(\gamma\sum_{l\neq 2}c_l\sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\Big), \quad \forall j \in \mathcal{Z}.\end{aligned} \tag{3.13}$$

Under the assumption of **Theorem 3.2** and the following inequality (proved next)

$$\begin{aligned} 2(\rho_j)^2 > \tfrac{s}{(1+s)^2}(3-\gamma)\rho_j\rho_j^n + \left(\tfrac{1}{1+s}\right)^2 (3-\gamma)\widetilde{U}_j^{(1)} \\ \cdot \left(\gamma \sum_{l\neq 2} c_l \sum_{k=-1}^{1} \alpha_{l,j+k}\rho_{j+k}\right), \quad \forall j \in \mathcal{Z}, \end{aligned} \tag{3.14}$$

holds, we will have $\alpha_e \geq 0$ (i.e., $e_j \geq 0$ for all $j \in \mathcal{Z}$). Otherwise, if $\alpha_e < 0$, we can derive a similar inconsistent relation $\alpha_e < \alpha_e$ from Eq.(3.14) and Eq.(3.13).

In the following, the inequality (3.14) will be proved. Since $\rho_j = \frac{s}{1+s}\rho_j^n + \frac{1}{1+s}\widetilde{U}_j^{(1)}$, we have

$$\begin{aligned} (\rho_j)^2 - \tfrac{s}{(1+s)^2}\rho_j\rho_j^n &= \tfrac{s^2}{(1+s)^2}(\rho_j^n)^2\left(1-\tfrac{1}{1+s}\right) \\ &+ \tfrac{s}{(1+s)^2}\rho_j^n\widetilde{U}_j^{(1)}\left(2-\tfrac{1}{1+s}\right) + \tfrac{1}{(1+s)^2}(\widetilde{U}_j^{(1)})^2 \\ &\geq \tfrac{1}{(1+s)^2}(\widetilde{U}_j^{(1)})^2. \end{aligned}$$

Thanks to the requirement $2 \leq \gamma \leq 3$, we further have

$$2(\rho_j)^2 - \frac{s(3-\gamma)}{(1+s)^2}\rho_j\rho_j^n > \frac{2}{(1+s)^2}(\widetilde{U}_j^{(1)})^2.$$

Therefore, we get

$$\begin{aligned} 2(\rho_j)^2 - \tfrac{s}{(1+s)^2}(3-\gamma)\rho_j\rho_j^n - \tfrac{3-\gamma}{(1+s)^2}\widetilde{U}_j^{(1)}\left(\gamma\sum_{l\neq 2} c_l \sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\right) \\ > \tfrac{2}{(1+s)^2}(\widetilde{U}_j^{(1)})^2 - \tfrac{3-\gamma}{(1+s)^2}\widetilde{U}_j^{(1)}\left(\gamma\sum_{l\neq 2} c_l \sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\right) \\ = \tfrac{1}{(1+s)^2}\widetilde{U}_j^{(1)}\left[2\widetilde{U}_j^{(1)} - (3-\gamma)\left(\gamma\sum_{l\neq 2} c_l \sum_{k=-1}^{1}\alpha_{l,j+k}\rho_{j+k}\right)\right] \\ \geq \tfrac{1}{(1+s)^2}(\gamma-1)(\gamma-2)\left(\widetilde{U}_j^{(1)}\right)^2 \geq 0, \end{aligned}$$

for $\gamma \geq 2$. This completes the proof of **Theorem 3.2**.

Remark. For the Beam scheme, due to the existence of the internal energy in the 2nd particle, the positivity can be guaranteed for any γ between 1 and 3 (See [17] for details). In some sense, the implicit Beam scheme will be more robust than the implicit Steger-Warming method.

Acknowledgments

This research was done during H.Z. Tang's visit to Hong Kong University of Science and Technology, and was supported by the National Natural Science Foundation of China, Grant No. 19901031, and the Laboratory of Computational

Physics of Beijing Institute of Applied Physics and Computational Mathematics. K. Xu was supported by the Research Grant Council of Hong Kong through HKUST6132/00P.

References

[1] A. Chalabi, On convergence of numerical schemes for hyperbolic conservation laws with stiff source terms, *Math. Comp.*, **66**(1997), 527–545.

[2] G.Q. Chen and P.G. Lefloch, Entropy flux-splittings for hyperbolic conservation laws Part I: General framework, *Comm. Pure Appl. Math.*, **48**(1995), 691–729.

[3] B. Einfeldt, C.D. Munz, P.L.Roe, and B. Sjogreen, On Godunov–type methods near low densities, *J. Comput. Phys.*, **93**(1991), 273–295.

[4] J.L. Estivalezes and P. Villedieu, High–order positivity–preserving kinetic schemes for the compressible Euler equations, *SIAM J. Numer. Anal.*, **33**(1996), 2050–2067.

[5] S.K. Godunov, A difference scheme for numerical computation of discontinuous solutions of hydrodynamics equations, *Mat. Sb.* **47**(1959), 271–306.

[6] J. Gressier, P. Villedieu and J.M. Moschetta, Positivity of flux vector splitting schemes, *J. Comput. Phys.* **155**(1999), 199–220.

[7] A. Harten, P.D. Lax, and B. van Leer, On upstream differencing and Godunov–type schemes for hyperbolic conservation laws, *SIAM Rev.* **25**(1983), 35–61.

[8] B. Khobalatte and B. Perthame, Maximum principle on the entropy and second–order kinetic schemes, *Math. Comp.* **62**(1994), 119–131.

[9] T. Linde and P.L. Roe, *On multidimensional positively conservative high–resolution schemes*, In *Barriers and Challenges in Computational Fluid Dynamics*, Hampton, VA, 1996, 299–313.

[10] S.H. Lui and K. Xu, Entropy analysis of kinetic flux vector splitting schemes for the compressible Euler equations, *Z. angew. Math. Phys.* **52** (2001), 62–78.

[11] B. Perthame, Second-order Boltzmann schemes for compressible Euler equations in one and two space dimensions, *SIAM J. Numer. Anal.* **29**(1992), 1–29.

[12] B. Perthame and C.W. Shu, On positivity preserving finite volume schemes for Euler equations, *Numer. Math.* **73**(1996), 119–130.

[13] D.I. Pullin, Direct simulation methods for compressible inviscid ideal gas flow, *J. Comput. Phys.* **34**(1980), 231–244.

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

[15] R. Sanders and K. Prendergast, The possible relation of the three–kiloparsec arm to explosions in the galactic nucleus, *Astrophysical J.* **188**(1974), 489–500.

[16] J.L. Steger and R.F. Warming, Flux–vector splitting of the inviscid gas–dynamic equations with applications to finite difference schemes, *J. Comput. Phys.* **40**(1981), 263–293.

[17] H.Z. Tang and K. Xu, Positivity analysis of an implicit kinetic scheme for the compressible Euler equations, preprint (2000).

[18] T. Tang and K. Xu, Gas-kinetic schemes for the compressible Euler equations: Positivity–preserving analysis, *Z. angew. Math. Phys.* **50**(1999), 258–281.

[19] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, 2nd Edition, Springer–Verlag, 1999.

[20] B. van Leer, Flux–vector splitting for the Euler equations, ICASE Report No. 82–30 (1982).

[21] B. van Leer, Flux Vector Splitting for 1990's, Invited lecture for the CFD symposium on Aeropropulsion, NASA Lewis Research Center, April 24-26 (1990), Cleveland, Ohio. NASA CP–3078, 1991.

[22] Kun Xu, *Gas-kinetic Schemes for Unsteady Compressible Flow Simulations*, VKI Fluid Dynamics Lecture Series, 1998-03 (1998).

Hua-Zhong Tang
State Key Laboratory of Scientific and Engineering Computing
Institute of Computational Mathematics
Chinese Academy of Sciences
P.O. Box 2719
Beijing 100080
P.R. China
e-mail: thz@lsec.cc.ac.cn

Kun Xu
Department of Mathematics
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon
Hong Kong
e-mail: makxu@uxmail.ust.hk

(Received: March 8, 2000; May 30, 2000)