Journal of Computational Physics **155,** 199–220 (1999)
Article ID jcph.1999.6337, available online at http://www.idealibrary.com on IDEAL®

Positivity of Flux Vector Splitting Schemes

Jérémie Gressier, Philippe Villedieu, and Jean-Marc Moschetta

École Nationale Supérieure de l'Aéronautique et de l'Espace, ONERA,
BP 4025, 31055 Toulouse Cedex 4, France
E-mail: gressier@onecert.fr, villedieu@onecert.fr, moscheta@supaero.fr

Received January 27, 1999; revised July 14, 1999

Over the last ten years, robustness of schemes has raised an increasing interest among the CFD community. One mathematical aspect of scheme robustness is the positivity preserving property. At high Mach numbers, solving the conservative Euler equations can lead to negative densities or internal energy. Some schemes such as the flux vector splitting (FVS) schemes are known to avoid this drawback. In this study, a general method is detailed to analyze the positivity of FVS schemes. As an application, three classical FVS schemes (Van Leer's, Hänel's variant, and Steger and Warming's) are proved to be positively conservative under a CFL-like condition. Finally, it is proved that for any FVS scheme, there is an intrinsic incompatibility between the desirable property of positivity and the exact resolution of contact discontinuities.

Key Words: stability and convergence of numerical methods; other numerical methods.

INTRODUCTION

In high speed flows computations, robust schemes are necessary to deal with intense shocks or rarefactions. As a result, numerical schemes are likely to produce negative density or internal energy after a finite time step. In highly accelerated flows, the total energy is mainly composed of kinetic energy. Yet, in conservative formulation, both total and kinetic energy are computed independently, and their difference yields the internal energy which may become negative. Computations then update the flow to non-physical states, and make the time integration fail.

In order to give some mathematical interpretation of schemes robustness or weakness in such severe configurations, it is useful to introduce the positivity property: a scheme is said to be positively conservative if, starting from a set of physically admissible states, it can only compute new states with positive densities and internal energies. Perthame [12] first proposed a scheme which satisfies this property. Afterwards, Einfeldt *et al.* [3] gave

0021-9991/99 $30.00

some results concerning Godunov-type schemes. They proved that the Godunov scheme [5] is positively conservative while Roe's scheme [16] is not, and they derived the HLLE method, a positive variant of the HLL schemes family of Harten *et al.* [7]. Later, Villedieu and Mazet [20] proved that Pullin's EFM kinetic scheme [15] (later renamed as KFVS by Deshpande [1]) is positively conservative under a CFL-like condition. Recently, Dubroca [2] proposed a positive variant of Roe's method. This study has to be distinguished from the Larrouturou [8] approach which has been used by Liou [11], where only the density positivity is addressed.

Since any scheme is positively conservative for a zero time step, it is absolutely essential to specify a time step condition when defining the positivity property.

Recently, Linde and Roe [9] extended the pioneering work of Perthame *et al.* [13, 14] and proved the remarkable theorem which states that given a first-order one-dimensional positively conservative scheme one can always build a second-order multidimensional positively conservative scheme for the Euler equations with the van Leer MUSCL approach. In a similar way, Estivalezes and Villedieu [4] have proposed a general framework to transform a positive FVS scheme into a positive multidimensional second-order accurate scheme with a variant of the so-called anti-diffusive flux approach. This is the reason why only first-order one-dimensional methods will be considered in the following. Although, in Linde and Roe's paper, the initial positivity definition includes a CFL-like condition, the final positivity condition which is derived to build the numerical flux of a positive scheme is not actually associated with a maximum allowable time step.

In this work, particular emphasis has been put on the CFL form of the time step condition which guarantees the positivity preserving property. In the following, all other time step conditions for which an arbitrary small time step might be required to update some particular admissible initial conditions will not be considered.

In Section 1, a method adapted for FVS schemes is detailed to provide a necessary and sufficient condition for positivity. Although some schemes, such as the flux vector splitting (FVS) schemes, are known to be robust in various practical situations, to the best of the authors' knowledge, their positivity property has not yet been proved in general. Using the framework derived in Section 1, the positivity of the Van Leer scheme [18] and the one of Steger and Warming [17] is proved in Section 2, and the maximal CFL-like condition is given.

Finally, in Section 3, it is proved that any FVS scheme, which has been designed to preserve stationary contact discontinuities, cannot satisfy the necessary conditions of positivity detailed in Section 1.

1. FVS SCHEMES AND POSITIVITY

The one-dimensional Euler equations can be written in conservation law form as

$$\frac{\partial \mathcal{U}}{\partial t} + \frac{\partial \mathcal{F}(\mathcal{U})}{\partial x} = 0, \tag{1a}$$

where

$$\mathcal{U} = \begin{pmatrix} \rho \\ \rho u \\ \rho E \end{pmatrix} \quad \text{and} \quad \mathcal{F}(\mathcal{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho u H \end{pmatrix} \tag{1b}$$

with the total energy $E = e + \frac{1}{2}u^2$, the total (or stagnation) enthalpy H such that $\rho H = \rho E + p$ and the pressure p, given by the pressure law $p = p(\rho, e)$. For sake of simplicity, this study has been restricted to the case of perfect gases for which the pressure law is given by

$$p = (\gamma - 1)\rho e, \tag{1c}$$

where γ is the ratio of specific heats: a constant such that $1 < \gamma < 3$.

Since one can formally extend any first-order one-dimensional positively conservative method to a second-order multidimensional positively conservative method (see [13, 14, 9]), we will restrict ourselves to the case of first-order schemes for the one-dimensional Euler equations in the following analysis. After a discretization of the integral form of Eq. (1a), conservative explicit methods can be expressed under the form

$$\mathbb{U}_i = \mathcal{U}_i - \frac{\Delta t}{\Delta x}[F_{i+1/2} - F_{i-1/2}], \tag{2}$$

where

- $\mathcal{U}_i$ is the average value over cell Ω_i of the vector of conservative variables ${}^T(\rho, \rho u, \rho E)$ at a given time step. $\mathbb{U}_i$ is the average value in the same sense at the following time step.
- Δx is the measure of cell Ω_i.
- $F_{i+1/2}$ is the numerical flux between the cells Ω_i and Ω_{i+1}. The numerical flux is a function $F_{i+1/2} = F(\mathcal{U}_i, \mathcal{U}_{i+1})$ of the states of both neighboring cells. The numerical flux must satisfy the consistency condition

$$F(\mathcal{U}, \mathcal{U}) = \mathcal{F}(\mathcal{U}) = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ u(\rho E + p) \end{pmatrix} \tag{3}$$

with the closure relation $p = (\gamma - 1)(\rho E - \frac{1}{2}\rho u^2)$ which is derived from Eq. (1c).

DEFINITION 1. For a given state $\mathcal{U}$, the characteristic wave speed $\lambda(\mathcal{U})$ is defined by

$$\lambda(\mathcal{U}) = |u| + \sqrt{\frac{\gamma p}{\rho}}. \tag{4}$$

For a given cell Ω_i, a local CFL number χ_i^{loc} is defined by

$$\chi_i^{loc} = \lambda(\mathcal{U}_i)\frac{\Delta t}{\Delta x}. \tag{5}$$

Remarks.

- The characteristic wave speed $\lambda(\mathcal{U})$ is the maximum wave speed in the flow and is naturally involved in stability conditions. This speed naturally appears in the linearized Euler equations since it is the spectral radius of the Jacobian matrix $\partial\mathcal{F}/\partial\mathcal{U}$.

• This definition is consistent with the well-known CFL condition which aims at ensuring linear stability of the explicit scheme given by Eq. (2). This condition can be written as

$$\Delta t \leq \chi \frac{\Delta x}{\max_{i\in\mathbb{Z}} \lambda(\mathcal{U}_i)}. \tag{6a}$$

It means that the time step must be small enough so that the fastest waves cannot travel across more than one cell during the integration process. Since the fastest wave velocity is approximated by $\lambda(\mathcal{U})$, the CFL number χ generally satisfies $0 < \chi < 1$. Using Definition 1, condition (6a) may be rewritten as

$$\max_{i\in\mathbb{Z}} \chi_i^{loc} \leq \chi. \tag{6b}$$

The discretized conservation equation Eq. (2) can then be rewritten with $\lambda_i = \lambda(\mathcal{U}_i)$,

$$\mathbb{U}_i = \mathcal{U}_i - \frac{\chi_i^{loc}}{\lambda_i}[F_{i+1/2} - F_{i-1/2}]. \tag{7}$$

1.1. *Physical States and Positive Solutions*

For physical reasons, the state $\mathcal{U}$ cannot take any arbitrary value in $\mathbb{R}^3$. It must satisfy the constraints

$$\rho > 0 \qquad \text{and} \qquad e > 0. \tag{8}$$

One can define $\Omega_{\mathcal{U}}$ as the space of physically admissible states. A state is physically admissible if its density ρ and its internal energy $\rho E - 1/2\rho u^2$ are positive. Therefore, the following definition can be given for the open set $\Omega_{\mathcal{U}}$ and its closure $\bar{\Omega}_{\mathcal{U}}$.

DEFINITION 2. The space of physically admissible states, also called positive states, is defined as

$$\Omega_{\mathcal{U}} = \left\{\mathcal{U} = {}^T(u_1, u_2, u_3) \;\middle|\; u_1 > 0 \text{ and } 2u_1u_3 - u_2^2 > 0\right\} \tag{9a}$$

$$\bar{\Omega}_{\mathcal{U}} = \left\{\mathcal{U} = {}^T(u_1, u_2, u_3) \;\middle|\; u_1 \geq 0, u_3 \geq 0 \text{ and } 2u_1u_3 - u_2^2 \geq 0\right\}. \tag{9b}$$

Remarks.

• It can be easily shown (see Lemma 2 in the Appendix) that $\Omega_{\mathcal{U}}$ and $\bar{\Omega}_{\mathcal{U}}$ are *convex cones*. This means that for Ω denoting either $\Omega_{\mathcal{U}}$ or $\bar{\Omega}_{\mathcal{U}}$, the following property holds

$$\forall \mathcal{U}_1, \mathcal{U}_2 \in \Omega, \forall\, \alpha_1, \alpha_2 > 0, \qquad \alpha_1\mathcal{U}_1 + \alpha_2\mathcal{U}_2 \in \Omega. \tag{10}$$

• Although vacuum is an admissible state, it has not been added to $\Omega_{\mathcal{U}}$ since it is not expected to be reached in practical computations. Nevertheless, it belongs to $\bar{\Omega}_{\mathcal{U}}$.

• According to Definition 2, $\Omega_{\mathcal{U}}$ is an open set. $\bar{\Omega}_{\mathcal{U}}$ is the closure of $\Omega_{\mathcal{U}}$.

• The true internal energy is calculated using $\rho e = u_3 - (1/2)(u_2^2/u_1)$. Yet, because of its simplicity, the expression in the definition will be used to prove its positivity.

DEFINITION 3. A scheme is said to be *positively conservative* if and only if there exists a constant χ, such that ensuring both the conditions

$$\bullet \ \forall i \in \mathbb{Z}, \qquad \mathcal{U}_i \in \Omega_{\mathcal{U}} \tag{11a}$$

$$\bullet \ \Delta t \leq \chi \frac{\Delta x}{\max_{i\in\mathbb{Z}} \lambda(\mathcal{U}_i)} \tag{11b}$$

implies

$$\forall i \in \mathbb{Z}, \qquad \mathbb{U}_i \in \Omega_{\mathcal{U}}. \tag{12}$$

Remarks.

- The definition means that a scheme is said to be positively conservative if it leaves the set of admissible state invariant under a CFL-like condition.
- If a scheme is positively conservative for a given CFL number χ, then it remains positively conservative for any CFL number $\chi' \leq \chi$. Indeed, it is a straightforward consequence of the property of convexity of $\Omega_{\mathcal{U}}$,

$$\mathcal{U} - \frac{\chi'}{\lambda}\Delta F = \frac{\chi'}{\chi}\Big(\mathcal{U} - \frac{\chi}{\lambda}\Delta F\Big) + \frac{\chi - \chi'}{\chi}\mathcal{U}. \tag{13}$$

- For $\Delta t = 0$, according to Eq. (2), one has $\forall i \in \mathbb{Z}$, $\mathbb{U}_i = \mathcal{U}_i \in \Omega_{\mathcal{U}}$ whatever the scheme and its flux function are. So, for any continuous flux function F, since $\Omega_{\mathcal{U}}$ is an open subset of $\mathbb{R}^3$, whatever initial conditions $\mathcal{U}_i$ are in $\Omega_{\mathcal{U}}$, one can find Δt small enough which will preserve positivity of states $\mathbb{U}_i$.

Consequently, the property of positivity does not rely on proving that it exists Δt such that $(\forall i \in \mathbb{Z}, \mathcal{U}_i \in \Omega_{\mathcal{U}} \Rightarrow \mathbb{U} \in \Omega_{\mathcal{U}})$, but it consists of proving that this time step is not too small compared to a Δt given by the stability condition (6a). Otherwise, one can find a situation in which a physical admissible state can only be obtained by a vanishing time step, which is not acceptable for practical gas dynamics applications.

- On the contrary, a scheme is said to be *non-positive* if

$$\forall \chi > 0, \exists (\mathcal{U})_{i\in\mathbb{Z}} \in \Omega_{\mathcal{U}}, \qquad \mathbb{U}_i \notin \Omega_{\mathcal{U}}. \tag{14}$$

For a non-positive scheme, one may have to use an extremely small time step to update the solution and may not be able to produce a physically admissible solution after a finite period of time.

1.2. *Positivity of FVS Schemes*

Flux vector splitting (FVS) schemes are built by adding the contributions of both cells located on either sides of a given interface. The numerical flux of any FVS method can be expressed as

$$F_{i+1/2} = F^+(\mathcal{U}_i) + F^-(\mathcal{U}_{i+1}). \tag{15}$$

The consistency condition Eq. (3) becomes

$$F^+(\mathcal{U}) + F^-(\mathcal{U}) = \mathcal{F}(\mathcal{U}), \tag{16}$$

where $\mathcal{F}$ is the exact Euler flux.

The aim of this section is to derive a necessary and sufficient condition to ensure the positivity of a given FVS scheme. This study has been restricted to a class of FVS schemes in which the fluxes $F^{\pm}$ satisfy the symmetry property

$$\overline{F^{-}(\mathcal{U})} = -F^{+}(\bar{\mathcal{U}}), \tag{17}$$

where $\bar{X}$ is the symmetric vector ${}^{T}(x_1, -x_2, x_3)$ of $X = {}^{T}(x_1, x_2, x_3)$. This property is a straightforward consequence of the flux isotropy: flux formulation is invariant by rotation of the coordinates system. Therefore, this requirement is not actually a real restriction since in practice all available FVS schemes satisfy the symmetry property.

For all FVS methods which satisfy the symmetry property, the F^{+} function is sufficient to define a FVS scheme since the F^{-} function can be computed from Eq. (16), and then, the numerical flux can be obtained from Eq. (15). Furthermore, the following notation is defined

$$F^{\star}(\mathcal{U}) = F^{+}(\mathcal{U}) - F^{-}(\mathcal{U}). \tag{18}$$

An additional assumption on numerical fluxes is necessary to proceed to the proof of Theorem 1. Since $\mathcal{U}$ can be expressed as a function of ρ, u, and a, $F^{\pm}$ is also a function of these three variables. Keeping the same notations when writing F^{+} in other variables, the assumption is expressed as

$$\forall u, a \in \mathbb{R} \times \mathbb{R}^{+}, \qquad \lim_{\rho \to 0} F^{\pm}(\rho, u, a) = 0. \tag{19}$$

In fact, $F^{\pm}(\mathcal{U})$ is generally an homogeneous function of ρ, and the previous assumption Eq. (19) is not restrictive. Obviously, $\mathcal{U}$ shares the same property.

THEOREM 1. *A given FVS scheme satisfying properties* (16), (17), *and* (19) *is positively conservative if and only if its $F^{\pm}$ functions satisfy both the properties*:

- $\forall \mathcal{U} \in \Omega_{\mathcal{U}}, \qquad F^{+}(\mathcal{U}) \in \bar{\Omega}_{\mathcal{U}}$ (20a)
- $\exists \chi > 0, \ \forall \mathcal{U} \in \Omega_{\mathcal{U}}, \qquad \mathcal{U} - \dfrac{\chi}{\lambda(\mathcal{U})} F^{\star}(\mathcal{U}) \in \bar{\Omega}_{\mathcal{U}}.$ (20b)

In that case, *the less restrictive positivity condition is expressed as*

$$\forall i \in \mathbb{Z}, \qquad \chi_i^{loc} < \chi_{opt}, \tag{21}$$

*where χ_{opt} is the greatest constant χ satisfying (20*b*).*

Remarks.

- If (F^{+}) satisfies condition (20a), then $(-F^{-})$ and $(F^{\star})$ belong to $\Omega_{\mathcal{U}}$.
- As it has been pointed out in Definition 3 of a positive scheme, such a FVS scheme is positively conservative while using any CFL number $\chi \leq \chi_{opt}$ by convexity considerations.
- The above double condition is not only a sufficient condition of positivity but also a necessary condition which can be very helpful to show that a given FVS method is not positively conservative.

Proof. The conservation Eq. (7) can be expressed as the sum of the contributions of three cells: in the case of FVS schemes, Eq. (7) is rewritten as

$$\mathbb{U}_i = \mathcal{U}_i - \frac{\chi_i^{loc}}{\lambda_i}[F^+(\mathcal{U}_i) + F^-(\mathcal{U}_{i+1}) - F^+(\mathcal{U}_{i-1}) - F^-(\mathcal{U}_i)] \tag{22a}$$

$$= \mathcal{U}_i - \frac{\chi_i^{loc}}{\lambda_i}[F^\star(\mathcal{U}_i) - \overline{F^+(\mathcal{U}_{i+1})} - F^+(\mathcal{U}_{i-1})] \tag{22b}$$

$$= \mathcal{W}_0(\mathcal{U}_i) + \frac{\chi_i^{loc}}{\lambda_i}\mathcal{W}_L(\mathcal{U}_{i-1}) + \frac{\chi_i^{loc}}{\lambda_i}\mathcal{W}_R(\mathcal{U}_{i+1}), \tag{22c}$$

where

$$\mathcal{W}_0(\mathcal{U}) = \mathcal{U} - \frac{\chi^{loc}}{\lambda}F^\star(\mathcal{U}) \tag{23a}$$

$$\mathcal{W}_L(\mathcal{U}) = F^+(\mathcal{U}) \tag{23b}$$

$$\mathcal{W}_R(\mathcal{U}) = -F^-(\mathcal{U}) = \overline{F^+(\mathcal{U})}. \tag{23c}$$

• *Conditions (20*a*) and (20*b*) are sufficient.* On one hand, using condition (20a) and that the symmetry operator keeps $\bar{\Omega}_\mathcal{U}$ invariant, one has $F^+ \in \bar{\Omega}_\mathcal{U} \Rightarrow \overline{F^+} \in \bar{\Omega}_\mathcal{U}$, and then $\mathcal{W}_L$ and $\mathcal{W}_R$ are physically admissible states.

On the other hand, $\mathcal{W}_0$ may be rewritten

$$\mathcal{W}_0(\mathcal{U}_i) = \mathcal{U}_i - \frac{\chi_i^{loc}}{\lambda_i}F^\star(\mathcal{U}_i) \tag{24a}$$

$$= \frac{\chi - \chi_i^{loc}}{\chi}\mathcal{U}_i + \frac{\chi_i^{loc}}{\chi}\left(\mathcal{U}_i - \frac{\chi}{\lambda_i}F^\star(\mathcal{U}_i)\right). \tag{24b}$$

Assuming that $\forall i \in \mathbb{Z},\ \chi_i^{loc} < \chi$ as a usual CFL condition, one has $(1 - \chi^{loc}/\chi)\mathcal{U}_i \in \Omega_\mathcal{U}$ and condition (20b) implies that the second term of Eq. (24b) belongs to $\bar{\Omega}_\mathcal{U}$. Hence (Lemma 3), $\mathcal{W}_0 \in \Omega_\mathcal{U}$. Using Lemmas 2 and 3 (see Appendix),

$$\mathcal{W}_L, \mathcal{W}_R \in \bar{\Omega}_\mathcal{U}, \qquad \mathcal{W}_0 \in \Omega_\mathcal{U} \Rightarrow \mathbb{U}_i \in \Omega_\mathcal{U} \tag{25}$$

$\forall i \in \mathbb{Z},\ \mathbb{U}_i$ is a physically admissible state and the scheme is positively conservative.

• *Condition (20*a*) is necessary.* If this condition is not satisfied, then

$$\exists\, \mathcal{U}_c \in \Omega_\mathcal{U}, \qquad F^+(\mathcal{U}_c) \notin \bar{\Omega}_\mathcal{U} \tag{26}$$

One can rewrite the updated state $\mathbb{U}_i$ with the following set of initial conditions

$$\begin{array}{cccccccc} & \mathcal{U}_c & & \mathcal{U}_c & \mathcal{U}_p & \mathcal{U}_p & \mathcal{U}_p & \\ \cdots & i-2 & & i-1 & i & i+1 & i+2 & \cdots \end{array}$$

$$\mathbb{U}_i = \underbrace{\mathcal{U}_p - \frac{\chi_i^{loc}}{\lambda_{\max}}F^+(\mathcal{U}_p)}_{\mathcal{W}_p} + \underbrace{\frac{\chi_i^{loc}}{\lambda_{\max}}F^+(\mathcal{U}_c)}_{\mathcal{W}_c}, \tag{27}$$

where $\lambda_{\max} = \max(\lambda_c, \lambda_p)$.

Let $\Delta_{\mathcal{U}} = \mathbb{R}^3 - \bar{\Omega}_{\mathcal{U}}$. Since $F^+(\mathcal{U}_c) \notin \bar{\Omega}_{\mathcal{U}}$ and $\Delta_{\mathcal{U}}$ is an open set, there exists a ball around $\mathcal{W}_c$, whose radius is not zero and included in $\Delta_{\mathcal{U}}$. Since $\mathcal{W}_c$ only depends on u_p and a_p through $\lambda_{\max}$ but not on ρ_p, one can make ρ_p decrease while keeping $\mathcal{W}_c$ constant. Then, using assumption (19), one can find small enough density ρ_p such that the updated state $\mathbb{U}_i$ is yet in the ball, hence not in $\Omega_{\mathcal{U}}$.

Hence, for all CFL numbers χ satisfying condition (20a), one can always find some initial conditions such that the non-positivity of F^+ could not be balanced and $\mathbb{U}_i \notin \Omega_{\mathcal{U}}$.

• *Condition (20*b*) is necessary.* If this condition is not satisfied, then

$$\forall \chi > 0, \exists\, \mathcal{U}_c \in \Omega_{\mathcal{U}}, \qquad \mathcal{U}_c - \frac{\chi^{loc}}{\lambda} F^\star(\mathcal{U}_c) \notin \bar{\Omega}_{\mathcal{U}}. \tag{28}$$

Then, one can write the updated state under the form $\mathbb{U}_i = \mathcal{W}_c + \mathcal{W}_p$ with $\mathcal{W}_p \in \Omega_{\mathcal{U}}$ and $\mathcal{W}_c \notin \bar{\Omega}_{\mathcal{U}}$. In the same way as in the first part of the present proof, for any CFL number χ, one can always find $\mathcal{U}_c$ satisfying Eq. (28) and then adjust densities of the neighboring cells small enough such that the non-positivity of $\mathcal{W}_c = \mathcal{U}_c - \chi^{loc}/\lambda F^\star(\mathcal{U}_c)$ could not be balanced.

The proof is completed.

Therefore, owing to the particular property of FVS schemes that yields separate contributions of the local cell $\mathcal{U}_i$ and its neighbors $\mathcal{U}_{i-1}$ and $\mathcal{U}_{i+1}$, the positivity of a given FVS scheme is ruled by two necessary and sufficient conditions.

2. POSITIVITY OF SOME CLASSICAL FVS SCHEMES

Some FVS schemes are already known to be positively conservative (EFM [20] and Perthame's kinetic scheme [12]). Some other classical FVS schemes such as the one of van Leer [18] or Steger and Warming [17] are known to be very robust and do not produce negative states. However, to the best of the authors' knowledge, their intrinsic positivity property has not yet been proved.

In this section, both conditions (20a) and (20b) will be used to prove that those schemes are positively conservative. Moreover, a maximum CFL number $\chi(M)$, which only depends on the local Mach number, will be expressed as a necessary and sufficient condition for positivity. Using the smallest value of $\chi(M)$ for all Mach numbers will provide a sufficient CFL condition for positivity which may be used in practical computations. Here are some practical details to describe the method which will be applied in the following

• First, $(F^+ \in \Omega_{\mathcal{U}})$ is a necessary condition to prove the positivity of a scheme. If this condition is not satisfied, one can always find some states $(\mathcal{U})_i$ for which $\mathcal{W}_0$ will not be able to balance the non-positivity of F^+ as demonstrated in Subsection 1.2. Positivity of $F^+ = {}^T(f_1, f_2, f_3)$ is proved in the same way as it is for a state through the evaluation of f_1 and $2 f_1 f_3 - f_2^2$.

• Then, a condition on the time step so that $\mathcal{U} - \Delta t/\Delta x F^\star(\mathcal{U}) \in \Omega_{\mathcal{U}}$ has to be extracted. If it can be expressed as a CFL condition, the scheme is shown to be *positively conservative*. If not, according to Theorem 1, the scheme is *non-positive*.

Condition (20b) can be written as $(\mathcal{W}_0 \in \Omega_{\mathcal{U}})$. It needs strict positivity of two terms: mass positivity conditions are generally straightforward to derive. However, internal energy positivity generally requires further algebra. In the case of FVS methods, this second condition

can be easily put under the quadratic form

$$a(M)\zeta(\chi^{loc}, M)^2 + 2b(M)\zeta(\chi^{loc}, M) + c(M) < 0, \tag{29a}$$

where $a(M)$, $b(M)$, $c(M)$ are scalar functions of the local Mach number M and $\zeta(\chi^{loc}, M)$ is a scalar function of both M and the dimensionless time step χ^{loc}. For the schemes considered in the present paper, the three following properties are satisfied: $a(M) > 0$, $c(M) < 0$, and $\zeta(\chi^{loc}, M) \geq 0$ if the mass positivity condition is satisfied. Therefore, the function $\zeta(\chi^{loc}, M)$ has to lie between the roots of the quadratic expression (29a). Since one root is negative, the positivity of internal energy is ensured whenever

$$\zeta(\chi^{loc}, M) < \zeta_{\max} = \frac{-b(M) + \sqrt{b(M)^2 - a(M)c(M)}}{a(M)}. \tag{29b}$$

It will be shown that $\zeta(\chi^{loc}, M)$ is an increasing monotone function of χ^{loc}. Hence, condition (29b) will automatically lead to a condition on the local CFL number χ^{loc}, which is expressed as

$$\chi^{loc} < \chi^{loc}_{\max}(M). \tag{29c}$$

The scheme positivity will be proved in two steps. First, $F^+(\mathcal{U})$ has to be an admissible state. Since this first condition does not involve the local CFL number, it should not lead to stringent conditions. Second, requiring positivity of $\mathcal{W}_0$ will lead to a time step condition which depends on the local Mach number. The final CFL-like condition which will be used to satisfy the positivity property Definition 3 will then be derived by computing the smallest value of the local CFL-like condition for all values of the Mach number.

To derive these conditions, let us define two dimensionless coefficients as functions of the local Mach number

$$K_E = \frac{E}{a^2} = \frac{1}{\gamma(\gamma-1)} + \frac{1}{2}M^2 \tag{30a}$$

$$K_H = \frac{H}{a^2} = \frac{1}{\gamma-1} + \frac{1}{2}M^2. \tag{30b}$$

2.1. *The Fully Upwind Case*

In supersonic areas, the numerical flux is fully upwind for almost every FVS scheme. It means that the numerical flux $F(\mathcal{U}_L, \mathcal{U}_R)$ is equal either to the real flux $\mathcal{F}(\mathcal{U}_L)$ or $\mathcal{F}(\mathcal{U}_R)$ according to the sign of the Mach number. The following analysis remains valid not only for FVS schemes but for all upwind schemes which produce full upwinding in supersonic areas. Nervetheless, although this property seems to be natural for FVS schemes, it does not have to be shared by flux difference splitting (FDS) schemes [10].

For FVS schemes, full upwinding requires that $F^+(\mathcal{U})$ is either null or equal to $\mathcal{F}(\mathcal{U})$ if the absolute local Mach number is greater than one. Furthermore, using the symmetry property, the upwind case with the Mach number greater than one will only be considered here without loss of generality.

LEMMA 1.

$$F(\mathcal{U}) \in \bar{\Omega}_{\mathcal{U}} \qquad \text{if and only if } M \geq \sqrt{\frac{\gamma-1}{2\gamma}} \tag{31a}$$

$$\mathcal{U} - \frac{\chi}{\lambda} F(\mathcal{U}) \in \Omega_{\mathcal{U}} \qquad \text{if and only if } \chi < \chi_{\max} = \frac{|M|+1}{|M|+\sqrt{(\gamma-1)/2\gamma}}. \tag{31b}$$

Remarks.

- The case $F^+ = \mathcal{F}(\mathcal{U})$ is included in this lemma. The other case (for which $F^+ = {}^T(0,0,0)$) always satisfies the conditions of Theorem 1 since the vacuum state $\mathcal{U} = {}^T(0,0,0)$ belongs to $\bar{\Omega}_{\mathcal{U}}$.
- Since most schemes (and particularly VL and SW schemes) are fully upwind for $M > 1$, condition (31a) is not restrictive.
- The condition of Eq. (31b) is necessary and sufficient. Nevertheless, a sufficient condition can be obtained by using the minimum of the local CFL numbers, which is

$$\chi_{opt} = \inf_{M \geq 1} \chi_{\max} = 1. \tag{32}$$

Consequently, *all schemes are positively conservative in regions where the numerical flux is fully upwind under the usual CFL condition* $\chi < 1$. Obviously, this result is not limited to the class of FVS schemes and equally applies to any numerical flux which is fully upwind in supersonic regions.

Proof. (1) Positivity of vector $\mathcal{F}(\mathcal{U})$. Following the method detailed in Subsection 1.2, $F^+(\mathcal{U})$ which derives into $\mathcal{F}(\mathcal{U})$ in supersonic areas has to be equivalent to an admissible state. This vector can be written as

$$\mathcal{F}(\mathcal{U}) = \rho a \begin{pmatrix} M \\ a\left[M^2 + \frac{1}{\gamma}\right] \\ a^2 M K_H \end{pmatrix}, \tag{33}$$

where K_H is defined by Eq. (30b). Mass positivity is straightforward since $M \geq 1$. Positivity of the quantity $(2 f_1 f_3 - f_2^2)$ leads to

$$\rho^2 a^4 \left[\frac{2M^2}{\gamma(\gamma-1)} - \frac{1}{\gamma^2}\right] \geq 0. \tag{34}$$

The flux $\mathcal{F}(\mathcal{U})$ is then an admissible state if

$$M \geq M_{\min} = \sqrt{\frac{\gamma-1}{2\gamma}}. \tag{35}$$

This condition is always satisfied since $M_{\min} < 1$ and full upwinding only appears in supersonic areas.

(2) Positivity of vector $\mathcal{W}_0$. Following the method described in the beginning of Section 2, developing mass and energy terms of $\mathcal{W}_0$ state will lead to a condition on the time step which will make the scheme positively conservative. The term $\mathcal{W}_0$ can be developed as

$$\mathcal{W}_0 = \mathcal{U} - \frac{\chi^{loc}}{\lambda}\mathcal{F}(\mathcal{U}) = \rho\left(1 - \frac{\chi^{loc}}{M+1}M\right)\begin{pmatrix} 1 \\ a\left[M - \frac{1}{\gamma}\zeta\right] \\ a^2\left[K_E - \frac{M}{\gamma}\zeta\right] \end{pmatrix}, \tag{36}$$

where $\zeta = \chi^{loc}/(1 + M - \chi^{loc}M)$ and K_E has been defined by Eq. (30a).

Mass positivity requires $1 - (\chi^{loc}/(M+1))M > 0$. ζ is then a positive function of χ^{loc} and the Mach number M. By developing $(2u_1u_3 - u_2^2)$, positivity of internal energy leads to the following condition: $\zeta^2 < \frac{2\gamma}{\gamma-1}$. Positivity conditions can be summarized as

$$\text{mass} \qquad \chi^{loc} < \frac{|M|+1}{|M|} \tag{37a}$$

$$\text{internal energy} \qquad \chi^{loc} < \chi_{\max} = \frac{|M|+1}{|M| + \sqrt{(\gamma-1)/2\gamma}}. \tag{37b}$$

Any fully upwind FVS scheme is then positively conservative in supersonic areas under condition (37b), which is the most stringent. Finally, one can check that $\chi_{opt} = \inf_{|M|>1} \chi_{\max}(M) = 1$.

The proof is completed.

Since the upwind case has been addressed, the previous analysis can be applied to all FVS schemes where flux expressions only differ in subsonic areas.

2.2. *Van Leer's Scheme*

The Van Leer scheme (VL) proposed in 1982 [18], and one of its variants (VLH), proposed by Hänel *et al.* [6] satisfy properties (16), (17), and (19). They yield a fully upwind numerical flux in supersonic areas. In subsonic areas, their numerical flux can be expressed under the common expression

$$F^{\pm} = \rho a K_M^{\pm}\begin{pmatrix} 1 \\ a\left[M + \frac{K_P^{\pm}}{\gamma}\right] \\ a^2 K_H^{\pm} \end{pmatrix}, \tag{38}$$

where $K_M^{\pm}$, $K_P^{\pm}$, and $K_H^{\pm}$ are defined by

$$K_M^{\pm} = \pm\frac{(M \pm 1)^2}{4} \tag{39a}$$

$$K_P^{\pm} = \pm(2 \mp M) \tag{39b}$$

$$K_H^{\pm} = \begin{cases} \frac{2}{\gamma^2-1}\left(1 \pm \frac{\gamma-1}{2}M\right)^2 & \text{(VL)} \\ K_H = \frac{1}{\gamma-1} + \frac{M^2}{2} & \text{(VLH).} \end{cases} \tag{39c}$$

These variants only differ from each other in the expression of their energy flux term. After convergence in time, VLH guarantees constancy of the total enthalpy field in the flow.

THEOREM 2. *The Van Leer scheme is positively conservative* $\forall \gamma > 1$. *The optimal CFL number is*

$$\chi_{opt} = \min\Big[\inf_{M\in[0;1]} \chi_{\max}^{VL}(M), 1\Big], \tag{40}$$

where $\chi_{\max}^{VL}(M)$ *is defined by Eq.* (47).

For Van Leer's scheme, $\chi_{opt} = 1$. *For Hänel's variant,* $\chi_{opt} = \min(1, \frac{2}{\gamma})$.

Remarks.

- This condition is necessary and sufficient. $\chi_{\max}^{VL}$ is a complicated function of the local Mach number whose expression strongly depends on the version considered for Van Leer's method (VL or VLH, see Eq. (47)).
- $\chi_{\max}^{VL}$ is defined by Eq. (47) in the subsonic range. In the supersonic range, the scheme is fully upwind and condition (31b) applies.
- Condition (40) is necessary and sufficient. Nevertheless, a sufficient condition can be obtained by using the minimum of the local CFL numbers (including the condition in the supersonic range), which is $\chi_{opt} = 1$ for usual gases where $1 < \gamma < 2$. This means that *Van Leer's original and modified methods are positively conservative under the usual CFL condition* $\chi < 1$.

Proof. (1) Positivity of vector F^+. To satisfy condition (20a), it is necessary to calculate the mass and the internal energy terms of the equivalent state of F^+. One has to prove that F^+ belongs to $\bar{\Omega}_{\mathcal{U}}$ which is the closure of the admissible states space. For both schemes (VL and VLH), the mass term is positive since they have the same expression $\rho a K_M^+$, which is unconditionally positive. On the contrary, the internal energy terms must be developed according to the expressions for K_H^+ associated with each variant. The term $(\rho a K_M^+)^2$ can be simplified because it does not affect the sign of the expression. The positivity of both schemes is ruled by the condition

$$2K_H^+ - \left(M + \frac{K_P^+}{\gamma}\right)^2 \geq 0. \tag{41}$$

- For the VL scheme, Eq. (41) leads to the condition

$$\frac{4}{\gamma^2(\gamma^2-1)}\left(1 + \frac{\gamma-1}{2}M\right)^2 \geq 0 \tag{42a}$$

which is positive $\forall\, M$ since $\gamma > 1$.

- For the VLH scheme, Eq. (41) leads to a parabolic function of M

$$\frac{1}{\gamma^2}\left[(2\gamma-1)M^2 - 4(\gamma-1)M + \frac{2\gamma^2}{\gamma-1} - 4\right] \geq 0 \tag{42b}$$

which is always positive since its minimum equals $2\gamma^2/(\gamma-1)(2\gamma-1)$, which is positive for $\gamma > 1$.

Both schemes then provide a numerical flux F^+ which corresponds to a physical state, without any condition. Hence, both VL and VLH schemes satisfy the first requirement for positivity. There remains to exhibit a CFL-like condition by analyzing the other term $\mathcal{W}_0$.

(2) Positivity of vector $\mathcal{W}_0$. Positivity analysis of vector $\mathcal{W}_0$ will lead to a necessary and sufficient condition on the time step to make the scheme positively conservative. Using Eq. (38), $\mathcal{W}_0$ vector may be written as

$$\mathcal{W}_0 = \mathcal{U} - \frac{\chi^{loc}}{\lambda} F^\star(\mathcal{U}), \tag{43a}$$

where

$$F^\star(\mathcal{U}) = F^+(\mathcal{U}) - F^-(\mathcal{U}) = \rho a \begin{pmatrix} K_M^\star \\ a\left[M K_M^\star + \frac{1}{\gamma} K_P^\star\right] \\ a^2 K_H^\star \end{pmatrix} \tag{43b}$$

with

$$K_M^\star = K_M^+ - K_M^- = \frac{M^2+1}{2} \tag{44a}$$

$$K_P^\star = K_M^+ K_P^+ - K_M^- K_P^- = \frac{1}{2} M(3 - M^2) \tag{44b}$$

$$K_H^\star = K_M^+ K_H^+ - K_M^- K_H^-. \tag{44c}$$

Vector $\mathcal{W}_0$ is then rewritten as

$$\mathcal{W}_0 = \rho \left(1 - \frac{\chi^{loc} K_M^\star}{1+M}\right) \begin{pmatrix} 1 \\ a\left[M - \frac{K_P^\star}{\gamma}\zeta\right] \\ a^2[K_E - \zeta(K_H^\star - K_M^\star K_E)] \end{pmatrix}, \tag{45}$$

where $\zeta = \chi^{loc}/(1 + M - \chi^{loc} K_M^\star)$ and K_E has been defined by Eq. (30a). ζ is positive since mass positivity requires $1 - \chi^{loc} K_M^\star/(1+M) > 0$. Following the method described in Section 2, internal energy positivity leads to a condition under the form of Eq. (29a) with the coefficients

$$a(M) = \left(\frac{K_P^\star}{\gamma}\right)^2 \tag{46a}$$

$$b(M) = K_H^\star - K_M^\star K_E - M\frac{K_P^\star}{\gamma} \tag{46b}$$

$$c(M) = -\frac{2}{\gamma(\gamma-1)}. \tag{46c}$$

Only $b(M)$ differs between the two variants VL and VLH, because of the definition of $K_H^\pm$. Calculations give

$$b(M) = \begin{cases} \frac{(M^2-1)^2}{2\gamma(\gamma+1)} & \text{(VL)} \\ \frac{(M^2-1)^2}{2\gamma^2} & \text{(VLH).} \end{cases} \tag{46d}$$

$\zeta_{\max}(M)$ can be calculated using Eq. (29b). The maximum local CFL number $\chi^{loc}_{\max}$ is then straightforward to obtain by inverting the $\zeta(\chi^{loc}, M)$ function. The internal energy is then positive under the condition

$$\chi^{loc} < \chi^{VL}_{\max} = \frac{(1+M)\zeta_{\max}(M)}{1+((M^2+1)/2)\zeta_{\max}(M)}, \tag{47}$$

since $\zeta_{\max}(M)$ is an intricate function of the local Mach number M. Expressions are not detailed but this limit is plotted as a function of M in Subsection 2.4.

(3) Computation of χ_{opt}. To use the same constant χ whatever the flow is, it is needed to compute the smallest value (for VL or VLH schemes)

$$\chi_{opt} = \inf_{M\in[0;+\infty]} \chi^{VL}_{\max}(M). \tag{48a}$$

Since both schemes are fully upwind in supersonic regions, Lemma 1 applies and

$$\inf_{M\in[1;+\infty]} \chi_{\max}(M) = 1. \tag{48b}$$

A study of the function $\chi^{VL}_{\max}$ has been performed. Calculations are tedious and are not presented here for the sake of simplicity. $\chi^{VL}_{\max}(M)$ is shown to be an increasing then decreasing function in $[0; 1]$. Hence, its smallest value is either $\chi^{VL}_{\max}(M=0)$ or $\chi^{VL}_{\max}(M=1)$. Since, $\chi^{VL}_{\max}$ joins the fully upwind condition at $M=1$, its value is greater than 1. Hence,

$$\chi_{opt} = \min\big(1, \chi^{VL}_{\max}(0)\big). \tag{48c}$$

$\chi^{VL}_{\max}(0)$ can easily be computed and gives

$$\chi^{VL}_{\max}(0) = \frac{\gamma+1}{\gamma} \tag{48d}$$

$$\chi^{VLH}_{\max}(0) = \frac{2}{\gamma}. \tag{48e}$$

Since, $\frac{\gamma+1}{\gamma} > 1$ for $\gamma > 1$, both optimal CFL conditions of the theorem follow.

The proof is completed.

2.3. *Steger and Warming's Scheme*

The Steger–Warming (SW) scheme [17] satisfies the assumptions (16), (17), and (19) too. Its $F^{\pm}$ functions are fully upwind in the supersonic regions. However, in the subsonic area, its expressions are slightly more intricate since they differ according to the sign of the local Mach number. When the Mach number is positive, vector $F^{+}(\mathcal{U})$ is expressed as

$$F^{+}(\mathcal{U}) = \frac{\rho a}{2\gamma}\begin{pmatrix} (2\gamma-1)M+1 \\ a[(2\gamma-1)M^2+(M+1)^2] \\ a^2\big[(\gamma-1)M^3+\frac{(M+1)^3}{2}+\frac{3-\gamma}{2(\gamma-1)}(M+1)\big] \end{pmatrix}. \tag{49a}$$

When negative, $F^+(\mathcal{U})$ vector is expressed as

$$F^+(\mathcal{U}) = \frac{\rho a}{2\gamma}(M+1)\begin{pmatrix} 1 \\ a[M+1] \\ a^2[K_H + M] \end{pmatrix}. \tag{49b}$$

$F^-(\mathcal{U})$ expressions can easily be calculated thanks to the consistency condition: $F^+(\mathcal{U}) + F^-(\mathcal{U}) = \mathcal{F}(\mathcal{U})$.

THEOREM 3. *The Steger and Warming scheme is positively conservative* $\forall\gamma$ *such that* $1 < \gamma < 3$. *The optimal CFL number is*

$$\chi_{opt} = \min\Big[\inf_{M\in[0;1]} \chi^{SW}_{\max}(M), 1\Big] = 1, \tag{50}$$

where $\chi^{SW}_{\max}(M)$ *is defined by Eq.* (55).

Remarks.

- $\chi^{SW}_{\max}$ is defined by Eq. (55) in the subsonic range. In the supersonic range, the scheme is fully upwind and Lemma 1 applies.
- Condition (50) is necessary and sufficient. $\chi^{SW}_{\max}$ is a complex function of the local Mach number (see Eq. (55)). Yet, a sufficient condition can be obtained by using the minimum of the local CFL numbers (including the condition in the supersonic range), which is $\chi_{opt} = 1$. Therefore, *the Steger and Warming method is positively conservative under the usual CFL condition* $\chi < 1$.

Proof. (1) Positivity of vector F^+. For both expressions (49a) and (49b), the mass term is unconditionally positive. Concerning the equivalent internal energy, both terms are developed and lead to expressions proportional to

$$(M+1)\frac{(3\gamma-1)M + 3 - \gamma}{\gamma - 1} \qquad \text{if } M \geq 0 \tag{51a}$$

$$\frac{3-\gamma}{\gamma-1} \qquad \text{if } M \leq 0 \tag{51b}$$

which are both positive providing that $1 < \gamma < 3$.
In the subsonic range, $\forall\,\mathcal{U} \in \Omega_{\mathcal{U}}$, $F^+(\mathcal{U}) \in \bar{\Omega}_{\mathcal{U}}$ and condition (20a) is satisfied.

(2) Positivity of vector $\mathcal{W}_0$. As it was done for VL and VLH schemes, the positivity of vector $\mathcal{W}_0$ will lead to a condition on the time step which will guarantee the scheme positivity. $\mathcal{W}_0$ can be expressed as

$$\mathcal{W}_0 = \mathcal{U} - \frac{\chi^{loc}}{\lambda} F^\star(\mathcal{U}), \tag{52a}$$

where

$$F^\star(\mathcal{U}) = F^+(\mathcal{U}) - F^-(\mathcal{U}) = \frac{\rho a}{\gamma}\begin{pmatrix} (\gamma-1)M + 1 \\ aM[(\gamma-1)M + 2] \\ a^2\left[\frac{\gamma-1}{2}M^3 + \frac{3}{2}M^2 + \frac{1}{\gamma-1}\right] \end{pmatrix}. \tag{52b}$$

$\mathcal{W}_0$ can then be rewritten as

$$\mathcal{W}_0 = \rho\left(1 - \frac{\chi^{loc}}{\gamma}\frac{1+(\gamma-1)M}{1+M}\right)\begin{pmatrix} 1 \\ aM(1-\zeta) \\ a^2\left[K_E - \zeta\frac{1-M+\gamma M^2}{\gamma}\right] \end{pmatrix}, \tag{52c}$$

where

$$\zeta = \frac{\chi^{loc}}{\gamma(1+M) - \chi^{loc}[1+(\gamma-1)M]}.$$

Mass positivity requires

$$\chi^{loc} < \frac{\gamma(1+M)}{1+(\gamma-1)M}. \tag{53}$$

Under this condition, ζ is positive. The internal energy term can be developed and leads to a general condition similar to Eq. (29a) where

$$a(M) = M^2 \tag{54a}$$

$$b(M) = \frac{1-M}{\gamma} \tag{54b}$$

$$c(M) = -\frac{2}{\gamma(\gamma-1)}. \tag{54c}$$

The maximum value of ζ is computed from Eq. (29b). The positivity condition is then given by

$$\chi^{loc} < \chi^{SW}_{\max} = \frac{\gamma(M+1)\zeta_{\max}(M)}{1+[1+(\gamma-1)M]\zeta_{\max}(M)} \tag{55}$$

in which $\chi_{\max}$ can easily be computed and is plotted in Subsection 2.4.

(3) Computation of χ_{opt}. The framework is here the same as it is for VL and VLH schemes. A study of the function $\chi^{SW}_{\max}$ has been performed. The optimal CFL number is shown to be

$$\chi_{opt} = \min\left(1, \chi^{SW}_{\max}(0)\right). \tag{56}$$

$\chi^{SW}_{\max}(0)$ can easily be computed and gives 1. Hence, the CFL condition of the theorem follows.

The proof is completed.

2.4. *Review of Positivity Conditions*

Results and positivity conditions are summarized in Tables I and II. Local necessary and sufficient conditions are given. It should be pointed out that, by itself, the positivity of vector F^+ is only a necessary condition and does not ensure the scheme positivity. The positivity of vector $\mathcal{W}_0$ leads to a maximum time step which has then to be put into a CFL-like form $\chi^{loc} < \chi_{opt}$. This is the case for VL, VLH, and SW schemes since $\chi_{opt} = \inf_M(\chi_{\max})$ is not zero.

It can be easily verified that the internal energy positivity conditions (Table II) are more stringent than the mass positivity conditions (Table I). Therefore, it is the internal energy

TABLE I
Mass Positivity Conditions

		VL and VLH	SW
F^+	Supersonic	Unconditionally positive	
	Subsonic	Unconditionally positive	
$\mathcal{W}_0$	Supersonic	$\chi^{loc} < \dfrac{\lvert M\rvert+1}{\lvert M\rvert}$	
	Subsonic	$\chi^{loc} < \dfrac{2(\lvert M\rvert+1)}{1+M^2}$	$\chi^{loc} < \dfrac{\gamma(\lvert M\rvert+1)}{1+(\gamma-1)\lvert M\rvert}$

positivity condition which actually rules the scheme positivity. Moreover, it means that zero values cannot be reached simultaneously by density and internal energy. Since expressions of $\chi_{\max}^{VL}$, $\chi_{\max}^{VLH}$, and $\chi_{\max}^{SW}$ are intricate, they are not detailed but these coefficients can be easily computed as a function of the local Mach number following Eqs. (47), (55), and associated notations.

The smallest values of these conditions have been computed and lead to the optimal CFL condition χ_{opt} which ensures that the scheme is positively conservative in all configurations. These constants χ_{opt} are summarized in Table III and lead to an optimal CFL number of one for usual gases where $1 < \gamma < 2$.

Since necessary and sufficient conditions have been derived, it can be interesting to plot the local CFL conditions. For usual values of γ in the range $[1; 2]$, the greatest allowable time steps are obtained in decreasing order with the VL, VLH, and SW schemes.

$\chi_{\max}$ functions are plotted in Fig. 1 in the case of $\gamma = 1.4$. It shows that at $M = 1$ the three conditions join the condition derived in the fully upwind case. Moreover, both VL and VLH conditions are differentiable at $M = 1$. The SW scheme yields the most severe condition while the VL scheme allows a greater local CFL condition in the subsonic range.

All three curves merge in the supersonic range where the CFL condition implies that χ should decrease to 1 for high Mach numbers (Fig. 1). As a consequence, a CFL number of one *a fortiori* ensures positivity of the three schemes. Yet, higher CFL numbers can be used with VL and VLH schemes if the flow is expected not to exceed a given Mach number. For example, according to Fig. 1, a CFL number of 1.45 (for $\gamma = 1.4$) can be used in subsonic flows although it would not maintain positivity with the SW scheme. Note that this condition only ensures the scheme positivity, but not its stability. Using too high CFL numbers might produce oscillations even though the updated solution would still be an admissible state.

TABLE II
Internal Energy Positivity Conditions

		VL or VLH	SW
F^+	Supersonic	$M \geq \sqrt{\dfrac{\gamma-1}{2\gamma}}$	
	Subsonic	$\gamma \geq 1$	$1 \leq \gamma \leq 3$
$\mathcal{W}_0$	Supersonic	$\chi^{loc} < \dfrac{\lvert M\rvert+1}{\lvert M\rvert+\sqrt{\frac{\gamma-1}{2\gamma}}}$	
	Subsonic	$\chi^{loc} < \chi_{\max}^{VL/VLH}$	$\chi^{loc} < \chi_{\max}^{SW}$

TABLE III
Optimal CFL Number χ_{opt}

VL	VLH	SW
1	$\min\left(1, \frac{2}{\gamma}\right)$	1

3. ACCURACY VERSUS POSITIVITY

Most FVS schemes have proved to be robust in many flow configurations. Some of them have been proved to be positively conservative [12, 20]. Others have been analyzed in this paper. But none of them are able to exactly resolve contact discontinuities since it remains a non-vanishing dissipation which smears out an initial discontinuity of densities.

Van Leer [19] pointed out that preventing numerical diffusion of contact discontinuities may lead to a marginally stable or unstable behavior for slow flows. Nevertheless, he concluded that the question would need more work.

In the present study, the question of linear stability is not tackled. But the strength of Theorem 1, since both conditions are necessary, leads to the following theorem.

THEOREM 4. *If a FVS scheme exactly preserves stationary contact discontinuities, then it cannot be positively conservative.*

Remarks.

- This theorem explains why no FVS schemes have been built so far to simultaneously yield their famous robustness and the vanishing numerical dissipation on contact waves.

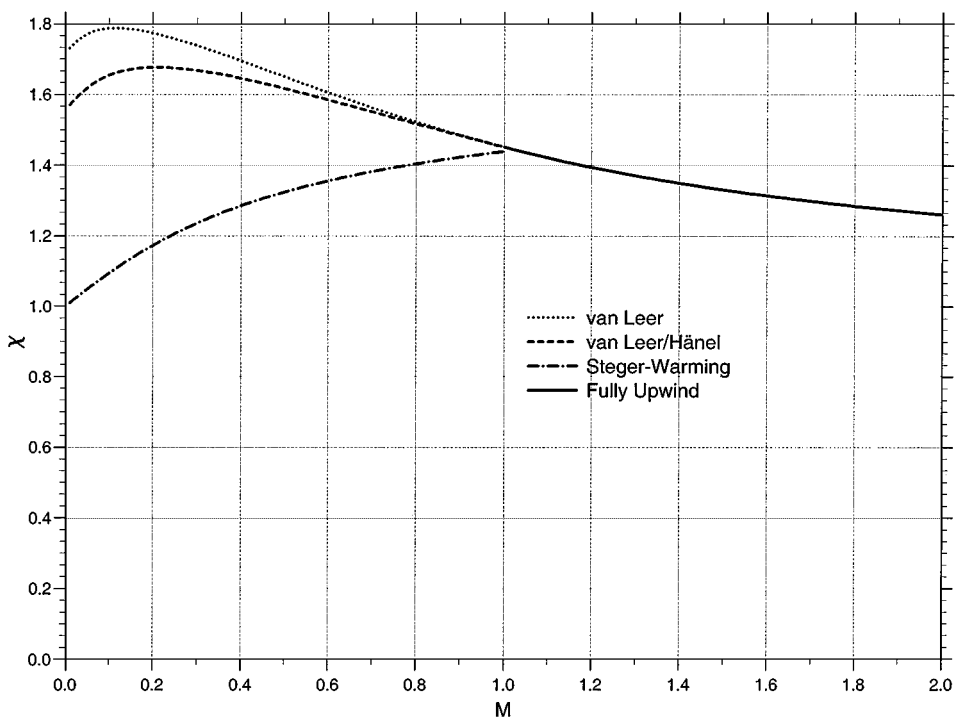

FIG. 1. Maximum CFL number χ^{loc} to ensure internal energy positivity, ($\gamma = 1.4$).

• FVS schemes are attractive because they are generally easy to implement, easy to make implicit, and lead to a low computational cost. However, the consequence of this theorem is that a scheme must include a hybrid technique with FDS schemes in order to satisfy both properties of robustness and accuracy.

Proof. Consider a FVS scheme given by its flux functions $F^{\pm}$ and assume it exactly preserves stationary contact discontinuities. Then, the interface flux between $\mathcal{U}_L = {}^T(\rho_L, 0, \frac{p}{\gamma-1})$ and $\mathcal{U}_R = {}^T(\rho_R, 0, \frac{p}{\gamma-1})$ must satisfy

$$F^+(\mathcal{U}_L) + F^-(\mathcal{U}_R) = \begin{pmatrix} 0 \\ p \\ 0 \end{pmatrix}. \tag{57}$$

Since ρ_L and ρ_R are independent variables, $F^+(\mathcal{U}_L)$ must be a function of only p. Hence, for all $\mathcal{U} = {}^T(\rho, 0, \frac{p}{\gamma-1})$,

$$F^+(\mathcal{U}) = \begin{pmatrix} f_1(p) \\ f_2(p) \\ f_3(p) \end{pmatrix}. \tag{58a}$$

Moreover, considering the symmetry property (17) and using $\bar{\bar{\mathcal{U}}} = \mathcal{U}$, one has $F^-(\mathcal{U}) = -\overline{F^+(\mathcal{U})}$. Then,

$$F^-(\mathcal{U}) = \begin{pmatrix} -f_1(p) \\ +f_2(p) \\ -f_3(p) \end{pmatrix}. \tag{58b}$$

Substituting expressions (58a) and (58b) in Eq. (57), one obtains $f_2(p) = p/2$. Moreover, $f_1(p)$ must be positive or null to satisfy the condition (20a) of positivity.

- If $f_1(p) = 0$, condition (20a) is not satisfied since $f_2(p)$ is not null.
- If $f_1(p) > 0$, then $\mathcal{W}_0 = \mathcal{U} - \frac{\chi^{loc}}{\lambda} F^\star(\mathcal{U})$ mass term may be expressed as

$$\rho - \frac{\chi^{loc}}{a} 2 f_1(p) = \rho - \sqrt{\rho}\left(2\chi^{loc}\frac{f_1(p)}{\sqrt{\gamma p}}\right). \tag{59}$$

Hence, for all functions $f_1(p)$ and for all $\chi^{loc} > 0$, one can find p and ρ such that expression (59) is negative.

Hence, if a FVS scheme has been designed to exactly preserve contact discontinuities, then it cannot satisfy both necessary conditions of Theorem 1.

The proof is completed.

4. CONCLUDING REMARKS

A general method to prove the positivity of FVS schemes has been detailed. It leads to two necessary and sufficient conditions on the flux vectors $F^{\pm}$.

It has been applied to standard FVS schemes, namely the Van Leer scheme, one of its variants, and Steger and Warming schemes Although these schemes have been known to

be robust, they are now proved to be positively conservative under a CFL condition of 1, for usual values of the specific heat ratio γ in the range [1; 2]. In particular, this shows that all these FVS schemes can be confidently applied to gas dynamics problems including real gas effects for which γ may range between 1.4 and 1.

Moreover, these conditions have been proved to be incompatible with the particular form of FVS schemes which would be able to exactly preserve stationary contact discontinuities. Hence, a robust FVS scheme cannot exactly compute contact discontinuities. In other words, an accurate and robust scheme must not be fully FVS. This drastically limits the capabilities of the class of FVS schemes.

APPENDIX A

LEMMA 2. *The set of admissible states $\Omega_{\mathcal{U}}$ and its closure $\bar{\Omega}_{\mathcal{U}}$ are convex cones, i.e.,*

$$\forall \mathcal{U}_1, \mathcal{U}_2 \in \Omega_{\mathcal{U}}, \forall \alpha_1, \alpha_2 > 0, \qquad \alpha_1 \mathcal{U}_1 + \alpha_2 \mathcal{U}_2 \in \Omega_{\mathcal{U}} \tag{60a}$$

$$\forall \mathcal{U}_1, \mathcal{U}_2 \in \bar{\Omega}_{\mathcal{U}}, \forall \alpha_1, \alpha_2 \geq 0, \qquad \alpha_1 \mathcal{U}_1 + \alpha_2 \mathcal{U}_2 \in \bar{\Omega}_{\mathcal{U}}. \tag{60b}$$

Proof. One can define an order relation denoted by $\succ$ which corresponds to $>$ for $\Omega_{\mathcal{U}}$ and $\geq$ for $\bar{\Omega}_{\mathcal{U}}$. Then, $\Omega_{\mathcal{U}}$ and $\bar{\Omega}_{\mathcal{U}}$ are defined by

$$\left\{ \mathcal{U} = {}^T(u_1, u_2, u_3) \,\middle|\, u_1 \succ 0, u_3 \succ 0 \text{ and } 2u_1u_3 - u_2^2 \succ 0 \right\}. \tag{61}$$

Let Ω be either $\Omega_{\mathcal{U}}$ or $\bar{\Omega}_{\mathcal{U}}$. The proof is completed in two steps

- For all $\mathcal{U} \in \Omega$, $\forall \alpha \in \mathbb{R}^+$, one obtains directly

$$\alpha u_1 \succ 0 \tag{62a}$$

$$\alpha u_3 \succ 0 \tag{62b}$$

$$2(\alpha u_1 \alpha u_3) - (\alpha u_2)^2 = \alpha^2 \left(2u_1u_3 - u_2^2\right) \succ 0. \tag{62c}$$

Then, $\alpha\mathcal{U} \in \Omega$. Hence, Ω is a cone.

- For all $\mathcal{U}, \mathcal{V} \in \Omega$, their components satisfy

$$u_1 \succ 0, \qquad u_3 \succ 0, \qquad 2u_1u_3 - u_2^2 \succ 0 \tag{63a}$$

$$v_1 \succ 0, \qquad v_3 \succ 0, \qquad 2v_1v_3 - v_2^2 \succ 0. \tag{63b}$$

Obviously,

$$u_1 + v_1 \succ 0, \qquad u_3 + v_3 \succ 0. \tag{64}$$

One has to prove the positivity of $\mathcal{U} + \mathcal{V}$ internal energy. If u_1 (resp. v_1) equals zero (only when belonging to $\bar{\Omega}_{\mathcal{U}}$), then u_2 (resp. v_2) equals zero and

$$2(u_1 + v_1)(u_3 + v_3) - (u_2 + v_2)^2 = \left(2v_1v_3 - v_2^2\right) + 2v_1u_3 \succ 0. \tag{65}$$

Otherwise (u_1 and $v_1 \neq 0$), one can develop

$$\begin{aligned} &2(u_1+v_1)(u_3+v_3)-(u_2+v_2)^2 \\ &\quad=\left(2u_1u_3-u_2^2\right)+\left(2v_1v_3-v_2^2\right)+2(u_1v_3+v_1u_3-u_2v_2) \\ &\quad\succ 2(u_1v_3+v_1u_3-u_2v_2) \end{aligned} \tag{66a}$$

and

$$\begin{aligned} &2(u_1v_3+v_1u_3-u_2v_2) \\ &\quad=\frac{u_1^2(2v_1v_3)+v_1^2(2u_1u_3)-2u_1v_1u_2v_2}{u_1v_1} \\ &\quad\succ\frac{u_1^2v_2^2+v_1^2u_2^2-2u_1v_1u_2v_2}{u_1v_1} \\ &\quad\succ\frac{(u_1v_2-v_1u_2)^2}{u_1v_1} \\ &\quad\succ 0. \end{aligned} \tag{66b}$$

Hence, $\mathcal{U}+\mathcal{V}\in\Omega$.

LEMMA 3.

$$\forall\mathcal{U}_1\in\Omega_{\mathcal{U}},\forall\mathcal{U}_2\in\bar{\Omega}_{\mathcal{U}},\forall\alpha_1>0,\forall\alpha_2\geq 0,\qquad \alpha_1\mathcal{U}_1+\alpha_2\mathcal{U}_2\in\Omega_{\mathcal{U}}. \tag{67}$$

Proof. The proof is similar to that of Lemma 2. $\mathcal{U}_1$ positivity yields strict inequalities which prove strict positivity of density and internal energy of $\mathcal{U}_1+\mathcal{U}_2$.

APPENDIX B. NOMENCLATURE

ρ	density	χ^{loc}	local CFL number
p	pressure	a	sound speed
M	Mach number	e	internal energy
H	total enthalpy	K_H	dimensionless coefficient H/a^2
E	total energy	K_E	dimensionless coefficient E/a^2
$\mathcal{U}$	state vector	$\mathbb{U}$	updated state vector
$\Omega_{\mathcal{U}}$	space of physical states	γ	ratio of specific heats
$\mathcal{F}$	physical flux vector	F	numerical flux vector

REFERENCES

1. S. M. Deshpande, *Kinetic Theory Based New Upwind Methods for Inviscid Compressible Flows*, AIAA Paper 86-0275, January 1986.
2. B. Dubroca, Positively conservative Roe's matrix for Euler equations, in *16th ICNMFD*, Lecture Notes in Physics (Springer-Verlag, New York/Berlin, 1998), p. 272.
3. B. Einfeldt, C. D. Munz, P. L. Roe, and B. Sjögreen, On Godunov-type methods near low densities, *J. Comput. Phys.* **92**, 273 (1991).
4. J. L. Estivalezes and P. Villedieu, High-order positivity-preserving kinetic schemes for the compressible Euler equations, *SIAM J. Numer. Anal.* **33**(5), 2050 (1996).

5. S. K. Godunov, A difference scheme for numerical computation of discontinuous solutions of hydrodynamics equations, *Math. Sb.* **47**(3), 271 (1959).
6. D. Hänel, R. Schwane, and G. Seider, *On the Accuracy of Upwind Schemes for the Solution of the Navier–Stokes Equations*, AIAA Paper 87-1105, 1987.
7. A. Harten, P. D. Lax, and B. Van Leer, On upstream differencing and Godunov-type schemes for hyperbolic conservation laws, *SIAM Rev.* **25**(1), 35 (1983).
8. B. Larrouturou, How to preserve the mass fractions positivity when computing compressible multi-component flows, *J. Comput. Phys.* **95**, 59 (1991).
9. T. Linde and P. L. Roe, *Robust Euler Codes*, AIAA Paper 97-2098, 1997.
10. T. Linde and P. L. Roe, On a mistaken notion of "proper upwinding," *J. Comput. Phys.* **142**, 611 (1998).
11. M. S. Liou, A sequel to AUSM : AUSM+, *J. Comput. Phys.* **129**, 364 (1996).
12. B. Perthame, Boltzmann type schemes for gas dynamics and the entropy property, *SIAM J. Numer. Anal* **27**(6), 1405 (1990).
13. B. Perthame and Y. Qiu, *A Variant of van Leer's Methods for Multidimensional Systems of Conservation Laws*, Technical Report 1562, INRIA, 1991.
14. B. Perthame and C. Shu, On positive preserving finite volume schemes for the compressible Euler equations, *Numer. Math.* **76**, 119 (1996).
15. D. I. Pullin, Direct simulation methods for compressible inviscid ideal gas flow, *J. Comput. Phys.* **34**, 231 (1980).
16. P. L. Roe, Approximate Riemann solvers, parameters vectors, and difference schemes, *J. Comput. Phys.* **43**, 357 (1981).
17. J. L. Steger and R. F. Warming, Flux vector splitting of the inviscid gasdynamics equations with application to finite-difference methods, *J. Comput. Phys.* **40**, 263 (1981).
18. B. van Leer, Flux vector splitting for the Euler equations, in *8th ICNMFD*, Lecture Notes in Physics (Springer–Verlag, New York/Berlin, 1982), Vol. 170, p. 505.
19. B. van Leer, *Flux Vector Splitting for the 1990's*, NASA CP-3078, 1991.
20. P. Villedieu and P. A. Mazet, Schémas cinétiques pour les equations d'Euler hors equilibre thermochimique, *Recherche Aérospatiale* **2**, 85 (1995).