JOURNAL OF COMPUTATIONAL PHYSICS 95, 59–84 (1991)

How to Preserve the Mass Fractions Positivity when Computing Compressible Multi-component Flows

B. LARROUTUROU

CERMICS, INRIA, Sophia-Antipolis, 06560 Valbonne, France

Received September 12, 1989; revised February 16, 1990

We are interested in the numerical investigation of the compressible flow of a gaseous mixture. Considering a hyperbolic system including the Euler equations for the mixture and a mass conservation equation for each species, we propose a new approximation scheme for the convective term of the species equations. This approximation relies on some properties of the exact solution of the Riemann problem for the multi-component system, and applies when an upwind Godunov-type scheme is used for the Euler equations. Its main interest lies in the fact that it preserves the positivity and monotonicity of the mass fractions of all species.

1. INTRODUCTION

The numerical simulation of flows of gaseous mixtures, and in particular of chemically reactive flows, has received increasing attention in the last years and is still the subject of numerous investigations. For some of these flows, such as transonic reactive flows (see, e.g., [7]), hypersonic flows (see, e.g., [4, 9]), or detonations (see, e.g., [19]), the hyperbolic aspects of the phenomenon play a major role. This is the origin of our interest in the following system of the "multi-component Euler equations," which contains the classical Euler equations written for the mixture and an additional "species equation":

$$\begin{aligned} \rho_t + (\rho u)_x &= 0, \\ (\rho u)_t + (\rho u^2 + p)_x &= 0, \\ E_t + [u(E+p)]_x &= 0, \\ (\rho Y)_t + (\rho u Y)_x &= 0 \end{aligned} \tag{1.1}$$

(see the definition of the notations below).

We are interested in the numerical solution of (1.1) using upwind difference schemes. Essentially two different strategies have been used in previous studies for this problem. Either one uses for the first three equations in (1.1) one of the numerous available schemes aimed at solving the Euler equations and one solves separately the species equation; or one considers the whole system (1.1) as a system

0021-9991/91 $3.00

of conservation laws and solves all equations in a coupled way by extending to (1.1) one of the above-mentioned "Euler schemes." Some details about these two existing approaches are presented in Section 2 below.

Our aim is to present a third approach for the solution of (1.1). This new approach is based on one property of the exact solution of the Riemann problem for (1.1), which is derived in Section 3. Besides its major computational simplicity, the proposed method has over the two existing approaches the advantage of preserving the maximum principle (and in particular the positivity) for the mass fractions of all species.

Let us now describe more precisely the setting of our work. We are interested in the description of the multi-dimensional flow of a compressible gaseous mixture of N species. But, for the sake of simplicity, we will consider in a first step the one-dimensional flow of a mixture of two gaseous species $\mathscr{S}_1$ and $\mathscr{S}_2$. In the absence of diffusive and reactive phenomena, this flow is described by system (1.1), where ρ is the density of the mixture, u is its velocity, p its pressure, E the total energy per unit volume of the mixture, and Y is the mass fraction of species $\mathscr{S}_1$. If one assumes that both species behave as perfect gases, then the pressure in (1.1) is given by the relation (see [22])

$$p = (\gamma - 1)(E - \tfrac{1}{2}\rho u^2), \tag{1.2}$$

where γ is the ratio of the specific heats *of the mixture*, given by

$$\gamma = \frac{YC_{v1}\gamma_1 + (1-Y)\,C_{v2}\gamma_2}{YC_{v1} + (1-Y)\,C_{v2}}, \tag{1.3}$$

(the subscripts 1 and 2 in this expression refer to species $\mathscr{S}_1$ and $\mathscr{S}_2$, respectively). As already said, the first three equations in (1.1) with (1.2) together form the classical Euler equations, written here for the mixture of both species. Note however that, if $\gamma_1 \neq \gamma_2$, the coefficient γ in (1.2) actually depends on Y, whereas it is constant in the classical (i.e., single-component) Euler equations. The fourth equation in (1.1) will be referred to as the species equation.

Remark 1. Using the first and fourth equation in (1.1), one can rewrite the species equation in nonconservative form as $Y_t + uY_x = 0$, which shows that the variable Y is purely convected by the flow. The latter nonconservative equation is the basis of the donor-cell approximation (2.1) below.

Remark 2. In fact the results presented in this paper apply to more general situations. We could equally well consider a mixture of N perfect gases: the last equation in (1.1) would then be replaced by $N-1$ equations $(\rho Y_k)_t + (\rho u Y_k)_x = 0$ for $1 \leqslant k \leqslant N-1$ (see Remarks 10, 13, and 17 below). We could also consider the case where the last equation in (1.1) describes the convection of different quantities which may or may not affect the value of γ (the variable Y could, for instance, represent a passive scalar, or the variables k and ε in the classical $k-\varepsilon$ turbulence

model [27]). Also, the numerical methods presented in this paper can be extended to compute the flow of a *real gas* mixture (see [9] for more details in this direction).

Let us now introduce the notations:

$$W=\begin{pmatrix}\rho\\ \rho u\\ E\\ \rho Y\end{pmatrix}=\begin{pmatrix}W^1\\ W^2\\ W^3\\ W^4\end{pmatrix},\qquad F=\begin{pmatrix}\rho u\\ \rho u^2+p\\ u(E+p)\\ \rho uY\end{pmatrix}=\begin{pmatrix}F^1\\ F^2\\ F^3\\ F^4\end{pmatrix}. \tag{1.4}$$

For the numerical solution of (1.1) (which can be written as $W_t+F_x=0$ in vector form), we will first consider explicit conservative schemes of the form

$$\frac{W_i^{n+1}-W_i^n}{\Delta t}+\frac{\phi_{i+1/2}^n-\phi_{i-1/2}^n}{\Delta x}=0, \tag{1.5}$$

where the numerical flux $\phi_{i+1/2}^n$ is evaluated using a "numerical flux function" Φ which defines the scheme under consideration:

$$\phi_{i+1/2}^n=\Phi(W_i^n, W_{i+1}^n). \tag{1.6}$$

The notations in (1.5) are classical: Δt and Δx represent the discretization steps in time and space, respectively, and i and n are the spatial and temporal subscripts. We will sometimes write

$$\phi_{i+1/2}^n=\begin{pmatrix}\phi_{i+1/2}^{1,n}\\ \phi_{i+1/2}^{2,n}\\ \phi_{i+1/2}^{3,n}\\ \phi_{i+1/2}^{4,n}\end{pmatrix}; \tag{1.6}$$

here, the superscript n denotes the time level, while the superscripts 1, 2, 3, 4 denote the different components of the numerical flux. When no confusion is possible, we will write $\phi_{i+1/2}$ instead of $\phi_{i+1/2}^n$ for the sake of simplicity.

The remainder of the paper is organized as follows. We recall some facts about the two existing numerical approaches in Section 2. The new positivity-preserving approximation of the species equation is then presented in Section 3, and some extensions are discussed in Section 4. Lastly, Section 5 is devoted to the presentation and comparison of some numerical results.

2. Two Existing Approaches

We begin by presenting in this section two existing approaches for the numerical solution of (1.1), and by discussing their advantages and drawbacks. We will then propose a third new approach in Section 3.

2.1. *The Uncoupled Approach*

A first and simple approach to the numerical solution of (1.1) has been used in, e.g., [17]. It consists in treating separately at each time step the Euler equations and the species equation in (1.1). For Euler equations, one uses one of the classical (single-component) "Euler schemes" (such as the approximate Riemann solver of Osher [26], or of Roe [28], or the flux vector splitting of Van Leer [32]), with a "frozen γ." This means that, for the evaluation of the first three components of $\phi_{i+1/2}$ at time t^n, one uses the flux of one of the above Euler schemes computed using a frozen value $\gamma_{i+1/2}=\gamma((Y_i^n+Y_{i+1}^n)/2)$. Beside this, one uses an upwind approximation of the donor-cell type for the species equation; defining $u_{i+1/2}^n=\frac{1}{2}(u_i^n+u_{i+1}^n)$, one evaluates the fourth component $\phi_{i+1/2}^4$ of $\phi_{i+1/2}$ by

$$\phi_{i+1/2}^4 = u_{i+1/2}^n \times \begin{cases} (\rho Y)_i^n & \text{if } u_{i+1/2}^n > 0, \\ (\rho Y)_{i+1}^n & \text{if } u_{i+1/2}^n < 0. \end{cases} \tag{2.1}$$

In the sequel, this uncoupled approximation will be referred to as "approach (A)."

2.2. *The Fully Coupled Approach*

A second type of approximation for system (1.1) has been used in [1, 2, 4, 11, 16], and is discussed in detail in [22]. It consists in seeing (1.1) as a whole, and in extending to this bigger system of conservation laws the upwind schemes (of Osher, Roe, Van Leer) which have been developed for the single-component case. This approach, where the Euler equations and the species equation are not separated, will be referred to in the sequel as "approach (B)."

Referring to [22] for the details, we just recall here the basic facts about system (1.1). The components of the flux vector F can be expressed as functions of the conservative variables W^l $(1 \leqslant l \leqslant 4)$, and the system $W_t + F(W)_x = 0$ is hyperbolic; the Jacobian matrix $A(W)=(\partial F^l/\partial W^m)$ has four real eigenvalues $\lambda_1 = u-c$, $\lambda_2 = u$, $\lambda_3 = u$, and $\lambda_4 = u+c$, where $c=\sqrt{\gamma p/\rho}$ is the sound speed, and four real eigenvectors.

It is then possible to extend to (1.1) the classical upwind schemes developed for the Euler equations. We now briefly describe these extended schemes, referring to [22] and the references therein for more details.

2.2.1. *The Multi-component Van Leer Scheme*

The extension of the continuously differentiable flux splitting of Van Leer [32] to multi-component flows was first derived in [22]. As in the single-component case, the numerical flux function has the form:

$$\Phi(W_L, W_R) = F_+(W_L) + F_-(W_R); \tag{2.2}$$

the split fluxes F_+ and F_- are defined by the identity $F(W) = F_+(W) + F_-(W)$ and the following expressions:

* if $u \geqslant c = \sqrt{\gamma p/\rho}$, $F_+(W) = F(W)$;
* if $-c \leqslant u \leqslant c$,

$$F_+(W) = \begin{pmatrix} F_+^1 \\ F_+^2 \\ F_+^3 \\ F_+^4 \end{pmatrix} = \begin{pmatrix} \dfrac{\rho}{4c}(u+c)^2 \\ F_+^1\left(u - \dfrac{u-2c}{\gamma}\right) \\ \dfrac{\gamma^2}{2(\gamma^2-1)}\dfrac{(F_+^2)^2}{F_+^1} \\ YF_+^1 \end{pmatrix} \tag{2.3}$$

(the first three components of F_+ have the same expression as in the single-component case, but with the nonconstant coefficient γ given by (1.3));

* if $u \leqslant -c$, $F_+(W) = 0$.

The following result, which says that this flux splitting can be used to define a stable conservative scheme, is proved in [22]:

PROPOSITION 1. *If the specific heat ratio γ_k of each species in the mixture satisfies the inequality $1 < \gamma_k < 3$, then all eigenvalues of the Jacobian matrix DF_+/DW (resp. DF_-/DW) are real and positive (resp. negative).*

2.2.2. *The Multi-component Roe Scheme*

The extension to mixtures of the conservative upwind scheme based on the approximate Riemann solver of Roe [28] has been derived in [1, 11]. The numerical flux function of the extended Roe scheme has the same expression as in the single-component case,

$$\Phi(W_L, W_R) = \frac{F(W_L) + F(W_R)}{2} + \frac{1}{2}|\tilde{A}|(W_L - W_R), \tag{2.4}$$

where $\tilde{A} = \tilde{A}(W_L, W_R)$ is a diagonalisable matrix which satisfies the property:

$$F(W_L) - F(W_R) = \tilde{A}(W_L - W_R). \tag{2.5}$$

When γ is constant, that is, when $\gamma_1 = \gamma_2$ in (1.3), then the matrix $\tilde{A}(W_L, W_R)$ is equal to the Jacobian matrix $A(\tilde{W})$, where the state $\tilde{W}$ is "Roe's average" of W_L and W_R. When $\gamma_1 \neq \gamma_2$, the matrix $A(\tilde{W})$ does no longer satisfy (2.5), and the matrix $\tilde{A}$ has to be chosen close to but different from $A(\tilde{W})$ (see [1, 11, 22] for the details).

2.2.3. *The Multi-component Osher Scheme*

The multi-component Osher scheme has been derived in [2]. As in the single-component case, the numerical flux function of the extended Osher scheme has the expression

$$\Phi(W_L, W_R) = \frac{F(W_L) + F(W_R)}{2} + \int_{W_L}^{W_R} |A(W)|\, dW, \tag{2.6}$$

where the integration is carried out along a path connecting W_L to W_R in the state-space. The integration path is piecewise parallel to the right eigenvectors of the flux Jacobian matrix A, and the evaluation of the integral in (2.6) requires the knowledge of the Riemann invariants associated with each eigenvector. As in the single-component case, the evaluation of this integral relies on the determination of the pressure on the middle part of this integration path (see [2] for the details).

2.3. *Discussion*

The main advantage of approach (A) is its simplicity: even in two or three space dimensions, following this approach to transform a single-component Euler code into a multi-component code requires a very small programming and computational effort. This approach also has the advantage of preserving the positivity of the mass fraction Y (the proof of this fact is similar to the one of Lemma 3 below). But this approach also has important drawbacks, which will be discussed in more detail below.

Compared to approach (A), all schemes of type (B), except the extended Van Leer scheme, are substantially more expensive from the computational point of view. This is especially true for the extended Roe or Steger and Warming schemes: 4×4 Jacobian matrices have to be handled (that is, $N+2 \times N+2$ matrices if one considers a mixture of N species), while only 3×3 matrices are involved in the corresponding schemes of type (A). Nevertheless, approach (B), which is more satisfactory from a theoretical point of view, actually gives better results than approach (A): the accuracy of the results is greatly improved (see Section 5 below).

However, we need to make precise here that the "discrete maximum principle" for the mass fraction does not hold with any of the two approaches (that is, the inequalities $0 \leqslant Y \leqslant 1$ are not preserved by these schemes), except for the multi-component Van Leer scheme (see Remark 9 below). But in practice, the mass fraction values obtained with approach (B) remain much closer to the interval [0, 1] than those obtained with approach (A) (see Section 5 and [6]).

Intuitively, one can feel the origin of the inaccuracy of the computed mass fraction values in approach (A) where the mass fraction Y is evaluated as the ratio of two conservative variables,

$$Y = \frac{W^4}{W^1} = \frac{\rho Y}{\rho}; \tag{2.11}$$

in approach (A), these two variables W^1 and W^4 are advanced in time using two different approximations (more precisely, two different upwinding techniques). On the opposite, a single global and coupled approach is used for all conservative variables W^l in the schemes of type (B).

This remark leads us to propose a third approach, which we present in the next section.

3. A Third Approach

In this section, we propose a third approach in order to improve the multi-component Roe and Osher schemes which do not preserve the maximum principle for the mass fraction.

Let us first recall that, in the so-called "Godunov-type" schemes based on approximate Riemann solvers (see [18]), the numerical flux $\phi^n_{i+1/2}$ is seen as an approximation of the flux of the exact solution of the Riemann problem constructed with the neighbouring states W^n_i and W^n_{i+1}. This is why we now examine the Riemann problem for system (1.1), before we present the new approach (C).

3.1. *The Multi-component Riemann Problem*

Two states W_L and W_R being given, we consider the Riemann problem:

$$\begin{aligned} &W_t + F(W)_x = 0 \qquad \text{for} \quad x \in \mathbb{R},\ t \geqslant 0, \\ &W(x,0) = \begin{cases} W_L & \text{if} \quad x < 0, \\ W_R & \text{if} \quad x > 0. \end{cases} \end{aligned} \tag{3.1}$$

This problem is solved in [1, 2, 22]. Its exact solution $W^{\mathscr{R}}(x, t)$ is presented on Fig. 1. It is of course self-similar (i.e., $W^{\mathscr{R}}(x, t)$ only depends on the ratio x/t), and consists, as in the single-component case, of four constant states $W_{(1)}$, $W_{(2)}$, $W_{(3)}$, $W_{(4)}$ separated by shocks, rarefaction waves, or a contact discontinuity. More precisely, as shown on Fig. 1, $W_{(1)} = W_L$ and $W_{(2)}$ are separated by a wave associated with the first characteristic field, that is, with the eigenvalue $\lambda_1 = u - c$, either a 1-shock or a 1-rarefaction wave; $W_{(2)}$ and $W_{(3)}$ are separated by a contact discontinuity (associated with the eigenvalue u); $W_{(3)}$ and $W_{(4)} = W_R$ are separated by a 4-wave, associated with $\lambda_4 = u + c$. Also, the pressure p and the velocity u are continuous across the contact discontinuity. Last but not least, the mass fraction Y remains constant across the 1-wave and the 4-wave (whatever these waves are, shocks or rarefactions) and only varies across the contact discontinuity.

For $\sigma \in \mathbb{R}$ and $t \geqslant 0$, we will denote $\mathscr{W}(\sigma; W_L, W_R)$ (or sometimes simply $\mathscr{W}(\sigma)$) the value of $W^{\mathscr{R}}(\sigma t, t)$, which is independent of t; and we will call $u^*(W_L, W_R)$ (or simply u^*) the speed of the contact discontinuity.

As we could expect, it appears from Fig. 1 that the maximum principle for the

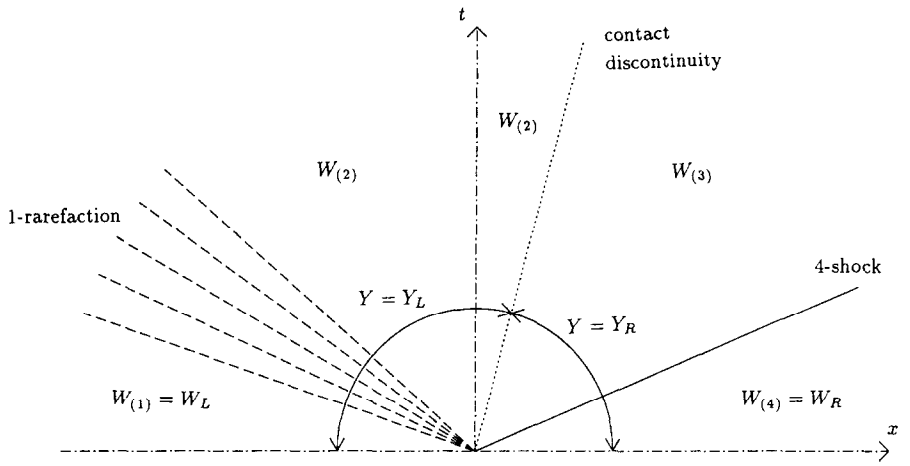

FIG. 1. The solution of the multi-component Riemann problem (3.1).

mass fraction is preserved for the exact solution of the Riemann problem: for any σ, we have $\min(Y_L, Y_R) \leqslant Y(\mathcal{W}(\sigma)) \leqslant \max(Y_L, Y_R)$. As a consequence, any numerical method where the evaluation of the numerical flux $\phi_{i+1/2}$ is evaluated using an *exact* Riemann solver (as in the original Godunov method [15], where $\phi^n_{i+1/2} = F[\mathcal{W}(0; W^n_i, W^n_{i+1})]$) will also satisfy the maximum principle for the mass fraction.

But, as we have said above, such is not the case for the numerical schemes based on the multi-component *approximate* Riemann solvers of Roe and Osher. This is exactly the essential point of this paper: we state in the next lemma the basic property that the exact solution of the Riemann problem satisfies, and that the approximate multi-component Riemann solvers should also satisfy in order to preserve the maximum principle. This property is the following:

LEMMA 1. *For any states W_L and W_R, the following equality holds:*

$$F^4[\mathcal{W}(0)] = F^1[\mathcal{W}(0)] \times \begin{cases} Y_L & \text{if } F^1[\mathcal{W}(0)] > 0, \\ Y_R & \text{if } F^1[\mathcal{W}(0)] < 0. \end{cases} \tag{3.2}$$

The proof begins with another lemma:

LEMMA 2. *For any states W_L and W_R, the quantities $F^1(\mathcal{W}(0; W_L, W_R))$ and $u^*(W_L, W_R)$ have the same sign.*

Proof of Lemma 2. It consists in considering the structure of the exact solution $\mathcal{W}$ of the Riemann problem and in using the classical properties of the rarefaction and shock waves. Let us assume, without loss of generality, that $u^* > 0$; we are going to show that $F^1(\mathcal{W}(0)) > 0$.

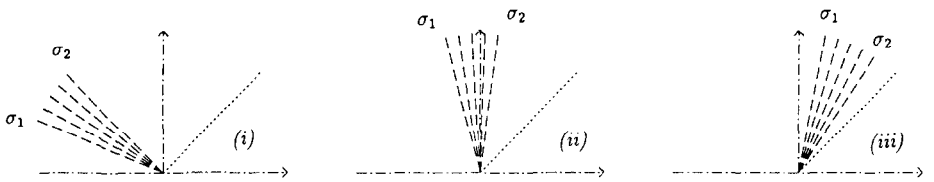

FIG. 2. The three different configurations for the 1-rarefaction.

(a) Consider first the situation where the 1-wave is a rarefaction wave, located in the region $\sigma_1 \leqslant x/t \leqslant \sigma_2$, with $\sigma_2 < u^*$. Since this wave is associated with the first eigenvalue $\lambda_1 = u - c$, we have (see [23])

$$u(\mathscr{W}(\sigma)) - c(\mathscr{W}(\sigma)) = \sigma, \qquad \forall \sigma \in [\sigma_1, \sigma_2]. \tag{3.3}$$

We then have to consider three different cases (see Fig. 2):

(i) If $\sigma_2 < 0$, then $\mathscr{W}(0) = W_{(2)}$, $u(\mathscr{W}(0)) = u^* > 0$; whence $F^1(\mathscr{W}(0)) > 0$.

(ii) If $\sigma_1 \leqslant 0 \leqslant \sigma_2$, applying (3.3) for $\sigma = 0$, we see that $u(\mathscr{W}(0)) > 0$; whence $F^1(\mathscr{W}(0)) > 0$.

(iii) If $\sigma_1 > 0$, then $\mathscr{W}(0) = \mathscr{W}(\sigma_1) = W_L$, and (3.3) yields $u(W_L) - c(W_L) = \sigma_1 > 0$; whence $u(\mathscr{W}(0)) = u(W_L) > 0$, and again $F^1(\mathscr{W}(0)) > 0$.

(b) Consider now the situation where the 1-wave is a shock wave propagating with the speed $\sigma_1 < u^*$. We again consider three different cases (see Fig. 3):

(i) If $\sigma_1 < 0$, then $u(\mathscr{W}(0)) = u(W_{(2)}) = u^* > 0$, and $F^1(\mathscr{W}(0)) > 0$.

(ii) If $\sigma_1 = 0$, the Rankine–Hugoniot relations for the steady 1-shock imply (see [23])

$$\rho u(W_L) = \rho u(W_{(2)}) = \rho(W_{(2)})\, u^* > 0, \tag{3.4}$$

and $F^1(\mathscr{W}(0))$ is well defined and positive since $F^1(W_L) = F^1(W_{(2)}) > 0$.

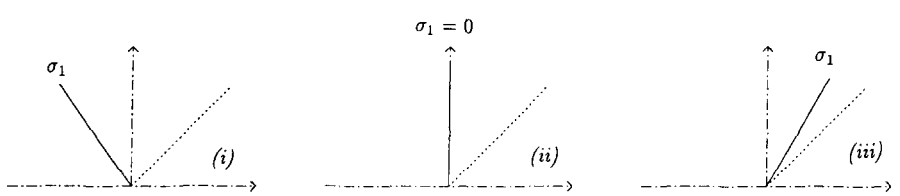

FIG. 3. The three different configurations for the 1-shock.

(iii) If $\sigma_1 > 0$, then $\mathscr{W}(0) = W_L$, and the Lax conditions for the 1-shock imply (see [23]):

$$u(W_L) - c(W_L) > \sigma_1 > 0. \tag{3.5}$$

This yields $u(W_L) > 0$, whence $F^1(\mathscr{W}(0)) > 0$. ∎

Proof of Lemma 1. For any $\sigma \in \mathbb{R}$, we have from Fig. 1,

$$Y(\mathscr{W}(\sigma)) = \begin{cases} Y_L & \text{if } \sigma < u^*, \\ Y_R & \text{if } \sigma > u^*. \end{cases} \tag{3.6}$$

Using the definitions of $F^1 = \rho u$ and $F^4 = \rho u Y$, we can write

$$F^4[\mathscr{W}(0)] = F^1[\mathscr{W}(0)] \times \begin{cases} Y_L & \text{if } u^* > 0, \\ Y_R & \text{if } u^* < 0, \end{cases} \tag{3.7}$$

and (3.2) follows from Lemma 2. ∎

Remark 3. It can be noticed that (3.2) and (3.7) remain valid if $F^1(\mathscr{W}(0)) = 0$ or if $u^* = 0$, since we then have $F^4(\mathscr{W}(0)) = F^1(\mathscr{W}(0)) = 0$ (we recall that, by definition of the contact discontinuity, u^* is also the velocity of the two states separated by this discontinuity: $u(W_{(2)}) = u(W_{(3)}) = u^*$).

Remark 4. We have assumed in the proof of Lemma 1 that a contact discontinuity actually exists in the structure of the exact solution $W^{\mathscr{R}}$. If this is not case, then $Y(x, t)$ is necessarily constant, $Y(x, t) \equiv Y_L = Y_R$, and (3.2) holds again.

3.2. *The Approximate Multi-component Riemann Flux*

We can now present the third approach (C), which is based on property (3.2).

Keeping in mind that, in the "Godunov-type" schemes, the numerical flux $\phi^n_{i+1/2}$ approximates the flux $F[\mathscr{W}(0; W^n_i, W^n_{i+1})]$, we can define a third approach (C) as follows for the Roe and Osher schemes: The first three components $\phi^1_{i+1/2}$, $\phi^2_{i+1/2}$, $\phi^3_{i+1/2}$ of the flux are evaluated as in approach (B), but the fourth component of the flux is defined by the next relation, which mimics (3.2):

$$\phi^4_{i+1/2} = \phi^1_{i+1/2} \times \begin{cases} Y^n_i & \text{if } \phi^1_{i+1/2} > 0, \\ Y^n_{i+1} & \text{if } \phi^1_{i+1/2} < 0. \end{cases} \tag{3.8}$$

Therefore, this approach (C) defines modified multi-component approximate Riemann solvers, which satisfy the property (3.2) of the exact Riemann solver. In comparison with approach (B), only the species mass fluxes are modified, so that approach (C) uses the same discrete mass fluxes for the species equation and for the continuity equation.

Since the approximations used for the variables W^1 and W^4 are now well

coupled, we may expect better result for the mass fraction $Y = W^4/W^1$. This is the object of the next lemma and of the following remarks:

LEMMA 3. *Under the CFL-like condition,*

$$\frac{\Delta t}{\Delta x}\left[\frac{\max(\phi^1_{i+1/2}, 0)}{\rho^n_i} - \frac{\min(\phi^1_{i-1/2}, 0)}{\rho^n_i}\right] \leqslant 1 \qquad \forall i, \tag{3.9}$$

the schemes of type (C) *preserve the maximum principle for the mass fraction Y: for all i and* $n \geqslant 0$,

$$\min_j Y^0_j \leqslant Y^n_i \leqslant \max_j Y^0_j. \tag{3.10}$$

Proof. Setting $\lambda = \Delta t/\Delta x$, we have

$$\rho^{n+1}_i = \rho^n_i - \lambda(\phi^1_{i+1/2} - \phi^1_{i-1/2}), \tag{3.11}$$

$$\rho^{n+1}_i Y^{n+1}_i = \rho^n_i Y^n_i - \lambda(\phi^4_{i+1/2} - \phi^4_{i-1/2}). \tag{3.12}$$

Let us temporarily assume that $\phi^1_{i+1/2} > 0$, $\phi^1_{i-1/2} > 0$. Then, (3.8), (3.11), and (3.12) imply

$$Y^{n+1}_i = \frac{Y^n_i(\rho^n_i - \lambda\phi^1_{i+1/2}) + Y^n_{i-1}(\lambda\phi^1_{i-1/2})}{(\rho^n_i - \lambda\phi^1_{i+1/2}) + (\lambda\phi^1_{i-1/2})}; \tag{3.13}$$

if (3.9) holds, (3.13) shows that Y^{n+1}_i is a convex linear combination of Y^n_i and Y^n_{i-1}.

It is easy to check that the same conclusion also holds with different hypotheses on the signs of $\phi^1_{i+1/2}$ and $\phi^1_{i-1/2}$, which concludes the proof. ∎

Remark 5. It is implicitly assumed in the above proof that ρ^{n+1}_i does not vanish. Therefore, rigorously speaking, the property (3.10) only holds if the "Euler scheme" used to evaluate ϕ^1, ϕ^2, ϕ^3 preserves the positivity of the density (this is, of course, the least one may ask to this "Euler scheme"!).

Remark 6. One may think at first glance that the preceding proof does not use any property of the approximate mass flux ϕ^1. This is not really the case. Indeed, it is easy to see that the CFL-like condition (3.9) which ensures the discrete maximum principle for the mass fraction may well be impossible to fulfill in practice if one uses a centered approximation $\phi^1_{i+1/2} = \frac{1}{2}[(\rho u)_i + (\rho u)_{i+1}]$ (take, for instance, $u_j \equiv \hat{u} > 0$ for all j, $\rho^n_i = \varepsilon > 0$, $\rho^n_{i+1} = 1$: (3.9) then imposes that $\hat{u}\,\Delta t/\Delta x \leqslant 2\varepsilon/(1+\varepsilon)$!). On the other hand, several numerical experiments have shown that the condition (3.9) is not more restrictive than the usual CFL condition when an upwind scheme is used to evaluate the first three components of $\phi_{i+1/2}$ (see [6]).

Remark 7. The proof of Lemma 3 shows that the definition (3.8) of $\phi^4_{i+1/2}$ is *exactly what is needed* in order to ensure the maximum principle for the mass

fraction. We repeat that schemes of type (A) or (B) do not in general satisfy (3.10) (see, however, Remark 9 below).

Remark 8. Arguing as in the proof of Lemma 3, one can show that the schemes of type (C) preserve the monotonicity of the mass fractions: if $Y_j^0 \leqslant Y_{j+1}^0$ for all j, then $Y_j^n \leqslant Y_{j+1}^n$ for all j and $n \geqslant 0$. More generally, these schemes are TVD (total variation diminishing) for the mass fraction Y.

Remark 9. In principle, method (C) can be used in conjunction with any conservative Euler scheme: one can use any of the classical schemes for the Euler equations, and then discretize the species equation using (3.8). But, from the very origin of (3.8), approach (C) is designed for schemes based on approximate Riemann solvers. In particular, there is no interest in using approach (C) with the Van Leer scheme. Indeed, we have seen that the multi-component Van Leer scheme gives the following expressions for the first and fourth components of the numerical flux:

$$\phi_{i+1/2}^1 = F_+^1(W_i^n) + F_-^1(W_{i+1}^n), \tag{3.14}$$

$$\phi_{i+1/2}^4 = Y_i F_+^1(W_i^n) + Y_{i+1} F_-^1(W_{i+1}^n). \tag{3.15}$$

Thus, the Van Leer scheme of type (B) already uses the same discrete mass fluxes $F_\pm^1$ for the continuity equation and the species equation, as the Roe and Osher schemes of type (C). Moreover, since $F_+^1(W) \geqslant 0$ and $F_-^1(W) \leqslant 0$ for any W, it is easy to see that the proof of Lemma 3 applies: the Van Leer scheme of type (B) also preserves the maximum principle for the mass fraction.

Remark 10. Consider the flow of a mixture of N species, with $N \geqslant 2$. Then, taking $W = (\rho, \rho u, E, \rho Y_1, ..., \rho Y_{N-1})$, one will preserve the inequalities $0 \leqslant Y_k \leqslant 1$ for $1 \leqslant k \leqslant N-1$ if approach (C) is used for the $N-1$ species equations. Moreover, it is easy to see that the last mass fraction Y_N, evaluated as $(Y_N)_i^n = 1 - \sum_{k=1}^{N-1} (Y_k)_i^n$ will also remain in the interval $[0, 1]$.

Remark 11. Consider now the case where a diffusive term and a consumption source term appear in the right-hand side of the species equations, which then has the form

$$(\rho Y)_t + (\rho Y u)_x = \mu Y_{xx} - K \rho Y, \tag{3.16}$$

(with $K > 0$), and assume that a fully explicit scheme is used to discretize this equation (see Remark 18 below for an implicit treatment of (3.16)):

$$\frac{(\rho Y)_i^{n+1} - (\rho Y)_i^n}{\Delta t} + \frac{\phi_{i+1/2}^4 - \phi_{i-1/2}^4}{\Delta x} = \mu \frac{Y_{i+1}^n - 2Y_i^n + Y_{i-1}^n}{\Delta x^2} - K \rho_i^n Y_i^n. \tag{3.17}$$

It is then easy to see that the maximum principle for the mass fraction is again preserved under the following stability restriction on the time step

$$\frac{\Delta t}{\Delta x}\left[\frac{\max(\phi_{i+1/2}^1, 0)}{\rho_i^n} - \frac{\min(\phi_{i-1/2}^1, 0)}{\rho_i^n}\right] + \frac{2}{\rho_i^n}\frac{\mu \Delta t}{\Delta x^2} + K \Delta t \leqslant 1 \qquad \forall i, \tag{3.18}$$

if the convective term is approximated using approach (C). This maximum principle now takes the form

$$\text{if} \quad \min_j Y_j^0 \geqslant 0, \quad \text{then } 0 \leqslant Y_i^n \leqslant \max_j Y_j^0 \qquad \forall i, \forall n. \tag{3.19}$$

4. Extensions

Before discussing in detail the advantages and drawbacks of approach (C) and comparing it to approaches (A) and (B) in Section 5, we now present how these types of schemes can be extended to second-order accuracy, to implicit time-stepping, or to multi-dimensional flows while keeping the basic property (3.10).

4.1. *Second-Order Accuracy*

Let us therefore consider second-order extensions of the preceding schemes. More precisely, we will first consider schemes which are second-order accurate in space but remain first-order accurate in time. Starting from the previous first-order accurate schemes, the second-order spatial accuracy is obtained by using piecewise linear variables instead of piecewise constant variables, following the "MUSCL" approach of Van Leer [31]. This method involves three steps:

(a) At each time step, starting from the values W_i^n, one first evaluates slopes s_i^n for all variables which are chosen to be piecewise linear. Several choices are possible at this stage (for instance, one can choose either the conservative variables ρ, ρu, E, ρY or the "physical variables" ρ, u, p, Y to vary linearly in each computational cell); here, we take Y (and not ρY) as a piecewise linear variable.

(b) Slope limiters are then used in order to avoid the creation of new extrema; here again different strategies exist to evaluate the limited slopes. We assume here that the slopes are constrained such that, taking the variable Y as an example, we have

$$\min_{|j-i| \leqslant 1} Y_j^n \leqslant Y_i^n \pm \frac{\Delta x}{2} s_i^n \leqslant \max_{|j-i| \leqslant 1} Y_j^n. \tag{4.1}$$

(c) The limited slopes are then used to evaluate cell-interface values $W_{i+1/2,\pm}^n$ (one sets $Y_{i+1/2,-}^n = Y_i^n + (\Delta x/2)\, s_i^n$, $Y_{i-1/2,+}^n = Y_i^n - (\Delta x/2)\, s_i^n$), and the solution is advanced in time according to relation (1.5), where we now take

$$\phi_{i+1/2} = \Phi(W_{i+1/2,-}^n, W_{i+1/2,+}^n). \tag{4.2}$$

This construction (a)–(c) can therefore be applied to any of the schemes presented in the preceding sections; in particular, we can construct a limited second-order accurate scheme of type (C) if in (4.2) we use a numerical flux function of type (C), that is with the fourth component ϕ^4 of the flux evaluated using (3.8). Then, we have the following:

LEMMA 4. *Under the more restrictive CFL-like condition,*

$$\frac{\Delta t}{\Delta x}\left[\frac{\max(\phi^1_{i+1/2}, 0)}{\rho^n_i} - \frac{\min(\phi^1_{i-1/2}, 0)}{\rho^n_i}\right] \leqslant \frac{1}{2} \qquad \forall i, \tag{4.3}$$

the spatially second-order accurate schemes of type (C) *constructed above preserve the maximum principle for the mass fraction* Y: *for all* i *and* $n \geqslant 0$, *we have*

$$\min_j Y^0_j \leqslant Y^n_i \leqslant \max_j Y^0_j. \tag{4.4}$$

Proof. We will assume that $\phi^1_{i+1/2} > 0$, $\phi^1_{i-1/2} > 0$ (the proof is very similar in the other cases). Using relations (3.11)–(3.12), we have

$$\begin{aligned} Y^{n+1}_i &= \frac{\rho^n_i Y^n_i - \lambda\phi^1_{i+1/2} Y^n_{i+1/2,-} + \lambda\phi^1_{i-1/2} Y^n_{i-1/2,-}}{\rho^n_i - \lambda\phi^1_{i+1/2} + \lambda\phi^1_{i-1/2}} \\ &= \frac{(\rho^n_i - \lambda\phi^1_{i+1/2})\left(\dfrac{\rho^n_i Y^n_i - \lambda\phi^1_{i+1/2} Y^n_{i+1/2,-}}{\rho^n_i - \lambda\phi^1_{i+1/2}}\right) + \lambda\phi^1_{i-1/2} Y^n_{i-1/2,-}}{(\rho^n_i - \lambda\phi^1_{i+1/2}) + \lambda\phi^1_{i-1/2}}; \end{aligned} \tag{4.5}$$

Let a and b be two real numbers such that $a \leqslant Y^n_j \leqslant b$ for all j. Since $Y^n_{i-1/2,-} \in [a, b]$ from (4.1), it suffices to check that

$$\hat{Y} = \frac{\rho^n_i Y^n_i - \lambda\phi^1_{i+1/2} Y^n_{i+1/2,-}}{\rho^n_i - \lambda\phi^1_{i+1/2}} \in [a, b]. \tag{4.6}$$

Using the definition of $Y^n_{1+1/2,-}$ and setting $X = \lambda(\phi^1_{i+1/2}/\rho^n_i)$, we obtain

$$\hat{Y} = Y^n_i - \frac{X}{1-X}\frac{\Delta x}{2} s^n_i, \tag{4.7}$$

and (4.6) holds from (4.1), since $0 \leqslant X \leqslant \frac{1}{2}$ from (4.3). ∎

Remark 12. The preceding proof clearly shows that it is preferable to use limited slopes for the variable Y and not for ρY when the approach (C) is used to discretize the species equation. Nevertheless, this strategy is almost but not completely sufficient to guarantee the preservation of the maximum principle for the mass fractions when a fully second-order accurate scheme (in both space and time) is used. This question is discussed in the Appendix.

Remark 13. In the case of a mixture of two species $\mathscr{S}_1$ and $\mathscr{S}_2$, the preceding proof shows that the mass fractions of both species will remain in the interval $[0, 1]$ when the above spatially second-order accurate method is used; since $Y_1 = Y \in [0, 1]$, then $Y_2 = 1 - Y \in [0, 1]$. But additional difficulties arise when one considers mixtures of N species with $N > 2$. Applying the above method to species k, we get $(Y_k)^n_i \in [0, 1]$ for all k, $1 \leqslant k \leqslant N-1$; but it is no more clear whether

$(Y_N)_i^n = 1 - \sum_{k=1}^{N-1} (Y_k)_i^n$ also is in the interval $[0, 1]$ (this comes from the fact that the limiters are by their very essence *nonlinear*). A way to remedy this difficulty is the following: instead of treating separately the Nth species, one could treat all species in the same way, setting $(Y_N)_i^n = (Y_N)_i^n + (\Delta x/2)(s_N)_i^n$ and applying the scheme of type (C) for all k, $1 \leqslant k \leqslant N$ (and not $N-1$). The maximum principle will be preserved if the slopes are constrained such that

$$\min_{|j-i| \leqslant 1} (Y_k)_j^n - (Y_k)_i^n \leqslant \frac{\Delta x}{2} |(s_k)_i^n| \leqslant \max_{|j-i| \leqslant 1} (Y_k)_j^n - (Y_k)_i^n \tag{4.8}$$

(which is equivalent to (4.1)), and

$$\sum_{k=1}^{N} (s_k)_i^n = 0, \tag{4.9}$$

for all i and n (the last relation imposes that the equality $\sum_{k=1}^{N} (Y_k)_i^n = 1$ is preserved).

4.2. *Implicit Time Stepping*

Let us now consider the extension to an implicit time stepping. As in the explicit case, the implicit (C) schemes will be based on the principle that the species equation is integrated after the Euler equations, using the discrete mass fluxes of the continuity equation.

Let us therefore consider as given a conservative implicit scheme for the Euler equations, which we write as

$$\frac{(W^l)_i^{n+1} - (W^l)_i^n}{\Delta t} + \frac{\phi_{i+1/2}^{l,n+1} - \phi_{i-1/2}^{l,n+1}}{\Delta x} = 0, \tag{4.10}$$

for $l = 1, 2, 3$, where in general $\phi_{i+1/2}^{l,n+1}$ depends on W_i^n, W_{i+1}^n, W_i^{n+1}, W_{i+1}^{n+1}: the values $(W^l)_i^{n+1}$ $(1 \leqslant l \leqslant 3)$ are computed from the values W_i^n by solving at each time step either a nonlinear or a linear discrete problem (see, e.g., [8, 13] for such implicit versions of the most classical upwind Euler schemes).

Then, we construct an implicit multi-component scheme of type (C) by adding the following discrete species equation:

$$\frac{(\rho Y)_i^{n+1} - (\rho Y)_i^n}{\Delta t} + \frac{\phi_{i+1/2}^{4,n+1} - \phi_{i-1/2}^{4,n+1}}{\Delta x} = 0, \tag{4.11}$$

where we take

$$\phi_{i+1/2}^{4,n+1} = \phi_{i+1/2}^{1,n+1} \times \begin{cases} Y_i^{n+1} & \text{if } \phi_{i+1/2}^{1,n+1} > 0, \\ Y_{i+1}^{n+1} & \text{if } \phi_{i+1/2}^{1,n+1} < 0. \end{cases} \tag{4.12}$$

Therefore the solution of a linear system is required at each time step to evaluate the mass fractions Y_i^{n+1}.

The next lemma shows that this scheme still has the property (3.10), without any restriction on the time step.

LEMMA 5. *For any value of the time step $\Delta t > 0$, the scheme* (4.10)–(4.12) *preserves the maximum principle for the mass fraction Y: for all i and $n \geqslant 0$,*

$$\min_j Y_j^0 \leqslant Y_i^n \leqslant \max_j Y_j^0. \tag{4.13}$$

Proof. The mass fractions Y_i^{n+1} at time level $n+1$ are obtained by solving the tridiagonal system

$$\sum_j A_{i,j}^{n+1} Y_j^{n+1} = \sum_j B_{i,j}^n Y_j^n, \tag{4.14}$$

where $B_{i,j}^n = \rho_i^n \delta_{i,j}$ (δ_{ij} is the Kronecker delta), and

$$\begin{aligned} A_{i,i}^{n+1} &= \rho_i^{n+1} + \lambda \max(\phi_{i+1/2}^{1,n+1}, 0) - \lambda \min(\phi_{i-1/2}^{1,n+1}, 0), \\ A_{i,i+1}^{n+1} &= \lambda \min(\phi_{i+1/2}^{1,n+1}, 0), \\ A_{i,i-1}^{n+1} &= -\lambda \max(\phi_{i-1/2}^{1,n+1}, 0). \end{aligned} \tag{4.15}$$

It is clear from these expressions that the matrix $A = (A_{i,j}^{n+1})$ is an M-matrix, i.e., that

$$\begin{aligned} A_{i,i}^{n+1} &> 0 \qquad \forall i, \\ A_{i,j}^{n+1} &\leqslant 0 \qquad \forall i \neq j, \\ \sum_j A_{i,j}^{n+1} &> 0 \qquad \forall i. \end{aligned} \tag{4.16}$$

These properties imply that the matrix A^{-1} has only positive elements (see, e.g., [33]): $(A^{-1})_{i,j} \geqslant 0$ for all i and j. Thus, writing $C = A^{-1}B$ (where $B = (B_{i,j}^n)$), we see that

$$C_{i,j} \geqslant 0 \qquad \text{for all } i \text{ and } j. \tag{4.17}$$

Now denoting U the vector whose components are all equal to 1, we see from (4.11)–(4.12) and (4.10) written for $l = 1$ that $AU = BU$, whence $CU = U$; that is,

$$\sum_j C_{i,j} = 1 \qquad \text{for all } i. \tag{4.18}$$

Since (4.14) implies that $Y_i^{n+1} = \sum C_{i,j} Y_j^n$, (4.18) and (4.19) show that Y_i^{n+1} is again a linear convex combination of the old mass fraction values Y_j^n, which ends the proof. ∎

Remark 14. The preceding proof is slightly too vague: we have not taken any boundary condition into account. In practice, the matrix A and the vector U are of finite size, and one has to check that the discrete maximum principle (4.13) holds with the actual discrete boundary conditions.

Remark 15. In practice, implementing the implicit scheme (4.10)–(4.12) may be far less simple than implementing the explicit schemes of type (C). Indeed, the implicit integration of the Euler equations often uses a linearized implicit flux (see, e.g., [8, 13]); that is, the fluxes $\phi^{l,n+1}_{i+1/2}$ in (4.10) are given as

$$\phi^{l,n+1}_{i+1/2} = \Phi^l(W^n_i, W^n_{i+1}) + \frac{\partial \Phi^l}{\partial W_i}(W^{n+1}_i - W^n_i) + \frac{\partial \Phi^l}{\partial W_{i+1}}(W^{n+1}_{i+1} - W^n_{i+1}), \tag{4.19}$$

which leads to solving at each time step a linear system for computing ρ^{n+1}_i, $(\rho u)^{n+1}_i$, E^{n+1}_i for all i. Once these values are known, one has to use them to evaluate the fluxes $\phi^{1,n+1}_{i+1/2}$ given by (4.19) before one can solve the species equation using the species fluxes (4.12) (see [24] for more details).

Remark 16. In some circumstances, such as for highly subsonic flows, it may be of interest to integrate explicitly the species equation while treating implicitly the Euler equations (see [10]). Then, to construct a semi-implicit multi-component scheme of type (C), we simply replace (4.12) by

$$\phi^{4,n+1}_{i+1/2} = \phi^{1,n+1}_{i+1/2} \times \begin{cases} Y^n_i & \text{if } \phi^{1,n+1}_{i+1/2} > 0, \\ Y^n_{i+1} & \text{if } \phi^{1,n+1}_{i+1/2} < 0. \end{cases} \tag{4.20}$$

This scheme again preserves the maximum principle (4.13) for the mass fraction, under a CFL-like condition (this condition is obtained by replacing the fluxes ϕ_1 by $\phi^{1,n+1}$ in (3.9)).

Remark 17. Consider again the flow of N species, with $N > 2$. Then, the matrix A appearing in the linear system which is to be solved at each time step for updating the mass fraction values is the same for all species equations; moreover, A is an M-matrix, which makes it easier to solve the linear systems.

Remark 18. Remark 11 also holds for implicit schemes. When the species equation has the form (3.16), the maximum principle (3.17) is preserved if the implicit approach (C) is used in conjunction with classical explicit or implicit approximations of the diffusive and reactive terms. More precisely, the maximum principle is preserved under appropriate stability conditions if the diffusive or the reactive term is integrated explicitly, and it is preserved for any value of the time step if all terms are treated implicitly. The latter is even true when a complex set of chemical reactions is considered (the proof of this fact is similar to the one in [14], where an upwind treatment of a *linear* convection operator is considered instead of the present approach (C)).

4.3. *Multi-dimensional Flows*

All the above methods which have been presented so far in a one-dimensional context can be extended to multi-dimensional flows. We simply recall here the basic features of this extension to two-dimensional geometries; to make the approach very general, we consider a two-dimensional unstructured triangular mesh (see Fig. 4), following the ideas of Dervieux and Fezoui (see [8, 12]); but all approaches (A), (B), and (C) can of course be used on structured quadrilateral finite-volume meshes.

Let us therefore consider the two-dimensional multi-component Euler equations $\hat{W}_t + \hat{F}_x + \hat{G}_y = 0$, with

$$\hat{W} = \begin{pmatrix} \rho \\ \rho u \\ \rho v \\ E \\ \rho Y \end{pmatrix}, \qquad \hat{F} = \begin{pmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E+p) \\ \rho uY \end{pmatrix}, \qquad \hat{G} = \begin{pmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E+p) \\ \rho vY \end{pmatrix}, \tag{4.21}$$

with:

$$p = (\gamma - 1)[E - \tfrac{1}{2}\rho(u^2 + v^2)], \tag{4.22}$$

where γ is again the local specific heat ratio (1.3) of the mixture.

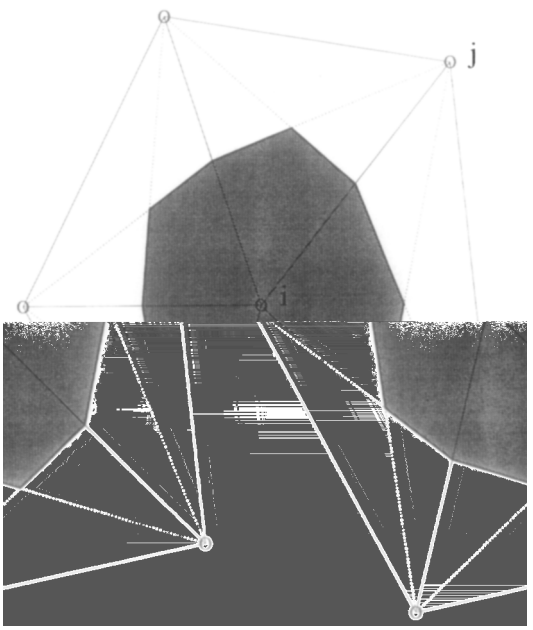

FIG. 4. The control volume C_i.

Then, we consider a dual partition of the domain in control volumes or cells: a cell C_i is constructed around each vertex S_i by means of the medians of the neighbouring triangles, as shown on Fig. 4.

Integrating the system $\hat{W}_t + \hat{F}_x + \hat{G}_y = 0$ on the control volume C_i, we obtain

$$\iint_{C_i} \hat{W}_t + \int_{\partial C_i} (\hat{F} v_i^x + \hat{G} v_i^y) = 0, \tag{4.23}$$

where $\mathbf{v}_i = (v_i^x, v_i^y)$ is the outward unit normal on ∂C_i. In fact, we rewrite (4.23) as

$$\iint_{C_i} \hat{W}_t + \sum_{j \in \mathscr{K}(i)} \int_{\partial C_{ij}} (\hat{F} v_i^x + \hat{G} v_i^y) = 0, \tag{4.24}$$

where $\mathscr{K}(i)$ is the set of neighbouring nodes of S_i, and where $\partial C_{ij} = \partial C_i \cap \partial C_j$. Then, the flux integrals in (4.24) are approximated using a 1D-like procedure, as explained below:

The fact that the system $\hat{W}_t + \hat{F}(\hat{W})_x + \hat{G}(\hat{W})_y = 0$ is a nonlinear hyperbolic system of conservation laws means that, for any $(\alpha, \beta) \in \mathbb{R}^2$, the matrix $\alpha(\partial \hat{F}/\partial \hat{W}) + \beta(\partial \hat{G}/\partial \hat{W})$ has five real eigenvalues,

$$\begin{aligned} \lambda_1 &= \alpha u + \beta v - \sqrt{\alpha^2 + \beta^2}\, c, \\ \lambda_2 &= \lambda_3 = \lambda_4 = \alpha u + \beta v, \\ \lambda_5 &= \alpha u + \beta v + \sqrt{\alpha^2 + \beta^2}\, c, \end{aligned} \tag{4.25}$$

with $c = \sqrt{\gamma p/\rho}$, and a complete set of real eigenvectors. Thus, we can extend all approximations defined in Section 2 for the one-dimensional flux vector F to the flux vector $\alpha \hat{F} + \beta \hat{G}$ (in other words, we use here the rotational invariance of the Euler equations). Consider, for instance, the Steger and Warming approximation. Given two values $\hat{W}_L$ and $\hat{W}_R$, and a vector $\boldsymbol{\eta} = (\eta^x, \eta^y)$, we define a numerical flux function Φ by

$$\Phi(\hat{W}_L, \hat{W}_R, \boldsymbol{\eta}) = (A_\eta)_+ (\hat{W}_L)\, \hat{W}_L + (A_\eta)_- (\hat{W}_R)\, \hat{W}_R, \tag{4.26}$$

where A_η is the matrix $\eta^x(\partial \hat{F}/\partial \hat{W}) + \eta^y(\partial \hat{G}/\partial \hat{W})$; thus, we define the vector $\mathbf{v}_{ij} = (v_{ij}^x, v_{ij}^y)$ by

$$v_{ij}^x = \int_{\partial C_{ij}} v_i^x, \qquad v_{ij}^y = \int_{\partial C_{ij}} v_i^y, \tag{4.27}$$

and we obtain a first-order accurate explicit upwind approximation of (4.24) by writing

$$\frac{\hat{W}_i^{n+1} - \hat{W}_i^n}{\Delta t} \operatorname{area}(C_i) + \sum_{j \in \mathscr{K}(i)} \Phi(\hat{W}_i^n, \hat{W}_j^n, \mathbf{v}_{ij}) = 0. \tag{4.28}$$

This extension to two space dimensions can of course be used for method (C). Then, the first four components of $\Phi(\hat{W}_i, \hat{W}_j, \mathbf{v}_{ij})$ are evaluated using an "Euler scheme," and the fifth component is given by

$$\Phi^5(\hat{W}_i, \hat{W}_j, \mathbf{v}_{ij}) = \Phi^1(\hat{W}_i, \hat{W}_j, \mathbf{v}_{ij}) \times \begin{cases} Y_i & \text{if } \Phi^1(\hat{W}_i, \hat{W}_j, \mathbf{v}_{ij}) > 0, \\ Y_j & \text{if } \Phi^1(\hat{W}_i, \hat{W}_j, \mathbf{v}_{ij}) < 0. \end{cases} \tag{4.29}$$

It is straightforward to check that this two-dimensional extension of method (C) still preserves the maximum principle for the mass fraction.

The above multi-dimensional approach (C) can of course be used when diffusive, viscous and reactive effects are added to the Euler equations (4.21). We refer to, e.g., [5, 21] for the details.

5. Numerical Results

Several numerical experiments, involving the first- or second-order accurate schemes of type (A), (B), and (C) have been done by D. Chargy, R. Abgrall, L. Fezoui, and the author, and detailed comparisons are presented in [6]. We simply illustrate here the main features of methods (A), (B), and (C) by showing one numerical example.

We consider the well-known shock tube of Sod [29], with two different species. More precisely, we solve the Riemann problem (3.1) with

$$\begin{aligned} &\rho_L = 1, && \rho_R = 0.125, \\ &u_L = 0, && u_R = 0, \\ &p_L = 1, && p_R = 0.1, \\ &Y_L = 1, && Y_R = 0, \end{aligned} \tag{5.1}$$

with $\gamma_1 = 1.4$, $\gamma_2 = 1.2$, and $C_{v1} = C_{v2}$. We have used 101 equally spaced mesh points in the interval $[-0.5, 0.5]$, and Δt is chosen at each time level according to the classical CFL condition:

$$\max_i(|u_i + c_i|, |u_i|, |u_i - c_i|)\frac{\Delta t}{\Delta x} = \text{CFL} \leqslant 1 \tag{5.2}$$

(we take CFL $= 0.75$).

Figures 5 and 6 show the results obtained at time $t = 0.21$ using the explicit Roe schemes of types (A), (B), and (C). These results are only first-order accurate: no particular attention should therefore be paid to the quality of the results for the Euler variables (density, velocity, and pressure), since no particular care of reducing the numerical diffusion has been taken. These results can be improved using any of the now classical devices for nonoscillatory second-order accurate schemes: slope

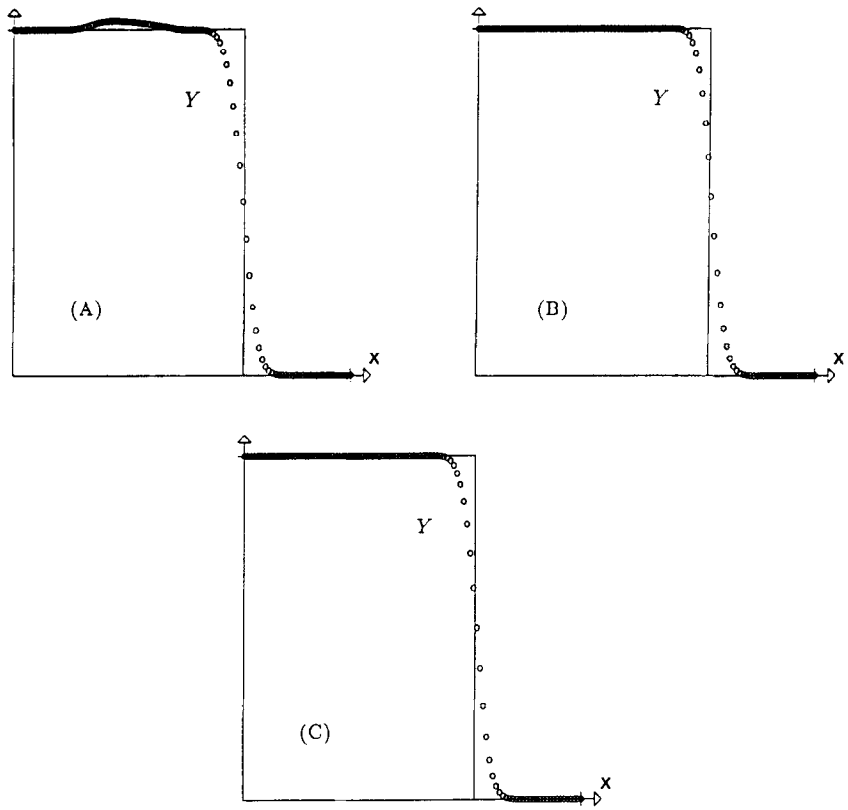

FIG. 5. Mass fraction profiles: computed and exact solutions.

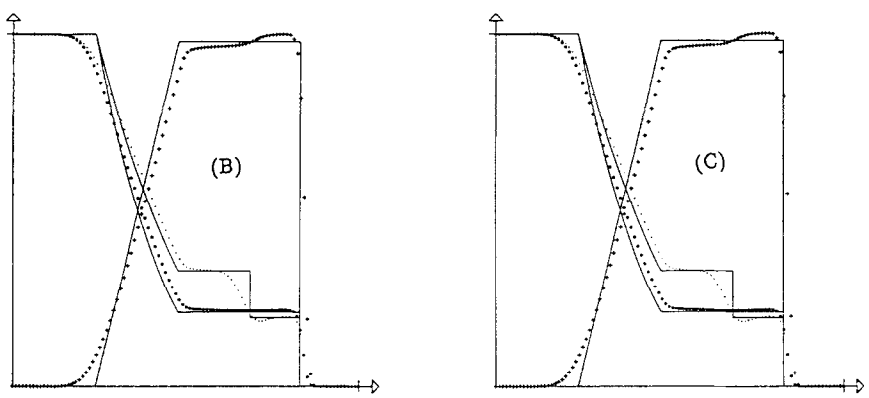

FIG. 6. Density, pressure, and velocity profiles: computed and exact solutions.

limiters, as briefly mentioned above, or TVD or FCT methods (see, e.g., [25, 34]). The quality of the results given by the Roe, Osher, or Van Leer schemes on the considered two-component shock tube problem are approximately the same as for the corresponding single-component shock tube; the only difference concerns the treatment of the contact discontinuity, where all schemes have difficulties in keeping a constant value for the velocity and pressure (see [1, 6, 22] for an attempt to explain this difficulty).

The mass fraction profiles are shown on Fig. 5. Clearly the results of methods (B) and (C) are far superior than those of approach (A), which gives very bad mass fraction values. Moreover, the computed mass fractions are slightly better with approach (C) than with approach (B). Notice indeed that values of Y which are slightly out of the interval [0, 1] appear in case (B), where $\max_i Y_i^n = 1 + 1.5 \times 10^{-6}$ (the error in this value is small but significant, since the computation has been made using double-precision real numbers; in other situations, negative values of Y of the order of 10^{-2} or even 10^{-1} appear when schemes of type (B) are used; see [6, 11]). Therefore, schemes of type (C), which preserve the maximum principle, actually give more accurate results for the mass fraction.

We should emphasize here that preserving the maximum principle for the mass fraction Y is far superior than preserving only the positivity of Y. Indeed, approach (A) preserves the positivity, but gives very poor results. On the other hand, approach (C) preserves the maximum principle.

Beside this, it appears on Fig. 6 that the results obtained in cases (B) and (C) for the hydrodynamical variables ρ, u, p are almost indistinguishable.

One could think that the schemes of type (C) are more expensive that the corresponding schemes of type (B), since applying method (C) consists in applying method (B) and then modifying the species mass fluxes using (3.8). But this is not the case, because only the first three components of the numerical flux ϕ have to be evaluated before one uses (3.8); in practice, the cost of method (C) is slightly smaller than the cost of method (B) (3 or 5% for an explicit calculation).

6. Conclusions

We have examined why the approximate Riemann solvers of Roe and Osher do not preserve the maximum principle for the mass fractions when applied to the computation of multi-component flows, and proposed a modification of these schemes, based on some property of the exact solution of the multi-component Riemann problem.

The basic idea behind this new approach is simple: when the same discrete mass fluxes are used to approximate the convective fluxes of the continuity equation and of all mass fraction equations, the resulting scheme preserves the maximum principle (and in particular the positivity) for all mass fractions.

This advantage may reveal many very useful applications (see the results in

[6, 24]), and in particular for reactive flows where the nonlinear chemical source terms precisely involve the mass fractions and where the appearance of negative mass fractions very often leads to nonlinear numerical instabilities.

APPENDIX

As already said in Remark 12, we examine in this Appendix the extension of approach (C) to fully second-order accurate (in both space and time) schemes.

In comparison with Section 4.1, the temporal second-order accuracy will now be obtained by evaluating the fluxes at the half time step $n+\frac{1}{2}$ (see, e.g., [12, 31]). Thus, the integration from time t^n to t^{n+1} is done using the five successive steps described below (steps (a) and (b) are the same as in Section 4.1):

(a) At each time step, starting from the values W_i^n, one first evaluate slopes s_i^n for all variables which are chosen to be piecewise linear. Again, we take here Y (and not ρY) as a piecewise linear variable.

(b) Slope limiters are then used in order to avoid the creation of new extrema; in particular, (4.1) still holds:

$$\min_{|j-i|\leqslant 1} Y_j^n \leqslant Y_i^n \pm \frac{\Delta x}{2} s_i^n \leqslant \max_{|j-i|\leqslant 1} Y_j^n. \tag{A.1}$$

(c) The limited slopes are used to evaluate cell-interface values $W_{i+1/2,\pm}^n$, and the solution is advanced in time over a half time step using a centered predictor, setting:

$$W_i^{n+1/2} = W_i^n - \frac{\lambda}{2}\left[F(W_{i+1/2,-}^n) - F(W_{i-1/2,+}^n)\right]. \tag{A.2}$$

(d) Next, we again use the same slopes s_i^n to evaluate cell interface values $W_{i+1/2,\pm}^{n+1/2}$ at the half time step (in particular, we set $Y_{i+1/2,-}^{n+1/2} = Y_i^{n+1/2} + (\Delta x/2)\, s_i^n$).

(e) Lastly, we complete the time step by evaluating W_i^{n+1}, using the upwind numerical flux function Φ,

$$\frac{W_i^{n+1} - W_i^n}{\Delta t} + \frac{\phi_{i+1/2}^{n+1/2} - \phi_{i-1/2}^{n+1/2}}{\Delta x} = 0, \tag{A.3}$$

where:

$$\phi_{i+1/2}^{n+1/2} = \Phi(W_{i+1/2,-}^{n+1/2},\ W_{i+1/2,+}^{n+1/2}). \tag{A.4}$$

Let us use this construction (a)–(e) with a numerical flux function of type (C), that is with the fourth component ϕ^4 of the flux evaluated using (3.8), and examine

whether the maximum principle is preserved for the mass fraction under the following CFL-like condition (which now replaces (4.3)):

$$\frac{\Delta t}{\Delta x}\left[\frac{\max(\phi_{i+1/2}^{1,n+1/2},0)}{\rho_i^n}-\frac{\min(\phi_{i+1/2}^{1,n+1/2},0)}{\rho_{i+1}^n}\right]\leqslant\frac{1}{2}\,\forall i. \tag{A.5}$$

We will assume that $\phi_{i+1/2}^{1,n+1/2}>0$ and $\phi_{i-1/2}^{1,n+1/2}>0$. Then, we have $\phi_{i\pm1/2}^{4,n+1/2}=\phi_{i\pm1/2}^{1,n+1/2}Y_{i\pm1/2,-}^{n+1/2}$, and we obtain

$$Y_i^{n+1}=\frac{(\rho_i^n-\lambda\phi_{i+1/2}^{1,n+1/2})\,\hat{Y}+\lambda\phi_{i-1/2}^{1,n+1/2}Y_{i-1/2,-}^{n+1/2}}{(\rho_i^n-\lambda\phi_{i+1/2}^{1,n+1/2})+\lambda\phi_{i-1/2}^{1,n+1/2}}, \tag{A.6}$$

where we now have set

$$\hat{Y}=\frac{\rho_i^nY_i^n-\lambda\phi_{i+1/2}^{1,n+1/2}Y_{i+1/2,-}^{n+1/2}}{\rho_i^n-\lambda\phi_{i+1/2}^{1,n+1/2}}. \tag{A.7}$$

Let again a and b be two real numbers such that $a\leqslant Y_j^n\leqslant b$ for all j; we have to examine if both values $Y_{i-1/2,-}^{n+1/2}$ and $\hat{Y}$ are in the interval $[a,b]$.

Let us introduce some notations. We set

$$X_{i+1/2}=\lambda\frac{\phi_{i+1/2}^{1,n+1/2}}{\rho_i^n}, \tag{A.8}$$

and

$$Z_{i+1/2}=\lambda\frac{(\rho u)_{i+1/2,-}^n}{\rho_i^n},\qquad Z_{i-1/2}=\lambda\frac{(\rho u)_{i-1/2,+}^n}{\rho_i^n}. \tag{A.9}$$

Then, we obtain

$$\rho_i^{n+1/2}=\rho_i^n-\tfrac{1}{2}\rho_i^n(Z_{i+1/2}-Z_{i-1/2}), \tag{A.10}$$

$$\rho_i^{n+1/2}Y_i^{n+1/2}=\rho_i^nY_i^n-\tfrac{1}{2}\rho_i^n(Z_{i+1/2}Y_{i+1/2,-}^n-Z_{i-1/2}Y_{i-1/2,+}^n), \tag{A.11}$$

and it is easy to check that

$$Y_i^{n+1/2}=Y_i^n-\frac{\Delta x}{4}s_i^n\frac{Z_{i+1/2}+Z_{i-1/2}}{1-\frac{1}{2}(Z_{i+1/2}-Z_{i-1/2})}, \tag{A.12}$$

$$Y_{i+1/2,-}^{n+1/2}=Y_i^n+\frac{\Delta x}{2}s_i^n\left[\frac{1-Z_{i+1/2}}{1-\frac{1}{2}(Z_{i+1/2}-Z_{i-1/2})}\right], \tag{A.13}$$

$$\hat{Y}=Y_i^n-\frac{\Delta x}{2}s_i^n\left[\frac{1-Z_{i+1/2}}{1-\frac{1}{2}(Z_{i+1/2}-Z_{i-1/2})}\right]\cdot\left(\frac{X_{i+1/2}}{1-X_{i+1/2}}\right). \tag{A.14}$$

We know from (A.5) and our assumptions on the sign of $\phi_{i+1/2}^{1,n+1/2}$ that

$0 \leqslant X_{i+1/2} \leqslant \frac{1}{2}$. Thus the term in parentheses in (A.14) is in the interval [0, 1]. If, moreover, we also have

$$0 \leqslant Z_{j \pm 1/2} \leqslant \tfrac{1}{2} \qquad \text{for all } j, \tag{A.15}$$

then it is easy to conclude from (A.13), (A.14), and (A.1) that the term in brackets in (A.13)–(A.14) also is in the interval [0, 1], whence $Y^{n+1/2}_{i+1/2,-} \in [a, b]$ and $\hat{Y} \in [a, b]$. But these conclusions do not hold any longer if $Z_{i-1/2} + Z_{i-1/2} < 0$ (the term in brackets in (A.13)–(A.14) then exceeds 1). Of course, since we have assumed that $\phi^{1,n+1/2}_{i+1/2} > 0$ and $\phi^{1,n+1/2}_{i-1/2} > 0$, the quantities $Z_{i+1/2}$ and $Z_{i-1/2}$ are likely to be positive. Therefore, we can only conclude (but this conclusion is confirmed by the numerical experiments; see [6]) that this second-order accurate scheme preserves *in general* the maximum principle for the mass fraction.

ACKNOWLEDGMENTS

I thank my colleagues R. Abgrall, D. Chargy, and L. Fezoui for their help in this work.

REFERENCES

1. R. ABGRALL, *Rech. Aérosp.* **6**, 31 (1988).
2. R. ABGRALL AND J. L. MONTAGNÉ, *Rech. Aérosp.* **4**, 1 (1989).
3. F. BENKHALDOUN, A. DERVIEUX, G. FERNANDEZ, H. GUILLARD, AND B. LARROUTUROU, "Some Finite-Element Investigations of Stiff Combustion Problems: Mesh Adaption and Implicit Time-Stepping," in *Mathematical Modelling in Combustion and Related Topics*, edited by C. Brauner and C. Schmidt-Lainé (NATO ASI Series E, Nijhoff, Dordrecht, 1988), p. 393.
4. G. V. CANDLER AND R. W. MCCORMACK, "The Computation of Hypersonic Flows in Chemical and Thermal Nonequilibrium," Paper No. 107, Third National Aero-Space Plane Technology Symposium (1987).
5. D. CHARGY, "Simulation numérique d'écoulements réactifs transsoniques," thesis, in preparation.
6. D. CHARGY, R. ABGRALL, L. FEZOUI, AND B. LARROUTUROU, "Comparisons of Several Numerical Schemes for Multi-component One-dimensional Flows," INRIA Report 1253, (1990).
7. D. CHARGY, A. DERVIEUX, AND B. LARROUTUROU, "Upwind adaptive finite-element investigations of two-dimensional transonic reactive flows," *Int. J. Num. Methods Fluids* **11**, 751 (1990).
8. A. DERVIEUX, "Steady Euler Simulations Using Unstructured Meshes," in *Partial Differential Equations of Hyperbolic Type and Applications*, edited by G. Geymonat (World Scientific, Singapore, 1987), p. 33.
9. J. A. DÉSIDÉRI, N. GLINSKY, AND E. HETTENA, "Hypersonic Reactive Flow Computations," *Comput. Fluids* **18** (2), 151–182 (1990).
10. G. FERNANDEZ AND H. GUILLARD, "Implicit Schemes for Subsonic Combustion Problems," in *Numerical Combustion*, edited by A. Dervieux and B. Larrouturou (Lecture Notes in Physics, Vol. 351, Springer-Verlag, New York/Berlin, 1989), p. 277.
11. G. FERNANDEZ AND B. LARROUTUROU, "Hyperbolic Schemes for Multi-component Euler Equations," in *Nonlinear Hyperbolic Equations—Theory, Numerical Methods, and Applications*, edited by J. Ballmann and R. Jeltsch (Notes on Numerical Fluid Mechanics, Vol. 24, Vieweg, Braunschweig, 1989), p. 128.
12. L. FEZOUI, INRIA Report 358 (1985).

13. L. FEZOUI AND B. STOUFFLET, *J. Comput. Phys.* **84**, No. 1, 174 (1989).
14. M. GHILANI AND B. LARROUTUROU, "Upwind Computation of Steady Planar Flames with Complex Chemistry," *Mod. Math. Anal. Num.* **25** (1), 67 (1991).
15. S. K. GODUNOV, *Math. Sb.* **47**, 271 (1959).
16. B. GROSSMANN AND P. CINNELLA, "Upwind Methods for Flows with Non-equilibrium Chemistry and Thermodynamics," in *Numerical Combustion*, edited by A. Dervieux and B. Larrouturou (Lecture Notes in Physics, Vol. 351, Springer-Verlag, New York/Berlin, 1989), p. 323.
17. A. HABBAL, A. DERVIEUX, H. GUILLARD, AND B. LARROUTUROU, INRIA Report 690 (1987).
18. A. HARTEN, P. D. LAX, AND B. VAN LEER, *SIAM Rev.* **25**, 35 (1983).
19. K. KAILASANATH, E. S. ORAN, AND J. P. BORIS, "Numerical Simulations of Flames and Detonations," in *Numerical Combustion*, edited by A. Dervieux and B. Larrouturou (Lecture Notes in Physics, Vol. 351, Springer-Verlag, New York/Berlin, 1989), p. 82.
20. B. LARROUTUROU, *Introduction to Combustion Modelling* (Springer Series in Computational Physics, 1991), to appear.
21. B. LARROUTUROU, "Recent Progress in Reactive Flow Computations," in *Computing Methods in Applied Sciences and Engineering*, edited by R. Glowinski (SIAM, Philadelphia, 1990), p. 249.
22. B. LARROUTUROU AND L. FEZOUI, "On the Equations of Multi-component Perfect or Real Gas Inviscid Flow," in *Nonlinear Hyperbolic Problems*, edited by Carasso Charrier Hanouzet Joly (Lecture Notes in Mathematics, Vol. 1402, Springer-Verlag, Heidelberg, 1989).
23. P. D. LAX, "Hyperbolic Systems of Conservation Laws and the Mathematical Theory of Shock Waves," CBMS Regional Conference Series in Applied Mathematics, Vol. 11 (SIAM, Philadelphia, 1972).
24. B. N'KONGA, H. GUILLARD, AND B. LARROUTUROU, in preparation.
25. E. S. ORAN AND J. P. BORIS, *Numerical Simulation of Reactive Flows* (Elsevier, New York, 1987).
26. S. OSHER AND F. SOLOMON, *Math. Comput.* **38**, 339 (1982).
27. W. RODI, *Turbulence Models and Their Applications in Hydraulics*, AIRH (1984).
28. P. L. ROE, *J. Comput. Phys.* **43**, 357 (1981).
29. G. A. SOD, *J. Comput. Phys.* **27**, 1 (1977).
30. J. L. STEGER AND R. F. WARMING, *J. Comput. Phys.* **40**, 263 (1981).
31. B. VAN LEER, *J. Comput. Phys.* **23**, 263 (1977).
32. B. VAN LEER, "Flux-Vector Splitting for the Euler Equations," in *Eighth International Conference on Numerical Methods in Fluid Dynamics*, edited by E. Krause (Lecture Notes in Physics, Vol. 170, Springer-Verlag, New York/Berlin, 1982), p. 507.
33. R. S. VARGA, *Matrix Iterative Analysis* (Prentice-Hall, Englewood Cliffs, NJ, 1962).
34. H. C. YEE, NASA Ames Technical Memorandum 89464 (1987).