Maximum principle on the entropy and minimal limitations for kinetic schemes

Brahim Khobalatte, Benoît Perthame

INRIA

UNITE DE RECHERCHE
INRIA-ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
BP105
78153 Le Chesnay Cedex
France
Tél.(1) 39 63 55 11

Rapports de Recherche

1992

ème anniversaire

N° 1628

Programme 6

Calcul Scientifique, Modélisation et
Logiciel numérique par Ordinateur

MAXIMUM PRINCIPLE ON THE ENTROPY AND MINIMAL LIMITATIONS FOR KINETIC SCHEMES

Brahim KHOBALATTE
Benoit PERTHAME

Février 1992

MAXIMUM PRINCIPLE ON THE ENTROPY AND MINIMAL LIMITATIONS FOR KINETIC SCHEMES

Brahim Khobalatte(*) and Benoit Perthame(*)(**)

(*) INRIA, Centre de Rocquencourt, Projet MENUSIN
BP. 105, 78173 Le Chesnay Cedex, France

(**) Université d'Orléans, Département de Mathématiques
BP. 6759, 45067 Orléans Cedex 2, France

Abstract :

We consider kinetic schemes for the multidimensional inviscid gaz dynamics equations (compressible Euler equations).We prove that the discrete maximum principle holds for a special convex entropy. This fixes the choice of the equilibrium functions necessary for kinetic schemes. We use this property to perform a second order oscillation free scheme where only one slope limitation (for three conserved quantities in 1d) is necessary. Numerical results assert the strong convergence of the scheme.

PRINCIPE DU MAXIMUM SUR L'ENTROPIE ET LIMITEURS POUR SCHEMAS CINETIQUES

Résumé :

Nous considérons des schémas cinétiques pour les équations de la dynamique des gaz non-visqueux (Equations d'Euler compressible). Nous prouvons le principe du maximum discret pour une entropie convexe particulière. Ceci fixe le choix de la fonction d'équilibre nécessaire aux schémas cinétiques. Nous utilisons ensuite cette propriété pour réaliser des schémas du second ordre, non-oscillants, où une seule limitation est nécessaire (pour trois quantités conservées en 1D). Des résultats numériques indiquent la convergence forte du schéma.

Key-words : Compressible Euler Equations - Upwind schemes - Kinetic schemes - Entropy property - Second order schemes.

Mots-clés : Equation d'Euler compressible - Schémas décentrés - Schémas cinétiques - Propriété d'entropie - Schémas d'ordre deux.

A.M.S. class. numbers : 35L64, 76N10, 65M93, 76P05.

Introduction

We consider the gas dynamics equations in one or two space dimensions

$$(1)\qquad \begin{cases} \partial_t \rho + div(\rho u) = 0 \\ \partial_t \rho u_j + div(\rho u_j u) + \partial_{x_j} p = 0, j = 1, 2, \\ \partial_t E + div[(E + p)u] = 0 \end{cases}$$

where $x = (x_1, x_2), u = (u_1, u_2)$ and the total energy $E = \rho|u|^2/2 + \rho T/(\gamma - 1)$ is related to the pressure by the relation $p = \rho T, 1 < \gamma \leq 2$ in dimension 2, $1 < \gamma \leq 3$ in dimension 1.

It is known that, because of shock waves, an entropy inequality has to be added to (1) (see Lax [3] for instance)

$$(2)\qquad \partial_t \rho S + div(\rho u S) \leq 0,$$

where the specific entropy can be choosen as

$$(3)\qquad S = \rho / T^{1/(\gamma-1)}.$$

As it was proved by Tadmor [10], the combination of (1) and (2) yields that S satisfies the maximum principle

$$(4)\qquad S(x, t+h) \leq Max\{S(y,t); |y - x| \leq \|u\|_\infty h\},$$

and, in 1D, Godunov and Lax-Friedrichs schemes preserv this property at the discretized level because they solve exactly the system (1). A reason why (4) should hold is that S satisfies the (meaningless) equation

$$\partial_t S + u.\nabla_x S \leq 0.$$

The purpose of this paper is to show that the property (4) is also satisfied for kinetic schemes in 1 or 2 dimensions (we do not consider higher dimensions here but the extension is straightforward). This requires to choose the equilibrium function in an appropriate way, in the class introduced by Perthame [6, 7], and to interpret the scheme as a discretization of a transport equation. Then, the property (4) follows from a variational principle. It is remarkable that the appropriate equilibrium function is not the maxwellian distribution.

It is natural to try to extend this property to second order accurate schemes. Then, it appears that a conservative second order reconstruction, following the method introduced by Van Leer [12], has to increase the specific entropy and we can only impose the maximum principle up to a second order error. This is achieved in reconstructing second order approximation $\varphi_{i+\frac{1}{2}}$ of $\varphi(x_{i+\frac{1}{2}})$ for $\varphi = \rho, u$ or S. To do so, we use *centered* predictions of $\Delta\varphi$ and we just impose

$$(5)\qquad 0 \leq S_{i+\frac{1}{2}}, S_{i-\frac{1}{2}} \leq Max(S_i, S_{i+1}, S_{i-1}) \quad \text{and} \quad \rho_{i+\frac{1}{2}} \geq 0$$

and the conservation of the quantities $\Psi = \rho, \rho u, E$ i.e.

$$2\Psi_i = \Psi_{i+\frac{1}{2}} + \Psi_{i-\frac{1}{2}}. \tag{6}$$

In practice, to realize (6), we have to relax (5) up to second order.

Numerical tests show that this limitation (5) alone is enough to prevent much of the oscillations in the fully second order scheme, at least for some classical tests. This is somewhat surprising since non-oscillatory schemes usually require as many limitations as conserved quantities, even though ENO theory ([2, 9] and the references therein) shows that some flexibility in the reconstruction is possible.

We would like to point out that the conservative entropy inequality (2) is well understood at the discrete level (Osher [5], Tadmor [11]) for general hyperbolic systems. But the maximum principle for the specific entropy (3) is not a consequence of (2) alone and it holds only for the particular case of gas dynamics (and related systems), therefore it requires a specific proof.

The paper is organized as follows. In the first section, we consider the 1D case and the 2D case, for a general mesh (rectangular, triangular, dual type) is treated in Section II. Then, we describe the second order scheme, how the limitation (4) is used and numerical tests in Section III.

I. The 1D case

The general form of a conservative scheme for (1) is

$$U_i^{n+1} - U_i^n + \sigma(F_{i+1/2}^n - F_{i-1/2}^n) = 0 \tag{7}$$

where $U_i^n = (\rho_i^n, (\rho u)_i^n, E_i^n)^t$ is the average, on the mesh $(x_{i-1/2}, x_{i+1/2})$ with uniform size Δx, of the vector $(\rho(x, n\Delta t), \rho u(x, n\Delta t), E(x, n\Delta t))$. The time step Δt is related to σ by

$$\sigma = \Delta t / \Delta x. \tag{8}$$

The class of kinetic schemes we are going to consider is given by a flux splitting

$$F_{i+1/2}^n = F^+(U_i^n) + F^-(U_{i+1}^n) \tag{9}$$

$$F^+(U) = \rho \int_{v \geq 0} v[(1, v, v^2/2)^t \chi(\frac{v-u}{\sqrt{T}}) + (0, 0, T)^t \zeta(\frac{v-u}{\sqrt{T}})] dv / \sqrt{T} \tag{10}$$

and F^- is obtained in integrating over $v \leq 0$, rather than $v \geq 0$. This flux is consistant as soon as $F^-(U) + F^+(U) = (\rho u, \rho u^2 + \rho T, (E+p)u)^t$ which is achieved when χ, ζ are even, nonnegative functions satisfying (see [7])

$$\int_{\mathbb{R}} (1, w^2) \chi(w) dx = (1, 1), \int_{\mathbb{R}} \zeta(w) dw = \lambda := \frac{1}{2}(3 - \gamma)/(\gamma - 1). \tag{11}$$

We are now going to specify χ and ζ

$$\chi(w) = \alpha(1 - w^2/\beta)_+^\lambda, \zeta(w) = \delta^{\frac{\gamma-1}{\gamma-3}}[\chi(w)]^{(\gamma+1)/(3-\gamma)} \tag{12}$$

where α, β, δ are choosen to satisfy (11) i.e.

$$\begin{cases} \alpha = [2\sqrt{\beta}\int_0^{\frac{\pi}{2}} cos^{2/(\gamma-1)}\theta d\theta]^{-1}, \\ \beta = \int_0^{\frac{\pi}{2}} cos^{2/(\gamma-1)}\theta d\theta / \int_0^{\frac{\pi}{2}} sin^2\theta cos^{2/(\gamma-1)}\theta d\theta, \\ \delta^{\frac{\gamma+1}{\gamma-3}} = \dfrac{\lambda\beta}{2(\lambda+1)}\alpha^{-1/\lambda}. \end{cases} \tag{13}$$

With this choice we can obtain a family of singular entropy inequalities. They correspond to the generalized convex functions $\rho\Pi(U)$, where Π is parametrized by $\eta > 0$

$$\Pi(U) = \begin{cases} 0 & \text{if} \quad \rho^{\gamma-1}/T < \eta, \\ 1 & \text{if} \quad \rho^{\gamma-1}/T = \eta, \\ +\infty & \text{otherwise.} \end{cases} \tag{14}$$

which are obtained as the limit as p tends to $+\infty$ of the convex entropies $\rho(\rho^{\gamma-1}/T\eta)^p$. The corresponding conservative entropy inequality has a flux splitting form

$$\begin{aligned} (\rho\Pi)_i^{n+1} \le (\rho\Pi)_i^n - \sigma G^+(U_i^n) + \sigma G^-(U_i^n) \\ + \sigma G^+(U_{i-1}^n) - \sigma G^-(U_{i+1}^n), \end{aligned} \tag{15}$$

where the entropy fluxes $G^\pm$ depend on η

$$G^\pm(U) = F_\rho^\pm(U)\Pi(U), \tag{16}$$

where $F_\rho^\pm$ is the mass flux in (7), (9). It has to be noted that $G^\pm$ has the sign $\pm$ and $\rho_i^n - \sigma F_\rho^+(U_i^n) + \sigma F_\rho^-(U_i^n) \ge 0$ with the above CFL condition. Therefore the r.h.s. of (15) is the sum of three nonnegative terms depending respectively of $U_i^n, U_{i-1}^n, U_{i+1}^n$. We use the convention that the r.h.s. is $+\infty$ whenever one of those three terms is $+\infty$.

We can now state our main result

Theorem 1. With the choice (12), the kinetic scheme (7)-(10) satisfies
(i) $\rho_i^{n+1} \ge 0, T_i^{n+1} \ge 0$ whenever $\rho_i^n, T_i^n \ge 0$,
(ii) the conservative entropy inequalities (15) for any $\eta > 0$,
(iii) the maximum principle on the specific entropy

$$S_i^{n+1} := \rho_i^{n+1}/(T_i^{n+1})^{1/(\gamma-1)} \le Max(S_i^n, S_{i+1}^n, S_{i-1}^n) \tag{17}$$

under the CFL condition $(|u_i^n| + \sqrt{\beta T_i^n})\sigma \le 1, \forall i$.

Remarks. 1) for $\gamma = 1.4$, we find $\beta = 7$, and thus our CFL condition is stricter then the classical one. But in practice we can use the classical one.

2) the classical Boltzmann scheme corresponds to $\chi = \alpha e^{-w^2/2}$ (see Deshpande [1]). It cannot be analyzed in those terms. Also the theory of kinetic formulations of isentropic systems developed in Lions-Perthame-Tadmor [4] requires the χ function in (12), but there, the entropies are much simpler than those developed in the proof below.

Proof of Theorem 1

1st step. The kinetic level. We first introduce the discretized transport equations

$$\bar{f}_i(v) - f_i^n(v) + \sigma[v_+ f_i^n(v) - v_- f_{i+1}^n(v) - v_+ f_{i-1}^n(v) + v_- f_i^n(v)] = 0, \tag{18}$$

$$\bar{g}_i(v) - g_i^n(v) + \sigma[v_+ g_i^n(v) - v_- g_{i+1}^n(v) - v_+ g_{i-1}^n(v) + v_- g_i^n(v)] = 0, \tag{19}$$

where $v_+ = max(0, v), v_+ - v_- = v$ and

$$f_i^n(v) = \rho_i^n \chi[(v - u_i^n)/\sqrt{T}_i^n], g_i^n(v) = \rho_i^n \sqrt{T}_i^n \zeta[(v - u_i^n)/\sqrt{T}_i^n] \tag{20}$$

As usual ([1], [7]) the finite difference scheme (7), (10) is deduced multiplying (18) by the vector $(1, v, v^2/2)^t$ and adding to it (19) multiplied by $(0, 0, 1)^t$ and integrating dv.

Indeed, this clearly follows from the identities

$$\begin{aligned} U_i^n &= \int_{\mathbb{R}} (f_i^n, v f_i^n, \frac{v^2}{2} f_i^n + g_i^n)^t dv, \\ U_i^{n+1} &:= \int (\bar{f}_i, v\bar{f}_i, \frac{v^2}{2}\bar{f}_i + \bar{g}_i)^t dv, \\ F^\pm(U_i^n) &= \pm \int v_\pm (f_i^n, v f_i^n, \frac{v^2}{2} f_i^n + g_i^n)^t dv, \end{aligned} \tag{21}$$

which follows from the only consistancy relations (11). Now, we have $\bar{f}_i \geq 0, \bar{g}_i \geq 0$, under the condition $\sigma|v| \leq 1$ for all v such that $f_i^n(v) \neq 0$, and this is exactly the CFL condition of theorem 1. This proves (i).

2nd step. The maximum principle. We notice that $h = (f^{\gamma+1} g^{\gamma-3})^{1/(\gamma-1)}$ is a convex function of f, g. Since $\bar{f}_i$ and $\bar{g}_i$ are also convex combinations of $f_i^n, f_{i+1}^n, f_{i-1}^n$ and $g_i^n, g_{i+1}^n, g_{i-1}^n$ (whenever σ satisfies the CFL condition), we thus have

$$\bar{h}_i \leq h_i^n(1 - \sigma v_+ - \sigma v_-) + h_{i+1}^n \sigma v_- + h_{i-1}^n \sigma v_+. \tag{22}$$

Now ,with the choice (12) of χ, ζ, h is just given by

$$h_i^n = \delta\,(S_i^n)^2 \mathbf{1}_{\{|v-u_i^n|^2 \leq \beta T_i^n\}} \tag{23}$$

and thus we obtain

$$\bar{h}_i \leq \Sigma := \delta Max(S_i^n, S_{i+1}^n, S_{i-1}^n)^2. \tag{24}$$

At this level we need a lemma similar to those of [6], and which proof is simple calculus of variation and thus is skept.

Lemma 2. Let $e = \rho\tau/(\gamma-1)$ such that

$$e = min\{\int_{\mathbb{R}}[\frac{v^2}{2}f(v)+g(v)]dv; f(v)\geq 0, g(v)\geq 0,$$
$$\int_{\mathbb{R}}(1,v)f(v)dv = (\rho,0),(\rho\geq 0), f^{\gamma+1}g^{\gamma-3}\leq \Sigma^{\gamma-1}\}.$$

Then $\rho/\tau^{1/(\gamma-1)} = \sqrt{\Sigma/\delta}$ and the minimum is achieved by $f = \frac{\rho}{\sqrt{\tau}}\chi(\frac{v-u}{\sqrt{\tau}}), g = \rho\sqrt{\tau}\zeta(\frac{v-u}{\sqrt{\tau}})$. □

Going back to the proof, we apply the lemma with $f = \bar{f}_i(v-u_i^{n+1}), g = \bar{g}_i(v-u_i^n)$ so that the constraints in the minimization problem are realized with $\rho = \rho_i^{n+1}$ and Σ given in (24). We thus have

$$(\gamma-1)\int\frac{v^2}{2}f+g = \rho_i^{n+1}T_i^{n+1}\geq e(\gamma-1) =$$
$$\rho_i^{n+1}(\rho_i^{n+1}/\sqrt{\Sigma/\delta})^{\gamma-1}$$

which exactly means $S_i^{n+1}\geq\sqrt{\Sigma/\delta}$ and (iii) is proved.

3rd step. Entropy inequality. As in the second step, let us introduce, for a fixed positive number η, and for $p\geq 1$ the function

$$k = f[\frac{(f^{\gamma+1}g^{\gamma-3})^{1/(\gamma-1)}}{\eta^2}]^p.$$

Since it is a convex function of f and g we also have

$$\bar{k}_i(v)\leq k_i^n(v)(1-\sigma v_+-\sigma v_-)+k_{i+1}^n(v)\sigma v_- + k_{i-1}^n(v)\sigma v_+.$$

we need now the

Lemma 3. The minimization problem

$$Min\{\int_{\mathbb{R}}f(f^{\gamma+1}g^{\gamma-3})^{p/(\gamma-1)}dv; f\geq 0, g\geq 0,$$
$$\int_{\mathbb{R}}(1,v)fdv = (\rho,0), \int_{\mathbb{R}}\frac{v^2}{2}f+g = \rho T/(\gamma-1)\} \tag{25}$$

admits a unique minimum

$$F_p = \frac{\rho\alpha_p}{\sqrt{T}}(1-\frac{v^2}{\beta_p T})_+^{(1+2p\lambda)/2p}, G_p = T\delta_p F_p\,(1-\frac{v^2}{\beta_p T})_+$$

where $\alpha_p, \beta_p, \delta_p$ are such that the constraints in (25) are satisfied. □

Again we skip the proof of this lemma which consists in writting the Euler-Lagrange equations associated with (25). As before, we use this lemma with $\rho = \rho_i^{n+1}, T = T_i^{n+1}, f = \bar{f}_i(v - u_i^{n+1}), g = \bar{g}_i(v - u_i^{n+1})$ and the corresponding minimizer F_p, G_p thus satisfy

$$
(26) \qquad \begin{aligned}
&\int_{\mathbb{R}} F_p[\frac{(F_p^{\gamma+1} G_p^{\gamma-3})^{1/(\gamma-1)}}{\eta^2}]^p \leq \int_{\mathbb{R}} \bar{k}_i(v) dv \\
&\leq \int_{\mathbb{R}} k_i^n(v)(1 - \sigma v_+ - \sigma v_-) + k_{i+1}^n(v)\sigma v_- + k_{i-1}^n(v)\sigma v_+] dv.
\end{aligned}
$$

Now we let p go to $+\infty$ and we find exactly the entropy inequality (15) since the r.h.s. of (26) goes to $\rho_i^{n+1}\Pi(U_i^{n+1})$ and

$$
\begin{aligned}
&\int_{\mathbb{R}} k_i^n(v)(1 - \sigma v_+ - \sigma v_-) dv \to (\rho_i^n - \sigma F_\rho^+(U_i^n) + \sigma F_\rho^-(U_i^n))\Pi(U_i^n) \\
&\int k_{i\pm1}^n(v) v_\pm dv \to \pm F_\rho^\pm(U_i^n)\Pi(U_i^n).
\end{aligned}
$$

This concludes the proof of Theorem 1. □

Let us end this section by some remarks on the entropy. First, let us notice that the choice

$$
\chi = \alpha_p(1 - \frac{w^2}{\beta_p})_+^{(1+2p\lambda)/2p}, \varphi = \delta_p \chi.(1 - \frac{w^2}{\beta_p})
$$

in the scheme (7)-(10) leads to an entropy inequality (for a regular entropy now) of the form (15) with

$$
\Pi(U) = \mu_p \rho(\rho/T^{1/(\gamma-1)})^{2p}, G^+ = \int_{v\geq 0} v F_p (F_p^{\gamma+1} G_p^{\gamma-3})^{p/(\gamma-1)} dv,
$$

with F_p, G_p defined in lemma 3 and some appropriate constant μ_p. The proof of this as well as the proof of (ii) in Theorem 1 follows in fact that of [6], but here we have a more general approach dealing with two functions f, g rather than two kinetic vriables v, I as in Deshpande [1]. Also we would like to emphasize that an exact entropy inequality is necessary to get a maximum principle on the specific entropy, and it is an open question if the proofs of Osher [5] or Tadmor [11] could be extended to get, for Roe or Osher schemes a maximum principle, or for kinetic schemes the entropy inequality.

II. The 2D case

We show here that our results can be naturally extended to the 2D equations discretized on an unstructured mesh. Our motivations and notations follow those introduced in Perthame-Qiu [8].

Let us consider a grid as shown in figure 1, which cells C_i have $L(i)$ edges $E_1, ...E_L(L = 3$ for triangles, 4 for rectangles and depends on i for dual type grids). We call ν_l the unit outward normal to $E_l, |E_l|$ the length of $E_l, |C_i|$ the area of C_i and $j(l)$ the index of the cell $C_{j(l)}$ neighboring C_i along E_l ($j(l)$ also depends on i, but we skip this dependance for simplicity).

We now set $U_i^n = (\rho, \rho u_1, \rho u_2, E)_i^n$ and we consider numerical schemes for the equations (1) under the form

$$
(27) \qquad \begin{cases} U_i^{n+1}|C_i| = U_i^n|C_i| - \Delta t \sum_{l=1}^{L(i)} |E_l| F_l.\nu_l, \\ F_l \nu_l = F^+(U_i^n).\nu_l + F^-(U_{j(l)}^n), \\ F^{\pm}(U) = \pm\rho \int_{\mathbb{R}^2} (v.\nu) \pm \ [(1, v, |v|^2/2)^t \chi(\frac{v-u}{\sqrt{T}}) \\ \qquad\qquad + (0,0,T)^t \varphi(\frac{v-u}{\sqrt{T}})] dv/T. \end{cases}
$$

The consistancy relations are now that the nonnegative even functions χ, φ satisfy

$$
(28) \qquad \int_{\mathbb{R}^2} (1, w_k w_l)\chi(w)dw = (1, \delta_{kl}), \int_{\mathbb{R}^2} \varphi(w)dw = \lambda := \frac{2-\gamma}{2(\gamma-1)},
$$

the general value of λ is $(2 + N - N\gamma)/2(\gamma - 1)$ in N dimensions, $1 < \gamma \le (N + 2)/N$. We choose now

$$
(29) \qquad \chi(w) = \alpha(1 - \frac{|w|^2}{\beta})_+^\lambda, \varphi(w) = \delta(1 - \frac{|w|^2}{\beta})_+^{\lambda+1},
$$

where again α, β, δ are the only constants which yields (28). And we obtain the

Theorem 4. The scheme (27)-(29) satisfies
(i) $\rho_i^{n+1} \ge 0, T_i^{n+1} \ge 0$ whenever $\rho_i^n \ge 0, T_i^n \ge 0$

(ii) the singular family of conservative entropy inequalities

$$
(30) \qquad \begin{aligned} (\rho\Pi)_i^{n+1}|C_i| \le [(\rho\Pi)_i^n|C_i| - \Delta t \sum_l |E_l| G^+(U_i^n)] \\ - \Delta t \sum_l |E_l| G^-(U_{j(l)}^n) \end{aligned}
$$

(iii) the maximum principle on the specific entropy

$$
(31) \qquad S_i^{n+1} \le Max(S_i^n, S_{j(1)}^n, \dots, S_{j(l)}^n),
$$

under the CFL condition $\Delta t \sum_l |E_l|(|u_i^n| + \sqrt{\beta T_i^n}) \le |C_i|$ for all i. □

In (30), Π is defined as before by formula (14) and

$$G^{\pm}(U) = F_{\rho}^{\pm}(U)\Pi(U). \tag{32}$$

Notice that the notation $\pm$ here differs from that of the 1D case, just because of the introduction of the normals, and because we have no natural orientation for an unstructured grid. Again, the r.h.s. of (30) is composed of $L+1$ nonnegative terms and we use the convention that it is $+\infty$ whenever one of those $L+1$ terms is $+\infty$.

We skip the proof of Theorem 4 which is a straightforward extension of that in Section I. The only new point is to introduce, following [8], the kinetic scheme

$$\bar{f}_i(v)|C_i| = f_i^n(v)(|C_i| - \Delta t \sum_l (v.\nu_l)_+|E_l|)$$
$$+ \sum_l (v.\nu_l)_-|E_l|f_{j(l)}^n(v)$$

(with similar formulae acting on g), together with the condition

$$f_i^n = \rho_i^n \chi(\frac{v-u_i^n}{\sqrt{T_i^n}})/T_i^n, g_i^n = \rho_i^n \varphi(\frac{v-u_i^n}{\sqrt{T_i^n}}). \tag{33}$$

then the exponents in (29) are uniquely recovered by the requirements that, $(fg^{\gamma-2})^{1/(\gamma-1)}$ being homogeneous to $S1]_{\{\dots\}}$, the minimum in Lemma 2, but with the constraint

$$fg^{\gamma-2} \leq \Sigma^{\gamma-1},$$

is achieved for our choice of χ, φ in (29).

III. Minimal limitations for second order schemes.

We go back to the 1D case and consider second order schemes in space and time obtained using slope reconstruction (see [12]) together with a Runge-Kutta scheme in time. Our purpose is to show that only few oscillations appear (see figure 2) with the above kinetic scheme, using *centered* slopes on ρ, u, T and limited so as to preserv the nonnegativity of ρ and T as in [7]. Moreover an additional limitation ensuring the maximum principle on the specific entropy up to second order is enough to damp all oscillations (see figure 3). This amounts to a single limitation of min-mod type, combining $D\rho, DT$ for three quantities. The results are better accurate than with a min-mod limitation on the three quantities as is shown in figure 4.

III.1. The second order scheme

Let us denote by $U_i^{n,\pm}$ the inner approximations in the mesh i of $U^n(x_{i+1/2} \pm \Delta x/2)$. The construction of ΔU is discussed latter.

Then, the second order, in space and time, scheme we use is

$$\begin{cases} \tilde{U}_i - U_i^n + \sigma(F^+(U_i^{n,+}) + F^-(U_{i+1}^{n,-}) - F^+(U_{i-1}^{n,+}) - F^-(U_i^{n,-})) = 0, \\ \hat{U}_i - \tilde{U}_i + \sigma(F^+(\tilde{U}_i^+) + F^-(\tilde{U}_{i+1}^-) - F^+(\tilde{U}_{i-1}^+) - F^-(\tilde{U}_i^-)) = 0, \\ U_i^{n+1} = (U_i^n + \hat{U}_i)/2. \end{cases} \tag{34}$$

This particular Runge-Kutta scheme will preserv nonnegativity, while we would be unable to prove it with other schemes. The reason is that $\tilde{U}$ and $\hat{U}$ will have nonnegative density and temperature and then a convex combination of them, as U_i^{n+1} will also, since ρ and ρT are concave functions of U.

III.2. Nonnegativity of ρ, T and limitations.

We now prove that the scheme (34), with light limitations on a centered prediction of the derivatives, preservs nonnegativity. We use the variables ρ, u and $\Sigma = \rho^\gamma / p = S^{\gamma-1}$ and set

$$(35) \qquad \begin{cases} \Delta\rho_i = sgn(\rho_{i+1} - \rho_{i-1}) Min(|\rho_{i+1} - \rho_{i-1}|/4, \rho_i), \\ \Delta u_i = sgn(u_{i+1} - u_{i-1}) Min(|u_{i+1} - u_{i-1}|/4, \sqrt{T_i/(\gamma-1)}), \\ \Delta\Sigma_i = sgn(\Sigma_{i+1} - \Sigma_{i-1}) Min(|\Sigma_{i+1} - \Sigma_{i-1}|/4, \Sigma_i/4). \end{cases}$$

Then, following the idea introduced in [7], we set (droping the exponent n)

$$(36) \qquad \rho_i^\pm = \rho_i \pm \Delta\rho_i, u_i^\pm = \underline{u}_i \pm \Delta u_i, \Sigma_i^\pm = \underline{\Sigma}_i \pm \Delta\Sigma_i.$$

where $\underline{u}_i$ and $\underline{\Sigma}_i$ are computed for conservation of momentum and energy by

$$\rho_i^+ u_i^+ + \rho_i^- u_i^- = 2\rho_i u_i, \rho_i^+ u_i^{+2}/2 + \frac{\rho_i^{+\gamma}}{\Sigma_i^+(\gamma-1)} + \rho_i^- u_i^{-2}/2 + \frac{\rho_i^{-\gamma}}{\Sigma_i^-(\gamma-1)} = 2E_i.$$

This is readily achieved for the *second order modifications* of u_i, Σ_i given by

$$(37) \qquad \underline{u}_i = u_i - \mu \Delta u_i, \mu = \Delta\rho_i/\rho_i.$$

and $\underline{\Sigma}_i$ is the positive root of the polynomial

$$(37') \qquad C\underline{\Sigma}^2 - (A+B)\underline{\Sigma} + (B-A)\Delta\Sigma_i - C\Delta\Sigma_i^2 = 0$$

where

$$A = (\rho_i - \Delta\rho_i)^\gamma, B = (\rho_i + \Delta\rho_i)^\gamma, C = 2\rho_i T_i + \frac{1}{2}(\mu^2 - 1)(\Delta u_i)^2/(\gamma - 1).$$

we add to the centered prediction (35) the supplementary limitation

$$(37'') \qquad |\Delta\Sigma_i| \le (A+B)/2C,$$

then it is easy to check that (37') has indeed a unique non-negative root $\underline{\Sigma}_i$. Also $\rho_i^\pm \ge 0$ and thus we obtain the

Theorem 5. The scheme (34)-(37) preservs the positivity of ρ and T, under the CFL condition $(|u_i^{n,\pm}| + \sqrt{\beta T_i^{n,\pm}})\sigma \le 1/2$. □

Proof of Theorem 5. First of all let us show that $\tilde{U}$ has nonnegative density and temperature.

We use the following kinetic scheme

$$(38) \qquad \bar{f}_i - (f_i^{n,+} + f_i^{n,-})/2 + \sigma[v^+ f_i^{n,+} - v^- f_{i+1}^{n,-} - v^+ f_{i-1}^{n,+} + v^- f_i^{n,-}] = 0,$$

with the same equation for g, and

$$f_i^{n,\pm} = \rho_i^{n,\pm} \chi(\frac{(v - u_i^{n,\pm})}{\sqrt{T_i^{n,\pm}}})/\sqrt{T_i^{n,\pm}}, g_i^{n,\pm} = \rho_i^{n,\pm}\xi(\ldots)\sqrt{T_i^{n,\pm}}.$$

We claim that (38) stems, using the same combination as in Section I, the scheme (34). Indeed, we just have to check that the second term of (38) gives U_i^n ; this occurs since

$$\int_{\mathbb{R}} [(1, v, v^2/2)(f_i^{n,+} + f_i^{n,-}) + (0,0,1)(g_i^{n,+} + g_i^{n,-})]dv =$$
$$= U_i^{n,+} + U_i^{n,-} = 2U_i^n,$$

thanks to (36). Now to check the nonnegativity of $\bar{f}$, we need $1/2 \geq \sigma|v|$, which gives the CFL number of one half. □

At this level, the limitations involved in (35) are very light giving however few oscillations (figure 2). Let us go one step further and consider the maximum principle on the entropy.

III.3. Limitation by maximum of entropy.

We still denote $\Sigma = \rho^\gamma/p$ and we now require to have the maximum principle on S or Σ. Therefore we impose the additional limitation in (35), (37")

$$(39) \qquad |\Delta\Sigma_i| \leq Max(\Sigma_i, \Sigma_{i+1}, \Sigma_{i-1}) - \Sigma_i.$$

This implies a maximum principle on $\Sigma_{i\pm1/2}$ up to a second order term because $\Sigma_{i\pm1/2}$ is given through $\underline{\Sigma}_i$ and not Σ_i in (36). It seems impossible to perform second order reconstruction, satisfying the conservativity requirements (36) and the maximum principle on Σ or S.

In figure 3, we show the numerical results obtained coupling the scheme (34)-(37) to the additional limitation (39) ; the oscillations around the contact discontinuity are damped completely and it remains only an overshoot before and after the shock waves. This is also true for other tests problems : Lax shock tube, blast waves problem.

Remark : The choices of ρ and u as primitive variables for the reconstruction is somewhat arbitrary here. Only Σ plays a particular role. Let us only point out that they lead to particularly simple computations, and they are natural in the kinetic schemes.

Acknowledgment : The second author wishes to thank R. Sanders for valuable discussions.

References

[1] S. Deshpande. A second order accurate, kinetic theory based, method for inviscid compressible flows. NASA Langley Technical Paper n° 2613 (1986).

[2] A. Harten, B. Engquist, S. Osher, Sr. Chakravarthy. Uniformly high order accurate essentially non oscillatory schemes, J.Comp. Phys. Vol. 71, (1987) p. 231.

[3] P.D. Lax. Hyperbolic systems of conservation laws and the mathematical theory of shock waves, CBMS Regional Conference Series in Applied Mathematics, 11, SIAM, Philadelphia (1972).

[4] P.L. Lions, B. Perthame, E. Tadmor. Work in preparation.

[5] S. Osher. Riemann Solvers, the entropy condition, and difference approximations. SIAM J. Numer. Anal. 21 (2) (1984) p. 217-235.

[6] B. Perthame. Boltzmann type schemes for gas dynamics and the entropy property. SIAM J. Num. Anal. 27 (6) (1990), p. 1405-1421.

[7] B. Perthame. Second order Boltzmann schemes for compressible Euler equations in one and two space variables. SIAM J. Num. Anal. 29 (1) (1992).

[8] B. Perthame, Y. Qiu. A new variant of Van Leer's method for multidimensional systems of conservation laws. INRIA report n° 1562.

[9] C.W. Shu, S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. J. Comp. Phys. Vol. 83 (1989) p. 32.

[10] E. Tadmor. A minimum entropy principle in the gas dynamics equations. Appl. Num. Math. 2 (1986) p. 211-219.

[11] E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. Math. Comp. 49 (1987), p. 91-103.

[12] B. Van Leer. Towards the ultimate conservative difference scheme V, A second order sequel of Godunov's method, J. Comp. Phys. (1979), p. 101-136.

average order of the scheme / type of reconstruction	energy	density	velocity
min_mod	.77	.87	.96
centered	.79	.87	.99
centered +limitation on entropy	.87	.91	1.00

figure 4

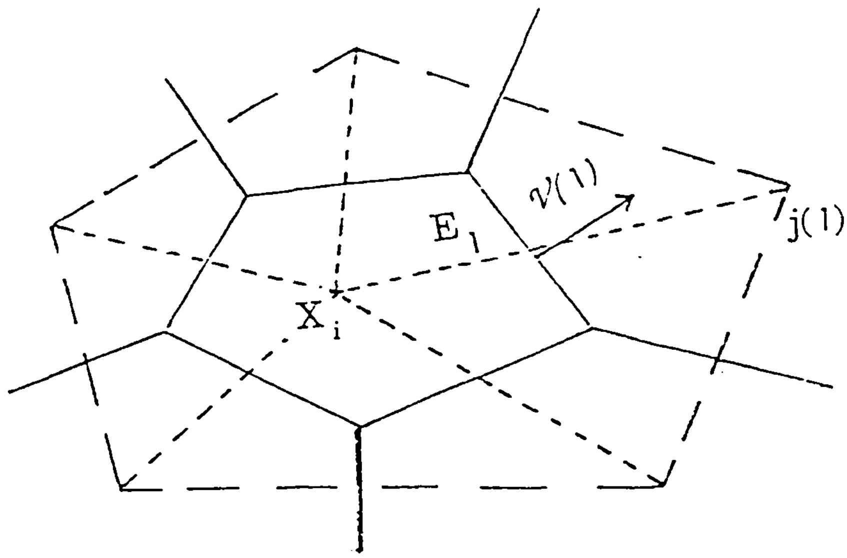
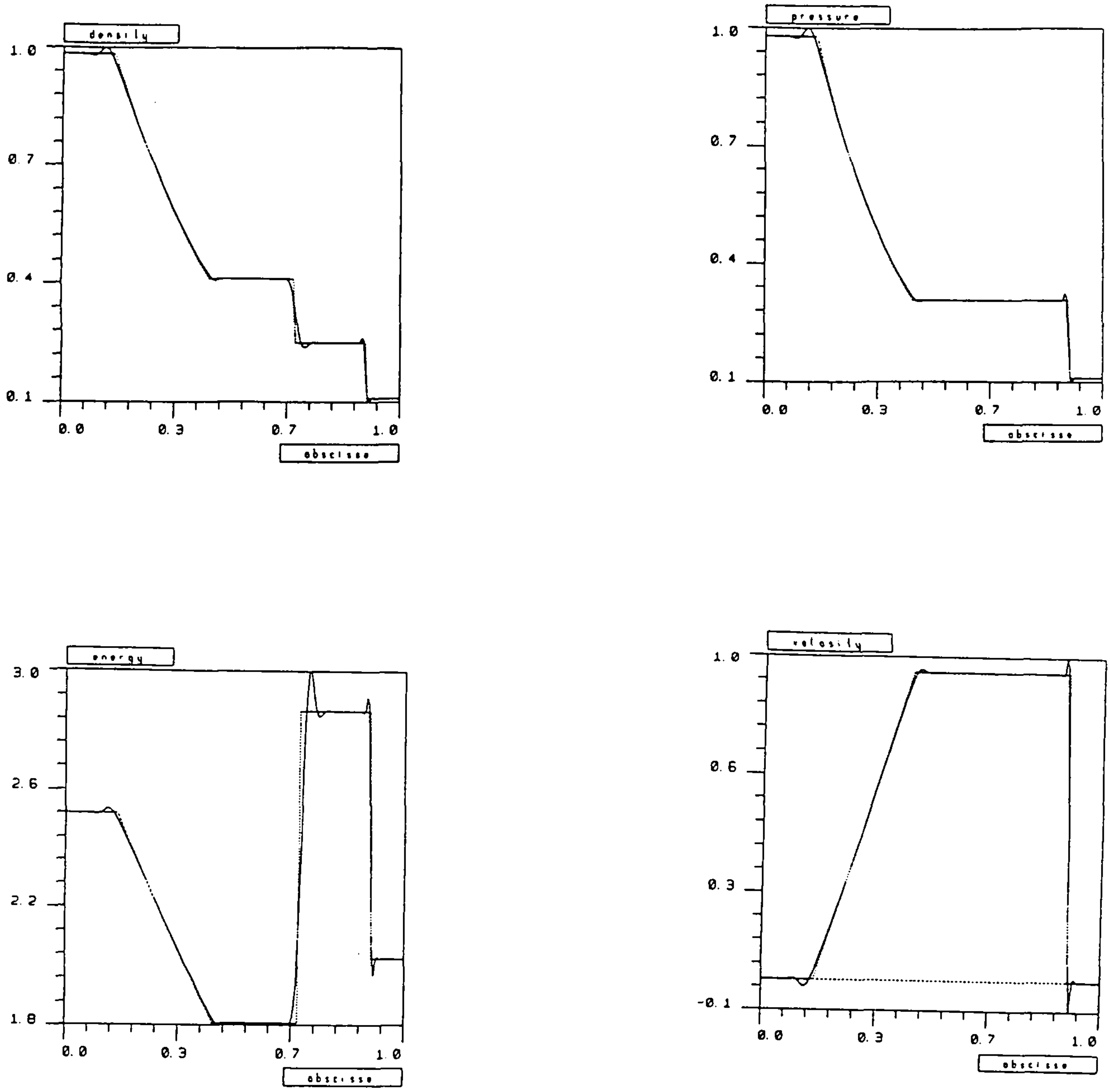


figure 1 : an example of unstructured mesh.



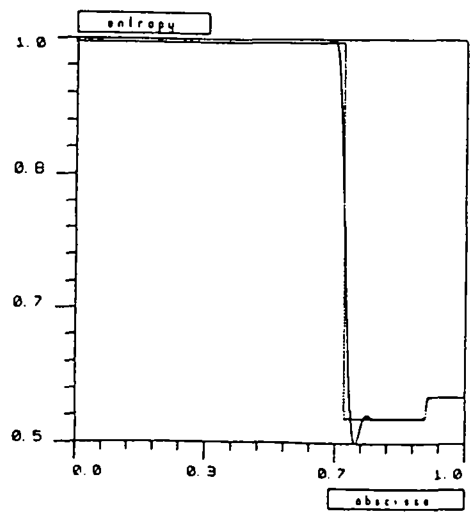
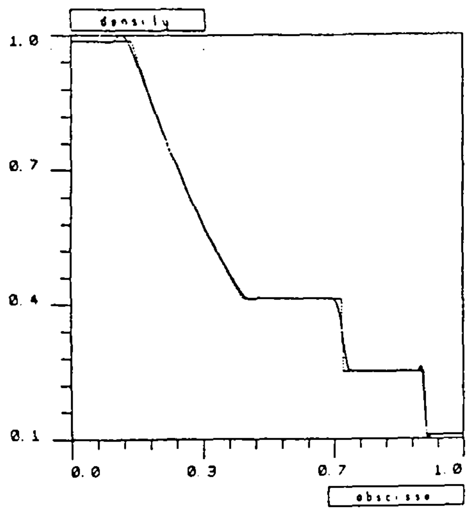
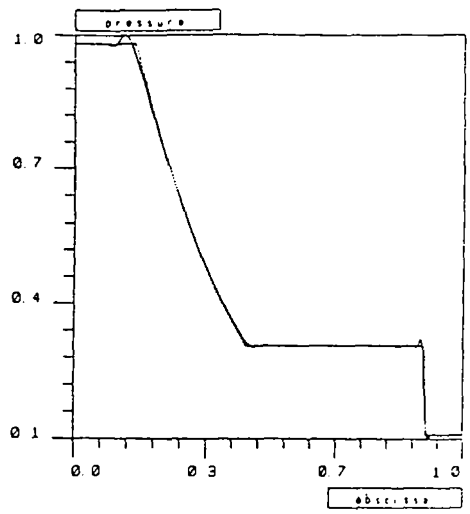
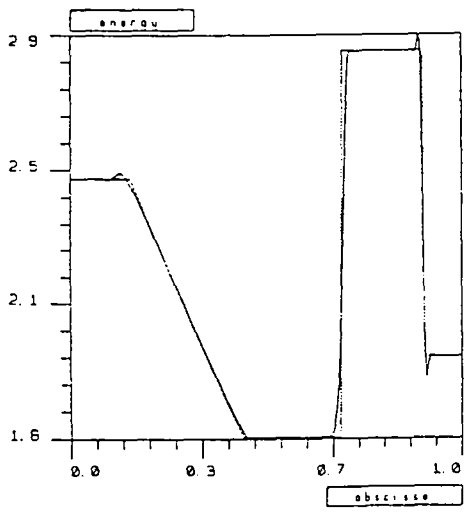
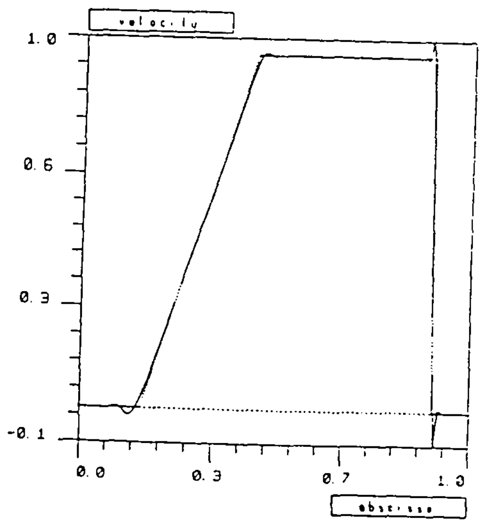


figure 2 : sod shock tube . 200 points .
second order scheme with
centered , non_limited slopes .









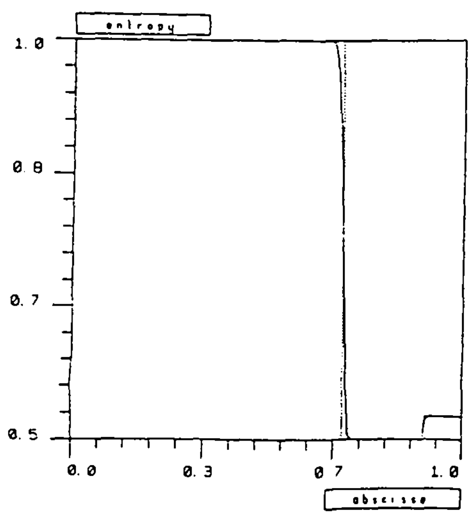


figure 3 : sod shock tube . 200 points . second order scheme with the only limitation on the entropy (6) .

ISSN 0249-6399