SIAM J. NUMER. ANAL.
Vol. 36, No. 3, pp. 738–771

CONVERGENCE OF A FINITE VOLUME EXTENSION OF THE NESSYAHU–TADMOR SCHEME ON UNSTRUCTURED GRIDS FOR A TWO-DIMENSIONAL LINEAR HYPERBOLIC EQUATION*

PAUL ARMINJON† AND MARIE-CLAUDE VIALLON‡

This paper is dedicated to the memory of Ami Harten and Haim Nessyahu.

Abstract. The nonoscillatory central difference scheme of Nessyahu and Tadmor is a Godunov-type scheme for one-dimensional hyperbolic conservation laws in which the resolution of Riemann problems at the cell interfaces is bypassed thanks to the use of the staggered Lax–Friedrichs scheme. Piecewise linear MUSCL-type (monotonic upstream-centered scheme for conservation laws) cell interpolants and slope limiters lead to an oscillation-free second-order resolution. Convergence to the entropic solution was proved in the scalar case.

After extending the scheme to a two-step finite volume method for two-dimensional hyperbolic conservation laws on unstructured grids, we present here a proof of convergence to a weak solution in the case of the linear scalar hyperbolic equation $u_t + \mathrm{div}(\vec{V}\,u) = 0$. Since the scheme is Riemann solver–free, it provides a truly multidimensional approach to the numerical approximation of compressible flows, with a firm mathematical basis.

Numerical experiments show the feasibility and high accuracy of the method.

Key words. finite volumes, staggered unstructured triangular grids, barycentric cells, MUSCL, least-squares limiters, L-infinity estimate, weak convergence, double ellipse

AMS subject classifications. 35L65, 65M12

PII. S0036142994278899

1. Introduction and description of the method.

1.1. Introduction. In recent papers [6], [7], [9], [10], we have presented for the scalar conservation equation $u_t + f(u)_x + g(u)_y = 0$ a two-step, two-dimensional finite volume method, inspired both by earlier work on unstructured triangular grids [3], [4], [1], [2] and by the nonoscillatory central differencing scheme of Nessyahu and Tadmor [25], which is a Godunov-type scheme for one-dimensional hyperbolic conservation laws, where the resolution of Riemann problems [30] at the cell interfaces is bypassed thanks to the use of the staggered form of the Lax–Friedrichs scheme; second-order oscillation-free resolution is obtained via the use of van Leer's piecewise linear MUSCL-type (monotonic upstream-centered scheme for conservation laws) cell interpolants combined with slope limiters [22], [23].

The construction of our finite volume scheme is based on a finite volume extension of the Lax–Friedrichs scheme using two specific grids at alternate time steps. Starting from an arbitrary finite element triangulation, we use the barycentric cells associated with this grid at odd time steps and a dual grid of quadrilateral cells at even time steps. Each time step can itself be viewed as a predictor-corrector process.

*Received by the editors December 16, 1994; accepted for publication (in revised form) June 17, 1998; published electronically March 30, 1999.

http://www.siam.org/journals/sinum/36-3/27889.html

†Dép. de math. et de stat. Université de Montréal, C.P. 6128, Succ. centre-ville, Montréal, PQ H3C 3J7, Canada (arminjon@crm.umontreal.ca). The research of this author was supported by a grant from the National Science and Engineering Research Council of Canada and by an FCAR grant from the government of the province of Québec.

‡Équipe d'analyse numérique, Université de Saint-Étienne, 23 rue Paul Michelon, 42023 Saint-Étienne Cedex 2, France (canon@univ-st-etienne.fr).

Results of some preliminary numerical experiments [9] using the first author's extension of the Nessyahu–Tadmor (NT) scheme to *rectangular* grids [5] (two-dimensional linear convection; discontinuous solution of Burgers' equation for discontinuous initial data, with shocks and rarefactions; diffraction of a planar shock wave around a 90º corner, Mach 3 wind tunnel with a forward facing step) confirmed the quality observed for the one-dimensional computations, while numerical experiments with the new finite volume method for *unstructured triangular* grids [7], [10] established the feasibility and high accuracy of the method. In [7], [33], we describe a comparison of our method with a discontinuous finite element method recently proposed by Jaffré and Kaddouri [18] for the problems of supersonic flow around a blunt body [28] and around a double ellipse [34].

For the one-dimensional scalar conservation law $u_t + f(u)_x = 0$, convergence to the unique entropic solution was obtained by Nessyahu and Tadmor in the case of a genuinely nonlinear equation, with the help of the total variation diminishing (TVD) property and a cell entropy inequality [25].

In this paper, we obtain an L^∞ bound which does not rely on an h-dependent limiter and, under the assumption of an h-dependent limiter, an estimate of the weighted total variation (see, e.g., [14] for similar estimates in a different context), which leads to L^∞-weak* convergence of the numerical solution to a weak solution of the linear scalar equation

$$u_t + \operatorname{div}(u\overrightarrow{V}) = 0, \qquad t \in [0,T],\ (x,y) \in \mathbb{R}^2, \tag{1.1a}$$

with initial condition

$$u(x,y,0) = u_0(x,y), \qquad (x,y) \in \mathbb{R}^2, \tag{1.1b}$$

where $u_0 \in L^\infty(\mathbb{R}^2)$ is a given function with compact support, and $\vec{V}(x,y) = \big(V_1(x,y), V_2(x,y)\big)^T$ is a given continuous vector function such that

$$\operatorname{div}\overrightarrow{V} = 0. \tag{1.1c}$$

Similar L^∞ bounds (not using h-dependent limiters) have already been given for unstructured two-dimensional grids and explicit or implicit finite volume schemes [8], [14] and for MUSCL-type finite volume schemes by Geiben-Wierse [17] and Liu [24]. In [15], Cockburn, Hou, and Shu have obtained a local maximum principle for their Runge–Kutta local projection discontinuous Galerkin methods, also defined for general triangulations.

In [36], a maximum principle for the case of rectangular grids is derived which is similar to that appearing for scalar equations on unstructured triangular grids in section 2 but uses more natural limiters. Reference [36] also contains several nontrivial numerical examples (e.g., a nonstrictly hyperbolic system).

Convergence of formally higher-order accurate MUSCL-type finite volume schemes on unstructured grids, even for nonlinear scalar conservation laws, was recently shown by Cockburn, Coquel, and Le Floch [16], Kröner, Noelle, and Rokyta [19], [20], and Noelle [26], [27]. The convergence results proved there are somewhat more general than our result since they treat the nonlinear case. Reference [16] also gives an error estimate, while [26] admits irregular families of grids, where assumption (1.2) of our paper may be relaxed; moreover, the above-mentioned papers require less restrictive CFL conditions.

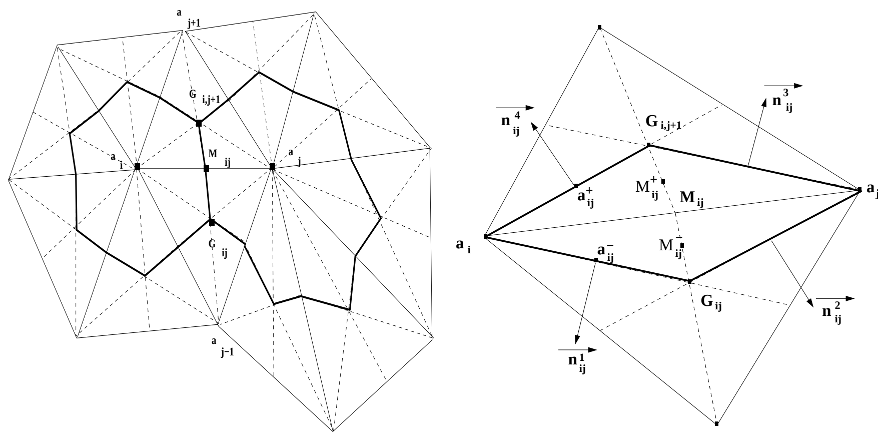

FIG. 1. *Barycentric cells around nodes* $\boldsymbol{a_i}, \boldsymbol{a_j}$*; quadrilateral cell* $\boldsymbol{a_i G_{ij} a_j G_{i,j+1}}$.

However, the additional difficulty which had to be dealt with in our work is that the scheme is a two-step scheme, which makes the analysis substantially more elaborate. The authors are currently working on an extension to nonlinear conservation equations.

In [37], the rectangular grid scheme is extended to the *incompressible* Euler equations for Cartesian grids, while in [12], [13], our finite volume method is developed into a staggered grid mixed finite volume/finite element method for the compressible Navier–Stokes equations on unstructured triangular grids, with applications to three test problems (supersonic flow around a flat plate, a NACA-0012 airfoil, and a double ellipse).

1.2. Description of the method and notation used in the paper. We introduce a triangulation $\mathcal{T}$ in $\mathbb{R}^2$ with the property that the intersection of two triangles is either empty or consists of one common vertex or side. We assume that there exist four positive constants a, b, c, and d such that the usual finite element nondegenerescence conditions

$$\begin{cases} ah \ \le \ell(I) \ \le bh & \text{for every side } I, \\ ch^2 \le A(K) \le dh^2 & \text{for every triangle } K \in \mathcal{T} \end{cases} \tag{1.2}$$

hold, where $\ell(I)$ denotes the length of side I and $A(K)$ the area of triangle K, respectively.

The two-dimensional generalization [6], [7] of the NT scheme is a two-step finite volume scheme defined with the help of two alternate grids. For the first grid, the nodes are the vertices a_i of the triangles $K \in \mathcal{T}$, and the finite volume cells are the barycentric cells C_i, obtained by joining the midpoints M_{ij} of the sides originating at node a_i to the centroids G_{ij} of the triangles of $\mathcal{T}$ which meet at a_i (Figure 1). For the second grid the nodes are the midpoints M_{ij} of the sides, while the cells are the quadrilaterals $L_{ij} = a_i G_{ij} a_j G_{i,j+1}$ obtained by joining two nodes a_i, a_j to the centroids of the two triangles of $\mathcal{T}$ of which $a_i a_j$ is a side. We use the following notation.

NOTATION.

a_i is the ith vertex.

M_{ij} is the midpoint of side $a_i a_j$.

n_i is the number of the nodes which are adjacent to a_i.
$G_{ij}(j = 1, \ldots, n_i)$ is the centroid of a triangle of which a_i is a vertex.
C_i is the barycentric cell constructed around a_i.
Γ_{ij} is the cell boundary element $G_{ij}M_{ij}G_{i,j+1}$.
$\partial C_i = \bigcup_{j=1}^{n_i} \Gamma_{ij}$ is the boundary of cell C_i.
L_{ij} is the quadrilateral cell with vertices a_i, G_{ij}, a_j, $G_{i,j+1}$:

$$q_{ij} = \frac{A(a_i G_{ij} M_{ij})}{A(L_{ij} \cap C_i)}, \quad r_{ij} = \frac{A(L_{ij} \cap C_i)}{A(C_i)}. \tag{1.3}$$

The unknowns are u_i^n, the numerical approximation of the exact value $u(a_i, t^n)$ at node a_i and time t^n $(n = 0, 2, 4, \ldots)$, and u_{ij}^{n+1}, the numerical approximation of $u(M_{ij}, t^{n+1})$, for each node index i and every j "neighbor of i." We choose a constant time step with $t^n = n\Delta t$, for $0 \leq n \leq L$ with $t^L = L\Delta t = T$.

To initialize the time marching process we let

$$u_i^0 = \frac{1}{A(C_i)} \iint_{C_i} u_0(x, y) dx dy. \tag{1.4}$$

The solution $u(x, y, t)$ of the Cauchy problem (1.1) is approximated by a cellwise, piecewise linear function. At time t^n (n even), starting from the known values u_i^n, we introduce for each cell C_i an *approximate gradient* $\overrightarrow{\Delta}_i^n$ (satisfying some specific conditions to be described later), and at every point $M(x, y)$ of cell C_i we define

$$u(x, y, t^n) \equiv u_{C_i}(x, y, t^n) = u_{C_i}^n(x, y) = u_i^n + \overrightarrow{a_i M} \cdot \overrightarrow{\Delta}_i^n \ (n = 0, 2, 4, \ldots). \tag{1.5}$$

Integrating this linear function on the quadrilateral cell L_{ij} leads to the first (and further odd-numbered) time step of the scheme:

$$u_{ij}^{n+1} = \frac{1}{2}(u_i^n + u_j^n) + \frac{1}{6}(\overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i^n} + \overrightarrow{\mathrm{Vect}_j} \cdot \overrightarrow{\Delta_j^n}) - \frac{\Delta t}{A(L_{ij})} \sum_{I \in \partial L_{ij}} u_I^{n+1/2} V(I), \tag{1.6}$$

where

$$\overrightarrow{\mathrm{Vect}_i} = \overrightarrow{a_i M_{ij}} + q_{ij} \overrightarrow{a_i G_{ij}} + (1 - q_{ij}) \overrightarrow{a_i G_{i,j+1}} = 2q_{ij} \overrightarrow{a_i M_{ij}^-} + 2(1 - q_{ij}) \overrightarrow{a_i M_{ij}^+}, \tag{1.7}$$

M_{ij}^- = midpoint of $G_{ij}M_{ij}$,
M_{ij}^+ = midpoint of $M_{ij}G_{i,j+1}$,
$\partial L_{ij} = \{a_i G_{ij},\ a_j G_{ij},\ a_j G_{i,j+1},\ a_i G_{i,j+1}\}$,
$\partial L_{ij} \cap C_i = \{a_i G_{ij},\ a_i G_{i,j+1}\}$,
$\partial L_{ij} \cap C_j = \{a_j G_{ij},\ a_j G_{i,j+1}\}$,
I is a side of quadrilateral L_{ij},
$\overrightarrow{n_I}$ = unit outer normal to L_{ij}, for $I \in \partial L_{ij}$,
$\overrightarrow{V_I}$ = average value of $\vec{V}$ along the side $I \in \partial L_{ij}$,
$V(I) = \overrightarrow{V_I} \cdot \overrightarrow{n_I}\ \ell(I)$ with $\ell(I)$ = length of I,

and

$$u_I^{n+1/2} = \begin{cases} u_i^n + \frac{1}{2}(\overrightarrow{I} - \Delta t \overrightarrow{V_I}) \cdot \overrightarrow{\Delta_i^n} & \text{if } I \in (\partial L_{ij}) \cap C_i, \\ u_j^n + \frac{1}{2}(\overrightarrow{I} - \Delta t \overrightarrow{V_I}) \cdot \overrightarrow{\Delta_j^n} & \text{if } I \in (\partial L_{ij}) \cap C_j. \end{cases} \tag{1.8}$$

In preparation for the second (even) time step, we now compute for each cell L_{ij} an approximate gradient $\overrightarrow{\Delta_{ij}^{n+1}}$ (satisfying specific conditions to be described later), and at each point $M(x, y)$ of the quadrilateral cell L_{ij} we define

$$u(x, y, t^{n+1}) = u^{n+1}(x, y) = u_{ij}^{n+1} + \overrightarrow{M_{ij}M} \cdot \overrightarrow{\Delta_{ij}^{n+1}} \quad (n = 0, 2, 4, \dots). \tag{1.9}$$

Integrating this linear function on the barycentric cell C_i leads to the second (even) time step of the scheme:

$$u_i^{n+2} = \sum_{j=1}^{n_i} r_{ij} u_{ij}^{n+1} + \frac{1}{3} \sum_{j=1}^{n_i} r_{ij} \overrightarrow{\mathrm{Vect}_{ij}} \cdot \overrightarrow{\Delta_{ij}^{n+1}} - \frac{\Delta t}{A(C_i)} \sum_{I \in \partial C_i} u_I^{n+3/2} V(I), \tag{1.10}$$

where

$$\begin{cases} \overrightarrow{\mathrm{Vect}_{ij}} & = \overrightarrow{M_{ij}a_i} + q_{ij}\ \overrightarrow{M_{ij}G_{ij}} + (1 - q_{ij})\ \overrightarrow{M_{ij}G_{i,j+1}} \\ & = 2q_{ij}\ \overrightarrow{M_{ij}a_{ij}^-} + 2(1 - q_{ij})\overrightarrow{M_{ij}a_{ij}^+}\ , \end{cases} \tag{1.11}$$

a_{ij}^- = midpoint of $a_i G_{ij}$,
a_{ij}^+ = midpoint of $a_i G_{i,j+1}$,
$\partial C_i = \{G_{ij}M_{ij}, M_{ij}G_{i,j+1}, j = 1, \dots, n_i\}$,
$(\partial C_i) \cap L_{ij} = \{G_{ij}M_{ij}, M_{ij}G_{i,j+1}\}$,
$\overrightarrow{n_I}$ = unit outer normal to C_i for $I \in \partial C_i$,
$\overrightarrow{V_I}$ = average value of $\overrightarrow{V}$ along the side $I \in \partial C_i$,
$V(I) = \overrightarrow{V_I} \cdot \overrightarrow{n_I}\ \ell(I)$,
and

$$u_I^{n+3/2} = u_{ij}^{n+1} + \frac{1}{2}(\overrightarrow{I} - \Delta t \overrightarrow{V_I}) \cdot \overrightarrow{\Delta_{ij}^{n+1}} \text{ for } I \in (\partial C_i) \cap L_{ij}. \tag{1.12}$$

The numerical solution of (1.1) is then defined by

$$u_{\mathcal{T},\Delta t}(x, y, t) = u(x, y, t^n) \text{ for } t^n \le t < t^{n+1}, \tag{1.13}$$

where $u(x, y, t^n)$ is given by (1.5) (n even) and (1.9) (n odd), respectively.

In section 2, we prove that if we consider a sequence $\{\mathcal{T}_k, \Delta t_k\}_{k \in \mathbb{N}}$ such that $\mathcal{T}_k$ satisfies (1.2), with $h = h_k$, where h_k and Δt_k tend to zero while $\frac{\Delta t_k}{h_k}$ remains bounded (CFL-like condition), the corresponding sequence of approximate solutions $\{u_{\mathcal{T}_k,\Delta t_k}\}$ defined by (1.13) is then bounded in $L^\infty(\mathbb{R}^2 \times \mathbb{R}^+)$.

Therefore there exists a subsequence, again written $\{u_{\mathcal{T}_k,\Delta t_k}\}$, which converges to some function u in $L^\infty(\mathbb{R}^2 \times \mathbb{R}^+)$-weak*.

In section 3, we obtain a so-called "weighted total variation" estimate (cf. [14]), weaker than an estimate on the total variation of the numerical solution but sufficient to prove (section 4) that the limit u of the above subsequence is indeed a weak solution of problem (2.1).

As correctly observed by one referee, the limiters used in the convergence proof allow no variation within cell L_{ij} if $u_i - u_j = 0$.

This might lead to a substantial loss of accuracy in problems where, e.g., the triangulation makes many ij segments parallel to the x-direction while the exact solution is x-independent. But these limiters are never used in practice. Limiters used in *actual numerical simulations*, described in [7], [33], are much less severe than those introduced in sections 2–4 *to prove convergence*, so that the overall accuracy of our method is not exposed to the degradation which would result from the use of these "theoretical" limiters.

While we intend in the future to try to design a limiter following, e.g., [15], [17], which would at the same time be truly multidimensional (as the limiters we use in [7], [33]) and allow a very high level of accuracy, we have concentrated here on a limiter which makes the convergence proof more accessible. In fact, in section 2 we give an example of one such possible choice of limiter, less restrictive than the one we use in the convergence proof presented here (Remark 2.5.5).

In section 5, we present a systematic comparison of our method with a discontinuous finite element method developed at INRIA [18] by Jaffré and Kaddouri, in a typical test selected from the numerical experiments described in [7], [10], [33], which include several comparisons with other methods and give a rather favorable overview of the properties of our method: whenever a comparison was possible, the capture of shocks was sharper, without breach of monotonicity, and the convergence history much faster; finally, the computing times were also significantly shorter.

2. An L^∞-estimate of the numerical solution. We shall prove that under an appropriate CFL condition, u_{ij}^{n+1} is a convex combination of u_i^n and u_j^n (n even) and u_i^{n+2} is a convex combination of the values u_{ij}^{n+1} at all adjacent midpoints M_{ij} $(1 \le j \le n_i)$. This will imply that

$$\|u^n\|_{L^\infty(\mathbb{R}^2)} \;\le\; \| \, u^0 \, \|_{L^\infty(\mathbb{R}^2)} < \infty,$$

since we have assumed that $u_0 \in L^\infty(\mathbb{R}^2)$.

2.1. Analysis of the first step of the scheme. The first step of the scheme consists in writing u_{ij}^{n+1} as a function of u_i^n and u_j^n according to (1.6). In order to prove that $|u_{ij}^{n+1}| \le \max\{|u_i^n|, |u_j^n|\}$, we shall write u_{ij}^{n+1} as a convex combination of u_i^n and u_j^n.

Let us first factor out u_i^n and u_j^n in (1.6) by multiplying (1.6) by $(u_j^n - u_i^n)/(u_j^n - u_i^n)$:

(2.1a)
$$\begin{aligned} u_{ij}^{n+1} = u_i^n &\left[\frac{1}{2} - \left(\frac{1}{6} \frac{\overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i^n} + \overrightarrow{\mathrm{Vect}_j} \cdot \overrightarrow{\Delta_j^n}}{u_j^n - u_i^n} - \frac{\Delta t}{A(L_{ij})} \sum_{I \in \partial L_{ij}} \frac{u_I^{n+1/2}}{u_j^n - u_i^n} V(I) \right) \right] \\ &+ u_j^n \left[\frac{1}{2} + \left(\frac{1}{6} \frac{\overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i^n} + \overrightarrow{\mathrm{Vect}_j} \cdot \overrightarrow{\Delta_j^n}}{u_j^n - u_i^n} - \frac{\Delta t}{A(L_{ij})} \sum_{I \in \partial L_{ij}} \frac{u_I^{n+1/2}}{u_j^n - u_i^n} V(I) \right) \right]. \end{aligned}$$

u_{ij}^{n+1} will therefore be a convex combination of u_i^n and u_j^n provided that

$$\text{(2.1b)} \qquad \left| \frac{1}{6} \frac{\overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i^n} + \overrightarrow{\mathrm{Vect}_j} \cdot \overrightarrow{\Delta_j^n}}{u_j^n - u_i^n} - \frac{\Delta t}{A(L_{ij})} \sum_{I \in \partial L_{ij}} \frac{u_I^{n+1/2}}{u_j^n - u_i^n} V(I) \right| \le \frac{1}{2}.$$

In the rest of this paragraph, we shall show that inequality (2.1b) holds under a CFL-type condition and some appropriate slope limitation for $\overrightarrow{\Delta_i^n}$. For simplicity we rewrite (2.1.b) as

$$\text{(2.1c)} \qquad \left| \frac{1}{6} T_1 - \frac{\Delta t}{A(L_{ij})} T_2 \right| \le \frac{1}{2}$$

and omit the time index n.

2.1.1. Estimate for T_1. We introduce a slope limitation for the gradients $\overrightarrow{\Delta}_i$ of the piecewise linear approximation (1.5):

$$\left\| \frac{\overrightarrow{\Delta}_i}{u_j - u_i} \right\| \leq \frac{6}{7b} \cdot \frac{\varepsilon}{h}, \qquad \varepsilon \geq 0, \qquad j \text{ neighbor of } i, \tag{2.2a}$$

$$\frac{\overrightarrow{a_i M_{ij}^-} \cdot \overrightarrow{\Delta}_i}{u_j - u_i} \geq 0, \qquad \frac{\overrightarrow{a_i M_{ij}^+} \cdot \overrightarrow{\Delta}_i}{u_j - u_i} \geq 0, \qquad j \text{ neighbor of } i. \tag{2.2b}$$

LEMMA 2.1. *Under conditions* (2.2), *we have* $|T_1| \leq \varepsilon$.

Proof. In order to interpret the slope limitation conditions (2.2), we observe that the value of our linear interpolant (1.5) in cell C_i at the point $M = M_{ij}^-$ (midpoint of $G_{ij} M_{ij}$) is given by

$$u_{C_i}(M_{ij}^-) = u_i + \overrightarrow{a_i M_{ij}^-} \cdot \overrightarrow{\Delta_i}. \tag{2.3a}$$

Condition (2.2b) then means that

$$\frac{u_{C_i}(M_{ij}^-) - u_i}{u_j - u_i} \geq 0, \tag{2.3b}$$

while condition (2.2a) implies

$$\frac{u_{C_i}(M_{ij}^-) - u_i}{u_j - u_i} \leq \|\overrightarrow{a_i M_{ij}^-}\| \frac{6}{7b} \cdot \frac{\varepsilon}{h}. \tag{2.3c}$$

On the other hand we have from (1.2)

$$\|\overrightarrow{a_i M_{ij}^-}\| \leq \frac{1}{2}\big(\|\overrightarrow{a_i G_{ij}}\| + \|\overrightarrow{a_i M_{ij}}\|\big) \leq \frac{1}{2}\left(\frac{2}{3} + \frac{1}{2}\right) bh = \frac{7}{12} bh$$

and therefore

$$0 \leq \frac{u_{C_i}(M_{ij}^-) - u_i}{u_j - u_i} \leq \frac{\varepsilon}{2} \tag{2.3d}$$

(with the same bounds for M_{ij}^+).

Choosing, for instance, $\varepsilon = 2$ then forces the piecewise linear cell values at M_{ij}^- and M_{ij}^+ to lie between u_i and u_j. (Specific conditions on ε will be described later.)

Writing T_1 as a difference,

$$T_1 = \frac{\overrightarrow{\text{Vect}_i} \cdot \overrightarrow{\Delta_i}}{u_j - u_i} - \frac{\overrightarrow{\text{Vect}_j} \cdot \overrightarrow{\Delta_j}}{u_i - u_j} \equiv T_{11} - T_{12}, \tag{2.4a}$$

where, by (1.7),

$$T_{11} = \frac{2\, q_{ij}\, \overrightarrow{a_i M_{ij}^-} \cdot \overrightarrow{\Delta_i} + 2(1 - q_{ij}) \overrightarrow{a_i M_{ij}^+} \cdot \overrightarrow{\Delta_i}}{u_j - u_i}, \tag{2.4b}$$

we see from (2.3) that $0 \leq T_{11} \leq \varepsilon$, and the same inequalities hold for T_{12}. T_1 is therefore the difference of two positive numbers each of which is less than ε. We conclude that $|T_1| \leq \varepsilon$. □

2.1.2. Estimate for T_2. Let $\|\overrightarrow{V}\|_\infty \equiv \sup\{\|\overrightarrow{V}_I\| \text{ for } I \in \partial L_{ij} \text{ or } I \in \partial C_i, \text{ for arbitrary nodes } i, j\}$. We introduce the following CFL condition:

$$\frac{\Delta t}{h}\|\overrightarrow{V}\|_\infty \leq \beta, \quad \beta > 0. \tag{2.5}$$

(Appropriate conditions on β will be specified later.)

LEMMA 2.2. *If* (2.2) *and the CFL condition* (2.5) *are satisfied, then*

$$\frac{\Delta t}{A(L_{ij})}|T_2| \leq \frac{2b}{c}\beta\Big(1 + \frac{4}{7}\varepsilon + \frac{6}{7b}\varepsilon\beta\Big).$$

Proof. From the definition introduced between (1.7) and (1.8), we have

$$\sum_{I \in \partial L_{ij}} V(I) = \sum_{I \in \partial L_{ij}} \overrightarrow{V}_I \cdot \overrightarrow{n_I}\ \ell(I) = \int_{\partial L_{ij}} \overrightarrow{V} \cdot \overrightarrow{n}\, d\sigma = \int_{L_{ij}} \operatorname{div} \vec{V}\, dA = 0,$$

since we have assumed that $\operatorname{div} \overrightarrow{V} = 0$; T_2 can thus be rewritten as

$$T_2 = \sum_{I \in \partial L_{ij}} \frac{u_I^{n+1/2} - u_i^n}{u_j^n - u_i^n} V(I) = \sum_{I \in (\partial L_{ij}) \cap C_i} \frac{1}{2} \frac{(\overrightarrow{I} - \Delta t \overrightarrow{V_I}) \cdot \overrightarrow{\Delta_i^n}}{u_j^n - u_i^n} V(I)$$
$$+ \sum_{I \in (\partial L_{ij}) \cap C_j} \left(1 + \frac{1}{2} \frac{(\overrightarrow{I} - \Delta t \overrightarrow{V_I}) \cdot \overrightarrow{\Delta_j^n}}{u_j^n - u_i^n}\right) V(I)$$

by (1.8) and after inserting $-u_j^n + u_j^n$ in the numerator of the summation on $\partial L_{ij} \cap C_j$. Applying (1.2), (2.2a), and the triangular inequality now gives

$$|T_2| \leq \left(4 \cdot \frac{1}{2}\left(\frac{2}{3}bh + \Delta t\|\overrightarrow{V}\|_\infty\right)\frac{6}{7b}\frac{\varepsilon}{h} + 2\right)\frac{2}{3}bh\|\overrightarrow{V}\|_\infty$$

so that (1.2) and (2.5) finally lead to

$$\frac{\Delta t}{A(L_{ij})}|T_2| \leq \frac{2b}{c}\beta\left(\frac{4\varepsilon}{7} + \frac{6\varepsilon\beta}{7b} + 1\right). \qquad \square$$

Remark 2.3. Condition (2.2a) can be replaced by a condition on the relative increments

$$\frac{u_{C_i}^n(M_{ij}^-) - u_i^n}{u_j^n - u_i^n}, \quad \frac{u_{C_i}^n(M_{ij}^+) - u_i^n}{u_j^n - u_i^n}, \quad \frac{u_I^{n+1/2} - u_i^n}{u_j^n - u_i^n} \quad \text{for } I \in \partial L_{ij}.$$

2.1.3. Conclusion.

LEMMA 2.4. *If conditions* (1.2), (2.2)*, and* (2.5) *hold and if* ε *and* β *satisfy*

$$P(\varepsilon, \beta) \equiv \frac{12}{7c}\varepsilon\beta^2 + \left(1 + \frac{4}{7}\varepsilon\right)\frac{2b}{c}\beta + \frac{\varepsilon - 3}{6} \leq 0, \tag{2.6}$$

then

$$|u_{ij}^{n+1}| \leq \max\{|u_i^n|, |u_j^n|\}. \tag{2.7}$$

Proof. From the remarks at the beginning of section 2, we know that (2.7) will be satisfied if (2.1c) holds. This follows directly from Lemmas 2.1 and 2.2. $\square$

Remark 2.5.

1. The case when $\varepsilon = 0$, i.e., $\overrightarrow{\Delta_i^n} = 0$, corresponds to our finite volume two-dimensional extension of the Lax–Friedrichs scheme [7], [10], for which condition (2.6) takes the form

$$P(\varepsilon,\beta) = \frac{2b}{c}\beta - \frac{1}{2} \le 0. \tag{2.8a}$$

In view of (2.5), this is a CFL condition:

$$\frac{\Delta t \|\overrightarrow{V}\|_\infty}{h} \le \frac{1}{4}\frac{c}{b}. \tag{2.8b}$$

2. The desired bounds for u_{ij}^{n+1} for the two-dimensional NT scheme are obtained if there exist $\varepsilon > 0$ and $\beta > 0$ such that $P(\varepsilon,\beta) \le 0$. If we consider the roots β_1 and β_2 of the quadratic polynomial $P(\cdot,\beta)$, then the product $\beta_1\beta_2$ is equal to $\frac{7c}{72}\cdot\frac{\varepsilon-3}{\varepsilon}$, while $\beta_1+\beta_2 = 7b(1+\frac{4}{7}\varepsilon)/6\varepsilon < 0$. A positive solution $\beta > 0$ of (2.6) will exist if and only if the discriminant of $P(\varepsilon,\beta)$ is positive and at least one of its roots is strictly positive, which will be true if $\varepsilon < 3$. On the other hand a solution of (2.6) with $\varepsilon > 0$ can clearly exist only if $\beta < \frac{1}{4}\frac{c}{b}$.

3. Condition (2.2b) can be omitted; we then get the bound $|T_1| \le 2\varepsilon$ instead of $|T_1| \le \varepsilon$, since the signs of the terms T_{11} and T_{12} in (2.4) are no longer necessarily the same. We then obtain a solution $\beta > 0$ if and only if $\varepsilon < \frac{3}{2}$.

4. Condition (2.2) can be replaced by a condition providing explicit bounds for the value of the numerical solution on the boundary of cell C_i:

$$0 \le \frac{u_{C_i}(M) - u_i}{u_j - u_i} \le \frac{\varepsilon}{2} \text{ for } M \in \left\{G_{ij}, M_{ij}, G_{i,j+1}\right\},\ j \text{ neighbor of } i, \tag{2.2$'$}$$

where

$$u_{C_i}(M) = u_i + \overrightarrow{a_iM}\cdot\overrightarrow{\Delta_i}.$$

Writing (2.2) in the form

$$0 \le \frac{\overrightarrow{a_iM}\cdot\overrightarrow{\Delta_i}}{u_j - u_i} = \|\overrightarrow{a_iM}\| \left\|\frac{\overrightarrow{\Delta_i}}{u_j-u_i}\right\| \cos(\overrightarrow{a_iM},\overrightarrow{\Delta_i}) \le \frac{\varepsilon}{2} \tag{2.2$''$}$$

and using a lower bound for the cosine obtained from (1.2), we get

$$\left\|\frac{\overrightarrow{\Delta}_i}{u_j - u_j}\right\| \le \frac{b^2}{ac}\frac{\varepsilon}{h}.$$

The inequality in Lemma 2.2 then takes the form

$$\frac{\Delta t}{A(L_{ij})}|T_2| \le \frac{b}{c}\beta\left(2+\varepsilon+\frac{2b^2}{ac}\beta\varepsilon\right),$$

which leads to the condition

$$P(\varepsilon,\beta) = \frac{2b^3}{ac^2}\varepsilon\beta^2 + (2+\varepsilon)\frac{b}{c}\beta + \frac{\varepsilon-3}{6} \le 0$$

in Lemma 2.4.

5. One drawback of limiter (2.2), with its consequence (2.3d), is that it allows no variation in the whole cell L_{ij} if $u_j = u_i$. It is possible to allow the solution to vary in L_{ij} by modifying the limiter as follows.

In view of (1.7), we can limit the variation of the cell value of u at M_{ij}, G_{ij}, and $G_{i,j+1}$ instead of M_{ij}^- and M_{ij}^+. For example, we can impose

$$\begin{cases} 0 \le \frac{u_{C_i}(M_{ij})-u_i}{u_j-u_i} \le \frac{\varepsilon}{2}, \\ u_{C_i}(G_{ij}) - u_i = \lambda_1(u_j - u_i) + \lambda_2(u_{j-1} - u_i), \qquad \lambda_1, \lambda_2 \ge 0\ ,\ \lambda_1 + \lambda_2 \le \frac{\varepsilon}{2}, \end{cases}$$

and some complementary conditions, which corresponds, if $0 \le \varepsilon \le 2$, to forcing $u_{C_i}(M_{ij})$ to be comprised between u_i and u_j and the value $u_{C_i}(G_{ij})$ to fall within the interval determined by $u_i,\ u_{j-1},\ u_j$. This would then lead to a maximum principle in the form

$$|u_{ij}^{n+1} \le \max\{|u_i|, |u_{j-1}|, |u_j|, |u_{j+1}|\}.$$

2.2. Analysis of the second step of the scheme. The second step of the scheme consists of writing u_i^{n+2} (n even) as a function of the neighboring values $u_{ij}^{n+1}(1 \le j \le n_i)$ at time t^{n+1}, with the help of (1.10). To show that $|u_i^{n+2}| \le \max_{1\le j\le n_i}\{|u_{ij}^{n+1}|\}$, we shall write u_i^{n+2} as a convex combination of the values u_{ij}^{n+1}. Starting from (1.10) and (1.11), we get

$$(2.9)\quad u_i^{n+2} = \sum_{j=1}^{n_i} r_{ij}u_{ij}^{n+1} + \frac{1}{3}\sum_{j=1}^{n_i} r_{ij}(2q_{ij}\overrightarrow{M_{ij}a_{ij}^-} + 2(1-q_{ij})\overrightarrow{M_{ij}a_{ij}^+})\cdot\overrightarrow{\Delta_{ij}^{n+1}}$$
$$- \frac{\Delta t}{A(C_i)}\sum_{j=1}^{n_i}\big(u_{M_{ij}G_{ij}}^{n+3/2}V(M_{ij}G_{ij}) + u_{M_{ij}G_{i,j+1}}^{n+3/2}V(M_{ij}G_{i,j+1})\big).$$

Multiplying the two terms depending on G_{ij} by $(u_{i,j-1}^{n+1} - u_{ij}^{n+1})\,/\,(u_{i,j-1}^{n+1} - u_{ij}^{n+1})$ and the terms depending on $G_{i,j+1}$ by $(u_{i,j+1}^{n+1} - u_{ij}^{n+1})\,/\,(u_{i,j+1}^{n+1} - u_{ij}^{n+1})$, we can factor out u_{ij}^{n+1} (whereby the summation index j is being shifted; the time index n is partly omitted for simplicity):

$$(2.10a)\quad u_i^{n+2} = \sum_{j=1}^{n_i}\frac{1}{3}u_{ij}^{n+1}\Bigg\{3r_{ij} + 2r_{ij}q_{ij}\frac{\overrightarrow{M_{ij}a_{ij}^-}\cdot\overrightarrow{\Delta_{ij}}}{u_{ij}-u_{i,j-1}} + 2r_{ij}(1-q_{ij})\frac{\overrightarrow{M_{ij}a_{ij}^+}\cdot\overrightarrow{\Delta_{ij}}}{u_{ij}-u_{i,j+1}}$$
$$+ 2r_{i,j+1}q_{i,j+1}\frac{\overrightarrow{M_{i,j+1}a_{i,j+1}^-}\cdot\overrightarrow{\Delta_{i,j+1}}}{u_{ij}-u_{i,j+1}} + 2r_{i,j-1}(1-q_{i,j-1})\frac{\overrightarrow{M_{i,j-1}a_{i,j-1}^+}\cdot\overrightarrow{\Delta_{i,j-1}}}{u_{ij}-u_{i,j-1}}$$
$$-\frac{3}{2}\,\frac{\Delta t}{A(C_i)}\Bigg(\frac{(\overrightarrow{I}-\Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_{ij}}}{u_{ij}-u_{i,j-1}}V(I)\Big|_{I=M_{ij}G_{ij}} + \frac{(\vec{I}-\Delta t\vec{V}_I)\cdot\overrightarrow{\Delta_{i,j+1}}}{u_{ij}-u_{i,j+1}}V(I)\Big|_{I=M_{i,j+1}G_{i,j+1}}$$
$$+ \frac{(\overrightarrow{I}-\Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_{ij}}}{u_{ij}-u_{i,j+1}}V(I)\Big|_{I=M_{ij}G_{i,j+1}} + \frac{(\overrightarrow{I}-\Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_{i,j-1}}}{u_{ij}-u_{i,j-1}}V(I)\Big|_{I=M_{i,j-1}G_{ij}}\Bigg)$$
$$- 3\frac{\Delta t}{A(C_i)}\big[V(M_{ij}G_{ij}) + V(M_{ij}G_{i,j+1})\big]\Bigg\}.$$

For quick reference to the individual terms of this expression, we rewrite (2.10a) as

$$(2.10b)\qquad u_i^{n+2} = \sum_{j=1}^{n_i}\frac{1}{3}u_{ij}^{n+1}\left\{3r_{ij} - T_3 - \frac{3}{2}\,\frac{\Delta t}{A(C_i)}T_4 - 3\frac{\Delta t}{A(C_i)}T_5\right\}.$$

The sum of the brackets for $j=1,\dots,n_i$ is equal to 3 since $\sum_{j=1}^{n_i} r_{ij}=1$, $\sum_{j=1}^{n_i} T_3 = \sum_{j=1}^{n_i} T_4 = 0$ due to pairwise cancellations, and $\sum_{j=1}^{n_i} T_5 = 0$ since div $\overrightarrow{V} = 0$. Therefore u_i^{n+2} will be a convex combination of the values u_{ij}^{n+1} at the neighboring staggered cells L_{ij}, provided that

$$T_3 + \frac{3}{2}\,\frac{\Delta t}{A(C_i)} T_4 + 3\,\frac{\Delta t}{A(C_i)} T_5 \le \; 3 r_{ij}. \tag{2.11}$$

2.2.1. Estimate for T_3. We introduce slope limitations for the gradients $\overrightarrow{\Delta}_{ij}$:

$$\left\| \frac{\overrightarrow{\Delta}_{ij}}{u_{ij}-u_{i,j-1}} \right\| \le \frac{2}{5b}\cdot\frac{\gamma}{h}\,, \quad \left\| \frac{\overrightarrow{\Delta}_{ij}}{u_{ij}-u_{i,j+1}} \right\| \le \frac{2}{5b}\cdot\frac{\gamma}{h}, \gamma \ge 0,\ j \text{ neighbor of } i, \tag{2.12}$$

from which we derive the following lemma.

LEMMA 2.6. *Under conditions* (2.12), *we have* $|T_3| \le \frac{2}{3} r_{ij}\gamma$.

Proof. To interpret the slope limitations (2.12) we write

$$u_{L_{ij}}(a_{ij}^-) = u_{ij} + \overrightarrow{M_{ij}a_{ij}^-}\cdot \overrightarrow{\Delta_{ij}}.$$

Since

$$\|\overrightarrow{M_{ij}a_{ij}^-}\| \le \frac{1}{2}\left(\|\overrightarrow{M_{ij}a_i}\| + \|\overrightarrow{M_{ij}G_{ij}}\|\right) \le \frac{1}{2}\left(\frac{1}{2}+\frac{1}{3}\right) bh = \frac{5}{12} bh,$$

we see that condition (2.12) implies

$$\left| \frac{u_{L_{ij}}(a_{ij}^-) - u_{ij}}{u_{i,j-1}-u_{ij}} \right| \le \frac{5}{12} bh \cdot \frac{2}{5b}\cdot\frac{\gamma}{h} = \frac{\gamma}{6}, \tag{2.13a}$$

as well as

$$\left| \frac{u_{L_{ij}}(a_{ij}^+) - u_{ij}}{u_{i,j+1}-u_{ij}} \right| \le \frac{\gamma}{6}. \tag{2.13b}$$

These two conditions, which restrict the variation of $u_{L_{ij}}(a_{ij}^-)$ and $u_{L_{ij}}(a_{ij}^+)$ about the value u_{ij}, are sufficient to establish Lemma 2.6. Indeed, we can write

(2.13c)

$$|T_3| \le 2r_{ij}q_{ij}\frac{\gamma}{6} + 2r_{ij}(1-q_{ij})\frac{\gamma}{6} + 2r_{i,j+1}q_{i,j+1}\frac{\gamma}{6} + 2r_{i,j-1}(1-q_{i,j-1})\frac{\gamma}{6} \le \frac{2}{3} r_{ij}\gamma$$

since $q_{ij}r_{ij} = (1-q_{i,j-1})r_{i,j-1}$ and $(1-q_{ij})r_{ij} = q_{i,j+1}r_{i,j+1}$. ∎

2.2.2. Estimate for T_4. We introduce the following CFL condition:

$$\frac{\Delta t \|\overrightarrow{V}\|_\infty}{h} \le \eta, \qquad \eta > 0. \tag{2.14}$$

LEMMA 2.7. *If the slope limitations* (2.12) *and the CFL condition* (2.14) *are satisfied, we have*

$$\frac{\Delta t}{A(C_i)} |T_4| \;\le\; \frac{b}{c}\cdot\frac{1}{n_i}\cdot \eta\gamma\left(\frac{8}{15}+\frac{8}{5b}\eta\right).$$

Proof. From (1.2) and (2.12) we get

$$|T_4| \;\leq\; 4\left(\frac{1}{3}bh + \Delta t\|\overrightarrow{V}\|_\infty\right)\frac{2}{5b}\;\frac{\gamma}{h}\;\frac{1}{3}bh\|\overrightarrow{V}\|_\infty,$$

and thus by (1.2)

$$\frac{\Delta t}{A(C_i)}|T_4| \;\leq\; 3\frac{\Delta t}{ch^2}\;\frac{1}{n_i}|T_4| \;\leq\; \frac{b}{c}\;\frac{1}{n_i}\eta\gamma\left(\frac{8}{15} + \frac{8}{5b}\eta\right). \qquad \square$$

Remark 2.8. Condition (2.12) can be replaced by a condition on the increment ratios

$$(2.12') \qquad \left|\frac{u_I^{n+3/2} - u_{ij}^{n+1}}{u_{ij}^{n+1} - u_{i,j-1}^{n+1}}\right|, \quad I = M_{ij}G_{ij}, \quad \text{and} \quad \left|\frac{u_I^{n+3/2} - u_{ij}^{n+1}}{u_{ij}^{n+1} - u_{i,j+1}^{n+1}}\right|, \quad I = M_{ij}G_{i,j+1},$$

whereby conditions (2.13) are conserved.

2.2.3. Conclusion.

LEMMA 2.9. *Assume conditions* (2.12) *and* (2.14) *are satisfied; if* γ *and* η *are such that*

$$(2.15) \qquad Q(\gamma,\eta) \equiv \frac{12}{5c}\frac{1}{n_i}\gamma\eta^2 + \frac{2}{n_i}\frac{b}{c}\left(3 + \frac{2}{5}\gamma\right)\eta + \left(\frac{2}{3}\gamma - 3\right)r_{ij} \leq 0$$

for each j *neighbor of* i*, then we have*

$$(2.16) \qquad |u_i^{n+2}| \leq \max_{1\leq j\leq n_i}\left\{|u_{ij}^{n+1}|\right\}.$$

Proof. Inequality (2.16) will hold if (2.11) is satisfied, which will be the case, in view of Lemmas 2.6 and 2.7, if

$$\frac{2}{3}r_{ij}\gamma + \frac{3}{2}\frac{b}{c}\frac{1}{n_i}\eta\gamma\left(\frac{8}{15} + \frac{8}{5b}\eta\right) + 9\frac{\Delta t}{ch^2}\frac{1}{n_i}\frac{2}{3}bh\|\overrightarrow{V}\|_\infty \leq 3r_{ij},$$

which, together with the CFL condition (2.14), is equivalent to (2.15). $\square$

Remark 2.10.

1. The particular case $\gamma = 0$, i.e., $\overrightarrow{\Delta_{ij}^{n+1}} = 0$, corresponds to the finite volume extension of the Lax–Friedrichs scheme, for which condition (2.15) takes the form

$$Q(0,\eta) = \frac{6}{n_i}\frac{b}{c}\eta - 3r_{ij} \leq 0$$

for each j neighbor of i, or equivalently

$$(2.17\text{a}) \qquad n_i r_{ij} \geq 2\frac{b}{c}\eta.$$

Bounds for the value of n_i, r_{ij} can be obtained from (1.2) and geometric considerations. We have defined $r_{ij} = \frac{A(L_{ij}\cap C_i)}{A(C_i)}$, which is therefore equal to the ratio of the area of two subtriangles such as $a_iG_{ij}M_{ij}$ and the area of $2n_i$ subtriangles covering C_i; applying (1.2), we have

$$\frac{1}{6}ch^2 \leq A(a_iG_{ij}M_{ij}) \leq \frac{1}{6}dh^2,$$

for which we obtain the bounds

$$\frac{c}{d} \leq n_i r_{ij} \leq \frac{d}{c}. \tag{2.17b}$$

From (2.17a)–(2.17b), condition (2.15) will therefore hold if $\eta < \frac{c^2}{2bd}$, so that we can choose the CFL condition

$$\frac{\Delta t \|\overrightarrow{V}\|_\infty}{h} \leq \frac{1}{2}\frac{c^2}{bd} \tag{2.17c}$$

for the second step of the finite volume Lax–Friedrichs scheme.

2. The two-dimensional finite volume extension of the NT scheme is obtained for $\gamma > 0$ and $\eta > 0$. To ensure that $Q(\gamma, \eta) \leq 0$ for arbitrarily small η, we must necessarily have $\gamma < \frac{9}{2}$ and $\eta < \frac{c^2}{2bd}$ (using arbitrarily small values of γ in (2.15)).

3. As in the case of the first time step, where condition (2.2) on the gradients $\overrightarrow{\Delta_i^n}$ could be replaced by a condition giving specific bounds for the value $u_{C_i}(M)$, $M \in \{G_{ij}, M_{ij}, G_{i,j+1}\}$ (Remark 2.5-4), we can introduce here instead of (2.12) the following conditions:

(2.12′)

$$0 \leq \frac{u_{L_{ij}}(M) - u_{ij}}{u_{i,j-1} - u_{ij}} \leq \frac{\gamma}{6} \text{ and } 0 \leq \frac{u_{L_{ij}}(M) - u_{ij}}{u_{i,j+1} - u_{ij}} \leq \frac{\gamma}{6}, \quad M \in \{a_i, G_{ij}, G_{i,j+1}\},$$

where

$$u_{L_{ij}}(M) = u_{ij} + \overrightarrow{M_{ij}M} \cdot \overrightarrow{\Delta_{ij}}.$$

For $\gamma \leq 6$, this means that the value of $u_{L_{ij}}(M)$ at a_i, G_{ij} and $G_{i,j+1}$ falls within the range of the values $u_{i,j-1}$, u_{ij}, and $u_{i,j+1}$.

2.3. L^∞-estimate of the solution after two time steps. Combining the results of Lemmas 2.4 and 2.9, we obtain the following lemma.

LEMMA 2.11. *We assume that condition* (2.2) *with* $\varepsilon < 3$ *(cf. Remark* 2.5.2*), condition* (2.12) *with* $\gamma < \frac{9}{2}$ *(Remark* 2.10.2*), and the CFL condition*

$$\frac{\Delta t \|\overrightarrow{V}\|_\infty}{h} \leq \min\{\beta, \eta\}, \tag{2.18}$$

where β *and* η *are chosen such that* $P(\varepsilon, \beta) \leq 0$, $Q(\gamma, \eta) \leq 0$, *are all satisfied. We then have the inequalities*

$$\sup_i |u_i^{n+2}| \leq \sup_{i;1\leq j\leq n_i} |u_{ij}^{n+1}| \leq \sup_i |u_i^n|. \tag{2.19}$$

In order to obtain a bound for $|u^n(x,y)|$ valid for any point $M(x,y)$ of the computation domain, it is sufficient to find a bound for $|u^n(G_{ij})|$ and $|u^n(M_{ij})|$ on one hand and a bound for $|u^{n+1}(a_i)|$ and $|u^{n+1}(G_{ij})|$ on the other hand (with n even and j neighbor of i). We introduce an additional condition on the gradients $\overrightarrow{\Delta_i^n}$ and $\overrightarrow{\Delta_{ij}^{n+1}}$, for each index i and each j neighbor of i:

$$\left\{\begin{array}{l} |u_{C_i}^n(M)| = |u_i^n + \overrightarrow{a_iM} \cdot \overrightarrow{\Delta_i^n}| \leq \sup_i |u_i^n|, \ M \in \{G_{ij}, M_{ij}, G_{i,j+1}\} \\ |u_{L_{ij}}^{n+1}(M)| = |u_{ij}^{n+1} + \overrightarrow{M_{ij}M} \cdot \overrightarrow{\Delta_{ij}^{n+1}}| \leq \sup_{i;1\leq j\leq n_i} |u_{ij}^{n+1}|, \\ \qquad M \in \{a_i, G_{ij}, a_j, G_{i,j+1}\} \end{array}\right\}. \tag{2.20}$$

Formula (2.20) is automatically satisfied if we choose ε and γ such that

$$\varepsilon \leq 2 \qquad \text{and} \qquad \gamma \leq 6 \tag{2.21}$$

in the case of the alternate slope limitations (2.2′)–(2.12″).

We then get the following fundamental L^∞-bound for the numerical solution.

THEOREM 2.12. *Under the hypothesis of Lemma* 2.11 *and condition* (2.20)*, we have*

$$\|u^n\|_{L^\infty(\mathbb{R}^2)} \leq \|u^0\|_{L^\infty(\mathbb{R}^2)} \leq \|u_0\|_{L^\infty(\mathbb{R}^2)}. \tag{2.22}$$

3. Estimate on the weighted total variation. We introduce an additional condition on the gradients $\overrightarrow{\Delta_i^n}$ and $\overrightarrow{\Delta_{ij}^{n+1}}$ for j neighbor of i: There exists a constant α ($0 < \alpha < 1$) such that

$$\|\overrightarrow{\Delta_i^n}\| \leq Ch^{\alpha-1}, \quad \|\overrightarrow{\Delta_{ij}^{n+1}}\| \leq Ch^{\alpha-1}, \tag{3.1}$$

where C is a constant independent of h and Δt.

In the rest of this paper, the hypothesis of Lemma 2.11, completed by conditions (2.20) and (2.21), as well as (3.1) and (3.15) below, will be referred to as "conditions (CFLCP)."

The aim of this section is to prove the following.

THEOREM 3.1. *Under the condition (CFLCP), if $h \leq 1$, there exists a constant C (independent of h and Δt) such that*

$$\sum_{i\in\mathfrak{I}} \sum_{\substack{n=0 \\ n \text{ even}}}^{L-2} A(C_i) \sum_{j=1}^{n_i} r_{ij}|u_i^n - u_j^n| \leq Ch^{\frac{\alpha}{2}-1}, \tag{3.2}$$

where $\mathfrak{I}$ is such that $\sum_{i\in\mathfrak{I}} A(C_i)$ is bounded.

Proof. We first write u_i^{n+2} as a function of u_i^n according to (1.10) and (1.6).

(3.3)

$$u_i^{n+2} = \sum_{j=1}^{n_i} r_{ij} \left(\frac{u_i^n + u_j^n}{2} + \frac{\overrightarrow{\text{Vect}_i} \cdot \overrightarrow{\Delta_i^n} + \overrightarrow{\text{Vect}_j} \cdot \overrightarrow{\Delta_j^n}}{6} - \frac{\Delta t}{A(L_{ij})} \sum_{I\in\partial L_{ij}} u_I^{n+1/2} V(I) \right) + \frac{1}{3}\sum_{j=1}^{n_i} r_{ij} \overrightarrow{\text{Vect}_{ij}} \cdot \overrightarrow{\Delta_{ij}^{n+1}} - \frac{\Delta t}{A(C_i)} \sum_{I\in\partial C_i} u_I^{n+3/2} V(I).$$

Observing that $r_{ij}\frac{A(C_i)}{A(L_{ij})} = \frac{1}{2}$ (Figure 1), we get

$$(u_i^{n+2} - u_i^n)A(C_i) + \Delta t \left(\sum_{I\in\partial C_i} u_I^{n+3/2} V(I) + \frac{1}{2} \sum_{j=1}^{n_i} \sum_{I\in\partial L_{ij}} u_I^{n+1/2} V(I) \right) - A(C_i) \left(\sum_{j=1}^{n_i} r_{ij} \frac{u_j^n - u_i^n}{2} + \frac{1}{6} \sum_{j=1}^{n_i} r_{ij} (\overrightarrow{\text{Vect}_i} \cdot \overrightarrow{\Delta_i^n} + \overrightarrow{\text{Vect}_j} \cdot \overrightarrow{\Delta_j^n}) + \frac{1}{3} \sum_{j=1}^{n_i} r_{ij} \overrightarrow{\text{Vect}_{ij}} \cdot \overrightarrow{\Delta_{ij}^{n+1}} \right) = 0. \tag{3.4}$$

Multiplying by u_i^n and observing that

$$(u_i^{n+2} - u_i^n)u_i^n = -\frac{1}{2}(u_i^n - u_i^{n+2})^2 - \frac{1}{2}(u_i^n)^2 + \frac{1}{2}(u_i^{n+2})^2,$$

we obtain, summing on $i \in \mathfrak{I}$ and even positive integers $n = 0, 2, \dots, L-2$,

$$\begin{aligned}(3.5a)\quad & -\frac{1}{2}\sum_{i,n} A(C_i)(u_i^n - u_i^{n+2})^2 + \frac{1}{2}\sum_{i,n} A(C_i)\big\{(u_i^{n+2})^2 - (u_i^n)^2\big\} \\ &+ \sum_{i,n} \Delta t\Bigg(\sum_{I\in\partial C_i} u_I^{n+3/2} u_i^n V(I) + \frac{1}{2}\sum_j \sum_{I\in\partial L_{ij}} u_I^{n+1/2} u_i^n V(I)\Bigg) \\ &- \sum_{i,n} u_i^n A(C_i)\Bigg(\sum_j r_{ij}\frac{u_j^n - u_i^n}{2} + \frac{1}{6}\sum_j r_{ij}(\overrightarrow{\mathrm{Vect}_i}\cdot\overrightarrow{\Delta_i^n} + \overrightarrow{\mathrm{Vect}_j}\cdot\overrightarrow{\Delta_j^n}) \\ &\qquad + \frac{1}{3}\sum_j r_{ij}\overrightarrow{\mathrm{Vect}_{ij}}\cdot\overrightarrow{\Delta_{ij}^{n+1}}\Bigg) = 0,\end{aligned}$$

which we decompose as

$$(3.5b)\qquad -\frac{1}{2}T_1 + \frac{1}{2}T_2 + T_3 - T_4 = 0$$

with

$$(3.5c)\qquad T_3 = \Delta t\sum_{i,n} u_i^n T_5 \quad\text{and}\quad T_4 = \sum_{i,n} u_i^n A(C_i) T_6. \qquad \square$$

3.1. Estimate for the term T_2. We have

$$T_2 = \sum_i A(C_i)\big((u_i^L)^2 - (u_i^0)^2\big) \geq -\sum_i A(C_i)(u_i^0)^2.$$

Since $u_0 \in L^\infty(\mathbb{R}^2)$ has compact support (by assumption), we have $\|u^0\|_{L^2(\mathbb{R}^2)} < \infty$. We can then write

$$\|u^0\|_{L^2(\mathbb{R}^2)} \leq Ch^{(\alpha-1)/2},$$

since $h \leq 1$, for a positive constant C. Therefore, with another constant again noted C,

$$(3.6)\qquad T_2 \geq -Ch^{\alpha-1}.$$

3.2. Estimate for T_1. From (3.5) and (3.1), we have

$$(3.7a)\qquad T_6 = \frac{1}{2}\sum_{j=1}^{n_i} r_{ij}(u_j^n - u_i^n) + O(h^\alpha).$$

Isolating $u_i^n - u_i^{n+2}$ from (3.4), we get $u_i^n - u_i^{n+2} = \frac{\Delta t}{A(C_i)}T_5 - T_6$ and thus

$$(3.7b)\qquad T_1 \equiv \sum_{i,n} A(C_i)(u_i^n - u_i^{n+2})^2 = \sum_{i,n}\frac{1}{A(C_i)}(\Delta t\, T_5 - A(C_i)T_6)^2,$$

where the term T_5 can be written

$$T_5 = \sum_{I\in\partial C_i} (u_I^{n+3/2} - u_i^n)V(I) + \frac{1}{2}\sum_{j=1}^{n_i}\sum_{I\in\partial L_{ij}} (u_I^{n+1/2} - u_i^n)V(I) \tag{3.7c}$$

$$\text{since} \quad \text{div } \overrightarrow{V} = 0 \quad \text{leads to} \sum_{I\in\partial C_i} V(I) = \sum_{I\in\partial L_{ij}} V(I) = 0. \tag{3.7d}$$

On the other hand, definitions (1.8) and (1.12) imply

$$\begin{aligned} T_5 = &\sum_j \sum_{I\in(\partial C_i)\cap L_{ij}} \big(u_{ij}^{n+1} - u_i^n + \frac{1}{2}(\overrightarrow{I} - \Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_{ij}^{n+1}}\big)V(I) \\ &+ \frac{1}{2}\sum_j \sum_{I\in(\partial L_{ij})\cap C_i} \frac{1}{2}(\overrightarrow{I} - \Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_i^n}V(I) \\ &+ \frac{1}{2}\sum_j \sum_{I\in(\partial L_{ij})\cap C_j} \big(u_j^n - u_i^n + \frac{1}{2}(\overrightarrow{I} - \Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_j^n}\big)V(I). \end{aligned} \tag{3.7e}$$

Replacing u_{ij}^{n+1} by its value (with the help of (1.6) and (1.8) using (3.7d)), we obtain the following expression for T_5:

$$\begin{aligned} T_5 = &O(h^{\alpha+1}) + \frac{1}{2}\sum_j (u_j - u_i)\left(\sum_{I\in(\partial L_{ij})\cap C_j} V(I) + \sum_{I\in(\partial C_i)\cap L_{ij}} V(I)\right) \\ &+ \Delta t\sum_j \frac{u_j - u_i}{A(L_{ij})}\left(\sum_{I\in(\partial C_i)\cap L_{ij}} V(I)\right)\left(\sum_{I\in(\partial L_{ij})\cap C_i} V(I)\right). \end{aligned} \tag{3.7f}$$

From the definition of $V(I)$ on ∂L_{ij} and ∂C_i and in view of (3.7d), we get

$$\sum_{I\in(\partial L_{ij})\cap C_j} V(I) = \sum_{I\in(\partial C_i)\cap L_{ij}} V(I) \text{ and } \sum_{I\in(\partial L_{ij})\cap C_i} V(I) = -\sum_{I\in(\partial C_i)\cap L_{ij}} V(I), \tag{3.7g}$$

so that

$$T_5 = O(h^{\alpha+1}) + \sum_j (u_j - u_i)\left\{\sum_{I\in(\partial C_i)\cap L_{ij}} V(I) - \frac{\Delta t}{A(L_{ij})}\left[\sum_{I\in(\partial C_i)\cap L_{ij}} V(I)\right]^2\right\}$$

or

$$T_5 = O(h^{\alpha+1}) + \sum_j (u_j - u_i)\left(\sum_{I\in(\partial C_i)\cap L_{ij}} V(I)\right)\left(1 - \frac{\Delta t}{A(L_{ij})}\sum_{I\in(\partial C_i)\cap L_{ij}} V(I)\right). \tag{3.8}$$

We now observe that (3.8) and (3.7a) lead to

$$\begin{aligned} \Delta t\ T_5 - A(C_i)T_6 = &O(h^{\alpha+2}) \\ &+ \sum_j (u_j - u_i)\left(\Delta t(\sum V(I))\left(1 - \frac{\Delta t}{A(L_{ij})}\sum V(I)\right) - \frac{1}{2}r_{ij}A(C_i)\right), \end{aligned}$$

where $\sum V(I)$ represents the summation for $I \in (\partial C_i) \cap L_{ij}$.

Introducing $x_{ij} \equiv \frac{\Delta t}{A(L_{ij})}\big(\sum V(I)\big)$, and observing that $r_{ij}A(C_i) = \frac{1}{2}A(L_{ij})$ (see Figure 1), we find

$$\Delta t\, T_5 - A(C_i)T_6 = O(h^{\alpha+2}) + \sum_j \frac{1}{4}A(L_{ij})(u_j - u_i)(4x_{ij}(1-x_{ij}) - 1). \tag{3.9a}$$

Using (3.7b), we then obtain the following expression for T_1.

$$T_1 = \sum_{i,n} \frac{1}{A(C_i)} \left(O(h^{\alpha+2}) - \sum_j \frac{1}{2} r_{ij} A(C_i)(u_j - u_i)(1 - 2x_{ij})^2 \right)^2. \tag{3.9b}$$

LEMMA 3.2. *Let $\mathcal{T}$ be a finite element triangulation satisfying condition* (1.2) *and $\{a_i : i \in \mathcal{I}\}$ a set of nodes such that $\sum_{i\in\mathcal{I}} A(C_i) \le A$, where A is a constant independent of h.*

a. *Then if $\{\alpha_i\}_{i\in\mathcal{I}}$ is a family of numbers such that $|\alpha_i| \le \alpha < \infty$ $(i \in \mathcal{I})$ we have the estimate*

$$\left|\sum_{i\in\mathcal{I}} \alpha_i\right| \le \sum_{i\in\mathcal{I}} |\alpha_i| = O(h^{-2}). \tag{3.9c}$$

b. *Similarly if $|\beta_n| < \beta < \infty$ $(0 \le n \le L)$ with $L\,\Delta t = T < \infty$, then*

$$\left|\sum_{n=0}^{L-1} \beta_n\right| \le \sum_{n=0}^{L-1} |\beta_n| = O(h^{-1}).$$

Proof. a. We have $\left|\sum_{i\in\mathcal{I}} \alpha_i A(C_i)\right| \le \sum |\alpha_i| A(C_i) \le \alpha \sum A(C_i) \le \alpha A = O(1)$ so that

$$\left|\sum \alpha_i\right| \le \sum |\alpha_i| \le \sum_{i\in\mathcal{I}} \frac{|\alpha_i| A(C_i)}{\min_i A(C_i)} = \frac{1}{\min_{i\in\mathcal{I}} A(C_i)} O(1) \le \frac{O(1)}{c' h^2} = O(h^{-2}),$$

since (1.2) leads to $c'h^2 \le A(C_i) \le d'h^2$ $(i \in \mathcal{I})$ for appropriate positive constants c', d'.

The proof of part b is quite similar. □

Expanding the square in (3.9b), we get the following estimate for the term T_1.

$$T_1 = O(h^{\alpha-1}) + \frac{1}{4}\sum_{i,n} A(C_i) \left[\sum_j r_{ij}(u_j - u_i)(1 - 2x_{ij})^2 \right]^2$$

$$\le O(h^{\alpha-1}) + \frac{1}{4}\sum_{i,n} A(C_i) \left[\sum_j r_{ij}(u_j - u_i)^2 (1 - 2x_{ij})^4 \right] \tag{3.9d}$$

by the Cauchy–Schwarz inequality and noting that $\sum_j r_{ij} = 1$.

3.3. Estimate for the difference $T_3 - T_4$. From (3.5c) we find $T_3 - T_4 = \sum_{i,n}(\Delta t\, T_5 - A(C_i)T_6)u_i^n$.

Applying (3.9a) and Lemma 3.2, and using the identities $(u_j - u_i)u_i = -\frac{1}{2}(u_i - u_j)^2 - \frac{1}{2}u_i^2 + \frac{1}{2}u_j^2$ and $2r_{ij}A(C_i) = A(L_{ij})$, we get

$$T_3 - T_4 = O(h^{\alpha-1}) - \frac{1}{4}\sum_{i,j,n} A(L_{ij})(1-2x_{ij})^2(u_j^n - u_i^n)u_i^n$$

$$(3.10) \qquad = O(h^{\alpha-1}) + \frac{1}{4}\sum_{i,n} A(C_i)\left[\sum_j r_{ij}(u_j - u_i)^2(1-2x_{ij})^2\right]$$

$$+ \frac{1}{8}\sum_{i,j,n} A(L_{ij})(1-2x_{ij})^2 u_i^2 - \frac{1}{8}\sum_{i,j,n} A(L_{ij})(1-2x_{ij})^2 u_j^2.$$

From the definitions of $V(I)$ on ∂C_i and x_{ij}, we have $x_{ji} = -x_{ij}$ so that reversing the order of summation and setting $i = j'$, $j = i'$ in the last term of (3.10) lead to

$$\sum_{i,j,n} A(L_{ij})(1-2x_{ij})^2 u_j^2 = \sum_{i,j,n} A(L_{ij})(1+2x_{ij})^2 u_i^2,$$

from which we deduce, for the last two terms of (3.10),

$$\frac{1}{8}\sum_{i,j,n} A(L_{ij})(1-2x_{ij})^2 u_i^2 - \frac{1}{8}\sum_{i,j,n} A(L_{ij})(1-2x_{ij})^2 u_j^2 = -\sum_{i,j,n} A(L_{ij})u_i^2 x_{ij}$$

$$= -\Delta t \sum_{i,j,n} u_i^2 \left[\sum_{I\in(\partial C_i)\cap L_{ij}} V(I)\right]$$

$$= -\Delta t \sum_{i,n} u_i^2 \sum_j \sum_{I\in(\partial C_i)\cap L_{ij}} V(I) = -\Delta t \sum_{i,n} u_i^2 \sum_{I\in\partial C_i} V(I) = 0$$

by (3.7d). The difference $T_3 - T_4$ can therefore be written

$$(3.11) \qquad T_3 - T_4 = O(h^{\alpha-1}) + \frac{1}{4}\sum_{i,n} A(C_i)\left[\sum_j r_{ij}(u_j - u_i)^2(1-2x_{ij})^2\right].$$

3.4. Preliminary estimate. Introducing the estimates (3.6), (3.9d), and (3.11) into (3.5b), we obtain

$$(3.12a) \qquad \begin{aligned} &\frac{-1}{8}\sum_{i,n} A(C_i)\left[\sum_j r_{ij}(u_j - u_i)^2(1-2x_{ij})^4\right] \\ &\quad + \frac{1}{4}\sum_{i,n} A(C_i)\left[\sum_j r_{ij}(u_j - u_i)^2(1-2x_{ij})^2\right] \le Ch^{\alpha-1}, \end{aligned}$$

which we write as

$$(3.12b) \qquad \begin{aligned} &\sum_{i,n} A(C_i)\sum_j r_{ij}(u_j - u_i)^2 P(x_{ij}) \le C\,h^{\alpha-1}, \\ &\text{where} \quad P(x_{ij}) = (1-2x_{ij})^2\big(2 - (1-2x_{ij})^2\big). \end{aligned}$$

We will try to find a condition ensuring that $P(x_{ij})$ has a strictly positive lower bound; this will enable us to omit the factor $P(x_{ij})$ in the second summation of (3.12b).

Using (1.2), the definition of $V(I)$, and (2.18) successively, we easily obtain

$$|x_{ij}| \le \frac{3\Delta t}{2ch^2} \sum_{I\in(\partial C_i)\cap L_{ij}} |V(I)| \le \frac{b}{c}\frac{\Delta t\|\vec{V}\|_\infty}{h} \le \frac{b}{c}\min\{\beta,\eta\}. \tag{3.13a}$$

Applying condition (2.8a) with $\varepsilon = 0$ and Remark 2.10.1,2 we get $\beta < \frac{c}{4b}$ and $\eta < \frac{c^2}{2bd}$, so that

$$|x_{ij}| \le \frac{b}{c}\min\left\{\frac{1}{4}\frac{c}{b}, \frac{c^2}{2bd}\right\} \le \min\left\{\frac{1}{4}, \frac{1}{2}\frac{c}{d}\right\}. \tag{3.13b}$$

This is still insufficient to ensure $P(x_{ij}) > 0$. Examining $P(x_{ij})$, we find that we must complement this condition with the restriction

$$\frac{1-\sqrt{2}}{2} < x_{ij} < \frac{1+\sqrt{2}}{2}. \tag{3.13c}$$

But from (3.13a) we have

$$-\frac{b}{c}\frac{\Delta t\|\vec{V}\|_\infty}{h} < x_{ij} < \frac{b}{c}\frac{\Delta t\|\vec{V}\|_\infty}{h}. \tag{3.13d}$$

Combining these inequalities leads to the condition

$$\frac{\Delta t\|\vec{V}\|_\infty}{h} < \frac{\sqrt{2}-1}{2}\frac{c}{d}. \tag{3.14}$$

The CFL condition (2.18) must therefore be further reinforced as follows:

$$\frac{\Delta t\|\vec{V}\|_\infty}{h} < \min\left\{\beta, \eta, \frac{\sqrt{2}-1}{2}\frac{c}{b}\right\}. \tag{3.15}$$

Under condition (3.15), $P(x_{ij})$ is necessarily strictly positive and can be made bounded away from zero if inequality (3.15) is strict, since (3.15) gives $|x_{ij}| < (\sqrt{2}-1)/2$ and thus (3.13c).

This guarantees $P(x_{ij}) > \varepsilon$ if $0 < \delta < |x_{ij}| < \frac{\sqrt{2}-1}{2} - \delta$.

Remark 3.3. A closer look at inequality (3.12a) allows a cancellation of the terms proportional to x_{ij}, and thus a slightly better CFL condition, obtained by replacing $\frac{\sqrt{2}-1}{2} \cong 0.207$ by $(\frac{\sqrt{5}}{4} - \frac{1}{2})^{1/2} \cong 0.243$.

3.5. Conclusion. Under condition (CFLCP) and (3.15) and by the previous argument, we have from (3.12b)

$$\sum_{i,n} A(C_i) \sum_j r_{ij}(u_j^n - u_i^n)^2 \le C\,h^{\alpha-1}.$$

Applying Schwarz's inequality and Lemma 3.2, we then get

$$\begin{aligned}\sum_{i,n} A(C_i) \sum_j r_{ij}|u_j^n - u_i^n| &\le \left(\sum_{i,j,n} A(C_i) r_{ij}(u_j^n - u_i^n)^2\right)^{1/2} \left(\sum_{i,j,n} A(C_i) r_{ij}\right)^{1/2} \\ &\le (C\,h^{\alpha-1})^{1/2}(C' h^{-1})^{1/2} \le D h^{\alpha/2-1}.\end{aligned}$$

This completes the proof of Theorem 3.1. □

4. Convergence in L^∞-weak*. In this section, we shall prove the following result.

THEOREM 4.1. *Under conditions (CFLCP),* (3.15), (4.4j), *and* (4.4l) *below, the sequence of numerical approximations* $\{u_{\mathcal{T},\Delta t}\}$ *converges, when* h *tends to* 0, *toward the weak solution of the initial value problem* (1.1), *in the space* $L^\infty(\mathbb{R}^2\times\mathbb{R}^+)$*-weak*.*

Proof. From Theorem 2.12, we have deduced the existence of a subsequence which converges to a function u in L^∞-weak*, and we must now prove that u is the unique weak solution of equation (1.1).

We assume that $h \le 1$ (since h tends to 0) and, following the classical approach, we consider a function $\varphi \in C_o^\infty(\mathbb{R}^2 \times [0,T))$. The numerical approximation to the solution of (1.1) is given by (1.5)–(1.9), which lead to

$$\frac{1}{2\Delta t}\big(u^{n+2}(x,y) - u^n(x,y)\big) = \frac{1}{2\Delta t}(u_i^{n+2} - u_i^n) + \frac{1}{2\Delta t}\overrightarrow{a_iM}\cdot\big(\overrightarrow{\Delta_i^{n+2}} - \overrightarrow{\Delta_i^{n}}\big) \quad \text{for } M=(x,y)\in C_i,\ n=0,2,4,\ldots. \tag{4.1}$$

We multiply (4.1) by $\varphi^n(x,y) = \varphi(x,y,t^n)$ and by $2\Delta t$, integrate on C_i, and sum for $n = 0,2,\ldots,L-2$ and all i to obtain

$$\sum_{\substack{i;n=0\\ n\text{ even}}}^{n=L-2} 2\Delta t\int_{C_i}\frac{u^{n+2}-u^n}{2\Delta t}\varphi^n = \sum_{\substack{i;n=0\\ n\text{ even}}}^{n=L-2}\int_{C_i}(u_i^{n+2}-u_i^n)\varphi^n + \sum_{\substack{i;n=0\\ n\text{ even}}}^{n=L-2} 2\Delta t\int_{C_i}\frac{\overrightarrow{a_iM}\cdot\big(\overrightarrow{\Delta_i^{n+2}}-\overrightarrow{\Delta_i^{n}}\big)}{2\Delta t}\varphi^n, \tag{4.2}$$

which we write as $A_1 = A_2 + A_3$. Applying the summation by parts formula [29]

$$\Delta t\sum_{n=r}^{s}\left(\frac{u^{n+1}-u^n}{\Delta t}\right)\varphi^n \equiv (D_+u,\varphi) = -\Delta t\sum_{n=r+1}^{s+1} u^n\frac{\varphi^n-\varphi^{n-1}}{\Delta t} + u^{s+1}\varphi^{s+1} - u^r\varphi^r,$$

where $(f,g)\equiv\sum_{n=r}^{s} f^n g^n\Delta t$, to the case of even n and $r=0$, $s=L-2$, and using the fact that $u^n \to u$ in L^∞-weak*, and $\varphi^L = 0$, we obtain

$$A_1 \underset{\Delta t\to 0}{\longrightarrow} -\int_0^T\int_{\mathbb{R}^2} u\frac{\partial\varphi}{\partial t} - \int_{\mathbb{R}^2} u_0\varphi^0, \tag{4.3}$$

as in the proof of the Lax–Wendroff theorem [21].

Let $\mathbb{B}$ be a compact set in $\mathbb{R}^2$ containing a neighborhood of the (spatial) support of φ and thus all barycentric cells C_i such that $\{\text{spatial support }(\varphi)\}\cap C_i \ne \phi$, for any $(\mathcal{T},\Delta t)$, provided that h is chosen small enough (which will be assumed). Let $\mathcal{I}$ be the set of those indices i such that $C_i \subset \mathbb{B}$. Applying summation by parts we obtain

$$A_3 = -\sum_{i\in\mathcal{I}}\int_{C_i}\overrightarrow{a_iM}\cdot\overrightarrow{\Delta_i^{0}}\varphi^0 - \sum_{\substack{i;n=2\\ n\text{ even}}}^{L} 2\Delta t\int_{C_i}\overrightarrow{a_iM}\cdot\overrightarrow{\Delta_i^{n}}\left(\frac{\varphi^n-\varphi^{n-2}}{2\Delta t}\right).$$

Using (3.1), Lemma 3.2, and the fact that $\sum_{i\in\mathcal{I}} A(C_i) < A[\text{supp}(\varphi)] < \infty$, we find

$$|A_3| = O(h^\alpha) \quad\text{so that}\quad \lim_{\substack{\Delta t\to 0\\ h\to 0}} A_3 = 0. \tag{4.4a}$$

We shall now examine the second summation A_2 in (4.2), which is more complex. From (4.2) and (3.4) we have

$$A_2 = \sum_{\substack{i;n=0 \\ n \text{ even}}}^{L-2} \int_{C_i} \left(T_6 - \frac{\Delta t}{A(C_i)} T_5 \right) \varphi^n, \tag{4.4b}$$

where T_5, T_6 are defined by (3.5).

4.1. Analysis of the first sum in A_2. Let us write

$$\sum_{\substack{i\in\mathfrak{I};n=0 \\ n \text{ even}}}^{L-2} \int_{C_i} T_6 \varphi^n = S_1 + S_2 + S_3. \tag{4.4c}$$

Lemma 4.2.

$$\lim_{h\to 0} S_1 \equiv \lim_{h\to 0} \sum_{\substack{i,n=0 \\ n \text{ even}}}^{L-2} \int_{C_i} \sum_j r_{ij} \frac{u_j - u_i}{2} \varphi^n dxdy = 0. \tag{4.4d}$$

Proof. Defining $\varphi_i^n \equiv \frac{1}{A(C_i)} \int_{C_i} \varphi^n(x,y)dxdy = \varphi(M_i^n)$ for suitable M_i^n, we have by symmetry considerations

$$S_1 = \frac{1}{2} \sum_{i,j,n} A(L_{ij} \cap C_i)(u_j - u_i)\varphi_i^n == \frac{1}{4} \sum_{i,j,n} A(L_{ij} \cap C_i)(u_j - u_i)(\varphi_i^n - \varphi_j^n)$$

and thus

$$S_1 = \frac{1}{4} \sum_{i,j,n} A(L_{ij} \cap C_i)(u_j - u_i)\overrightarrow{M_j^n M_i^n} \cdot \overrightarrow{grad}\varphi^n(P_{ij}) \quad \text{with} \quad P_{ij} \in [M_i^n M_j^n].$$

Using 3.2 and Lemma 3.2 we obtain

$$|S_1| \leq Ch \sum_{i,j,n} A(L_{ij} \cap C_i)|u_j - u_i| \leq Ch^{\frac{\alpha}{2}},$$

which proves Lemma 4.2. ∎

We now examine the second term of (4.4c),

$$S_2 = \sum_{i,j,n} \int_{C_i} \frac{1}{6} r_{ij} (\overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i} + \overrightarrow{\mathrm{Vect}_j} \cdot \overrightarrow{\Delta_j})\varphi^n, \tag{4.4e}$$

where $\overrightarrow{\mathrm{Vect}_i}$ is defined by (1.7)–(1.3). From the definition of r_{ij}, φ_i^n we have

$$S_2 = \frac{1}{6} \sum_{i,j,n} A(L_{ij} \cap C_i)(\overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i} + \overrightarrow{\mathrm{Vect}_j} \cdot \overrightarrow{\Delta_j})\varphi_i^n$$

$$= \frac{1}{6} \sum_{i,j,n} A(L_{ij} \cap C_i)\ \overrightarrow{\mathrm{Vect}_i} \cdot \overrightarrow{\Delta_i}(\varphi_i^n + \varphi_j^n)$$

$$\begin{aligned}
&= \frac{1}{3}\sum_{i,j,n} A(L_{ij}\cap C_i)\overrightarrow{\mathrm{Vect}_i}\cdot\overrightarrow{\Delta_i}\ \varphi_i^n + \frac{1}{6}\sum_{i,j,n} A(L_{ij}\cap C_i)\overrightarrow{\mathrm{Vect}_i}\cdot\overrightarrow{\Delta_i}(\varphi_j^n-\varphi_i^n) \\
&\equiv S_{21}+S_{22},
\end{aligned}\tag{4.4f}$$

where we have written $\varphi_i^n = 2\varphi_i^n - \varphi_i^n$ and used symmetry arguments $(i \leftrightarrow j)$.

By Lemma 3.2, (1.7), (3.1), and the mean value theorem, we have

$$S_{22} = O(h^{-2-1+2+1+(\alpha-1)+1}) = O(h^{\alpha})\tag{4.4g}$$

so that S_{22} tends to zero as $h\to 0$ since $0<\alpha<1$. We must now examine

$$S_{21} = \frac{1}{3}\sum_{i,j,n} A(L_{ij}\cap C_i)\overrightarrow{\mathrm{Vect}_i}\cdot\overrightarrow{\Delta_i}\ \varphi_i^n.$$

Let $\widetilde{a}_i$ be the centroid of cell C_i, and $\widetilde{M}_{ij}$ the centroid of cell L_{ij}.

LEMMA 4.3.

$$\overrightarrow{a_i\widetilde{a}_i} = \frac{1}{3}\sum_j r_{ij}\overrightarrow{\mathrm{Vect}_i},\tag{4.4h}$$

$$A(L_{ij})\overrightarrow{M_{ij}\widetilde{M}_{ij}} = \frac{1}{3}A(L_{ij}\cap C_i)(\overrightarrow{\mathrm{Vect}_{ij}}+\overrightarrow{\mathrm{Vect}_{ij}}).\tag{4.4i}$$

Proof. Denoting by $\widetilde{G_{ij}}$ (resp., $\widetilde{G_{i,j+1}}$) the centroid of triangle $a_iG_{ij}M_{ij}$ (resp., $a_iM_{ij}G_{i,j+1}$) and letting $M=(x,y)\in\mathbb{R}^2$, we have by (1.7)

$$\begin{aligned}
\overrightarrow{\mathrm{Vect}_i} &= q_{ij}(\overrightarrow{a_iM_{ij}}+\overrightarrow{a_iG_{ij}}+(1-q_{ij})(\overrightarrow{a_iM_{ij}}+\overrightarrow{a_iG_{i,j+1}}) \\
&= \frac{3}{A(L_{ij}\cap C_i)}\left\{A(a_iG_{ij}M_{ij})\overrightarrow{a_i\widetilde{G}_{ij}}+A(a_iM_{ij}G_{i,j+1})\overrightarrow{a_i\widetilde{G}_{i,j+1}}\right\} \\
&= \frac{3}{A(L_{ij}\cap C_i)}\left\{\int_{a_iG_{ij}M_{ij}}\overrightarrow{a_iM}dxdy+\int_{a_iM_{ij}G_{i,j+1}}\overrightarrow{a_iM}dxdy\right\} \\
&= \frac{3}{A(L_{ij}\cap C_i)}\int_{L_{ij}\cap C_i}\overrightarrow{a_iM}dxdy
\end{aligned}$$

and therefore

$$\sum_j r_{ij}\overrightarrow{\mathrm{Vect}_i} = \frac{3}{A(C_i)}\sum_j\int_{L_{ij}\cap C_i}\overrightarrow{a_iM}dxdy = \frac{3}{A(C_i)}\int_{C_i}\overrightarrow{a_iM}dxdy = 3\overrightarrow{a_i\widetilde{a}_i}.$$

The proof of (4.4.i) is similar.

$$\begin{aligned}
\overrightarrow{\mathrm{Vect}_{ij}} &= q_{ij}(\overrightarrow{M_{ij}a_i}+\overrightarrow{M_{ij}G_{ij}})+(1-q_{ij})(\overrightarrow{M_{ij}a_i}+\overrightarrow{M_{ij}G_{i,j+1}}) \\
&= \frac{3}{A(L_{ij}\cap C_i)}\left\{A(a_iG_{ij}M_{ij})\overrightarrow{M_{ij}\widetilde{G}_{ij}}+A(a_iM_{ij}G_{i,j+1})\overrightarrow{M_{ij}\widetilde{G}_{i,j+1}}\right\} \\
&= \frac{3}{A(L_{ij}\cap C_i)}\int_{L_{ij}\cap C_i}\overrightarrow{M_{ij}M}dxdy,
\end{aligned}$$

from which we deduce

$$\frac{1}{3}A(L_{ij}\cap C_i)(\overrightarrow{\mathrm{Vect}_{ij}}+\overrightarrow{\mathrm{Vect}_{ji}}) = \int_{L_{ij}\cap C_i}\overrightarrow{M_{ij}M}\,dxdy + \int_{L_{ij}\cap C_j}\overrightarrow{M_{ij}M}\,dxdy$$
$$= \int_{L_{ij}}\overrightarrow{M_{ij}M}\,dxdy = A(L_{ij})\overrightarrow{M_{ij}\widetilde{M}_{ij}},$$

which proves Lemma 4.3. □

We shall now show that under an additional condition the term S_2 tends to zero as $h\to 0$.

LEMMA 4.4. *Under conditions* (1.2) *and* (3.1), *if* $\alpha > \frac{1}{2}$ *and if there exists* $C>0$ *such that*

$$\left|\overrightarrow{a_i\widetilde{a}_i}\cdot\overrightarrow{\triangle_i^n}\right| \le Ch^{2\alpha} \quad \textit{for all } i\ \in \mathfrak{I},\ n=0,\cdots,L-2\ (n\ \textit{even}), \tag{4.4j}$$

then $\lim_{h\to 0} S_2 = 0$.

Proof. By (4.4g), it suffices to show that $\lim_{h\to 0} S_{21}=0$. We have

$$S_{21} = \frac{1}{3}\sum_{i,n}A(C_i)\left(\sum_j r_{ij}\overrightarrow{\mathrm{Vect}_i}\cdot\overrightarrow{\triangle_i^n}\right)\varphi_i^n = \sum_{i,n}A(C_i)\overrightarrow{a_i\widetilde{a}_i}\cdot\overrightarrow{\triangle_i^n}\ \varphi_i^n,$$

so that by Lemma 3.2 and (4.4j), $|S_{21}|\le Ch^{2\alpha-1}$ and thus $\lim_{h\to 0}S_{21}=0$ since $2\alpha-1>0$. □

Remark 4.5. In the case of a regular or structured grid, one can construct cells for which $a_i \equiv \overrightarrow{\widetilde{a}_i}$, and condition (4.4j) is then trivially satisfied. Otherwise, this condition can be interpreted as imposing that a_i and $\overrightarrow{\widetilde{a}_i}$ should be "close enough" or else that we exert, at each time step, a certain control on the direction of the gradient vector $\overrightarrow{\triangle_i^n}$. Condition (4.4j) can also be written as $u^n_{C_i}(\widetilde{a}_i)-u^n_{C_i}(a_i)=O(h^{2\alpha})$, which is a regularity condition on the piecewise linear reconstruction $u^n_{C_i}$.

We now examine the third term of (4.4c):

$$\begin{aligned}
S_3 &= \sum_{i,j,n}\int_{C_i}\frac{1}{3}r_{ij}\overrightarrow{\mathrm{Vect}_{ij}}\cdot\overrightarrow{\triangle_{ij}^{n+1}}\ \varphi^n\\
&= \frac{1}{6}\sum_{i,j,n}A(L_{ij}\cap C_i)\left\{\overrightarrow{\mathrm{Vect}_{ij}}\cdot\overrightarrow{\triangle_{ij}^{n+1}}\ \varphi_i^n + \overrightarrow{\mathrm{Vect}_{ji}}\cdot\overrightarrow{\triangle_{ij}^{n+1}}\ \varphi_j^n\right\}\\
&= \frac{1}{6}\sum_{i,j,n}A(L_{ij}\cap C_i)(\overrightarrow{\mathrm{Vect}_{ij}}+\overrightarrow{\mathrm{Vect}_{ji}})\cdot\overrightarrow{\triangle_{ij}^{n+1}}\varphi_i^n\\
&\quad + \frac{1}{6}\sum_{i,j,n}A(L_{ij}\cap C_i)(\overrightarrow{\mathrm{Vect}_{ji}}\cdot\overrightarrow{\triangle_{ij}^{n+1}})(\varphi_j^n-\varphi_i^n)\\
&\equiv S_{31}+S_{32}.
\end{aligned} \tag{4.4k}$$

Proceeding as before, we can show that $S_{32}=O(h^\alpha)$, and therefore $\lim_{h\to 0}S_{32}=0$. In order to handle S_{31}, we shall need the following.

LEMMA 4.6. *Under conditions* (1.2) *and* (3.1), *if* $\alpha>\frac{1}{2}$ *and if there exists* $C>0$ *such that*

$$\left|\overrightarrow{M_{ij}\widetilde{M}_{ij}}\cdot\overrightarrow{\triangle_{ij}^{n+1}}\right| \le Ch^{2\alpha} \quad i,j\in\mathfrak{I},\ n=0,2,\ldots,L-2\ (n\ \textit{even}), \tag{4.4l}$$

then $\lim_{h\to 0} S_{31} = 0$ *and therefore* $\lim_{h\to 0} S_3 = 0$.

Proof. Applying Lemma 4.3, (4.4l), and Lemma 3.2, we have

$$|S_{31}| = \left| \frac{1}{2} \sum_{i,j,n} A(L_{ij}) \overrightarrow{M_{ij}\widetilde{M}_{ij}} \cdot \overrightarrow{\triangle_{ij}^{n+1}}\ \varphi_i^n \right| \le Ch^{2\alpha-1}$$

and thus $\lim_{h\to 0} S_{31} = 0$. □

Collecting the results of Lemmas 4.2, 4.5, and 4.6 we see that we have shown that $\lim_{h\to 0} S_i = 0$ for $i = 1, \ldots, 3$, and the first sum in A_2 therefore tends to zero as $h \to 0$, i.e.,

$$\lim_{h\to 0} \sum_{\substack{i;n=0 \\ n \text{ even}}}^{n=L-2} \int_{C_i} T_6\ \varphi^n = 0. \tag{4.5}$$

4.2. Preliminary elements for the analysis of the second sum in A_2. Let I be a side of a triangle $K \in \mathcal{T}$, $\overrightarrow{n_I}$ the outer normal to I, and x_I an arbitrary given real number associated with side I. As described in [14], one can construct a function $\overrightarrow{F_K}(x,y)$ such that

(i)

$$\overrightarrow{F_K}(x,y) \cdot \overrightarrow{n_I} = x_I \quad \text{for all} \quad (x,y) \in I, \tag{4.6a}$$

and

(ii) div $\overrightarrow{F_K}(x,y)$ takes a constant value (depending on $\overrightarrow{F_K}$ and thus on the three parameters $\{x_I\}_{I\in\partial K}$) for all $(x,y) \in \mathbb{R}^2$. These functions can be written as (see [14])

$$\overrightarrow{F_K}(x,y) = \sum_{I\in\partial K} x_I \overrightarrow{F_{K,I}}(x,y), \tag{4.6b}$$

where

$$\overrightarrow{F_{K,I}}(x,y) \cdot \overrightarrow{n_J} = \begin{cases} 1 & \text{if} \quad I = J \\ \\ 0 & \text{if} \quad I \ne J \end{cases} \qquad \text{for } I, J \in \partial K. \tag{4.6c}$$

Under condition (1.2), $\|\overrightarrow{F_{K,I}}\|_{L^\infty(\mathbb{R}^2)}$ is bounded by a constant depending only on a, b, c, and d.

Moreover, for every function $\overrightarrow{V} : \mathbb{R}^2 \to \mathbb{R}^2$ and any triangle $K \in \mathcal{T}$, one can write

$$\overrightarrow{V} = \sum_{I\in\partial K} (\overrightarrow{V} \cdot \overrightarrow{n_I}) \overrightarrow{F_{K,I}}. \tag{4.6d}$$

4.3. Analysis of the second sum in A_2. The intersection $L_{ij} \cap C_i$ of cells L_{ij} and C_i (Figure 1) can be decomposed in the triangles $a_i G_{ij} M_{ij} = K_{ij}^r$ or K^r and $a_i M_{ij} G_{i,j+1} = K_{ij}^\ell$ or K^ℓ. With the above notations, we define

$$\overrightarrow{F_{K^\ell}}(x,y) = \sum_{I\in\partial K^\ell} u_I^{n+1/2} \overrightarrow{V_I} \cdot \overrightarrow{n_I} \overrightarrow{F_{K,I}}(x,y), \tag{4.7a}$$

(4.7b) $$\overrightarrow{G_{K^\ell}}(x,y) = \sum_{I\in\partial K^\ell} u_I^{n+3/2}\overrightarrow{V_I}\cdot\overrightarrow{n_I}\overrightarrow{F_{K,I}}(x,y),$$

with similar definitions for the functions $\overrightarrow{F_{K^r}}$ and $\overrightarrow{G_{K^r}}$. We also define the average values

(4.8a) $$u_I^{n+1/2} = u_i^n \quad \text{if} \quad I = a_iM_{ij}\ , \quad u_j^n \quad \text{if} \quad I = a_jM_{ij}, \qquad \frac{u_i^n + u_j^n}{2} \quad \text{if} \quad I = G_{ij}M_{ij} \quad \text{or} \quad M_{ij}G_{i,j+1},$$

(4.8b) $$u_I^{n+3/2} = u_{ij}^{n+1} \quad \text{if} \quad I \in \{a_iG_{ij},\ a_iM_{ij},\ a_iG_{i,j+1},\ a_jG_{ij},\ a_jM_{ij},\ a_jG_{i,j+1}\}.$$

Applying (4.6) and (4.7) now gives

$$\sum_{I\in\partial L_{ij}} u_I^{n+1/2}V(I) = \sum_{K\in L_{ij}}\sum_{I\in\partial K}\int_I u_I^{n+1/2}\overrightarrow{V_I}\cdot\overrightarrow{n_I}d\sigma = \sum_{K\in L_{ij}}\int_{\partial K}\overrightarrow{F_K}(x,y)\cdot\overrightarrow{n}(x,y)d\sigma,$$

where the set of triangles "$K \in L_{ij}$" is equal to $\left\{K_{ij}^\ell,\ K_{ij}^r,\ K_{ji}^\ell,\ K_{ji}^r\right\}$.

We can therefore write

(4.9a) $$\sum_{I\in\partial L_{ij}} u_I^{n+1/2}V(I) = \sum_{K\in L_{ij}}\int_K \operatorname{div}\overrightarrow{F_K}(x,y)dxdy = \sum_{K\in L_{ij}} A(K)\operatorname{div}\overrightarrow{F_K}.$$

In the same manner, we can show that

(4.9b) $$\sum_{I\in\partial C_i} u_I^{n+3/2}V(I) = \sum_j\left(A(K_{ij}^\ell)\operatorname{div}\ \overrightarrow{G_{K_{ij}^\ell}} + A(K_{ij}^r)\operatorname{div}\ \overrightarrow{G_{K_{ij}^r}}\right).$$

4.4. Limit of A_2. With the help of (4.4b) and (4.9), the second sum in A_2 can be written (for even n)

(4.10)
$$-\sum_{i;n=0}^{n=L-2}\frac{\Delta t}{A(C_i)}\int_{C_i} T_5\varphi^n = -\sum_{i;n=0}^{L-2}\Delta t\int_{C_i}\sum_j\left\{\frac{1}{2}r_{ij}(q_{ij}\operatorname{div}\ \overrightarrow{F_{K_{ij}^r}} + (1-q_{ij})\operatorname{div}\ \overrightarrow{F_{K_{ij}^\ell}}\right.$$
$$\left. + q_{ij}\ \operatorname{div}\ \overrightarrow{F_{K_{ji}^\ell}} + (1-q_{ij})\operatorname{div}\ \overrightarrow{F_{K_{ji}^r}}) + r_{ij}\left(q_{ij}\ \operatorname{div}\ \overrightarrow{G_{K_{ij}^r}} + (1-q_{ij})\operatorname{div}\ \overrightarrow{G_{K_{ij}^\ell}}\right)\right\}\varphi^n$$
$$= \sum_{i;n=0}^{L-2}\Delta t\int_{C_i}\sum_j\left\{\frac{1}{2}r_{ij}\left(q_{ij}(\overrightarrow{F_{K_{ij}^r}} + \overrightarrow{F_{K_{ji}^\ell}}) + (1-q_{ij})(\overrightarrow{F_{K_{ij}^\ell}} + \overrightarrow{F_{K_{ji}^r}})\right)\right.$$
$$\left. + r_{ij}(q_{ij}\overrightarrow{G_{K_{ij}^r}} + (1-q_{ij})\overrightarrow{G_{K_{ij}^\ell}})\right\}\cdot\overrightarrow{\nabla}\varphi^n.$$

We now define the function $\overrightarrow{R^n}(x,y)$ by $\overrightarrow{R^n}(x,y) = u^n(x,y)\overrightarrow{V}(x,y)$ and observe that

(4.11) $$\sum_{\substack{n=0\\ n\ \text{even}}}^{L-2} 2\Delta t\int_{\mathbb{R}^2}\overrightarrow{R^n}(x,y)\cdot\overrightarrow{\nabla}\varphi^n \underset{\Delta t\to 0}{\longrightarrow}\int_0^T\int_{\mathbb{R}^2} u\overrightarrow{V}\cdot\overrightarrow{\nabla}\varphi.$$

To complete the analysis of the second sum in A_2, we need the following lemma.

LEMMA 4.7.

$$\sum_{\substack{n=0\\ n\ even}}^{L-2} \Delta t \int_{C_i} \left\{ \sum_j \left\{ \frac{1}{2} r_{ij}\big(q_{ij}(\overrightarrow{F_{K_{ij}^r}} + \overrightarrow{F_{K_{ji}^\ell}}) + (1-q_{ij})(\overrightarrow{F_{K_{ij}^\ell}} + \overrightarrow{F_{K_{ji}^r}})\big) + r_{ij}\big(q_{ij}\overrightarrow{G_{K_{ij}^r}} + (1-q_{ij})\overrightarrow{G_{K_{ij}^\ell}}\big) \right\} - 2\overrightarrow{R^n} \right\} \cdot \overrightarrow{\nabla}\varphi^n \longrightarrow 0 \tag{4.12}$$

in L^∞-weak as $\Delta t \to 0$.*

From (4.11) and Lemma 4.7, we conclude that the second sum in A_2, represented by (4.10), satisfies

$$-\sum_{\substack{i;n=0\\ n\ \text{even}}}^{L-2} \frac{\Delta t}{A(C_i)} \int_{C_i} T_5 \varphi^n \longrightarrow \int_0^T \int_{\mathbb{R}^2} u\overrightarrow{V} \cdot \overrightarrow{\nabla}\varphi \qquad \text{as} \qquad \Delta t \to 0. \tag{4.13}$$

Proof of Lemma 4.7. We first observe that

$$\sum_j \left(\frac{1}{2} r_{ij}\big(2q_{ij} + 2(1-q_{ij})\big) + r_{ij}(q_{ij} + 1 - q_{ij}) \right) = 2\sum_j r_{ij} = 2,$$

which enables us to distribute $\overrightarrow{R^n}$ onto each term in the sum in (4.12) according to the coefficients $\frac{1}{2}r_{ij}q_{ij}$, $\frac{1}{2}r_{ij}q_{ij}$, $\frac{1}{2}r_{ij}(1-q_{ij})$, $\frac{1}{2}r_{ij}(1-q_{ij})$, $r_{ij}q_{ij}$, $r_{ij}(1-q_{ij})$; it will therefore be sufficient to show that, typically, terms of the form

$$\sum_{i,n} \Delta t \int_{C_i} \sum_j r_{ij} q_{ij} \big(\overrightarrow{F_K} - \overrightarrow{R^n}\big) \cdot \overrightarrow{\nabla}\varphi^n \longrightarrow 0 \quad \text{as} \quad \Delta t \to 0. \tag{4.14}$$

Each of the six terms appearing in (4.12) will be handled in the same manner. From (4.7) and (4.6d), we have

$$\begin{aligned} \overrightarrow{F_K} - \overrightarrow{R^n} &= \sum_{I\in\partial K} u_I^{n+1/2} \overrightarrow{V_I} \cdot \overrightarrow{n_I}\ \overrightarrow{F_{K,I}} - \sum_{I\in\partial K} u^n \overrightarrow{V} \cdot \overrightarrow{n_I}\ \overrightarrow{F_{K,I}} \\ &= \sum_{I\in\partial K} \big(u_I^{n+1/2} - u^n\big)\overrightarrow{V_I} \cdot \overrightarrow{n_I}\ \overrightarrow{F_{K,I}} + \sum_{I\in\partial K} u^n \big(\overrightarrow{V_I} - \overrightarrow{V}\big) \cdot \overrightarrow{n_I}\ \overrightarrow{F_{K,I}}, \end{aligned}$$

so that we can split (4.14) into two parts, the second of which,

$$\sum_{i;n=0}^{L-2} \Delta t \int_{C_i} \sum_j r_{ij} q_{ij} u^n \Big(\sum_{I\in\partial K} \big(\overrightarrow{V_I} - \overrightarrow{V}\big) \cdot \overrightarrow{n_I}\ \overrightarrow{F_{K,I}} \Big) \cdot \overrightarrow{\nabla}\varphi^n,$$

tends to zero as $\Delta t \to 0$ since $\sum r_{ij}q_{ij} = \frac{1}{2}$; u^n, $\overrightarrow{\nabla}\varphi^n$, $\overrightarrow{F_{K,I}}$ are bounded; $i \in \mathcal{I}$, where $\mathcal{I}$ is the set of indices i such that $C_i \subset B$, so that $\sum_{i\in\mathcal{I}} A(C_i)$ is bounded; and $\|\overrightarrow{V_I} - \overrightarrow{V}\|$ tends to zero for every side I contained in the compact set B containing the support of φ.

For the first part of (4.14), we get

$$\begin{aligned}
&\Big|\sum_{i,n}\Delta t\int_{C_i}\sum_j r_{ij}q_{ij}\Big(\sum_{I\in\partial K}(u_I^{n+1/2}-u^n)\overrightarrow{V_I}\cdot\overrightarrow{n_I}\ \overrightarrow{F_{K,I}}\Big)\cdot\overrightarrow{\nabla}\varphi^n\Big|\\
&\quad\le\sum_{i,n}\Delta t\int_{C_i}\sum_j r_{ij}q_{ij}\sum_{I\in\partial K}\big|u_I^{n+1/2}-u_i^n\big|\ \big|\overrightarrow{V_I}\cdot\overrightarrow{n_I}\big|\ \big\|\overrightarrow{F_{K,I}}\big\|_{L^\infty}\big\|\overrightarrow{\nabla}\varphi^n\big\|_{L^\infty}\\
&\quad+\sum_{i,n}\Delta t\int_{C_i}\sum_j r_{ij}q_{ij}\sum_{I\in\partial K}\big|\overrightarrow{a_iM}\cdot\overrightarrow{\Delta_i^n}\big|\ \big|\overrightarrow{V_I}\cdot\overrightarrow{n_I}\big|\ \big\|\overrightarrow{F_{K,I}}\big\|_{L^\infty}\big\|\overrightarrow{\nabla\varphi^n}\big\|_{L^\infty}\\
&\quad=A_4+A_5.
\end{aligned}$$

With (3.1), A_5 clearly tends to zero since $i\in\mathcal{I}$.

Let us now examine the term A_4 and more precisely the difference $\big|u_I^{n+1/2}-u_i^n\big|$ for $I\in\partial K$. The treatment along the various triangle edges is similar; for example, if $I\in\{a_jG_{ij},a_jG_{i,j+1}\}$, $\big|u_I^{n+1/2}-u_i^n\big|\le\big|\frac{1}{2}(\overrightarrow{I}-\Delta t\overrightarrow{V_I})\cdot\overrightarrow{\Delta_j^n}\big|+\big|u_j^n-u_i^n\big|\le Ch^\alpha+\big|u_j^n-u_i^n\big|$ by (1.8) and (3.1).

Since $q_{ij}<1$, we obtain from the CFL condition and Theorem 3.1

$$\begin{aligned}
A_4&\le\sum_{i,n}\Delta tA(C_i)\Big(\sum_j r_{ij}|u_i^n-u_j^n|\Big)\big\|\overrightarrow{V}\big\|_{L^\infty}\big\|\overrightarrow{\nabla}\varphi\big\|_{L^\infty}\Big(\sum_{I\in\partial K}\big\|\overrightarrow{F_{K,I}}\big\|_{L^\infty}\Big)+\tilde{C}h^\alpha\\
&\le Ch^{\alpha/2}+\tilde{C}h^\alpha.
\end{aligned}$$

A_4 therefore tends to zero as $\Delta t\to 0$, completing the proof of Lemma 4.7. ◻

Remark 4.8. The terms proportional to $\overrightarrow{G_K}-\overrightarrow{R^n}$ are handled in a similar way. We have

$$|u_I^{n+3/2}-u_i^n|\le Ch^\alpha+|u_{ij}^{n+1}-u_i^n|\le Ch^\alpha+|u_i^n-u_j^n|,$$

since under conditions (CFLCP), u_{ij}^{n+1} is a convex combination of u_i^n and u_j^n.

4.5. Conclusion. Collecting the results of (4.3), (4.4a), and (4.13), we find that the L^∞-weak* limit u of the subsequence $\{u_{\mathcal{T}_k,\Delta t_k}\}$ described at the end of section 1 satisfies

$$\text{(4.15a)}\qquad \int_0^T\int_{\mathbb{R}^2}u\,\frac{\partial\varphi}{\partial t}\;+\;\int_0^T\int_{\mathbb{R}^2}u\overrightarrow{V}\cdot\overrightarrow{\nabla}\varphi\;+\;\int_{\mathbb{R}^2}u^o\varphi^o=0,$$

thus establishing that u is a weak solution of problem (1.1) and completing the proof of Theorem 4.1, which guarantees the convergence of our two-dimensional finite volume generalization of the nonoscillatory central difference scheme of Nessyahu and Tadmor. ◻

In [7], [10], [12], [13], [33], we describe a variety of numerical experiments with our finite volume scheme and an extension to a mixed finite volume/finite element method for the compressible Navier–Stokes equations, including comparisons with other well-established methods. These comparisons show the high level of accuracy and efficiency provided by our scheme.

5. Numerical experiments. In this section we have selected one of the numerical experiments presented in [7], [33], [10], [11], the case of supersonic flow around a double ellipse.

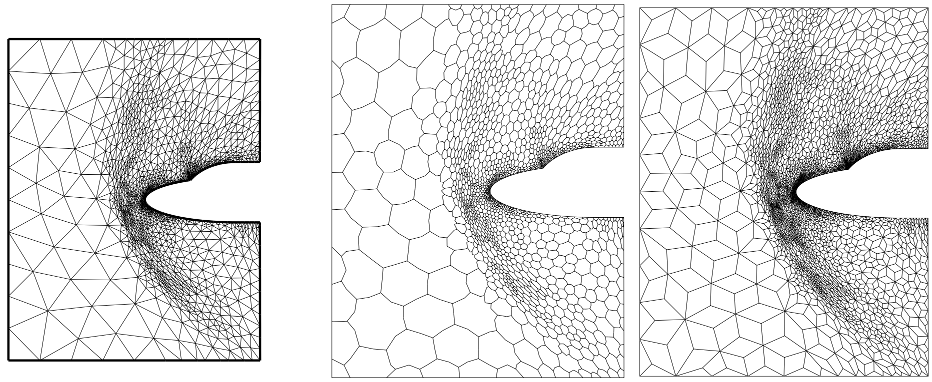

FIG. 2. *Euler flow around a double ellipse. Original grid, barycentric cells C_i, and quadrilateral cells L_{ij}.*

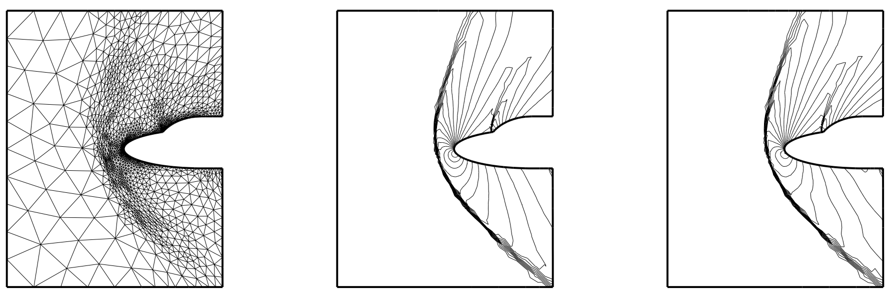

FIG. 3. *Double ellipse: Initial mesh (1558 vertices) and solution (pressure and Mach contours) (FV).*

Example. Supersonic flow past a double ellipse at 20° of angle of attack and $M_\infty = 2$.

For this problem, inspired by [34] but with Mach number $M_\infty = 2$ instead of the range of hypersonic Mach numbers considered there and 20° of angle of attack, the geometry is a double ellipse; it can be defined by

$$\begin{cases} x \le 0 & \begin{cases} z \le 0, & \left(\frac{x}{0.06}\right)^2 + \left(\frac{z}{0.015}\right)^2 = 1, \\ z \ge 0, & \left(\frac{x}{0.035}\right)^2 + \left(\frac{z}{0.025}\right)^2 = 1, \end{cases} \\ 0 \le x \le 0.016 & \begin{cases} z \ge 0, & z = 0.025, \\ z \le 0, & z = -0.015. \end{cases} \end{cases}$$

For this steady flow problem we compared our finite volume method with a discontinuous finite element method recently proposed by Jaffré and Kaddouri [18] and which seems to be fairly competitive; we used the same three meshes with both methods. For the initial mesh (1558 vertices, Figure 2), both methods give comparable results (Figures 3–6), albeit with very unequal computing times (see below). Notice that the C_p curves can be nearly superposed, which is an indication that both methods are indeed doing some reasonable calculation. The same is true for the pressure and Mach contours of both methods, with perhaps a very small advantage for our finite

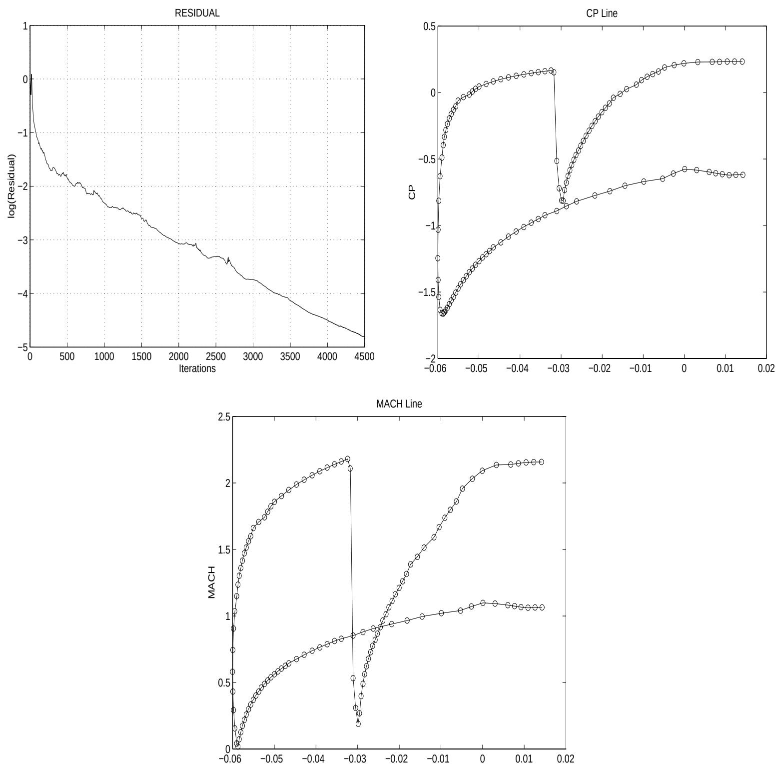

FIG. 4. *Residual and* C_p *and Mach body cuts (initial mesh* 1558 *vertices) (FV).*

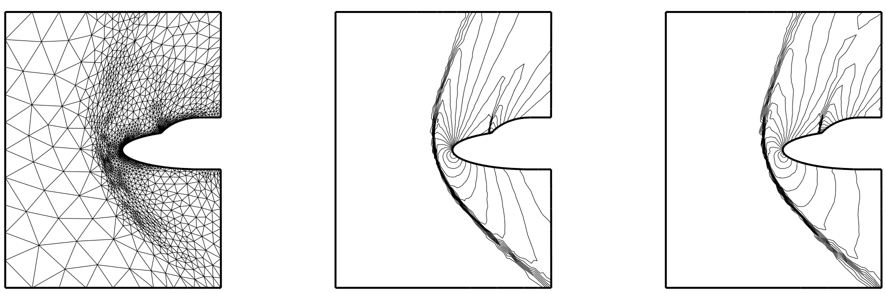

FIG. 5. *Double ellipse: Initial mesh (*1558 *vertices) and solution (pressure and Mach contours) (DFE).*

volume (FV) method which gives slightly sharper shocks and somewhat smoother level contours.

For the intermediate mesh (2792 vertices), the advantage offered by the FV method becomes a little more obvious in Figures 7 and 8, where the breaches of monotonicity are more important with the discontinuous finite element (DFE) method (lower part of the bow shock). Moreover the pressure and Mach contours are more regular with the FV method.

The final mesh (5055 vertices, Figures 9, 10, and 11) shows a clear advantage for the FV method, which gives a nearly perfect shock resolution with very smooth

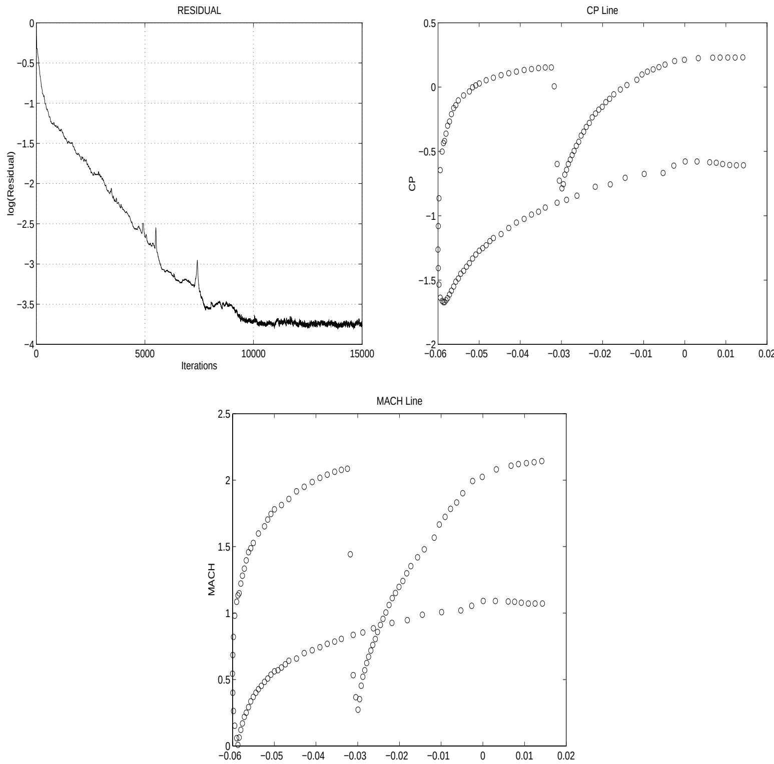

FIG. 6. *Residual and C_p and Mach body cuts (initial mesh* 1558 *vertices) (DFE).*

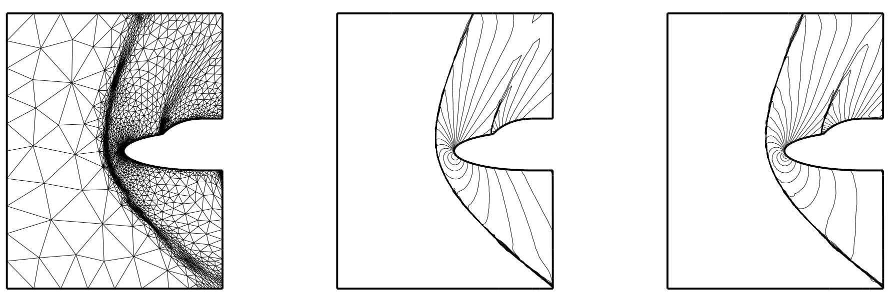

FIG. 7. *Double ellipse: First adaptation (*2792 *vertices) and solution (pressure and Mach contours) (FV).*

contours, while the DFE method shows a breach of monotonicity in the lower part of the bow shock.

As was the case with the initial mesh, the C_p curves can again be nearly exactly superposed, while the Mach line of the FV method is slightly higher, for the left part of the upper curve, than with the DFE method, a fact which is confirmed by Tables 1 and 2.

The major difference between the two methods appears to lie in the convergence history and computing times. Figures 4, 6, and 12 show a clear advantage for our

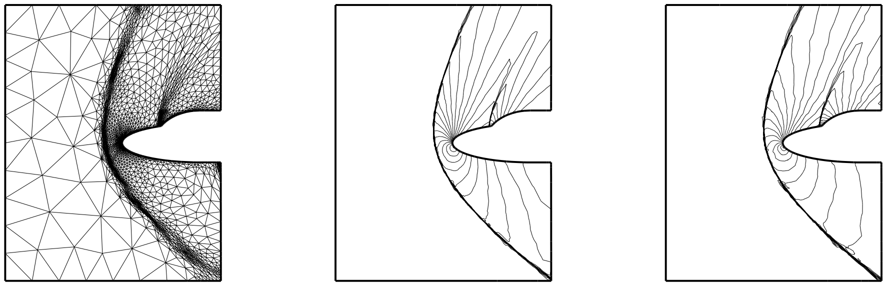

FIG. 8. *Double ellipse: First adaptation (*2792 *vertices) and solution (pressure and Mach contours) (DFE).*

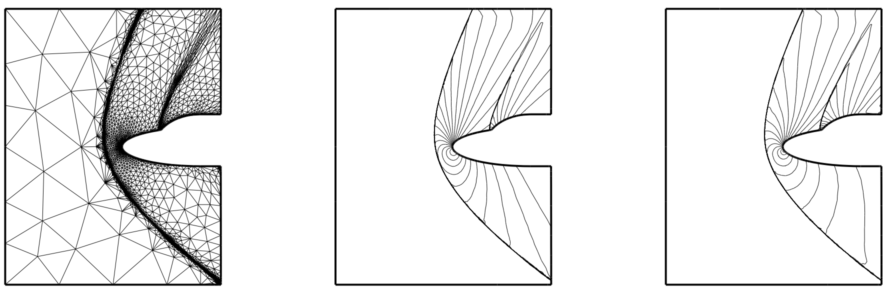

FIG. 9. *Double ellipse: Final mesh (*5055 *vertices) and solution (pressure and Mach contours) (FV).*

TABLE 1
The maximal and minimal values of pressure and Mach number (FV).

FV	Pressure		Mach	
Initial mesh	min = $6.1671760e^{-2}$	max = 1.009705	min = $1.750865e^{-2}$	max = 2.253697
2nd mesh	min = $6.1395669e^{-2}$	max = 1.006208	min = $5.3774943e^{-3}$	max = 2.266636
Final mesh	min = $6.0346086e^{-2}$	max = 1.009427	min = $5.209895e^{-3}$	max = 2.270716

TABLE 2
The maximal and minimal values of pressure and Mach number (DFE).

DFE	Pressure		Mach	
Initial mesh	min = $6.2501445e^{-2}$	max = 1.014265	min = $8.2084965e^{-3}$	max = 2.190479
2nd mesh	min = $6.2390134e^{-2}$	max = 1.007068	min = $2.0193825e^{-3}$	max = 2.216612
Final mesh	min = $6.3052103e^{-2}$	max = 1.007425	min = $1.3435918e^{-2}$	max = 2.211899

finite volume method for the initial mesh (1558 vertices). Computing times (CPU: 3564 for FV and 48288 for DFE) confirm the advantage of the proposed FV method.

Finally, let us mention that all calculations have been performed on a Silicon Graphics Station of the Centre de Recherches Mathématiques, Université de Montréal (model Challenge, 100 Mhz, 6 processors).

Acknowledgments. The authors would like to express their warmest gratitude to Professor C. Carasso and the University of St. Etienne for providing moral and material support to their collaboration, and to the reviewers for their numerous com-

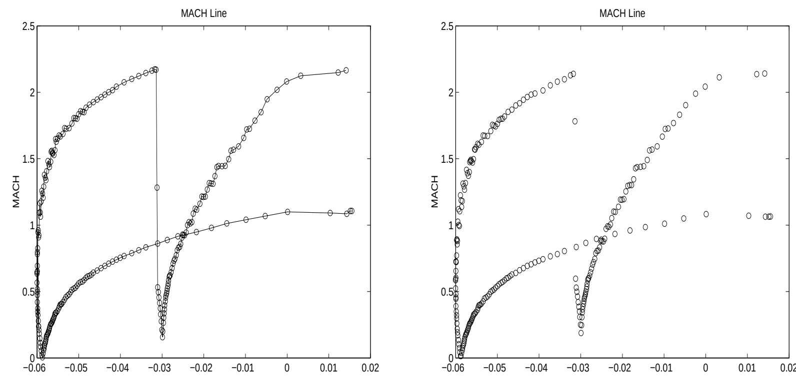

FIG. 10. *C_p and Mach body cuts for the final mesh (5055 vertices) (FV, left, and DFE, right).*

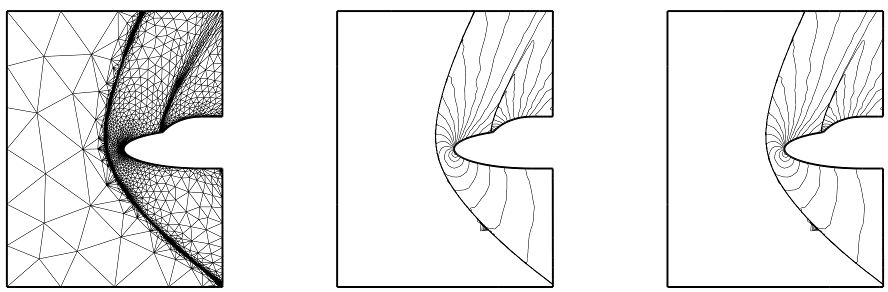

FIG. 11. *Double ellipse: Final mesh (5055 vertices) and solution (pressure and Mach contours)(DFE).*

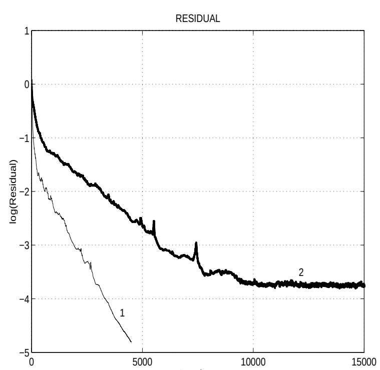

FIG. 12. *Residual for initial mesh (1558 vertices) ((1) = FV − (2) = DFE).*

ments and corrections, which resulted from a very professional examination of our work. We thank the editors of *The International Journal of Computational Fluid Dynamics* and *Computational Fluid Dynamics Review* for kindly allowing us to borrow the example in section 5 from our earlier papers and A. Madrane for performing the numerical computations.

REFERENCES

[1] F. Angrand, A. Dervieux, L. Loth, and G. Vijayasundaram, *Simulations of Euler transonic flows by means of explicit finite-element type schemes*, INRIA Res. Rep. 250, Le Chesnay, France, 1983.

[2] F. Angrand and A. Dervieux, *Some explicit triangular finite element schemes for the Euler equations*, Internat. J. Numer. Methods Fluids, 4 (1984), pp. 749–764.

[3] P. Arminjon, A. Dervieux, L. Fezoui, H. Steve, and B. Stoufflet, *Non-oscillatory schemes for multidimensional Euler calculations with unstructured grids*, in Proceedings 2nd International Conference on Hyperbolic Problems, Aachen, Germany, 1988, R. Jeltsch and J. Ballmann, eds., Notes Numer. Fluid Mech., 24, Vieweg, Braunschweig, 1988, pp. 1–10.

[4] P. Arminjon and A. Dervieux, *Construction of TVD-like artificial viscosities on two-dimensional arbitrary FEM grids*, J. Comput. Phys., 106 (1993), pp. 176–198.

[5] P. Arminjon, *Two-dimensional finite volume extension of the Nessyahu-Tadmor scheme for conservation laws*, private communication to A. Jameson, Princeton University, Princeton, NJ, 1992.

[6] P. Arminjon and M.C. Viallon, *Généralisation du schéma de Nessyahu-Tadmor pour une équation hyperbolique à deux dimensions d'espace*, C. R. Acad. Sci., Paris, Sér. I Math., 320 (1995), pp. 85–88.

[7] P. Arminjon, M.C. Viallon, and A. Madrane, *A finite volume extension of the Lax-Friedrichs and Nessyahu-Tadmor schemes for conservationlaws on unstructured grids*, revised version with numerical applications, Internat. J. Comp. Fluid Dynamics, 9 (1997), pp. 1–22.

[8] P. Arminjon and A. Dervieux, *Incremental non-oscillatory finite volume schemes for compressible flows*, in Proceedings 1st International Conference Computational Fluid Dynamics Society of Canada, Montréal, 1993, R. Baliga and W.G. Habashi, eds., CFD Society of Canada, 1993, pp. 91–102.

[9] P. Arminjon, D. Stanescu, and M.C. Viallon, *A two-dimensional finite volume extension of the Lax-Friedrichs and Nessyahu-Tadmor schemes for compressible flows*, in A Collection of Technical Papers, Proceedings of the 6th International Symposium on Comp. Fluid Dynamics, Lake Tahoe, NV, M. Hafez, ed., University of California, Davis, CA, 4, pp. 7–14.

[10] P. Arminjon, M.C. Viallon, and A. Madrane, (1995). *From Lax-Friedrichs to a multidimensional finite volume extension of the Nessyahu-Tadmor scheme for compressible flows*, in Proceedings International Conference on Numerical Methods for the Euler and Navier-Stokes Equations, Centre de Recherches Mathématiques, Université de Montréal, 1995, P. Arminjon and A. Dervieux, eds., final version of the Proceeding to be published in the American Math. Soc.–Centre de Recherches Mathématiques Conference Proceedings Series.

[11] P. Arminjon, A. Madrane, and M.C. Viallon, *Comparison of a finite volume version of the Lax-Friedrichs and Nessyahu-Tadmor schemes and a discontinuous finite element method for compressible flows on unstructured grids*, in Godunov's method for Gas Dynamics, Proceedings of symposium honoring S.K. Godunov, 1997, University of Michigan, Ann Arbor, MI, B. van Leer, ed., J. Comput. Phys., to appear.

[12] P. Arminjon and A. Madrane, *A mixed finite volume/finite element method for* 2*-dimensional compressible Navier-Stokes equations on unstructured grids*, in Proceedings 7th International Conference on Hyperbolic Problems, ETH Zürich, M. Fey and R. Jeltsch, eds., 1998.

[13] P. Arminjon and A. Madrane, *A staggered Lax-Friedrichs type mixed finite volume/finite element method for the simulation of viscous compressible flows on unstructured triangular grids*, AIAA Journal, to appear.

[14] S. Champier and T. Gallouët, *Convergence d'un schéma décentré amont pour une équation hyperbolique linéaire sur un maillage triangulaire*, RAIRO Modél. Math. Anal. Numer., 26 (1992), pp. 835–853.

[15] B. Cockburn, S. Hou, and C.W. Shu, *The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws.* IV*: The multidimensional case*, Math. Comp., 54 (1990), pp. 545–581.

[16] B. Cockburn, F. Coquel, and P. Le Floch, *An error estimate for finite volume methods for multi-dimensional conservation laws*, Math. Comp., 63 (1994), pp. 77–103.

[17] M. Geiben, *Convergence of MUSCL-Type Finite Volume Schemes on Unstructured Triangular Grids*, SFB 256 Preprint 318, University of Bonn, Germany, 1993.

[18] J. Jaffré and L. Kaddouri, *Discontinuous finite elements for the Euler equations*, in Proceedings 3rd International Conference on Hyperbolic Problems, Uppsala, Sweden, B. Engquist and B. Gustafsson, eds., Studentlitteratur, Lund, Sweden, 1991, pp. 602–610.

[19] D. Kröner, S. Noelle, and M. Rokyta, *Convergence of higher order upwind finite volume schemes on unstructured grids for scalar conservation laws in several space dimensions*, SFB 256 Preprint 368, Univ. of Bonn, Germany; Numer. Math., 71 (1993), pp. 527–560.

[20] D. Kröner and M. Rokyta, *Convergence of upwind finite volume schemes for scalar conservation laws in two dimensions*, SIAM J. Numer. Anal., 31 (1994), pp. 324–343.

[21] P.D. Lax and B. Wendroff, *Systems of conservation laws*, Comm. Pure Appl. Math., 13 (1960), pp. 217–237.

[22] B. van Leer, *Towards the ultimate conservative difference scheme* II*: Monotonicity and conservation combined in a second order scheme*, J. Comput. Phys., 14 (1974), pp. 361–370.

[23] B. van Leer, *Towards the ultimate conservative difference scheme* IV*: A new approach to numerical convection*, J. Comput. Phys., 23 (1977), pp. 276–299.

[24] X.D. Liu, *A maximum principle satisfying modification of triangle based adaptive stencils for the solution of scalar hyperbolic conservation laws*, SIAM J. Numer. Anal, 30 (1993), pp. 701–716.

[25] H. Nessyahu and E. Tadmor, *Non-oscillatory central differencing for hyperbolic conservation laws*, J. Comput. Physics, 87 (1990), pp. 408–463.

[26] S. Noelle, *Convergence of higher order finite volume schemes on irregular grids*, Adv. Comput. Math., 3 (1995), pp. 197–218.

[27] S. Noelle, *A note on entropy inequalities and error estimates for higher order accurate finite volume schemes on irregular families of grids*, Math. Comp., 65 (1996), pp. 1155–1163.

[28] J. Peraire, M. Vahdati, K. Morgan, and O.C. Zienkiewicz, *Adaptive remeshing for compressible flow computations*, J. Comput. Phys., 72 (1987), pp. 449–466.

[29] R.D. Richtmyer and K.W. Morton, *Difference Methods for Initial Value Problems*, 2nd ed., Interscience and John Wiley, New York, 1967.

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

[31] M.C. Viallon and P. Arminjon, *Convergence du schéma de Nessyahu-Tadmor sur des maillages non-structurés pour une équation hyperbolique linéaire bidimensionnelle*, Rapport de recherche 183, Equipe d'Analyse Numérique, Universités de Lyon–St-Étienne, 1994.

[32] M.C. Viallon and P. Arminjon, *Convergence of a finite volume extension of the Nessyahu-Tadmor Scheme on unstructured grids for a two-dimensional linear hyperbolic equation*, Res. Report CRM-2239, Centre de Recherches Mathématiques, Univ. de Montréal, Montréal, Canada, revised version, 1996.

[33] P. Arminjon, M.C. Viallon, A. Madrane, and H. Kaddouri, *Discontinuous finite elements and a* 2*-dimensional finite volume generalization of the Lax-Friedrichs and Nessyahu-Tadmor schemes for compressible flows on unstructured grids*, in Computational Fluid Dynamics Review 1998, Vol. 1, M. Hafez and K. Oshima, eds., World Scientific, 1998, pp. 241–261.

[34] INRIA and GAMNI-SMAI, *Workshop on Hypersonic Flows for Reentry Problems*, Problem 6: Flow over a double ellipse, test case 6.1: Non-Reactive Flows, Antibes, France, 1990.

[35] M.J. Castro Diaz and F. Hecht, *Anisotropic Surface Mesh Generation*, INRIA Res. Rep. 2672, INRIA, Le Chesnay, France, 1995.

[36] G.S. Jiang and E. Tadmor, *Non-oscillatory central schemes for multidimensional hyperbolic conservation laws*, UCLA CAM Report 96-36, UCLA, Los Angeles, CA, 1996, SIAM J. Sci. Comput., 19 (1998), pp. 1892–1917.

[37] D. Levy and E. Tadmor, *Non-oscillatory central schemes for the incompressible* 2*-D Euler equations*, Math Res. Lett., 4 (1997), pp. 1–20.