RAIRO
MODÉLISATION MATHÉMATIQUE
ET ANALYSE NUMÉRIQUE

GUY CHAVENT
BERNARDO COCKBURN

The local projection $P^0 - P^1$-discontinuous-Galerkin finite element method for scalar conservation laws

RAIRO – Modélisation mathématique et analyse numérique, tome 23, n° 4 (1989), p. 565-592.

<http://www.numdam.org/item?id=M2AN_1989__23_4_565_0>

MATHEMATICAL MODELLING AND NUMERICAL ANALYSIS
MODÉLISATION MATHÉMATIQUE ET ANALYSE NUMÉRIQUE
(Vol 23, n° 4, 1989, p 565-592)

THE LOCAL PROJECTION P^0-P^1-DISCONTINUOUS-GALERKIN FINITE ELEMENT METHOD FOR SCALAR CONSERVATION LAWS (*)

by Guy CHAVENT (1) and Bernardo COCKBURN (2)

Abstract — *In this paper we introduce the Local Projection P^0P^1-Discontinuous Galerkin finite elemente method ($\Lambda\Pi P^0P^1$-scheme) for solving numerically scalar conservation laws This is an explicit method obtained by modifying the explicit Discontinuous Galerkin method introduced by G Chavent and G Salzano [3], via a simple local projection based on the monotonicity-preserving projections introduced by van Leer [13] The resulting scheme is an extension of Godunov scheme that verifies a local maximum principle, and is TVDM (total variation diminishing in the means) Convergence to a weak solution is proven We display numerical evidence that the scheme is an entropy scheme of order one even when discontinuities are present*

Resume — *Nous proposons une méthode d'éléments finis discontinus P^0P^1 avec projection locale pour le calcul des lois de conservation scalaires C'est un schéma explicite obtenu en modifiant la méthode de Galerkin discontinue explicite, introduite par G Chavent et G Salzano [3], a l'aide d'une simple projection locale basée sur les projections introduites par Van Leer [13] qui garde ses propriétés de conservation de la monotonicité Le schéma correspondant est une extension du schéma de Godunov qui vérifie un principe du maximum local, et est DVTM (diminue la variation totale sur les moyennes) Nous démontrons la convergence vers une solution faible, et fournissons des résultats numériques montrant que le schema est entropique d'ordre un même en presence de discontinuite*

1. INRODUCTION

In this paper we introduce and analyze a new finite element method, the local projection P^0P^1-Discontinuous Galerkin method ($\Lambda\Pi P^0P^1$-scheme), devised to solve numerically the scalar conservation law

$$\begin{aligned} &\partial_t u + \partial_x f(u) = 0\,, && \text{on } (0,T)\times\mathbb{R}\,,\\ &u(t=0) = u_0\,, && \text{in } \mathbb{R}, \end{aligned} \tag{1.1}$$

where the nonlinear function f is assumed to be C^1, and the initial data u_0 is assumed to belong to the space $L^1(\mathbb{R})\cap BV(\mathbb{R})$. This finite element

(*) Received in December 1987

(1) INRIA, Domaine de Voluceau, Rocquencourt, B.P 105, 78153 Le Chesnay Cedex, France , and CEREMADE, Université Paris-Dauphine, 75775 Paris Cedex 16, France

(2) IMA, University of Minnesota, 514 Vincent Hall, Minneapolis, Minnesota 55455, USA

M² AN Modelisation mathematique et Analyse numérique 0764-583X/89/04/565/28/$ 4 10

method is a predictor-corrector method whose prediction is given by the explicit P^0P^1-Discontinuous-Galerkin method introduced by G Chavent and G Salzano in [3], and whose correction is obtained by means of a very simple local projection, that we shall call $\Lambda\Pi$, based on the monotonicity-preserving projection introduced by Van Leer in [13] The basic idea of this method is to write the approximate solution u_h as the sum of a piecewise-constant function $\bar{u}_h$, and a function $\tilde{u}_h$ whose restriction to each element has zero-mean, and to consider the method as *a finite difference scheme for the means* $\bar{u}_h$ The function $\tilde{u}_h$ is considered as a parameter The local projection $\Lambda\Pi$ acts on the parameter $\tilde{u}_h$, and is constructed in order to preserve the conservativity, and enforce the stability of *the scheme for the means* $\bar{u}_h$ In the extreme case in which the parameter $\tilde{u}_h$ is set identically equal to zero by the local projection $\Lambda\Pi$, our scheme reduces to the well known Godunov scheme In the general case, *the scheme for the means* keeps the local maximum principle verified by Godunov scheme, and is TVD (total variation diminishing) Thus, the $\Lambda\Pi P^0P^1$-scheme is conservative, positive, and TVDM, i e total variation diminishing *in the means* We show that these properties, together with some properties of the local projection $\Lambda\Pi$, imply the existence of a subsequence converging to a weak solution of (1 1) Our numerical results indicate that if the *cfl*-number is mildly small enough, the scheme converges to the entropy solution with a rate of convergence equal to 1 in the $L^\infty(0, T, L^1_{\text{loc}})$-norm even in the presence of discontinuities

In 74 Le Saint and Raviart [9] introduced the Discontinuous-Galerkin method for solving the neutron transport equation

$$\mu\, \partial_t u + \nu\, \partial_x u + \sigma u = g$$

They choose their approximate function to be piecewise a polynomial of at most degree $k \geq 0$ in each of the variables t, and x In this way they obtained an implicit scheme, but they did not had to solve it globally Indeed, they proved that it is possible to solve it locally due to the fact that the direction of the propagation of the information, (μ, ν), is always the same In the general case, this is no longer true, for the local direction of propagation, $(1, f'(u))$, depends on values that have not been calculated yet ! To overcome this difficulty, in 1978 G Chavent and G Salzano [3] modified this method and obtained an explicit scheme that we shall call the P^0P^1-Discontinuous-Galerkin method In this method the t- and x-directions are treated in a different way the approximate solution is taken to be piecewise constant in time, and piecewise linear in space The two main advantages of the method are that it is explicit, and that it is very easy to generalize to the case of several space dimensions However, the scheme has a very restrictive stability condition — as we shall prove later —, and it

may not converge to the entropy solution in the case in which the nonlinearity f is nonconvex — as the numerical evidence we shall display indicates. In 1984 one of the authors [4] modified the scheme and obtained a scheme called the G-1/2 scheme, for which the convergence to the entropy solution was proven in the general case. A further development of the ideas involved in the construction of this scheme lead to the theory of quasi-monotone schemes for which $L^\infty(0, T; L^1(\mathbb{R}))$-error estimates have been obtained ; see [5]. The scheme we now introduce can be considered as a simplification of the initial G-1/2 scheme. This simplification leads to a very simple, and much cheaper algorithm, but complicates enormously the proof of its convergence. At each time step the $\Lambda\Pi P^0P^1$-scheme consists of two phases : in the first, a prediction is obtained by using the unchanged P^0P^1-method ; in the second, a correction is obtained by applying the local projection $\Lambda\Pi$ to it. This projection depends on a parameter, $\theta \in [0, 1]$, (θ may vary from element to element, but we have performed our numerical experiments with $\theta \equiv$ *constant*) and is based on the monotonicity-preserving local projections introduced by Van Leer in [13] : for $\theta \equiv 1$ the $\Lambda\Pi$ projection coincides with the one defined in [13, (66)] (thus, the $\Lambda\Pi P^0P^1$-scheme can be considered as a Discontinuous-Galerkin finite element version of the schemes introduced in [13]). One of the main contributions of this work is that we have proved that in fact the use of the local projection $\Lambda\Pi$ — originally devised in order to produce positive and monotonicity-preserving schemes — renders the scheme under consideration a TVDM scheme whose approximate solution verifies a local maximum principle ; see Proposition 3.2. These two properties allow us to conclude that the scheme is indeed total variation bounded (TVB) and that it generates a subsequence converging in $L^\infty(0, T; L^1_{\text{loc}}(\mathbb{R}))$ to a weak solution of (1.1) ; see Theorem 3.3. The problem of proving that the weak solution is indeed the entropy solution is still open. A result in this direction is the proof of the convergence of MUSCL-type semidiscrete schemes in the case of a convex (or concave) nonlinearity by Osher in [10]. Also, Johnson and Pitkaranta [7] have analized the Discontinuous-Galerkin method in the linear case.

An outline of the paper follows. In Section 2 we define the P^0P^1-Discontinuous-Galerkin method, we obtain the L^2 cfl-stability condition for the linear case, and display some numerical experiences that show the typical behavior of the method. In Section 3 we define the local-projection P^0P^1-Discontinuous-Galerkin method, we obtain some stability properties, prove the convergence to a weak solution, and test it in the same examples the P^0P^1-Discontinuous-Galerkin method was tested. We end with some concluding remarks in Section 4. In what follows, the P^0P^1-Discontinuous-Galerkin method will be referred to simply by the P^0P^1-scheme, and the local-projection P^0P^1-Discontinuous-Galerkin by the $\Lambda\Pi P^0P^1$-scheme.

2. THE P^0P^1-SCHEME

2.1. Preliminaries

As usual, the sets $\{t^n\}_{n=1,\ ,N}$, and $\{x_{i+1/2}\}_{i\in\mathbb{Z}}$ are partitions of $[0, T]$, and $\mathbb{R}$, respectively. We set $\Delta t^n = t^{n+1} - t^n$, and $\Delta x_i = x_{i+1/2} - x_{i-1/2}$, and denote by J^n, and I_i the intervals (t^n, t^{n+1}), and $(x_{i-1/2}, x_{i+1/2})$, respectively. Finally, K_i^n stands for the element $J^n \times I_i$, and h for $\sup_i\{\Delta x_i\}$. The space of functions of $L^1(\mathbb{R}) \cap BV(\mathbb{R})$ whose restriction to each interval I_i is linear will be denoted by W_h. By $I(a_1, \ldots, a_m)$ we shall denote the closed intervall $[\min\{a_1, \ldots, a_m\}, \max\{a_1, \ldots, a_m\}]$.

We shall need the following equality :

$$\begin{aligned} &-\int_{K_i^n} u \cdot \partial_t\varphi + \int_{\partial K_i^n} u \cdot \varphi \cdot n_t \\ &-\int_{K_i^n} f(u) \cdot \partial_x\varphi + \int_{\partial K_i^n} f(u) \cdot \varphi \cdot n_x = 0\ , \quad \forall\varphi \in C^{\,1}(K_i^n)\ , \end{aligned} \tag{2.1}$$

where $n = (n_x, n_t)$ is the outward unit normal to ∂K_i^n. This equality is obtained by simply multiplying (1.1) by φ and formally integrating by parts.

We recall the definition of the Godunov flux h^G associated to the function h :

$$\begin{aligned} h^G(w, w) &= h(w)\ , \\ h^G(w, v) &= h(\xi)\ , \end{aligned} \tag{2.2}$$

where ξ is *any* point $\in I(w, v)$ such that :

$$(h(\xi) - h(c)) \cdot \operatorname{sign}(w - v) \leq 0\ , \quad \forall c \in I(w, v)\ .$$

See Osher [11], and Brenier and Osher [1] for further details.

We shall also need the concept of viscosity of a numerical flux f^{fe} associated to f:

$$vis^{fe}(w, v) = \begin{cases} 0\ , & \text{if } w = v\ , \\ \dfrac{f(w) - 2\,.\,f^{fe}(w, v) + f(v)}{w - v}\ , & \text{otherwise}\ . \end{cases} \tag{2.3}$$

See for example Tadmor [12].

Finally, we define the *cfl*-number as follows :

$$cfl = \sup_{i\in\mathbb{Z},\, n=1,\ ,N} \frac{\Delta t^n}{\Delta x_i} \cdot \|f'\|_{L^\infty(C(u_0))} \tag{2.4}$$

where $C(u_0)$ is the convex hull of the range of the initial data u_0.

2.2. Definition of the P^0P^1-Scheme

Before defining the P^0P^1-scheme we need to introduce the finite dimensional space V_h. A function u_h is said to belong to V_h if :

(*a*) In each element K_l^n the approximate solution u_h is constant in time and linear in space : $u_h\big|_{K_l^n} \in P^0(J^n) \times P^1(I_l)$; i.e.,

$$u_h(t, x) = \bar{u}_l^n \varphi^0(s) + \tilde{u}_l^n \varphi^1(s) , \quad (t, x) \in K_l^n ,$$

where

$$\begin{aligned} &s = (x - x_l)/\Delta x_l , \\ &\varphi^0(s) = 1 , \\ &\varphi^1(s) = 2s , \qquad \forall s \in \left(-\frac{1}{2}, \frac{1}{2}\right) ; \end{aligned} \tag{2.5a}$$

(*b*) The *trace* of u_h in ∂K is chosen as follows :

$$\begin{aligned} u_h(t^n, x) &= \xi^n(x) , \qquad x \in I_l , \\ u_h(t, x_{l+1/2}) &= \xi_{l+1/2}(t) , \quad t \in J^n , \end{aligned} \tag{2.5b}$$

where

$$\begin{aligned} \xi^n(x) &= u_h(t^n + 0, x) , \\ f(\xi_{l+1/2}(t)) &= f^G(u_h(t, x_{l+1/2} + 0), u_h(t, x_{l+1/2} - 0)) , \end{aligned} \tag{2.5b}$$

where f^G denotes the Godunov flux associated to the function f, respectively ; see (2.2).

We can now define the P^0P^1-scheme as follows :

(*a*) Set $u_h(t, x) = \mathbb{P}_h(u_0)(x)$ for $(t, x) \in J^0 \times \mathbb{R}$, where $\mathbb{P}_h$ is the L^2-projection on the space W_h ;

(*b*) For $t \in [t^1, T)$ the approximate solution u_h is determined by the following variational formulation (compare with (2.1)) :

$$\begin{aligned} &- \int_{K_l^n} u_h \cdot \partial_t \varphi + \int_{\partial K_l^n} \xi \cdot \varphi \cdot n_t \\ &- \int_{K_l^n} f(u_h) \cdot \partial_x \varphi + \int_{\partial K_l^n} f(\xi) \cdot \varphi \cdot n_x = 0 , \\ &\qquad \forall \varphi \in P^0(J^n) \times P^1(I_l) . \end{aligned} \tag{2.6}$$

We want to stress the fact *to chose the trace of* u_h *on* $J^n \times \{x_{l+1/2}\}$, $\xi_{l+1/2}(t)$, *in this finite element framework is equivalent to chose an upwinding technique in the framework of finite difference schemes.* Also, *the way of choosing the trace of* u_h *along* $\{t^n\} \times I_l$, $\xi^n(x)$, *determines whether the*

scheme is explicit or implicit. Indeed, it is easy to check that with the following choice of $\xi^n(x)$

$$\xi^n(x) = u_h(t^n - 0, x) , \quad x \in I_i ,$$

the scheme (2.6) is implicit, even if u_h is piecewise constant in time !

As this point it is convenient to point out that our variational formulation (2.6) is strongly related to the one used by Le Saint and Raviart [9, (3.11)] to introduce the Discontinuous Galerkin method. In fact, after a simple integration by parts (2.6) can be rewritten as follows :

$$\int_{K_i^n} [\partial_t u_h + \partial_x f(u_h)] \cdot \varphi + \int_{\partial K_i^n} [\xi - u_h, f(\xi) - f(u_h)] \cdot n = 0 ,$$

$$\forall \varphi \in P^0(J^n) \times P^1(I_i) , \tag{2.6'}$$

that has the same formal structure than [9, (3.11)].

In terms of the degrees of freedom $\{\bar{u}_i^n, \tilde{u}_i^n\}_{i \in \mathbb{Z}, n = 0, \ldots, N}$ the $P^0 P^1$-scheme (2.5)-(2.6) reads :

(*a*) The degrees of freedom of the initial data are computed as follows :

$$\bar{u}_i^0 = \int_{I_i} u_0(s)\, ds/\Delta x_i ,$$

$$\tilde{u}_i^0 = 6 \int_{I_i} (s - x_i)\, u_0(s)\, ds/\Delta x_i^2 ; \tag{2.7a}$$

(*b*) For $n = 0, \ldots, N-1$ the degrees of freedom of u_h^{n+1} are obtained as the solution of :

$$(\bar{u}_i^{n+1} - \bar{u}_i^n)/\Delta t^n + (f_{i+1/2}^{P^0 P^1, n} - f_{i-1/2}^{P^0 P^1, n})/\Delta x_i = 0 ,$$

$$(\tilde{u}_i^{n+1} - \tilde{u}_i^n)/\Delta t^n + 3(f_{i+1/2}^{P^0 P^1, n} + f_{i-1/2}^{P^0 P^1, n})/\Delta x_i -$$

$$- 6 \left\{ \int \int_{K_i^n} f(u_h(t, x))\, dt\, dx/(\Delta t^n \Delta x_i) \right\} \Delta x_i = 0 , \tag{2.7b}$$

where $f_{i+1/2}^{P^0 P^1, n} = f(\xi_{i+1/2}(t^n))$, and the integral over K_i^n is approximated by Simpson's quadrature rule.

2.3. The linear case : the L^2-stability condition

In the linear case $f(u) = u$, we have $f_{i+1/2}^{P^0 P^1, n} = \bar{u}_i^n + \tilde{u}_i^n$, and

$$6 \left\{ \iint_{K_i^n} f(u_h(t, x))\, dt\, dx/(\Delta t^n \Delta x_i) \right\} = \bar{u}_i^n .$$

Thus, in the case of a uniform grid the scheme reads as follows :

(*a*) Compute the degrees of freedom of $u_{0,h}$ by (2.2*c*) ;

(*b*) For $n = 0, \ldots, N-1$ compute the degrees of freedom of u_h^{n+1} as follows :

$$\bar{u}_i^{n+1} = (1 - cfl)\,\bar{u}_i^n + cfl\,\bar{u}_{i-1}^n - cfl\,(\tilde{u}_i^n - \tilde{u}_{i-1}^n),$$
$$\tilde{u}_i^{n+1} = 3\,cfl\,(\bar{u}_i^n - \bar{u}_{i-1}^n) + (1 - 3\,cfl)\,\tilde{u}_i^n - 3\,cfl\,\tilde{u}_{i-1}^n,$$

where $cfl = \Delta t/\Delta x$. Note that when $cfl = 1$ we have

$$\bar{u}_i^{n+1} = \bar{u}_{i-1}^n - [\tilde{u}_i^n - \tilde{u}_{i-1}^n],$$
$$\tilde{u}_i^{n+1} = \tilde{u}_i^n + [3\,cfl\,(\bar{u}_i^n - \bar{u}_{i-1}^n) - 2\,\tilde{u}_i^n - 4\,\tilde{u}_{i-1}^n],$$

and so, the scheme *does not follow the characteristics*, for the expressions between brackets are not identically zero. Moreover, the scheme is unstable in L^2 if the *cfl* is $O(1)$. However, if *cfl* is allowed to decrease as $h \downarrow 0$, the scheme can be made L^2-stable. More precisely, we have the following result.

PROPOSITION 2.1 : *Let* $\{u_h\}_{h \downarrow 0}$ *be the sequence of approximate solutions determined by the* P^0P^1*-scheme* (2.2) *in the case* $f(u) = u$. *Then, the scheme is* L^2*-stable, i.e., there exists a constant* C *independent of the discretization parameters, and the initial data* u_0 *such that*

$$\|u_h\|_{L^2(\mathbb{R})} \leq C \cdot \|u_0\|_{L^2(\mathbb{R})},$$

if and only if $cfl = O(h^{1/2})$ *as* $h \downarrow 0$.

The proof of this stability result is given in the appendix ; see also [2]. Roughly speaking, this result tells us that the scheme has serious difficulties in following the information at the right speed, and so, it must be helped by letting the numerical speed $(\Delta x/\Delta t)$ go to infinity as $O(h^{-1/2})$ as h goes to zero. As a direct consequence of this result, and the well known Equivalence Theorem of Lax, the P^0P^1-scheme converges strongly in $L^\infty(0, T ; L^2(\mathbb{R}))$ to the (unique) solution of (1.1).

However, in the nonlinear case the scheme may fail to converge to the entropy solution of (1.1), even under the condition $cfl = O(h^{1/2})$, as the numerical results of next Subsection show.

2.4. Some Numerical Experiments

In this Subsection we test the P^0P^1-scheme in six different problems for which we can calculate the exact solution. To avoid the possible influence of the boundary conditions on the behavior of the approximate solution, we have taken a fixed space domain $\Omega = (0, L)$ on which we have imposed periodic boundary conditions. Our test problems can then be defined by

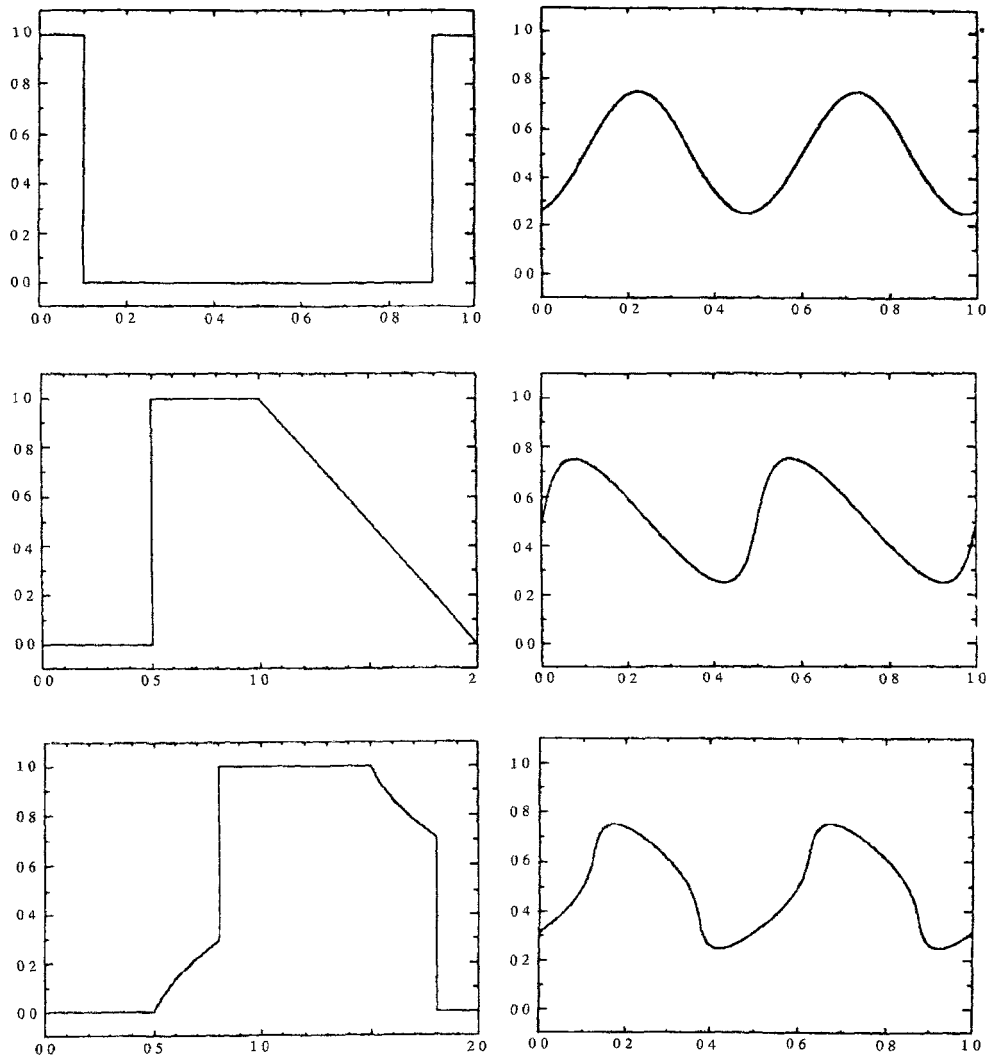
giving L, the final time T, the nonlinearity f, and the initial data u_0 on Ω ; see the table below.

TABLE 2.1
Definition of the test problems.

problem	Ω	T	$f(u)$	$u_0(x)$
1	(0,1)	0.5	u	$\begin{cases} 1, & \text{if } x \in (0.4, 0.6), \\ 0, & \text{otherwise.} \end{cases}$
2	(0,2)	0.5	$u(1-u)$	$\begin{cases} 1, & \text{if } x \in (0.5, 1.5), \\ 0, & \text{otherwise.} \end{cases}$
3	(0,2)	0.5	$\frac{1}{2}\frac{u^2}{u^2+(1-u)^2}$	$\begin{cases} 1, & \text{if } x \in (0.5, 1.5), \\ 0, & \text{otherwise.} \end{cases}$
4	(0,1)	0.1	u	$\frac{1}{2}(1+\frac{1}{2}sin(4\pi x))$
5	(0,1)	0.1	$u(1-u)$	$\frac{1}{2}(1+\frac{1}{2}sin(4\pi x))$
6	(0,1)	0.1	$\frac{1}{2}\frac{u^2}{u^2+(1-u)^2}$	$\frac{1}{2}(1+\frac{1}{2}sin(4\pi x))$

The solutions to these test problems are shown on figures 2.1. Let us point out, that problems 1, 2 and 3, for which the solution is not « smooth », have their « smooth » counterpart in problems 4, 5 and 6, respectively. We have constructed our three first test problems trying to obtain a reasonably wide set of singularities : the solution of problem 1 presents two contact discontinuities (i.e., the characterictics run parallel to them) ; the one of problem 2 (a Burgers problem) has a stationary discontinuity, a sonic point (i.e., a point u for which $f'(u) = 0$), and two moving « corners », i.e., two points at which the space derivative is discontinuous ; finally, the solution of problem 3 (a Buckley-Leveret problem) displays a couple of « rarefaction waves » (i.e., smooth regions) followed by shocks (i.e., the characteristics run into them).

We want to stress the fact that contact discontinuities are the most difficult to approximate. Roughly speaking, this is due to the following heuristic argument. If the scheme has not enough viscosity strong oscillations may appear around them, and if the scheme has too much viscosity the



Figures 2.1. — The exact solutions of the test problems.

approximate solution will be smoothed out without opposition from the characteristics. In the nonlinear case the characteristics may run into the shock and in this way they counterbalance the smoothing effect of the viscosity of the scheme.

In all the experiments we have used a uniform grid characterised by the discretization parameters Δt, and Δx. We have taken Δx very small, i.e.,

$\Delta x = \dfrac{1}{1\,024}$, in order to be in the asymptotic regime of the error, and we have used the linear L^2-stability condition of Proposition 2.1 to relate Δt, and Δx :

$$\Delta t = c_0 \cdot \Delta x^{3/2} . \tag{2.8}$$

For each of the test problems we have calculated two kinds of $L^1(\Omega')$-errors at time $t = T$,

$$\begin{aligned} e_{0,\Omega',T}(\Delta t, \Delta x) &= \left\| u(T) - \bar{u}_h(T) \right\|_{L^1(\Omega')}, \\ e_{1,\Omega',T}(\Delta t, \Delta x) &= \left\| u(T) - \bar{u}_h(T) \right\|_{L^1(\Omega')}, \end{aligned} \tag{2.9a}$$

in order to see the influence of he non-piecewise constant part of the approximate solution, $\tilde{u}_h$, in the *representation* of u_h. (We shall make this point more precise in the discussion of our numerical results below.) Also, we have estimated the rates of convergence α_0, and α_1 as follows

$$\begin{aligned} \alpha_{0,\Omega',T}(\Delta t, \Delta x) &= \ln\left(\frac{e_{0,\Omega'}(\Delta t/8, \Delta x/4)}{e_{0,\Omega'}(\Delta t, \Delta x)} \right) / \ln(4), \\ \alpha_{1,\Omega',T}(\Delta t, \Delta x) &= \ln\left(\frac{e_{1,\Omega'}(\Delta t/8, \Delta x/4)}{e_{1,\Omega'}(\Delta t, \Delta x)} \right) / \ln(4). \end{aligned} \tag{2.9b}$$

Note that if Δx is divided by 4, and Δt by 8 the stability condition (2.8) is verified. In problem 3 we have taken $f(u) = 0$, $\forall u < 0$, and $f(u) = 0.5$, $\forall u > 1$. In the table 2.2 we show the errors and their respective rates of convergence.

First, let us point out that in the three first problems the effect of $\tilde{u}_h$ in the error in negligible, whereas it is really important in the last three problems where the solutions are smooth. This indicates that *globally* the contribution of $\tilde{u}_h$ to the representation of the approximate solution u_h *is negligible when discontinuities in the solution or in its space derivative are present*, but it is important if the solution is smooth, for it reduces the error of the means, e_0, in $O(h^{1/2})$!

From table 2.2 we can also see that the rate of convergence is around 3/4 for the contact discontinuities of problem 1. In this case, the contact discontinuities of the exact solution at time $t = T = 0.5$ are located at $x = 0.1$, and $x = 0.9$, and the error is concentrated around them, as can be seen in the table below.

From table 2.2 we also see that the rate of convergence is 1 for problem 2. The exact solution has a stationary discontinuity at $x = 0.5$, a sonic point at $x = 1.5$, and two « corners » located at $x = 1$, and $x = 2$ for $t = T = 0.5$. Since the discontinuity is stationary (... and it is placed at the boundary of an

TABLE 2.2

L^1-errors and rates of convergence for the P^0P^1-scheme.

The quantities e_0, and e_1 are the errors $e_{0,\Omega',T}(\Delta t, \Delta x)$, and $e_{1,\Omega',T}(\Delta t, \Delta x)$, respectively, defined by (2.9*a*). The quantities α_0, and α_1 are the corresponding rates of convergence $\alpha_{0,\Omega',T}(\Delta t, \Delta x)$, and $\alpha_{1,\Omega',T}(\Delta t, \Delta x)$, respectively, defined by (2.9*b*). For all the problems $\Delta x = \frac{1}{1\,024}$, and Δt has been calculated from (2.8), where the constant c_0 was set equal to 1/2. The set Ω' has been taken equal to Ω defined in the table 2.1.

problem	$10^4 \cdot e_0$	α_0	$10^4 \cdot e_1$	α_1
1	58	0.7690	58	0.7717
2	3.5	1.0685	2.2	1.1149
3	747	0.0321	750	0.0406
4	4.3	1.0127	0.19	1.4990
5	4.2	1.0043	.04	1.6599
6	4.3	1.0124	0.24	1.4844

TABLE 2.3

Concentration of the error around the contact discontinuities.

The solution of the linear problem 1 presents two contact discontinuities located at $x = 0.1$, and at $x = 0.9$ at $t = T = 0.5$. In this case $\Omega = [0, 1]$.

Ω'	$\frac{e_{0,\Omega',T}(\Delta t,\Delta x)}{e_{0,\Omega,T}(\Delta t,\Delta x)}$	$\frac{e_{1,\Omega',T}(\Delta t,\Delta x)}{e_{1,\Omega,T}(\Delta t,\Delta x)}$
(0.0,0.2)	0.5063	0.4996
(0.8,1.0)	0.4937	0.5004

element !) it has been approximated extremely well by the scheme. In fact, almost all the error is essentially concentrated in between the « corner » points of the exact solution. In the table 2.4 we show the distribution of the error in that region.

Finally we see from table 2.2 that the scheme seems not to converge to the entropy solution of problem 3. In fact it does converge to the weak solution displayed in the figure 2.2. It is easy to check that its shocks travel at a lower

TABLE 2.4

Distribution of the error between the « corner » points.

The approximate solution of problem 2 posseses two « corner » points, i.e. points at which the space derivative is discontinuous, located at $x = 1.0$, and $x = 2.0$ for $t = T = 0.5$, and a sonic point at $x = 1.5$. Note how the error accumulates around the sonic point, and, more strongly, around the « corner » points.

Ω'	$\frac{e_{0,c_0,\Omega',T}(\Delta x)}{e_{0,c_0,\Omega,T}(\Delta x)}$	$\frac{e_{0,c_0,\Omega',T}(\Delta x)}{e_{0,c_0,\Omega,T}(\Delta x)} / \frac{\lvert\Omega'\rvert}{\lvert\Omega\rvert}$	$\frac{e_{1,c_0,\Omega',T}(\Delta x)}{e_{1,c_0,\Omega,T}(\Delta x)}$	$\frac{e_{1,c_0,\Omega',T}(\Delta x)}{e_{1,c_0,\Omega,T}(\Delta x)} / \frac{\lvert\Omega'\rvert}{\lvert\Omega\rvert}$
(0.9,1.0)	0.1077	2.1535	0.1746	3.4914
(1.0,1.1)	0.1033	2.0668	0.1510	3.0201
(1.1,1.4)	0.2130	1.4199	0.1162	0.7745
(1.4,1.6)	0.1520	1.5201	0.1165	1.1651
(1.6,1.9)	0.2130	1.4199	0.1162	0.7745
(1.9,2.0)	0.1033	2.0668	0.1510	3.0201
(0.0,0.1)	0.1077	2.1535	0.1746	3.4914

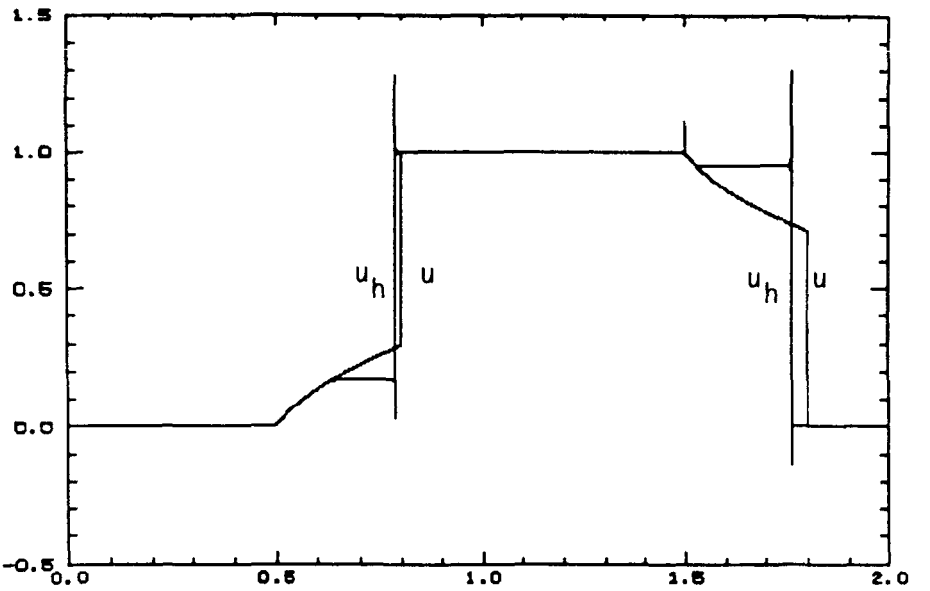


Figure 2.2. — The approximate solution determined by the $P^0 P^1$-scheme, u_h, does not converge to the entropy solution of the Buckley-Leverett problem 3, u.

speed than the shocks of the entropy solution. Thus, when the nonlinearity f is nonconvex the scheme may fail to catch the entropy shocks : they travel too quickly for the P^0P^1-scheme !

2.5. The Viscosity of the flux $f^{P^0P^1}$

In order to explain the behavior of the P^0P^1-scheme and motivate the modification that gives rise to the $\Lambda\Pi P^0P^1$-scheme we shall end this Section with some important remarks on the viscosity of the flux $f^{P^0P^1}$. By definition (2.3) the viscosity of the flux $f^{P^0P^1}$ is given by

$$vis_{i+1/2}^{P^0P^1,n} = \begin{cases} 0, & \text{if } \bar{u}_{i+1}^n = \bar{u}_i^n , \\ \dfrac{f(\bar{u}_{i+1}^n) - 2 \cdot f_{i+1/2}^{P^0P^1,n} + f(\bar{u}_i)}{\bar{u}_{i+1} - \bar{u}_i} , & \text{otherwise} . \end{cases} \tag{2.10}$$

Taking into account the definition of the viscosity associated to the Godunov flux we can rewrite this expression as follows :

$$vis_{i+1/2}^{P^0P^1,n} = vis_{i+1/2}^{G,n} + \delta vis_{i+1/2}^n ,$$

where

$$\delta vis_{i+1/2}^n = \begin{cases} 0, & \text{if } \bar{u}_{i+1}^n = \bar{u}_i^n , \\ 2\dfrac{f_{i+1/2}^{G,n} - f_{i+1/2}^{P^0P^1,n}}{\bar{u}_{i+1}^n - \bar{u}_i^n} , & \text{otherwise} . \end{cases}$$

In this way we can interpret $\delta vis_{i+1/2}^n$ as the amount of viscosity that the P^0P^1-flux is subtracting from the Godunov viscosity $vis_{i+1/2}^{G,n}$. If such amount is equal to zero the flux $f^{P^0P^1}$ reduces to the Godunov one, and the *scheme for the means* becomes Godunov scheme. If it is positive and not too big the scheme for the means will behave as a standard three-point monotone finite difference scheme. If it is negative, and too big in absolute value the viscosity associated to the flux $f^{P^0P^1}$ may be negative, equal to zero (as for the well known centered scheme), or positive but « too small », and we may be obliged to let the *cfl*-number to go to zero as $h \downarrow 0$ in order to achieve stability (in fact, the centered scheme is L^2-stable if and only if $cfl = O(h)$ as $h \downarrow 0$!). Proposition 2.1 indicates that this is the case for the P^0P^1-scheme ; roughly speaking, its viscosity is « too small ». As the viscosity is essentially a speed, this explains why the P^0P^1-scheme is « too slow ».

This fact is indeed quite reasonable as the following heuristic argument shows. For the sake of simplicity, let us consider the linear case $f(u) = u$. Assuming that $\bar{u}_{i+1}^n \neq \bar{u}_i^n$ we have

$$\delta vis_{i+1/2}^n = -2 \frac{\tilde{u}_i^n}{\bar{u}_{i+1}^n - \bar{u}_i^n}.$$

From this expression we can immediately conclude that the flux $f_{i+1/2}^{P^0P^1,n}$ produces *less* (resp., *more*) viscosity than the Godunov flux when $\tilde{u}_i^n$ and $(\bar{u}_{i+1}^n - \bar{u}_i^n)$ have the *same* (resp., *opposite*) sign. As the degrees of freedom $\bar{u}_i^n$, and $\tilde{u}_i^n$ are supposed to approximate $u(t^n, x_i)$, and $\partial_x u(t^n, x_i)\,\Delta x_i/2$, respectively, we expect $\delta vis_{i+1/2}^n$ to behave essentially like -1 ! Note that if we force $\tilde{u}_i^n$ to be $(\bar{u}_{i+1}^n - \bar{u}_i^n)/2$ then $\delta vis_{i+1/2}^n \equiv -1$. In this case the scheme for the means is nothing but the centered scheme ! Fortunately, the P^0P^1-scheme choses $\tilde{u}_i^n$ in a wiser way, and has a less restrictive *cfl*-condition that is, however, far from being optimal.

In order to improve the behavior of the P^0P^1-scheme we are going to exert a strong control on the size of the viscosity associated to $f^{P^0P^1}$. This will be done by means of the local projection $\Lambda\Pi$ we define and analyze in the next Section.

3. THE $\Lambda\Pi P^0P^1$-SCHEME

In this Section we first introduce the operator $\Lambda\Pi$, that is a local projection with which we shall improve the P^0P^1-scheme. Then, we define and analyze the $\Lambda\Pi P^0P^1$-scheme. Finally we test it on the same problems we tested the P^0P^1-scheme.

3.2. Definition of the operators $\Lambda\Pi$

We shall define the operators $\Lambda\Pi$ in three steps.

Let us first introduce the following projection operator :

$$\mathbb{P}_{[a,b]}(c) = \begin{cases} b, & \text{if } c > b, \\ c, & \text{if } c \in [a,b], \\ a, & \text{if } c < a. \end{cases} \tag{3.1a}$$

Next, to each function w_h of V_h, and each sequence $\{\theta_i\}_{i\in\mathbb{Z}}$ of real numbers we associate the set of intervals $\{\mathbb{I}_i\}_{i\in\mathbb{Z}}$ defined as follows

$$\begin{aligned} \mathbb{I}_i &= I(0, \eta_{i-1/2}^+) \cap I(0, \eta_{i+1/2}^-), \\ \eta_{i-1/2}^+ &= \theta_i \cdot (\bar{w}_i - \bar{w}_{i-1}), \\ \eta_{i+1/2}^- &= \theta_i \cdot (\bar{w}_{i+1} - \bar{w}_i). \end{aligned} \tag{3.1b}$$

Finally, we define the operator $\Lambda\Pi(w_h)$ as follows :

$$\Lambda\Pi(w_h) : V_h \to V_h , \qquad u_h \mapsto u_h^* , \tag{3.1c}$$

where, the degrees of freedom of u_h^* are defined as follows :

$$\bar{u}_i^* = \bar{u}_i ,$$
$$\tilde{u}_i^* = \mathbb{P}_{\mathbb{I}_i}^{\infty}(\tilde{u}_i) .$$

3.3. Properties of the operators $\Lambda\Pi$

We have the following result.

PROPOSITION 3.1 : *Let $w_h \in V_h$, and let $\{\theta_i\}_{i \in \mathbb{Z}}$ be an arbitrary sequence of real numbers. Set $u_h^* = \Lambda\Pi(w_h)(u_h)$, where $u_h \in V_h$, and $\Lambda\Pi(w_h)$ is defined by* (3.1). *Then*

(a) $$\Lambda\Pi(w_h)\, u_h^* = u_h^* ,$$

(b) $$\int_{\mathbb{R}} u_h^* = \int_{\mathbb{R}} \bar{w}_h .$$

If moreover we have that $\theta_i \in [0, 1]$, for $i \in \mathbb{Z}$, then

(c) $$u_h^*(x) \in I(\bar{w}_{i-1}, \bar{w}_i, \bar{w}_{i+1}), \quad \textit{for } x \in I_i .$$

The first property states that $\Lambda\Pi(w_h)$ is indeed a (locally defined) projection, and the conservativity property (b) that the « mass » of the projection u_h^* is equal to the mass of the *projecting* function w_h. Note that the positivity property (c) relates the resulting projection u_h^* solely with the piecewise constant part of w_h, $\bar{w}_h$

Proof: The properties (a, b) follow directly from the definition of the operator under consideration. By (3.1*c*), and the fact that $\theta_i \in [0, 1]$ we have

$$\bar{w}_i^* = \bar{w}_i ,$$
$$\tilde{u}_i^* \in I(\bar{w}_i - \bar{w}_{i-1}, 0, \bar{w}_{i+1} - \bar{w}_i) .$$

From this, and (2.5*a*) the maximum principle (c) follows easily.

The monotonicity-preserving projection used by Van Leer [13, (66)] uses $\theta_i \equiv 1$.

3.4. Definition of the $\Lambda\Pi P^0 P^1$-Scheme

Now we can define the $\Lambda\Pi P^0 P^1$-scheme as follows :

(a) Compute u^0 as follows : first, compute u_h^0 by (2.7*a*) ; then, compute $u_h^{*,0} = \Lambda\Pi(u_h^0)(u_h^0)$; finally, set $u_h^0 = u_h^{*,0}$.

(b) For $n = 1, \ldots, N-1$ compute u_h^{n+1} as follows : first, compute u_h^{n+1} given by the $P^0 P^1$-scheme, as in (2.7*b*) ; then, compute $u_h^{*,n+1} = \Lambda\Pi(u_h^{n+1})(u_h^{n+1})$; finally, set $u_h^{n+1} = u_h^{*,n+1}$.

3.5. Stability and Convergence Properties

The stability properties of our scheme are a direct consequence of the properties of the numerical flux $f^{P^0 P^1}$, and the properties of the family of local projections $\Lambda\Pi$.

PROPOSITION 3.2 : *Set* $\theta^* = \sup_{i \in \mathbb{Z}} \theta_i$, *and suppose that* $\theta^* \in [0, 1]$. *Then, for* $cfl \in \left[0, \dfrac{1}{1+\theta^*}\right]$ *the approximate solution* u_h *defined by the* $\Lambda\Pi P^0 P^1$*-scheme verifies the local maximum principles*

$$\bar{u}^{n+1} \in I(\bar{u}_{i-1}^n, \bar{u}_i^n, \bar{u}_{i+1}^n), \tag{3.2a}$$

$$u_h(t^{n+1}, x) \in I(\bar{u}_{i-2}^n, \bar{u}_{i-1}^n, \bar{u}_i^n, \bar{u}_{i+1}^n, \bar{u}_{i+2}^n), \quad \forall x \in I_i . \tag{3.2b}$$

Moreover, for $cfl \in [0, 1/2]$ *the scheme is TVDM (total variation diminishing in the means) ; i.e.,*

$$\|\bar{u}_h^{n+1}\|_{BV(\mathbb{R})} \leq \|\bar{u}_h^n\|_{BV(\mathbb{R})} . \tag{3.2c}$$

Proof : Consider the equations for the means :

$$\bar{u}_i^{n+1} = \bar{u}_i^n - (\Delta t^n / \Delta x_i)(f_{i+1/2}^{P^0 P^1, n} - f_{i-1/2}^{P^0 P^1, n}) .$$

If $(\bar{u}_{i+1}^n - \bar{u}_i^n) \cdot (\bar{u}_{i-1}^n - \bar{u}_i^n) \geq 0$, then $\tilde{u}_i^n = 0$ — by the definition of the $\Lambda\Pi$ projection, and we can apply to this equation the analysis of the standard Godunov scheme, see LeRoux [8], to obtain

$$\bar{u}^{n+1} \in I(\bar{u}_{i-1}^n + \tilde{u}_{i-1}^n, \bar{u}_i^n, \bar{u}_{i+1}^n - \tilde{u}_{i+1}^n), \quad \text{for} \quad cfl \in [0, 1] .$$

Again, the definition of the $\Lambda\Pi$ projection guarantees that this maximum principle implies (3.2*a*).

Now, if $(\bar{u}_{i+1}^n - \bar{u}_i^n) \cdot (\bar{u}_{i-1}^n - \bar{u}_i^n) < 0$, following Harten [6] and Osher [10, Lemma 2.4] we rewrite the equation for the means as follows :

$$\bar{u}_i^{n+1} = \bar{u}_i^n + C_i^n(\bar{u}_{i+1}^n - \bar{u}_i^n) + D_i^n(\bar{u}_{i-1}^n - \bar{u}_i^n) ,$$

where

$$C_i^n = \left(1 - \frac{\tilde{u}_{i+1}^n - \tilde{u}_i^n}{\bar{u}_{i+1}^n - \bar{u}_i^n}\right) \times \left[-\frac{\Delta t^n}{\Delta x_i} \cdot \frac{f^G(\bar{u}_{i+1}^n - \tilde{u}_{i+1}^n, \bar{u}_i^n + \tilde{u}_i^n) - f^G(\bar{u}_i^n - \tilde{u}_i^n, \bar{u}_i^n + \tilde{u}_i^n)}{(\bar{u}_{i+1}^n - \tilde{u}_{i+1}^n) - (\bar{u}_i^n - \tilde{u}_i^n)}\right],$$

and

$$D_i^n = \left(1 + \frac{\tilde{u}_{i-1}^n - \tilde{u}_i^n}{\bar{u}_{i-1}^n - \bar{u}_i^n}\right) \times \left[-\frac{\Delta t^n}{\Delta x_i} \cdot \frac{f^G(\bar{u}_i^n - \tilde{u}_i^n, \bar{u}_i^n + \tilde{u}_i^n) - f^G(\bar{u}_i^n - \tilde{u}_i^n, \bar{u}_{i-1}^n + \tilde{u}_{i-1}^n)}{(\bar{u}_{i-1}^n + \tilde{u}_{i-1}^n) - (\bar{u}_i^n + \tilde{u}_i^n)}\right].$$

To obtain (3.2*a*) in this case it is enough to prove that C_i^n and $D_i^n \in [0, 1]$. First, note that as the flux $f^{P^0P^1}$ is nonincreasing in its first argument, and nondecreasing in its second one the expression between brackets in the definition of C_i^n and D_i^n are non-negative numbers.

Second, by the definition of the local projection $\Lambda\Pi$ we have

$$\frac{\tilde{u}_{i+1}^n}{\bar{u}_{i+1}^n - \bar{u}_i^u} \in [0, \theta^*],$$

$$\frac{\tilde{u}_i^n}{\bar{u}_{i+1}^n - \bar{u}_i^u} \in [0, \theta^*],$$

and this implies that

$$\frac{\tilde{u}_{i+1}^n - \tilde{u}_i^n}{\bar{u}_{i+1}^n - \bar{u}_i^u} \in [-\theta^*, \theta^*].$$

From this it is easy to see that C_i^n and $D_i^n \in [cfl \cdot (1 - \theta^*), cfl \cdot (1 + \theta^*)]$. This proves (3.2*a*). To prove the maximum principle (3.2*b*) we simply combine (3.2*a*) with the following property :

$$u_h(t^{n+1}, x) \in I(\bar{u}_{i-1}^{n+1}, \bar{u}_i^{n+1}, \bar{u}_{i+1}^{n+1}), \quad \forall x \in I_i,$$

which is a direct consequence of the definition of the $\Lambda\Pi$ projection.

Finally, to obtain (3.2*c*) we only have to prove that $C_i^n + D_{i+1}^n \leqslant 1$; see Harten [6]. But

$$C_i^n \leqslant \left(1 - \frac{\tilde{u}_{i+1}^n - \tilde{u}_i^n}{\bar{u}_{i+1}^n - \bar{u}_i^n}\right) cfl,$$

$$D_{i+1}^n \leqslant \left(1 + \frac{\tilde{u}_{i+1}^n - \tilde{u}_i^n}{\bar{u}_{i+1}^n - \bar{u}_i^n}\right) cfl,$$

and so $C_i^n + D_{i+1}^n \leqslant 2\, cfl$. This proves the result.

We point out that if $\theta^* \in [0, 1/2]$, as in [10], it is possible to improve the estimate (3.2*b*) and obtain the following one :

$$\|u_h^{n+1}\|_{BV(\mathbb{R})} \leq \|u_h^n\|_{BV(\mathbb{R})} . \tag{3.2b'}$$

This property ensures the compactness in $L^\infty(0, T ; L^1_{\text{loc}})$ of the sequence $\{u_h\}_{h \downarrow 0}$. However, by allowing θ^* to lie in $[0, 1]$ we do not loose this property, for the compactness in $L^\infty(0, T ; L^1_{\text{loc}})$ of the sequence $\{\bar{u}_h\}_{h \downarrow 0}$ *implies* the one of $\{u_h\}_{h \downarrow 0}$ as we shall see in the next convergence result.

THEOREM 3.3 : *Under the hypothesis of Proposition 3.2, the sequence* $\{u_h\}_{h \downarrow 0}$ *generated by the* $\Lambda\Pi P^0 P^1$*-scheme has a subsequence converging strongly in* $L^\infty(0, T ; L^1_{\text{loc}}(\mathbb{R}))$ *to a weak solution of* (1.1).

Proof: By Proposition 3.2 the sequence $\{\bar{u}_h\}_{h \downarrow 0}$ is bounded in the space $L^\infty(0, T ; L^1(\mathbb{R}) \cap BV(\mathbb{R}))$. Also, note that the flux $f^{P^0 P^1}$ *as a function of the means* is consistent with f, for we have

$$\begin{aligned} f^{P^0 P^1}_{i+1/2} &= f^G(\bar{u}_{i+1} - \tilde{u}_{i+1}, \bar{u}_i + \tilde{u}_i) \\ &= f^G(\bar{u}, \bar{u}) \\ &= f(\bar{u}) , \end{aligned}$$

whenever $\bar{u}_{i+1} = \bar{u}_i$ (remember that in this situation $\tilde{u}_{i+1}$, as well as $\tilde{u}_i$, are set equal to zero by the $\Lambda\Pi$-projection, see (3.1)). These two facts, together with the fact that the scheme for the means is written in conservation form :

$$(\bar{u}_i^{n+1} - \bar{u}_i^n)/\Delta t^n + (f^{P^0 P^1, n}_{i+1/2} - f^{P^0 P^1, n}_{i-1/2})/\Delta x_i = 0 ,$$

imply, by a standard argument, the convergence of a subsequence, $\{\bar{u}_{h'}\}_{h' \downarrow 0}$ to a weak solution of (1.1), u.

Moreover, as we have

$$\|\tilde{u}_h^{n+1}\|_{L^1(\mathbb{R})} \leq \theta^* \cdot \|u_0\|_{BV(\mathbb{R})} \cdot h ,$$

we have that not only $\{\bar{u}_{h'}\}_{h' \downarrow 0}$, but $\{u_{h'}\}_{h' \downarrow 0}$ converges to the limit u. This completes the proof.

We end this Subsection by pointing out that if in the definition of the numerical flux $f^{P^0 P^1}$ (2.5*b*) the Godunov flux is replaced by any two-point monotone flux both Proposition 3.2 and Theorem 3.3 remain valid, modulo a possible trivial change in the *cfl* condition.

3.6. Some Numerical Experiments

In this Subsection we test the $\Lambda\Pi P^0 P^1$-scheme in the same test problems in which we tested the $P^0 P^1$-scheme. We have considered the cases

$\theta_i \equiv 0, 1/2, 1$, in order to have an idea of the influence of the size of θ^*. (In this paper no attempt has been made to define θ_i as a function of u_h *and* the nonlinearity f). We recall that the $\Lambda\Pi P^0 P^1$-scheme with $\theta_i \equiv 0$ is nothing but Godunov scheme. Also, we have considered the cases $cfl = 1/2$, and $cfl = 1/8$ to see how this influence depends on the *cfl*-number. We have set $\Delta x = \frac{1}{1\,024}$, as for the P^0P^1-scheme.

Our numerical results are shown in the tables 3.1, 3.2 below. We have not displayed this time the error of the means, e_0, for it possesses essentially the same rate of convergence than the one of the error e_1, and it is also of the same order of magnitude. The rate of convergence has been estimated as follows :

$$\alpha_{1,\Omega',T}(\Delta t, \Delta x) = \ln \frac{e_{1,\Omega'}(\Delta t/2, \Delta x/2)}{e_{1,\Omega'}(\Delta t, \Delta x)} / \ln (2) .$$

In the case of the problems 4, 5, and 6 that have a smooth solution we can see that the best results have been obtained for $\theta_i \equiv 1/2$. Also, when the *cfl*-number diminishes from 1/2 to 1/8, the differences between the cases $\theta_i \equiv 1/2$, and $\theta_i \equiv 1$ become negligeable.

For the problem 1, the contact discontinuities has been better approximated when $\theta_i \equiv 1$. Moreover, it is interesting to note that when the *cfl*-number decreases, the performance of the scheme gets worse in the cases

TABLE 3.1

L^1-errors and rates of convergence for the $\Lambda\Pi P^0 P^1$-scheme for $cfl = 1/2$.

The quantity e_1 is the error $e_{0,\Omega',T}(\Delta t, \Delta x)$ defined by (2.9a). The quantity α_1 is the corresponding rate of convergence $\alpha_{1,\Omega',T}(\Delta t, \Delta x)$, defined above. For problems 1, 2, 3 we took $\Delta x = \frac{1}{1\,024}$, and $\Delta x = \frac{1}{1\,000}$ for problems 4, 5, 6. The set Ω' has been taken equal to Ω defined in the table 2.1.

	$\theta_i \equiv 0$		$\theta_i \equiv 1/2$		$\theta_i \equiv 1$	
problem	$10^4 \cdot e_1$	α_1	$10^4 \cdot e_1$	α_1	$10^4 \cdot e_1$	α_1
1	249	0.4996	14.25	0.9945	10.08	1.0000
2	23.96	0.8065	16.18	0.6815	187.2	0.1322
3	41.77	0.8465	6.95	1.1084	44.67	0.3578
4	6.27	0.9972	6.26	0.9940	14.19	0.8351
5	5.59	0.9711	1.21	0.9431	1.61	0.8611
6	8.57	0.9707	7.22	0.9858	16.26	0.7423

TABLE 3.2

L^1-errors and rates of convergence for the $\Lambda\Pi P^0 P^1$-scheme for $cfl = 1/8$.

	$\theta_t \equiv 0$		$\theta_t \equiv 1/2$		$\theta_t \equiv 1$	
problem	$10^4 \cdot e_1$	α_1	$10^4 \cdot e_1$	α_1	$10^4 \cdot e_1$	α_1
1	330	0.4998	62.62	0.8238	9.60	1.0001
2	32.06	0.7867	1.67	0.8003	23.73	0.0140
3	51.78	0.8385	13.66	0.9149	6.99	0.9540
4	10.96	0.9950	1.63	1.0738	1.77	0.9140
5	6.41	0.9722	0.31	1.1617	0.32	1.0454
6	13.53	0.9647	1.72	0.9246	1.84	1.0124

$\theta_t \equiv 0$, and $\theta_t \equiv 1/2$, but remains essentially the same when $\theta_t \equiv 1$. This observation led us to try to measure the deterioration of the contact discontinuities. We do that by studying how the measure of the set in which the approximate solution belongs to the interval [0.01 0.99] evolves with respect to the discretization parameter Δx, and the time t. More precisely, we set

$$\mu(u_h(t)) = \text{measure of } \{x : \bar{u}_h(t, x) \in [0.01, 0.99]\} ,$$

and we assume that $\mu(u_h(t))$ behaves like $(\Delta x)^{\alpha'} \cdot t^{\beta}$. We estimate α', and β as follows :

$$\alpha'(T, \Delta t, \Delta x) = \ln \left(\frac{\mu(u_h(T ; \Delta t/2, \Delta x/2))}{\mu(u_h(T ; \Delta t, \Delta x))} \right) / \ln (2) ,$$

$$\beta(T, \Delta t, \Delta x) = \ln \left(\frac{\mu(u_h(T/2 ; \Delta t, \Delta x))}{\mu(u_h(T ; \Delta t, \Delta x))} \right) / \ln (2) .$$

The results are shown in the table below.

We see that in fact $\alpha' \simeq \alpha$, as expected. Not also that in all the cases $\alpha' + \beta = 1$! This means that the more α' is smaller than 1, the more the approximation of the discontinuity deteriorates with time ; moreover, $\alpha' = 1$ implies there is no deterioration of the discontinuity. These results indicate that the smallest deterioration of the contact discontinuities occurs when $\theta_t \equiv 1$. Moreover, at least for $cl f = 1/2, 1/8$, *there seems to be no deterioration of the approximation of the discontinuity with time.*

TABLE 3 3

Deterioration of the approximation to the contact discontinuities

The quantities α', and β are the rates $\alpha'(T, \Delta t, \Delta x)$, and $\beta(T, \Delta t, \Delta x)$, respectively, defined above. We have taken $\Delta x = \frac{1}{1\,024}$, and $cfl = 1/2$, $= 1/8$.

θ_i	α'	β	α'	β
$\equiv 0$	0 49098635	0.50901365	0.47916784	0.52083216
$\equiv 1/2$	1.00000000	0.00000000	0.83399005	0.16600995
$\equiv 1$	1.00000000	0 00000000	1.00000000	0.00000000

For problem 2, where the nonlinearity is strictly concave, the choice $\theta_i \equiv 1/2$ seems to be the best. In the case $\theta_i \equiv 1$ the low rates of convergence indicate that the approximate solution is converging to a weak solution that is not the entropy one. See figures 3.1.

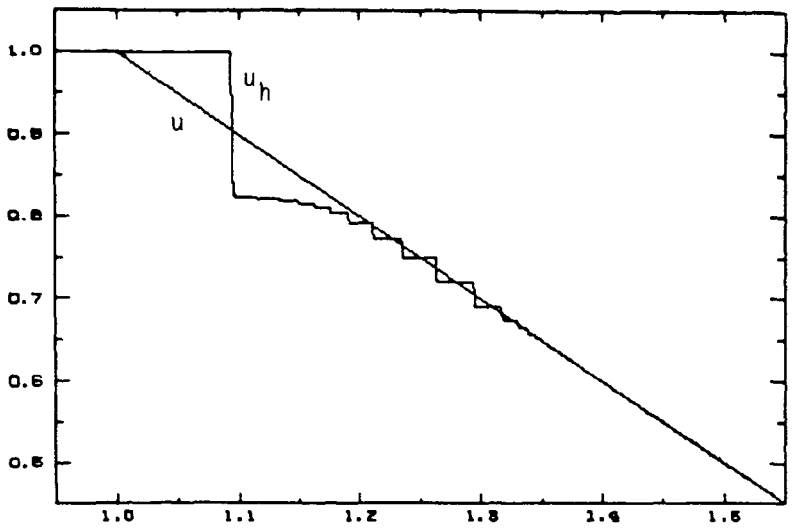


Figure 3.1a. — Convergence of the approximate solution determined by the $\Lambda\Pi P^0$ P^1-scheme with $\theta_i \equiv 1$, and $cfl = 1/2$ to a nonentropy weak solution of the Burgers problem 2.

The tendency of the P^0P^1-scheme to create nonentropy shocks can be seen here. From Table 3 2 we see that this phenomenon persists with $cfl = 1/8$ A more restrictive local projection, i e a smaller θ^*, is needed to counterbalance it, see next figure

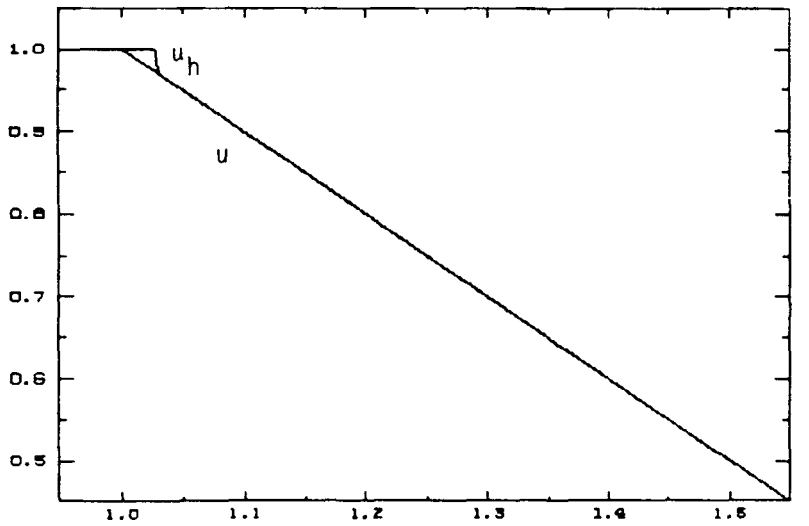


Figure 3.1*b*. — Convergence of the approximate solution determined by the $\Lambda\Pi P^0 P^1$-scheme with $\theta_i \equiv 1/2$, and $cfl = 1/2$ to the solution of the Burgers problem 2.

In this case the rate of convergence is 0.68, see Table 3.1. Note how the error accumulates around the corner points. The convergence is much better for $cfl = 1/8$, see next figure.

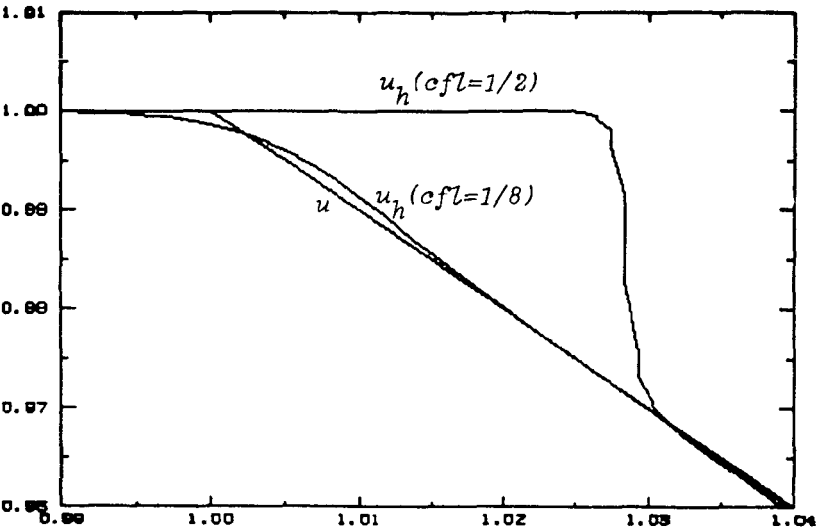


Figure 3.1*c*. — (Zoom on figure 3.1*b*) Convergence of the approximate solution determined by the $\Lambda\Pi P^0 P^1$-scheme with $\theta_i \equiv 1/2$, to the solution of the Burgers problem 2.

The approximation of the « corner points improves when cfl diminishes. The approximate solution convergences faster for $cfl = 1/8$ (the rate is 0.80 see Table 3.2) than for $cfl = 1/2$ (the rate is only 0.68, see Table 3.1).

Something similar seems to happen in problem 3, in the case $cfl = 1/2$; see figures 3.2. In this case the choice $\theta_t \equiv 1/2$ is definetly the best. However, for $cfl = 1/8$, the choice $\theta_t \equiv 1$ is the best. In figure 3.2*a* we show that in the case $cfl = 1/2$ and $\theta_t \equiv 1$ the $\Lambda\Pi P^0P^1$-scheme converges to a weak solution that is not the entropy one. We want to stress the fact that without computing the actual $L^1 =$ errors it would be impossible to detect this phenomenon, for the nonentropy shock of u_h is extremely near to the entropy one ! (Compare the scales of figures 2.2 and 3.2). In figure 3.2*b* we show that this situation is remediated by considering a smaller *cfl* number.

We end this Section by concluding that for the smooth cases, $\theta \simeq 1/2$ seems to be the best choice for $cfl \simeq 1/2$. However, the difference between the choices $\theta_t \equiv 1/2$, and $\theta_t \equiv 1$ becomes negligeable for $cfl = 1/8$. The scheme in these cases is a first order-accurate one. For approximating contact discontinuities the choice $\theta_t \equiv 1$ is the best. It also seems to be the optimal choice for small *cfl* and Buckley-Leverett type problems. However,

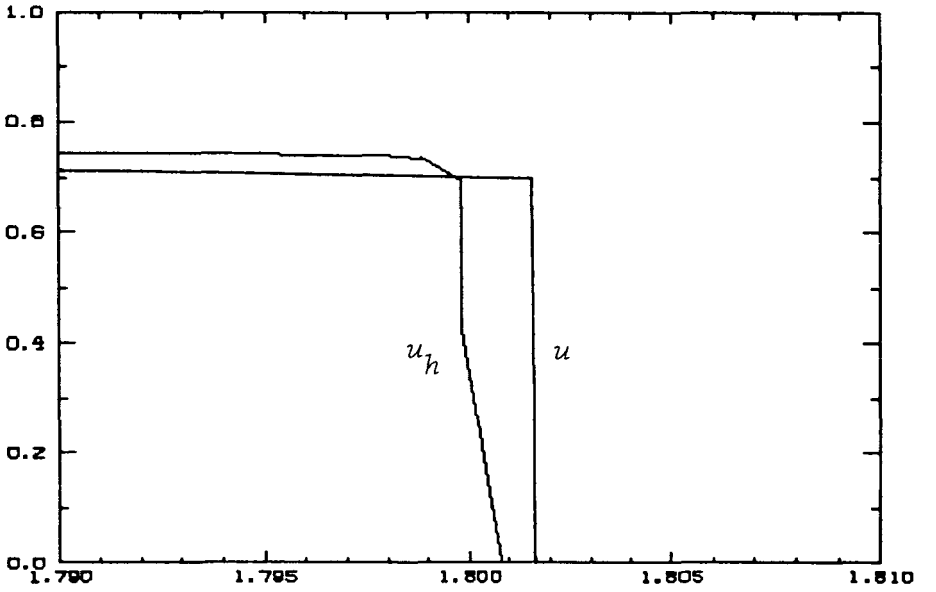


Figure 3.2a. — Detail of the convergence of the approximate solution determined by the $\Lambda\Pi P^0P^1$-scheme with $\theta_t \equiv 1$, and $cfl = 1/2$ to a non entropy solution of the Buckley-Leverett problem 3.

Note how the approximate solution is unable to catch the entropy shock. (However, the improvement with respect to the behavior of the P^0P^1-scheme is dramatic, see Table 2.2, and figure 2.2. In fact both the exact and the approximate solution would appear undistinguishable if ploted with the same scales of figure 2.2). This situation is much better for $cfl = 1/8$, see next figure

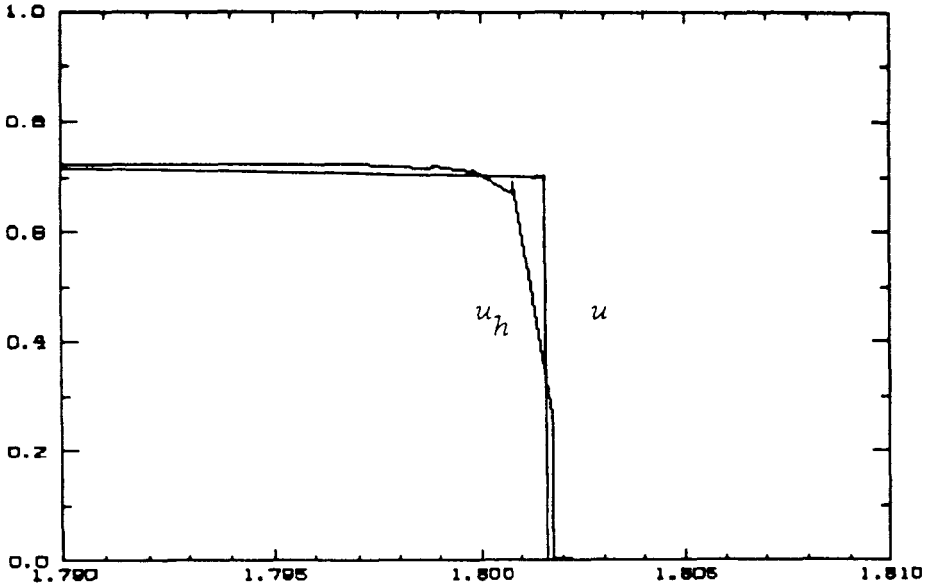


Figure 3.2*b*. — Detail of the convergence of the approximate solution determined by the $\Lambda\Pi P^0 P^1$-scheme with $\theta_i \equiv 1$, and $cfl = 1/8$ to the entropy solution of the Buckley-Leverett problem 3.

In this case the rate of convergence seems to be optimal : it is 0.95, see Table 3 2 Note that the shock has been captured in a single element.

for concave (or convex) nonlinearities this choice seems to give an approximation to a nonentropy solution ! (... as did the $P^0 P^1$-scheme). In this case, the choice $\theta_i \equiv 1/2$ is the best.

These results indicate that with an appropriate choice of the quantities θ_i (that must depend on the approximate solution u_h as well as on *the nonlinearity f*) the $\Lambda\Pi P^0 P^1$-scheme behaves as a first order accurate entropy scheme even in the presence of discontinuities.

4. CONCLUSION

We have introduced and analyzed the $\Lambda\Pi P^0 P^1$-scheme for the scalar conservation law (1.1). This is a finite element scheme obtained by a simple modification of the explicit discontinuous Galerkin scheme used by G. Chavent and G. Salzano [3], via a local projection based on one of the monotonicity-preserving projections introduced by van Leer [13]. The resulting scheme verifies a local maximum principle, and is also TVDM

(total variation diminishing in the means), a new property that allow us to prove the existence of a subsequence converging to a weak solution of (1.1). Our numerical results indicate that the scheme does converge to the entropy solution for small *cfl,* and is first order accurate even in the presence of discontinuities.

APPENDIX

PROOF OF PROPOSITION 2.1

We shall proceed in several steps. As usual, we assume that $\Delta t^n \equiv \Delta t$, and that $\Delta x \equiv h$. We shall only outline the proof. The reader is refered to [2] for details.

1. The Discrete Fourier Transform

Let u_h be an element of the space $W_h \cap L^2(\mathbb{R})$. We define its Discrete Fourier Transform (DFT) as follows :

$$[u_h](\theta) = \sum_{\iota \in \mathbb{Z}} [u_h]_\iota \, e^{j\iota\theta} h \,, \quad \forall \theta \in [-\pi, \pi] \,,$$

where $[u_h]_\iota^t = [\bar{u}_\iota, \tilde{u}_\iota / \sqrt{3}]$, and $j^2 = -1$. It is easy to verify that the DFT is an isometry from $W_h \cap L^2(\mathbb{R})$ to the space of 2π-periodic functions in $L^2(-\pi, \pi ; \mathbb{R}^2)$. In particular we have

$$\|u_h\|_{L^2(\mathbb{R})} = \| [u_h] \|_{L^2(-\pi, \pi, \mathbb{R}^2)} .$$

2. THE AMPLIFICATION MATRIX

We can rewrite the P^0P^1-scheme as follows

$$[u_h^{n+1}]_\iota = A_0 [u_h^n]_\iota + A_1 [u_h^n]_{\iota - 1} \,,$$

where A_0, and A_1 are two by two matrices. We then obtain easily that

$$[u_h]^{n+1} = G(\theta, cfl) \, [u_h]^n \,,$$

where the amplification matrix $G(\theta, cfl)$ is given by

$$\begin{aligned} G(\theta, cfl) &= A_0 + A_1 \cdot e^{j\theta} \,, \\ &= Id + cfl \cdot \Gamma(\theta) \,, \end{aligned}$$

and

$$\Gamma(\theta) = \begin{pmatrix} -1(-e^{j\theta}) & -\sqrt{3(1-e^{j\theta})} \\ \sqrt{3(1-e^{j\theta})} & -3\,.\,(1+e^{j\theta}) \end{pmatrix}.$$

It is then possible to show that,

$$\|u_h(T)\|_{L^2(\mathbb{R})} \leq M\,.\,\|u_0\|_{L^2(\mathbb{R})},$$

where

$$M = \sup_{n \leq T/\Delta t} \sup_{\theta \in [-\pi, \pi]} \|G(\theta, cfl)\|,$$

and

$$\|G\| = (\rho(G^* G))^{1/2},$$

with $\rho(A)$ being the spectral radius of the matrix A. In this way our problem is reduced to obtain the bound M.

3. ESTIMATING M

It can be shown that the matrix $G(\theta, cfl)$ is always diagonalisable. Let $P(\theta, cfl)$ be such that $P^{-1}GP$ is diagonal. Then it can be proved that

$$M \in [\nu^{-1}, \nu]\,.\,\sup_{n \leq T/\Delta t}(\rho(cfl))^n,$$

where

$$\nu = \sup_{cfl \in [0,1],\, \theta \in [-\pi, \pi]} \text{Condition of } P\ (\theta, cfl),$$
$$\rho(cfl) = \sup_{\theta \in [-\pi, \pi]} \rho(G(\theta, cfl)).$$

4. ESTIMATING THE EIGENVALUES OF G

The eigenvalues of the matrix G are given by

$$\lambda_{\pm} = 1 + cfl\,.\,\left(\frac{1}{2}\operatorname{tr}\Gamma \pm \left(\frac{1}{2}\operatorname{tr}\Gamma\right)^2 - \det\Gamma\right).$$

It can be proven that for any given $\varepsilon > 0$ there exists a cfl^* such that $\forall cfl \in [0, cfl^*]$:

$$|\lambda_-(\theta)| \leq 1, \quad \forall \theta \in [-\pi, \pi],$$
$$|\lambda_+(\theta)| \leq 1, \quad \forall \theta \in [-\pi, \pi] \backslash [-\varepsilon, \varepsilon].$$

Moreover, the modulus of the eigenvalue λ_+ in a small neighborhood of $\theta = 0$ is strictly bigger than zero, except for $\theta = 0$. More precisely, in such a neighborhood λ_+ can be expanded as follows :

$$\lambda_+(\theta) = 1 + cfl \cdot \left[-\frac{1}{72}\theta^4 + O(\theta^6) \right] + j \cdot cfl \cdot [-\theta + O(\theta^3)] .$$

From this, the following expression follows easily :

$$\sup_{\theta \in [-\varepsilon, \varepsilon]} |\lambda_+(\theta)| = 1 + \frac{9}{2} cfl^3 + o(cfl^3) .$$

In this way there exists $c_0 > 0$, and a cfl^* such that $\forall cfl \in [0, cfl^*]$:

$$\rho(cfl) \in [1 + c_0^{-1} cfl^3, 1 + c_0 cfl^3] .$$

5. CONCLUSION

All this imply that

$$M \in [\nu^{-1} \cdot \sup_{n \leq T/\Delta t}(1 + c_0^{-1} cfl^3)^n, \nu \cdot \sup_{n \leq T/\Delta t}(1 + c_0 cfl^3)^n] ,$$

and this interval remains bounded if and only if $cfl^3 = O(\Delta t)$, i.e., if and only if $cfl = O(h^{1/2})$. This proves the result.

REFERENCES

[1] Y. BRENIER and S. OSHER, *Approximate Riemman Solvers and Numerical Flux Functions,* ICASE report n° 84-63 (1984).

[2] G. CHAVENT and B. COCKBURN, *Convergence et Stabilité des Schémas LRG,* INRIA report.

[3] G. CHAVENT and G. SALZANO, *A finite Element Method for the 1D Water Flooding Problem with Gravity,* J. Comp. Phys., 45 (1982), pp. 307-344.

[4] B. COCKBURN, *Le Schéma G-k/2 pour les Lois de Conservation Scalaires,* Congrès National d'Analyse Numérique (1984), pp. 53-56.

[5] B. COCKBURN, *The Quasi-Monotone schemes for Scalar Conservation Laws,* IMA Preprint Series n° 263, 268 and 277. To appear in SIAM J. Numer. Anal.

[6] A. HARTEN, *On a class of high-resolution total-variation-stable finite-difference schemes,* SIAM J. Numer. Anal., 21 (1984), pp. 1-23.

[7] C. JOHNSON and J. PITKARANTA, *An Analysis of the Discontinuous Galerkin Method for a Scalar Hyperbolic Equation,* Math. of Comp., 46 (1986), pp. 1-26.

[8] A. Y. LEROUX, *A Numerical Conception of Entropy for Quasi-Linear Equations,* Math. of Comp., 31 (1977), pp. 848-872.

[9] P LESAINT and P A RAVIART, *On a Finite Element Method for Solving the Neutron Transport Equation, Mathematical Aspects of Finite Element in Partial Differential Equations,* Academic Press, Ed Carl de Boor, pp 89-145

[10] S OSHER, *Convergence of Generalized MUSCL Schemes,* SIAM J Numer Anal , 22 (1984), pp 947-961

[11] S OSHER, *Riemman Solvers, the Entropy Condition and Difference Approximations,* SIAM J Numer Anal , 21 (1984), pp 217-235

[12] E TADMOR, *Numerical Viscosity and the Entropy Condition for Conservative Difference Schemes,* Math Comp , 43 (1984), pp 369-381

[13] B VAN LEER, *Towards the Ultimate Conservative Scheme, II Monotonicity and Conservation Combined in a Second Order Scheme,* J Comput Phys , 14 (1974), pp 361-370