



Convergence of the Finite Volume Method for scalar conservation laws with discontinuous flux function

Sébastien Martin, Julien Vovelle

► To cite this version:

Sébastien Martin, Julien Vovelle. Convergence of the Finite Volume Method for scalar conservation laws with discontinuous flux function. ESAIM: Mathematical Modelling and Numerical Analysis, 2008, 42 (5), pp.699-727. 10.1051/m2an:2008023 . hal-00149007v1

HAL Id: hal-00149007

<https://hal.science/hal-00149007v1>

Submitted on 24 May 2007 (v1), last revised 20 Nov 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CONVERGENCE OF THE FINITE VOLUME METHOD FOR SCALAR CONSERVATION LAWS WITH DISCONTINUOUS FLUX FUNCTION

SÉBASTIEN MARTIN¹ AND JULIEN VOVELLE²

Abstract. We show the convergence of the Finite Volume method applied to the computation of the entropy solution of the Cauchy-Dirichlet Problem for a scalar conservation law with discontinuous (with respect to the unknown) flux function.

Résumé. On prouve la convergence de la methode Volume Fini, appliquée au calcul de la solution entropique du Problème de Cauchy-Dirichlet pour une loi de conservation scalaire dont le flux est discontinu en l'inconnue.

1991 Mathematics Subject Classification. 35L65 (76M12).

May 23, 2007.

1. INTRODUCTION

Let $\Omega \subset \mathbb{R}^d$ be an open Lipschitz subset of \mathbb{R}^d , $d \in \mathbb{N}^*$, let $A: \mathbb{R} \rightarrow \mathbb{R}$ be a function which is continuous except at some points of a finite set \mathcal{S} , let $u_0 \in L^\infty(\Omega)$; set $Q := (0, T) \times \Omega$. We consider the evolution problem

$$P(u_0) \begin{cases} u_t + \operatorname{div}(A(u)) = 0, & \text{on } Q, \\ u(0, \cdot) = u_0, & \text{on } \Omega, \\ u = 0, & \text{on } (0, T) \times \partial\Omega. \end{cases}$$

In this paper, the flux function A is allowed to have discontinuities of first type on a finite subset \mathcal{S} of \mathbb{R} . More precisely, let $s \in \mathcal{S}$, for $i = 1, \dots, d$, let

$$A_i(s^+) = \lim_{r \searrow s} A_i(r), \quad A_i(s^-) = \lim_{r \nearrow s} A_i(r),$$

and we assume that, at least for some i , $A_i(s^+) \neq A_i(s^-)$. In this paper, we propose an implicit finite volume method for the resolution of the problem, address its numerical analysis (well-posed character and convergence of the discrete problem) and present some numerical results.

Keywords and phrases: Finite Volume scheme, conservation law, discontinuous flux

¹ Université Paris-Sud XI / Laboratoire de Mathématiques - CNRS UMR 8628, Bât. 425, F-91405 Orsay Cedex, France

² Antenne de Bretagne de l'ENS Cachan et IRMAR / Campus de Ker Lann, 35170-Bruz, France

Theoretical background

The mathematical analysis of scalar conservation laws with a discontinuous flux function starts with the work of Carrillo [Car99, Car03]. In case where the flux function is at least continuous, there has been many works (we refer to Andreianov, Benilan and Kruřkov [BK96, ABK00] in case the flux function is continuous, and to the seminal works of Oleinik, Volpert and Kruřkov [Ol59, Vol67, Kru70] in case the flux function is at least Lipschitz continuous).

In case where the flux function is discontinuous, or continuous (at least has not the local Lipschitz regularity), there are phenomena of propagation (of the values of the initial datum) at infinite speed, even in an L^∞ framework. To be well-posed, the problem has therefore to be considered on a bounded domain. This in turn raises the question of the boundary conditions. For conservation laws with a “regular” flux function (at least locally Lipschitz continuous), the question has been examined first by Bardos, Leroux and Nedelec [BLN79], second by Otto [Ot96] in a more general framework. In [Car03], as in our present work, the boundary datum is homogeneous, i.e. constant (taken to be 0 after an obvious step of normalization by translation).

Numerical analysis

The analysis of the approximation of non-linear hyperbolic problems thanks to the finite volume method began in the mid 80’s, involving several authors as, for example, Cockburn, Coquel, LeFloch [CCL95], Szepessy [Sze91], Vila [Vil94], Kröner, Rokyta, Wierse [KRW96], Eymard, Gallouët, Herbin [EGH00]. In space dimension greater than 2, a major difficulty in the proof of convergence of the numerical approximation is the lack of strong compactness estimates for this approximation, like BV estimates for example (they may indeed not exist [Des04]). In the L^∞ framework, one is therefore led to use generalized notions of solution, as measure-valued solution (this goes back to DiPerna [DiP85]) or entropy process solution [EGH00].

Organization of the paper

The paper is organized as follows:

- In Section 2, we recall the main results (due to Carrillo [Car03]) on $P(u_0, f)$. This leads us to introduce an equivalent problem $\tilde{P}(u_0, f)$ which will be useful for the numerical analysis. We also outline the new difficulties that arise when dealing with discontinuous flux functions, in terms of numerical procedure: in particular, we state that classical Finite Volume methods cannot be applied to compute the numerical solution of $P(u_0, f)$. This leads us in a natural way to define and analyze an alternate numerical method.
- In Section 3, we present an implicit Finite Volume scheme and state related mathematical results, in particular the well-posed character of the scheme and the convergence of the method.
- In Section 4, numerical tests are presented: in particular, we focus on Riemann problems which allows us to concentrate on the analysis of certain shocks and rarefaction waves peculiar to the framework of discontinuous flux functions. We also extend the method to non-autonomous flux functions to a physical application that arises in lubrication theory.

Remark 1.1.

- In our numerical examples, the scalar conservation laws are source term free, therefore we have worked with such conservation laws, i.e. without source term. Notice that the extension to such an equation as

$$u_t + \operatorname{div}(A(u)) = f,$$

with, say, $f \in L^1(0, T; L^\infty(\Omega))$ is not straightforward (in particular the proof of Proposition 2.4 is more tricky in case $f \neq 0$, see [Car03]).

- We use the result of comparison-uniqueness of Carrillo [Car03] (see Section 3.3.3) and are therefore in the same framework. In particular the boundary conditions are homogeneous. The possible extension to non-homogeneous boundary conditions is questionable when they are spatially varying.

2. AN EQUIVALENT FORMULATION

We denote $\gamma_s = A(s^+) - A(s^-)$ and we assume that the flux has the following structure:

$$\begin{cases} A = \phi + J \\ \phi \in C(\mathbb{R}, \mathbb{R}^d), \quad \phi_i(0) = 0, \quad i = 1, \dots, d \\ J(r) = \sum_{s \in \mathcal{S}, s \leq 0} \gamma_s \text{sign}^-(r - s) + \sum_{s \in \mathcal{S}, s \geq 0} \gamma_s \text{sign}^+(r - s), \end{cases}$$

where $\gamma_s \in \mathbb{R}^d$ for any $s \in \mathcal{S}$ and

$$\text{sign}^+(r) = \begin{cases} 1 & \text{if } r > 0 \\ [0, 1] & \text{if } r = 0 \\ 0 & \text{if } r < 0 \end{cases}, \quad \text{sign}^-(r) = \begin{cases} 0 & \text{if } r > 0 \\ [-1, 0] & \text{if } r = 0 \\ -1 & \text{if } r < 0 \end{cases}.$$

In the following, we set $N = \text{card}(\mathcal{S})$.

The notion of “entropy solution” is used to select a solution [Kru70]. In the framework of a conservation law set on a bounded domain, we have to use “semi Kružkov entropy-flux pairs”. Thus, $u \mapsto (u - k)^\pm$ are the so-called “semi Kružkov entropies” defined by

$$(u - k)^+ = \begin{cases} u - k & \text{if } u \geq k, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad (u - k)^- = (k - u)^+.$$

For a regular flux function ϕ , functions $\text{sgn}_0^\pm(u - k)(\phi(u) - \phi(k))$ are the corresponding “semi Kružkov fluxes”, where $u \mapsto \text{sgn}_0^\pm(u)$ is the derivative of the function $u \mapsto u^\pm$ with value 0 at point 0. Of course, when dealing with discontinuous flux functions, the definition of entropy solution has to be slightly modified:

Definition 2.1. Let $u_0 \in L^\infty(\Omega)$. The function $(u, (\chi_j)_{j \in \mathcal{S}}) \in L^\infty(Q) \times L^\infty(Q)^N$ is an entropy solution of $P(u_0)$ if and only if

- (i) $\chi_{s < 0} \in \text{sign}^-(u - s)$, $\chi_{s > 0} \in \text{sign}^+(u - s)$, $\chi_0 \in \text{sign}^+(u) + \text{sign}^-(u)$ if $0 \in \mathcal{S}$,
- (ii) $\text{esslim}_{t \rightarrow 0} \|u(t, \cdot) - u_0\|_{L^1(\Omega)} = 0$,
- (iii) $u_t + \text{div} \left(\phi(u) + \sum_{s \in \mathcal{S}} \gamma_s \chi_s \right) = 0$ in $\mathcal{D}'(Q)$,
- (iv) for any $\varphi \in \mathcal{D}((0, T) \times \mathbb{R}^d)$, $\varphi \geq 0$, for any $k \in \mathbb{R}^\pm$ and for any $k_s \in \text{sign}_0^\pm(k - s)$,

$$\int_Q (u - k)^\pm \varphi_t + \left(\text{sign}_0^\pm(u - k) (\phi(u) - \phi(k)) + \sum_{s \in \mathcal{S}^\pm(k)} \gamma_s (\chi_s - k_s)^\pm \right) \cdot \nabla \varphi \geq 0,$$

with $\mathcal{S}^+(k) = \{s \in \mathcal{S}, s \geq k\}$ and $\mathcal{S}^-(k) = \{s \in \mathcal{S}, s \leq k\}$.

As proposed by Carrillo [Car03], we set

$$v := u + \sum_{s \in \mathcal{S}} \chi_s.$$

Then, $v \in (I + \mathcal{G})(u)$ where \mathcal{G} is a maximal monotone operator in \mathbb{R} :

$$\mathcal{G}(r) = \sum_{s \in \mathcal{S}^-(0)} \text{sign}^-(r - s) + \sum_{s \in \mathcal{S}^+(0)} \text{sign}^+(r - s),$$

and the function $u := g(v)$, where $g := (I + \mathcal{G})^{-1}$, satisfies

- (i) $I - g$ is nondecreasing, bounded and continuous,
- (ii) $g \in C^{0,1}(\mathbb{R})$, $g(0) = 0$, g is nondecreasing and $D(g^{-1}) = D(I + \mathcal{G}) = \mathbb{R}$.

However, the way to deal with the mathematical analysis led by Carrillo is based on a reformulation of Problem P(u_0), which is proved to be equivalent to

$$\tilde{P}(u_0) \begin{cases} g(v)_t + \operatorname{div}(B(v)) = 0, & \text{on } Q, \\ g(v)(0, \cdot) = u_0, & \text{on } \Omega, \\ g(v) = 0, & \text{on } (0, T) \times \partial\Omega. \end{cases}$$

where $u = g(v)$, $B(v) = A \circ g$, B being a Lipschitz continuous function (to be further discussed). According to Carrillo [Car99, Car03], we have the following definition:

Definition 2.2. Let $u_0 \in L^\infty(\Omega)$. A function $v \in L^\infty(Q)$ is an entropy solution of $\tilde{P}(u_0)$ if and only if

- (i) $\operatorname{esslim}_{t \rightarrow 0} \|g(v(t, \cdot)) - u_0\|_{L^1(\Omega)} = 0$,
- (ii) $(g(v))_t + \operatorname{div}(B(v)) = 0$ in $\mathcal{D}'(Q)$,
- (iii) for any $\varphi \in \mathcal{D}((0, T) \times \mathbb{R}^d)$, $\varphi \geq 0$, for any $k \in \mathbb{R}^\pm$,

$$\int_Q (g(v) - g(k))^\pm \varphi_t + \operatorname{sign}_0^\pm(v - k) (B(v) - B(k)) \cdot \nabla \varphi \geq 0.$$

The following two propositions describe how the definition can be enlarged, first by including the initial datum in the entropy inequality, second by considering a smaller set of constants k . This will be used in the proof of convergence of the numerical approximation.

Proposition 2.3. A function $v \in L^\infty(Q)$ is an entropy solution of $\tilde{P}(u_0)$ if and only if it satisfies the weak equation: for all $\varphi \in \mathcal{D}([0, T) \times \Omega)$,

$$\int_Q g(v) \varphi_t + B(v) \cdot \nabla \varphi + \int_\Omega u_0 \varphi(0, \cdot) = 0 \quad (1)$$

and the entropy condition: for any $\varphi \in \mathcal{D}([0, T) \times \mathbb{R}^d)$, $\varphi \geq 0$, for any $k \in \mathbb{R}^\pm$,

$$\int_Q (g(v) - g(k))^\pm \varphi_t + \operatorname{sign}_0^\pm(v - k) (B(v) - B(k)) \cdot \nabla \varphi + \int_\Omega (u_0 - g(k))^\pm \varphi(0, \cdot) \geq 0. \quad (2)$$

Proposition 2.4. A function $v \in L^\infty(Q)$ is an entropy solution of $\tilde{P}(u_0)$ if and only if it satisfies items (i)-(ii) of Definition 2.2 and the entropy condition

- (iii') for any $\varphi \in \mathcal{D}((0, T) \times \mathbb{R}^d)$, $\varphi \geq 0$, for any $k \in \mathbb{R}^\pm \setminus g^{-1}(\mathcal{S})$,

$$\int_Q (g(v) - g(k))^\pm \varphi_t + \operatorname{sign}_0^\pm(v - k) (B(v) - B(k)) \cdot \nabla \varphi \geq 0.$$

The proof of Proposition 2.3 and Proposition 2.4 are given in Appendix A.

The direct analysis of $\tilde{P}(u_0)$ is easier than the one of P(u_0) since regularity of the flux has been gained. Here, we propose a simple and alternative proof of the regularity of the flux:

Proposition 2.5. The flux B is a Lipschitz-continuous function (with a Lipschitz constant which will be denoted \mathcal{L}) w.r.t. v .

Proof. We express $\{\operatorname{sign}_0^\pm(u - s)\}_{s \in \mathcal{S}}$ as a function of v , by using the equality $u := g(v)$. Let us consider that the elements of \mathcal{S} are ordered as

$$s_{-k} < s_{-k+1} < \cdots < s_{-1} (< s_0 = 0) < s_1 < \cdots < s_k.$$

Then, by means of constructions we get, for any $s \in \mathcal{S}$, $\Psi_s(v) := \text{sign}_0^\pm(g(v) - s)$ with

$$\begin{aligned} \forall s_i \geq 0, \quad \Psi_{s_i}(v) &:= \begin{cases} 1 & \text{if } v \geq s_i + i \\ v - (s_i + i - 1) & \text{if } s_i + i - 1 < v < s_i + i \\ 0 & \text{if } v \leq s_i + i - 1 \end{cases}, \\ \forall s_{-i} \leq 0, \quad \Psi_{s_{-i}}(v) &:= \begin{cases} 0 & \text{if } v \geq s_{-i} - i \\ v - (s_{-i} - i) & \text{if } s_{-i} - i - 1 < v < s_{-i} - i \\ -1 & \text{if } v \leq s_{-i} - i - 1 \end{cases}. \end{aligned}$$

It appears that $\{\Psi_s\}_{s \in \mathcal{S}}$ are piecewise linear functions w.r.t. v . Additionally, g being also piecewise linear, the function

$$B := v \mapsto \phi \circ g(v) + \sum_{s \in \mathcal{S}, s \leq 0} \gamma_s \text{sign}_0^-(g(v) - s) + \sum_{s \in \mathcal{S}, s \geq 0} \gamma_s \text{sign}_0^+(g(v) - s)$$

is a Lipschitz-continuous function. \square

As a consequence of Proposition 2.5, we can define a map from $L^1(Q)$ into $L^1(Q) \times L^\infty(Q)^N$ by

$$\mathcal{T}(v) = (g(v), \{\Psi_s(v)\}_{s \in \mathcal{S}}).$$

The equivalence between $P(u_0)$ and $\tilde{P}(u_0)$ is stated in Theorem 2.6 and existence and uniqueness results are given in Theorem 2.7:

Theorem 2.6 (Carrillo, 2003). *Let v be an entropy solution of $\tilde{P}(u_0)$ then $\mathcal{T}(v)$ is an entropy solution of $P(u_0)$. Reciprocally, let $(u, \{\chi_s\}_{s \in \mathcal{S}})$ be an entropy solution of $P(u_0)$, then $v := u + \sum_{s \in \mathcal{S}} \chi_s$ is an entropy solution of $\tilde{P}(u_0)$.*

Theorem 2.7 (Carrillo, 2003). *There is a unique $g(v)$ such that v is an entropy solution of $\tilde{P}(u_0)$.*

In next sections, we focus on a Finite Volume method, which enable the computation of the entropy solution of $P(u_0)$. We show that the scheme is well-posed and convergent to the entropy solution. To this purpose, we use the regularity properties of the flux B . Thus, the formulation $\tilde{P}(u_0)$ is still the key-point of the method. Nevertheless, as a foreword of all the forthcoming sections, let us simply state that explicit (in time) finite volume schemes would not allow us to compute the solution for the following reasons:

- Explicit Finite Volume methods cannot be applied to compute the solution of $P(u_0)$ due to the lack of regularity of the flux: in particular, the implementation of explicit schemes is based on a Courant-Friedrichs-Lewy condition for which the Lipschitz constant of the flux plays a crucial role in terms of stability. Yet, the Lipschitz constant of $u \mapsto A(u)$ is not even defined.
- When dealing with Problem $\tilde{P}(u_0)$, one may think that difficulties have vanished since $v \mapsto B(v)$ is Lipschitz-continuous. However the difficulty is now that the inverse of g is non-univoque so that the value of v cannot be deduced from the value of $g(v)$.

Consequently, another strategy (use of implicit in time Finite Volume method) has been used in order to solve numerically Problem $P(u_0)$ or, alternatively, $\tilde{P}(u_0)$.

3. FINITE VOLUME SCHEME

3.1. Mesh and scheme

Mesh

Let \mathcal{T} be a family of disjoint connected polygonal subsets of Ω (called control volumes) such that $\bar{\Omega}$ is the union of the closures of the elements of this family and such that the common interface of two control volumes

is included in an hyperplane of \mathbb{R}^d . Let h be the size of the mesh: $h := \sup\{\text{diam}(K), K \in \mathcal{T}\}$. Notice that $h < +\infty$ for the set Ω is bounded. We suppose that the mesh has the following regularity property: there exists $\alpha > 0$ such that

$$\begin{cases} \alpha h^d \leq |K|, \\ |\partial K| \leq \alpha^{-1} h^{d-1}, \end{cases} \quad \forall K \in \mathcal{T}, \quad (3)$$

where $|K|$ is the d -dimensional Lebesgue measure of K and $|\partial K|$ is the $(d-1)$ -dimensional Lebesgue measure of ∂K . If two control volumes K and L have an edge in common, i.e. $K|L := \overline{K} \cap \overline{L} \neq \emptyset$, we say that L is a neighbor of K , denoted, quite abusively, $L \in \partial K$. We denote by n_{KL} the outward unit normal to K on $K|L$. To each edges σ of a control volume K which is part of $\partial\Omega$ is associated a fictitious control volume L , considered to belong to ∂K . We denote by \mathcal{T}^* the set of control volumes extended by this process.

Numerical fluxes

The numerical fluxes are computed by means of numerical flux functions $F_{KL}^n: \mathbb{R}^2 \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, $K \in \mathcal{T}$, $L \in \partial K$ with the following properties of

- consistency:

$$\forall v \in \mathbb{R}, \quad F_{KL}^n(v, v) = |K|L|B(v) \cdot n_{K|L}, \quad (4)$$

- conservation:

$$\forall (v, w) \in \mathbb{R}^2, \quad F_{KL}^n(v, w) = -F_{LK}^n(w, v), \quad (5)$$

- monotony:

$$\text{for all } w \in \mathbb{R}, \quad F_{KL}^n(\cdot, w) \text{ is a non-decreasing function on } \mathbb{R}, \quad (6)$$

- regularity:

$$\text{for all } w \in \mathbb{R}, \quad \frac{1}{|K|L|} F_{KL}^n(\cdot, w) \text{ is a } \mathcal{L}\text{-Lipschitz continuous function on } \mathbb{R}. \quad (7)$$

Notice that the properties above are supposed to be satisfied whatever $K \in \mathcal{T}$, $L \in \partial K$. In particular, by Eq. (5), we have:

- for all $v \in \mathbb{R}$, $F_{KL}^n(v, \cdot)$ is a non-increasing function on \mathbb{R} ,
- for all $v \in \mathbb{R}$, $\frac{1}{|K|L|} F_{KL}^n(v, \cdot)$ is a \mathcal{L} -Lipschitz continuous function on \mathbb{R} .

Also notice that we have the following remarkable identities:

$$\sum_{L \in \partial K} F_{KL}^n(v, v) = 0, \quad \forall v \in \mathbb{R}, \quad \forall n \in \mathbb{N}, \quad \forall K \in \mathcal{T}, \quad (8)$$

and, if $(a_K)_{K \in \mathcal{T}^*}$ is given,

$$\sum_{K \in \mathcal{T}} \sum_{L \in \partial K} F_{KL}^n(a_K, a_L) = \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial\Omega} F_{KL}^n(a_K, a_L), \quad \forall n \in \mathbb{N}. \quad (9)$$

Eq. (8) is a direct consequence of Eq. (4); the left hand-side of the equation is indeed

$$\int_{\partial K} B(v) \cdot n,$$

where n is the outward unit normal to K , and this last quantity is 0 by Green Formula. To prove Eq. (9), we first notice that any sum of the type

$$\sum_{K \in \mathcal{T}} \sum_{L \in \partial K} t_{KL}$$

can be decomposed as

$$\begin{aligned}
\sum_{K \in \mathcal{T}} \sum_{L \in \partial K \setminus \partial \Omega} t_{KL} + \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} t_{KL} &= \sum_{L \in \mathcal{T}} \sum_{K \in \partial L \setminus \partial \Omega} t_{KL} + \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} t_{KL} \\
&= \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \setminus \partial \Omega} t_{LK} + \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} t_{KL} \\
&= \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} t_{LK} + \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (t_{KL} - t_{LK}),
\end{aligned}$$

so that

$$\sum_{K \in \mathcal{T}} \sum_{L \in \partial K} (t_{KL} - t_{LK}) = \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (t_{KL} - t_{LK}).$$

In particular, we have the following lemma.

Lemma 3.1. *Let $t: \mathcal{T}^* \times \mathcal{T}^* \rightarrow \mathbb{R}$ be given. There is the identity*

$$\sum_{K \in \mathcal{T}} \sum_{L \in \partial K} t_{KL} = \frac{1}{2} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (t_{KL} + t_{LK}) + \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} t_{KL},$$

where, given a control volume $K \in \mathcal{T}$, ∂K_{int} is the set of the edges of K which are not contained in $\partial \Omega$.

We apply this last result to $t_{KL} := F_{KL}^n(a_K, a_L)$, which satisfies $t_{KL} = -t_{LK}$ by the property of conservation (5), to get Eq. (9).

Remark 3.2. The numerical fluxes are supposed to be globally Lipschitz although only their subsequent locally Lipschitz character will be of use, for the reason that, *a priori*, the numerical solution of the Finite Volume scheme stays in a ball of $L^\infty(\Omega)$, as proved in Lemma 3.6.

Finite Volume scheme

The entropy solution of $P(u_0)$ is approached with the help of the implicit Finite Volume scheme defined by

$$u_K^0 = \frac{1}{|K|} \int_K u_0(x) dx, \quad K \in \mathcal{T}, \quad (10)$$

$$g(v_K^{n+1}) = u_K^n - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1}), \quad n \in \mathbb{N}, \quad K \in \mathcal{T}, \quad (11)$$

$$u_K^{n+1} = g(v_K^{n+1}), \quad n \in \mathbb{N}, \quad K \in \mathcal{T}. \quad (12)$$

In Eq. (11), v_L^{n+1} is defined to be 0 if $L \in \mathcal{T}^* \setminus \mathcal{T}$.

Definition 3.3. A sequence $(u, v)_K^n$ is a sub-solution (resp. a super-solution) of Eq. (11)–(12) if Eq. (12) holds true while Eq. (11) is relaxed as an inequality with sign \leq (resp. sign \geq) and, also, $\tilde{v}_L^n \leq 0$ (resp. $\tilde{v}_L^n \geq 0$) for all fictitious control volume $L \in \mathcal{T}^* \setminus \mathcal{T}$.

Notation: The characteristic function of a set X is denoted by $\mathbf{1}_X$. Given a set of values (u_K^n) , $n \in \mathbb{N}$, $K \in \mathcal{T}$, we let u^n denote the piecewise constant function on Ω defined by

$$u^n = \sum_{K \in \mathcal{T}} u_K^n \mathbf{1}_K,$$

and also denote by u_h the piecewise constant function on Q defined by

$$u_h = \sum_{n \in \mathbb{N}} u^{n+1} \mathbf{1}_{[n\delta t, (n+1)\delta t)}.$$

3.2. Resolution of the Finite Volume scheme

A priori estimates

Lemma 3.4. *Assume $(u, v)_K^n$ is solution to Eq. (11)–(12) and $(\tilde{u}, \tilde{v})_K^n$ is sub-solution to Eq. (11)–(12). Then we have the L^1 -estimates*

$$\int_{\Omega} (\tilde{u}^{n+1} - u^{n+1})^+ \leq \int_{\Omega} (\tilde{u}^n - u^n)^+, \forall n \in \mathbb{N}. \quad (13)$$

Proof. By monotony of the numerical fluxes, the right hand-side of (11) defines a function which is non-decreasing with respect to the variable v_L^{n+1} . It is also non-decreasing with respect to the variable u_K^n , therefore we have

$$g(v_K^{n+1}) \leq u_K^n \top \tilde{u}_K^n - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1} \top \tilde{v}_L^{n+1})$$

where $a \top b := \max(a, b)$ for a, b real numbers. Similarly, we have

$$g(\tilde{v}_K^{n+1}) \leq u_K^n \top \tilde{u}_K^n - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(\tilde{v}_K^{n+1}, v_L^{n+1} \top \tilde{v}_L^{n+1}).$$

The value of $v_K^{n+1} \top \tilde{v}_K^{n+1}$ is either v_K^{n+1} or \tilde{v}_K^{n+1} but in any case, in virtue of Eq. (12) and of the fact that g is non-decreasing, we have

$$g(v_K^{n+1}) \top g(\tilde{v}_K^{n+1}) \leq u_K^n \top \tilde{u}_K^n - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1} \top \tilde{v}_K^{n+1}, v_L^{n+1} \top \tilde{v}_L^{n+1}).$$

Subtract both terms of the identity (11) to this inequality, and use the formula $(a - b)^+ = a \top b - b$ for a, b real numbers to deduce that

$$(\tilde{u}_K^{n+1} - u_K^{n+1})^+ \leq (\tilde{u}_K^n - u_K^n)^+ - \frac{\delta t}{|K|} \sum_{L \in \partial K} (F_{KL}^n(v_K^{n+1} \top \tilde{v}_K^{n+1}, v_L^{n+1} \top \tilde{v}_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_L^{n+1})).$$

Multiply by $|K|$, then sum this last inequality over $K \in \mathcal{T}$ and use Eq. (9) to get

$$\int_{\Omega} (\tilde{u}^{n+1} - u^{n+1})^+ \leq \int_{\Omega} (\tilde{u}^n - u^n)^+ - \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top \tilde{v}_K^{n+1}, v_L^{n+1} \top \tilde{v}_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_L^{n+1})).$$

By definition of v_L^{n+1} for $L \in \mathcal{T}^* \setminus \mathcal{T}$, the last term here is

$$-\delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top \tilde{v}_K^{n+1}, 0) - F_{KL}^n(v_K^{n+1}, 0))$$

which is non-negative by monotony of the numerical fluxes. This proves the lemma. \square

An immediate consequence of Lemma 3.4 is the following estimate, which we apply then to prove an L^∞ -estimate.

Corollary 3.5. *Assume $(u, v)_K^n$ is solution to Eq. (11)–(12) and $(\tilde{u}, \tilde{v})_K^n$ is sub-solution to Eq. (11)–(12). Then we have*

$$\int_{\Omega} (\tilde{u}^n - u^n)^+ \leq \int_{\Omega} (\tilde{u}^0 - u^0)^+. \quad (14)$$

Lemma 3.6. *Let $T > 0$ be given. Any solution u_h of Eq. (10)–(12) satisfies*

$$\|u_h\|_{L^\infty(\Omega \times (0, T))} \leq \|u_0\|_{L^\infty(\Omega)}. \quad (15)$$

Remark 3.7. We have $v_h \in (I + \mathcal{G})(u_h)$ where \mathcal{G} is a maximal monotone graph which sends bounded subsets of \mathbb{R} on bounded subsets of \mathbb{R} . In particular, the a priori estimate on the L^∞ -norm of u_h gives an a priori estimate on the L^∞ -norm of v_h .

Proof. Notice that, given $\alpha \leq 0$, there exists $\beta \leq 0$ such that $\alpha = g(\beta)$. For $K \in \mathcal{T}^*$, set

$$\tilde{u}_K^n := -\|u_0\|_{L^\infty(\Omega)},$$

and let $\tilde{v}_K^n \leq 0$ be such that Eq. (12) holds. Since \tilde{u}_K^n actually does not depend on K , we can choose \tilde{v}_K^n independent on K also. Then we have

$$\sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1}) = 0$$

for all $K \in \mathcal{T}$. Consequently, $(\tilde{u}, \tilde{v})_K^n$ is a sub-solution to Eq. (11)–(12). If u is a solution to Eq. (10)–(12), we deduce from Corollary 3.5 that

$$\int_{\Omega} (\tilde{u}^n - u^n)^+ \leq \int_{\Omega} (\tilde{u}^0 - u^0)^+$$

for $n \leq T/\delta t$. Since $(\tilde{u}^0 - u^0)^+ = 0$, we have $(\tilde{u}^n - u^n)^+ = 0$, which gives $u_h \geq -\|u_0\|_{L^\infty(\Omega)}$. We apply this result to $-u_h$ to conclude to (15). \square

Theorem 3.8. Assume $u_0 \in L^\infty(\Omega)$, then the Finite Volume scheme (10)–(12) admits a solution (u_h, v_h) , where u_h is uniquely determined.

Proof. Uniqueness is a consequence of Lemma 3.4. As in [EGH00], we use the topological degree to prove existence. Let $n \in \mathbb{N}$ be given, suppose that the existence of u^1, \dots, u^n satisfying Eq. (11)–(12) has been proved. Let $T > 0$, $T \geq (n+2)\delta t$, let P denote the number of elements of \mathcal{T} . We see Eq. (11)–(12) as an equation

$$(g - G)(v^{n+1}) = 0 \tag{16}$$

in \mathbb{R}^P (g acting component by component on \mathbb{R}^P). By Lemma 3.6 and Remark 3.7, Eq. (16) has no solution in $\partial B(0, R')$ for a given $R' > 0$ (depending on $\|u_0\|_{L^\infty(\Omega)}$). Here, $B(0, R')$ is the open ball of center 0, radius R' in \mathbb{R}^P . Actually, the proof of Lemma 3.6 readily adapts to show that, more generally, the equation

$$(g - \lambda G)(u^{n+1}) = 0, \quad \lambda \in [0, 1]$$

has no solution in $\partial B(0, R')$. By homotopy invariance of the topological degree, we have $\deg(g - G, 0, B(0, R')) = \deg(g, 0, B(0, R'))$. The latter is not 0 since g maps \mathbb{R} onto \mathbb{R} , therefore the Finite Volume scheme admits a solution. \square

3.3. Convergence of the Finite Volume method

3.3.1. Entropy inequalities

It is the conjunction of local discrete entropy inequalities (Lemma 3.9) with the consistency of the scheme (the latter proved with the help of the so-called weak BV inequality, see [EGH00] and Lemma 3.11) which shows that the solution u_h of the Finite Volume scheme is an approximate entropy solution to $P(u_0)$, as asserted in Theorem 3.10.

Lemma 3.9. Let (u_h, v_h) be the solution to Eq. (10)–(12). Then, for all $k \in \mathbb{R}$, we have: for all $n \in \mathbb{N}$, for all $K \in \mathcal{T}$ s.t. $\partial K \cap \partial\Omega = \emptyset$,

$$(u_K^{n+1} - g(k))^+ \leq (u_K^n - g(k))^+ - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k), \tag{17}$$

and Eq. (17) still holds true if $\partial K \cap \partial\Omega \neq \emptyset$ and $k \geq 0$.

Proof. The proof is very similar to the one of Lemma 3.4. By a monotony argument, we deduce from Eq. (11) that

$$u_K^{n+1} \leq u_K^n \top g(k) - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1} \top k).$$

Let $K \in \mathcal{T}$ s.t. $\partial K \cap \partial \Omega = \emptyset$. By Eq. (8) and a monotony argument, we have

$$g(k) = g(k) - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(k, k) \leq u_K^n \top g(k) - \frac{\delta t}{|K|} \sum_{L \in \partial K} F_{KL}^n(k, v_L^{n+1} \top k),$$

and the last inequality still holds if $\partial K \cap \partial \Omega \neq \emptyset$ and $k \geq 0$. By Eq. (12), and the monotony of g , we then deduce Eq. (17) according to the cases $v_K^{n+1} \leq k$ or $v_K^{n+1} > k$. \square

Theorem 3.10. *Assume that Eq. (3) holds. A solution (u_h, v_h) to Eq. (10)–(12) satisfies the following approximate entropy inequalities: for all real non-negative k , for all non-negative test function $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d)$,*

$$\int_Q (u_h - g(k))^+ \varphi_t + \text{sgn}_0^+(v_h - k)(B(v_h) - B(k)) \cdot \nabla \varphi + \int_\Omega (u_0 - g(k))^+ \varphi(0, \cdot) \geq \eta(h, \delta t, \varphi) \quad (18)$$

where $\eta(h, \delta t, \varphi) \rightarrow 0$ when $h, \delta t \rightarrow 0$. The approximate entropy inequality (18) is also satisfied for $k \leq 0$ if $\varphi \in \mathcal{D}([0, +\infty) \times \Omega)$ is non-negative.

Proof. We will prove Eq. (18) in the case $k \geq 0$, $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d)$ non-negative (the proof in the case where k is any real and φ vanishes on $\partial \Omega$ is actually more easy). For $K \in \mathcal{T}$, $n \in \mathbb{N}$, denote by φ_K^n be the mean value of φ over the space-time cell $[n\delta t, (n+1)\delta t) \times K$:

$$\varphi_K^n := \frac{1}{\delta t |K|} \int_{n\delta t}^{(n+1)\delta t} \int_K \varphi(t, x) dx dt.$$

Since u_h and v_h are piecewise constant functions, the first sum in Eq. (18) reads

$$\begin{aligned} & \int_Q (u_h - g(k))^+ \varphi_t + \text{sgn}_0^+(v_h - k)(B(v_h) - B(k)) \cdot \nabla \varphi \\ &= \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \int_{n\delta t}^{(n+1)\delta t} \int_K (u_K^{n+1} - g(k))^+ \varphi_t + \text{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \nabla \varphi \\ &= \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \left\{ (u_K^{n+1} - g(k))^+ \left(\int_K \varphi((n+1)\delta t, \cdot) - \int_K \varphi(n\delta t, \cdot) \right) \right. \\ & \quad \left. + \text{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \sum_{L \in \partial K} \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L} \right. \\ &= \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} ((u_K^n - g(k))^+ - (u_K^{n+1} - g(k))^+) \int_K \varphi(n\delta t, \cdot) - \sum_{K \in \mathcal{T}} (u_K^0 - g(k))^+ \int_K \varphi(0, \cdot) \\ & \quad \left. + \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} \text{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L}. \right. \end{aligned}$$

Use the discrete entropy inequalities (Lemma 3.9), i.e. multiply the inequality (17) by

$$\int_K \varphi(n\delta t, \cdot)$$

and sum over n, K , to see that the left hand-side of Eq. (18) is bounded from below by $\eta(h, \delta t, \varphi) := \eta^0(h, \delta t, \varphi) + \eta^c(h, \delta t, \varphi)$ where

$$\begin{aligned}\eta^0(h, \delta t, \varphi) &= \int_{\Omega} [(u_0 - g(k))^+ - (u_{0,h} - g(k))]^+ \varphi(0, \cdot), \\ \eta^c(h, \delta t, \varphi) &= \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} \operatorname{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L} \\ &\quad + \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) \frac{\delta t}{|K|} \int_K \varphi(n\delta t, \cdot).\end{aligned}$$

Since $u \mapsto (u - g(k))^+$ is a 1-Lipschitz continuous function, the first error term $\eta^0(h, \delta t, \varphi)$ is bounded by

$$\|\varphi(\cdot, 0)\|_{L^\infty(\Omega)} \|u_0 - u_{0,h}\|_{L^1(\Omega)},$$

therefore it tends to 0 with h . To prove that $\eta^c(h, \delta t, \varphi)$ tends to 0 with $h, \delta t$, we write $\operatorname{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) = B(v_K^{n+1} \top k) - B(k)$ and use Lemma 3.1 to get

$$\begin{aligned}& \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} \operatorname{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L} \\ &= \frac{1}{2} \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (B(v_K^{n+1} \top k) - B(v_L^{n+1} \top k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L} \\ &\quad + \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (B(v_K^{n+1} \top k) - B(k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L}.\end{aligned}$$

Let

$$\varphi_{K|L}^n := \frac{1}{\delta t |K|L|} \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi$$

denote the mean value of φ over $(n\delta t, (n+1)\delta t) \times K|L$. The property of consistency (4) shows that we have

$$\begin{aligned}& \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} \operatorname{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L} \\ &= \frac{1}{2} \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) - F_{KL}^n(v_L^{n+1} \top k, v_L^{n+1} \top k)) \varphi_{K|L}^n \\ &\quad + \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) - F_{KL}^n(k, k)) \varphi_{K|L}^n.\end{aligned}$$

In order to compare both terms in $\eta^c(h, \delta t, \varphi)$, first we transform the previous equality as

$$\begin{aligned}
& \sum_{n \in \mathbb{N}} \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} \operatorname{sgn}_0^+(v_K^{n+1} - k)(B(v_K^{n+1}) - B(k)) \cdot \int_{n\delta t}^{(n+1)\delta t} \int_{K|L} \varphi n_{K|L} \\
&= \frac{1}{2} \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k)) \varphi_{K|L}^n \\
&\quad - (F_{KL}^n(v_L^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k)) \varphi_{K|L}^n \\
&\quad + \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) - F_{KL}^n(k, k)) \varphi_{K|L}^n
\end{aligned}$$

and, second, setting

$$\varphi_K^n := \frac{1}{|K|} \int_K \varphi(n\delta t, \cdot),$$

we transform the second term of $\eta^c(h, \delta t, \varphi)$ as follows (we use, successively, the discrete version of the Stokes Formula on a control volume (8) and the formula of discrete integration by parts of Lemma 3.1):

$$\begin{aligned}
& \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) \varphi_K^n \\
&= \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)) \varphi_K^n \\
&= \frac{1}{2} \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)) \varphi_K^n \\
&\quad - (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_L^{n+1} \top k, v_L^{n+1} \top k)) \varphi_L^n \\
&\quad + \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)) \varphi_K^n.
\end{aligned}$$

Then $\eta^c(h, \delta t, \varphi)$ is decomposed as the sum of three terms $\eta_j^c(h, \delta t, \varphi)$, $j = 1, 2, 3$ where

$$\begin{aligned}
\eta_1^c(h, \delta t, \varphi) &= \frac{1}{2} \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)) (\varphi_K^n - \varphi_{K|L}^n), \\
\eta_2^c(h, \delta t, \varphi) &= -\frac{1}{2} \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_L^{n+1} \top k, v_L^{n+1} \top k)) (\varphi_L^n - \varphi_{K|L}^n), \\
\eta_3^c(h, \delta t, \varphi) &= \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) - F_{KL}^n(k, k)) \varphi_{K|L}^n \\
&\quad + (F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)) \varphi_K^n.
\end{aligned}$$

If $L \in \partial K \cap \partial \Omega$, then $v_L^{n+1} \top k = 0 \top k = k$ (recall that $k \geq 0$), hence, by monotony of the numerical flux, $F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) \geq F_{KL}^n(k, k)$ and we have

$$\eta_3^c(h, \delta t, \varphi) \geq \sum_{n \in \mathbb{N}} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} (F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) - F_{KL}^n(k, k)) (\varphi_{K|L}^n - \varphi_K^n).$$

We still denote by $\eta_j^c(h, \delta t, \varphi)$ the right hand-side of this inequality (this does not affect Eq. (18)). The differences $\varphi_{K|L}^n - \varphi_K^n$ are of order $h + \delta t$ [EGH00]:

$$\forall K \in \mathcal{T}, L \in \partial K, \quad |\varphi_{K|L}^n - \varphi_K^n| \leq C \|\varphi\|_{W^{1,\infty}(Q)}(h + \delta t).$$

Under the following weak BV -estimate (see Lemma 3.11), we then deduce that $\eta_j^c(h, \delta t, \varphi) = \mathcal{O}(\sqrt{h + \delta t})$, $j = 1, 2$ and $\eta_3^c(h, \delta t, \varphi) = \mathcal{O}(h + \delta t)$ (see [EGH00, Vov02] for the detail of the constants involved in these estimates). This shows that $\eta(h, \delta t, \varphi)$ tends to 0 when $h, \delta t \rightarrow 0$ and achieves the proof of Theorem 3.10. \square

3.3.2. Weak BV estimate

Lemma 3.11. *Assume that Eq. (3) holds. A solution (u_h, v_h) to Eq. (10)–(12) satisfies the following inequality: for all $N \in \mathbb{N}$, $k \geq 0$,*

$$\sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} |F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)| \leq \frac{C(\alpha, N \delta t)}{\sqrt{h + \delta t}}. \quad (19)$$

Proof. The proof is a variation on the proof of the BV estimate given in [EGH00]. Set $T := (N + 1)\delta t$. Let $R := \|u_0\|_{L^\infty(\Omega)}$. The a priori estimate (15) shows that $\|u_h\|_{L^\infty(Q)} \leq R$. Since $u_h = g(v_h)$, there also exists $R' \geq R$ such that $\|v_h\|_{L^\infty(Q)} \leq R'$. We denote by C various constants which depend on T , Ω , α and R' . Multiply the discrete equation (11) by $|K|v_K^{n+1}$ and sum the result over $n \leq N, K \in \mathcal{T}$ to obtain

$$\begin{aligned} & \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} v_K^{n+1} \left(\sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1}) \right) \\ &= - \sum_{n \leq N} \sum_{K \in \mathcal{T}} |K| (u_K^{n+1} - u_K^n) v_K^{n+1} \\ &= - \sum_{n \leq N} \sum_{K \in \mathcal{T}} |K| (g(v_K^{n+1}) - g(v_K^n)) v_K^{n+1} \end{aligned}$$

We change the indexation of the sum over n according to the formula

$$\sum_{n \leq N} H(v^{n+1}, v^n) = \sum_{n, m \in \Delta_N} H(v^m, v^n), \quad \Delta_N := \{n, m \leq N, |n - m| \leq 1, v^n < v^m\}, \quad (20)$$

which is valid for any $H(v, w) \in C(\mathbb{R}^2)$ which vanishes for $v = w$. We obtain

$$\sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} v_K^{n+1} \left(\sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - |K| f_K^n \right) = - \sum_{n, m \in \Delta_N} \sum_{K \in \mathcal{T}} |K| (g(v_K^m) - g(v_K^n)) v_K^m.$$

Then, for $w > v$, we use the inequality

$$(g(w) - g(v))w = w \int_v^w g'(s) ds \geq \int_v^w s g'(s) ds = - \int_v^w g + [s g(s)]_v^w$$

to obtain

$$\sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} v_K^{n+1} \left(\sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1}) \right) \leq \sum_{n, m \in \Delta_N} \sum_{K \in \mathcal{T}} |K| \int_{v_K^n}^{v_K^m} g - [s g(s)]_{v_K^n}^{v_K^m}.$$

By a new change of the indexation (the converse of (20)), we have

$$\begin{aligned} \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} v_K^{n+1} \left(\sum_{L \in \partial K} F_{KL}^n(v_K^{n+1}, v_L^{n+1}) \right) &\leq \sum_{n \leq N} \sum_{K \in \mathcal{T}} |K| \left(\int_{v_K^n}^{v_K^{n+1}} g - [sg(s)]_{v_K^n}^{v_K^{n+1}} \right) \\ &= \sum_{K \in \mathcal{T}} |K| \left(\int_{v_K^0}^{v_K^{N+1}} g - [sg(s)]_{v_K^0}^{v_K^{N+1}} \right) \leq C. \end{aligned} \quad (21)$$

Using discrete divergence equation (8), we infer

$$\sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} v_K^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_K^{n+1})) \leq C.$$

We then use the formula of discrete integration by parts of Lemma 3.1 to obtain

$$\begin{aligned} \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} [v_K^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_K^{n+1})) \\ - v_L^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_L^{n+1}, v_L^{n+1}))] \\ \leq C - 2 \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} v_K^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_K^{n+1})). \end{aligned}$$

The numerical fluxes being $\mathcal{L}|K|L|$ -Lipschitz continuous (see Eq. (7)), there is a bound

$$|v_K^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_K^{n+1}))| \leq C|K|L|.$$

Since

$$\sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} |K|L| = |\partial \Omega|,$$

we obtain the bound

$$\begin{aligned} \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} [v_K^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_K^{n+1}, v_K^{n+1})) \\ - v_L^{n+1} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_L^{n+1}, v_L^{n+1}))] \leq C. \end{aligned} \quad (22)$$

Now we proceed as in the beginning of the proof: we change the indexation over K, L (analogously to Eq. (20)) by the formula

$$\sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} H(v_K, v_L) = \sum_{(K, L) \in \mathcal{E}_{\text{int}}} H(v_K, v_L), \quad \mathcal{E}_{\text{int}} := \{K, L \in \mathcal{T}, v_K > v_L\}, \quad (23)$$

which is valid for any $H(v, w) \in C(\mathbb{R}^2)$ which vanishes for $v = w$, and, for $w > v$, use the identity

$$\begin{aligned} w (F_{KL}^n(w, v) - F_{KL}^n(w, w)) - v (F_{KL}^n(w, v) - F_{KL}^n(v, v)) \\ = \int_v^w \frac{d}{ds} (s (F_{KL}^n(w, v) - F_{KL}^n(s, s))) ds \\ = - \int_v^w s \frac{d}{ds} F_{KL}^n(s, s) ds + \int_v^w (F_{KL}^n(w, v) - F_{KL}^n(s, s)) ds \end{aligned}$$

to deduce from Eq. (22) the estimate

$$\begin{aligned} \sum_{n \leq N} \delta t \sum_{(K,L) \in \mathcal{E}_{\text{int}}} \int_{v_L^{n+1}}^{v_K^{n+1}} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(s, s)) ds \\ \leq C + \sum_{n \leq N} \delta t \sum_{(K,L) \in \mathcal{E}_{\text{int}}} \int_{v_L^{n+1}}^{v_K^{n+1}} s \frac{d}{ds} F_{KL}^n(s, s) ds. \end{aligned} \quad (24)$$

Use again a change of indexation (23) and the discrete integration formula in Lemma 3.1 to reduce the sum in the right hand-side of Eq. (24) to a sum over the boundary of Ω , as follows:

$$\begin{aligned} & \sum_{n \leq N} \delta t \sum_{(K,L) \in \mathcal{E}_{\text{int}}} \int_{v_L^{n+1}}^{v_K^{n+1}} s \frac{d}{ds} F_{KL}^n(s, s) ds \\ &= \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K_{\text{int}}} \int_{v_L^{n+1}}^{v_K^{n+1}} s \frac{d}{ds} F_{KL}^n(s, s) ds \\ &= \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K} \int_0^{v_K^{n+1}} s \frac{d}{ds} F_{KL}^n(s, s) ds - \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} \int_0^{v_K^{n+1}} s \frac{d}{ds} F_{KL}^n(s, s) ds \\ &= - \sum_{n \leq N} \delta t \sum_{K \in \mathcal{T}} \sum_{L \in \partial K \cap \partial \Omega} \int_0^{v_K^{n+1}} s \frac{d}{ds} F_{KL}^n(s, s) ds \text{ by Eq. (8)}. \end{aligned}$$

The last term here is bounded by a constant C . From Eq. (24), we eventually deduce

$$\sum_{n \leq N} \delta t \sum_{(K,L) \in \mathcal{E}_{\text{int}}} \int_{v_L^{n+1}}^{v_K^{n+1}} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(s, s)) ds \leq C,$$

and, by monotony of $v \mapsto F_{KL}^n(v, s)$,

$$\sum_{n \leq N} \delta t \sum_{(K,L) \in \mathcal{E}_{\text{int}}} \int_{v_L^{n+1}}^{v_K^{n+1}} (F_{KL}^n(v_K^{n+1}, v_L^{n+1}) - F_{KL}^n(v_K^{n+1}, s)) ds \leq C. \quad (25)$$

Set temporarily $\Psi(s) := -F_{KL}^n(v_K^{n+1}, s)$: this is a non-decreasing, Lipschitz continuous function. If it is (strictly) increasing, then we have, for $w > v$, and by the change of variable $\sigma := \Psi(s) - \Psi(v)$ (with inverse $s = \Phi(\sigma)$):

$$\int_v^w (\Psi(s) - \Psi(w)) ds = \int_0^{\Psi(w) - \Psi(v)} \sigma \Phi'(\sigma) d\sigma \geq \int_0^{\Psi(w) - \Psi(v)} \sigma \frac{1}{\text{Lip}(\Psi)} d\sigma = \frac{1}{2\text{Lip}(\Psi)} (\Psi(w) - \Psi(v))^2. \quad (26)$$

If Ψ is not strictly increasing, then we obtain Eq. (26) by approaching Ψ by $s \mapsto \Psi(s) + \varepsilon s$. By monotony of Ψ , we have furthermore

$$\int_v^w (\Psi(s) - \Psi(w)) ds \geq \frac{1}{2\text{Lip}(\Psi)} (\Psi(w) - \Psi((v \top k) \perp w))^2$$

for every $k \in \mathbb{R}$. Noticing that

$$F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k) = \Psi(w) - \Psi((v \top k) \perp w)$$

for $v = v_L^{n+1}$, $w = v_K^{n+1}$, we eventually deduce from Eq. (25) the estimate

$$\sum_{n \leq N} \delta t \sum_{(K,L) \in \mathcal{E}_{\text{int}}} \frac{1}{2\mathcal{L}|K|L|} |F_{KL}^n(v_K^{n+1} \top k, v_L^{n+1} \top k) - F_{KL}^n(v_K^{n+1} \top k, v_K^{n+1} \top k)|^2 \leq C.$$

The Cauchy-Schwarz inequality then gives Eq. (19). This concludes the proof of the Lemma. \square

3.3.3. Convergence

Theorem 3.12. *Let (\mathcal{T}_n) be a sequence of meshes of size h_n converging to 0, for which the condition of uniformity (3) holds uniformly in n . Let (u^n, v^n) be the numerical solution defined by the scheme (10)–(12). Let v be an entropy solution to Problem $\tilde{\mathbf{P}}(u_0, f)$. Then, for every $1 \leq p < +\infty$, (u^n) converges strongly in $L^p(Q)$ to the entropy solution to $g(v)$.*

Proof. We give the sketch of the proof of convergence. It is based on two results. First, the following result of (weak) compactness, based on the uniform L^∞ estimate on (v^n) (see Remark 3.7) and proved in [EGH00].

Proposition 3.13. *There exists a function $v^\infty \in L^\infty(Q \times [0, 1])$ and a subsequence still denoted by (v^n) such that*

$$\psi(v^n) \rightarrow \int_0^1 \psi(v^\infty(\cdot, \alpha)) d\alpha \text{ in } L^\infty(Q) \text{ w} - *$$

for every $\psi \in C(\mathbb{R})$.

We use this result of compactness to pass to the limit in the approximate entropy inequality (18) to obtain the following generalized entropy inequality: for all $k \geq 0$, for all non-negative test function $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d)$,

$$\begin{aligned} \int_0^1 \int_Q (g(v^\infty(x, t, \alpha)) - g(k))^+ \varphi_t + \text{sgn}_0^+(v^\infty(x, t, \alpha) - k) (B(v^\infty(x, t, \alpha)) - B(k)) \cdot \nabla \varphi \, dx \, dt \, d\alpha \\ + \int_\Omega (u_0(x) - g(k))^+ \varphi(0, x) \, dx \geq 0. \end{aligned} \quad (27)$$

Similarly, we show the generalized entropy inequality: for all $k \leq 0$, for all non-negative test function $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d)$,

$$\begin{aligned} \int_0^1 \int_Q (g(v^\infty(x, t, \alpha)) - g(k))^- \varphi_t + \text{sgn}_0^-(v^\infty(x, t, \alpha) - k) (B(v^\infty(x, t, \alpha)) - B(k)) \cdot \nabla \varphi \, dx \, dt \, d\alpha \\ + \int_\Omega (u_0(x) - g(k))^- \varphi(0, x) \, dx \geq 0. \end{aligned} \quad (28)$$

Using also the fact that Eq. (18) is satisfied for any $k \in \mathbb{R}$ if $\varphi \in \mathcal{D}([0, +\infty) \times \Omega)$, we have Eq. (27)–(28) for any $k \in \mathbb{R}$, $\varphi \in \mathcal{D}([0, T] \times \Omega)$. Taking k respectively close enough to $-\infty$ and $+\infty$, this shows that v^∞ is a generalized weak solution in the sense that

$$\int_0^1 \int_Q (g(v^\infty) \varphi_t + B(v^\infty) \cdot \nabla \varphi) + \int_\Omega u_0 \varphi(0, \cdot) = 0 \quad (29)$$

for all $\varphi \in \mathcal{D}([0, T] \times \Omega)$. A straightforward adaptation of Proposition 2.3 (just add the sum with respect to α) then shows that we have (compare with Definition 2.2):

- (i) $\text{esslim}_{t \rightarrow 0} \|g(v^\infty(t, \cdot)) - u_0\|_{L^1(\Omega \times (0, 1))} = 0$,
- (ii) $\partial_t \int_0^1 g(v^\infty)(\cdot, \cdot, \alpha) \, d\alpha + \text{div} \int_0^1 B(v^\infty)(\cdot, \cdot, \alpha) \, d\alpha = 0$ in $\mathcal{D}'(Q)$,

(iii) for any $\varphi \in \mathcal{D}((0, T) \times \mathbb{R}^d)$, $\varphi \geq 0$, for any $k \in \mathbb{R}^\pm$,

$$\int_Q \int_0^1 (g(v^\infty) - g(k))^\pm \varphi_t + \text{sign}_0^\pm(v^\infty - k) (B(v^\infty) - B(k)) \cdot \nabla \varphi \geq 0.$$

The second major tool is the result of comparison proved by Carrillo [Car03], which gives the uniqueness of $g(v)$ for v entropy solution to $\tilde{P}(u_0)$. Let \tilde{v}^∞ be an element of $L^\infty(Q \times [0, 1])$ satisfying the same properties (i)-(ii)-(iii) as v^∞ . An elementary adaptation of the proof of Theorem 4.1 and Corollary 4.2 of Carrillo [Car03] (it is indeed elementary, just supplement with the sum with respect to α, β ; we refer to [EGH00] for such an extension of the comparison inequalities for entropy solutions to entropy process solutions) shows that

$$\int_0^1 \int_0^1 \int_Q |g(v^\infty(x, t, \alpha)) - g(\tilde{v}^\infty(x, t, \beta))| \varphi_t dx dt d\alpha d\beta \leq 0$$

for every non-negative test function $\varphi \in \mathcal{D}([0, +\infty))$. In particular, taking $\tilde{v}^\infty = v^\infty$ and $\varphi(t) = T - t$, T arbitrary, we see that $g(v^\infty)(x, t, \alpha) = g(v^\infty)(x, t, \beta)$ for a.e. $(x, t, \alpha, \beta) \in Q \times [0, 1]^2$. This implies that

$$g(v^\infty)(x, t, \alpha) = \int_0^1 g(v^\infty)(x, t, \beta) d\beta,$$

i.e. that $g(v^\infty)$ does not depend on α :

$$g(v^\infty(x, t, \alpha)) = g(w(x, t)), \quad w(x, t) := \int_0^1 v^\infty(x, t, \alpha) d\alpha.$$

Since $B = A \circ g$, also $B(v^\infty)$ does not depend on α . Besides, if $k \in \mathbb{R} \setminus g(\mathcal{S})$, then

$$\text{sign}_0^\pm(v^\infty(x, t, \alpha) - k) = \text{sign}_0^\pm(g(v^\infty)(x, t, \alpha) - g(k))$$

does not depend on α and we deduce from Eq. (27)–(28) that: for any $\varphi \in \mathcal{D}((0, T) \times \mathbb{R}^d)$, $\varphi \geq 0$, for any $k \in \mathbb{R}^\pm \setminus g^{-1}(\mathcal{S})$,

$$\int_Q (g(w) - g(k))^\pm \varphi_t + \text{sign}_0^\pm(w - k) (B(w) - B(k)) \cdot \nabla \varphi \geq 0.$$

By Proposition 2.4, w is therefore an entropy solution to $\tilde{P}(u_0)$. Furthermore, there is strong convergence of (u^n) to $g(w)$. Indeed, we have (taking $\psi(v) = g(v)$ in Proposition 3.13), $u^n \rightarrow g(w)$ in $L^\infty(Q)$ $w - *$, and, thus, $u^n \rightarrow g(w)$ in $L^2(Q)$ weak. Taking $\psi(v) = g(v)^2$ in Proposition 3.13, we obtain the convergence of the norms in $L^2(Q)$ and, therefore, the strong convergence in $L^2(Q)$. The uniform bound in $L^\infty(Q)$ then gives the strong convergence in any $L^p(Q)$, $p \in [1, +\infty)$. \square

4. NUMERICAL VALIDATION AND APPLICATIONS

4.1. Numerical procedure

Let us focus on one-dimensional scalar conservation laws

$$u_t + [A(u)]_x = 0 \tag{30}$$

with for instance $A(u) := \text{sign}^+(u - \alpha) \phi(u)$ or $A(u) := (1 - \text{sign}^+(u - \alpha)) \phi(u)$, the function ϕ being regular. Initial and boundary conditions have to be considered, as described before. Taking the new unknown $v = u + \chi_\alpha$

allows to express u as a function of the reduced unknown v :

$$u := g_\alpha(v) = \begin{cases} v & \text{if } v \leq \alpha \\ 1 & \text{if } \alpha < v \leq 1 + \alpha \\ v - 1 & \text{if } v > 1 + \alpha \end{cases}$$

and the scalar conservation laws reduce to

$$[g_\alpha(v)]_t + [B_\alpha(v)]_x = 0, \quad (31)$$

with corresponding boundary and initial conditions. We emphasize that, although the well-posed character of the problem has been showed by Carrillo [Car03], the behavior of the solution for hyperbolic scalar conservation laws with discontinuous flux functions has never been observed. Here, we will present the numerical results related to classical problems which have been modified in the sense that the flux has become discontinuous.

Algorithms

- *Numerical flux.* Considering the flux B (for convenience, g_α will be denoted g), we denote by B^+ (resp. B^-) a nondecreasing (resp. non-increasing) function such that $B = B^+ + B^-$. In the sequel, the implicit scheme related to Eq. (31) is an upwind three-point scheme

$$g(v_i^{n+1}) = g(v_i^n) - \lambda (\mathcal{B}(v_i^{n+1}, v_{i+1}^{n+1}) - \mathcal{B}(v_{i-1}^{n+1}, v_i^{n+1})),$$

with the following numerical flux:

$$\mathcal{B}(u, v) := B^+(u) + B^-(v).$$

- *Newton algorithm.* Each step of the above scheme is solved using a classical Newton algorithm (in order to solve some equation of the type $\mathcal{F}(X) = 0$) with a stopping test based on the ℓ^∞ -norm (of vector $\mathcal{F}(X)$). Moreover, in the Newton procedure, the Jacobian matrix which is involved in each step is tridiagonal (for a three-point scheme) so that numerous efficient algorithms are available to evaluate the solution of the linear system: in particular, the Gauss-Thomas algorithm has been used to solve the linear systems.

Numerical parameters

Solutions to scalar conservation laws have been computed, corresponding to classical models whose flux has been only modified by the introduction of a threshold (Heaviside graph), as described at the beginning of the section. Thus $A(u)$ may take one of the following forms:

$$\text{sign}^+(u - \alpha) \phi(u), \quad (1 - \text{sign}^+(u - \alpha)) \phi(u), \quad \phi(u) + 1 - \text{sign}^+(u - \alpha),$$

for which the flux ϕ may be $\phi(u) = u$ (advection), $\phi(u) = u^2/2$ (Burgers), $\phi(u) = u(1 - u)$ (traffic flow). Numerical parameters which have been used in the following subsection are presented in TABLE 1.

4.2. Riemann problems and the convex-hull construction

Consider the Riemann problem related to a scalar conservation law with a regular flux ϕ . We recall that a discontinuity of the problem, connecting the left-hand state u_L and the right-hand state u_R with shock speed σ , is called a classical shock if it satisfies the Rankine-Hugoniot relations

$$\frac{\phi(u_R) - \phi(u_L)}{u_R - u_L} = \sigma$$

| | Δx | Δt | N | T | λ | ε |
|------------------|------------|------------|-----|-----|-----------|---------------|
| Advection | 0.00125 | 0.000625 | 800 | 0.5 | 0.5 | 10^{-8} |
| Burgers | 0.00250 | 0.002500 | 160 | 0.4 | 1.0 | 10^{-8} |
| Traffic flow | 0.00250 | 0.005000 | 200 | 1.0 | 2.0 | 10^{-8} |
| Lubrication flow | 0.00125 | 0.002500 | 400 | 1.0 | 2.0 | 10^{-8} |

TABLE 1. Numerical parameters: Δx (spatial step), Δt (time step), N (number of time steps), T (final time), λ (CFL parameter), ε (precision in the Newton procedure).

and the Oleinik condition

$$\frac{\phi(u) - \phi(u_L)}{u - u_L} > \sigma > \frac{\phi(u) - \phi(u_R)}{u - u_R}$$

for all u between u_L and u_R . The Oleinik condition means in particular that the graph of ϕ is lying below (resp. above) the line connecting u_L to u_R when $u_R < u_L$ (resp. $u_R > u_L$). Thus, the entropy-satisfying solution to a non-convex Riemann problem can be determined from the graph of $u \mapsto \phi(u)$ in a simple manner. If $u_L > u_R$, construct the convex hull of the set

$$\{(u, z), u_L \leq u \leq u_R \text{ and } z \geq \phi(u)\}.$$

The convex hull is the smallest convex set containing the original set. If $u_L < u_R$, then we look instead at the convex hull of the set of points below the graph

$$\{(u, z), u_R \leq u \leq u_L \text{ and } z \leq \phi(u)\}.$$

This construction allows to build the solution of the Riemann problem and locate shock waves and rarefaction waves (see Chapter 16 in [Lev02], for instance, for further details). We will see that the same idea works for fluxes with discontinuities of first type. The numerical tests for Riemann problems always assume that

$$\alpha \in [\min(u_L, u_R), \max(u_L, u_R)],$$

in order to describe the influence of the discontinuity (else the equation $\partial_t u + (A(u))_x = 0$ reduces either to $u_t = 0$ or $u_t + (\phi(u))_x = 0$).

Advection equation

The conservation law with the flux $A(u) = u + 1 - \text{sign}^+(u)$ was considered, from the theoretical point of view, by Dias, Figueira and Rodrigues [DFR05] as a limit case of a phase transition problem. We focus on the Riemann problem with $u_L = -0.5$ and $u_R = 1$, the initial discontinuity being located at $x = 0.5$. Numerical results are given by FIG.1. Conservative thresholding effects may be observed: a discontinuity between -0.5 and 0 and another one between 0 and 1 have formed and evolved at different speeds. In fact, the analysis can be totally described by the characteristics method and the convex-hull construction of the flux $u \mapsto u + 1 - \text{sign}^+(u)$ with $u_L = -0.5$ and $u_R = 1$ (see FIG.2). If we look at the lower boundary of this set, we see that it is composed of straight line segments:

- segment (\mathcal{S}_1) : from $(u_L, 1 + u_L)$ to $(0, 0)$,
- segment (\mathcal{S}_2) : from $(0, 0)$ to (u_R, u_R) .

The straight line (\mathcal{S}_1) (resp. (\mathcal{S}_2)) represent a shock jumping from u_L to 0 (resp. 0 to u_R) and the slope of the line segment is equal to the shock speed σ :

$$\sigma = \begin{cases} \frac{1 + u_L}{u_L} & \text{for the shock } (\mathcal{S}_1), \\ 1 & \text{for the shock } (\mathcal{S}_2). \end{cases}$$

Notice that the speed of the shock (\mathcal{S}_1) is negative for $u_L \in (-1, 0)$, zero for $u_L = -1$, positive for $u_L \in (-\infty, -1)$. However, (\mathcal{S}_1) evolves at a slower speed than (\mathcal{S}_2) .

Burgers equation

Considering the flux $A(u) = \text{sign}^+(u - 1) u^2/2$, we focus on the Riemann problem with $u_L = 0$ and $u_R > 1$, the initial discontinuity being located at $x = 0.2$. Two tests are provided:

- Data (A) correspond to $u_R = 1.50$ (see FIG.3).
- Data (B) correspond to $u_R = 3.00$ (see FIG.5).

It may be observed that if $1 < u_R \leq 2$ there is no rarefaction wave (although such a wave exists for the classical Burgers equation): the solution is instead composed of a stationary shock between 0 and 1 and a moving shock between 1 and u_R . If $u_R > 2$, we obtain a stationary shock between 0 and 1, an evolutive shock between 1 and 2 and a rarefaction wave between 2 and u_R . As in the previous example, the analysis can be totally described by the convex-hull construction of the flux $u \mapsto \text{sign}^+(u - 1) u^2/2$ with $u_L = 0$ and $u_R > 1$ (see FIG.4 and 6). The construction provides the following analysis:

- if $1 < u_R \leq 2$, then the solution is determined by two shocks:
 - a shock (\mathcal{S}_1) from u_L to 1, with a shock speed $\sigma = 0$,
 - a shock (\mathcal{S}_2) from 1 to u_R , with a shock speed $\sigma = \frac{u_R}{u_R - 1}$.
- if $u_R > 2$, then the solution is determined by two shocks and a rarefaction wave:
 - a shock (\mathcal{S}_1) from u_L to 1, with a shock speed $\sigma = 0$,
 - a compound wave is formed: a shock (\mathcal{S}_2) from 1 to 2, with a shock speed $\sigma = 2$ is immediately followed by a rarefaction wave (\mathcal{R}) between 2 and u_R (as the lower boundary of the convex-hull construction follows the graph of $u \mapsto A(u)$ which is strictly convex).

Traffic flow equation

We choose the following flux $A(u) = (1 - \text{sign}^+(u - \alpha)) u(1 - u)$ with $\alpha = 0.75$, $u_L = 1$ and $u_R = 0.5$. Numerical results are given by FIG.7. Again, they may be recovered using the convex-hull construction (see FIG.8): in particular, we observe

- a shock between u_R and α which propagates at speed

$$\sigma = -\frac{\alpha(1 - \alpha)}{u_R - \alpha} < 0,$$

- a rarefaction wave between α and u_L .

Notice that this example may model, in a very simple manner, some uniform obstacle which prevents cars densities from reaching high values: even starting from a full-saturated configuration, the density is locally thresholded at the value α in a finite time.

4.3. Application to lubrication theory: a multi-fluid model

Physical motivation

Consider two Newtonian fluids with viscosities $\nu := 1$ (for the so-called reference fluid) and $\nu := \varepsilon$. A thin film flow with two different (non-miscible) fluids in journal bearings of infinite width (i.e. including shearing effects, as usual in lubrication theory) may be described by a modified Buckley-Leverett equation, see [BMV06], which takes the form:

$$u_t + (Qf_\varepsilon(u) + v_0 h(x) u(1 - f_\varepsilon(u)))_x = 0,$$

where u denotes the saturation of the reference fluid, h is the converging-diverging gap between the rigid surfaces which enclose the bi-fluid mixture, v_0 is the shearing velocity of the device and Q is the input flow. Notice that f_ε is typically S -shaped and satisfies $f_\varepsilon(0) = 0$ and $f_\varepsilon(1) = 1$ (as in the classical Buckley-Leverett equation)

but its profile strongly depends on ε : for convenience and without loss of generality, we may assume that f_ε is defined by

$$f_\varepsilon(u) := \begin{cases} 0, & \text{if } u \in [0, 1 - \varepsilon) \\ \frac{u - (1 - \varepsilon)}{\varepsilon}, & \text{if } u \in [1 - \varepsilon, 1) \end{cases}$$

The study and simulation of this model for small values of ε are of high interest for the understanding of cavitation phenomena, which cannot be avoided in lubricated regimes: cavitation is defined as the rupture of the continuous liquid lubricated film due to the formation of gas bubbles. In that case, we may observe “saturated” areas (the set $[u = 1]$) and “cavitated” or “unsaturated” areas (the set $[0 \leq u < 1]$, corresponding to a liquid / gas mixture). In this setting, the viscosity ratio is about $\varepsilon \sim 10^{-3}$. Of course, the numerical computation of the corresponding entropy solution by using explicit schemes becomes very difficult since the CFL condition is $\lambda \varepsilon \leq 1$. Implicit schemes are needed and, at the limit (for even smaller values of ε), we may proceed as follows: as the entropy solution lies in $[0, 1]$ (for initial and boundary data in $[0, 1]$), f_ε can be extended by 0 on $(-\infty, 0)$ and 1 on $(1, +\infty)$ and we have the following convergence:

$$f_\varepsilon \rightarrow \text{sign}^+(\cdot - 1) \quad \text{in } L^p(\mathbb{R}), \quad 1 \leq p < +\infty.$$

The problem to be considered is

$$u_t + (Q \text{sign}^+(u - 1) + h(x) u (1 - \text{sign}^+(u - 1)))_x = 0$$

with appropriate data, as will be stated further.

Numerical results

Unlike all the previous cases, the flux

$$A(\cdot, u) = Q \text{sign}^+(u - 1) + h u (1 - \text{sign}^+(u - 1))$$

is not autonomous, as $h' \neq 0$ (we recall that h has typically a converging-diverging profile) which leads to balance effects: in particular, constant functions are not stationary solutions, as will be observed in the numerical simulation. However, all the previous theoretical analysis may be adapted to non-autonomous fluxes. Let us fix the data:

- the normalized gap h is defined as: $h(x) = (2 \cdot x - 1)^2 + 0.5$,
- the shear velocity is normalized: $v_0 = 1$,
- the initial and boundary conditions are taken constant and equal to $\bar{u} = 0.5$ (i.e. the initial configuration is unsaturated and the device is partially supplied in liquid lubricant),
- the input flow is taken to $Q = \bar{u} v_0 h(0)$ (which means that the global mass-flow input is only contributed by the liquid phase, see [BMV06]).

Numerical results are given by FIG.9–10. It is observed that, due to the non-autonomous contribution, constant states are not preserved. In particular, starting from a full cavitated setting, we observe that a saturated region appears, located at the minimal gap region, as the entropy solution converges to

$$\min \left(1, \frac{Q}{h} \right)$$

However, similarly to the results obtained in [MVo07] for regular fluxes, the large-time behavior of the entropy solution highly depends on the initial condition.

4.4. Concluding remarks

In this paper, we have investigated the numerical analysis and computation of the entropy solution of scalar conservation laws with discontinuous fluxes. For this, we have used implicit finite volume methods because explicit schemes cannot be used in this framework:

- Explicit schemes applied to $P(u_0)$ are unadapted for obvious reasons related to the lack of regularity of the flux;
- Explicit schemes applied to $\tilde{P}(u_0)$ are also unadapted, despite the regularity of the flux, because the inverse of g is non-univoque. Still, one could think that computations led with some chosen inverse of g would allow us to recover the correct numerical solution: more precisely, assume that $\mathcal{S} = \{\alpha\}$, with $\alpha > 0$. Then g^{-1} is multivalued (see further) and one could use, instead, any function g_τ^{-1} with $\tau \in [0, 1]$ in a numerical explicit procedure:

$$g^{-1}(v) := \begin{cases} v & \text{if } v < \alpha, \\ [\alpha, \alpha + 1] & \text{if } v = \alpha, \\ v + 1 & \text{if } v > \alpha, \end{cases} \quad g_\tau^{-1}(v) := \begin{cases} v & \text{if } v < \alpha, \\ \alpha + \tau & \text{if } v = \alpha, \\ v + 1 & \text{if } v > \alpha, \end{cases}$$

In fact, for the advection equation, Burgers equation or traffic flow equation, the corresponding numerical solution (for different fixed values of τ) is not correct, in most of cases, and instabilities of the numerical solution with respect to the spatial mesh have been observed.

The wrong behavior of explicit methods evidences the need to implement implicit methods which are proved to be rigorous from a mathematical point of view.

APPENDIX A. PROOF OF PROPOSITION 2.3 AND PROPOSITION 2.4

A.1. Proof of Proposition 2.3

Let $v \in L^\infty(Q)$ be an entropy solution of $\tilde{P}(u_0, f)$. Let $\varphi \in \mathcal{D}([0, T] \times \mathbb{R}^d)$ and, for $h > 0$, let ω_h be the cut-off function defined by $\omega_h(t) = \min(1, t/h)$. Using $\varphi\omega_h$ as a test function, we obtain (by use of the Lebesgue dominated convergence Theorem)

$$\int_Q (g(v) - g(k))^\pm \varphi_t + \text{sign}_0^\pm(v - k) (B(v) - B(k)) \cdot \nabla \varphi + \frac{1}{h} \int_0^h \int_\Omega (g(v) - g(k))^\pm \varphi + o(1) \geq 0$$

when $h \rightarrow 0$. Since $\text{esslim}_{t \rightarrow 0} \|g(v(t, \cdot)) - u_0\|_{L^1(\Omega)} = 0$, we obtain

$$\int_Q (g(v) - g(k))^\pm \varphi_t + \text{sign}_0^\pm(v - k) (B(v) - B(k)) \cdot \nabla \varphi + \int_\Omega (u_0 - g(k))^\pm \varphi(0, \cdot) \geq 0$$

at the limit $h \rightarrow 0$, which is Eq. (2). Similarly, from the weak equation $g(v)_t + \text{div}(B(v)) = 0$ in $\mathcal{D}'(Q)$, we deduce the weak formulation including the initial datum (1).

Conversely, assume that the weak equation and the weak entropy inequalities (1) and (2) are satisfied. This implies in particular that the weak entropy inequalities (2) are satisfied whatever the sign of k , as soon as the test function is compactly supported in $[0, T] \times \Omega$ (use the identity $(g(v) - g(k))^+ = (g(v) - g(k))^- + (g(v) - g(k))$). Let a non-negative $\theta \in \mathcal{D}(\Omega)$ be fixed and let $k \in \mathbb{R}$. Using Eq. (2) with the multiplicative test-function $\varphi(t, x) = \alpha(t)\theta(x)$ ($\alpha \in \mathcal{D}(0, T)$ non-negative), we obtain $K' \leq L$ in $\mathcal{D}'(0, T)$ where

$$K(t) := \int_\Omega (g(v) - g(k))^\pm \theta, \quad L(t) := \int_\Omega \text{sign}_0^\pm(v - k) (B(v) - B(k)) \cdot \nabla \theta.$$

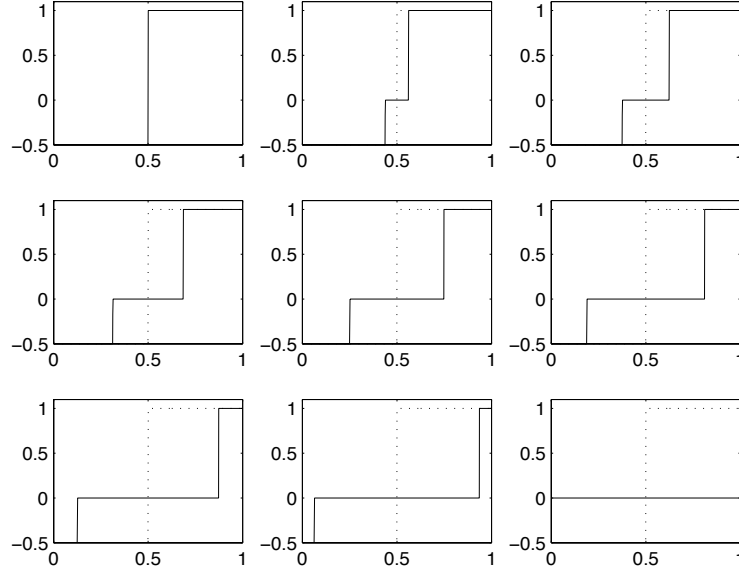


FIGURE 1. Entropy solution u of the advection equation at different time steps: $t = 100 * j * \Delta t$ ($j = 0, \dots, 8$). The solution (resp. solution at $t = 0$) is represented with plain (resp. dotted) lines.

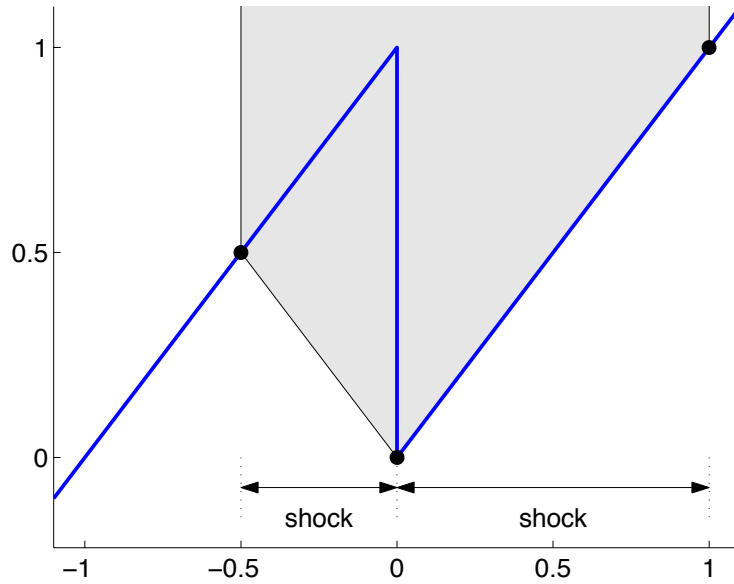


FIGURE 2. Convex-hull construction for the flux $u \mapsto u + 1 - \text{sign}^+(u)$ with the Riemann data $u_L = -0.5$ and $u_R = 1$

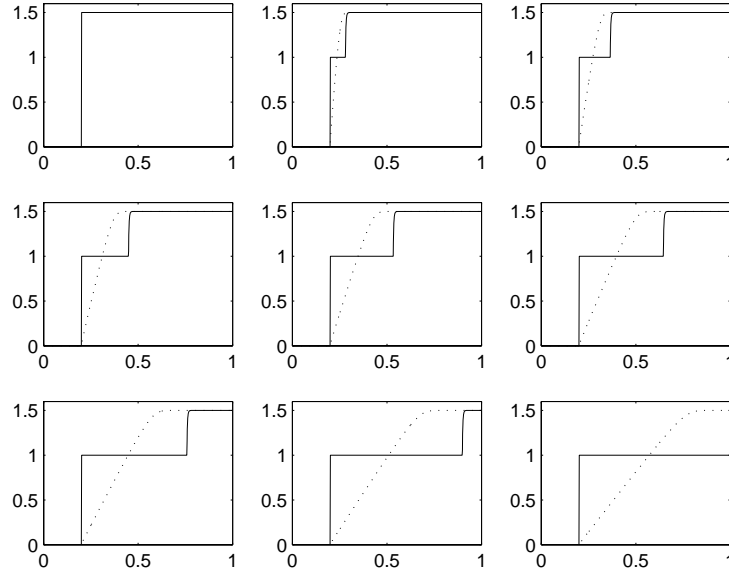


FIGURE 3. Entropy solution u of the Burgers equation with data (A) at different time steps: $t = 20 * j * \Delta t$ ($j = 0, \dots, 8$). The solution of the discontinuous (resp. continuous) problem is represented with plain (resp. dotted) lines.

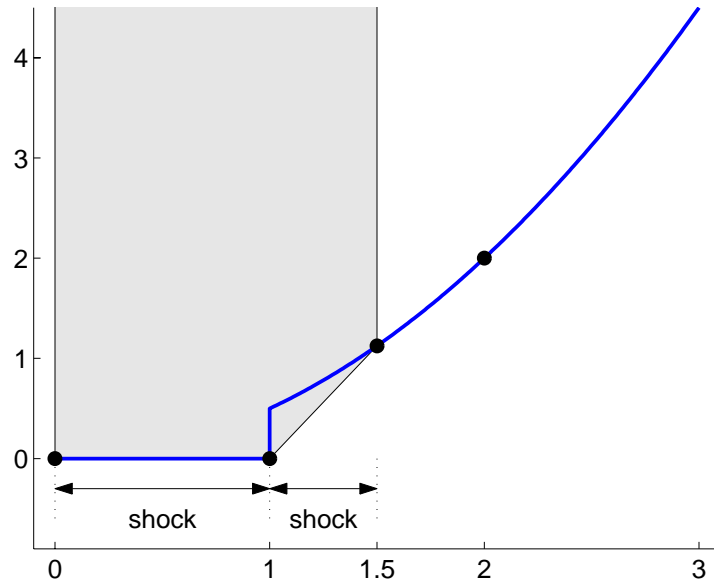


FIGURE 4. Convex-hull construction for the flux $u \mapsto \text{sign}^+(u-1)u^2/2$ with the Riemann data $u_L = 0$ and $u_R = 2$

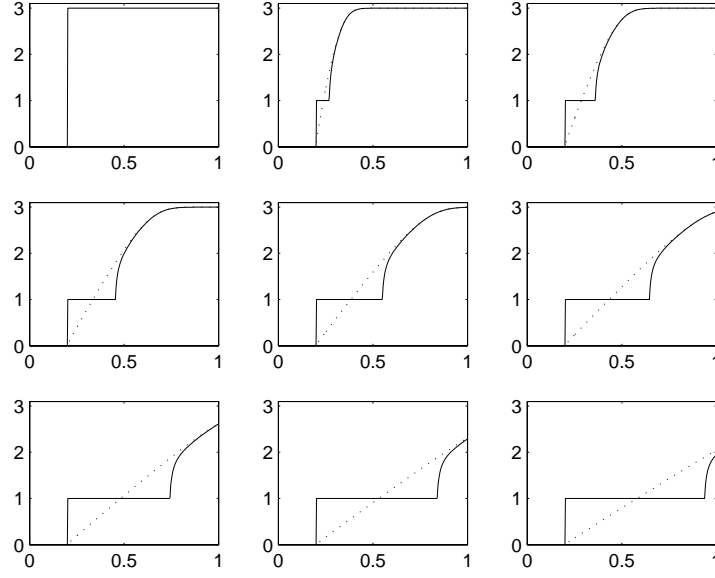


FIGURE 5. Entropy solution u of the Burgers equation with data (B) at different time steps: $t = 20 * j * \Delta t$ ($j = 0, \dots, 8$). The solution of the discontinuous (resp. continuous) problem is represented with plain (resp. dotted) lines.

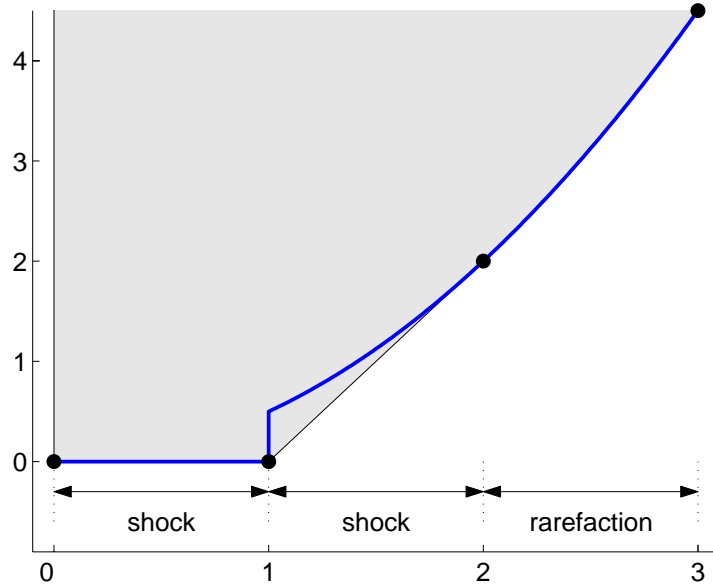


FIGURE 6. Convex-hull construction for the flux $u \mapsto \text{sign}^+(u-1)u^2/2$ with the Riemann data $u_L = 0$ and $u_R = 3$

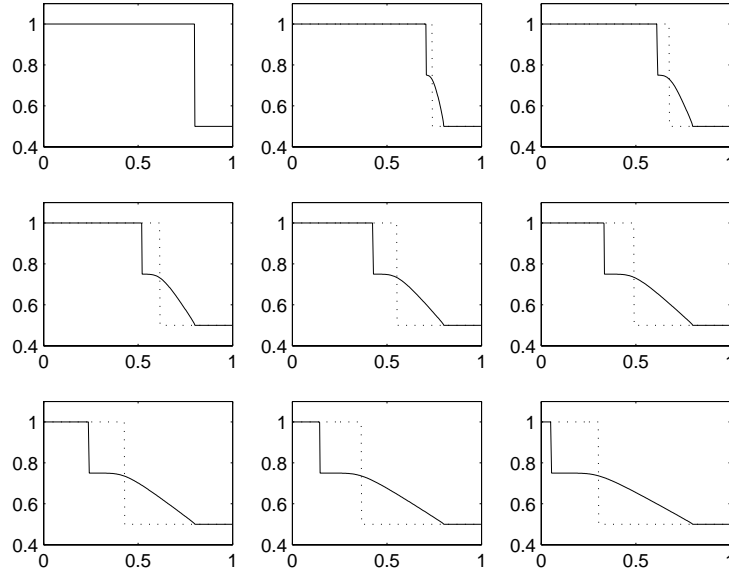


FIGURE 7. Entropy solution u of the traffic flow equation at different time steps: $t = 25*j*\Delta t$ ($j = 0, \dots, 8$). The solution of the discontinuous (resp. continuous) problem is represented with plain (resp. dotted) lines.

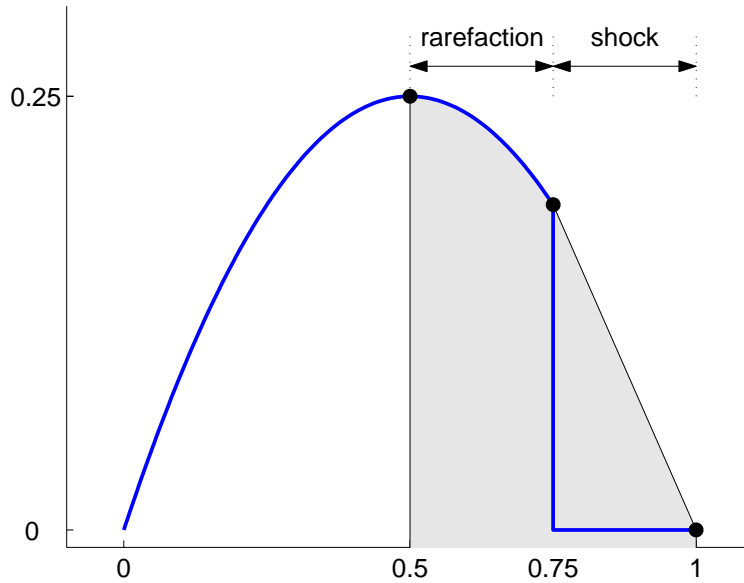


FIGURE 8. Convex-hull construction for the flux $u \mapsto (1 - \text{sign}^+(u - 0.75))u(1 - u)$ with the Riemann data $u_L = 0.5$ and $u_R = 1$

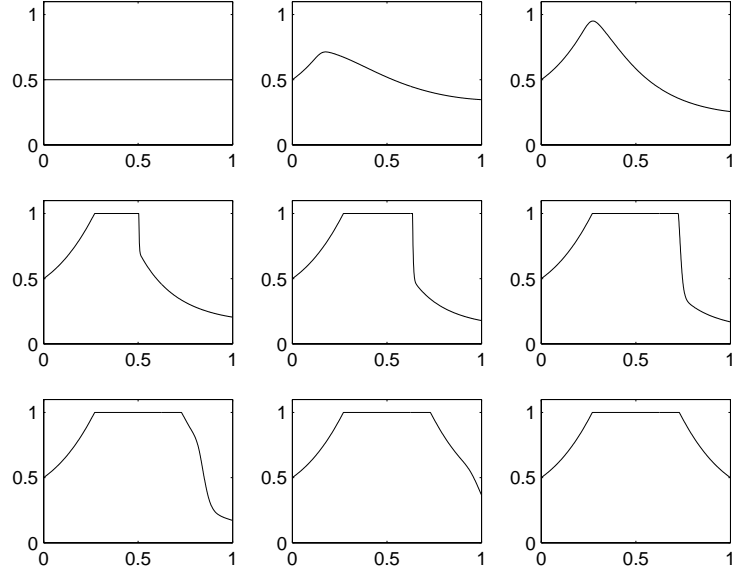


FIGURE 9. Entropy solution u of the non-autonomous Buckley-Leverett equation at different time steps: $t = 50 * j * \Delta t$ ($j = 0, \dots, 8$).

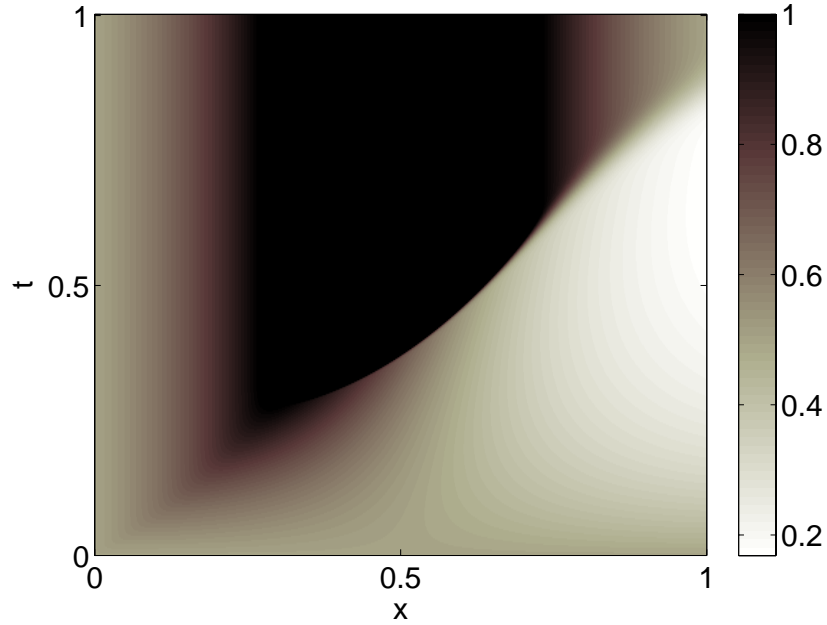


FIGURE 10. Entropy solution u of the non-autonomous Buckley-Leverett equation.

The distribution L is of order 0, therefore K' also, i.e. $K \in BV(0, T)$. This proves that $\text{esslim}_{t \rightarrow 0} K(t)$ exists (in $[0, +\infty]$ since $K \geq 0$). By Eq. (2) tested with $\varphi(t, x) = (1 - \omega_h(t))\theta(x)$, we obtain

$$\text{esslim}_{t \rightarrow 0} \int_{\Omega} (g(v(t, \cdot)) - g(k))^{\pm} \theta \leq \int_{\Omega} (u_0 - g(k))^{\pm} \theta.$$

The limit above is taken for t converging to 0 in a set $S(k, \theta)$ of full measure in $[0, T]$. Since the countable intersection of such sets is also of full measure in $[0, T]$, we have

$$\text{esslim}_{t \rightarrow 0} \int_{\Omega} |g(v(t, \cdot)) - w| \leq \int_{\Omega} |u_0 - w|.$$

for every $w \in L^{\infty}(Q)$ of the form

$$w = \sum_{i \in \mathbb{N}} g(k_i) \theta_i$$

where $k_i \in \mathbb{R}$, $\theta_i \in \mathcal{D}(\Omega)$ having disjoint support from θ_j if $i \neq j$. Since u_0 can be approximated by a sequence of functions w^n as above in $L^1(\Omega)$, the result follows:

$$\text{esslim}_{t \rightarrow 0} \int_{\Omega} |g(v(t, \cdot)) - u_0| = 0.$$

A.2. Proof of Proposition 2.4

If v satisfies the entropy inequality (iii) of Definition 2.2 for $k \in \mathbb{R}_{\pm} \setminus g^{-1}(\mathcal{S})$ and $k_* \in g^{-1}(\mathcal{S})$ with, say, $k_* \geq 0$, then $g^{-1}(\{g(k)\}) = [k_-, k_+]$ where $k_+ \geq 0$. Let (k_n) be a sequence of $\mathbb{R}_{+} \setminus g^{-1}(\mathcal{S})$ monotonically decreasing which converges to k_+ . Passing to the limit in the entropy inequality (iii) written with $k = k_n$, we obtain

$$\int_Q (g(v) - g(k_+))^+ \varphi_t + \text{sign}_0^+(v - k_+) (B(v) - B(k_+)) \cdot \nabla \varphi \geq 0.$$

Since $g(k_+) = g(k_*)$ and since $B = A \circ g$, we have $(g(v) - g(k_+))^+ = (g(v) - g(k_*))^+$ and

$$\text{sign}_0^+(v - k_+) (B(v) - B(k_+)) = \text{sign}_0^+(v - k_+) (B(v) - B(k_*)) = \text{sign}_0^+(v - k_*) (B(v) - B(k_*)),$$

therefore v satisfies the entropy inequality

$$\int_Q (g(v) - g(k_*))^+ \varphi_t + \text{sign}_0^+(v - k_*) (B(v) - B(k_*)) \cdot \nabla \varphi \geq 0.$$

REFERENCES

- [ABK00] B. P. Andreianov, P. B nilan, and S. N. Kruzhkov, *L¹-theory of scalar conservation law with continuous flux function*, J. Funct. Anal. **171** (2000), no. 1, 15–33.
- [BLN79] C. Bardos, A.-Y. Leroux and J.-C. Nedelec, *First order quasilinear equations with boundary conditions*, Comm. in P.D.E. **4** (1979), 1017–1034.
- [BMV06] G. Bayada, S. Martin and C. V zquez, *About a generalized Buckley-Leverett equation and lubrication multifluid flow*, Euro. J. Applied. Math. **17** (2006), no. 5, 491–524.
- [BK96] Ph. Benilan and S.N. Kruzhkov, *Conservation laws with continuous flux functions*, NoDEA Nonlinear Differential Equations Appl. **3** (1996), no. 4, 395–419.
- [Car99] J. Carrillo, *Entropy solutions for nonlinear degenerate problems*, Arch. Rational Mech. Anal. **147** (1999), no 4, 269–361.
- [Car03] J. Carrillo, *Conservation laws with discontinuous flux functions and boundary condition*, J. Evol. Eq. **3** (2003), no. 2, 687–705.
- [CCL95] B. Cockburn, F. Coquel and P.G. LeFloch, *Convergence of the finite volume method for multidimensional conservation laws*, SIAM J. Numer. Math. **32** (1995), no. 3, 775–796.

- [Des04] B. Després, *An explicit a priori estimate for a finite volume approximation of linear advection on non-Cartesian grids*, SIAM J. Numer. Anal. **42** (2004), no. 2, 484–504 (electronic).
- [DFR05] J.-P. Dias, M. Figueira, and J.-F. Rodrigues, *Solutions to a scalar discontinuous conservation law in a limit case of phase transitions*, J. Math. Fluid Mech. **7** (2005), no. 2, 153–163.
- [DiP85] R.J. DiPerna, *Measure-valued solutions to conservation laws*, Arch. Rational Mech. Anal. **88** (1985), no. 3, 223–270.
- [EGH00] R. Eymard, T. Gallouët and R. Herbin, *Finite volume methods*, Handbook of numerical analysis, Vol. VII, North-Holland, Amsterdam, 2000, pp. 713–1020.
- [KRW96] D. Kröner, M. Rokyta, and M. Wierse, *A Lax-Wendroff type theorem for upwind finite volume schemes in 2D*, East-West Journal of Numerical Mathematics (1996).
- [Kru70] S.N. Kružkov, *First order quasilinear equations with several independent variables*, Mat. Sb. (N.S.) **81** (1970), no. 123, 228–255.
- [Lev02] R.J. LeVeque, *Finite volume methods for hyperbolic problems*, Cambridge Texts in Applied Mathematics, Cambridge University Press, Cambridge, 2002. xx+558 pp.
- [MV07] S. Martin and J. Vovelle, *Large-time behaviour of the entropy solution of a scalar conservation law with boundary conditions*, Quart. Appl. Math., to appear (2007).
- [Ol59] O. Oleinik, *Uniqueness and stability of the generalized solution of the Cauchy problem for a quasi-linear equation*, Uspehi Mat. Nauk **14** (1959), no. 2 (86), 165–170.
- [Ot96] F. Otto, *Initial boundary value problem for a scalar conservation law*, C. R. Acad. Sci. Paris **322** Série I (1996) no. 8, 729–734.
- [Sze91] A. Szepessy, *Convergence of a streamline diffusion finite element method for scalar conservation laws with boundary conditions*, RAIRO Modél. Math. Anal. Numér. **25** (1991), no. 6, 749–782.
- [Vil94] J.-P. Vila, *Convergence and error estimates in finite volume schemes for general multidimensional scalar conservation laws. I. Explicit monotone schemes*, RAIRO Modél. Math. Anal. Numér. **28** (1994), no. 3, 267–295.
- [Vol67] A. I. Vol’pert, *Spaces bv and quasilinear equations*, Mat. Sb. (N.S.) **73** (**115**) (1967), 255–302.
- [Vov02] J. Vovelle, *Convergence of finite volume monotone schemes for scalar conservation laws on bounded domains*, Num. Math. **90** (2002), no. 3, 563–596.