



Analysis and approximation of a scalar conservation law with a flux function with discontinuous coefficients

Nicolas Seguin, Julien Vovelle

► To cite this version:

Nicolas Seguin, Julien Vovelle. Analysis and approximation of a scalar conservation law with a flux function with discontinuous coefficients. *Mathematical Models and Methods in Applied Sciences*, 2003, 13 (2), pp.221-257. 10.1142/S0218202503002477 . hal-01376535

HAL Id: hal-01376535

<https://hal.science/hal-01376535>

Submitted on 20 Nov 2018

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.

ANALYSIS AND APPROXIMATION OF A SCALAR CONSERVATION LAW WITH A FLUX FUNCTION WITH DISCONTINUOUS COEFFICIENTS

NICOLAS SEGUIN* and JULIEN VOVELLE†

*LATP – UMR CNRS 6632, Université de Provence,
 39 rue Joliot Curie, F-13453 Marseille Cedex 13, France*

**seguin@cmi.univ-mrs.fr*

†julien.vovelle@cmi.univ-mrs.fr

Received 25 September 2001

Revised 12 July 2002

Communicated by P. Degond

We study here a model of conservative nonlinear conservation law with a flux function with discontinuous coefficients, namely the equation $u_t + (k(x)u(1-u))_x = 0$. It is a particular entropy condition on the line of discontinuity of the coefficient k which ensures the uniqueness of the entropy solution. This condition is discussed and justified. On the other hand, we perform a numerical analysis of the problem. Two finite volume schemes, the Godunov scheme and the VFRoe-ncv scheme, are proposed to simulate the conservation law. They are compared with two finite volume methods classically used in an industrial context. Several tests confirm the good behavior of both new schemes, especially through the discontinuity of permeability k (whereas a loss of accuracy may be detected when industrial methods are performed). Moreover, a modified MUSCL method which accounts for stationary states is introduced.

Keywords: Conservation law; discontinuous coefficient; finite volume; resonance.

AMS Subject Classification: 35L65, 35R05, 76M12

1. Introduction

Consider the Cauchy problem associated with the following conservation law with discontinuous coefficient:

*Also at DPT MFTT, EDF – Recherche et Développement, 6 quai Watier, F-78401 Chatou cedex, France.

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(k(x)u(1-u)) = 0 & x \in \mathbb{R}, t \in \mathbb{R}_+, \\ u(0, x) = u_0(x), \\ k(x) = \begin{cases} k_L & \text{if } x < 0 \\ k_R & \text{if } x > 0, \end{cases} \end{cases} \quad (1.1)$$

with $k_L, k_R > 0$ and $k_L \neq k_R$.

This problem can be seen as a model for the governing equation of the saturation of a fluid in a gravity field flowing in a one-dimensional porous media with discontinuous permeability k . It is also one of the simplest examples of resonant system^{18,20}: indeed, it can be written as the system

$$u_t + (kg(u))_x = 0, \quad k_t = 0 \quad (1.2)$$

whose eigenvalues 0 and $kg'(u)$ can intersect each other if the function g' vanishes for a certain value of u . Resonant systems, including those of (1.2) have been studied by several authors. We refer to Refs. 21, 22, 27, 30 and references therein.

One of the main difficulty in the analysis of Problem (1.1) is the correct definition of a solution. Let us briefly justify this assertion: consider the Cauchy problem

$$u_t + (f(x, u))_x = 0, \quad t > 0, x \in \mathbb{R} \quad u(0, x) = u_0(x) \quad x \in \mathbb{R}. \quad (1.3)$$

Suppose that the function f is not continuous with respect to x . Actually, suppose that

$$f(x, u) = \begin{cases} -u & \text{if } x < 0 \\ u & \text{if } x > 0. \end{cases}$$

An easy computation (thanks to the method of the characteristics) ensures that the solution u of Problem (1.3) is known in the domain $\{|x| > t\}$ but is not determined at all in the domain $\{|x| < t\}$. Otherwise, the problem (1.3) cannot be well-posed for every function f ! In the case where the flux writes $f(x, u) = k(x)g(u)$ (k being discontinuous) and when the function g is convex (or concave) some definitions of solutions have been proposed. First, in Ref. 22, Klingenberg and Risebro define a weak solution which is shown to be unique and stable²¹ under a wave entropy condition. Secondly, in Ref. 31, Towers define a notion of entropy solution and prove the uniqueness of piecewise smooth entropy solutions. The definition of Towers (see Definition 1) is a global definition (in the sense that the entropy conditions are enclosed in the weak formulation and not required as local conditions) and, therefore, adapted to the study of the convergence of approximations to Problem (1.1). In Ref. 31 the “entropy condition” on the line of discontinuity of the function k is given (see Eq. (2.13)) which actually plays a central role in the study of the well-posedness of Problem (1.1). A justification for this condition is given here (see Sec. 3).

Assuming that the function $g : u \mapsto u(1-u)$ is concave and, therefore, has only one global maximum, is not insignificant. The study of Problem (1.1) or (1.2) in the case where the function g has more than one local extremum remains

difficult. Some steps forward have been made by Klingenberg, Risebro and Towers in particular.^{23,32} Nevertheless, one of the stake in the study of Problem (1.1) in the case where the function g has more than one global optimum, probably remains the design of a global definition of entropy solution.

The understanding of the Riemann problem is one of the bases of all the theory of one-dimensional systems of conservation laws. It is also a useful tool to make the problem (1.1) more intelligible (the question of uniqueness of a solution in particular) while it is necessary to define properly the Godunov scheme. The Riemann problem for resonant systems has been studied, near a hyperbolic singularity, by Isaacson, Marchesin, Plohr and Temple.^{17,19} The Riemann problem for scalar conservation laws as (1.1) has been completely solved by Gimse and Risebro.^{4,13} In Appendix A we present a new way to visualize the solution of the Riemann problem (see Figs. A.1 and A.2).

Different approximations of Problem (1.1) have been studied. In Ref. 22, Klingenberg and Risebro prove the convergence of the front-tracking method applied to (1.1) to the weak solution defined in Ref. 22. The first result of convergence of a numerical scheme has been established by Temple³⁰ who proves the convergence of the Glimm scheme associated to Problem (1.2). In Refs. 27 and 28, Lin, Temple and Wang prove the convergence to a weak solution of the Godunov method applied to Problem (1.2). Some finite volume schemes associated to Problem (1.1) have also been studied by Towers. In Ref. 31, the author proves the convergence of the Godunov scheme and the Engquist–Osher scheme, by considering a discretization of k staggered with respect to that of u . Here, we study the qualitative behavior of the Godunov scheme (with a discretization of k superimposed on that of u), of an approximate Godunov scheme and show that they behave much better than two schemes classically implemented in industrial codes. Notice that this study is close to some works that have been performed to investigate the behavior of the finite volume method applied to (systems of) conservation laws with source terms by Greenberg and Le Roux,¹¹ or Gallouët, Hérard, Seguin.⁷

A somewhat related subject is the study of a transport equation with a discontinuous coefficient.² Also notice that, forgetting the conservative form of Eq. (1.1), this latter can be rewritten

$$\frac{\partial u}{\partial t} + k(x) \frac{\partial}{\partial x} (u(1-u)) = -(k_R - k_L) \delta_0(x) u(1-u).$$

This underlines its relation with the study of the behavior, as $\varepsilon \rightarrow 0$, of the solution u_ε of the problem

$$\frac{\partial u_\varepsilon}{\partial t} + \frac{\partial}{\partial x} A(u_\varepsilon) + z_\varepsilon(x) B(u_\varepsilon) = 0,$$

where $z_\varepsilon \rightarrow \delta_0$, performed by Vasseur.³³

Besides, degenerate parabolic equations of the kind $u_t + (k(x)g(u))_x - A(u)_{xx} = 0$, where the coefficient k is discontinuous have been studied by Karlsen, Risebro and Towers (see Ref. 24 and references therein).

Eventually, Godlewski and Raviart recently provide a study of the consistence and the numerical approximation of Problem (1.1) when a condition of continuity is

set on the variable u .¹⁵ Notice that it is a different viewpoint from that of Karlsen, Klingenberg, Risebro, Towers and us: we impose a condition of continuity of the flux, $kg(u)$, at the interface $\{x = 0\}$.

This paper is organized as follows. In Sec. 2, we give an overview of the notion of entropy solution to Problem (1.1) as defined by Towers and prove that uniqueness holds for general L^∞ solutions. In Sec. 3, we discuss the entropy condition (2.13). Then, we present several finite volume schemes (Sec. 4), which may be seen as the adaptation to space-dependent flux of the well-balanced scheme¹¹ (initially designed for conservation laws with source terms). The first one is called the Godunov scheme, since it is based on the exact solution of the Riemann problem associated with the conservation law. This Riemann problem may be linearized to simplify its resolution, which leads to the second scheme studied here, namely the VFRoe-ncv scheme. Moreover, two other schemes are introduced, which are derived from methods classically used in the industrial context. Furthermore, a higher order method based on the MUSCL formalism is described. Indeed, stationary solutions are no longer uniform states, because the permeability k is not constant. Hence, in the implementation of the slope limiters, the characterization of stationary states has to be taken into account. The higher order method may then be implemented to simulate the convergence in time towards stationary solutions.

In Sec. 5, several numerical tests are given. First, some solutions of Riemann problems are computed to compare the four “first-order” schemes. Afterwards, quantitative results are shown. Measurements of rates of convergence when the mesh is refined are presented, with or without MUSCL reconstruction. Convergence in time towards stationary states is also performed.

Finally, a new visualization of the Riemann problem, additional tests to illustrate the resonance phenomenon and some BV estimates are available in the Appendices.

2. Entropy Solution

In the case where the function k is regular, the accurate notion of solution to the problem (1.1), namely the one which ensures existence and uniqueness, is the notion of entropy solution. Here, also, a notion of entropy solution can be defined. According to Towers,³¹ we have

Definition 1. Let $u_0 \in L^\infty(\mathbb{R})$, with $0 \leq u_0 \leq 1$ a.e. on \mathbb{R} . A function u of $L^\infty(\mathbb{R}_+ \times \mathbb{R})$ is said to be an entropy solution of the problem (1.1) if it satisfies the following entropy inequalities: for all $\kappa \in [0, 1]$, for all non-negative function $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}_+ \times \mathbb{R})$,

$$\begin{aligned} &\int_0^\infty \int_{\mathbb{R}} |u(t, x) - \kappa| \varphi_t(t, x) + k(x) \Phi(u(t, x), \kappa) \varphi_x(t, x) \, dx dt \\ &\quad + \int_{\mathbb{R}} |u_0(x) - \kappa| \varphi(0, x) \, dx + |k_L - k_R| \int_0^\infty g(\kappa) \varphi(t, 0) \, dt \geq 0, \end{aligned} \tag{2.4}$$

where Φ denotes the entropy flux associated with the Kruzkov entropy,

$$\Phi(u, \kappa) = \operatorname{sgn}(u - \kappa)(g(u) - g(\kappa)),$$

with

$$g(u) = u(1 - u)$$

and

$$\operatorname{sgn}(a) = \begin{cases} +1 & \text{if } a > 0, \\ 0 & \text{if } a = 0, \\ -1 & \text{if } a < 0. \end{cases}$$

Remark 1. This definition can be adapted in the case where the function k has finitely many jumps. It would be interesting to study the case where, more generally, the function k has a bounded total variation. However, we restricted our study to the case where the function k has one jump, keeping in view that this framework is rich enough to point out the different phenomena that occur in the numerical analysis of such problems.

Remark 2. Notice that this definition is pertinent only if the flux function g is concave (or convex). To our knowledge, no definition of solution to Problem (1.1) of this type has been given in the case where the flux g has two or more local maxima; however, the resolution of the Riemann problem, the convergence of the front-tracking method and of the smoothing method to the same solution has been proved in Ref. 23 while the convergence of an Engquist–Osher scheme to a weak solution has been proved by Towers in Ref. 32.

Actually, the flux function $u \mapsto u(1 - u)$ being considered, existence and uniqueness hold for an entropy solution as defined in Definition 1. More precisely, the following theorem can be proved.

Theorem 1. *Suppose that $u_0: \mathbb{R} \rightarrow \mathbb{R}$ is a measurable function such that $0 \leq u_0 \leq 1$ a.e. on \mathbb{R} . Then there exists a unique entropy solution u of the problem (1.1) in $L^\infty((0, T) \times \mathbb{R})$. The solution u satisfies $0 \leq u \leq 1$ a.e. on $(0, T) \times \mathbb{R}$. Besides, if the function $v \in L^\infty((0, T) \times \mathbb{R})$ is another entropy solution of Problem (1.1) with initial condition $v_0 \in L^\infty(\mathbb{R}, [0, 1])$ then, for every $R, T > 0$, the following result of comparison holds:*

$$\int_0^T \int_{(-R, R)} |u(t, x) - v(t, x)| \, dx dt \leq T \int_{(-R-KT, R+KT)} |u_0(x) - v_0(x)| \, dx, \quad (2.5)$$

where $K = \max\{k_L, k_R\}$.

A natural way to obtain the existence of an entropy solution to Problem (1.1) is to check the consistency of Definition 1 with the usual definition of an entropy solution given in the case where the function k is regular^{25,35}: let $(k_\varepsilon)_\varepsilon$ be a sequence of approximation of the function k such that: $\forall \varepsilon > 0$, the function k_ε is a regular

function, it is monotone non-decreasing or non-increasing, according to the sign of $k_R - k_L$ and it verifies

$$\begin{cases} k_\varepsilon(x) = k_L & \text{if } x \leq -\varepsilon, \\ k_\varepsilon(x) = k_R & \text{if } \varepsilon \leq x. \end{cases}$$

For any initial condition $u_0 \in L^\infty(\mathbb{R}; [0, 1])$ there exists a unique entropy solution u_ε of the problem (2.6):

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(k_\varepsilon(x)g(u)) = 0 & x \in \mathbb{R}, t \in \mathbb{R}_+, \\ u(0, x) = u_0(x), \end{cases} \tag{2.6}$$

which satisfies (2.7) and $0 \leq u_\varepsilon \leq 1$ a.e. Moreover, the sequence (u_ε) converges in $L^1_{\text{loc}}([0, T] \times \mathbb{R})$ to a function $u \in L^\infty([0, T] \times \mathbb{R}, [0, 1])$. To begin with, this result can be proved up to a subsequence, and when the initial condition additionally satisfies $u_0 \in BV(\mathbb{R})$. The compactness of the sequence (u_ε) is deduced from the following list of arguments. First, for every $\kappa \in [0, 1]$, the distribution

$$|u_\varepsilon - \kappa|_t + (k_\varepsilon \Phi(u_\varepsilon, \kappa))_x + k'_\varepsilon(x)g(\kappa)\text{sgn}(u_\varepsilon - \kappa)$$

is non-positive and, therefore, is a bounded measure. The distributions $|u_\varepsilon - \kappa|_t$ and $k'_\varepsilon(x)g(\kappa)\text{sgn}(u_\varepsilon - \kappa)$ are also bounded measures because, respectively, the distribution $\partial_t u_\varepsilon$ is a bounded measure (since $u_0 \in BV(\mathbb{R})$) and $TV(k_\varepsilon) \leq |k_R - k_L|$. Eventually, and consequently, the distribution $(k_\varepsilon \Phi(u_\varepsilon, \kappa))_x$ is also a bounded measure. From these facts, one can deduce that the function $\Phi(u_\varepsilon, \kappa)$ is BV with a BV -norm uniformly bounded with respect to ε (a rigorous proof of this estimate is given in Appendix C). Notice that this estimate on $\Phi(u_\varepsilon, \kappa)$ remains true even when the function g has more than one local maximum. Now, Helly's Theorem¹⁰ ensures that a subsequence of $(\Phi(u_\varepsilon, \kappa))$ is converging in $L^1_{\text{loc}}([0, T] \times \mathbb{R})$. The convergence of a subsequence of (u_ε) is deduced from the dominated convergence theorem and from the fact that the function $\Phi(\cdot, 1/2)$ is an invertible function with a continuous inverse. Thus, the function $\Phi(\cdot, 1/2)$ plays the role of a Temple function.³⁰ These tools are those used by Towers in Ref. 31 to prove the convergence of a Godunov scheme to the entropy solution of Problem (1.1). In Ref. 32, the same author studies the convergence of an Engquist–Osher scheme associated to problem (1.2) in the case where the flux-function g is not necessarily convex and introduces a new Temple function, also in order to get compactness on the sequence of numerical approximations.

These estimates *via* the use of a Temple function play a central role; then, the fact that every limit u of a subsequence of (u_ε) is an entropy solution to Problem (1.1) is quite natural in view of Definition 1. Indeed, let $\kappa \in [0, 1]$ and ϕ be a non-negative function of $\mathcal{C}^\infty_c(\mathbb{R}_+ \times \mathbb{R})$. Suppose that T is such that $\phi(t, x) = 0$ for every $(t, x) \in [T, +\infty) \times \mathbb{R}$. For every $\varepsilon > 0$, the function u_ε satisfies the following entropy

inequality:

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{R}} |u_\varepsilon(t, x) - \kappa| \varphi_t(t, x) + k(x) \Phi(u_\varepsilon(t, x), \kappa) \varphi_x(t, x) \, dx dt \\ & + \int_{\mathbb{R}} |u_0(x) - \kappa| \varphi(0, x) \, dx - \int_0^\infty \int_{\mathbb{R}} k'_\varepsilon(x) \operatorname{sgn}(u_\varepsilon(t, x) - \kappa) g(\kappa) \phi(t, x) \, dx dt \geq 0. \end{aligned} \quad (2.7)$$

As u_ε converges to u in $L^1_{\text{loc}}([0, T] \times \mathbb{R})$, the first term of (2.7) converges to the first term of (2.4) when $\varepsilon \rightarrow 0$ so that one has to focus on the study of the last term. The estimate $|\operatorname{sgn}(u_\varepsilon - \kappa)| \leq 1$ yields $\int_0^\infty \int_{\mathbb{R}} k'_\varepsilon(x) \operatorname{sgn}(u_\varepsilon - \kappa) g(\kappa) \phi \, dx dt \leq I_\varepsilon$ where $I_\varepsilon = \int_0^\infty \int_{\mathbb{R}} |k'_\varepsilon(x)| g(\kappa) \phi \, dx dt$. To conclude, one uses the fact that the monotony of the function k_ε is set by the sign of $k_L - k_R$ and several integrations by parts to get

$$\begin{aligned} I_\varepsilon &= \operatorname{sgn}(k_R - k_L) \int_0^\infty \int_{\mathbb{R}} k'_\varepsilon(x) g(\kappa) \phi \, dx dt \\ &= -\operatorname{sgn}(k_R - k_L) \int_0^\infty \int_{\mathbb{R}} k_\varepsilon(x) g(\kappa) \partial_x \phi \, dx dt \\ &\rightarrow -\operatorname{sgn}(k_R - k_L) \int_0^\infty \int_{\mathbb{R}} k(x) g(\kappa) \partial_x \phi \, dx dt \\ &= \operatorname{sgn}(k_R - k_L) (k_R - k_L) \int_0^\infty g(\kappa) \phi(t, 0) \, dt \\ &= |k_R - k_L| \int_0^\infty g(\kappa) \phi(t, 0) \, dt. \end{aligned}$$

To sum up, when the initial condition satisfies $u_0 \in BV(\mathbb{R})$, the sequence (u_ε) is compact in $L^1_{\text{loc}}([0, T] \times \mathbb{R})$ and has at least one adherence value u , which is an entropy solution of Problem (1.1). It is the result of comparison exposed in Theorem 1 which ensures that the whole sequence $(u_\varepsilon)_\varepsilon$ converges to u . Notice that this result of comparison (2.5) also yields the existence of an entropy solution of Problem (1.1) when the initial condition is merely a function of $L^\infty(\mathbb{R}; [0, 1])$. Indeed, suppose $u_0 \in L^\infty(\mathbb{R}; [0, 1])$ and set

$$u_0^\alpha = \rho_\alpha \star (\chi_{(-1/\alpha, 1/\alpha)} u_0),$$

where ρ_α is a classical sequence of mollifiers. Then $u_0^\alpha \in L^\infty(\mathbb{R}; [0, 1]) \cap BV(\mathbb{R})$ and $\lim_{\alpha \rightarrow 0} u_0^\alpha = u_0$ in $L^1_{\text{loc}}(\mathbb{R})$. Therefore, if u^α denotes the corresponding entropy solution, then the sequence (u^α) is a Cauchy sequence in $L^1_{\text{loc}}(\mathbb{R}_+ \times \mathbb{R})$ since the comparison

$$\int_0^T \int_{(-R, R)} |u^\alpha(t, x) - u^{\alpha'}(t, x)| \, dx dt \leq T \int_{(-R-KT, R+KT)} |u_0^\alpha(x) - u_0^{\alpha'}(x)| \, dx$$

holds for every $R, T > 0$. Consequently, this sequence (u^α) is convergent, and denoting by u its limit in $L^1_{\text{loc}}(\mathbb{R}_+ \times \mathbb{R})$, the function u is an entropy solution of Problem (1.1).

Thus, the result of comparison (2.5) not only entails uniqueness or continuous dependence on the data, but is also a key point to show the existence of an entropy solution in the general framework of L^∞ functions. How to prove it?

The classical proof of uniqueness of Kruzkov²⁵ applies without changes to prove that, if u and v are two entropy solutions of Problem (1.1), if φ is a non-negative function of $\mathcal{C}^\infty_c([0, T] \times \mathbb{R})$ which vanishes in a neighborhood of the line $\{x = 0\}$ of discontinuity of the function k , then the following inequation holds

$$\begin{aligned} & \int_0^\infty \int_{\mathbb{R}} |u(t, x) - v(t, x)| \varphi_t(t, x) + k(x) \Phi(u(t, x), v(t, x)) \varphi_x(t, x) \, dx dt \\ & + \int_{\mathbb{R}} |u_0(x) - v_0(x)| \phi(0, x) \, dx \geq 0. \end{aligned} \tag{2.8}$$

In order to remove this additional hypothesis on the test function φ , a particular entropy condition on the line of discontinuity of the function k will be used. Consider any non-negative function ψ of $\mathcal{C}^\infty_c([0, T] \times \mathbb{R})$ and, for $\varepsilon > 0$, set $\varphi(t, x) = \psi(t, x) (1 - \omega_\varepsilon(x))$ in (2.8), the cutoff function ω_ε being defined by

$$\omega_\varepsilon(x) = \begin{cases} 0 & \text{if } 2\varepsilon < |x| \\ \frac{-|x| + 2\varepsilon}{\varepsilon} & \text{if } \varepsilon \leq |x| \leq 2\varepsilon \\ 1 & \text{if } |x| < \varepsilon \end{cases}.$$

Passing to the limit $\varepsilon \rightarrow 0$ in the inequality obtained in this way, one gets

$$\int_0^\infty \int_{\mathbb{R}} |u - v| \psi_t + k(x) \Phi(u, v) \psi_x \, dx dt + \int_{\mathbb{R}} |u_0 - v_0| \psi(0) \, dx - J \geq 0, \tag{2.9}$$

where

$$J = \lim_{\varepsilon \rightarrow 0} \int_0^\infty \int_{\mathbb{R}} k(x) \Phi(u, v) \psi \omega'_\varepsilon(x) \, dx dt.$$

The last term J of (2.9) will turn to be non-negative (so that (2.8) will indeed right for any non-negative function $\varphi \in \mathcal{C}^\infty_c([0, T] \times \mathbb{R})$). In order to estimate this term J , one uses the result of existence of strong traces for solutions of nondegenerate conservation laws by Vasseur.³³ In Ref. 33, it is proved that, given a conservation law

$$u_t + (A(u))_x = 0$$

set on a domain $\Omega \subset \mathbb{R}_+ \times \mathbb{R}^d$ with a flux satisfying: for all $(\tau, \zeta) \in \mathbb{R}_+ \times \mathbb{R}^d \setminus \{(0, 0)\}$, the measure of the set $\{\xi; \tau + \zeta \cdot A(\xi) = 0\}$ is zero, then the entropy solution u has strong traces on $\partial\Omega$. Applying this result to Problem (1.1) on the sets $\Omega^- = (0, +\infty) \times (-\infty, 0)$ and $\Omega^+ = (0, +\infty) \times (0, +\infty)$ respectively, we get:

Lemma 1. *Let $u \in L^\infty(\mathbb{R}_+ \times \mathbb{R})$ be an entropy solution to Problem (1.1) with initial condition $u_0 \in L^\infty(\mathbb{R})$, $0 \leq u_0 \leq 1$ a.e. Then the function u admits strong*

traces on the line $\{x = 0\}$, that is: there exists some functions γu^- and γu^+ in $L^\infty(0, +\infty)$ such that, for every compact K of $(0, +\infty)$,

$$\operatorname{ess\,lim}_{s \rightarrow 0^\pm} \int_K |u(t, s) - \gamma u^\pm| dt = 0. \quad (2.10)$$

Coming back to the definition of the cutoff function ω_ε , one gets

$$J = \int_0^\infty (k_L \Phi(\gamma u^-(t), \gamma v^-(t)) - k_R \Phi(\gamma u^+(t), \gamma v^+(t))) \psi(t, 0) dt.$$

As already mentioned, the sign of J is actually determined, considering that a Rankine–Hugoniot relation and an entropy inequality occur on the line of discontinuity of the function k (see Eqs. (2.12) and (2.13)). As usual, the Rankine–Hugoniot relation is derived from the weak formulation of Problem (1.1) (every entropy solution is a weak solution) whereas the entropy condition is a consequence of the entropy inequality (2.4) set with the parameter κ equal to the point where the function g reaches its maximum, that is $\kappa = 1/2$ here. Precisely: set $\varphi := \varphi \omega_\varepsilon$ in (2.4) and pass to the limit $\varepsilon \rightarrow 0$ in the inequality obtained in this way. Using Lemma 1 again, this yields the inequality

$$\int_0^\infty (k_L \Phi(\gamma u^-(t), \kappa) - k_R \Phi(\gamma u^+(t), \kappa)) \varphi(t, 0) dt + |k_L - k_R| \int_0^\infty g(\kappa) \varphi(0, t) dt \geq 0$$

for every $\kappa \in [0, 1]$, or, still:

$$\forall \kappa \in [0, 1], \quad \text{for a.e. } t > 0, \quad k_L \Phi(\gamma u^-(t), \kappa) - k_R \Phi(\gamma u^+(t), \kappa) + |k_L - k_R| g(\kappa) \geq 0. \quad (2.11)$$

By choosing successively $\kappa = 0$ and $\kappa = 1$ in (2.11), one derives the Rankine–Hugoniot relation

$$k_L g(\gamma u^-) = k_R g(\gamma u^+). \quad (2.12)$$

By choosing $\kappa = 1/2$ in (2.11), one derives the following “entropy condition”: denote by a the derivative of the function g , then for a.e. $t > 0$,

$$a(\gamma u^+(t)) > 0 \Rightarrow a(\gamma u^-(t)) \geq 0. \quad (2.13)$$

(We refer to Ref. 31 for the rigorous proof of this result.) This condition (2.13) can be seen as a limit of entropy conditions (cf. Sec. 3). This justifies the denomination which has been given. As the study of the Riemann problem suggests it, this condition has to be specified in order to distinguish between several potential solutions. For example, with $k_L = 2$ and $k_R = 1$, if the initial condition u_0 is defined by

$$u_0(x) = \begin{cases} \frac{1}{8} & \text{if } x < 0, \\ \frac{1}{2} - \frac{\sqrt{2}}{8} & \text{if } x > 0, \end{cases}$$

then the stationary function $u(t, x) = u_0(x)$ seems to be an acceptable solution of Problem (1.1): it is an entropy solution apart from the line of discontinuity of k while it satisfies the Rankine–Hugoniot relation (2.12) on this line. Nevertheless, it is not the admissible solution to Problem (1.1) (see Appendix A for the description of this latter). Technically, the discussion of the respective positions of γu^- , γu^+ , γv^- and γv^+ , combined with the use of Eqs. (2.12) and (2.13) allows to prove that $J \geq 0$ (see the proof of Theorem 4.6 in Ref. 31). It is then classical to conclude to (2.5).

3. Derivation of the Entropy Condition (2.13)

Suppose that $u \in L^\infty \cap BV((0, T) \times \mathbb{R})$ is a solution of Problem (1.1), in accordance with a definition that we would like to determine. It is rather natural to suppose that the function u is, at least, a weak solution, that is to say satisfies:

$$\int_0^\infty \int_{\mathbb{R}} u \varphi_t + k(x) g(u) \varphi_x = 0 \tag{3.14}$$

for all $\varphi \in \mathcal{C}_c^\infty([0, +\infty[\times \mathbb{R})$. We also suppose that the function u is a “classical” entropy solution outside the line $\{x = 0\}$, which means that for all $\kappa \in \mathbb{R}$, for all non-negative function $\varphi \in \mathcal{C}_c^\infty([0, +\infty[\times \mathbb{R})$ such that $\varphi(t, 0) = 0, \forall t \in [0, T]$,

$$\begin{aligned} &\int_0^\infty \int_{\mathbb{R}} |u(t, x) - \kappa| \varphi_t(t, x) + k(x) \Phi(u(t, x), \kappa) \varphi_x(t, x) \, dx dt \\ &\quad + \int_{\mathbb{R}} |u_0(x) - \kappa| \varphi(0, x) \, dx \geq 0. \end{aligned} \tag{3.15}$$

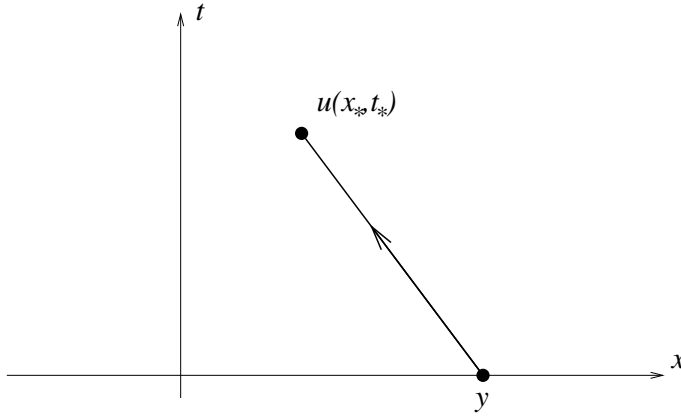
Denoting by γu^- and γu^+ the traces of the function u on $\{x = 0-\}$ and $\{x = 0+\}$ respectively, we deduce from (3.14) the following Rankine–Hugoniot condition:

$$k_L g(\gamma u^-) = k_R g(\gamma u^+). \tag{3.16}$$

Actually, if $u \in L^\infty \cap BV((0, T) \times \mathbb{R})$ satisfies (3.16) and (3.15) then Eq. (3.14) holds. Nevertheless, these two conditions are inadequate for a complete characterization of the solution: the study of the Riemann problem shows that there may be more than one single function in $L^\infty \cap BV((0, T) \times \mathbb{R})$ satisfying both (3.16) and (3.15). Thus, condition (3.16) has to be enforced with another condition, interpreted as an entropy condition on the line $\{x = 0\}$. Two possible approaches of the question are exposed here.

3.1. The characteristics method approach

Suppose that the values of the solution u are sought through the use of the characteristics method. Let t_\star be in $(0, T)$ and x_\star be in \mathbb{R} , for example $x_\star > 0$. Two cases have to be distinguished.


 Fig. 1. $a(u(t_*, x_*)) \leq 0$.

First case: $a(u(t_*, x_*)) \leq 0$

In the (x, t) -plane the equation of the half characteristic line is

$$x = x_* + k_R a(u(t_*, x_*))(t - t_*), \quad t \leq t_*.$$

Therefore, it does not intersect the line $\{x = 0\}$. The value of $u(t_*, x_*)$ is given by the value of the initial condition u_0 at the foot of the characteristic denoted by y in Fig. 1.

Second case: $a(u(t_*, x_*)) > 0$

The half characteristic line may intersect the line $\{x = 0\}$. Suppose that it happens at time $t = \tau$. Denoting by $\gamma u^+ = u(\tau, 0+)$ and $\gamma u^- = u(\tau, 0-)$ the traces of the function u , one has $u(t_*, x_*) = \gamma u^+$ (and $a(\gamma u^+) > 0$) for the solution u is constant along the characteristic lines. Thus, the aim is to determine γu^+ . The Rankine–Hugoniot condition (3.16) ensures $k_R g(\gamma u^+) = k_L g(\gamma u^-)$, which provides two possible values of γu^- , one such that $a(\gamma u^-) < 0$ and the other such that $a(\gamma u^-) \geq 0$. Besides, the calculus along the characteristics should be pursued, now starting from the point $(\tau, 0-)$. If $a(\gamma u^-) \geq 0$, then it is possible: the equation of the half characteristic line is

$$x = k_L a(u^-)(t - \tau), \quad t \leq \tau;$$

its slope is non-negative, thus it intersects the line $\{t = 0\}$ at a point y and $u(t_*, x_*) = u_0(y)$. It is this configuration which is described in Fig. 2.

If $a(u^+) < 0$, there exists an indetermination and this lack of information makes the calculus of the value $u(t_*, x_*)$ impossible by the characteristics method.

3.2. Interaction of waves

Again, a function $u \in L^\infty \cap BV((0, T) \times \mathbb{R})$ is supposed to satisfy (3.15) and (3.16). To understand which entropy condition could be imposed on the line $\{x = 0\}$,

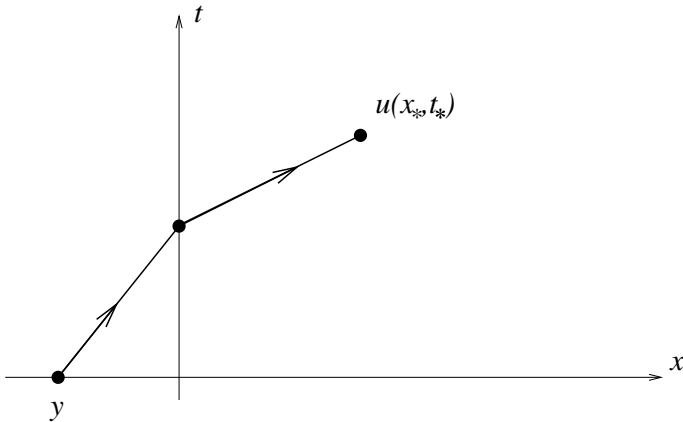


Fig. 2. $a(u(t_*, x_*)) > 0$.

that line is “thickened”: let ε be a positive number, we consider the continuous approximation k_ε of the function k define by:

$$k_\varepsilon(x) = \begin{cases} k_L & \text{if } x \leq -\varepsilon, \\ \frac{k_R - k_L}{2\varepsilon}x + \frac{k_R + k_L}{2} & \text{if } -\varepsilon \leq x \leq \varepsilon, \\ k_R & \text{if } \varepsilon \leq x \end{cases}$$

and seek for a *stationary* solution of the problem

$$\begin{cases} \frac{\partial}{\partial t}u_\varepsilon + \frac{\partial}{\partial x}(k_\varepsilon u_\varepsilon(1 - u_\varepsilon)) = 0, \\ u_\varepsilon(x \leq -\varepsilon) = u_L, \\ u_\varepsilon(x \geq \varepsilon) = u_R. \end{cases}$$

Let us denote by K_0 the quantity

$$K_0 = k_L \gamma u^-(1 - \gamma u^-) = k_R \gamma u^+(1 - \gamma u^+). \tag{3.17}$$

Then the function $u_\varepsilon = u_\varepsilon(x)$ has to satisfy the equation

$$k_\varepsilon(x)u_\varepsilon(x)(1 - u_\varepsilon(x)) = K_0, \quad x \in [-\varepsilon; \varepsilon]. \tag{3.18}$$

This equation defines two curves, whose parametrizations are denoted by u_1 and u_2 , such that (see Fig. 3):

$$\begin{aligned} x \in [-\varepsilon; \varepsilon], \quad u_1(x) &= \frac{1}{2} - \frac{\sqrt{k_\varepsilon(x)^2 - 4k_\varepsilon(x)K_0}}{2k_\varepsilon(x)}, \\ x \in [-\varepsilon; \varepsilon], \quad u_2(x) &= \frac{1}{2} + \frac{\sqrt{k_\varepsilon(x)^2 - 4k_\varepsilon(x)K_0}}{2k_\varepsilon(x)}. \end{aligned} \tag{3.19}$$

It is clear that:

$$\forall x \in [-\varepsilon; \varepsilon], \quad 0 \leq u_1(x) \leq \frac{1}{2} \leq u_2(x) \leq 1.$$

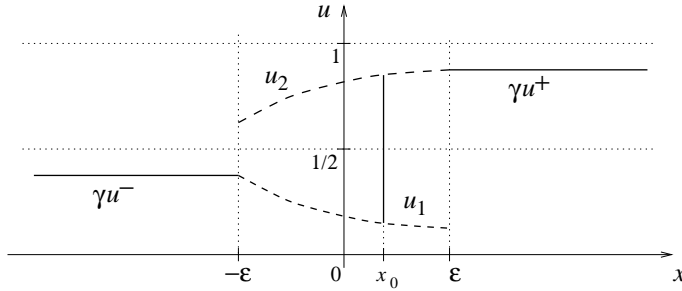


Fig. 3. Stationary connection between γu^- and γu^+ with $k(x) = k_\varepsilon(x)$.

The numbers γu^- and γu^+ satisfy Eq. (3.17); consequently, if $(1 - 2\gamma u^+)(1 - 2\gamma u^-) \geq 0$, that is to say:

$$a(\gamma u^+)a(\gamma u^-) \geq 0, \quad (3.20)$$

then γu^+ and γu^- can be linked by one of the two curves u_1 or u_2 . Now, suppose that condition (3.20) is not fulfilled. Then, in order to link γu^- to γu^+ , one has to introduce a stationary discontinuity between the two curves so that the following Rankine–Hugoniot condition may be satisfied:

$$k_\varepsilon(x_0)u_\varepsilon(x_0^-)(1 - u_\varepsilon(x_0^-)) = k_\varepsilon(x_0)u_\varepsilon(x_0^+)(1 - u_\varepsilon(x_0^+)),$$

the point x_0 denoting the point where the discontinuity occurs (see Fig. 3). As the function $g: u \mapsto u(1 - u)$ is concave, this discontinuity is admissible (in the entropy sense) if $u(x_0^+) > u(x_0^-)$. Letting ε tend to zero, this yields the condition

$$\gamma u^+ > \gamma u^-. \quad (3.21)$$

Eventually, we are led to the following condition:

$$\begin{cases} \text{either } a(\gamma u^+)a(\gamma u^-) \geq 0, \\ \text{either } a(\gamma u^+)a(\gamma u^-) < 0 \text{ and } \gamma u^+ > \gamma u^-. \end{cases}$$

This condition implies condition (2.13) (and is equivalent to it, in fact). Indeed, suppose $a(\gamma u^+) > 0$. Then, either $a(\gamma u^-) \geq 0$ holds, and in that case condition (3.20) is fulfilled, either $a(\gamma u^-) < 0$, and in that case

$$0 \leq \gamma u^+ < 1/2 < \gamma u^- \leq 1,$$

which contradicts condition (3.21). Notice that, here, condition (2.13) is derived by considering the limit of an entropy condition: this justifies the denomination of “entropy” condition to design the condition (2.13).

To conclude, condition (2.13) can be interpreted as the admissibility condition for the superposition of a stationary shock with a contact discontinuity. In several applications, shallow-water equations with topography (see Refs. 26 and 29) for instance, the construction of the solution of the Riemann problem is closely related to this condition. In fact, it is the resonance of the studied system which permits

to select the solution of the related Riemann problem. This kind of phenomenon may also be observed in the study of two-phase flows, with two-fluid two-pressure models.

4. Numerical Methods

All the methods presented in this section are finite volume methods (see Refs. 6 and 14). For the sake of simplicity, the presentation is restricted to regular meshes (though all methods may be naturally extended to irregular meshes). Let Δx be the space step, with $\Delta x = x_{i+1/2} - x_{i-1/2}$, $i \in \mathbb{Z}$, and let Δt be the time step, with $\Delta t = t^{n+1} - t^n$, $n \in \mathbb{N}$. Besides, let u_i^n denote the approximation of $\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} u(t^n, x) dx$. Integrating Eq. (1.1) over the cell $]x_{i-1/2}; x_{i+1/2}[\times [t^n; t^{n+1})$ yields:

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x} (\phi_{i+1/2}^n - \phi_{i-1/2}^n),$$

where $\phi_{i+1/2}^n$ is the numerical flux through the interface $\{x_{i+1/2}\} \times [t^n; t^{n+1})$. Let us emphasize that the permeability $k(x)$ is approximated by a piecewise constant function:

$$k_i = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} k(x) dx, \quad i \in \mathbb{Z}. \tag{4.22}$$

The numerical flux $\phi_{i+1/2}^n$ depends on k_i , k_{i+1} , u_i^n and u_{i+1}^n , and a consistency criterion is imposed:

$$u_i^n = u_{i+1}^n = u_0 \quad \text{and} \quad k_i = k_{i+1} = k_0 \implies \phi_{i+1/2}^n = k_0 u_0 (1 - u_0).$$

Moreover, a C.F.L. condition is associated with the time step Δt to ensure the stability of the scheme. Notice that all the methods presented here rely on conservative schemes, since the problem is conservative. The four schemes introduced are three-point schemes, as mentioned above. A higher order extension is also presented (five-point schemes), in order to increase the accuracy of the methods and their rates of convergence (when $\Delta x \rightarrow 0$).

4.1. Scheme 1

The first scheme is defined by the following numerical flux:

$$\phi_{i+1/2}^n = \frac{2k_i k_{i+1}}{(k_i + k_{i+1})} \frac{u_i^n (1 - u_{i+1}^n)}{(u_i^n + (1 - u_{i+1}^n))}. \tag{4.23}$$

The C.F.L. condition associated with this scheme is the classical condition which limits the time step Δt , according to the maximal speed of waves, computed on each cells:

$$\lambda_{\text{MAX}} \frac{\Delta t}{\Delta x} < \frac{1}{2}, \quad \text{where } \lambda_{\text{MAX}} = \max_{i \in \mathbb{Z}, n \in \mathbb{N}} (k_i (1 - 2u_i^n)). \tag{4.24}$$

For this scheme, the design of the fluxes is based on methods usually implemented in industrial codes.

4.2. Scheme 2

Physical considerations drive the conception of the second scheme:

$$\phi_{i+1/2}^n = \frac{k_i u_i^n k_{i+1} (1 - u_{i+1}^n)}{k_i u_i^n + k_{i+1} (1 - u_{i+1}^n)}. \quad (4.25)$$

Indeed, the physical variables are ku and $k(1-u)$, rather than k and u . The C.F.L. condition is the same as condition (4.24), associated with scheme 1.

4.3. The Godunov scheme

The Godunov scheme¹² is based on the resolution of the Riemann problem at each interface of the mesh. The application of the Godunov scheme to the framework of space-dependent flux may be seen as an extension of the works of Greenberg and Le Roux to deal with conservation laws with source terms (see Refs. 11 and 26). The Godunov method applied to resonant systems like (1.2) has been studied by Lin, Temple and Wang.^{27,28} A specific Godunov scheme associated to Problem (1.1) has been defined by Towers by considering a discretization of k staggered with respect to that of u .³¹ Here, we consider the Godunov method applied to the 2×2 system (1.2).

$$\begin{cases} \frac{\partial k}{\partial t} = 0, & \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(ku(1-u)) = 0, & t > t^n, x \in \mathbb{R}, \\ u(0, x) = \begin{cases} u_i^n & \text{if } x < x_{i+1/2} \\ u_{i+1}^n & \text{if } x > x_{i+1/2} \end{cases}, & k(x) = \begin{cases} k_i & \text{if } x < x_{i+1/2} \\ k_{i+1} & \text{if } x > x_{i+1/2} \end{cases}. \end{cases} \quad (4.26)$$

Let $u_{i+1/2}^n((x - x_{i+1/2})/(t - t^n); k_i, k_{i+1}, u_i^n, u_{i+1}^n)$ be the exact solution of this Riemann problem (see Appendix A for an explicit presentation of the solution). Since the function k is discontinuous through the interface $\{x_{i+1/2}\} \times [t^n; t^{n+1})$, the solution $u_{i+1/2}^n$ is also discontinuous through this interface. However, as the problem is conservative, the flux is continuous through this interface, and writes:

$$\begin{aligned} \phi_{i+1/2}^n &= k_i g(u_{i+1/2}^n(0^-; k_i, k_{i+1}, u_i^n, u_{i+1}^n)) \\ &= k_{i+1} g(u_{i+1/2}^n(0^+; k_i, k_{i+1}, u_i^n, u_{i+1}^n)), \end{aligned} \quad (4.27)$$

where $g(u) = u(1-u)$. Here, the C.F.L. condition is based on the maximal speed of waves associated with each local Riemann problem.

Remark 3. Since the problem involves a homogeneous conservation law, the definition of the flux is not ambiguous. Indeed, in the framework of shallow-water equations with topography for instance (see Refs. 7 and 26), the source term “breaks” the conservativity of the system and the flux becomes discontinuous through the interface. Moreover, the approximation of the topography by a piecewise constant function introduces a product of distributions, which is not defined (contrary to what happen in the current framework, where all jumps relations are well defined).

4.4. The VFRoe-ncv scheme

We present herein an approximate Godunov scheme, based on the exact solution of a linearized Riemann problem. A VFRoe-ncv scheme is defined by a change of variables (see Refs. 1 and 8), the new variable is denoted by $\theta(k, u)$. The choice of θ is motivated by the properties of the VFRoe-ncv scheme induced from this change of variables. Here, we set $\theta(k, u) = ku(1 - u)$ to be the new variable. If v is defined by $v(t, x) = \theta(k(x), u(t, x))$, then the VFRoe-ncv scheme is based on the exact resolution of the following linearized Riemann problem:

$$\begin{cases} \frac{\partial v}{\partial t} + (\hat{k}(1 - 2\hat{u}))\frac{\partial v}{\partial x} = 0, & t > t^n, \ x \in \mathbb{R}, \\ v(0, x) = \begin{cases} \theta(k_i, u_i^n) & \text{if } x < x_{i+1/2} \\ \theta(k_{i+1}, u_{i+1}^n) & \text{if } x > x_{i+1/2} \end{cases}, \end{cases} \tag{4.28}$$

where $\hat{k} = (k_i + k_{i+1})/2$ and $\hat{u} = (u_i^n + u_{i+1}^n)/2$.

The (nonlinear) initial problem thus becomes a classical convection equation, and the VFRoe-ncv scheme is reduced to the well-known upwind scheme for problem (4.28). Hence, as the Godunov scheme, the flux (which is represented by v) is continuous through the interface $\{x_{i+1/2}\} \times [t^n; t^{n+1})$ (this property is provided by the choice of θ ; another choice would lead to a jump of the flux v through the local interface). If $v_{i+1/2}^n((x - x_{i+1/2})/(t - t^n); k_i, k_{i+1}, u_i^n, u_{i+1}^n)$ is the exact solution of the Riemann problem (4.28), the numerical flux of the VFRoe-ncv scheme is:

$$\phi_{i+1/2}^n = v_{i+1/2}^n(0; k_i, k_{i+1}, u_i^n, u_{i+1}^n). \tag{4.29}$$

The C.F.L. condition is determined by the speed waves generated by the local Riemann problems, like the Godunov scheme (but the speed computed by the two methods are *a priori* different). Since the VFRoe-ncv scheme relies on the resolution of a linearized Riemann problem, an entropy fix has to be applied to avoid the occurrence of non-entropic shock when dealing with sonic rarefaction wave.¹⁶

Remark 4. For some initial conditions, the VFRoe-ncv scheme may compute exactly the same numerical results as the Godunov scheme. Indeed, if each local Riemann problem can be solved by an upwinding on v (i.e. no sonic point arises), both methods compute the same numerical flux (since it is completely defined by one of the two initial states).

4.5. A higher order extension

Classical methods to increase the accuracy and the rate of convergence (when $\Delta x \rightarrow 0$) of finite volume schemes call for a piecewise linear reconstruction by cell. A second-order Runge–Kutta method (also known as the Heun scheme) is associated with the reconstruction to approximate time derivatives. The linear reconstruction introduced here lies within a framework introduced by B. Van Leer in Ref. 34,

namely MUSCL (monotonic upwind schemes for conservation laws), with the minmod slope limiter. This formalism is usually applied to homogeneous conservation laws (see Ref. 8 for numerical results and measurements, dealing with the Euler system). However, a classical reconstruction applied in our context may penalize results provided by the initial algorithm, when simulating convergence in time (i.e. when $t \rightarrow +\infty$) towards stationary states. This phenomenon will be pointed out later by numerical results, and a method to avoid it is proposed (see Ref. 7 for a presentation of this method adapted to the approximation of the shallow-water equations with topography).

A MUSCL scheme may be described by the following three steps algorithm:

- (1) Let $\{u_i^n\}_{i \in \mathbb{Z}}$ be a piecewise constant approximation, compute $u^{\text{lin}}(t^n, x)$, a piecewise linear function.
- (2) Solve the conservation law with $u^{\text{lin}}(t^n, x)$ as initial condition and obtain $u^{\text{lin}}(t^{n+1}, x)$.
- (3) Compute $\{u_i^{n+1}\}_{i \in \mathbb{Z}}$, averaging $u^{\text{lin}}(t^{n+1}, x)$.

The method proposed in this section only modifies the first step. It requires the use of the minmod slope limiter and preserves its main properties (see Ref. 14 and references therein). See Remark 5 for details about the computation of steps 2 and 3.

In the following of this section, we drop the time dependence for all variables. Let $\{w_i\}_{i \in \mathbb{Z}}$ be a variable constant on each cell $I_i = [x_{i-1/2}; x_{i+1/2}]$ and $x_i = (x_{i+1/2} + x_{i-1/2})/2$. Let $\delta_i(w)$ be the slope associated with w_i on the cell I_i . Moreover, let $w_i^{\text{lin}}(x)$, $x \in I_i$, the linear function defined by:

$$w_i^{\text{lin}}(x) = w_i - \delta_i(x - x_i), \quad x \in I_i.$$

To compute the slope $\delta_i(w)$, the minmod slope limiter is used:

$$\delta_i(w) = \begin{cases} s_{i+1/2}(w) \min(|w_{i+1} - w_i|, |w_i - w_{i-1}|) / \Delta x & \text{if } s_{i-1/2}(w) = s_{i+1/2}(w), \\ 0 & \text{otherwise,} \end{cases} \quad (4.30)$$

where $s_{i+1/2}(w) = \text{sgn}(w_{i+1} - w_i)$. This linear reconstruction fulfills:

Proposition 1. *If w^{cst} and w^{lin} respectively denote the functions defined by the constant and linear piecewise approximations of w ,*

$$\begin{aligned} w^{\text{cst}}(x) &= w_i & i \in \mathbb{Z} \text{ such that } x \in I_i, \\ w^{\text{lin}}(x) &= w_i^{\text{lin}} & i \in \mathbb{Z} \text{ such that } x \in I_i, \end{aligned}$$

then w^{lin} , defined by the minmod slope limiter (4.30) from w^{cst} , verifies

$$|w^{\text{lin}}|_{BV(\mathbb{R})} = |w^{\text{cst}}|_{BV(\mathbb{R})}. \quad (4.31)$$

When dealing with a classical scalar conservation law $\partial_t u + \partial_x f(u) = 0$, the reconstruction is performed on the conservative variable $w = u$. Hence, Proposition 1 ensures that the method does not introduce oscillations on the modified variable.

Though the equation studied here is a conservation law, some properties deeply differ from the classical framework $\partial_t u + \partial_x f(u) = 0$, due to the space-dependence of the flux. Indeed, focusing on stationary states, the characterization of the latter is quite different; whereas piecewise constant functions represent stationary states for $\partial_t u + \partial_x f(u) = 0$, their form closely depends on the shape of $k(x)$ here, since stationary states have to verify $\partial_x(ku(1-u)) = 0$. Using the classical finite volume formalism, this equation becomes the following discrete equation:

$$\forall i \in \mathbb{Z}, \quad k_i u_i (1 - u_i) = k_{i+1} u_{i+1} (1 - u_{i+1}). \tag{4.32}$$

Thus, if the previous reconstruction with $w = u$ is computed in the current framework, a discrete stationary state (4.32) is not maintained by the whole algorithm (even assuming that the initial scheme is able to maintain it). Furthermore, convergence in time towards a stationary state may be altered (some numerical tests illustrate this numerical phenomenon in the following).

A way to avoid this problem is to take into account relation (4.32) in the linear reconstruction. Recalling the notation $v(x) = \theta(k(x), u(x)) = k(x)u(x)(1-u(x))$, (4.32) becomes $v_i = v_{i+1}$, $\forall i \in \mathbb{Z}$. A natural idea is to use the linear reconstruction with $w = v$, to ensure that oscillations do not occur when dealing with stationary states. Indeed, such a reconstruction implies the constraint (4.31) on the total variation of v (Proposition 1). However, the change of variables from (k, u) to (k, v) is not invertible, which forbids the use of this reconstruction (see Remark 6 for more details). Moreover, Proposition 1 on u is lost if the minmod reconstruction is only computed on v .

We propose here to keep the linear reconstruction on u with the minmod slope limiter (4.30), and to add a BV-like constraint on v (following equality (4.31)). Let us first introduce the function ϑ_i defined for $x \in I_i$ by

$$\vartheta_i(x) = \begin{cases} \frac{2(x_i - x)}{\Delta x} \theta(k_i, u^{\text{lin}}(x_{i-1/2}^+)) \\ \quad + \frac{2(x - x_{i-1/2})}{\Delta x} \theta(k_i, u^{\text{lin}}(x_i)) & \text{if } x \in]x_{i-1/2}; x_i], \\ \frac{2(x - x_i)}{\Delta x} \theta(k_i, u^{\text{lin}}(x_{i+1/2}^-)) \\ \quad + \frac{2(x_{i+1/2} - x)}{\Delta x} \theta(k_i, u^{\text{lin}}(x_i)) & \text{if } x \in]x_i; x_{i+1/2}[\end{cases}$$

and the function ϑ defined for $x \in \mathbb{R}$ by

$$\vartheta(x) = \vartheta_i(x) \quad i \in \mathbb{Z} \text{ such that } x \in I_i.$$

This function represents the linear interpolation provided by the values of u^{lin} at each interface of the mesh and at each center of cells (see Fig. 4). The function ϑ is linear on each interval $]x_{i-1/2}; x_i[$ and $]x_i; x_{i+1/2}[$, $\forall i \in \mathbb{Z}$. Moreover, the function ϑ is discontinuous at each interface $x_{i+1/2}$ and continuous at each center of cell x_i .

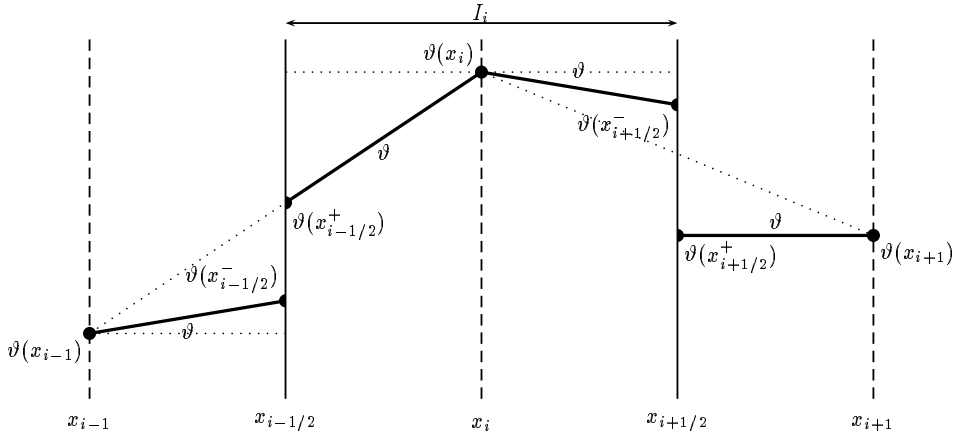


Fig. 4. A BV-like reconstruction to deal with stationary states.

We want to impose that the reconstruction on u verifies the *a posteriori* criterion:

$$|\vartheta|_{BV(\mathbb{R})} = |\theta(k^{\text{cst}}, u^{\text{cst}})|_{BV(\mathbb{R})}, \quad (4.33)$$

where $\theta(k^{\text{cst}}, u^{\text{cst}})$ represents the function v computed from the piecewise constant approximations of k and u . Equality (4.33) may be seen as the counterpart of equality (4.31) for v . In practice, Eq. (4.33) becomes:

$$\begin{aligned} \forall i \in \mathbb{N}, \quad 0 \leq |\vartheta(x_i) - \vartheta(x_{i-1/2}^-)| &\leq |\vartheta(x_i) - \vartheta(x_{i-1})|/2, \\ \forall i \in \mathbb{N}, \quad 0 \leq |\vartheta(x_{i+1/2}^+) - \vartheta(x_i)| &\leq |\vartheta(x_{i+1}) - \vartheta(x_i)|/2. \end{aligned} \quad (4.34)$$

In fact, (4.34) only implies (4.33). If condition (4.34) is not fulfilled, the slope is reset to $\delta_i(u) = 0$ (for instance, function ϑ represented in Fig. 4 fulfills conditions (4.34)).

Thus, one may sum up the reconstruction on u with the following algorithm:

- (i) Computation of $\delta_i(u)$ by the minmod slope limiter:

$$\delta_i(u) = \begin{cases} s_{i+1/2}(u) \min(|u_{i+1} - u_i|, |u_i - u_{i-1}|)/\Delta x & \text{if } s_{i-1/2}(u) \\ & = s_{i+1/2}(u), \\ 0 & \text{otherwise.} \end{cases}$$

- (ii) If condition (4.34) is not fulfilled by the linear approximation of u provided by this computation of $\delta_i(u)$, the slope is reset to $\delta_i(u) = 0$, otherwise, the slope is not modified.

Hence, using this method, the two constraints on the total variation of u and v , namely (4.31) and (4.33), are fulfilled. The algorithm of reconstruction is not optimal focusing on condition (4.33), since $|\vartheta|_{BV(\mathbb{R})} \leq |\theta(k^{\text{lin}}, u^{\text{lin}})|_{BV(\mathbb{R})}$, but it minimizes the number of logical tests.

Some numerical results are presented in the next section. Rates of convergence when the mesh is refined and measurements of the convergence to stationary states are provided. These tests are performed without reconstruction, with the minmod reconstruction on u , with the modified reconstruction, and with the reconstruction on v for the VFRoe-ncv scheme (see Remark 6).

Remark 5. A modification of the minmod slope limiter has been proposed here to take into account stationary states when the MUSCL formalism is applied. Hence, steps 2 and 3 are not modified and properties enumerated in Ref. 14 related to the minmod limiter are maintained. However, most of these properties are obtained using generalized Riemann problems in step 2 and exact finite volume integration in step 3. Here, classical Riemann problems have been used to compute numerical fluxes $\phi_{i+1/2}^n$ and the Heun scheme is used to approximate time derivatives.

Remark 6. Since the change of variables from (k, u) to (k, v) is not invertible, the reconstruction on v cannot be performed with a standard scheme. However, due to the specific form of the VFRoe-ncv scheme presented in Sec. 4.4, one may compute the linear reconstruction with the minmod slope limiter (4.30) on v with this scheme. Some numerical tests are presented in the following to illustrate the validity of this first idea.

5. Numerical Experiments

Several numerical tests are presented here, as well as qualitative results and quantitative results. All the schemes introduced are performed using no reconstruction, reconstruction on u , modified reconstruction and reconstruction on v with the VFRoe-ncv scheme.

5.1. Qualitative results

The following two tests correspond to Riemann problems. The length of the domain is 10 m. The mesh is composed of 100 cells and the C.F.L. condition is set to 0.45 for all schemes (recall however that the computation of speed of waves is different according to the scheme). The variables u and $v = ku(1-u)$ are plotted, in order to estimate the behavior of all schemes through the interface $\{x/t = 0\}$. Notice that no reconstruction has been used here.

The initial conditions of the first Riemann problem are $k_L = 2$, $k_R = 1$, $u_L = 0.5$ and $u_R = 0.3$. The results of Fig. 5 are plotted at $t = 4$ s. The analytic solution of this Riemann problem may be found in Appendix A. The numerical approximations provided by schemes 1 and 2 are very close to each other. One may notice that these two schemes are very diffusive, and that the intermediate state on the left of the interface is not well approximated (though the difference with the exact solution tends to 0 when the mesh is refined). On the other hand, the results provided by the VFRoe-ncv scheme and the Godunov scheme are very similar, and more accurate

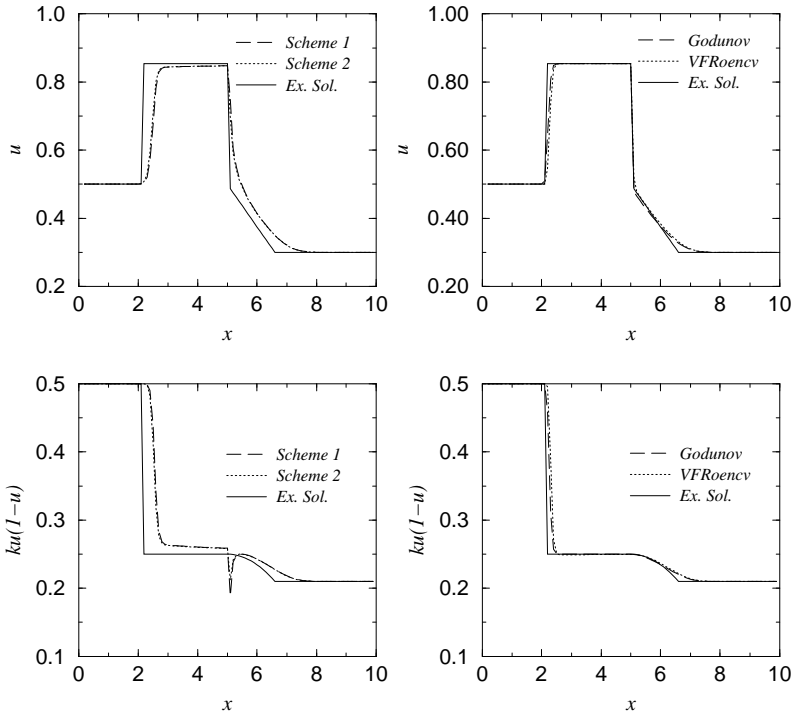


Fig. 5. 100 cells — $k_L = 2$, $k_R = 1$, $u_L = 0.5$ and $u_R = 0.3$.

than the results provided by the latter two schemes. Moreover, contrary to schemes 1 and 2, the VFRoe-ncv scheme and the Godunov scheme do not introduce oscillation on v at the interface $\{x/t = 0\}$.

The initial conditions of the second Riemann problem are $k_L = 2$, $k_R = 1$, $u_L = 0.95$ and $u_R = 0.8$. The results at $t = 2$ s are presented in Fig. 6. Same comments may be made about schemes 1 and 2, in particular about their behavior at the interface $\{x/t = 0\}$, where an undershoot is detected on the variable v (its length represents about 60% of $v_R - v_L$ for scheme 1 and about 25% of $v_R - v_L$ for scheme 2). Furthermore, the two schemes (in particular scheme 1) introduce a loss of monotonicity at the end of the rarefaction wave. Approximations provided by the VFRoe-ncv scheme and the Godunov scheme are better. Moreover, the two schemes exactly compute here the same results. As noticed in the previous test, the contact discontinuity is perfectly approximated (no point is introduced in the discontinuity) and the monotonicity of the rarefaction wave is maintained.

5.2. Quantitative results

We study in this section the ability of the schemes to converge towards the entropy solution. The first case provides measurements of the rates of convergence of the

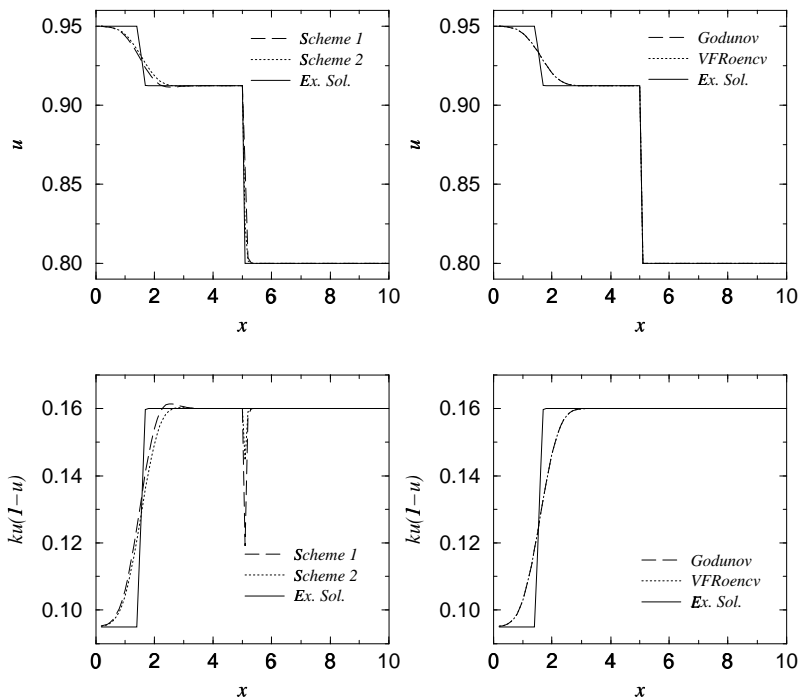


Fig. 6. 100 cells — $k_L = 2$, $k_R = 1$, $u_L = 0.95$ and $u_R = 0.8$.

methods in the L^1 norm when the mesh is refined. The second test gives some results about the variation in time in the L^2 norm of the numerical approximations when dealing with a transient simulation which converges towards a stationary state.

5.2.1. Convergence related to the space step

The computations of this test are based on the first of the two Riemann problems just discussed. The simulations are stopped at $t = 4$ s. The main interest of this test is that the analytic solution is composed of a shock wave, a contact discontinuity and a rarefaction wave (see Fig. 5). Some measurements of the numerical error provided by the methods when Δx tends to 0 are exposed. Let us define the L^1 norm of the numerical error by $\Delta x \sum_{i=1,\dots,N} |u_i^{\text{app}} - u^{\text{ex}}(x_i)|$. Several meshes are considered: involving 1000, 3000, 10,000 and 30,000 cells. Figure 7 represents, in a logarithmic scale, the profiles of the error provided by the different schemes with the different reconstructions. Moreover, Table 1 enumerates the different rates of convergence computed between the meshes with 10,000 and 30,000 nodes. The four schemes have been tested (a) without reconstruction, (b) with the minmod reconstruction on u , (c) with the modified reconstruction, and (d) with the minmod reconstruction on v for the VFRoe-nv scheme. As expected, all the methods converge towards the entropy solution (see Fig. 7). Notice first that, with any reconstruction, the Godunov

scheme is much more accurate than schemes 1 and 2. Indeed, if the Godunov scheme is performed on a mesh with N cells, schemes 1 and 2 must be performed on a mesh containing between $7N$ and $8N$ cells, in order to obtain the same L^1 error (with the same C.F.L. condition for all schemes), notice that the CPU time required is about the same for all schemes. Comparing the VFRoe-ncv scheme and the Godunov scheme, one may remark that the results are very close to each other when no reconstruction is active. On the other hand, if a reconstruction is added to the latter two schemes, the Godunov scheme becomes twice more accurate than the VFRoe-ncv scheme (and remains about four times more accurate than schemes 1 or 2). Notice that the difference between the reconstruction on u and the modified reconstruction is not significant.

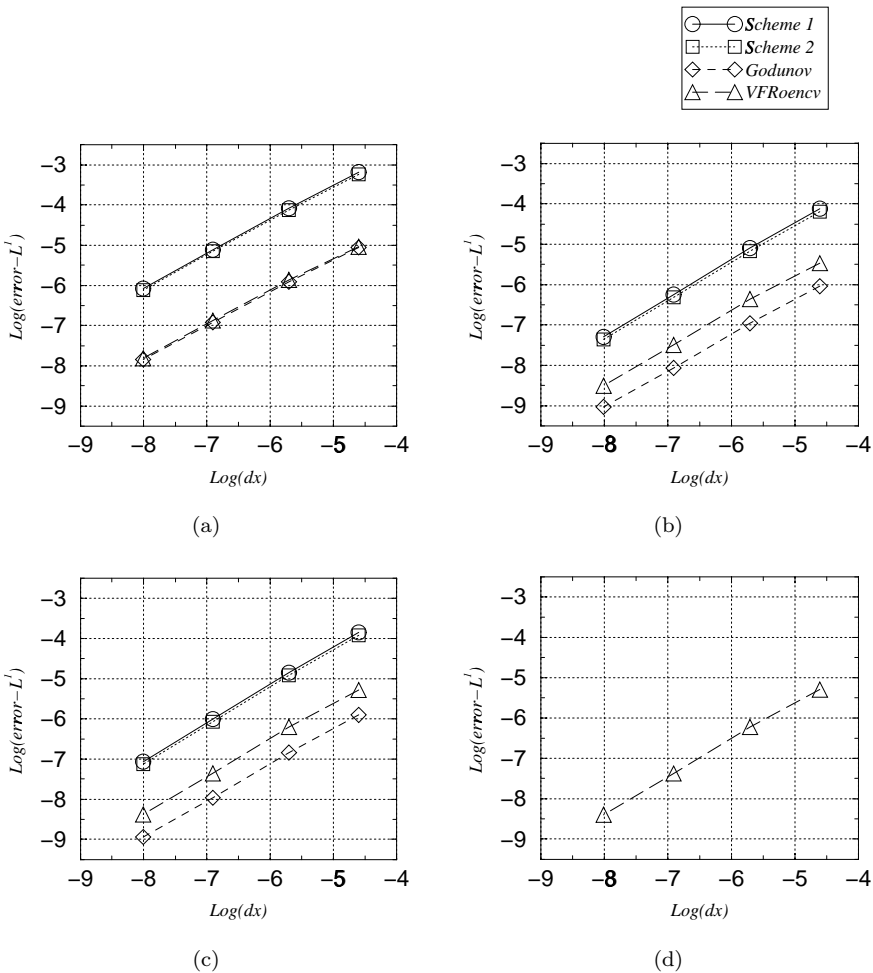


Fig. 7. Profiles of convergence for variable u . (a) no reconstruction, (b) reconstruction on u , (c) modified reconstruction, (d) reconstruction on v .

Table 1. Rates of convergence for variable u .

Reconstruction	Scheme 1	Scheme 2	Godunov	VFRoe-ncv
No reconstruction	0.88	0.87	0.82	0.85
Reconstruction on u	0.96	0.95	0.87	0.92
Modified reconstruction	0.96	0.96	0.89	0.93
Reconstruction on v	—	—	—	0.93

In Table 1 the rates of convergence provided by all the methods are listed. A reconstruction increase the rates of convergence. Moreover, the reconstruction on u and the modified reconstruction provide similar rates. Notice that, the more accurate a method is, the less important the convergence rate is, but differences are not significant. As above, the results provided by the reconstruction on u and the modified reconstruction are very similar (the reconstruction on v with the VFRoe-ncv scheme provides the same behavior). Hence, the computation of one of the reconstructions enables one to obtain much more accurate results (see Fig. 7) and to increase the rate of convergence associated with a scheme.

5.2.2. Convergence towards a stationary state

The test studied herein simulates a transient flow which converges towards a stationary state when t tends to $+\infty$. The initial conditions are:

$$k(x) = \begin{cases} 2 & \text{if } 0 \leq x \leq 2.5 \\ \frac{25 - 2x}{10} & \text{if } 2.5 < x < 7.5 \\ 1 & \text{if } 7.5 \leq x \leq 10 \end{cases}$$

and

$$u_0(x) = \begin{cases} 0.9 & \text{if } 0 \leq x \leq 2.5, \\ \frac{1 + \sqrt{0.28}}{2} & \text{if } 2.5 < x \leq 10. \end{cases}$$

The stationary solution writes

$$u(t = +\infty, x) = \begin{cases} 0.9 & \text{if } 0 \leq x \leq 2.5, \\ \frac{1}{2} + \frac{\sqrt{k(x)^2 - 0.72k(x)}}{2k(x)} & \text{if } 2.5 < x < 7.5, \\ \frac{1 + \sqrt{0.28}}{2} & \text{if } 7.5 \leq x \leq 10. \end{cases} \tag{5.35}$$

The mesh used for this test contains 100 cells, the C.F.L. condition is set to 0.45. We define the L^2 norm of the variation in time by $(\Delta x \sum_{i=1,\dots,100} |u_i^{n+1} - u_i^n|^2)^{1/2}$. The variation along the time is represented in Fig. 8 for all methods and reconstructions. Notice that all methods converge towards the stationary state (5.35) without any

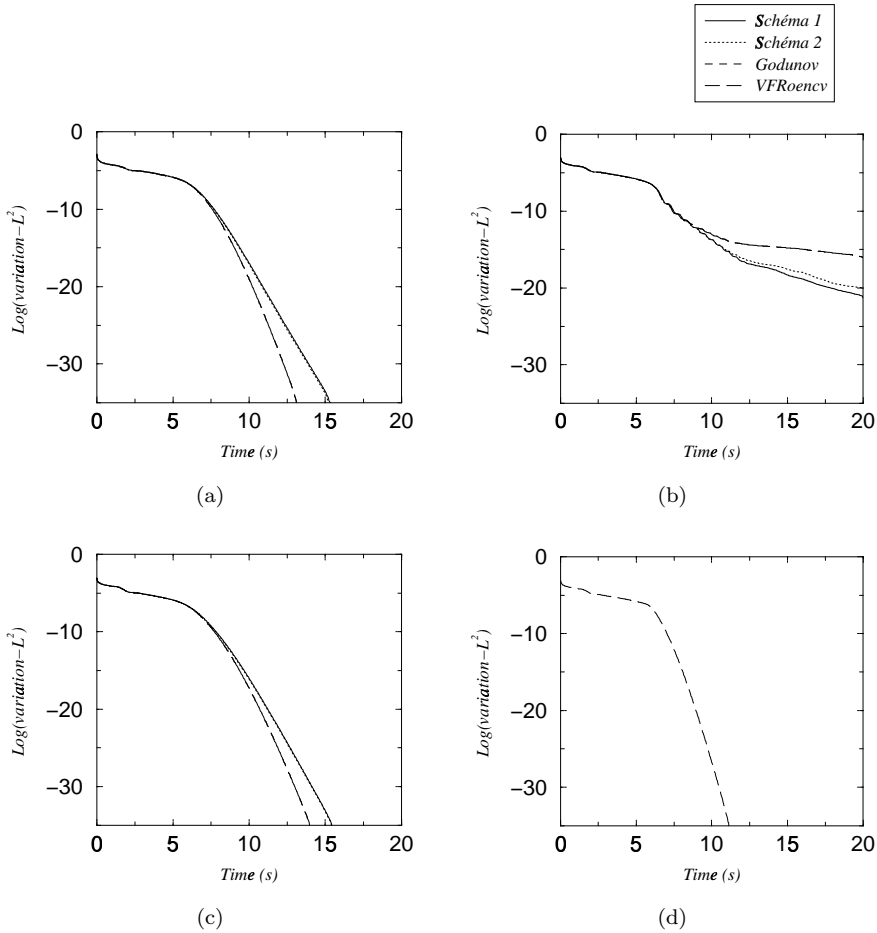


Fig. 8. Variation in time for variable u . (a) no reconstruction, (b) reconstruction on u , (c) modified reconstruction, (d) reconstruction on v .

reconstruction (see Fig. 8(a)). Moreover, the profiles computed by the VFRoe-ncv scheme and the Godunov scheme are superposed, and provide rates of convergence more important than the profiles computed by schemes 1 and 2 (which are also both superposed). If the classical MUSCL reconstruction on u is performed (Fig. 8(b)), rates of convergence are altered (all schemes nonetheless converge), in particular for the VFRoe-ncv scheme and the Godunov scheme. On the other hand, when the modified version of the MUSCL reconstruction is used (Fig. 8(c)), the behavior of all algorithms is better, and the rates of convergence are close to (but slightly less important than) the rates obtained without any reconstruction. Hence, taking into account stationary states in the MUSCL reconstruction enables to improve the behavior of the methods when dealing with convergence towards steady states. To confirm it, see Fig. 8(d). It represents the results computed by the VFRoe-ncv

scheme with the minmod reconstruction on v . Here, the rates of convergence are the greatest of those obtained by all other approximations. Similar results are obtained when the stationary solution contains a stationary shock. Similar simulations (convergence towards a steady state) have been performed with the VFRoe-ncv scheme for shallow-water equations with topography.⁷ Nevertheless, in that case, the gap between the results computed with the classical reconstruction and the modified version is quite important. Indeed, the classical minmod reconstruction provides non-convergent results (the L^2 variation in time does not decrease) whereas the modified reconstruction gives convergent and accurate results. This phenomenon may be due to the non-conservativeness of the whole system of shallow-water equations with topography.

6. Conclusion

A scalar conservation is studied here. Its flux depends not only on the conservative variable u but also on the space variable x *via* the permeability k which is a discontinuous function. Thanks to the conservative form of the equation, some jump relations are classically defined (whereas a product of distribution may occur when dealing with source terms or non-conservative equations). Existence and uniqueness of the entropy solution hold. The particular entropy admissibility criterion (namely condition (2.13)) has been discussed. Notice that this condition is independent of the choice of the integration path on k , contrary to the non-conservative framework⁵ (here, the path is only assumed to be regular and monotone). Several ways have been presented to recover this condition (or to clarify it). One of them deals with the superposition of two waves, namely the discontinuity on $\{x = 0\}$ and a shock wave. For this purpose, the jump of k is replaced by a linear connection k_ε in $] -\varepsilon; +\varepsilon[$. Stationary states are investigated inside this “thickness”. These solutions may be totally smooth, or they may involve a stationary shock wave. Obviously, this shock wave must agree with the entropy condition (which is classical, since k_ε is continuous). Hence, passing to the limit, the condition on the stationary shock wave remains for the solution u and is equivalent to condition (2.13). This point of view may be very useful in the study of some problems. For instance, focusing on shallow-water equations with topography, a similar phenomenon occurs when dealing with a discontinuous bottom (though the system is restricted to smooth topography, a bottom step is used to study the associated Riemann problem). Indeed, the solution of the Riemann problem (completely described in Ref. 29) may be defined without any ambiguity only as soon as the resonant case is completely understood. Furthermore, such a problem of resonance arises in two-phase flows, with a two-fluid two-pressure approach.⁹ Since these two problems are non-conservative, the derivation of the limit $[\varepsilon \rightarrow 0]$ is not so clear and the condition on the discontinuity becomes no more than a conjecture.

Several numerical schemes have also been provided, in order to simulate the conservation law. Two finite volume schemes have been introduced, derived from

those used in the industrial context. Two other schemes have been proposed, following the ideas of Greenberg and Le Roux,¹¹ the first one based on the exact solution of the Riemann problem and the other based on the exact solution of a linearized Riemann problem, respectively the Godunov scheme¹² and the VFRoe-nvcv scheme.¹ Some qualitative and quantitative tests confirm the good behavior of the latter two schemes, in comparison with the others two. A great difference of accuracy may be detected, especially through the discontinuity of permeability. Notice, however, that the rates of convergence (when the mesh is refined) of all the schemes are very close to each other. To increase the accuracy and the speed of convergence of all schemes, a higher order method is also presented. Though a classical MUSCL method can fulfill all these requirements, if we focus on convergence towards stationary states, the convergence may be perturbed by a classical reconstruction (even lost for shallow-water equations with topography). Hence, a modification of the reconstruction is proposed and tested (see Ref. 7 for the application to shallow-water equations with topography). The good behavior is recovered, without any significant loss of accuracy.

Appendix A. The Riemann Problem

We present in this appendix the exact solution of the following Riemann problem:

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(ku(1-u)) = 0 & x \in \mathbb{R}, t \in \mathbb{R}_+, \\ u(t=0, x) = \begin{cases} u_L & \text{if } x < 0 \\ u_R & \text{if } x > 0 \end{cases}, \quad k(x) = \begin{cases} k_L & \text{if } x < 0 \\ k_R & \text{if } x > 0 \end{cases}, \end{cases} \quad (\text{A.1})$$

where $k_L, k_R \in \mathbb{R}_+$ and $u_L, u_R \in [0; 1]$.

Appendix A.1. Properties of the solution of the Riemann problem

The Riemann problem (A.1) may be seen as two separated problems, $\{t \geq 0; x < 0\}$ and $\{t \geq 0; x > 0\}$, coupled by some interface conditions at $\{t \geq 0; x = 0\}$. Let us define $u^- = u(t > 0, x = 0^-)$ and $u^+ = u(t > 0, x = 0^+)$ ($u^-, u^+ \in [0; 1]$). The entropy solution u of (A.1) is self-similar (which implies that u^- and u^+ are constant), and verifies the following relations:

In $t \geq 0, x < 0$:

- u is the (unique) entropy solution of:

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(k_L u(1-u)) = 0 & t \in \mathbb{R}_+, x \in \mathbb{R}_-^*, \\ u(t=0, x) = u_L & x \in \mathbb{R}_-^*, \\ u(t, x=0^-) = u^- & t \in \mathbb{R}_+. \end{cases}$$

- If u contains a rarefaction wave, $u(t, x) \geq \frac{1}{2}$, $t \in \mathbb{R}_+$, $\forall x \in \mathbb{R}^*$.
- If u contains a shock wave, $u_L + u^- \geq 1$.

In $t \geq 0$, $x > 0$:

- u is the (unique) entropy solution of:

$$\begin{cases} \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(k_R u(1-u)) = 0 & t \in \mathbb{R}_+, x \in \mathbb{R}_+^*, \\ u(t=0, x) = u_R & x \in \mathbb{R}_+^*, \\ u(t, x=0^+) = u^+ & t \in \mathbb{R}_+. \end{cases}$$

- If u contains a rarefaction wave, $u(t, x) \leq \frac{1}{2}$, $t \in \mathbb{R}_+$, $\forall x \in \mathbb{R}_+^*$.
- If u contains a shock wave, $u_R + u^+ \leq 1$.

In $t \geq 0$, $x = 0$:

- $k_L u^-(1-u^-) = k_R u^+(1-u^+)$.
- Either $u^-, u^+ \leq \frac{1}{2}$ or $u^-, u^+ \geq \frac{1}{2}$ or $u^- < u^+$.

Thanks to all these properties, one may now describe explicitly the solution u of the whole Riemann problem (A.1).

Appendix A.2. The explicit form of the solution of the Riemann problem

We first present the construction of the solution when the permeability is constant $k(x) = k_0$. Let u_l and u_r be two different states in $[0; 1]$. One may link u_l and u_r by a rarefaction wave or by a shock wave.

If $u_l > u_r$: u_l and u_r are linked by a rarefaction wave defined by:

$$u(t, x) = \begin{cases} u_l & \text{if } x/t \leq k_0(1-2u_l), \\ \frac{k_0 - x/t}{2k_0} & \text{if } k_0(1-2u_l) < x/t < k_0(1-2u_r), \\ u_r & \text{if } x/t \geq k_0(1-2u_r). \end{cases} \quad (\text{A.2})$$

If $u_l < u_r$: u_l and u_r are linked by a shock wave defined by:

$$u(t, x) = \begin{cases} u_l & \text{if } x/t \leq k_0(1-(u_l+u_r)), \\ u_r & \text{if } x/t > k_0(1-(u_l+u_r)). \end{cases} \quad (\text{A.3})$$

So, the construction of the solution of Riemann problem (A.1) is reduced to the determination of u^- and u^+ . Indeed, since $k(x)$ is constant in $\{t \geq 0; x < 0\}$ and in $\{t \geq 0; x > 0\}$, the previous characterization gives the profile of u (by (A.2) or (A.3)). We first focus on the case $k_L > k_R$.

If $u_R < 1/2$:

- $0 \leq u_L < 1/2$ and $k_L g(u_L) \leq k_R g(u_R)$:
 - $u^- = u_L$,
 - u^+ is the smallest root of $k_L g(u_L) = k_R g(u^+)$, and u^+ and u_R are linked by a shock wave (defined by (A.3), with $u_l = u^+$ and $u_r = u_R$).
- $k_L g(u_L) \leq k_R g(u_R)$ and $k_L g(u_L) \leq k_R g(1/2)$:
 - $u^- = u_L$,
 - u^+ is the smallest root of $k_L g(u_L) = k_R g(u^+)$, and u^+ and u_R are linked by a rarefaction wave (defined by (A.2), with $u_l = u^+$ and $u_r = u_R$).
- $k_L g(u_L) > k_R g(1/2)$:
 - u^- is the greatest root of $k_L g(u^-) = k_R g(1/2)$, and u_L and u^- are linked by a shock wave (defined by (A.3), with $u_l = u_L$ and $u_r = u^-$),
 - $u^+ = 1/2$, and u^+ and u_R are linked by a rarefaction wave (defined by (A.2), with $u_l = u^+$ and $u_r = u_R$).
- $1/2 < u_L \leq 1$ and $k_L g(u_L) \leq k_R g(u_R)$:
 - u^- is the greatest root of $k_L g(u^-) = k_R g(1/2)$, and u_L and u^- are linked by a rarefaction wave (defined by (A.2), with $u_l = u_L$ and $u_r = u^-$),
 - $u^+ = 1/2$, and u^+ and u_R are linked by a rarefaction wave (defined by (A.2), with $u_l = u^+$ and $u_r = u_R$).

The previous cases are represented in Fig. A.1 (where CD denotes the contact discontinuity at the interface of the Riemann problem).

If $u_R > 1/2$:

- $0 \leq u_L < 1/2$ and $k_L g(u_L) \leq k_R g(u_R)$:
 - $u^- = u_L$,
 - u^+ is the smallest root of $k_L g(u_L) = k_R g(u^+)$, and u^+ and u_R are linked by a shock wave (defined by (A.3), with $u_l = u^+$ and $u_r = u_R$).
- $k_L g(u_L) > k_R g(u_R)$:
 - u^- is the greatest root of $k_L g(u^-) = k_R g(u_R)$, and u_L and u^- are linked by a shock wave (defined by (A.3), with $u_l = u_L$ and $u_r = u^-$),
 - $u^+ = u_R$.
- $1/2 < u_L \leq 1$ and $k_L g(u_L) \leq k_R g(u_R)$:
 - u^- is the greatest root of $k_L g(u^-) = k_R g(u_R)$, and u_L and u^- are linked by a rarefaction wave (defined by (A.2), with $u_l = u_L$ and $u_r = u^-$),
 - $u^+ = u_R$.

The previous cases are represented in Fig. A.2. The case where $u_R = 1/2$ may be directly deduced from the two previous cases.

Furthermore, if $k_L < k_R$, the construction of the solution u is the same as above. Indeed, setting $1 - u_L$ instead of u_L and $1 - u_R$ instead of u_R , and exchanging k_L and k_R , the solution may be constructed in the same way.

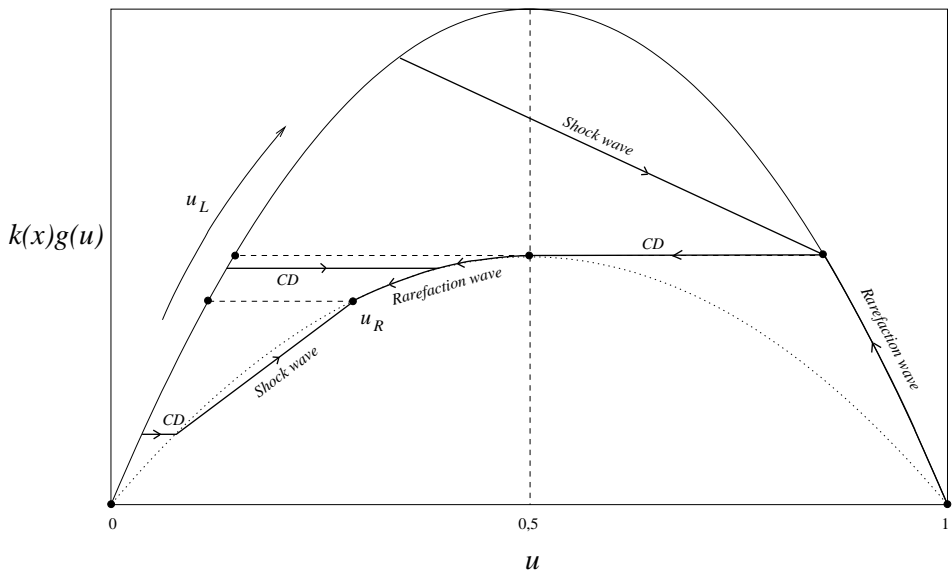


Fig. A.1. $k_L > k_R$ and $u_R < 1/2$.

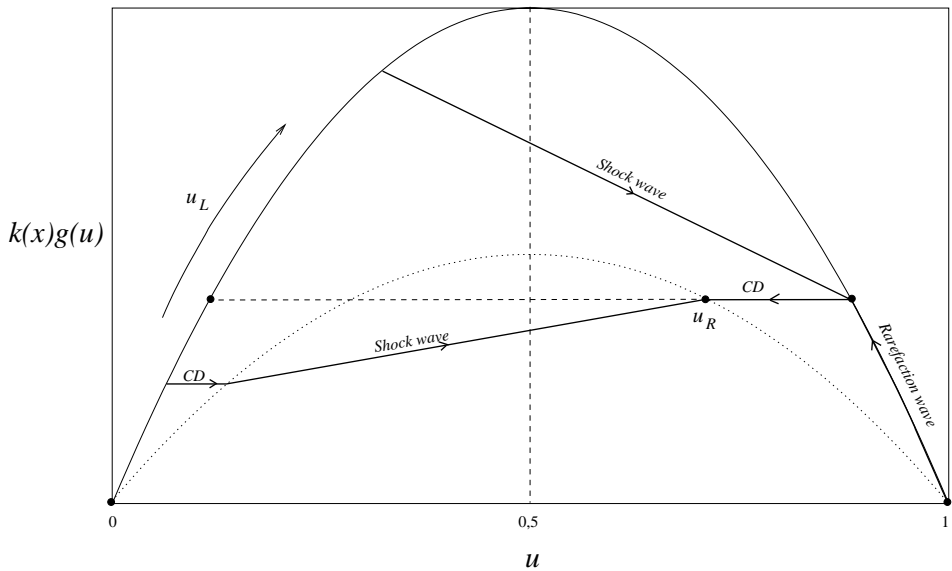


Fig. A.2. $k_L > k_R$ and $u_R > 1/2$.

Remark A.1. Let us emphasize on some interesting cases of the previous description. Assume that $k_L > k_R$. Choosing u_L such that $k_L g(u_L) > k_R g(1/2)$, one cannot find a state u_R separated from u_L just by a simple wave (namely a contact discontinuity). Hence, another wave (a shock wave or a rarefaction wave) must be introduced to allow the construction of the solution. For classical non-resonant

systems (Euler equations for instance), such a problem does not occur. Now, we focus on the two last cases described when $u_R < 1/2$. The profile of u (see Figs. 5 and B.2) seems to be composed of three different waves. In fact, two of them compose only one wave, and the constant state between them belongs to the same wave too. The third corresponds to the discontinuity of k .

Appendix B. Approximation of the Resonance Phenomenon

Here, we present two interesting tests computed using the Godunov scheme (4.27) with the higher order extension on a domain of 10 m divided in 500 cells.

Appendix B.1. A wave reflecting on a discontinuity of the permeability

The initial condition for this test is

$$k(x) = \begin{cases} 1 & \text{if } 0 < x < 7.5 \\ 2 & \text{if } 7.5 < x < 10 \end{cases} \quad \text{and} \quad u_0(x) = \begin{cases} 0.45 & \text{if } 0 < x < 2.5, \\ 0.3 & \text{if } 2.5 < x < 7.5, \\ 0.5 + \frac{\sqrt{2.32}}{4} & \text{if } 7.5 < x < 10. \end{cases}$$

The discontinuity of u_0 at $x = 2.5$ m generates a rarefaction wave moving to the right while the discontinuity in $x = 7.5$ m is in agreement with conditions (2.12) and (2.13). At time $t = 12$ s, the rarefaction wave starts to imping on the discontinuity in $x = 7.5$ m and for $t > 12$ s, it is reflected and becomes a shock wave, see Fig. B.1.

Appendix B.2. A bifurcation test case

Here, the initial condition is

$$k(x) = \begin{cases} 2 & \text{if } 0 < x < 6 \\ 1 & \text{if } 6 < x < 10 \end{cases} \quad \text{and} \quad u_0(x) = \begin{cases} 0.4 & \text{if } 0 < x < 2, \\ 0.13 & \text{if } 2 < x < 6, \\ 0.5 - \frac{\sqrt{0.0952}}{2} & \text{if } 6 < x < 10. \end{cases}$$

It generates a rarefaction wave at $x = 2$ m which moves to the right while the discontinuity in $x = 6$ m is in agreement with conditions (2.12) and (2.13). When the rarefaction wave meets the other discontinuity, it is separated into “two” waves: a shock wave which is reflected towards the left and a rarefaction wave bounded by the speeds 0 and $\sqrt{0.0952}$ (see Fig. B.2). Note that these “two” waves correspond to the same eigenvalue, namely $k(1 - 2u)$. It is called a “bifurcation” phenomenon.

Appendix C. BV Estimates

Lemma C.1. *Suppose $u_0 \in L^\infty \cap BV(\mathbb{R})$ and $0 \leq u_0 \leq 1$ a.e. on \mathbb{R} . Then the solution u_ε of Problem (2.6) satisfies the following maximum principle*

$$0 \leq u_\varepsilon(t, x) \leq 1 \quad \text{for a.e. } (t, x) \in \mathbb{R}_+ \times \mathbb{R} \quad (\text{C.1})$$

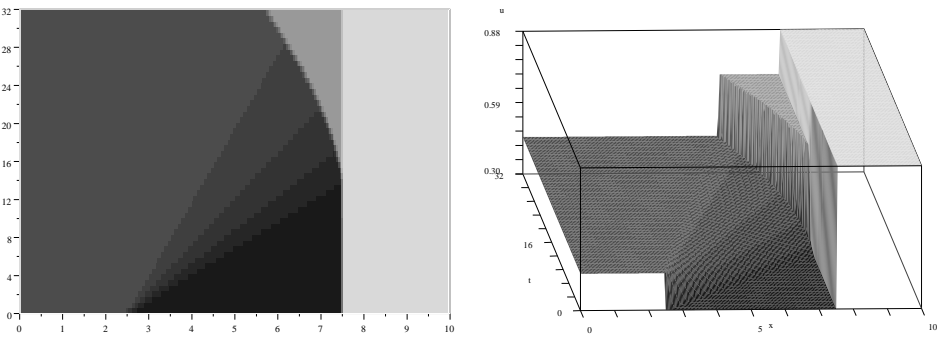


Fig. B.1. A rarefaction wave reflecting on a discontinuity of the permeability.

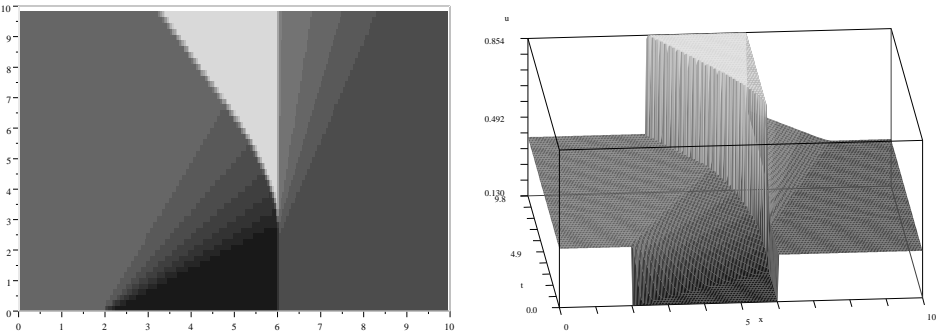


Fig. B.2. A rarefaction wave dividing in two waves.

as well as the following BV estimate: for any $T > 0$, for any $\kappa \in [0, 1]$, there exists a constant $C > 0$, depending on T , k_L , k_R and not on ε such that

$$|\Phi(u_\varepsilon, \kappa)|_{BV((0,T) \times \mathbb{R})} \leq C(|u_0|_{BV(\mathbb{R})} + |k_\varepsilon|_{BV(\mathbb{R})}). \tag{C.2}$$

Here, we suppose that u_0 is smooth with compact support: $u_0 \in \mathcal{C}_c^\infty(\mathbb{R})$, and that the bound $0 \leq u_0 \leq 1$ still holds. Let v^μ denotes the solution of the viscous approximation of Problem (2.6), that is

$$v_t + (k_\varepsilon(x)g(v))_x - \mu v_{xx} = 0, \tag{C.3}$$

with initial condition u_0 . Then, as (C.3) is a parabolic equation, the solution v^μ is smooth. We state several results put together in the following lemma.

Lemma C.2. (i) Let w^μ be another smooth solution of Eq. (C.3) with initial condition w_0 , such that $g(w^\mu(t, \pm\infty)) = 0$. Then the following result of comparison holds

$$\int_{\mathbb{R}} (v^\mu(t, x) - w^\mu(t, x))^+ dx \leq \int_{\mathbb{R}} (u_0(x) - w_0(x))^+ dx \quad \text{for every } t \geq 0. \tag{C.4}$$

- (ii) As the initial condition u_0 , the solution v^μ satisfies $0 \leq v^\mu \leq 1$.
 (iii) For every $T > 0$, $R > 0$, there exists a constant $C_{T,R}$ depending only on T and R such that

$$\mu \int_0^T \int_{-R}^R |v_x^\mu|^2 dx dt \leq C_{T,R}. \quad (\text{C.5})$$

Proof of Lemma C.2. For the proof of the first point see Dafermos,³ pp. 92–93. For the sake of completeness, we give me details:

Let η_α denote the smooth approximation of the function $v \mapsto v^+$ defined by

$$\eta_\alpha(v) = \begin{cases} 0 & \text{if } v \leq 0, \\ v^2/4\alpha & \text{if } 0 \leq v \leq 2\alpha, \\ v - \alpha & \text{if } 2\alpha \leq v. \end{cases}$$

Multiplying the equation

$$(v^\mu - w^\mu)_t + (k_\varepsilon g(v^\mu) - k_\varepsilon g(w^\mu))_x = \mu(v^\mu - w^\mu)_{xx}$$

by $\eta'_\alpha(v^\mu - w^\mu)$ and noting

$$\begin{aligned} A^\mu &= \partial_t \eta_\alpha(v^\mu - w^\mu) + \partial_x (\eta'_\alpha(v^\mu - w^\mu) [k_\varepsilon g(v^\mu) - k_\varepsilon g(w^\mu)]) \\ &\quad - \eta''_\alpha(v^\mu - w^\mu) [k_\varepsilon g(v^\mu) - k_\varepsilon g(w^\mu)] \partial_x (v^\mu - w^\mu) \end{aligned}$$

we get

$$A^\mu = \mu \partial_{xx} \eta_\alpha(v^\mu - w^\mu) - \mu \eta''_\alpha(v^\mu - w^\mu) [\partial_x (v^\mu - w^\mu)]^2 \leq \mu \partial_{xx} \eta_\alpha(v^\mu - w^\mu).$$

Integrating this last inequality on $(0, t) \times \mathbb{R}$ yields

$$\begin{aligned} &\int_{\mathbb{R}} \eta_\alpha(v^\mu(t) - w^\mu(t)) dx - \int_{\mathbb{R}} \eta_\alpha(u_0(t) - w_0(t)) dx \\ &\leq \int_0^t \int_{\mathbb{R}} \eta''_\alpha(v^\mu - w^\mu) [k_\varepsilon g(v^\mu) - k_\varepsilon g(w^\mu)] \partial_x (v^\mu - w^\mu) dx dt \\ &\leq \text{Lip}(g) \int_0^t \int_{\mathbb{R}} \eta''_\alpha(v^\mu - w^\mu) k_\varepsilon |v^\mu - w^\mu| \cdot |\partial_x (v^\mu - w^\mu)| dx dt. \end{aligned}$$

Notice that $g(v^\mu(t, \pm\infty)) = 0$ because $g(0) = 0$ and $v(t, \cdot)$ decreases rapidly to zero when $x \rightarrow \pm\infty$ (for the initial condition u_0 has a compact support), while $g(w^\mu(t, \pm\infty)) = 0$ by hypothesis. Eventually, letting α tend to zero yields (C.4).

We use this result of comparison to prove the L^∞ estimate on v^μ . As $g(1) = 0$, the constant function 1 is a solution to (C.3) with initial condition 1, therefore, one has

$$\int_{\mathbb{R}} (v^\mu(t, x) - 1)^+ dx \leq \int_{\mathbb{R}} (u_0(x) - 1)^+ dx \quad \text{for every } t \geq 0.$$

As $0 \leq u_0 \leq 1$ by hypothesis, there holds $(u_0 - 1)^+ = 0$ and $v^\mu(t, x) \leq 1$. Considering the function $s \mapsto s^-$ instead of $s \mapsto s^+$ and choosing $w^\mu = 0$ would give, in the same way, $v^\mu \geq 0$. Then the function v^μ satisfies the maximum principle (C.1).

Once an L^∞ estimate on the function v^μ is available, it is not really difficult to prove the energy estimate (C.5) (multiply Eq. (C.3) by v^μ and integrate by parts).

Now we turn to the BV -estimate on $\Phi(v^\mu, \kappa)$ and, for this purpose, we first give a bound on the L^1 -norm of v_t^μ . For h a positive number, the function $w^\mu(t, x) = v^\mu(t + h, x)$ is a solution of Eq. (C.3) with initial condition $v^\mu(h, \cdot)$. Using the result of comparison (C.4) with $s \mapsto |s|$ instead of $s \mapsto s^+$ (it is still true), we get

$$\int_{\mathbb{R}} |v^\mu(t + h, x) - v^\mu(t, x)| dx \leq \int_{\mathbb{R}} |v^\mu(h, x) - v^\mu(0, x)| dx \quad \text{for every } t \geq 0.$$

Dividing the result by h and letting h tend to $0+$, we get

$$\int_{\mathbb{R}} |v_t^\mu(t, x)| dx \leq \int_{\mathbb{R}} |v_t^\mu(0, x)| dx \quad \text{for every } t \geq 0.$$

Besides, $v_t^\mu(0, x) = -k'_\varepsilon(x)g(u_0(x)) - k_\varepsilon(x)g'(u_0(x))u'_0(x) + \mu u''_0(x)$, and, since $\sup\{|g(u)|; u \in [0, 1]\} = 1/4$, $\sup\{|g'(u)|; u \in [0, 1]\} = 1$, and $|k_\varepsilon|_{BV(\mathbb{R})} = \int_{\mathbb{R}} |k'_\varepsilon(x)| dx$, we have

$$\int_{\mathbb{R}} |v_t^\mu(t, x)| dx \leq \frac{1}{4}|k_\varepsilon|_{BV(\mathbb{R})} + \max\{k_L, k_R\}|u_0|_{BV(\mathbb{R})} + \mu \int_{\mathbb{R}} |u''_0(x)| dx. \quad (\text{C.6})$$

Let $\kappa \in [0, 1]$. Multiplying Eq. (C.3) by $\text{sgn}(v^\mu - \kappa)$ yields $\partial_x(k_\varepsilon \Phi(v^\mu, \kappa)) \leq S_1^\mu + S_2^\mu + S^\mu$ in $\mathcal{D}'((0, T) \times \mathbb{R})$, with

$$S_1^\mu = -\partial_t |v^\mu - \kappa|, \quad S_2^\mu = k'_\varepsilon \text{sgn}(v^\mu - \kappa)g(\kappa) \quad \text{and} \quad S^\mu = \mu \partial_{xx} |v^\mu - \kappa|.$$

We evaluate each distribution S^μ against a test-function φ of $\mathcal{C}_c^\infty((0, T) \times \mathbb{R})$ such that $0 \leq \varphi \leq 1$. From the estimate on the L^1 -norm of v_t^μ , we deduce

$$\langle S_1^\mu, \varphi \rangle \leq T \left(\frac{1}{4}|k_\varepsilon|_{BV(\mathbb{R})} + \max\{k_L, k_R\}|u_0|_{BV(\mathbb{R})} + \mu \int_{\mathbb{R}} |u''_0(x)| dx \right). \quad (\text{C.7})$$

Moreover, we have

$$\langle S_2^\mu, \varphi \rangle \leq \frac{T}{4}|k_\varepsilon|_{BV(\mathbb{R})}, \quad (\text{C.8})$$

and, from the energy estimate (C.5) and from the Cauchy-Schwarz inequality, we deduce

$$\langle S^\mu, \varphi \rangle \leq C_\varphi \|\varphi\|_{L^2(\mathbb{R})} \sqrt{\mu}, \quad (\text{C.9})$$

where C_φ depends on the support of φ .

Now, it is known that $\lim_{\mu \rightarrow 0} v^\mu = u_\varepsilon$ in $L^1_{\text{loc}}((0, +\infty) \times \mathbb{R})$. Therefore, each of the preceding distributions converges in $\mathcal{D}'((0, T) \times \mathbb{R})$. From (C.9), we deduce

$S^\mu \rightarrow 0$ so that there holds $\partial_x(k_\varepsilon \Phi(u_\varepsilon, \kappa)) \leq S_1 + S_2$ with

$$\langle S_1, \varphi \rangle \leq T \left(\frac{1}{4} |k_\varepsilon|_{BV(\mathbb{R})} + \max\{k_L, k_R\} |u_0|_{BV(\mathbb{R})} \right),$$

$$\langle S_2, \varphi \rangle \leq \frac{T}{4} |k_\varepsilon|_{BV(\mathbb{R})},$$

for every φ in $\mathcal{C}_c^\infty((0, T) \times \mathbb{R})$ such that $0 \leq \varphi \leq 1$. Therefore $k_\varepsilon \Phi(u_\varepsilon, \kappa) \in BV((0, T) \times \mathbb{R})$ and

$$\begin{aligned} |k_\varepsilon \Phi(u_\varepsilon, \kappa)|_{BV((0, T) \times \mathbb{R})} &\leq 2 \sup\{\langle \partial_x(k_\varepsilon \Phi(u_\varepsilon, \kappa)), \varphi \rangle; \varphi \in \mathcal{C}_c^\infty((0, T) \times \mathbb{R}; [0, 1])\} \\ &\leq T(|k_\varepsilon|_{BV(\mathbb{R})} + 2 \max\{k_L, k_R\} |u_0|_{BV(\mathbb{R})}). \end{aligned}$$

Writing $\partial_x(k_\varepsilon \Phi(u_\varepsilon, \kappa)) = k_\varepsilon \partial_x(\Phi(u_\varepsilon, \kappa)) + k'_\varepsilon \Phi(u_\varepsilon, \kappa)$ and using the estimates $|k'_\varepsilon \Phi(u_\varepsilon, \kappa)|_{L^1((0, T) \times \mathbb{R})} \leq T |k_\varepsilon|_{BV(\mathbb{R})}$ and $k_\varepsilon \geq \min\{k_L, k_R\}$, we get $\Phi(u_\varepsilon, \kappa) \in BV((0, T) \times \mathbb{R})$ and

$$|\Phi(u_\varepsilon, \kappa)|_{BV((0, T) \times \mathbb{R})} \leq \frac{2T}{\min\{k_L, k_R\}} (|k_\varepsilon|_{BV(\mathbb{R})} + \max\{k_L, k_R\} |u_0|_{BV(\mathbb{R})}).$$

By classical results of approximation, this is still true if $u_0 \in BV(\mathbb{R})$ and $0 \leq u_0 \leq 1$ a.e. This ends the proof of Lemma C.1.

Acknowledgments

N.S. was supported by EDF grant under contract C02770/AEE2704. Computational facilities were provided by EDF. Both authors would like to thank Thierry Gallouët and Jean-Marc Hérard for pointing this problem out to investigate and for very fruitful discussions.

References

1. T. Buffard, T. Gallouët and J.-M. Hérard, *A sequel to a rough Godunov scheme: Application to real gases*, *Comput. Fluids* **29** (2000) 813–847.
2. F. Bouchut and F. James, *One-dimensional transport equations with discontinuous coefficients*, *Nonlinear Anal.* **32** (1998) 891–933.
3. C. M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics* (Springer-Verlag, 2000).
4. S. Diehl, *On scalar conservation laws with point source and discontinuous flux function*, *SIAM J. Math. Anal.* **26** (1995) 1425–1451.
5. G. Dal Maso, P. G. Lefloch and F. Murat, *Definition and weak stability of nonconservative products*, *J. Math. Pures Appl.* **74** (1995) 483–548.
6. R. Eymard, T. Gallouët and R. Herbin, *Finite Volume Methods, Handbook of Numerical Analysis* (North-Holland, 2000), Vol. VII, pp. 713–1020.
7. T. Gallouët, J.-M. Hérard and N. Seguin, *Some approximate Godunov schemes to compute shallow-water equations with topography*, *Comput. Fluids* **32** (2003) 479–513.
8. T. Gallouët, J.-M. Hérard and N. Seguin, *Some recent finite volume method to compute Euler equations using real gas EOS*, *Int. J. Numer. Methods Fluids* **39** (2002) 1073–1138.

9. T. Gallouët, J.-M. Hérard and N. Seguin, *Numerical modelling of two-phase flows using the two-fluid two-pressure approach*, EDF-DRD Report HI-81/01/043/A (2001).
10. E. Giusti, *Minimal Surfaces and Functions of Bounded Variation* (Birkhäuser, 1984).
11. J. M. Greenberg and A. Y. Leroux, *A well-balanced scheme for the numerical processing of source terms in hyperbolic equations*, *SIAM J. Numer. Anal.* **33** (1996) 1–16.
12. S. K. Godunov, *A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics*, *Mat. Sb.* **47** (1959) 271–306.
13. T. Gimse and N. H. Risebro, *Riemann problems with a discontinuous flux function*, in *Third Int. Conf. Hyperbolic Problems* (Uppsala, 1990), Studentlitteratur, Lund, 1991, Vols. I, II, pp. 488–502.
14. E. Godlewski and P.-A. Raviart, *Numerical Approximation of Hyperbolic Systems of Conservation Laws* (Springer-Verlag, 1996).
15. E. Godlewski and P.-A. Raviart, *The numerical coupling of nonlinear hyperbolic systems of conservation laws: I. The scalar case*, preprint, 2001.
16. A. Harten and J. M. Hyman, *Self-adjusting grid methods for one-dimensional hyperbolic conservation laws*, *J. Comput. Phys.* **50** (1983) 235–269.
17. E. Isaacson, D. Marchesin, B. Plohr and B. Temple, *The Riemann problem near a hyperbolic singularity: The classification of solutions of quadratic Riemann problems. I*, *SIAM J. Appl. Math.* **48** (1988) 1009–1032.
18. E. Isaacson and B. Temple, *Analysis of a singular hyperbolic system of conservation laws*, *J. Differential Equations* **65** (1986) 250–268.
19. E. Isaacson and B. Temple, *The Riemann problem near a hyperbolic singularity. II, III*, *SIAM J. Appl. Math.* **48** (1988) 1287–1301; 1302–1318.
20. E. Isaacson and B. Temple, *Nonlinear resonance in inhomogeneous systems of conservation laws*, in *Mathematics of Nonlinear Science* (Phoenix, AZ, 1989) (Amer. Math. Soc., 1990), pp. 63–77.
21. R. A. Klausen and N. H. Risebro, *Stability of conservation laws with discontinuous coefficients*, *J. Differential Equations* **157** (1999) 41–60.
22. C. Klingenberg and N. H. Risebro, *Convex conservation laws with discontinuous coefficients. Existence, uniqueness and asymptotic behavior*, *Comm. Partial Differential Equations* **20** (1995) 1959–1990.
23. C. Klingenberg and N. H. Risebro, *Stability of a resonant system of conservation laws modeling polymer flow with gravitation*, *J. Differential Equations* **170** (2001) 344–380.
24. K. H. Karlsen, N. H. Risebro and J. D. Towers, *On a nonlinear degenerate parabolic transport–diffusion equation with a discontinuous coefficient*, No. 13.
25. S. N. Kružkov, *First-order quasilinear equations with several independent variables*, *Mat. Sb.* **81** (1970) 228–255.
26. A. Y. Le Roux, *Riemann solvers for some hyperbolic problems with a source term*, *Actes du 30ème Congrès d'Analyse Numérique: CANum '98* (Arles, 1998) (Soc. Math. Appl. Indust., 1999), pp. 75–90 (electronic).
27. L. Lin, J. B. Temple and J. Wang, *A comparison of convergence rates for Godunov's method and Glimm's method in resonant nonlinear systems of conservation laws*, *SIAM J. Numer. Anal.* **32** (1995) 824–840.
28. L. W. Lin, B. Temple and J. H. Wang, *Suppression of oscillations in Godunov's method for a resonant non-strictly hyperbolic system*, *SIAM J. Numer. Anal.* **32** (1995) 841–864.
29. N. Seguin, *Génération et validation de Rozavel, un code équilibre en hydraulique 2D*, mémoire de D.E.A., Université de Bordeaux I, June 1999, in French.
30. B. Temple, *Global solution of the Cauchy problem for a class of 2×2 nonstrictly*

- hyperbolic conservation laws, *Adv. Appl. Math.* **3** (1982) 335–375.
31. J. D. Towers, *Convergence of a difference scheme for conservation laws with a discontinuous flux*, *SIAM J. Numer. Anal.* **38** (2000) 681–698 (electronic).
32. J. D. Towers, *A difference scheme for conservation laws with a discontinuous flux: The nonconvex case*, *SIAM J. Numer. Anal.* **39** (2001) 1197–1218 (electronic).
33. A. Vasseur, *Strong traces for solutions of multidimensional scalar conservation laws*, *Arch. Rational Mech. Anal.* **160** (2001) 181–193.
34. B. van Leer, *Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method*, *J. Comput. Phys.* **32** (1979) 101–136.
35. A. I. Vol'pert, *Spaces b_v and quasilinear equations*, *Mat. Sb.* **73** (1967) 255–302.