A kinetic scheme for the Saint-Venant system with a source term.

Benoît Perthame, Chiara Simeoni

A kinetic scheme for the Saint-Venant system with a source term

B. Perthame, C. Simeoni

Département de Mathématiques et Applications
École Normale Supérieure
45, rue d'Ulm - 75230 Paris Cedex 05 - France
e-mails: Benoit.Perthame@ens.fr, Chiara.Simeoni@ens.fr

Abstract

The aim of this paper is to present a numerical scheme to compute Saint-Venant equations with a source term, due to the bottom topography, in a one-dimensional framework, which satisfies the following theoretical properties: it preserves the steady state of still water, satisfies an entropy inequality, preserves the non-negativity of the height of water and remains stable with a discontinuous bottom. This is achieved by means of a kinetic approach to the system, which is the departing point of the method developed here. In this context, we use a natural description of the microscopic behaviour of the system to define numerical fluxes at the interfaces of an unstructured mesh. We also use the concept of cell-centered conservative quantities (as usual in the finite volume method) and upwind interfacial sources as advocated by several authors. We show, analytically and also by means of numerical results, that the above properties are satisfied.

Key-words: Saint-Venant system, finite volume method, upwind interfacial sources, kinetic schemes.

1 Introduction

The Saint-Venant equations, a particular case of shallow water equations, are commonly used to describe physical situations such as flows in rivers or coastal areas. The one-dimensional version is well adapted for ideal rectangular rivers. It allows to describe the flow, at time $t \geq 0$ and at point $x \in \mathbb{R}$,

through the height of water $h(t,x) \geq 0$ and its velocity $u(t,x) \in \mathbb{R}$, by the hyperbolic system

$$\frac{\partial h}{\partial t} + \frac{\partial (hu)}{\partial x} = 0, \tag{1.1}$$

$$\frac{\partial (hu)}{\partial t} + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) + gh\frac{\partial Z}{\partial x} = 0, \tag{1.2}$$

where g denotes the gravity intensity and $Z(x)$ is the bottom height; therefore $h + Z$ is the level of the water surface (in what follows, we also denote the discharge by $q = hu$).

These equations were originally written by A. de Saint-Venant in [21] and more complete systems can be derived from the Navier-Stokes equations (see [8] and its references). In fact, the system (1.1)-(1.2) corresponds to a particularly simple case: other terms can be added to the right-hand side in order to take into account frictions on the bottom and the surface; a more general system can also be stated for rivers with variable sections.

The bottom topography introduces a source term in equation (1.2), influencing the unknowns of the problem. Hence, analytical properties of the system of isentropic Euler equations are deeply modified in comparison with the homogeneous model. For instance, a well-known problem is the occurrence of different kinds of steady states.

Several methods for solving hyperbolic systems of conservation laws with source terms have been investigated. The main problem is related to the approximation of such a source term, to assure the numerical preservation of properties fulfilled by the continuous model. A classical approach consists in using finite volume schemes (refer to [9], [16] and [6]), the finite volume method displaying the remarkable property of being water height conservative. But of course other methods are possible, such as finite elements (see [2] and the references therein).

A difficulty arising specifically with the Saint-Venant system is that of forcing the scheme to preserve steady states given by a lake at rest ($u = 0$, $h + Z = C^{st}$), as has been pointed out by several authors. To treat these particular problems, a specific numerical approach is needed. Here it is based on two concepts: first, the conservative quantities are cell-centered, as usual for finite volume schemes; second, as introduced by Roe [20], the source terms are upwinded at the cell interfaces. Initially for scalar conservation laws, Greenberg and LeRoux [14], then Gosse and LeRoux [10],[11] introduced the notion of well-balanced scheme and they use a reformulation of the source terms by means of non-conservative products to derive their numerical fluxes; this kind of numerical processing has been recently extended by Gosse [12],[13] to hyperbolic systems of balance laws. A kinetic scheme, which introduces

the notion of reflexions on bottom jumps and maintains steady states, is presented in Botchorishvili, Perthame and Vasseur [4] , which is proved to be convergent. Still using the interface values, instead of the cell-averages, for the source terms, Jin proposes in [15] a rather simple method for capturing the steady state solutions with a high order accuracy. Quite recently, various approaches appeared to build stable schemes which preserve the steady states: previous schemes have been modified for this purpose by Bermudez and Vasquez [3]; the Godunov scheme for an appropriately extended system is developed in [17]; an appropriate linearized Godunov scheme preserving all steady states has also been obtained by Gallouët, Hérard and Seguin [7].

But to our knowledge none of these methods are proved to satisfy all the stability properties: (i) water height remains nonnegative, (ii) the energy (entropy) inequality is satisfied, (iii) it preserves the steady states of still water. Certainly, Godunov schemes as modified in [14] or [7] can do that, at the expense of a fixed point, but this has not been proved.

In this paper, we consider a particular class of numerical schemes to compute the Saint-Venant equations, based on the kinetic interpretation of the system, which is presented in [1]. These kinetic schemes have many good properties that other solvers have difficulty in achieving; in particular, they are able to treat the case of a vacuum ($h = 0$ here, corresponding to dry soils, when the system loses hyperbolicity) and satisfy the properties (i), (ii), (iii) above. We refer to Perthame [19] for a survey of the theoretical properties of these schemes. We only note that kinetic schemes are a simple way to generate efficient building blocks (interpolations at the interfaces) in finite volume methods; they do not involve rarefied flows except for the technique used in proving their theoretical properties.
We present a numerical scheme for the system (1.1)-(1.2) by using a new kinetic solver, which exhibits all the advantages of this specific approach. We propose a way to take the source term directly into account in the definition of the numerical fluxes, whose structure relies on a natural description of the microscopic behaviour of the system through a potential barrier which is the bottom $Z(x)$.

The paper is organized in four sections. In the second section, we recall some properties of the Saint-Venant equations and we explain the kinetic approach to this system. In the third, we illustrate the kinetic scheme "with reflexions" and we demonstrate the properties (i), (ii), (iii). Several test problems for the flat and non-flat bottom cases are reported in the last section. We leave the extension of this method to higher order accuracy for a future work. Implementation in two spatial dimensions is also in progress (see [1] for a first attempt).

2 Preliminaries about the Saint-Venant equations

We recall here some well-known properties of the shallow water system. We take them into account to develop our numerical method so as to be coherent with the physical model. Then, by analogy with the Euler equations of compressible gas dynamics, we link the macroscopic Saint-Venant system to a microscopic description of the fluid, on which the method proposed in this paper is based.

2.1 Properties of the system

First of all, the system is naturally posed for $h(t,x) \geq 0$ and the water height h can indeed vanish (flooding zones, dry soils, tidal flats); this fact leads to a theoretical and numerical difficulty, because the system loses hyperbolicity at $h = 0$.

Another fundamental property is related to the *entropy inequality* of the Saint-Venant system, satisfied by the weak solutions, defined as in the following theorem.

Theorem 2.1. *The system (1.1)-(1.2) is strictly hyperbolic for $h > 0$. It admits a mathematical entropy, which is also the physical energy,*

$$E(h,u,Z) = h\frac{u^2}{2} + \frac{g}{2}h^2 + gZh, \tag{2.1}$$

which satisfies the "entropy inequality"

$$\frac{\partial E}{\partial t} + \frac{\partial}{\partial x}[u(E + \frac{g}{2}h^2)] \leq 0. \tag{2.2}$$

We do not prove this theorem, which relies on the classical theory of hyperbolic equations (see Serre [22], Dafermos [5]) and simple algebraic calculations. We remark only that for smooth solutions the inequality (2.2) becomes an equality.

Also, the system admits a family of smooth steady states characterized by the relations

$$hu = C_1, \tag{2.3}$$

$$\frac{u^2}{2} + g(h + Z) = C_2, \tag{2.4}$$

where C_1 and C_2 are two arbitrary constants. In particular, the simplest is the traditional steady state of a lake at rest, given by $u = 0$, $h + Z = C^{st}$.

2.2 Kinetic approach

We pass to explain how it is possible to introduce a kinetic approach for the Saint-Venant system. We describe it for the one-dimensional problem, but the construction is similar in two dimensions.

We consider a real function χ defined on $\mathbb{R}$, with the following properties,

$$\chi(\omega) = \chi(-\omega) \geq 0, \quad \int_{\mathbb{R}} \chi(\omega) d\omega = 1, \quad \int_{\mathbb{R}} \omega^2 \chi(\omega) d\omega = \frac{g}{2} \tag{2.5}$$

and define the density of particles $M(t, x, \xi)$ by a so-called *Gibbs equilibrium*

$$M(t, x, \xi) = M(h, \xi - u) = \sqrt{h(t,x)} \chi \left(\frac{\xi - u(t,x)}{\sqrt{h(t,x)}} \right). \tag{2.6}$$

These definitions allow to obtain a kinetic representation of the system.

Theorem 2.2. *The pair of functions (h, hu) is a strong solution of the Saint-Venant system (1.1)-(1.2) if and only if $M(h, \xi - u)$ satisfies the kinetic equation*

$$\frac{\partial M}{\partial t} + \xi \cdot \frac{\partial M}{\partial x} - g \frac{\partial Z}{\partial x} \cdot \frac{\partial M}{\partial \xi} = Q(t, x, \xi), \tag{2.7}$$

for some "collision term" $Q(t, x, \xi)$ which satisfies, for a.e. (t, x),

$$\int_{\mathbb{R}} Q d\xi = 0, \qquad \int_{\mathbb{R}} \xi Q d\xi = 0. \tag{2.8}$$

Proof. The proof relies on a very obvious computation. The two Saint-Venant equations are obtained by taking the moments of the kinetic equation (2.7) in $d\xi$, against 1, ξ and ξ^2 respectively: the right-hand side vanishes according to (2.8) and the left-hand sides coincide exactly thanks to hypothesis (2.5). These are consequences of the easy relations,

$$h = \int_{\mathbb{R}} M(h, \xi - u) d\xi, \tag{2.9}$$

$$hu = \int_{\mathbb{R}} \xi M(h, \xi - u) d\xi, \tag{2.10}$$

$$hu^2 + \frac{g}{2} h^2 = \int_{\mathbb{R}} \xi^2 M(h, \xi - u) d\xi, \tag{2.11}$$

directly obtained from the microscopic equilibrium (2.6). □

This theorem produces a very useful consequence: the non-linear shallow water system can be viewed as a single linear equation on a non-linear quantity M, for which it is easier to find simple numerical schemes with good theoretical properties.
We note that this form is much weaker than the *kinetic formulation* proposed by Lions, Perthame and Tadmor in [18], which represents all the entropies of the system.

We characterize the function χ which defines the density of particles $M(t,x,\xi)$ in the kinetic approach; in particular, we justify the interpretation of such a density as the microscopic equilibrium of the system, the *Gibbs equilibrium*. These facts are stated in the following propositions.

Lemma 2.3. *The minimum of the energy*

$$\mathcal{E}(f) = \int_{\mathbb{R}} \left[\frac{\xi^2}{2} f(\xi) + \frac{\pi^2 g^2}{6} f^3(\xi) + gZf(\xi) \right] d\xi, \tag{2.12}$$

under the constraints

$$f \geq 0, \quad \int_{\mathbb{R}} f(\xi) d\xi = h, \quad \int_{\mathbb{R}} \xi f(\xi) d\xi = hu,$$

is attained by the function $M(h, \xi - u) = \sqrt{h}\chi\left(\frac{\xi-u}{\sqrt{h}}\right)$, *with* χ *defined by*

$$\chi(\omega) = \frac{\sqrt{2}}{\pi\sqrt{g}} \left(1 - \frac{\omega^2}{2g}\right)_+^{\frac{1}{2}}. \tag{2.13}$$

Remark 2.4. The cubic term in the functional (2.12) takes the internal energy into account. In one dimension, it results from the transverse translational energy. Indeed the corresponding two-dimensional variational problem, for $Z = 0$, gives

$$\frac{1}{2}h(u^2+v^2) + \frac{g}{2}h^2 = \min \left\{ \int_{\mathbb{R}^2} \frac{|\xi|^2}{2} f(\xi) d\xi; \right.$$
$$\left. f \geq 0, \int_{\mathbb{R}^2} f(\xi) d\xi = h, \int_{\mathbb{R}^2} \xi f(\xi) d\xi = (hu, hv) \right\}$$

and we deduce $f(\xi_1) = \int_{\mathbb{R}} g(\xi_1, \xi_2)\ d\xi_2$.

Proof. Because of the constraints, it is sufficient to minimize the functional

$$\mathcal{E}_0(f) = \int_{\mathbb{R}} \left[\xi^2 f(\xi) + \frac{\pi^2 g^2}{3} f^3(\xi) \right] d\xi.$$

Since $\mathcal{E}_0(f)$ is a convex functional, the formula for M (and thus for χ) follows directly from the Euler-Lagrange equation associated to the minimization problem, for $f > 0$,

$$\xi^2 + \pi^2 g^2 f^2 = \lambda + \mu\xi,$$

where $\lambda(h,u)$ and $\mu(h,u)$ are Lagrange multipliers. One readily checks by convexity that it is a strict minimizer. □

Recalling the formula (2.1), we see that the minimum considered in Lemma 2.3 is given by

$$\mathcal{E}(M(h,\xi-u)) = E(h,u,Z),$$

again an immediate consequence of the relations stated in (2.9)-(2.11) and by the choice of the specific value $\frac{\pi^2 g^2}{6}$ in the energy. Hence, the properties of the function χ are consistent with the kinetic approach to the system, as introduced above.

We conclude this section by pointing out another motivation for the choice of χ in Lemma 2.3.

Lemma 2.5. *The function* $\chi(\omega) = \frac{\sqrt{2}}{\pi\sqrt{g}}\left(1 - \frac{\omega^2}{2g}\right)_+^{\frac{1}{2}}$ *is the only choice such that* $M(h,\xi-u) = \sqrt{h}\chi\left(\frac{\xi-u}{\sqrt{h}}\right)$ *satisfies the equation*

$$\xi \cdot \frac{\partial M}{\partial x} - g\frac{\partial Z}{\partial x} \cdot \frac{\partial M}{\partial \xi} = 0 \tag{2.14}$$

on all steady states given by a lake at rest,

$$u(t,x) = 0, \quad h(t,x) + Z(x) = H, \quad \forall t \geq 0.$$

Proof. Exploiting the hypotheses, we compute

$$\begin{aligned}
\frac{\partial M}{\partial x} &= \frac{1}{2\sqrt{h}}\frac{\partial h}{\partial x}\left[\chi\left(\frac{\xi}{\sqrt{h}}\right) - \frac{\xi}{\sqrt{h}}\chi'\left(\frac{\xi}{\sqrt{h}}\right)\right], \\
\frac{\partial Z}{\partial x} &= -\frac{\partial h}{\partial x}, \quad \frac{\partial M}{\partial \xi} = \chi'\left(\frac{\xi}{\sqrt{h}}\right),
\end{aligned}$$

so that the equation (2.14) becomes

$$\frac{\xi}{2\sqrt{h}}\frac{\partial h}{\partial x}\chi\left(\frac{\xi}{\sqrt{h}}\right) - \frac{\xi^2}{2h}\frac{\partial h}{\partial x}\chi'\left(\frac{\xi}{\sqrt{h}}\right) + g\frac{\partial h}{\partial x}\chi'\left(\frac{\xi}{\sqrt{h}}\right) = 0.$$

Let $\omega = \frac{\xi}{\sqrt{h}}$, the last relation can be rewritten as

$$\frac{1}{2}\frac{\partial h}{\partial x}\left[\omega\chi(\omega) + (2g - \omega^2)\chi'(\omega)\right] = 0.$$

A characterization for χ is therefore given by the equation

$$\omega\chi(\omega) + (2g - \omega^2)\chi'(\omega) = 0,$$

that admits, under the constraints (2.5), the unique solution

$$\chi(\omega) = \left(2g - \omega^2\right)_+^{\frac{1}{2}}.$$

□

Remark 2.6. *We note that the difficulty in preserving such steady states at the kinetic level might explain why the Maxwellian case does not work well (see Xu [24]).*

3 The kinetic scheme with reflections

We present a finite volume scheme for the one-dimensional Saint-Venant system, based on the kinetic approach described in Section 2, which has the property of preserving the steady state of a lake at rest. Also, it preserves the stability properties of the usual kinetic solvers and satisfies a precise in-cell entropy inequality.

3.1 The formulas

We consider a uniform mesh of $\mathbb{R}$, whose vertices are denoted x_i, $i \in \mathbb{Z}$. Let $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ be the control volume (cell), with $x_{i+\frac{1}{2}} = \frac{x_{i+1}+x_i}{2}$, and we denote the space-step by $\Delta x = length(C_i)$, so that $x_i = i\Delta x$, $i \in \mathbb{Z}$. We also consider a discretization in time by introducing a time-step Δt and we set $t_n = n\Delta t$, $n \in \mathbb{N}$.
If $Z(x)$ is the function describing the bottom height, its piecewise constant representation is given by $\bar{Z}(x) = Z_i \mathbb{1}_{C_i}(x)$, with $Z_i = \frac{1}{\Delta x}\int_{C_i} Z(x)dx$, for example.

We start from the microscopic equation (2.7), to perform a discretization directly on the density of particles

$$f_i^{n+1}(\xi) - M_i^n(\xi) + \frac{\Delta t}{\Delta x}\xi\left(M^-_{i+\frac{1}{2}}(\xi) - M^+_{i-\frac{1}{2}}(\xi)\right) = 0, \tag{3.1}$$

where the interface equilibrium densities $M^{\pm}_{i+\frac{1}{2}}$ are defined later. As usual, the "collision term" $Q(t,x,\xi)$ in the kinetic representation (2.7) of Saint-Venant equations, which relaxes the kinetic density to an equilibrium M, is neglected in the numerical scheme; at each time-step we project $f^n_i(\xi)$ on $M^n_i(\xi)$, which is a way of performing all collisions at once and to recover the *Gibbs equilibrium* without computing it.

Note that the fluxes can also be written as

$$M^-_{i+\frac{1}{2}}(\xi) = M_{i+\frac{1}{2}}(\xi) + \Big(M^-_{i+\frac{1}{2}}(\xi) - M_{i+\frac{1}{2}}(\xi)\Big)$$

and the quantity $\delta M^-_{i+\frac{1}{2}}(\xi) = M^-_{i+\frac{1}{2}}(\xi) - M_{i+\frac{1}{2}}(\xi)$ holds for the discrete contribution of the force term $h\frac{\partial Z}{\partial x}$ in the system, for negative velocities; indeed, $\delta M^-_{i+\frac{1}{2}}(\xi) = 0$ for $\xi \geq 0$ in the scheme to be presented below. This is the principle of the Interfacial Upwind Sources method: the source is not treated as a volumic term but at the interfaces and it is upwinded.

Now, we integrate the equation (3.1) in $d\xi$ against 1 and ξ, with notation

$$U^{n+1}_i = (h^{n+1}_i, (hu)^{n+1}_i), \tag{3.2}$$

$$h^{n+1}_i = \int_{\mathbb{R}} f^{n+1}_i(\xi)d\xi, \quad (hu)^{n+1}_i = \int_{\mathbb{R}} \xi f^{n+1}_i(\xi)d\xi \tag{3.3}$$

and we obtain the macroscopic scheme

$$U^{n+1}_i - U^n_i + \frac{\Delta t}{\Delta x}\left[\mathbb{F}^-_{i+\frac{1}{2}} - \mathbb{F}^+_{i-\frac{1}{2}}\right] = 0. \tag{3.4}$$

The numerical fluxes are thus given by the kinetic fluxes

$$\mathbb{F}^-_{i+\frac{1}{2}} = \int_{\mathbb{R}} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M^-_{i+\frac{1}{2}}(\xi)d\xi, \tag{3.5}$$

$$\mathbb{F}^+_{i-\frac{1}{2}} = \int_{\mathbb{R}} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M^+_{i-\frac{1}{2}}(\xi)d\xi. \tag{3.6}$$

In order to take the neighboring cells into account by means of a natural interpretation of the microscopic features of the system, we formulate a peculiar discretization for the fluxes in (3.1), computed by the upwind formulas

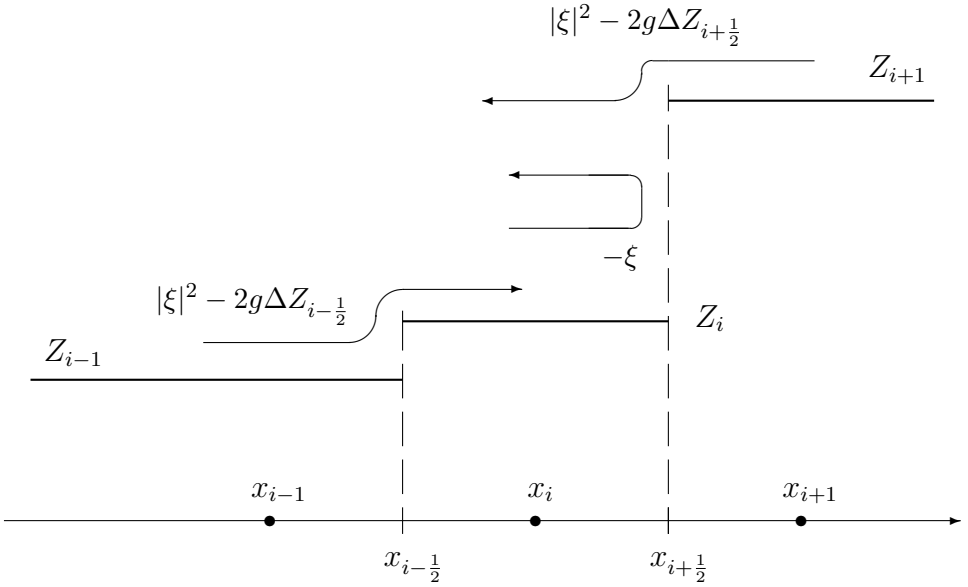
$$M^-_{i+\frac{1}{2}}(\xi) = M^n_i(\xi)\mathbb{1}_{\xi\geq 0} + M^n_{i+\frac{1}{2}}(\xi)\mathbb{1}_{\xi\leq 0}, \tag{3.7}$$

$$M^+_{i-\frac{1}{2}}(\xi) = M^n_{i-\frac{1}{2}}(\xi)\mathbb{1}_{\xi\geq 0} + M^n_i(\xi)\mathbb{1}_{\xi\leq 0}, \tag{3.8}$$

where we define

$$M^n_{i+\frac{1}{2}}(\xi) = M^n_i(-\xi)\mathbb{1}_{|\xi|^2\leq 2g\Delta Z_{i+\frac{1}{2}}} + M^n_{i+1}\left(-\sqrt{|\xi|^2 - 2g\Delta Z_{i+\frac{1}{2}}}\right)\mathbb{1}_{|\xi|^2\geq 2g\Delta Z_{i+\frac{1}{2}}},$$
$$M^n_{i-\frac{1}{2}}(\xi) = M^n_i(-\xi)\mathbb{1}_{|\xi|^2\leq 2g\Delta Z_{i-\frac{1}{2}}} + M^n_{i-1}\left(\sqrt{|\xi|^2 - 2g\Delta Z_{i-\frac{1}{2}}}\right)\mathbb{1}_{|\xi|^2\geq 2g\Delta Z_{i-\frac{1}{2}}}.$$

The figure below illustrates the typical situation occurring in a cell C_i of the mesh, centered at the point $x_i \in \mathbb{R}$; without loss of generality, we consider here the case of an increasing bottom slope, so that the bottom jumps are positive and negative for the neighboring cells C_{i-1} and C_{i+1} respectively.



The effect of the source term is made explicit by treating it as a physical potential. The definitions (3.7)-(3.8) are thus a mathematical formalization to describe the physical microscopic behaviour of the system: contributions to the value $f_i^{n+1}(\xi)$ are also given by particles in C_{i+1} and in C_{i-1} at time t_n, with kinetic energy sufficient to surpass the potential difference (speeded up or down through the potential jump) and by particles coming at velocity $-\xi$, reflected on the bottom jumps according to classical mechanics, when their energy is too small (i.e. $|\xi|^2 \leq 2g\Delta Z_{i+\frac{1}{2}}$).

Remark 3.1. We see immediately that *the kinetic scheme (3.4)-(3.6) is water height conservative*. In fact, still referring to the figure, we compute the first component of the numerical fluxes at the interface $x_{i+\frac{1}{2}}$ of the mesh

by the formulas (3.5)-(3.6),

$$(\mathbb{F}_h)^-_{i+\frac{1}{2}} = \int_{\xi\geq 0} \xi M_i^n(\xi)d\xi + \int_{\xi\leq 0} \xi M_i^n(-\xi)\mathbb{1}_{|\xi|^2\leq 2g\Delta Z_{i+\frac{1}{2}}}d\xi + \int_{\xi\leq 0} \xi M_{i+1}^n\left(-\sqrt{|\xi|^2-2g\Delta Z_{i+\frac{1}{2}}}\right)\mathbb{1}_{|\xi|^2\geq 2g\Delta Z_{i+\frac{1}{2}}}d\xi$$

and

$$(\mathbb{F}_h)^+_{i+\frac{1}{2}} = \int_{\xi\leq 0} \xi M_{i+1}^n(\xi)d\xi + \int_{\xi\geq 0} \xi M_i^n\left(\sqrt{|\xi|^2-2g\Delta Z_{i+\frac{1}{2}}}\right)d\xi,$$

so that a simple change of variable $|\xi'|^2 = |\xi|^2 - 2g\Delta Z_{i+\frac{1}{2}}$, $\xi' d\xi' = \xi d\xi$ allows to conclude that

$$(\mathbb{F}_h)^-_{i+\frac{1}{2}} = (\mathbb{F}_h)^+_{i+\frac{1}{2}}, \quad \forall i \in \mathbb{Z}.$$

The conservation of water height and momentum is also obvious for the system with a flat bottom: the continuous system (1.1)-(1.2) becomes homogeneous ($\frac{\partial Z}{\partial x} = 0$) and we obtain a conservative scheme, with the flux-splitting form of the standard kinetic scheme.
We emphasize that to show the property of the numerical scheme (3.4)-(3.6) to be consistent cannot be achieved in the classical manner. Because of the presence of the source term and the choice to process it implicitly, this question is much more delicate here.

3.2 Properties of the numerical scheme

We establish some theoretical properties of the numerical scheme introduced in the previous subsection, which represent the discrete analogue of the main properties of the Saint-Venant system stated in Section 2.

Theorem 3.2. *We assume the CFL condition*

$$\Delta t \ max\left(|u_i^n| + \sqrt{2gh_i^n}\right) \leq \Delta x. \tag{3.9}$$

Then, (i) the kinetic scheme (3.4)-(3.6) keeps the water height positive, i.e. $h_i^n \geq 0$ if this is the case initially; (ii) it satisfies the conservative in-cell entropy inequality,

$$E_i^{n+1} - E_i^n + \frac{\Delta t}{\Delta x}\left[\eta_{i+\frac{1}{2}}^n - \eta_{i-\frac{1}{2}}^n\right] \leq 0,$$

with the discrete entropy fluxes given in the formulas (3.11)-(3.12) below and the discrete energy

$$E_i^n = h_i^n\frac{|u_i^n|^2}{2} + \frac{g}{2}(h_i^n)^2 + gZ_ih_i^n;$$

(iii) the scheme (3.4)-(3.6) preserves the steady states of the system given by a lake at rest,

$$u_i^n = 0, \quad h_i^n + Z_i = H, \qquad \forall i \in \mathbb{Z}, \quad \forall n \in \mathbb{N}.$$

Proof. To prove the first stability property (i) of the scheme, we come back to the kinetic interpretation and we proceed by induction. We assume that $h_i^n \geq 0$, $\forall i \in \mathbb{Z}$, and we prove that $h_i^{n+1} \geq 0$, $\forall i \in \mathbb{Z}$; since

$$h_i^{n+1} = \int_{\mathbb{R}} f_i^{n+1}(\xi) d\xi,$$

it is sufficient to prove that $f_i^{n+1}(\xi) \geq 0$. We introduce the quantities

$$\xi_+ = max(0, \xi), \quad \xi_- = max(-\xi, 0), \quad \sigma = \frac{\Delta t}{\Delta x},$$

so that we can rewrite the microscopic scheme (3.1), (3.7)-(3.8) in the form

$$\begin{aligned} f_i^{n+1}(\xi) &= M_i^n(\xi) - \sigma\xi \left(M^-_{i+\frac{1}{2}}(\xi) - M^+_{i-\frac{1}{2}}(\xi) \right) \\ &= (1 - \sigma|\xi|)\, M_i^n(\xi) + \sigma\xi_- \Big[M_i^n(-\xi) \mathbb{1}_{|\xi|^2 \leq 2g\Delta Z_{i+\frac{1}{2}}} \\ &\qquad + M_{i+1}^n \left(-\sqrt{|\xi|^2 - 2g\Delta Z_{i+\frac{1}{2}}} \right) \mathbb{1}_{|\xi|^2 \geq 2g\Delta Z_{i+\frac{1}{2}}} \Big] \qquad (3.10) \\ &+ \sigma\xi_+ \Big[M_i^n(-\xi) \mathbb{1}_{|\xi|^2 \leq 2g\Delta Z_{i-\frac{1}{2}}} \\ &\qquad + M_{i-1}^n \left(\sqrt{|\xi|^2 - 2g\Delta Z_{i-\frac{1}{2}}} \right) \mathbb{1}_{|\xi|^2 \geq 2g\Delta Z_{i-\frac{1}{2}}} \Big]. \end{aligned}$$

Since the function χ has a compact support, it follows that

$$M_j^n(\xi) = 0 \qquad \text{if} \quad |\xi - u_j^n| \geq \sqrt{2gh_j^n};$$

we deduce that

$$f_i^{n+1}(\xi) \geq 0 \qquad \text{if} \quad |\xi - u_j^n| \geq \sqrt{2gh_j^n}, \quad \forall j \in \mathbb{Z},$$

as a sum of non-negative quantities.
Now, if $M_i^n(\xi) \neq 0$ then $|\xi - u_i^n| \leq \sqrt{2gh_i^n}$ and

$$|\xi| \leq |\xi - u_i^n| + |u_i^n| \leq \sqrt{2gh_i^n} + |u_i^n|.$$

We use the CFL condition to conclude that $\sigma|\xi| \leq 1$, so that $f_i^{n+1}(\xi)$ is a convex combination of three non-negative quantities and thus $f_i^{n+1}(\xi) \geq 0$,

$\forall \xi \in \mathbb{R}$, $\forall i \in \mathbb{Z}$.
For the entropy inequality (ii), the conclusion results from the relation (3.10), which describes $f_i^{n+1}(\xi)$ as a convex combination of density functions. Recalling the definition (2.12) for the energy convex functional $\mathcal{E}(f)$, we calculate it on the previous convex formula and, thanks to the relations stated in the proof of Theorem 2.2, we obtain

$$\mathcal{E}(f_i^{n+1}) - E_i^n + \frac{\Delta t}{\Delta x}\left[\eta_{i+\frac{1}{2}}^n - \eta_{i-\frac{1}{2}}^n\right] \leq 0,$$

where the entropy fluxes have the expressions

$$\begin{aligned}\eta_{i+\frac{1}{2}}^n = &\int_{\xi\geq 0}\left[\frac{\xi^3}{2}M_i^n(\xi) + \frac{\pi^2 g^2}{6}\xi\left[M_i^n(\xi)\right]^3 + gZ_i\,\xi M_i^n(\xi)\right]d\xi \\ &+ \int_{\xi\leq 0}\left[\frac{\xi^3}{2}M_{i+\frac{1}{2}}^n(\xi) + \frac{\pi^2 g^2}{6}\xi\left[M_{i+\frac{1}{2}}^n(\xi)\right]^3 + gZ_i\,\xi M_{i+\frac{1}{2}}^n(\xi)\right]d\xi,\end{aligned} \tag{3.11}$$

$$\begin{aligned}\eta_{i-\frac{1}{2}}^n = &\int_{\xi\geq 0}\left[\frac{\xi^3}{2}M_{i-\frac{1}{2}}^n(\xi) + \frac{\pi^2 g^2}{6}\xi\left[M_{i-\frac{1}{2}}^n(\xi)\right]^3 + gZ_i\,\xi M_{i-\frac{1}{2}}^n(\xi)\right]d\xi \\ &+ \int_{\xi\leq 0}\left[\frac{\xi^3}{2}M_i^n(\xi) + \frac{\pi^2 g^2}{6}\xi\left[M_i^n(\xi)\right]^3 + gZ_i\,\xi M_i^n(\xi)\right]d\xi.\end{aligned} \tag{3.12}$$

Next, we use Lemma 2.3 to deduce

$$E_i^{n+1} = \mathcal{E}(M_i^{n+1}) \leq \mathcal{E}(f_i^{n+1}).$$

Finally, we can again give a direct proof of the last statement (iii) at the microscopic level. We emphasize that this approach is also justified by the result stated in Lemma 2.5. From the formula (3.1) for the numerical scheme, it is enough to prove that

$$M_{i+\frac{1}{2}}^-(\xi) = M_{i-\frac{1}{2}}^+(\xi), \quad \forall \xi \in \mathbb{R}$$

and (iii) follows: indeed, this implies $f_i^{n+1}(\xi) = M_i^n(\xi)$, which also gives $h_i^{n+1} = h_i^n$, $u_i^{n+1} = u_i^n$, $\forall i \in \mathbb{Z}$.
According to the definition (3.7)-(3.8), we can distinguish two cases of the previous equality, for $\xi \geq 0$ and $\xi \leq 0$; since these cases present the same difficulty, we only consider the case $\xi \geq 0$. We also remark that, exploiting the hypothesis $u_i = 0$, we have

$$M_i^n(\xi) = \frac{\sqrt{2}}{\pi\sqrt{g}}\sqrt{h_i^n}\left(1 - \frac{\xi^2}{2gh_i^n}\right)^{\frac{1}{2}}$$

and

$$M^n_{i\pm1}\left(\mp\sqrt{|\xi|^2-2g\Delta Z_{i\pm\frac{1}{2}}}\right)=\frac{\sqrt{2}}{\pi\sqrt{g}}\sqrt{h^n_{i\pm1}}\left(1-\frac{\xi^2-2g\Delta Z_{i\pm\frac{1}{2}}}{2gh^n_{i\pm1}}\right)^{\frac{1}{2}}.$$

Next, for the case $\xi\geq 0$ and $|\xi|^2\leq 2g\Delta Z_{i-\frac{1}{2}}$, the result is obvious. There remains the case $\xi\geq 0$ and $|\xi|^2\geq 2g\Delta Z_{i-\frac{1}{2}}$, for which the statement reduces to verifying the equality

$$\sqrt{h^n_i}\left(1-\frac{\xi^2}{2gh^n_i}\right)^{\frac{1}{2}}=\sqrt{h^n_{i-1}}\left(1-\frac{\xi^2-2g\Delta Z_{i-\frac{1}{2}}}{2gh^n_{i-1}}\right)^{\frac{1}{2}}.$$

Thanks to the hypothesis $h^n_i+Z_i=h_{i-1}+Z_{i-1}$, $\forall i\in\mathbb{Z}$, it follows that $\Delta Z_{i-\frac{1}{2}}=h^n_i-h^n_{i-1}$, so that a simple algebraic computation completes the proof. □

4 Numerical implementation

To proceed to the actual implementation of the scheme (3.4)-(3.6), we have to compute the numerical fluxes explicitly. Since their expressions are not always immediate to calculate, it is necessary to use an approximation technique for some of them; in this section we indicate some fundamental properties and we give a possible appropriate approximation.

4.1 Computation of the integrals

According to the kinetic representation of the Saint-Venant system, the density of particles $M^n_i(\xi)$ in the formulas (3.7)-(3.8) is defined by

$$M^n_i(\xi)=\sqrt{h^n_i}\chi\left(\frac{\xi-u^n_i}{\sqrt{h^n_i}}\right),$$

which represents the discrete analogue of the microscopic *Gibbs equilibrium* considered in Section 2. With this definition, the formula (3.5) becomes

$$\begin{aligned}\mathbb{F}^-_{i+\frac{1}{2}}&=\int_{\xi\geq0}\xi\begin{pmatrix}1\\ \xi\end{pmatrix}M^n_i(\xi)d\xi+\int_{\xi\leq0}\xi\begin{pmatrix}1\\ \xi\end{pmatrix}M^n_{i+\frac{1}{2}}(\xi)d\xi\\&=\sqrt{h^n_i}\int_{\xi\geq0}\xi\begin{pmatrix}1\\ \xi\end{pmatrix}\chi\left(\frac{\xi-u^n_i}{\sqrt{h^n_i}}\right)d\xi\\&+\sqrt{h^n_i}\int_{\xi\leq0}\xi\begin{pmatrix}1\\ \xi\end{pmatrix}\chi\left(\frac{-\xi-u^n_i}{\sqrt{h^n_i}}\right)\mathbb{1}_{|\xi|^2\leq2g\Delta Z_{i+\frac{1}{2}}}d\xi\end{aligned}$$

$$+\sqrt{h^n_{i+1}}\int_{\xi\le 0}\xi\begin{pmatrix}1\\ \xi\end{pmatrix}\chi\left(\frac{-\sqrt{|\xi|^2-2g\Delta Z_{i+\frac12}}-u^n_{i+1}}{\sqrt{h^n_{i+1}}}\right)\mathbb{1}_{|\xi|^2\ge 2g\Delta Z_{i+\frac12}}d\xi.$$

We consider a change of variable $\xi'=-\sqrt{|\xi|^2-2g\Delta Z_{i+\frac12}}$, $\xi' d\xi'=\xi d\xi$, in the third term to obtain

$$\begin{aligned}&\int_{\xi\le 0}\xi\begin{pmatrix}1\\ \xi\end{pmatrix}\chi\left(\frac{-\sqrt{|\xi|^2-2g\Delta Z_{i+\frac12}}-u^n_{i+1}}{\sqrt{h^n_{i+1}}}\right)\mathbb{1}_{|\xi|^2\ge 2g\Delta Z_{i+\frac12}}d\xi\\ &=\int_{\xi'\le 0}\xi'\begin{pmatrix}1\\ -\sqrt{|\xi'|^2+2g\Delta Z_{i+\frac12}}\end{pmatrix}\chi\left(\frac{\xi'-u^n_{i+1}}{\sqrt{h^n_{i+1}}}\right)\mathbb{1}_{|\xi'|^2\ge -2g\Delta Z_{i+\frac12}}d\xi';\end{aligned}$$

then, a simple computation allows to conclude that

$$\begin{aligned}\mathbb{F}^-_{i+\frac12}&=h^n_i\int_{\omega\ge-\frac{u^n_i}{\sqrt{h^n_i}}}(\omega\sqrt{h^n_i}+u^n_i)\begin{pmatrix}1\\ \omega\sqrt{h^n_i}+u^n_i\end{pmatrix}\chi(\omega)d\omega\\ &\quad-h^n_i\int(\omega\sqrt{h^n_i}+u^n_i)\begin{pmatrix}1\\ -(\omega\sqrt{h^n_i}+u^n_i)\end{pmatrix}\\ &\qquad\times\chi(\omega)\mathbb{1}_{0\le\omega\sqrt{h^n_i}+u^n_i\le\sqrt{2g(\Delta Z_{i+\frac12})_+}}d\omega\\ &\quad+h^n_{i+1}\int_{\omega\le\frac{-\sqrt{2g(\Delta Z_{i+\frac12})_-}-u^n_{i+1}}{\sqrt{h^n_{i+1}}}}(\omega\sqrt{h^n_{i+1}}+u^n_{i+1})\\ &\qquad\times\begin{pmatrix}1\\ -\sqrt{|\omega\sqrt{h^n_{i+1}}+u^n_{i+1}|^2+2g\Delta Z_{i+\frac12}}\end{pmatrix}\chi(\omega)d\omega,\end{aligned}$$

where we have used the classical algebraic notations

$$(\Delta Z_{i+\frac12})_+=max(0,\Delta Z_{i+\frac12}),\quad (\Delta Z_{i+\frac12})_-=max(-\Delta Z_{i+\frac12},0).$$

Similar manipulations in the formula (3.6) lead to

$$\begin{aligned}\mathbb{F}^+_{i-\frac12}&=-h^n_i\int(\omega\sqrt{h^n_i}+u^n_i)\begin{pmatrix}1\\ -(\omega\sqrt{h^n_i}+u^n_i)\end{pmatrix}\\ &\qquad\times\chi(\omega)\mathbb{1}_{-\sqrt{2g(\Delta Z_{i-\frac12})_+}\le\omega\sqrt{h^n_i}+u^n_i\le 0}d\omega\\ &\quad+h^n_{i-1}\int_{\omega\ge\frac{\sqrt{2g(\Delta Z_{i-\frac12})_-}-u^n_{i-1}}{\sqrt{h^n_{i-1}}}}(\omega\sqrt{h^n_{i-1}}+u^n_{i-1})\\ &\qquad\times\begin{pmatrix}1\\ \sqrt{|\omega\sqrt{h^n_{i-1}}+u^n_{i-1}|^2+2g\Delta Z_{i-\frac12}}\end{pmatrix}\chi(\omega)d\omega\end{aligned}$$

$$+ h_i^n \int_{\omega \le -\frac{u_i^n}{\sqrt{h_i^n}}} (\omega\sqrt{h_i^n} + u_i^n) \left(\frac{1}{\omega\sqrt{h_i^n} + u_i^n} \right) \chi(\omega) d\omega.$$

Remark 4.1. *As observed earlier, we are not able to compute all the integrals in the previous formulas. In fact, the choice of function χ in Section 2, which is necessary to achieve the properties of the numerical scheme stated in Theorem 3.2, leads to integrals which do not generically have an explicit primitive function.*

We distinguish three terms, for each component, in the formula of $\mathbb{F}^-_{i+\frac{1}{2}}$ stated above. We point out that similar integrals characterize the expression of $\mathbb{F}^+_{i-\frac{1}{2}}$, only with changes of sign in the domains of integration; thus we proceed to describe the flux $\mathbb{F}^-_{i+\frac{1}{2}}$.
Recalling that the function χ is defined as in (2.13), some elementary manipulations lead to obtain

$$\mathbb{F}^h_{i+\frac{1}{2}} = \frac{2\sqrt{2g}}{\pi} \left[(h_i^n)^{\frac{3}{2}} \mathbb{I}_1^h(Fr_i) + (h_i^n)^{\frac{3}{2}} \mathbb{I}_2^h(Fr_i, K^-_{i+\frac{1}{2}}) + (h_{i+1}^n)^{\frac{3}{2}} \mathbb{I}_3^h(Fr_{i+1}, K^+_{i+\frac{1}{2}}) \right]$$

and

$$\mathbb{F}^q_{i+\frac{1}{2}} = \frac{4g}{\pi} \left[(h_i^n)^2 \mathbb{I}_1^q(Fr_i) + (h_i^n)^2 \mathbb{I}_2^q(Fr_i, K^-_{i+\frac{1}{2}}) + (h_{i+1}^n)^2 \mathbb{I}_3^q(Fr_{i+1}, K^+_{i+\frac{1}{2}}) \right],$$

where we introduce the dimensionless numbers

$$Fr_i = \frac{u_i^n}{\sqrt{2gh_i^n}}, \quad K^-_{i+\frac{1}{2}} = \frac{\Delta Z_{i+\frac{1}{2}}}{h_i^n}, \quad K^+_{i+\frac{1}{2}} = \frac{\Delta Z_{i+\frac{1}{2}}}{h_{i+1}^n},$$

with, dropping the indices when no ambiguity is possible,

$$\mathbb{I}_1^h = \int_{\omega \ge -Fr} \omega \left(1 - \omega^2\right)_+^{\frac{1}{2}} d\omega + Fr \int_{\omega \ge -Fr} \left(1 - \omega^2\right)_+^{\frac{1}{2}} d\omega,$$

$$\mathbb{I}_2^h = - \int_{-Fr \le \omega \le \sqrt{(K^-)_+} - Fr} \omega \left(1 - \omega^2\right)_+^{\frac{1}{2}} d\omega - Fr \int_{-Fr \le \omega \le \sqrt{(K^-)_+} - Fr} \left(1 - \omega^2\right)_+^{\frac{1}{2}} d\omega,$$

$$\mathbb{I}_3^h = \int_{\omega \le -\sqrt{(K^+)_-} - Fr} \omega \left(1 - \omega^2\right)_+^{\frac{1}{2}} d\omega + Fr \int_{\omega \le -\sqrt{(K^+)_-} - Fr} \left(1 - \omega^2\right)_+^{\frac{1}{2}} d\omega,$$

$$\mathbb{I}_1^q = \int_{\omega \geq -Fr} \omega^2 \left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega + 2\ Fr \int_{\omega \geq -Fr} \omega \left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega + Fr^2 \int_{\omega \geq -Fr} \left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega,$$

$$\mathbb{I}_2^q = \int_{-Fr \leq \omega \leq \sqrt{(K^-)_+}-Fr} \omega^2 \left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega + 2\ Fr \int_{-Fr \leq \omega \leq \sqrt{(K^-)_+}-Fr} \omega \left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega + Fr^2 \int_{-Fr \leq \omega \leq \sqrt{(K^-)_+}-Fr} \left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega$$

and, finally,

$$\mathbb{I}_3^q = -\int_{\omega \leq -\sqrt{(K^+)_-}-Fr} (\omega+Fr)\sqrt{(\omega+Fr)^2+K^+}\left(1-\omega^2\right)_+^{\frac{1}{2}} d\omega.$$

Note that almost all the previous terms reduce to the same basic forms, except the last term which is quite different from the others and we will treat it later on. Classical techniques of integration, along with the usual goniometric equalities, allow us to conclude that

$$\int_a^b \omega(1-\omega^2)^{\frac{1}{2}} d\omega = -\frac{1}{3}(1-\omega^2)^{\frac{3}{2}}\Big|_a^b,$$

$$\int_a^b (1-\omega^2)^{\frac{1}{2}} d\omega = \frac{1}{2}\left(\arccos\omega - \omega\sqrt{1-\omega^2}\right)\Big|_b^a,$$

$$\int_a^b \omega^2(1-\omega^2)^{\frac{1}{2}} d\omega = -\frac{1}{3}\omega(1-\omega^2)^{\frac{3}{2}}\Big|_a^b + \frac{1}{12}\left[\frac{3}{2}\arccos\omega + \omega\sqrt{1-\omega^2}\left(\omega^2-\frac{5}{2}\right)\right]\Big|_b^a.$$

The choice of limits in these integrals is made according to the support of the function χ, so that

$$-1 \leq a = a(Fr, K^-, K^+, \pm 1) \leq 1,$$
$$-1 \leq b = b(Fr, K^-, K^+, \pm 1) \leq 1;$$

we specify the values of a and b at the moment of writing the final procedures in the actual implementation of the numerical method.
Then, a short computation leads to the following results:

$$\mathbb{I}_1^h = \frac{1}{3}(1-\alpha^2)^{\frac{3}{2}} + \frac{1}{2}Fr\arccos\alpha - \frac{1}{2}Fr\alpha\sqrt{1-\alpha^2},$$

$$\mathbb{I}_1^q = \frac{1}{3}(1-\alpha^2)^{\frac{3}{2}}(2Fr+\alpha) + \frac{1}{2}\arccos\alpha\left(\frac{1}{4}+Fr^2\right) + \frac{1}{2}\alpha\sqrt{1-\alpha^2}\left(\frac{1}{6}\alpha^2 - \frac{5}{12} - Fr^2\right),$$

where $\alpha = min\{1, max\{-1, -Fr\}\}$;

$$\mathbb{I}_2^h = \frac{1}{3}(1-\beta^2)^{\frac{3}{2}} - \frac{1}{3}(1-\alpha^2)^{\frac{3}{2}} + \frac{1}{2}Fr(\arccos\beta - \arccos\alpha) + \frac{1}{2}Fr(\alpha\sqrt{1-\alpha^2} - \beta\sqrt{1-\beta^2}),$$

$$\mathbb{I}_2^q = \frac{1}{3}(1-\alpha^2)^{\frac{3}{2}}(2Fr+\alpha) - \frac{1}{3}(1-\beta^2)^{\frac{3}{2}}(2Fr+\beta) + \frac{1}{2}\left(\frac{1}{4}+Fr^2\right)(\arccos\alpha - \arccos\beta) + \frac{1}{2}\beta\sqrt{1-\beta^2}\left(\frac{5}{12} - \frac{1}{6}\beta^2 + Fr^2\right) - \frac{1}{2}\alpha\sqrt{1-\alpha^2}\left(\frac{5}{12} - \frac{1}{6}\alpha^2 + Fr^2\right),$$

where $\alpha = min\{1, max\{-1, -Fr\}\}$ and $\beta = max\left\{-1, min\{\sqrt{(K^-)_+} - Fr, 1\}\right\}$;

$$\mathbb{I}_3^h = -\frac{1}{3}(1-\beta^2)^{\frac{3}{2}} + \frac{1}{2}Fr(\pi - \arccos\beta) - \frac{1}{2}Fr\beta\sqrt{1-\beta^2},$$

where $\beta = max\left\{-1, min\{-\sqrt{(K^+)_-} - Fr, 1\}\right\}$.

We now return to the most complicated term $\mathbb{I}_3^q$. Setting $\alpha = -1$ and $\beta = max\left\{-1, min\{-\sqrt{(K^+)_-} - Fr, 1\}\right\}$, we can rewrite it as

$$\mathbb{I}_3^q = \int_\alpha^\beta f(\omega, Fr, K^+)d\omega,$$

with

$$f(\omega, Fr, K^+) = -(\omega + Fr)\sqrt{(\omega+Fr)^2 + K^+}\left(1-\omega^2\right)_+^{\frac{1}{2}}.$$

The presence of the square root makes it impossible to compute immediately; we need to formulate a suitable approximation, preserving the main theoretical features of the real integral. We propose a rather natural choice, based on a numerical method of integration, by means of a quadrature formula: in particular, comparison tests lead us to prefer a classical *repeated midpoint formula*,

$$\mathbb{I}_3^q(Fr, K^+) \simeq \mathbb{I}^*(Fr, K^+) = \frac{\beta-\alpha}{N}\sum_{j=1}^{N} f\left(\alpha + (j-\frac{1}{2})\frac{\beta-\alpha}{N}\right),$$

where N is chosen in order to assure the best compromise between the accuracy of the numerical method and a reasonable computating time. Of course faster algorithms are possible but at this level we are only interested in testing the method and we leave it for further extensions to improve the computational performance.

4.2 Some numerical tests

We conclude these notes with some numerical examples, that illustrate the results stated in the previous sections, in order to confirm that the properties of the Saint-Venant system (1.1)-(1.2) are preserved by the numerical scheme (3.4)-(3.6) introduced in this paper and to evaluate its performance on other classical test cases.
We check the properties of the scheme on different test cases for which analytical solutions of the equations are available and on more realistic applications (most of the experiments simulated here come from a workshop on dam-break wave simulation).

4.2.1 We begin with a non-stationary test case, a dam-break problem in a rectangular channel with flat bottom ($Z = 0$). The initial conditions are

$$u(0,x) = 0,$$
$$h(0,x) = \begin{cases} h_l & \text{for } x \leq 0 \\ h_r & \text{for } x > 0, \end{cases}$$

where $h_l > h_r$ in order to be consistent with the physical phenomenon of a dam-break from the left to the right.
Note that this case corresponds to a Riemann problem for the simpler homogeneous model of system (1.1)-(1.2) and we can compare the numerical solution with the exact solution (plotted with a dotted line), computed by the classical theory (see Dafermos [5], Serre [22]).
The channel length is L = $2000m$ and the computational domain is chosen to be symmetric around the point x = 0; the mesh size is Ds = L/100 and the time-step Dt is computed according to the CFL condition (3.9), in order to verify numerically that the water height positivity is preserved.
The Figures (1)-(2) and (3)-(4) present respectively the results observed at time T = $200s$ for a dam-break on a wet bed ($h_l = 1m$, $h_r = 0.5m$) and at time T = $150s$ for a dam-break on a dry soil ($h_l = 1m$, $h_r = 0m$).

Figure (1): DAM−BREAK on a WED BED − final water level

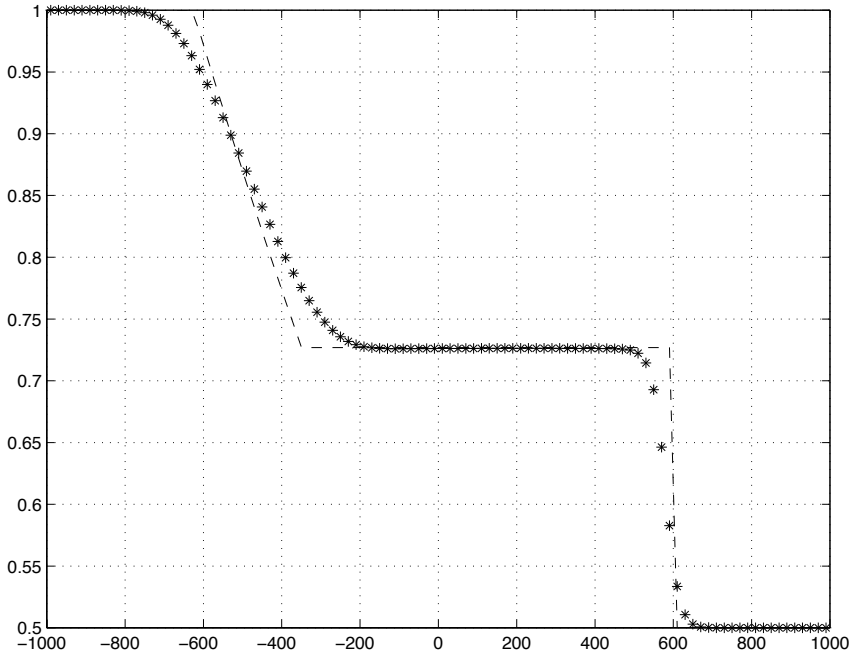
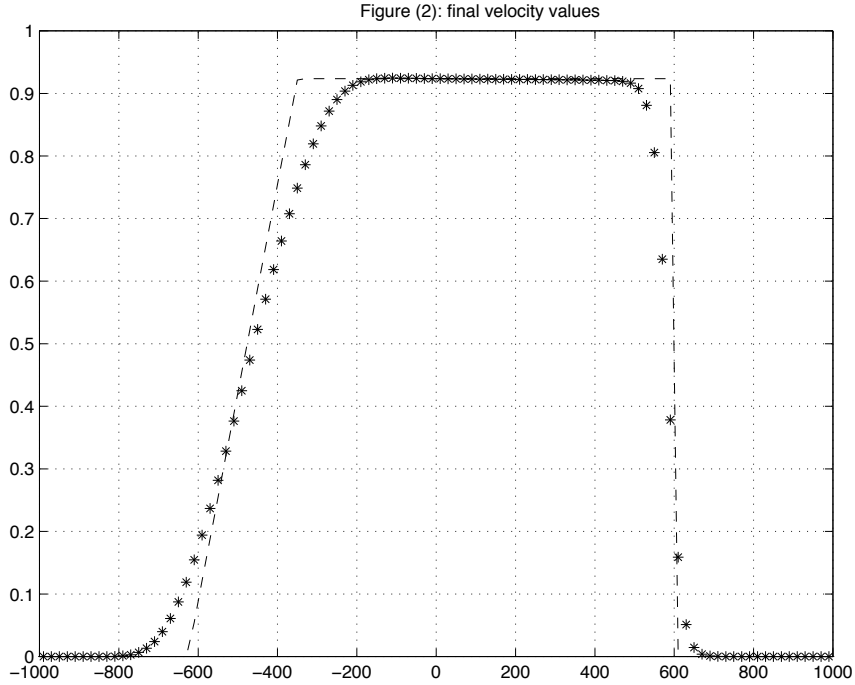
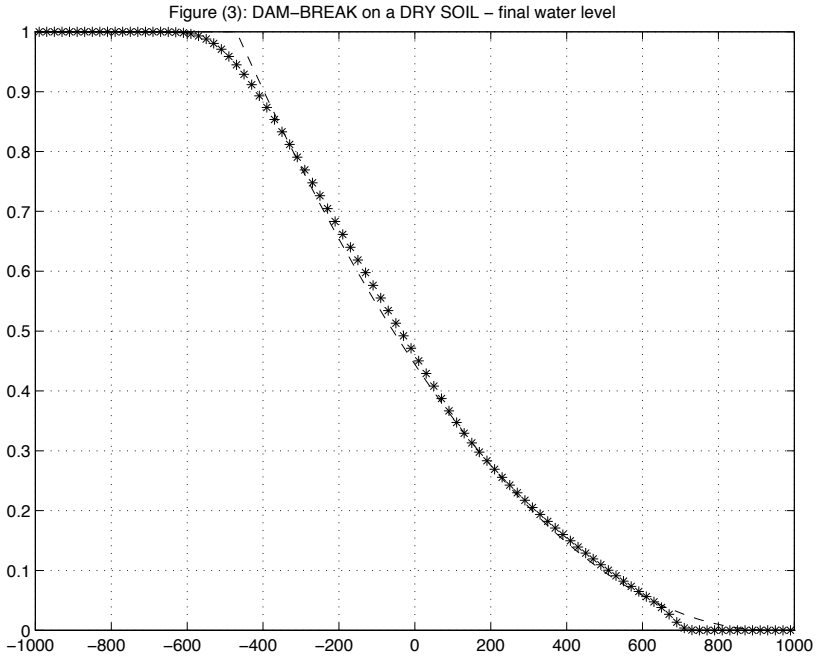
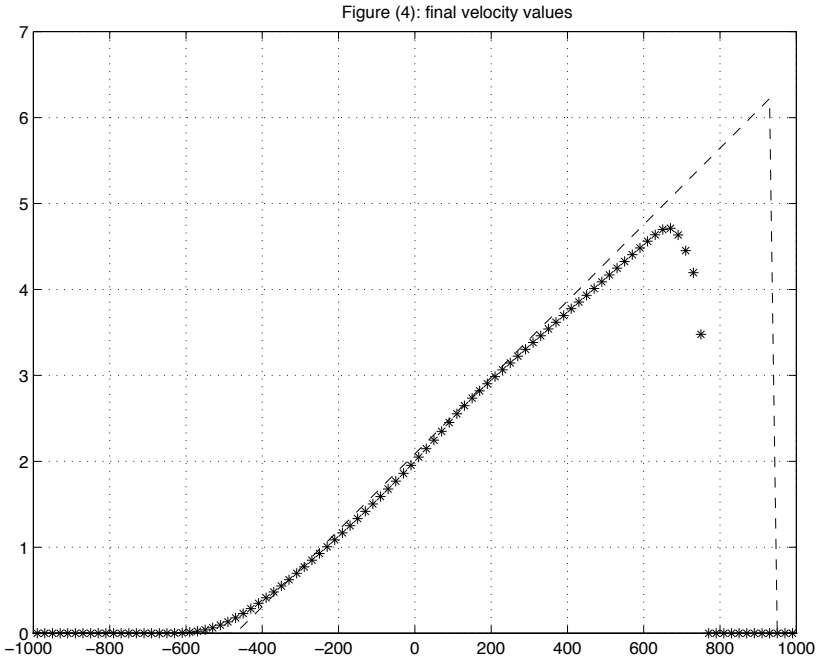


Figure (2): final velocity values







4.2.2 We now consider a test case concerning the steady state of a lake at rest, on a non-trivial topography, in order to validate the numerical scheme on a steady flow: we show that this steady state is preserved, up to the accuracy in the approximation of the integral we have discussed at the end of Subsection 4.1.

Figure (5): LAKE AT REST – final water level

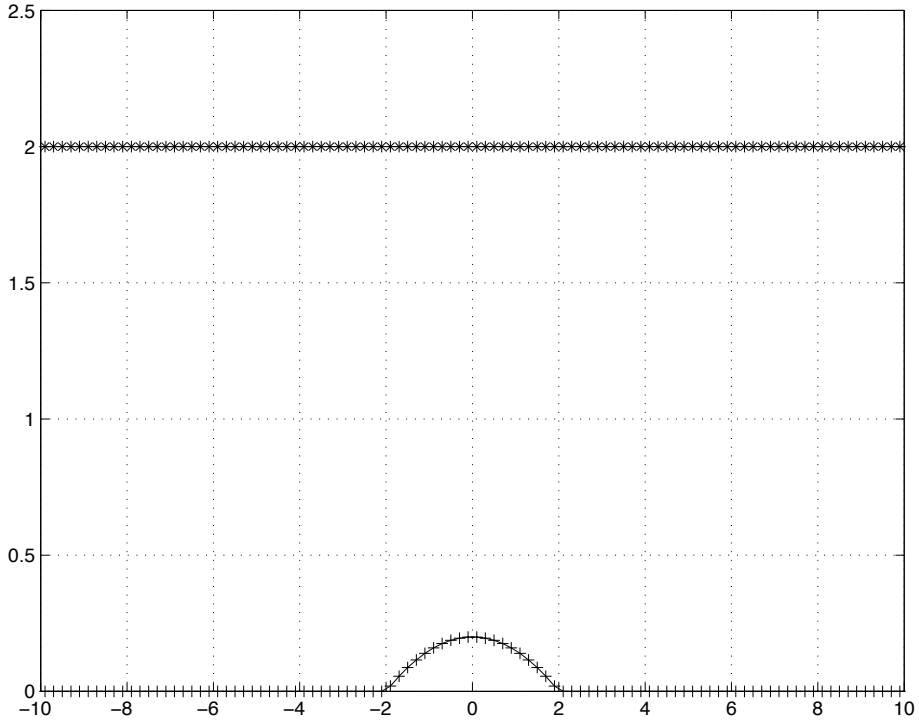


Figure (6): final velocity values

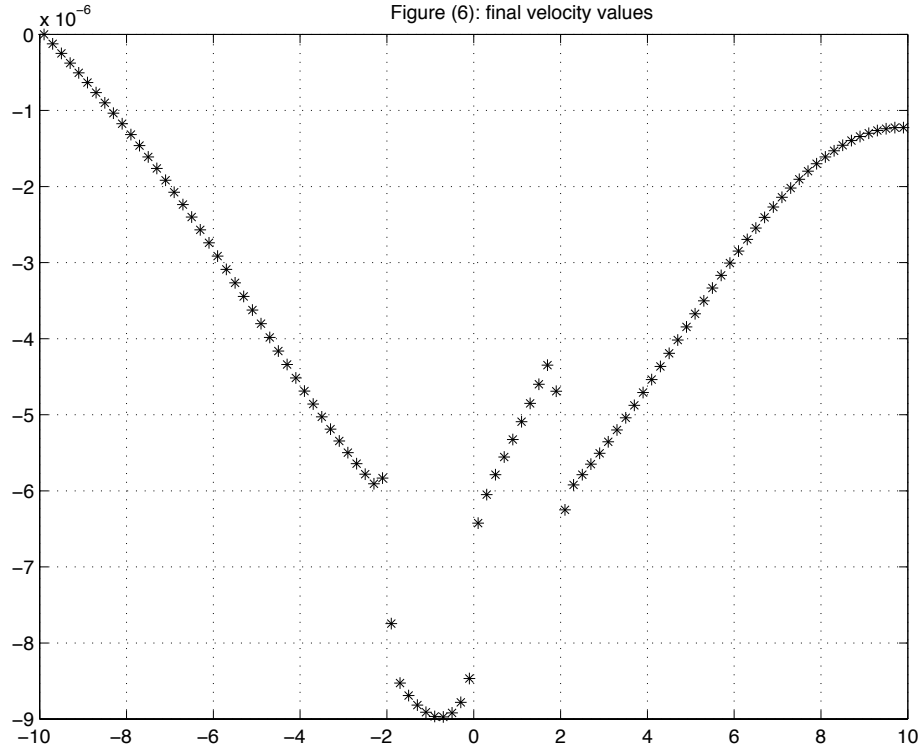


Figure (7): LAKE AT REST – final water level

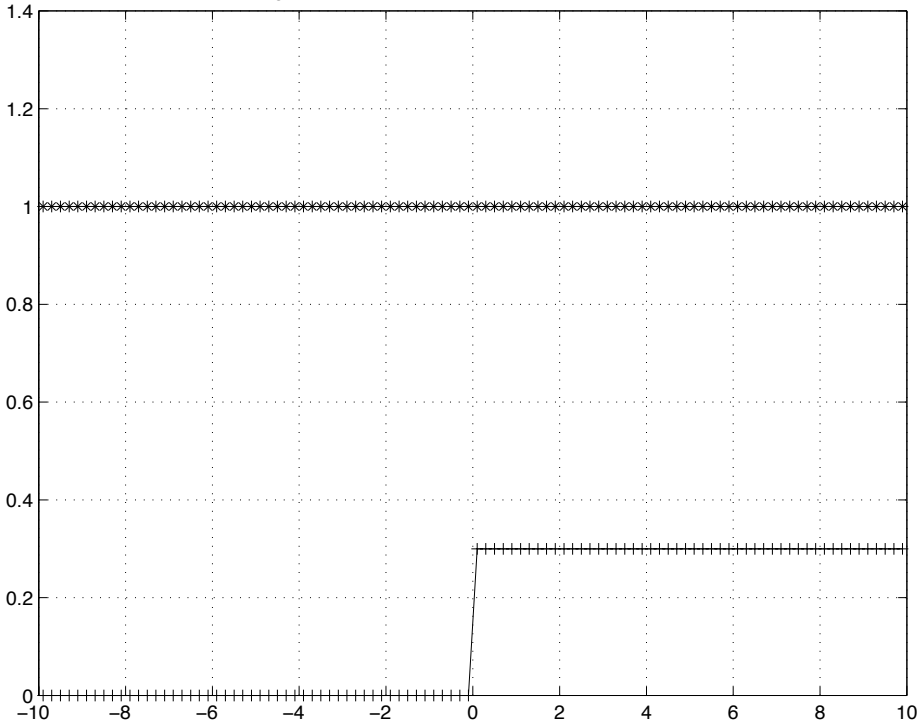
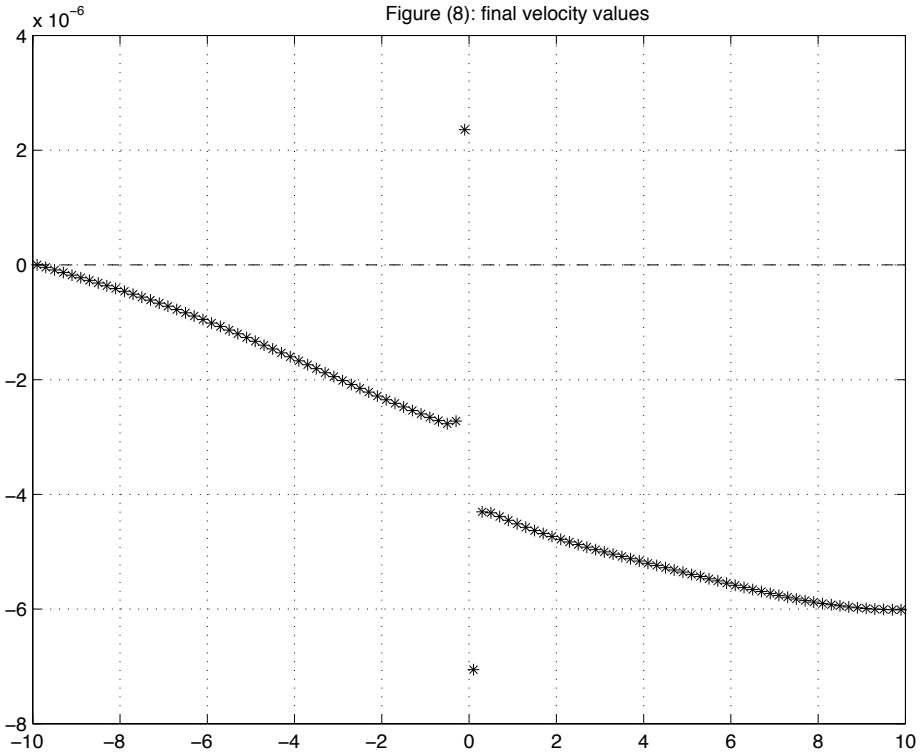


Figure (8): final velocity values



The initial conditions are

$$u(0,x) = 0,$$
$$h(0,x) + Z(x) = H,$$

with $H = 2m$ and we set the same computational parameters as in the first test case, except for the channel length L $= 20m$. The solution plotted in the Figures (5)-(6) corresponds to a channel with a parabolic bump on the bottom, described by the function

$$Z(x) = \left(0.2 - 0.05 * x^2\right)_+ ;$$

in the Figures (7)-(8) we present the same test case on a different topography, given by a discontinuous step

$$Z(x) = \begin{cases} Z_l & \text{for } x \leq 0 \\ Z_r & \text{for } x > 0, \end{cases}$$

with $Z_l = 0m$ and $Z_r = 0.3m$ (note that the geometry of the source term is not regular here, which is not in agreement with the assumptions of the classical theory).

4.2.3 Our purpose in the following test cases is to study the convergence in time towards a more general steady state. We consider a rectangular channel with the same geometry as in Figure (5) and we compute the steady state occurring since a constant discharge is imposed at the upstream boundary condition. We compare the numerical solution with the analytical solution, provided by the means of the formulas (2.3)-(2.4) in Section 1.
According to the boundary and initial conditions, the flow may be subcritical (or fluvial), transcritical without shock (the flow becomes torrential at the top of the bump and the outflow is torrential) and transcritical with shock (the flow becomes torrential at the top of the bump and the outflow is fluvial). We impose an upstream boundary condition Q_{in} on the discharge and a downstream boundary condition H_{out} on the water level, as follows:

- subcritical flow $\begin{cases} Q_{in} = 4.42m^2/s \\ H_{out} = 2m, \end{cases}$
- transcritical flow without shock $\begin{cases} Q_{in} = 1.53m^2/s \\ H_{out} = 0.66m, \end{cases}$

if the outflow is subcritical (remark that no condition is imposed on the downstream limit when the outflow becomes supercritical),

- transcritical flow with shock $\begin{cases} Q_{in} = 0.18m^2/s \\ H_{out} = 0.33m. \end{cases}$

Figure (9): RIVER at FLUVIAL FLOW – final water level

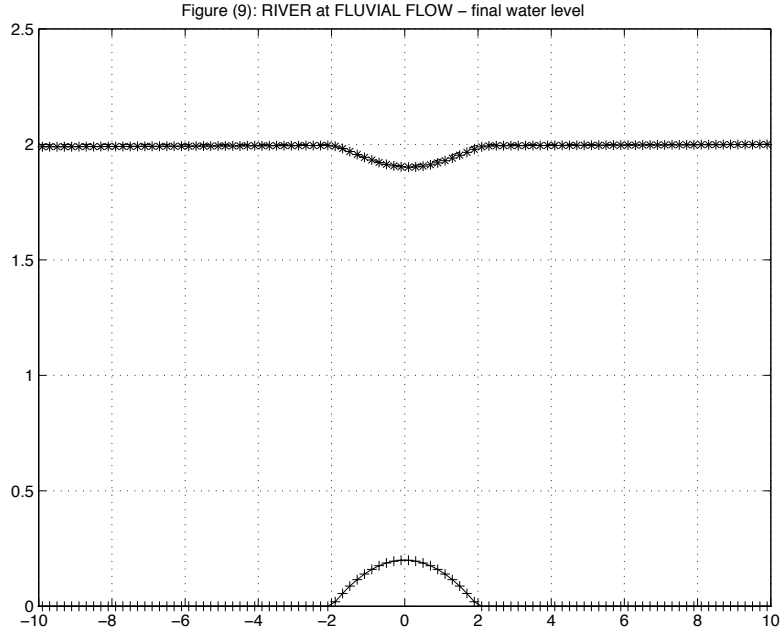
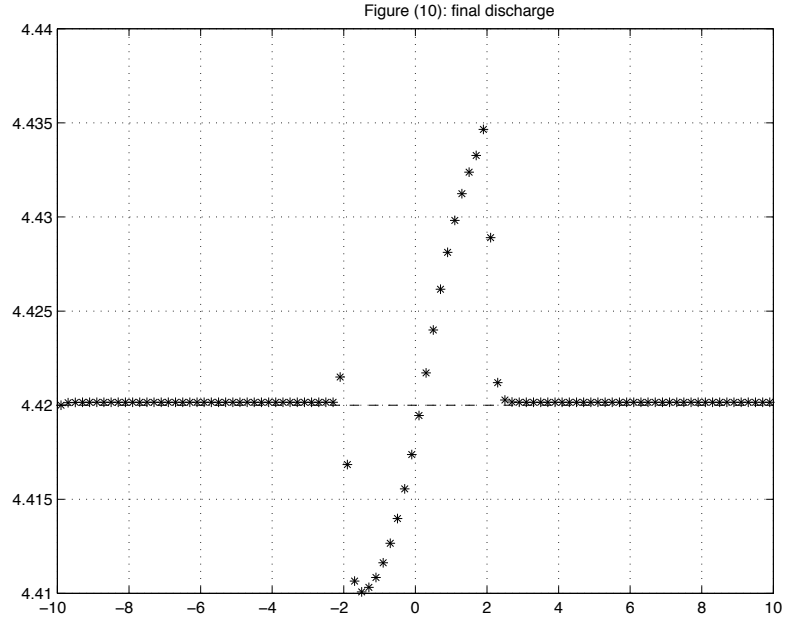


Figure (10): final discharge



For all these cases, the initial conditions are

$$u(0,x) = 0,$$
$$h(0,x) + Z(x) = H,$$

where H is the constant level of the water surface prescribed downstream. All solutions are plotted at time T = 200s (analytical solutions are also plotted

with a dotted line) and a mesh with Ds = L/100 is used, but of course some results can improve according to the mesh refinement.

Figure (11): RIVER at TORRENTIAL FLOW without shoch – final water level

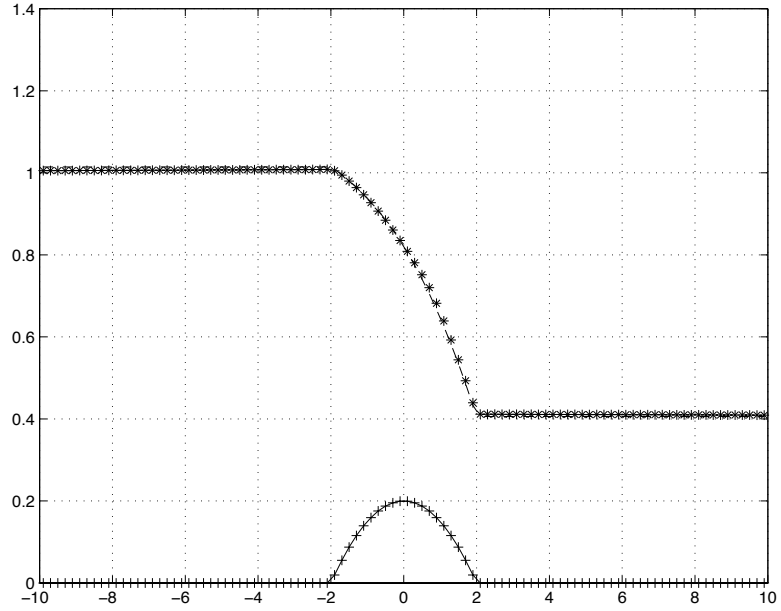
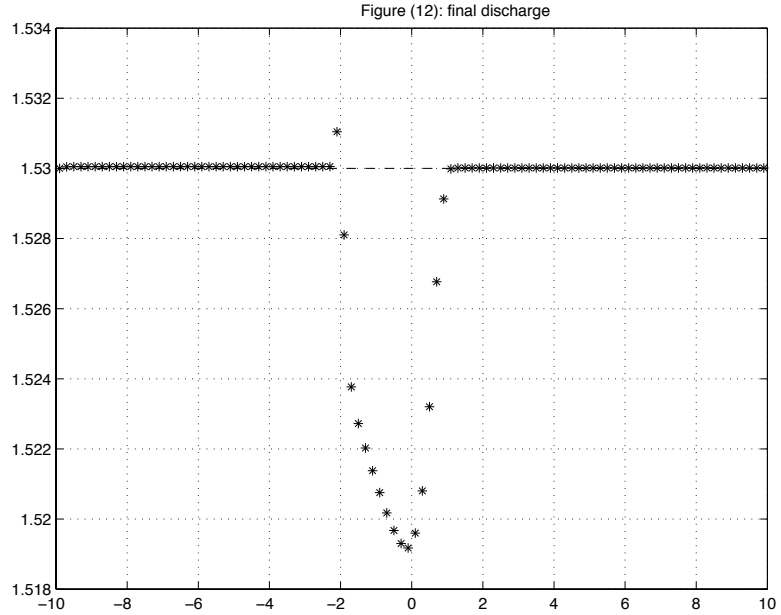
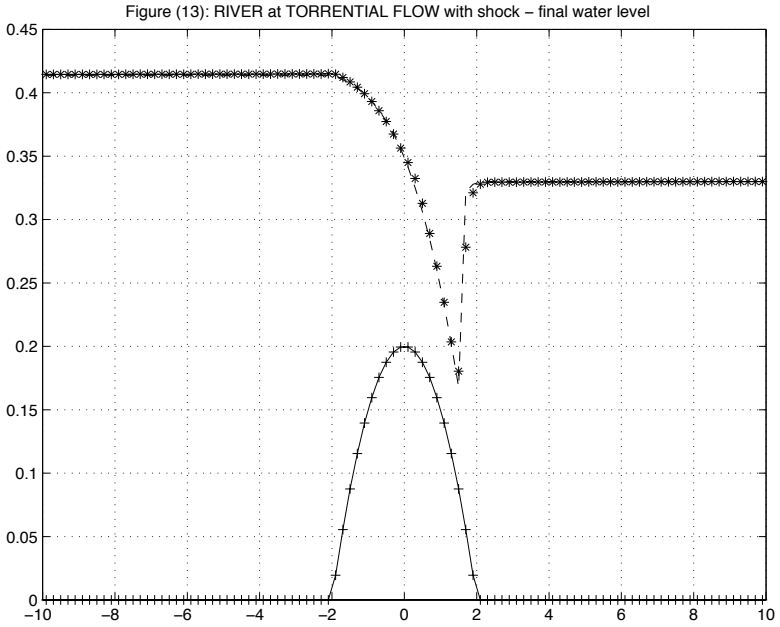
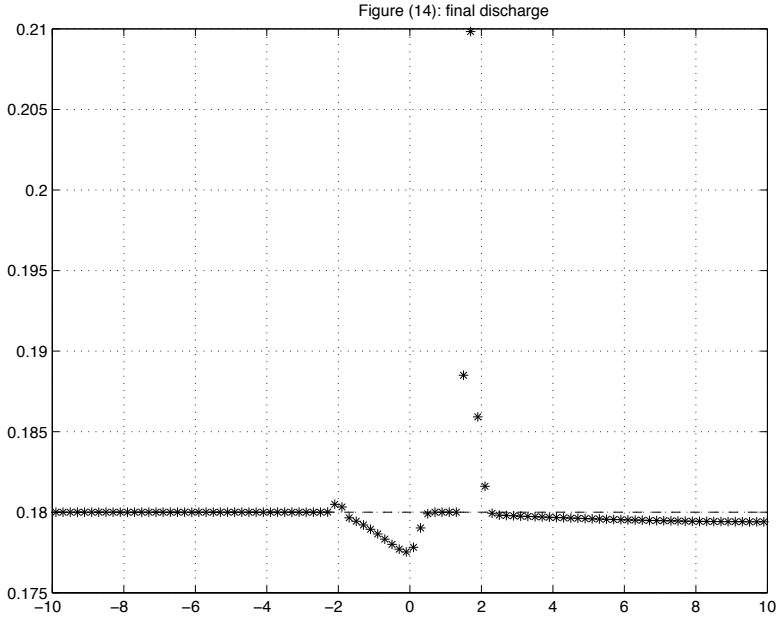


Figure (12): final discharge



Notice that Figures (10), (12) and (14) refer to steady states generally not at rest ($u \neq 0$) and thus property (iii) does not apply. To the best of our knowledge, only the results in [7] are comparable to these tests.





4.2.4 The last test case we deal with in this paper is the quasi-stationary case proposed by LeVeque in [17], to compute small perturbations of the steady state of a lake at rest. According to the parameters fixed by the author, as bottom topography we take

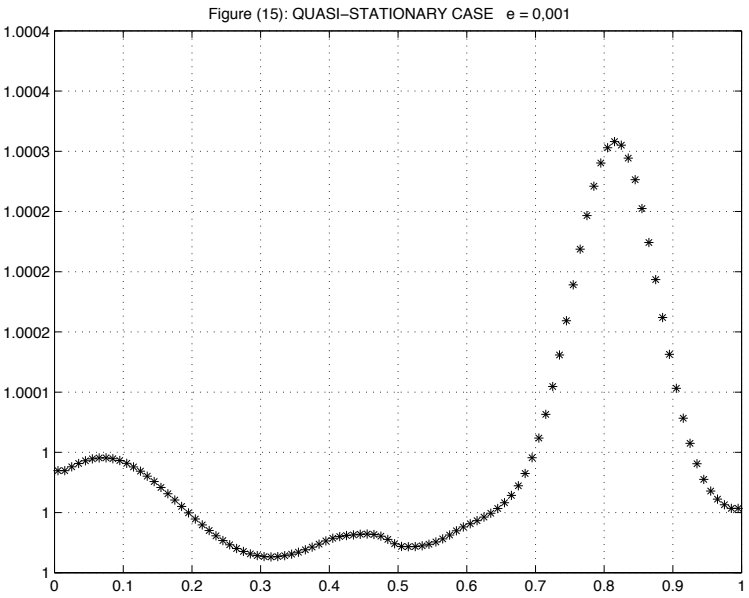
$$Z(x) = \Big(0.25 \left(\cos\left(\pi(x - 0.5)/0.1\right) + 1\right) \Big)_+ ,$$

centered in a rectangular channel of length L = $1m$ and the mesh size is Ds = L/100. The initial conditions are

$$u(0,x) = 0,$$
$$h(0,x) + Z(x) = H + \epsilon(x),$$

with $H = 1m$ and we consider a perturbation of the water surface

$$\epsilon(x) = \text{e } \mathbb{1}_{0.1<x<0.2}.$$



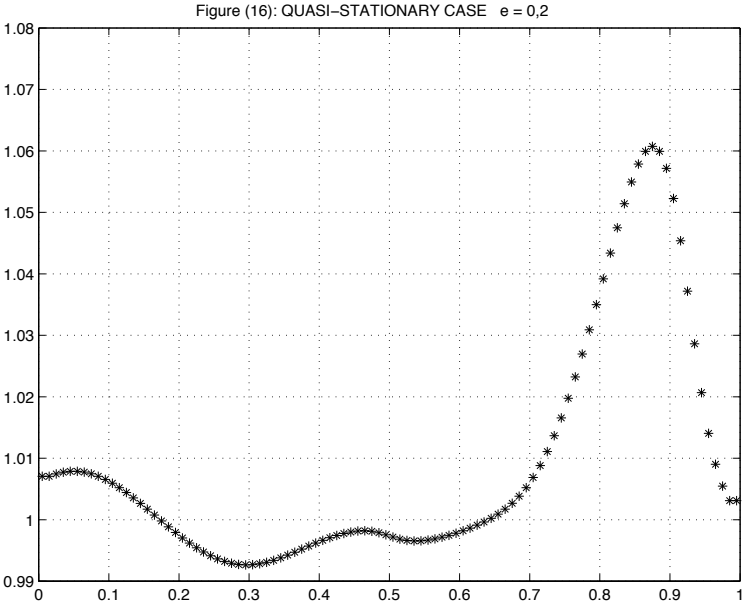


Figure (15) and Figure (16) show the water surface given by the numerical solution at time T = $0.7s$ (the usual simplification g=1 is also assumed in this example), respectively for e = 10^{-3} and e = 0.2.
We only remark that perturbations in quasi-steady problems are computed by our scheme with the same resolution as would be expected if calculating small perturbations about constant states for the homogeneous system ($\frac{\partial Z}{\partial x} = 0$).

References

[1] Audusse E., Bristeau M.O. and Perthame B., Kinetic schemes for Saint-Venant equations with source terms on unstructured grids, *INRIA Report* RR-3989 (2000)

[2] Arvanitis C., Katsaounis T. and Makridakis C., Adaptive finite element relaxation schemes for hyperbolic conservation laws, *M2AN Math. Model. Numer. Anal.*, **35** (2001), no. 1, 17-33

[3] Bermudez A. and Vasquez M.E., Upwind methods for hyperbolic conservation laws with source terms, *Comput. Fluids*, **23** (1994), no. 8, 1049-1071

[4] Botchorishvili R., Perthame B. and Vasseur A., Equilibrium Schemes for Scalar Conservation Laws with Stiff Sources, *INRIA Report* RR-8931 (2000)

[5] Dafermos C.M., *Hyperbolic conservation laws in continuum physics*, vol. GM 325 Springer-Verlag, Berlin, 1999

[6] Eymard R., Gallouët T. and Herbin R., *Finite volume methods*, Handbook of numerical analysis, vol. VIII, P.G.Ciarlet and J.L.Lions editors, Amsterdam, North-Holland, to appear

[7] Gallouët T., Hérard J.M. and Seguin N., Some approximate Godunov schemes to compute shallow-water equations with topography, *AIAA-2001* (2000)

[8] Gerbeau J.F. and Perthame B., Derivation of viscous Saint-Venant system for laminar shallow water; numerical validation, *Discrete Contin. Dynam. Systems*, to appear

[9] Godlewski E. and Raviart P.A., *Numerical approximation of hyperbolic systems of conservation laws*, Applied Mathematical Sciences **118**, New York, Springer-Verlag, 1996

[10] Gosse L. and LeRoux A.Y., A well-balanced scheme designed for inhomogeneous scalar conservation laws, *C.R. Acad. Sci. Paris Sér.I Math.*, **323** (1996), no. 5, 543-546

[11] Gosse L., A priori error estimate for a well-balanced scheme designed for inhomogeneous scalar conservation laws, *C.R. Acad. Sci. Paris Sér.I Math.*, **327** (1998), no. 5, 467-472

[12] Gosse L., A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms, *Comput. Math. Appl.*, **39** (2000), no. 9-10, 135-159

[13] Gosse L., A well-balanced scheme using non-conservative products designed for hyperbolic systems of conservation laws with source terms, *Math. Mod. Meth. Appl. Sci.*, to appear

[14] Greenberg J.M. and LeRoux A.Y., A well balanced scheme for the numerical processing of source terms in hyperbolic equations, *SIAM J. Num. Anal.*, **33** (1996), 1-16

[15] Jin S., A steady-state capturing method for hyperbolic systems with geometrical source terms, *M2AN Math. Model. Numer. Anal.*, to appear

[16] Le Vêque R.J., *Numerical Methods for Conservation Laws*, Lectures in Mathematics, ETH Zurich, Birkhauser, 1992

[17] Le Vêque R.J., Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm, *J. Comput. Phys.*, **146** (1998), no. 1, 346-365

[18] Lions P.L., Perthame B. and Tadmor E., Kinetic formulation of the Isentropic Gas Dynamics and p-Systems, *Commun. Math. Phys.*, **163** (1994), no. 2, 415-431

[19] Perthame B., An introduction to kinetic schemes for gas dynamics, *An introduction to recent developments in theory and numerics for conservation laws*, L.N. in Computational Sc. and Eng. **5**, D.Kroner, M.Ohlberger and C.Rohde editors, Springer, 1998

[20] Roe P.L., Upwind differenced schemes for hyperbolic conservation laws with source terms, *Proc. Conf. Hyperbolic Problems*, Carasso, Raviart and Serre editors, Springer, 1986, pp. 41-51

[21] de Saint-Venant A.J.C., Théorie du mouvement non-permanent des eaux, avec application aux crues des rivière at à l'introduction des marées dans leur lit, *C.R. Acad. Sci. Paris*, **73** (1871), 147-154

[22] Serre D., *Systèmes hyperboliques de lois de conservation, Tomes I et II*, Diderot Eds, Paris 1996

[23] Stroud A.H., *Numerical Quadrature and Solution of Ordinary Differential Equations*, Applied Mathematical Sciences **10**, New York-Heidlberg, Springer-Verlag, 1974, pp. 106-122

[24] Xu K., Unsplitting BGK-type schemes for the shallow water equations, *Internat. J. Modern Phys. C.*, **10** (1999), no.4, 505-516