TASK QUARTERLY 17 No 1–2, 63–90

KINETIC-INDUCED MOMENT SYSTEMS FOR THE SAINT-VENANT EQUATIONS

DIANA CRISTINA GIL MONTOYA[1,2]
AND JENS STRUCKMEIER[2]

[1]*Faculty of Applied Physics and Mathematics,*
Gdansk University of Technology,
Narutowicza 11/12, 80-233 Gdansk, Poland

[2]*Department of Mathematics, University of Hamburg,*
Bundesstrasse 55, 20146 Hamburg, Germany
dianacristina.gilmontoya@uni-hamburg.de
jens.struckmeier@uni-hamburg.de

(Received 20 December 2012; revised manuscript received 17 January 2013)

Abstract: Based on the relation between kinetic Boltzmann-like transport equations and nonlinear hyperbolic conservation laws, we derive kinetic-induced moment systems for the spatially one-dimensional shallow water equations (the Saint-Venant equations). Using Chapman-Enskog-like asymptotic expansion techniques in terms of the relaxation parameter of the kinetic equation, the resulting moment systems are asymptotically closed without the need for an additional closure relation. Moreover, the new second order moment equation for the (asymptotically) third order system may act as a monitoring function to detect shock and rarefaction waves, which we confirm by a number of numerical experiments.

Keywords: Saint-Venant equations, shallow water equations, Boltzmann equation, hyperbolic conservation laws, kinetic models and representations, relaxation systems, shock and rarefaction waves

1. Introduction

The Saint-Venant system (SVE) is the one-dimensional case of the shallow water equations, that describe the dynamics of a fluid flow under a free surface, *e.g.*, open channels, rivers, coastal areas, *etc.* Its importance stems from the fact that it is used for the prediction of tides in the ocean, floods on rivers, surges in channels, tsunamis, among other types of phenomena corresponding to shallow water flows.

On the other hand, analytical solutions of the system are, in general, not available, since in reality droughts and topography changes take place and new

terms have to be included in the system. Hence, it has become more relevant to use accurate numerical methods involving adaptive discretization techniques. The construction of these methods becomes more efficient, if it is possible to know or to detect in advance the location, where discontinuities in the solution will occur.

Using the relation between the SVE and kinetic transport equations it is possible to compute a hierarchy of kinetic-induced moment (relaxation) systems with a relaxation parameter given by ε, such that, in the limit $\varepsilon \to 0$ the system will tend to an equilibrium given by the inviscid SVE. Using asymptotic expansions for the moments of the kinetic equation in terms of ε, we are able to derive moment systems which are automatically closed without the need for an additional closure relation. In particular, the second order moment $W(x,t)$ will behave as δ-function at discontinuities, which might be used as a monitoring function in the construction of adaptive mesh refinement (AMR) models.

The paper is organized as follows: Chapter 2 contains a short review on the Saint-Venant equations, subsequently, the definition of kinetic representations is explained, which is the basis of our work. The application to the SVE and the computation of a third order moment system (relaxation system) based on a Chapman-Enskog-like expansion with its characteristic structure will be presented in Chapter 3; it is shown that it works as a monitoring function to detect discontinuities. Finally, a number of numerical experiments for both systems are given in Chapter 4.

2. The Saint-Venant system

2.1. Preliminaries

The Saint-Venant Equations (SVE) are a nonlinear hyperbolic system of partial differential equations, that describe the evolution of a height of water $h(t,x)$ and its horizontal velocity $u(t,x)$, at a time $t \geq 0$ and at a point $x \in \mathbf{R}$. The evolution is governed (in inviscid form) by the continuity and momentum equations given by:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{1}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}\left(hu^2 + \frac{g}{2}h^2\right) = 0 \tag{2}$$

where g stands for the gravitational acceleration and hu for the flow discharge. For details on the derivation of the inviscid[1] SVE see Leveque [1] and Stoker [2].

The previous system can be written in quasilinear form as:

$$\begin{pmatrix} h \\ hu \end{pmatrix}_{\mathrm{t}} + M(h,u)\begin{pmatrix} h \\ hu \end{pmatrix}_{\mathrm{x}} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3}$$

where the matrix $M(h,u)$ is the Jacobian of the flux matrix, and is given by:

$$M(h,u) = \begin{pmatrix} 0 & 1 \\ -u^2 + gh & 2u \end{pmatrix} \tag{4}$$

1. zero viscosity

with eigenvalues $\lambda_{1,2} = u(x,t) \mp \sqrt{gh(x,t)}$, which are distinct and real as long as $h > 0$, therefore, the system is considered strictly hyperbolic. As we can see, if $h = 0$ (vanishing water height), the system is not strictly hyperbolic anymore. In our case we will have two characteristic curves and two waves corresponding to a shock or a rarefaction wave, these curves are determined by the eigenvalues of the Jacobian of the flux matrix $M(h,u)$.

The corresponding eigenvectors are given by:

$$r^1 = \begin{pmatrix} 1 \\ u - \sqrt{\frac{g}{h}} \end{pmatrix} \tag{5}$$

$$r^2 = \begin{pmatrix} 1 \\ u + \sqrt{\frac{g}{h}} \end{pmatrix} \tag{6}$$

The SVE are completed by the *entropy function* $E(h,u)$, corresponding to the total energy of the system, with $q = [h, hu]^T$:

$$E(h,u) = \underbrace{\frac{1}{2}hu^2}_{\text{kinetic energy}} + \underbrace{\frac{g}{2}h^2}_{\text{potential energy}} \tag{7}$$

This is an obvious choice since an entropy function should be a conserved quantity, in this case the energy, whenever the unknown $q(x,t)$ is smooth, but which has a source or a sink at discontinuities. In our case, the energy will decrease in an admissible shock but will increase across an expansion shock.

Consequently, we can also define an *entropy flux* together with the *entropy inequality*, which, for weak solutions, reads:

$$\frac{\partial E}{\partial t} + \frac{\partial}{\partial x}\left[u(E + \frac{g}{2}h^2)\right] \leq 0 \tag{8}$$

and becomes an equality if the solution $q = [h, hu]^T$ is smooth. This condition will be useful to guarantee the convergence to the physically correct solution in the numerical experiments. For more details on the properties and definition of entropy functions see Leveque [1].

2.2. Kinetic Representation

The *Kinetic Formulation* as defined by Lions [3], is an equivalent formulation of a conservation laws system based on an appropriate transport equation such that:

- it contains a full family of entropy inequalities,
- it involves an additional variable ξ, termed kinetic or microscopic velocity,
- its ξ-moments recover the original equations and their entropy conditions.

A weaker version of this is the so-called *Kinetic Representation*, since it uses only the single entropy coming from the total energy (the motivation for this

can be seen in [4], section 1.7), and it is an equivalent system to the SVE (1) and (2). It is based on the condition:

$$\begin{pmatrix} h \\ hu \\ hu^2+\frac{g}{2}h^2 \end{pmatrix} = \int_{\mathbf{R}} \begin{pmatrix} 1 \\ \xi \\ \xi^2 \end{pmatrix} M(h,\xi-u)\,\mathrm{d}\xi \tag{9}$$

termed the *representation formula*, which is valid under the following conditions.

Consider the function $\chi(\omega)$ on $\mathbf{R}$, together with the properties:

$$\chi(\omega)=\chi(-\omega)\geq 0 \quad \text{(even non-negative function)} \tag{10}$$

$$\int_{\mathbf{R}} \chi(\omega)\,\mathrm{d}\omega = 1 \quad \text{and} \quad \int_{\mathbf{R}} \omega^2\chi(\omega)\,\mathrm{d}\omega = \kappa \tag{11}$$

and define $M(h,\xi-u)$, the density of particles as:

$$M(h,\xi-u)=\sqrt{h}\,\chi\left(\frac{\xi-u}{\sqrt{h}}\right) \tag{12}$$

The proof of the representation formula is just a simple computation of the right hand side of (9). Now, we can check by making integrations w.r.t. ξ against weights 1 and ξ, and using the values in (9) that $Q(t,x,\xi)$, which is a collision term defined as:

$$\frac{\partial}{\partial t}M(h,\xi-u)+\xi\frac{\partial}{\partial x}M(h,\xi-u) := Q(t,x,\xi) \tag{13}$$

satisfies the conservation relations:

$$\int_{\mathbf{R}} Q\,\mathrm{d}\xi = 0 \quad \text{and} \quad \int_{\mathbf{R}} \xi Q\,\mathrm{d}\xi = 0 \tag{14}$$

It is possible to consider (13) as the limit when $\varepsilon\to 0$ of a Boltzmann-type equation with a Bhatnagar-Gross-Krook (BGK) relaxation term:

$$\frac{\partial}{\partial t}f(t,x,\xi)+\xi\frac{\partial}{\partial x}f(t,x,\xi) = \frac{1}{\varepsilon}[M(h,\xi-u)-f(t,x,\xi)] \tag{15}$$

$$h(t,x) = \int_{\mathbf{R}} f(t,x,\xi)\,\mathrm{d}\xi \tag{16}$$

$$hu(t,x) = \int_{\mathbf{R}} \xi f(t,x,\xi)\,\mathrm{d}\xi \tag{17}$$

where

$$Q = \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}[M(h,\xi-u)-f(t,x,\xi)] \tag{18}$$

As mentioned in Perthame [4], the rigorous proof of this consideration is still an open problem.

3. Kinetic induced Moment System

3.1. Moment System from the Kinetic Model

Until now, all of the previous theory is well-known and it is going to form the basis of our work. Now, we want to derive a new system based on the SVE with

the use of the kinetic representation explained in Section 2.2, which will provide useful information for numerical treatments and the general kinetic theory.

3.1.1. Boltzmann-type equation and BGK relaxation model

The Boltzmann equation describes the statistical distribution of the density of particles f which is formed by the effects of the free advection of particles (left hand side), and the collisions between them which involve the exchange of energy and momentum (right hand side). One of the methods for expressing the collision term in a simpler way (and the method used here) is the Bhatnagar-Gross-Krook (BGK) model, where $M(\rho,\xi-u)$ is termed an *equilibrium function*, $f(t,x,\xi)$ is the density of particles, which at time t and position x moves with a velocity ξ, and $\varepsilon \ll 1$ is the mean free path[2]:

$$\frac{\partial}{\partial t} f(t,x,\xi) + \xi \frac{\partial}{\partial x} f(t,x,\xi) = \frac{1}{\varepsilon}[M(h,\xi-u) - f(t,x,\xi)] \tag{19}$$

A more in-depth study on the Boltzmann equation and the BGK model can be found in Struchtrup [5].

3.1.2. Moments of $f(t,x,\xi)$

The moments derived from the Boltzmann equation are an alternative description for f, since f is difficult to compute and contains a lot of information that is not necessarily useful. The moments are weighted averages of the distribution function, much more approachable than f itself, and form a complete set of partial differential equations (see [5]); not all the moments must be considered, only those that yield relevant information.

Consider the Equations (16) and (17), the zeroth and the first moments of the function $f(t,x,\xi)$ are already known to be:

$$W_0 = \int_{\mathbf{R}} f(t,x,\xi)\,\mathrm{d}\xi = h(t,x) \tag{20}$$

$$W_1 = \int_{\mathbf{R}} \xi f(t,x,\xi)\,\mathrm{d}\xi = hu(t,x) \tag{21}$$

The subsequent moments are defined as:

$$W_k = \int_{\mathbf{R}} \xi^k f(t,x,\xi)\,\mathrm{d}\xi \qquad \text{for } k = 2,3,4,\ldots \tag{22}$$

3.1.3. Equilibrium function $M(h,\xi-u)$

A system is in equilibrium, if there are no changes taking place in time and all the forces are balanced, *i.e.*, it is in a homogeneous steady state. The first three equilibrium values corresponding to the first three moments ($k=0,1,2$) of $M(h,\xi-u)$ are already given by the kinetic representation formula (9); starting

2. Average distance traveled by a particle between collisions

from this it is possible to find with a little bit of algebra the general formula for any k:

$$W_k\,|_E = \int_{\mathbf{R}} \xi^k\, M(h,\xi-u)\,\mathrm{d}\xi = h\,u^k + \frac{k(k-1)}{4}\,g\,h^2\,u^{k-2} \quad \forall k \tag{23}$$

3.1.4. Moment System

Multiplying both sides of (19) by the weights 1, ξ and ξ^k, and integrating over the microscopic velocity, we obtain the *moment equations*, which are infinitely many relations equivalent to the SVE. As a result, additional unknown quantities appear, and our aim is to find a finite number of moment equations that allow us to obtain a closed system. Following multiplication and subsequent integration we obtain:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{24}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x} W_2 = 0 \tag{25}$$

$$\frac{\partial}{\partial t} W_k + \frac{\partial}{\partial x} W_{k+1} = \frac{1}{\varepsilon}\left[h\,u^k + \frac{k(k-1)}{4}\,g\,h^2\,u^{k-2} - W_k\right] \tag{26}$$

First order non-equilibrium system

Following the technique proposed by Gil Montoya, Struchtrup and Struckmeier [6], we define the first order non-equilibrium values as the difference between the scalar variables and their equilibrium values; they will be useful since they vanish in equilibrium. These values are defined as:

$$W_k^{(1)} = W_k - W_k\,|_E \qquad \text{for } k = 2,3,4,\ldots \tag{27}$$

We can see that $W_0^{(1)} = W_1^{(1)} = 0$ and $W_2^{(1)} = W_2 - (hu^2 + \frac{g}{2}h^2)$. The corresponding first order non-equilibrium moment system is given by:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{28}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x} W_2^{(1)} + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) = 0 \tag{29}$$

$$\begin{aligned}\frac{\partial}{\partial t} W_k^{(1)} + \frac{\partial}{\partial t}(hu^k + \frac{k(k-1)}{4}\,gh^2u^{k-2}) + \frac{\partial}{\partial x} W_{k+1}^{(1)} \\ + \frac{\partial}{\partial x}(hu^{k+1} + \frac{k(k+1)}{4}\,gh^2u^{k-1}) = -\frac{1}{\varepsilon} W_k^{(1)}\end{aligned} \tag{30}$$

where the last equation holds for $k = 2,3,4,\ldots$.

The following step serves to eliminate the second time derivative in (30). After some algebraic computations we arrive at:

$$\begin{aligned}\frac{\partial}{\partial t} W_k^{(1)} + \frac{\partial}{\partial x} W_{k+1}^{(1)} + \frac{k}{4} gh^2 \frac{\partial}{\partial x}(u^{k-1}) - \frac{k(k-1)(k-2)}{4} g^2h^2u^{k-3}\frac{\partial h}{\partial x} \\ - ku^{k-3}[u^2 + \frac{(k-1)(k-2)}{4} gh]\frac{\partial}{\partial x} W_2^{(1)} = -\frac{1}{\varepsilon} W_k^{(1)}\end{aligned} \tag{31}$$

Expanding the moments in (31) according to the classical Chapman-Enskog approach in the form:

$$W_k^{(1)} = \varepsilon W_{k,1}^{(1)} + \varepsilon^2 W_{k,2}^{(1)} \qquad \text{for } k = 2,3,4,\ldots \tag{32}$$

and taking the first order terms we obtain:

$$W_{2,1}^{(1)} = -\frac{g}{2}h^2\frac{\partial u}{\partial x} \tag{33}$$

$$W_{k,1}^{(1)} = -\frac{k}{4}gh^2\frac{\partial}{\partial x}(u^{k-1}) + \frac{k(k-1)(k-2)}{4}g^2h^2u^{k-3}\frac{\partial h}{\partial x} \tag{34}$$

Equation (34) can be written as:

$$W_{k,1}^{(1)} = \frac{k(k-1)}{2}u^{k-3}[uW_{2,1}^{(1)} + \frac{(k-2)}{2}g^2h^2\frac{\partial h}{\partial x}] \tag{35}$$

Therefore, the zeroth order system yields the inviscid SVE:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{36}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}\left(hu^2 + \frac{g}{2}h^2\right) = 0 \tag{37}$$

and the first order system yields the viscous SVE:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{38}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}\left(hu^2 + \frac{g}{2}h^2\right) + \frac{\partial}{\partial x}W_2^{(1)} = 0 \tag{39}$$

where $W_2^{(1)} = \varepsilon W_{2,1}^{(1)} = -\varepsilon\frac{g}{2}h^2\frac{\partial u}{\partial x}$ is the diffusion term.

Now a third variable $W_2^{(1)}$ has appeared, but we only have two equations, in order to have a closed system we need to find a third equation.

Second order non-equilibrium system

$$W_k^{(2)} = W_k^{(1)} - \frac{k(k-1)}{2}u^{k-3}[uW_2^{(1)} + \frac{\varepsilon(k-2)}{2}g^2h^2\frac{\partial h}{\partial x}] \quad \text{for } k = 3,4,5,\ldots \tag{40}$$

with $W_k^{(1)} = \varepsilon W_{k,1}^{(1)}$.

The corresponding second order non-equilibrium moment system is given by:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{41}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}W_2^{(1)} + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) = 0 \tag{42}$$

$$\frac{\partial}{\partial t}W_2^{(1)} + \frac{\partial}{\partial x}W_3^{(1)} + \frac{g}{2}h^2\frac{\partial u}{\partial x} - 2u\frac{\partial}{\partial x}W_2^{(1)} = -\frac{1}{\varepsilon}W_2^{(1)} \tag{43}$$

$$\frac{\partial}{\partial t}W_2^{(1)} + \frac{\partial}{\partial x}W_3^{(2)} + \frac{3\varepsilon}{2}g^2\frac{\partial}{\partial x}\left(h^2\frac{\partial h}{\partial x}\right) + \left(\frac{g}{2}h^2 + 3W_2^{(1)}\right)\frac{\partial u}{\partial x} + u\frac{\partial}{\partial x}W_2^{(1)} = -\frac{1}{\varepsilon}W_2^{(1)} \tag{44}$$

Equation (44) results from (31) following the substitution $k=2$ since $W_k^{(2)}$ holds only for $k=3,4,5,\ldots$, with $W_3^{(1)}$ replaced according to (40). The final equation of the system is as follows:

$$\begin{aligned}&\frac{\partial}{\partial t}W_k^{(2)}+\frac{k(k-1)}{2}\frac{\partial}{\partial t}(u^{k-2}W_2^{(1)}+\frac{\varepsilon(k-2)}{2}g^2h^2u^{k-3}\frac{\partial h}{\partial x})+\frac{\partial}{\partial x}W_{k+1}^{(2)}\\&+\frac{k(k+1)}{2}\frac{\partial}{\partial x}(u^{k-1}W_2^{(1)}+\frac{\varepsilon(k-1)}{2}g^2h^2u^{k-2}\frac{\partial h}{\partial x})+\frac{k}{4}gh^2\frac{\partial}{\partial x}(u^{k-1})\\&-\frac{k(k-1)(k-2)}{4}g^2h^2u^{k-3}\frac{\partial h}{\partial x}-ku^{k-3}[u^2+\frac{(k-1)(k-2)}{4}gh]\frac{\partial}{\partial x}W_2^{(1)}\\&\qquad=-\frac{1}{\varepsilon}[W_k^{(2)}+\frac{k(k-1)}{2}u^{k-3}(uW_2^{(1)}+\frac{\varepsilon(k-2)}{2}g^2h^2\frac{\partial h}{\partial x})]\end{aligned}\tag{45}$$

which holds only for $k=3,4,5,\ldots$.

The following step serves to eliminate the second time derivative in (45). After some algebraic computations we arrive at:

$$\begin{aligned}&\frac{\partial}{\partial t}W_k^{(2)}+\frac{\partial}{\partial x}W_{k+1}^{(2)}-\frac{k(k-1)}{2}u^{k-2}\frac{\partial}{\partial x}W_3^{(2)}\\&-\frac{k(k-1)(k-2)}{2h}u^{k-3}\left[\left(\frac{\partial}{\partial x}W_2^{(1)}+gh\frac{\partial h}{\partial x}\right)\left(\varepsilon(k-3)\frac{g^2}{2}h^2u^{-1}\frac{\partial h}{\partial x}+W_2^{(1)}\right)\right.\\&\left.+\frac{g}{2}h^2\left(\varepsilon gh^2\frac{\partial^2u}{\partial x^2}+\frac{\partial}{\partial x}W_2^{(1)}\right)\right]=-\frac{1}{\varepsilon}W_k^{(2)}\end{aligned}\tag{46}$$

For convenience we want to indicate the variables that are already expanded, and thus we write $h=\widehat{h}$, $u=\widehat{u}$ and $W_2^{(1)}=\varepsilon\widehat{W_2^{(1)}}$. The system now reads:

$$\frac{\partial\widehat{h}}{\partial t}+\frac{\partial}{\partial x}(\widehat{h}\widehat{u})=0\tag{47}$$

$$\frac{\partial}{\partial t}(\widehat{h}\widehat{u})+\varepsilon\frac{\partial}{\partial x}\widehat{W_2^{(1)}}+\frac{\partial}{\partial x}(\widehat{h}\widehat{u}^2+\frac{g}{2}\widehat{h}^2)=0\tag{48}$$

$$\begin{aligned}&\frac{\partial}{\partial t}\varepsilon\widehat{W_2^{(1)}}+\frac{\partial}{\partial x}W_3^{(2)}+\frac{3\varepsilon}{2}g^2\frac{\partial}{\partial x}\left(\widehat{h}^2\frac{\partial\widehat{h}}{\partial x}\right)+\left(\frac{g}{2}\widehat{h}^2+3\varepsilon\widehat{W_2^{(1)}}\right)\frac{\partial\widehat{u}}{\partial x}\\&\qquad+\widehat{u}\frac{\partial}{\partial x}\varepsilon\widehat{W_2^{(1)}}=-\widehat{W_2^{(1)}}\end{aligned}\tag{49}$$

$$\begin{aligned}&\frac{\partial}{\partial t}W_k^{(2)}+\frac{\partial}{\partial x}W_{k+1}^{(2)}-\frac{k(k-1)}{2}\widehat{u}^{k-2}\frac{\partial}{\partial x}W_3^{(2)}\\&-\frac{k(k-1)(k-2)}{2\widehat{h}}\widehat{u}^{k-3}\left[\left(\frac{\partial}{\partial x}\varepsilon\widehat{W_2^{(1)}}+g\widehat{h}\frac{\partial\widehat{h}}{\partial x}\right)\left(\varepsilon(k-3)\frac{g^2}{2}\widehat{h}^2\widehat{u}^{-1}\frac{\partial\widehat{h}}{\partial x}\right.\right.\\&\left.\left.+\varepsilon\widehat{W_2^{(1)}}\right)+\frac{g}{2}\widehat{h}^2\left(\varepsilon g\widehat{h}^2\frac{\partial^2\widehat{u}}{\partial x^2}+\frac{\partial}{\partial x}\varepsilon\widehat{W_2^{(1)}}\right)\right]=-\frac{1}{\varepsilon}W_k^{(2)}\end{aligned}\tag{50}$$

Again, expanding the moments in (50) in the form:

$$W_k^{(2)}=\varepsilon^2W_{k,2}^{(2)}+\varepsilon^3W_{k,3}^{(2)}\qquad\text{for } k=3,4,5,\ldots\tag{51}$$

and taking the first order terms we obtain:

$$W_{3,2}^{(2)}=\frac{3}{2}g\left[2\widehat{W_2^{(1)}}\frac{\partial\widehat{h}}{\partial x}+\widehat{h}\frac{\partial}{\partial x}\widehat{W_2^{(1)}}+g\widehat{h}^3\frac{\partial^2\widehat{u}}{\partial x^2}\right]\tag{52}$$

$$W_{k,2}^{(2)} = \frac{k(k-1)(k-2)}{4} g\widehat{u}^{k-3}\left[2\widehat{W_2^{(1)}}\frac{\partial\widehat{h}}{\partial x} + \widehat{h}\frac{\partial}{\partial x}\widehat{W_2^{(1)}} + g\widehat{h}^3\frac{\partial^2\widehat{u}}{\partial x^2} + (k-3)g^2\widehat{h}^2\widehat{u}^{-1}\left(\frac{\partial\widehat{h}}{\partial x}\right)^2\right] \tag{53}$$

Equation (53) can be written as:

$$W_{k,2}^{(2)} = \frac{k(k-1)(k-2)}{4}\widehat{u}^{k-3}\left[\frac{2}{3}W_{3,2}^{(2)} + (k-3)g^3\widehat{h}^2\widehat{u}^{-1}\left(\frac{\partial\widehat{h}}{\partial x}\right)^2\right] \tag{54}$$

Third order system

If we now define $W_3^{(2)} = \varepsilon^2\widehat{W_3^{(2)}}$, since it is already expanded, we can obtain a third order system using (47), (48) and (49):

$$\frac{\partial\widehat{h}}{\partial t} + \frac{\partial}{\partial x}(\widehat{h}\widehat{u}) = 0 \tag{55}$$

$$\frac{\partial}{\partial t}(\widehat{h}\widehat{u}) + \varepsilon\frac{\partial}{\partial x}\widehat{W_2^{(1)}} + \frac{\partial}{\partial x}(\widehat{h}\widehat{u}^2 + \frac{g}{2}\widehat{h}^2) = 0 \tag{56}$$

$$\frac{\partial}{\partial t}\widehat{W_2^{(1)}} + \varepsilon\frac{\partial}{\partial x}\widehat{W_3^{(2)}} + \frac{3}{2}g^2\frac{\partial}{\partial x}\left(\widehat{h}^2\frac{\partial\widehat{h}}{\partial x}\right) + \left(\frac{g}{2\varepsilon}\widehat{h}^2 + 3\widehat{W_2^{(1)}}\right)\frac{\partial\widehat{u}}{\partial x} + \widehat{u}\frac{\partial}{\partial x}\widehat{W_2^{(1)}} = -\frac{1}{\varepsilon}\widehat{W_2^{(1)}} \tag{57}$$

with

$$\widehat{W_3^{(2)}} = \frac{3}{2}g\left[2\widehat{W_2^{(1)}}\frac{\partial\widehat{h}}{\partial x} + \widehat{h}\frac{\partial}{\partial x}\widehat{W_2^{(1)}} + g\widehat{h}^3\frac{\partial^2\widehat{u}}{\partial x^2}\right] \tag{58}$$

3.2. Moment System of third order

For simplicity we re-write $\widehat{h} = h$, $\widehat{u} = u$, $\widehat{W_2^{(1)}} = W$ and replace the value of $\widehat{W_3^{(2)}}$ found in the previous section. In this way we shall arrive at a third order closed system (showed below), which will be the focus of the present work:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{59}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) + \varepsilon\frac{\partial W}{\partial x} = 0 \tag{60}$$

$$\frac{\partial W}{\partial t} + \left(\frac{g}{2\varepsilon}h^2 + 3W\right)\frac{\partial u}{\partial x} + u\frac{\partial W}{\partial x} = -\frac{3}{2}\varepsilon g\frac{\partial}{\partial x}\left[\frac{g}{\varepsilon}h^2\frac{\partial h}{\partial x} + 2W\frac{\partial h}{\partial x} + h\frac{\partial W}{\partial x} + gh^3\frac{\partial^2 u}{\partial x^2}\right] - \frac{1}{\varepsilon}W \tag{61}$$

Equation (61) can be written in balance form as:

$$\frac{\partial}{\partial t}(\varepsilon W + hu^2 + \frac{g}{2}h^2) + \frac{\partial}{\partial x}(3\varepsilon uW + hu^3 + \frac{3g}{2}uh^2) = -\frac{3g}{2}\varepsilon^2\frac{\partial}{\partial x}\left[\frac{g}{\varepsilon}h^2\frac{\partial h}{\partial x} + 2W\frac{\partial h}{\partial x} + h\frac{\partial W}{\partial x} + gh^3\frac{\partial^2 u}{\partial x^2}\right] - W \tag{62}$$

3.2.1. Inviscid system

The respective homogeneous-inviscid system for (59), (60) and (62) is:

$$\begin{pmatrix} h \\ hu \\ \varepsilon W + hu^2 + \frac{g}{2}h^2 \end{pmatrix}_t + M(h,u,W) \begin{pmatrix} h \\ hu \\ \varepsilon W + hu^2 + \frac{g}{2}h^2 \end{pmatrix}_x = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \tag{63}$$

with the Jacobian of the flux matrix equal to

$$M(h,u,W) = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -3\frac{u(\varepsilon W + hu^2 + \frac{g}{2}h^2)}{h} + 4u^3 & 3\frac{\varepsilon W + hu^2 + \frac{g}{2}h^2}{h} - 6u^2 & 3u \end{pmatrix} \tag{64}$$

The system is strictly hyperbolic for $h > 0$, and the eigenvalues of $M(h,u,W)$ are given by:

$$\lambda_1 = u - \sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}} \tag{65}$$

$$\lambda_2 = u \tag{66}$$

$$\lambda_3 = u + \sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}} \tag{67}$$

Hence, we will have 3 characteristic curves and 3 waves, where information can travel at the fluid velocity, or move as acoustic waves at speeds $c = \mp\sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}}$. The 1-wave will move to the left, while the 3-wave will move to the right; between these two waves the velocity is constant (u_m), and the 2-wave will appear with velocity $\lambda_2 = u_m$.

Then the corresponding eigenvectors will read:

$$r^1 = \begin{pmatrix} 1 \\ u - \sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}} \\ (u - \sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}})^2 \end{pmatrix} \tag{68}$$

$$r^2 = \begin{pmatrix} 1 \\ u \\ u^2 \end{pmatrix} \tag{69}$$

$$r^3 = \begin{pmatrix} 1 \\ u + \sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}} \\ (u + \sqrt{\frac{3g}{2}h + 3\varepsilon\frac{W}{h}})^2 \end{pmatrix} \tag{70}$$

Shock and Rarefaction waves

The characteristic fields 1 and 3 are *genuinely non-linear* since $\nabla\lambda^p \cdot r^p \neq 0$, for $p = 1,3$; the 1-wave and 3-wave will deform into shock or rarefaction waves, given that λ^1 and λ^3 vary along the integral curves of r^1 and r^3, respectively. The behavior of both waves is similar to the one described in Chapter 2 for the Saint-Venant System.

For the system to have physical meaning, the weak solution must satisfy the entropy condition (8).

Contact discontinuity

In contrast to fields 1 and 3, $\nabla\lambda^2 \cdot r^2 = 0$ and the second field is termed a *linearly degenerate* field. Neither shock nor rarefaction waves can occur in the 2-characteristic field, instead *contact discontinuities* appear.

Contact discontinuities are linear discontinuities that propagate to the characteristic speed (66) on each side, without distorting, given that $\lambda_2 = u$ is constant along the integral curve of r^2; the height will jump on the discontinuity as will variables that depend on $h(x,t)$. Characteristics will be parallel to the wave in the $x-t$ plane, rather than impinging on it.

3.2.2. Shock and rarefaction waves detector

In the formal limit $\varepsilon \to 0$, the system (59)–(61) reads:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{71}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) = 0 \tag{72}$$

$$W = -\frac{g}{2}h^2\frac{\partial u}{\partial x} \tag{73}$$

Therefore, as ε goes to zero, and considering that $h > 0$, we have that if:

$$\frac{\partial u}{\partial x} = 0 \qquad W = 0 \tag{74}$$

$$\frac{\partial u}{\partial x} < 0 \qquad W > 0 \tag{75}$$

$$\frac{\partial u}{\partial x} > 0 \qquad W < 0 \tag{76}$$

Additionally, we could distinguish between shock and rarefaction waves, since:

$$\frac{\partial u}{\partial x} \to -\infty \qquad W \to \infty \qquad \text{(shock wave)}$$

$$\frac{\partial u}{\partial x} \to \infty \qquad W \to -\infty \qquad \text{(rarefaction wave)}$$

As we saw in Section 2.1, the generic solution for the shallow water system consists of two waves – depending on the initial data, each one will be either a shock wave or a rarefaction wave.

In order to prove that our third order moment system (61) may act as an indicator for the detection of shock and rarefaction waves, consider a discontinuity located at $s(t)$ with $h(x,0) \equiv h_0$ and $\big(u(s(t^-),t) = -u(s(t^+),t)\big)$, $u(s(t^-),t) > 0$; such that we have a two-shock Riemann solution. The speed of the shock waves will be given by the Rankine-Hugoniot conditions.

Hence, from (73) we can infer that for $\varepsilon \ll 1$ the solution of $W(x,t)$ tends to a δ-function located at the point of the discontinuity $s(t)$. Scaling the variable x as:

$$\eta = \frac{x - s(t)}{\varepsilon} \tag{77}$$

and introducing the scaled functions:

$$\widehat{h}(\eta,t) = h(\varepsilon\eta + s(t),t) \tag{78}$$

$$\widehat{u}(\eta,t) = u(\varepsilon\eta + s(t),t) \tag{79}$$

$$\widehat{W}(\eta,t) = \varepsilon W(\varepsilon\eta + s(t),t) \tag{80}$$

we turn the system (59)–(61) turn into:

$$\varepsilon\frac{\partial\widehat{h}}{\partial t} + (\widehat{u} - \dot{s}(t))\frac{\partial\widehat{h}}{\partial\eta} + \widehat{h}\frac{\partial\widehat{u}}{\partial\eta} = 0 \tag{81}$$

$$\varepsilon\frac{\partial}{\partial t}(\widehat{h}\widehat{u}) + (\widehat{u} - \dot{s}(t))\frac{\partial}{\partial\eta}(\widehat{h}\widehat{u}) + \widehat{h}\frac{\partial}{\partial\eta}(\frac{1}{2}u^2 + g\widehat{h}) + \frac{\partial\widehat{W}}{\partial\eta} = 0 \tag{82}$$

$$\begin{aligned}\varepsilon\frac{\partial\widehat{W}}{\partial t} + (\widehat{u} - \dot{s}(t))\frac{\partial\widehat{W}}{\partial\eta} + \left(\frac{g}{2}\widehat{h}^2 + 3\widehat{W}\right)\frac{\partial\widehat{u}}{\partial\eta} = -\frac{3g}{2}\frac{\partial}{\partial\eta}\Big[g\widehat{h}^2\frac{\partial\widehat{h}}{\partial\eta} + \\ 2\widehat{W}\frac{\partial\widehat{h}}{\partial\eta} + \widehat{h}\frac{\partial\widehat{W}}{\partial\eta} + g\widehat{h}^3\frac{\partial^2\widehat{u}}{\partial\eta^2}\Big] - \widehat{W}\end{aligned} \tag{83}$$

Again, taking the formal limit $\varepsilon \to 0$, we obtain the leading order of (81) and (82) as:

$$\frac{\partial\widehat{u}}{\partial\eta} = (\dot{s}(t) - \widehat{u})\frac{1}{\widehat{h}}\frac{\partial\widehat{h}}{\partial\eta} \tag{84}$$

$$\frac{\partial\widehat{W}}{\partial\eta} = (\dot{s}(t) - \widehat{u})\widehat{h}\frac{\partial\widehat{u}}{\partial\eta} - g\widehat{h}\frac{\partial\widehat{h}}{\partial\eta} \tag{85}$$

Moreover, Equations (59) and (60) can be written as:

$$\frac{\partial h}{\partial t} + \frac{\partial}{\partial x}(hu) = 0 \tag{86}$$

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) + \frac{1}{\varepsilon}\frac{\partial\widehat{W}}{\partial\eta} = 0 \tag{87}$$

since $\frac{\partial W}{\partial x} = \frac{1}{\varepsilon^2}\frac{\partial\widehat{W}}{\partial\eta}$. Now, replacing (85) into (87), we obtain:

$$\frac{\partial}{\partial t}(hu) + \frac{\partial}{\partial x}(hu^2 + \frac{g}{2}h^2) + \frac{1}{\varepsilon}(\dot{s}(t) - \widehat{u})\widehat{h}\frac{\partial\widehat{u}}{\partial\eta} - g\widehat{h}\frac{\partial\widehat{h}}{\partial\eta} = 0 \tag{88}$$

$$\frac{\partial}{\partial t}(hu) + u\frac{\partial}{\partial x}(hu) + \dot{s}(t)h\frac{\partial u}{\partial x} = 0 \tag{89}$$

$$u\frac{\partial h}{\partial t} + h\frac{\partial u}{\partial t} + u\frac{\partial}{\partial x}(hu) + \dot{s}(t)h\frac{\partial u}{\partial x} = 0 \tag{90}$$

And replacing $\frac{\partial h}{\partial t}$ according to (86) leads to:

$$\frac{\partial u}{\partial t} + \dot{s}(t)\frac{\partial u}{\partial x} = 0 \tag{91}$$

This demonstrate that in the limit $\varepsilon \to 0$, the system (59)–(61), yields the correct shock propagation.

4. Numerical example

We will use Godunov's Method together with the Roe linearisation and the Harten-Hyman Entropy Fix method (HHE) to get the numerical results from both of our systems: the inviscid Saint-Venant Equations (1) and (2), and the third order moment system (59), (60), (62). The finite volume method is implemented inside the Clawpack software for conservation laws [7].

The computational domain is given by $[-15,15]$ with $M = 500$ cells in the x direction, with an initial time step $\Delta t = 0.1$; the time step would be updated accordingly with the desired cfl condition $cfl = 0.9$ and the maximum value permitted $cfl\_max = 1$. The final time will be $t\_final = 3$.

4.1. Approximate Riemann Solvers

Define a function $\widehat{Q}_{i-\frac{1}{2}}(x/t)$ that approximates the true similarity solution of the Riemann problem with data Q_{i-1} and Q_i. The obvious choice is to use an approximation based on the solution of the linear problem, therefore, this function will consist of a set of m waves $W^p_{i-\frac{1}{2}}$ propagating at some speed $s^p_{i-\frac{1}{2}}$:

$$Q_i - Q_{i-1} = \sum_{p=1}^{m} W^p_{i-\frac{1}{2}} \tag{92}$$

Hence, we can use the waves and speed from the approximative solution to define:

$$A^- \Delta Q_{i-\frac{1}{2}} = \sum_{p=1}^{m} (s^p_{i-\frac{1}{2}})^- W^p_{i-\frac{1}{2}} \tag{93}$$

$$A^+ \Delta Q_{i-\frac{1}{2}} = \sum_{p=1}^{m} (s^p_{i-\frac{1}{2}})^+ W^p_{i-\frac{1}{2}} \tag{94}$$

where the minus and plus sign in the speed represents the $min(\lambda,0)$ and the $max(\lambda,0)$ respectively. We can subsequently use the fluctuations in the Godunov's method:

$$Q^{n+1}_i = Q^n_i - \frac{\Delta t}{\Delta x}\left[A^+ \Delta Q_{i-\frac{1}{2}} + A^- \Delta Q_{i+\frac{1}{2}}\right] \tag{95}$$

4.1.1. Linearised Riemann Solvers and the Roe Linearisation

Replace the non-linear problem by a linear one, in which at each cell interface we get:

$$\frac{\partial \widehat{q}}{\partial t} + \widehat{A}_{i-\frac{1}{2}} \frac{\partial \widehat{q}}{\partial x} = 0 \tag{96}$$

The matrix $\widehat{A}_{i-\frac{1}{2}}$:

- must be diagonalizable, with real eigenvalues,
- must satisfy $\widehat{A}_{i-\frac{1}{2}} \to f'(\overline{q})$ as $Q_{i-1}, Q_i \to \overline{q}$, since the new matrix is an approximation of the original Jacobian matrix in a neighborhood of Q_{i-1}, Q_i.

Hence, the approximate Riemann solution consists of m waves proportional to the eigenvectors $\widehat{r}^p_{i-\frac{1}{2}}$ of the matrix $\widehat{A}_{i-\frac{1}{2}}$, propagating with a speed equal to the eigenvalues $\widehat{s}^p_{i-\frac{1}{2}} = \widehat{\lambda} i - \frac{1}{2}^p$. Now, it is possible to solve the linear system:

$$Q_i - Q_{i-1} = \sum_{p=1}^{m} \alpha^p_{i-\frac{1}{2}} \widehat{r}^p_{i-\frac{1}{2}} \tag{97}$$

in order to obtain the values of $\alpha^p_{i-\frac{1}{2}}$ and compute the waves $W^p_{i-\frac{1}{2}} = \alpha^p_{i-\frac{1}{2}} \widehat{r}^p_{i-\frac{1}{2}}$. There are many choices for the linearised matrix, one of them and the one studied in this thesis is the Roe Linearisation, which for the SVE problem will be demonstrated to display certain desired properties.

In the case when the exact Riemann solution consists of a single shock, the approximate solution should agree with the exact one. Therefore it must hold that: "if Q_{i-1} and Q_i are connected by a single wave $W^p = Q_i - Q_{i-1}$ in the exact Riemann solution, then W^p should also be an eigenvector of $\widehat{A}_{i-\frac{1}{2}}$". Consequently, the Rankine-Hugoniot condition with $q_l = Q_{i-1}$ and $q_r = Q_i$ must be satisfied by both problems at the cell interface. Then:

$$f(Q_i) - f(Q_{i-1}) = \dot{s}(Q_i - Q_{i-1}) = \widehat{A}_{i-\frac{1}{2}}(Q_i - Q_{i-1}) \tag{98}$$

since the single shock is an eigenvector of the linearised matrix. This property will guarantee that (93) and (94) yield a conservative method, considering that the following condition is satisfied:

$$\left[A^- \Delta Q_{i-\frac{1}{2}} + A^+ \Delta Q_{i-\frac{1}{2}}\right] = f(Q_i) - f(Q_{i-1}) \tag{99}$$

A way to obtain a suitable matrix $\widehat{A}_{i-\frac{1}{2}}$ is to integrate the Jacobian matrix over an appropriate path in state space[3] between Q_{i-1} and Q_i:

$$\widehat{A}_{i-\frac{1}{2}} = \int_0^1 f'(q(\zeta))\mathrm{d}\zeta \tag{100}$$

where $q(\zeta) = Q_{i-1} + (Q_i - Q_{i-1})\zeta$ is a straight-line path parametrized for $0 \le \zeta \le 1$; nevertheless, the previous integral is not always easy to calculate. We can introduce a change of variable to facilitate this integration.

Assume that $z(q)$ is an invertible mapping so we know $q(z)$ as well and f as a function of z. The integration path will become $z(\zeta) = Z_{i-1} + (Z_i - Z_{i-1})\zeta$, where $Z_j = z(Q_j)$ for $j = i-1, i$. We can now write:

3. Graph of each possible state (h, hu) for the SVE

$$f(Q_i)-f(Q_{i-1})=\int_0^1 \frac{df(z(\zeta))}{d\zeta}d\zeta=\Big[\int_0^1 \frac{df(z(\zeta))}{dz}d\zeta\Big](Z_i-Z_{i-1})= \widehat{C}_{i-\frac{1}{2}}(Z_i-Z_{i-1}) \tag{101}$$

$$Q_i-Q_{i-1}=\int_0^1 \frac{dq(z(\zeta))}{d\zeta}d\zeta=\Big[\int_0^1 \frac{dq(z(\zeta))}{dz}d\zeta\Big](Z_i-Z_{i-1})= \widehat{B}_{i-\frac{1}{2}}(Z_i-Z_{i-1}) \tag{102}$$

replacing the left hand side of (101) by the right hand side of (98) and using (102) we obtain our desired matrix $\widehat{A}_{i-\frac{1}{2}}=\widehat{C}_{i-\frac{1}{2}}\widehat{B}^{-1}_{i-\frac{1}{2}}$.

4.1.2. The Harten-Hyman Entropy Fix

In the case of transonic rarefaction waves ($f'(q_l)<0<f'(q_r)$), the entropy condition will be violated and the use of (93) and (94) would lead to a wrong solution. Therefore, in the case where $\lambda^p<0$ to the left of the wave while $\lambda^p>0$ to the right of the wave, it is necessary to modify the fluctuations by performing an *entropy fix*.

Define a transonic rarefaction in the k^{th} wave as ($\lambda^k_l<0<\lambda^k_r$), where $\lambda^k_{l,r}$ is the k^{th} eigenvalue of the Jacobian matrix computed in the states $q^k_{l,r}$ to right and left:

$$q^k_l=Q_{i-1}+\sum_{p=1}^{k-1}W^p, \qquad q^k_r=q^k_l+W^k \tag{103}$$

where $W^p=W^p_{i-\frac{1}{2}}$.

We want to replace the single wave W^k with speed $\widehat{\lambda}^k$ by two waves; $W^k_l=\beta W^k$ propagating at speed λ^k_l and $W^k_r=(1-\beta)W^k$ propagating at speed λ^k_r. In order to retain conservation, it must hold that $\lambda^k_l W^k_l+\lambda^k_r W^k_r=\widehat{\lambda}^k W^k$, and we can compute β as:

$$\beta=\frac{\lambda^k_r-\widehat{\lambda}^k}{\lambda^k_r-\lambda^k_l} \tag{104}$$

A more practical way to adjust the fluctuations in (93) and (94) is to replace the positive and negative parts of the speeds ($\widehat{s}^k=\widehat{\lambda}^k$) in the k^{th} field by:

$$(\widehat{\lambda}^k)^-\equiv\beta\lambda^k_l \tag{105}$$

$$(\widehat{\lambda}^k)^+\equiv(1-\beta)\lambda^k_r \tag{106}$$

which adds up to $\widehat{\lambda}^k$ and will be non-zero in the transonic case.

For details on the Finite Volume Method and the Roe Linearisation the Reader is referred to Leveque [1].

4.2. Inviscid case of the SVE

The numerical results for this hyperbolic system are included in the shallow water Clawpack package, here we use the same Roe Solver as used in Chapter 15 of [1].

4.2.1. Riemann Solver

Roe Linearisation

Using $z = h^{-\frac{1}{2}} q$ as a parameter vector, we can define:

$$\begin{bmatrix} z^1 \\ z^2 \end{bmatrix} = \begin{bmatrix} \sqrt{h} \\ \sqrt{h}u \end{bmatrix} \tag{107}$$

then

$$q(z) = \begin{bmatrix} (z^1)^2 \\ z^1 z^2 \end{bmatrix} \Longrightarrow \frac{\partial q}{\partial z} = \begin{bmatrix} 2z^1 & 0 \\ z^2 & z^1 \end{bmatrix} \tag{108}$$

In a similar way we calculate:

$$f(z) = \begin{bmatrix} z^1 z^2 \\ (z^2)^2 + \frac{1}{2} g (z^1)^4 \end{bmatrix} \Longrightarrow \frac{\partial f}{\partial z} = \begin{bmatrix} z^2 & z^1 \\ 2g(z^1)^3 & 2z^2 \end{bmatrix} \tag{109}$$

obtaining

$$\widehat{A}_{i-\frac{1}{2}} = \begin{bmatrix} 0 & 1 \\ -\widehat{u}^2 + g\overline{h} & 2\widehat{u} \end{bmatrix} \tag{110}$$

where $\overline{h}$ is the arithmetic average and $\widehat{u}$ is the *Roe Average*:

$$\overline{h} = \frac{1}{2}(h_{i-1} + h_i) \tag{111}$$

$$\widehat{u} = \frac{\sqrt{h_{i-1}}u_{i-1} + \sqrt{h_i}u_i}{\sqrt{h_{i-1}} + \sqrt{h_i}} \tag{112}$$

The eigenvalues and the eigenvectors of the newly obtained matrix would be given by:

$$\widehat{\lambda}^1 = \widehat{u} - \widehat{c} \qquad \widehat{\lambda}^2 = \widehat{u} + \widehat{c} \tag{113}$$

$$\widehat{r}^1 = \begin{bmatrix} 1 \\ \widehat{u} - \widehat{c} \end{bmatrix} \qquad \widehat{r}^2 = \begin{bmatrix} 1 \\ \widehat{u} + \widehat{c} \end{bmatrix} \tag{114}$$

with $\widehat{c} = \sqrt{g\overline{h}}$. Then:

$$Q_i - Q_{i-1} = \begin{bmatrix} \delta^1 \\ \delta^2 \end{bmatrix} = \alpha^1_{i-\frac{1}{2}} \widehat{r}^1 + \alpha^2_{i-\frac{1}{2}} \widehat{r}^2 = W^1_{i-\frac{1}{2}} + W^2_{i-\frac{1}{2}} \tag{115}$$

Inverting the matrix of right eigenvectors we can compute the α coefficients by solving the linear system to the left, obtaining:

$$\alpha^1_{i-\frac{1}{2}} = \frac{(\widehat{u} + \widehat{c})\delta^1 - \delta^2}{2\widehat{c}} \tag{116}$$

$$\alpha^2_{i-\frac{1}{2}} = \frac{-(\widehat{u} - \widehat{c})\delta^1 + \delta^2}{2\widehat{c}} \tag{117}$$

We now have all the information to compute the waves and to update the fluctuations (93) and (94) in the Godunov's Method using the eigenvalues as the speeds:

$$s_l = \widehat{u} - \widehat{c} \tag{118}$$

$$s_r = \widehat{u} + \widehat{c} \tag{119}$$

Sonic Entropy Fix

We want to identify the transonic rarefactions in order to modify the fluctuations. Consider hr and ur – the values to the right of the discontinuity; and

hl, ul – the values to the left. For simplicity we do not use subscripts to denote the right and the left states, since later we must include the position on the grid as a subscript.

- **Check the 1-wave:**
  First we check the 1-wave (with velocity λ_1) in the left state $(i-1)$:

$$s_0 = ur_{i-1} - \sqrt{g * hr_{i-1}} \tag{120}$$

  If s_0 and s_l are greater than zero then there are no waves propagating to the left and we have a fully supersonic case; otherwise, we must verify if transonic waves occur. Define s_1 as the speed corresponding to λ_1 to the right of the 1-wave:

$$\begin{bmatrix} hr_1 \\ hur_1 \end{bmatrix} = \begin{bmatrix} hr_{i-1} \\ hr_{i-1} * ur_{i-1} \end{bmatrix} + W^1_{i-\frac{1}{2}} \tag{121}$$

$$s_1 = \frac{hur_1}{hr_1} - \sqrt{g * hr_1} \tag{122}$$

  If $s_0 < 0$ and $s_1 > 0$, then there is a transonic rarefaction in the 1-wave and the new speed will be given as in (105) and (106) by:

$$s_1^- = s_0\left(\frac{s_1 - s_l}{s_1 - s_0}\right) \tag{123}$$

  On the other hand, if this is not the case we have two possibilities; that $s_l < 0$ and the 1-wave is propagating to the left with $s_1^- = s_l$ or if none of the previous conditions hold, then the wave is propagating to the right.

- **Check the 2-wave:**
  We now check the 2-wave corresponding to λ_2 in the right state i:

$$s_3 = ul_i - \sqrt{g * hl_i} \tag{124}$$

  Define s_2 as the speed corresponding to λ_2 to the left of the 2-wave:

$$\begin{bmatrix} hl_2 \\ hul_2 \end{bmatrix} = \begin{bmatrix} hl_i \\ hl_i * ul_i \end{bmatrix} + W^2_{i-\frac{1}{2}} \tag{125}$$

$$s_2 = \frac{hul_2}{hl_2} - \sqrt{g * hl_2} \tag{126}$$

  If $s_2 < 0$ and $s_3 > 0$, then there is a transonic rarefaction in the 2-wave and the new speed will be given as in (105) and (106) by:

$$s_2^- = s_2\left(\frac{s_3 - s_r}{s_3 - s_2}\right) \tag{127}$$

  On the other hand, if this is not the case, we have two possibilities; that $s_r < 0$ and the 2-wave is propagating to the left with $s_2^- = s_r$ or, if none of the previous conditions hold, then it is propagating to the right.

In order to compute the fluctuations, we must define the total flux difference which is $td = \sum_{p=1}^{m} s^p_{i-\frac{1}{2}} W^p_{i-\frac{1}{2}}$. Then, finally:

$$A^- \Delta Q_{i-\frac{1}{2}} = s_1^- W^1_{i-\frac{1}{2}} + s_2^- W^2_{i-\frac{1}{2}} \tag{128}$$

$$A^+ \Delta Q_{i-\frac{1}{2}} = td - A^- \Delta Q_{i-\frac{1}{2}} \tag{129}$$

4.2.2. Initial and boundary conditions

For the boundary conditions Clawpack has already a number of options that can be specified in the input file; in our case we will apply a *non-reflecting outflow* using a zero-order extrapolation, in the following manner:

$$\begin{aligned} Q^n_0 &= Q^n_1 \qquad & Q^n_{-1} &= Q^n_1 \\ Q^n_{M+1} &= Q^n_M \qquad & Q^n_{M+2} &= Q^n_M \end{aligned} \tag{130}$$

where the interior values are given by $Q^n_1, \ldots, Q^n_M$, and the boundaries are represented by $x_1 = -15$ and $x_{M+1} = 15$. Since the solution will be comprised of waves propagating to the right and left, this approach will allow a non-incoming signal, avoiding spurious reflections in the left and right boundaries.

The computational domain is extended with two *ghost cells* at the end of both boundaries, in order to have the neighboring points needed for the computation of the fluxes. Furthermore, these cells will be useful for the discretization of the derivatives in the source term of the third order moment system. On the other hand, the initial conditions are not included in the software, this data will be supplied by us for each of the variables:

$$h(x,0) = \tan^{-1}(x) + 7.5 \quad \forall x \tag{131}$$

$$u(x,0) = 0 \quad \forall x \tag{132}$$

4.3. Moment system of the third order

By analogy with the Euler equations from gas dynamics presented in Chapter 15 of [1], it is possible to define in a similar way a Roe-solver for the homogeneous system (63). Then, the inhomogeneous part, together with the diffusion term, would be solved by applying a fractional-step method.

The same computational data as in the SVE is used, including results for three different values of the relaxation term, $\varepsilon = 0.01, 0.1, 1$.

4.3.1. Riemann Solver for the homogeneous system

Roe Linearisation

For the solution of the conservation law problem, we define $H = \frac{3\varepsilon W + hu^2 + \frac{3g}{2}h^2}{h}$ and the parameter vector as:

$$\begin{bmatrix} z^1 \\ z^2 \\ z^3 \end{bmatrix} = \begin{bmatrix} \sqrt{h} \\ \sqrt{h}u \\ \sqrt{h}H \end{bmatrix} \tag{133}$$

Then:

$$q(z)=\begin{bmatrix}(z^1)^2\\ z^1z^2\\ \frac{1}{3}z^1z^3+\frac{2}{3}(z^2)^2\end{bmatrix}\Longrightarrow\frac{\partial q}{\partial z}=\begin{bmatrix}2z^1 & 0 & 0\\ z^2 & z^1 & 0\\ \frac{1}{3}z^3 & \frac{4}{3}z^2 & \frac{1}{3}z^1\end{bmatrix} \tag{134}$$

In a similar way we calculate:

$$f(z)=\begin{bmatrix}z^1z^2\\ \frac{1}{3}z^1z^3+\frac{2}{3}(z^2)^2\\ z^2z^3\end{bmatrix}\Longrightarrow\frac{\partial f}{\partial z}=\begin{bmatrix}z^2 & z^1 & 0\\ \frac{1}{3}z^3 & \frac{4}{3}z^2 & \frac{1}{3}z^1\\ 0 & z^3 & z^2\end{bmatrix} \tag{135}$$

obtaining

$$\widehat{A}_{i-\frac{1}{2}}=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ -\widehat{u}\widehat{H}+2\widehat{u}^3 & \widehat{H}-4\widehat{u}^2 & 3\widehat{u}\end{bmatrix} \tag{136}$$

where $\widehat{u}$ and $\widehat{H}$ are the *Roe Averages*:

$$\widehat{u}=\frac{\sqrt{h_{i-1}}u_{i-1}+\sqrt{h_i}u_i}{\sqrt{h_{i-1}}+\sqrt{h_i}} \tag{137}$$

$$\widehat{H}=\frac{\sqrt{h_{i-1}}H_{i-1}+\sqrt{h_i}H_i}{\sqrt{h_{i-1}}+\sqrt{h_i}} \tag{138}$$

and $H_i=\frac{3\varepsilon W_i+h_iu_i^2+\frac{3g}{2}h_i^2}{h_i}$. The eigenvalues and the eigenvectors are given by:

$$\widehat{\lambda}^1=\widehat{u}-\widehat{c}\qquad\widehat{\lambda}^2=\widehat{u}\qquad\widehat{\lambda}^3=\widehat{u}+\widehat{c} \tag{139}$$

$$\widehat{r}^1=\begin{bmatrix}1\\ \widehat{u}-\widehat{c}\\ (\widehat{u}-\widehat{c})^2\end{bmatrix}\qquad\widehat{r}^2=\begin{bmatrix}1\\ \widehat{u}\\ (\widehat{u})^2\end{bmatrix}\qquad\widehat{r}^3=\begin{bmatrix}1\\ \widehat{u}+\widehat{c}\\ (\widehat{u}+\widehat{c})^2\end{bmatrix} \tag{140}$$

with $\widehat{c}=\sqrt{\widehat{H}-\widehat{u}^2}$. Then:

$$Q_i-Q_{i-1}=\begin{bmatrix}\delta^1\\ \delta^2\\ \delta^3\end{bmatrix}=\alpha^1_{i-\frac{1}{2}}\widehat{r}^1+\alpha^2_{i-\frac{1}{2}}\widehat{r}^2+\alpha^3_{i-\frac{1}{2}}\widehat{r}^3=W^1_{i-\frac{1}{2}}+W^2_{i-\frac{1}{2}}+W^3_{i-\frac{1}{2}} \tag{141}$$

Inverting the matrix of right eigenvectors we can compute the α coefficients by solving the linear system, to obtain:

$$\alpha^1_{i-\frac{1}{2}}=\frac{1}{2\widehat{c}^2}((\widehat{u}^2+\widehat{c}\widehat{u})\delta^1-(\widehat{c}+2\widehat{u})\delta^2+\delta^3) \tag{142}$$

$$\alpha^2_{i-\frac{1}{2}}=\frac{1}{\widehat{c}^2}((\widehat{c}^2-\widehat{u}^2)\delta^1+2\widehat{u}\delta^2-\delta^3) \tag{143}$$

$$\alpha^3_{i-\frac{1}{2}}=\frac{1}{2\widehat{c}^2}((\widehat{u}^2-\widehat{c}\widehat{u})\delta^1+(\widehat{c}-2\widehat{u})\delta^2+\delta^3) \tag{144}$$

We now have all the information to update the fluctuations (93) and (94) using the eigenvalues as the speeds.

Sonic Entropy Fix

The procedure to detect transonic rarefactions is very similar to the one described in the SVE case, with the difference that in this case we have to use the structure of the new eigenvalues for the computation of the speeds.

Also, now the system has three waves, therefore we must also check the middle one, whose speed is given by the velocity u, if it is less than zero then it will be propagating to the right, otherwise it will be propagating to the left; in this case a transonic wave will not occur.

4.3.2. Fractional-Step Method for the balance law system

In the case of Balance law equations[4], instead of solving the entire system, an easier approach is to use fractional-step methods. Consider the following system:

$$\frac{\partial q}{\partial t}+\frac{\partial}{\partial x}f(q)=\gamma(q,q_x,q_{xx},\ldots) \tag{145}$$

The idea is to split the complete system into two sub-problems that can be solved separately. The first problem will consist in the homogeneous conservation law system, which can be solved by using a finite volume method with a time step Δt, later this solution will be used as initial condition in the discretization of the second problem, which will change accordingly with the nature of the source term. Therefore, the solution of the previous system will be equivalent to solving:

$$\frac{\partial q}{\partial t}+\frac{\partial}{\partial x}f(q)=0 \tag{146}$$

$$\frac{\partial q}{\partial t}=\gamma(q,q_x,q_{xx},\ldots) \tag{147}$$

Using the numerical solution of (63) as initial condition, we can now solve the second problem resulting from the splitting of (59), (60), (62) accordingly with (146) and (147), which is given by:

$$\begin{pmatrix} h \\ hu \\ \varepsilon W+hu^2+\frac{g}{2}h^2 \end{pmatrix}_{\mathrm{t}} = \begin{pmatrix} 0 \\ 0 \\ -\frac{3g}{2}\varepsilon^2\frac{\partial}{\partial x}\left[\frac{g}{\varepsilon}h^2\frac{\partial h}{\partial x}+2W\frac{\partial h}{\partial x}+h\frac{\partial W}{\partial x}+gh^3\frac{\partial^2 u}{\partial x^2}\right]-W \end{pmatrix} \tag{148}$$

The previous system can be written in an equivalent form as:

$$\begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}_{\mathrm{t}} = \begin{pmatrix} 0 \\ 0 \\ -3g\varepsilon\frac{\partial}{\partial x}\left[\frac{\partial q_1}{\partial x}\left[q_3-\frac{(q_2)^2}{q_1}\right]+\frac{q_1}{2}\frac{\partial}{\partial x}\left[q_3-\frac{(q_2)^2}{q_1}-\frac{g}{2}(q_1)^2\right]+ \right. \\ \left. \frac{g}{2}\varepsilon(q_1)^3\frac{\partial^2}{\partial x^2}\left[\frac{q_2}{q_1}\right]\right]-\frac{1}{\varepsilon}\left[q_3-\frac{(q_2)^2}{q_1}-\frac{g}{2}(q_1)^2\right] \end{pmatrix} \tag{149}$$

4. A conservation law system with source terms

For the discretization of this system we will use an explicit method with a forward difference for the time derivative at t^n and central finite differences for the space derivatives at position x_i:

$$Q_i^{n+1} = Qi^* + \Delta t(\gamma(Qi^*)) \tag{150}$$

$$\frac{\partial q}{\partial x} = \frac{Q_{i+1} - Q_{i-1}}{2\Delta x} \tag{151}$$

$$\frac{\partial^2 q}{\partial x^2} = \frac{Q_{i+1} - 2Q_i + Q_{i-1}}{(\Delta x)^2} \tag{152}$$

$$\frac{\partial^3 q}{\partial x^3} = \frac{Q_{i+2} - 2Q_{i+1} + 2Q_{i-1} - Q_{i-2}}{2(\Delta x)^3} \tag{153}$$

where Qi^* is the solution of the homogeneous system at each time step. The *ghost cells* defined previously are used here for the computations at the boundaries.

4.3.3. Initial and Boundary Conditions

For the boundary condition, we will use again a non-reflecting outflow; as for the initial conditions, they will be the same as in the SVE regarding the variables h and u, together with $W(x,0) = 0 \quad \forall x$.

4.4. Results

Now we can compare the numerical solutions for both systems and demonstrate that in practice $W(x,t)$ will work as a singularity detector, also that the solutions for the height and velocity from the moment system of the third order will tend to the ones in the original SVE. The solutions are shown at times $t = 0$, 1.875, 3 and $\varepsilon = 0.01$, 0.1, 1 (see Figures 1–5). For the initial time, only the results of the first ε are presented, since they are equal for the others. The blue graphics correspond to the SVE and the red ones to the third order system.

As we could see in Figure 1, the initial condition for the height of the water $h(x,t)$ simulates the dam-break problem, which considers a wall separating the two different levels of water with zero velocity.

As time starts running, the wall is removed and water starts moving to the left side of the $x-axis$ originating two different types of waves. The front wave will steepen into a shock wave and the back wave will spread-out as a rarefaction wave.

The formation of both waves can be observed in the Saint-Venant Equations as well as in the third order moment system, proving – as we were expecting – that not only the solutions of the last system will approximate the original SVE, but also that as we decrease the value of ε the approximation will improve. On the other hand, the function $W(x,t)$ indeed behaves as a δ-function at discontinuities, going to ∞ at the point of the shock wave, to zero in the regions where the velocity is constant and to $-\infty$ in the region of the rarefaction wave. Also, we should notice that the position of the singularities is not approximated exactly, this can cause some problems for example when using the third order moment system to construct adaptive discretization techniques.

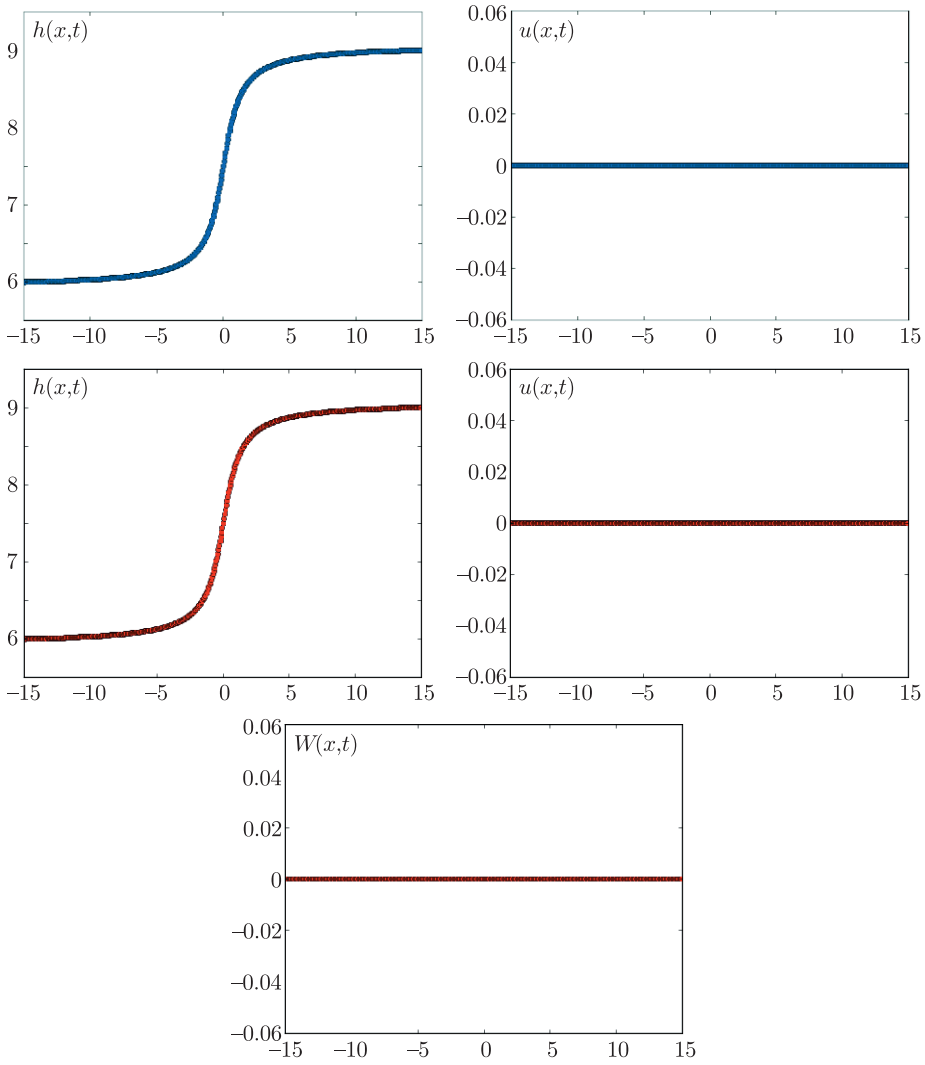

Figure 1. Initial conditions ($t=0$) for the SVE (blue) and the third order moment system (red)

In order to measure the differences at a fixed time, between the numerical solutions of $h(x,t)$ and $u(x,t)$ in both systems, we will use the discrete l_1-norm since for conservation laws the integral of the conserved quantities is of great relevance (as seen in previous chapters); and the l_∞-norm which helps us to prove point-wise convergence. We will use Q_m^n to denote the solutions of the third order moment system and q_m^n for their counterparts obtained from the SVE:

$$l_1^n = \frac{1}{M+1}\sum_{m=0}^{M} |\, Q_m^n - q_m^n \,| \tag{154}$$

$$l_\infty^n = \max_{m=0,\ldots,M} |\, Q_m^n - q_m^n \,| \tag{155}$$

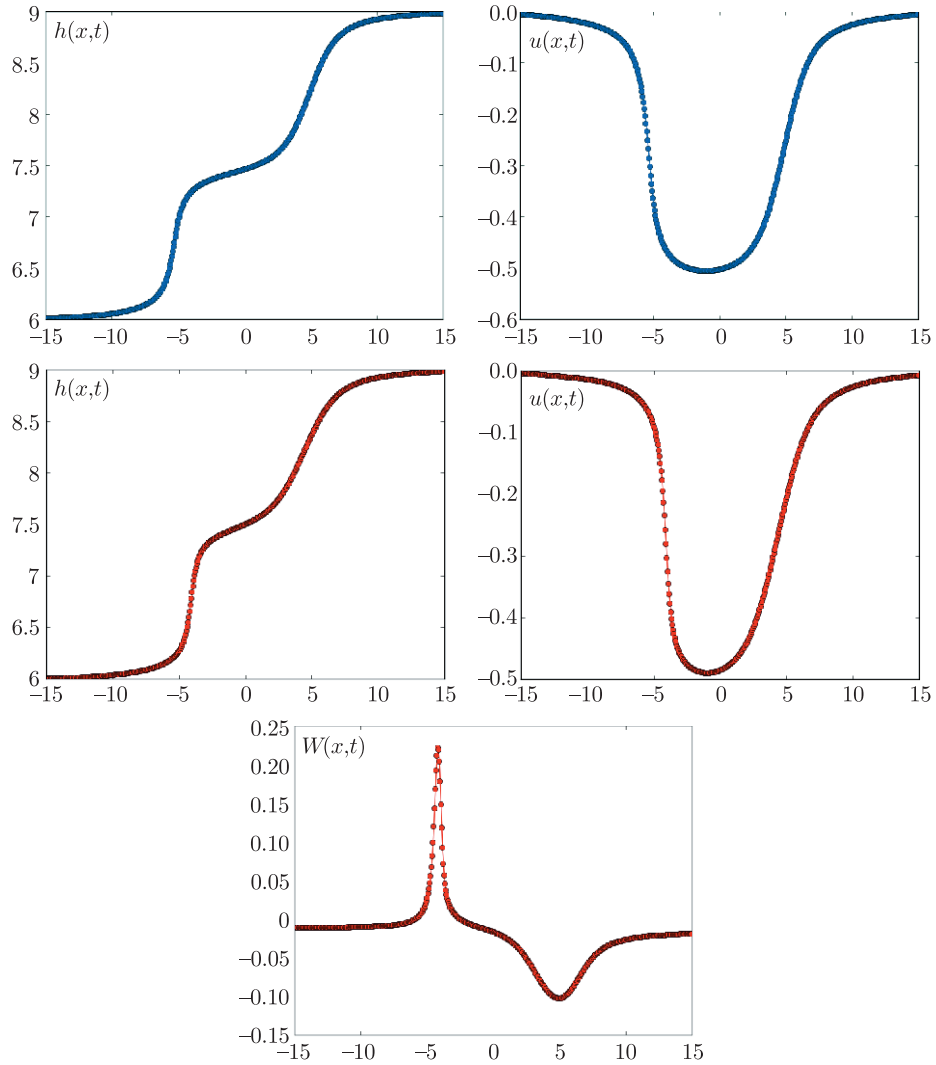

Figure 2. Numerical results at $t = 1.875$ for the SVE (blue) and the third order moment system (red) with $\varepsilon = 0.01$

In Table 1 the results for $t = 0.75$, 1.875, 3 and $\varepsilon = 0.01$ are shown, while Table 2 shows the values where the l_∞-norm is achieved. The errors are calculated only for the smallest value of ε since, as was clear from the figures, as we increase the relaxation term, the differences between the SVE results and the third order moment system will increase. Capital letters are used for the results from the third order moment system and lower case letters for the ones from the SVE.

The previous results allow us to see that as time advances, the l_p errors are increasing, and at each time step the l_∞-norm is achieved in the positions where the discontinuities are formed; these large differences are in part due to the fact that we are not in the formal limit of ε and shifts in the exact

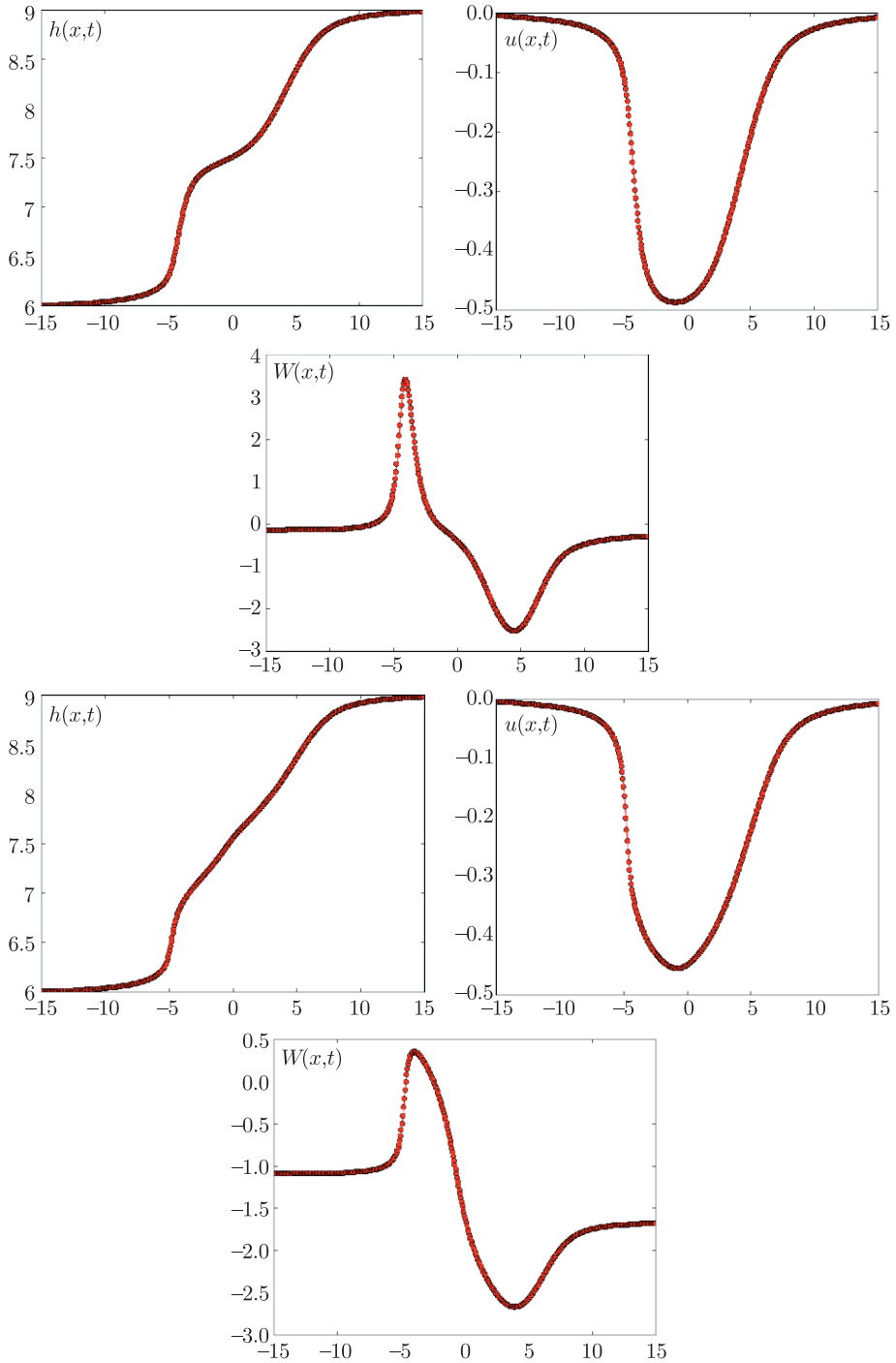

Figure 3. Numerical results for the third order moment system at $t = 1.875$ with $\varepsilon = 0.1$ (top) and $\varepsilon = 1$ (bottom)

Table 1. l_p errors between solutions of $h(x,t)$ and $u(x,t)$ for $\varepsilon = 0.01$

	$h(x,t)$		$u(x,t)$	
t	l_1^n	l_∞^n	l_1^n	l_∞^n
0.75	0.0242	0.2404	0.0126	0.1088
1.87	0.0762	0.7402	0.0315	0.2937
3	0.1283	1.0321	0.0499	0.4041

Table 2. Positions and values where the l_∞-norm is achieved for $\varepsilon = 0.01$

t	x_∞	H_∞	h_∞	U_∞	u_∞
0.75	−1.98	6.66	6.90	−0.1720	-0.281
1.87	−4.74	6.36	7.10	−0.1200	-0.414
3	−7.56	6.19	7.22	−0.0719	-0.476

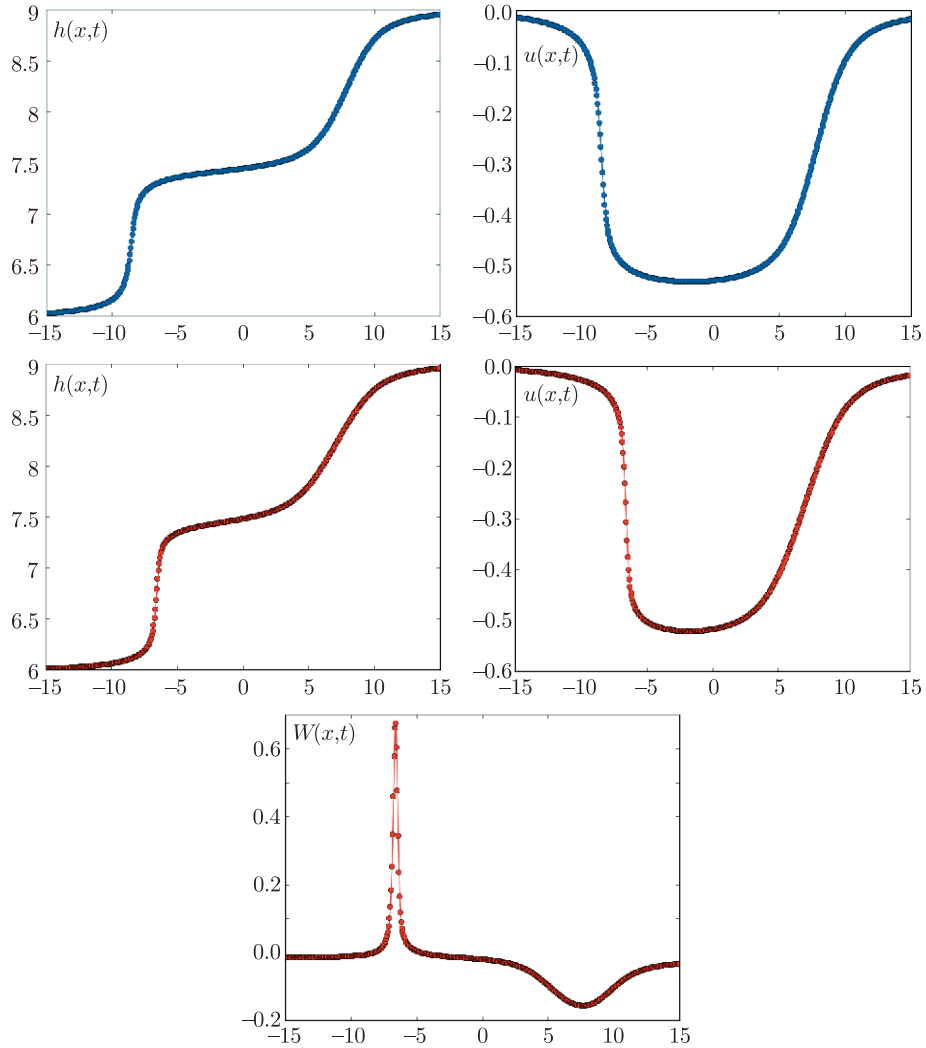

Figure 4. Numerical results at $t = 3$ for the SVE (blue) and the third order moment system (red) with $\varepsilon = 0.01$

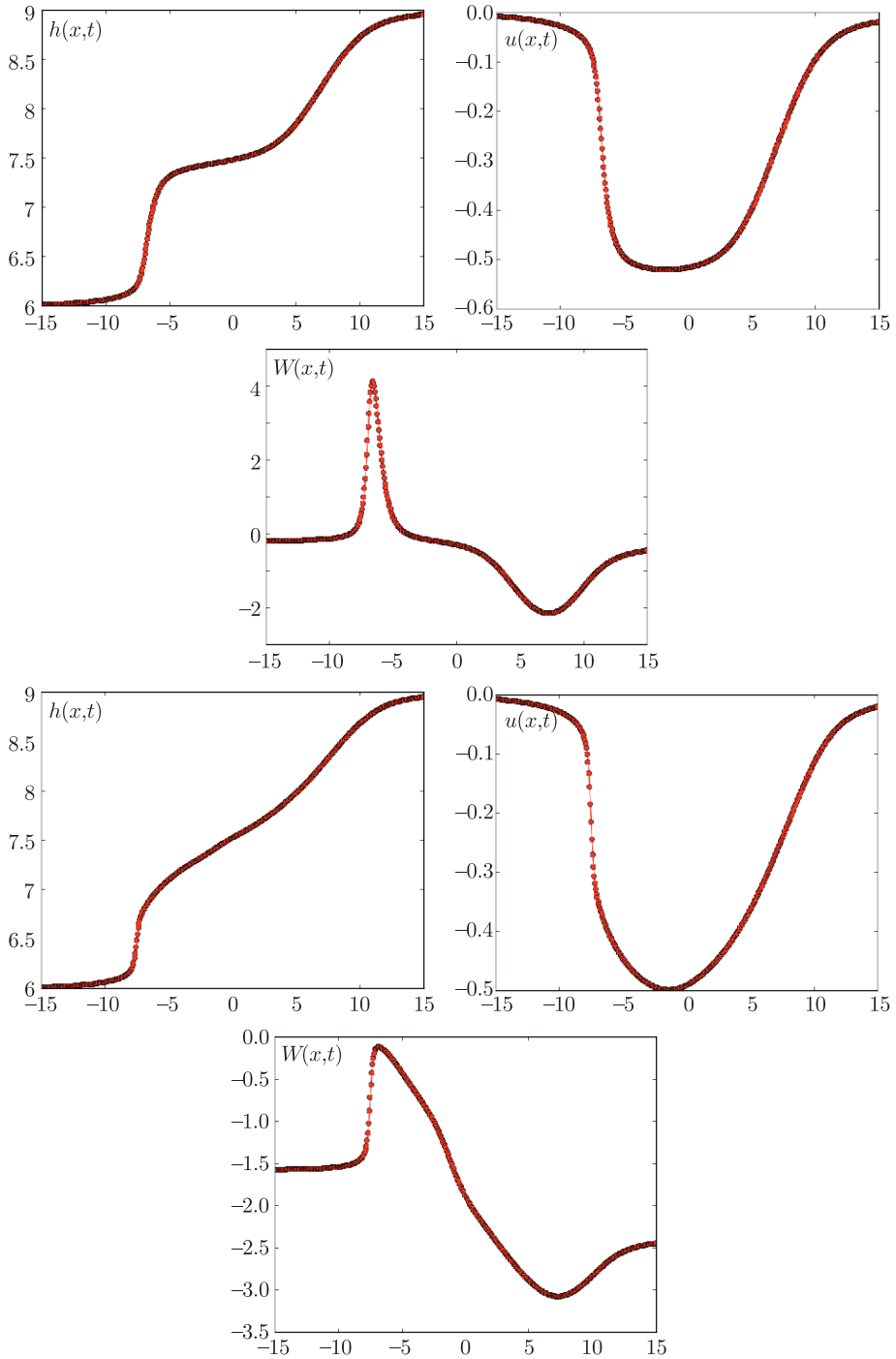

Figure 5. Numerical results for the third order moment system at $t = 3$ with $\varepsilon = 0.1$ (top) and $\varepsilon = 1$ (bottom)

location of the singularities will occur, becoming larger in time; another important reason for these differences comes from the use of an explicit scheme for the discretization of the source term which might not be the best approximation for the problem.

Nevertheless, in a general outlook the results obtained are very promising, and for $\varepsilon \ll 1$ the third order system represents a fine approximation for the Saint-Venant equations until now, as we can see with the l_1-norm. However, considerations in the approximation of the source term should be made together with a deeper study on the convergence of the third order moment to the SVE.

5. Conclusions and future work

Starting from a Boltzmann-like kinetic transport equation with a BGK relaxation term and using the moments of the density of particles $f(x,t,\xi)$, it was possible to reconstruct a relaxation model, such that, in the limit of $\varepsilon \to 0$, the density of particles will tend to an equilibrium function. This equilibrium function can be constructed, with the help of the *kinetic representation*, to give first three moment equations that will form an equivalent system to the Saint-Venant Equations.

As $\varepsilon \to 0$, the system approaches the SVE together with a higher order moment $W(x,t)$, that may act as a shock and rarefaction waves detector, given that at discontinuities it behaves like a δ-function.

On the other hand, the numerical results allow us to see that the theoretical predictions are confirmed, given that as ε is decreased, the solutions of $h(x,t)$, $u(x,t)$ and $W(x,t)$ are become closer to the equilibrium values. Nonetheless, the numerical errors show us that the convergence of the third order moment system to the SVE has to be improved, since large differences arise in the positions of the singularities. Hence, a more elaborated discretization for the source term should be implemented, for example using implicit schemes instead of an explicit one. Also, a deeper study of the system under different initial conditions using the Riemann invariants can give us more information on the behaviour of the solution.

Until now, we have achieved very good results for the one-dimensional case of the Saint-Venant Equations; moreover, we have built the basis for the construction of an adaptive discretization method that will allows to improve the time of the computations and the accuracy of the solutions. Along with this, considerations on the change of topography and dry states ($h(x,t) = 0$) can be made together with the two-dimensional case.

References

[1] Leveque R J 2002 *Finite Volume Methods for Hyperbolic Problems*, *Applied Mathematics*, Cambridge
[2] Stoker J J 1958 *Water Waves. The Mathematical Theory with Applications*, Wiley
[3] Lions P L, Perthame B and Tadmor E 1994 *Commun. Math. Phys.* **163** 415
[4] Perthame B 2002 *Kinetic Formulation of Conservation Laws*, Oxford University Press, NY

[5] Struchtrup H 2005 *Macroscopic Transport Equations for Rarefied Gas Flows*, Springer-Verlag, Berlin
[6] Gil Montoya D C, Struchtrup H and Struckmeier J 2013 *A Kinetic-Induced Moment System for Burgers Equation*, Hamburg (in preparation)
[7] Leveque R J, Berger M J, Langseth J O, Calhoun D, Mitran S, Blossey P, George D, Mandli K, Swierczewski C and Ketcheson D 2012 *Clawpack Software 4.6*, www.clawpack.org