A well-balanced solver for the Saint Venant Equations with variable cross-section

Raul Borsche

March 25, 2013

Abstract

In this paper we construct a numerical solver for the Saint Venant equations. Special attention is given to the balancing of the source terms, including the bottom slope and variable cross-sectional profiles. Therefore a special discretization of the pressure law is used, in order to transfer analytical properties to the numerical method. Based on this approximation a well-balanced solver is developed, assuring the C-property and depth positivity. The performance of this method is studied in several test cases focusing on accurate capturing of steady states.

1 Introduction

The Saint Venant equations are used for the modeling of the flow in rivers [1, 10, 20, 28, 27], as well as in sewer systems [6, 7, 19, 25]. In both cases accurate and stable numerical solvers are needed in order to resolve the involved flow phenomena. For the development of suitable numerical methods many ideas can be transferred from the similar shallow water equations. Here a large variety of schemes [21, 29] is available, addressing several different aspects like e.g. well-balancing of the bottom topography [2, 3, 15, 24, 31]. Nevertheless some aspects of the Saint Venant equations are not covered immediately by this correspondence. The two properties considered here are the approximation of the pressure law and the well-balancing of source terms, which include a variation in the cross-sectional profile.

For the approximation of the pressure law two common approaches exist. Often either the pressure law is approximate by some suitable functions like in [15, 19] or some additional solver is used in order to determine the correct pressure associated to a given cross-sectional area. In this paper an approach similar to the ideas of the Preissmann slot [10, 25] is considered. Instead of approximating the pressure law directly, the shape of the profile is fitted by piecewise linear functions. This leads to a sufficiently accurate approximation of the pressure law and additionally preserves several relations between the involved quantities for the numerical method. This is of special importance for the well-balancing of the source terms.

There has been a remarkable effort in developing numerical schemes for the shallow water equations, in order to incorporate the stiff source term of varying bottom topography [2, 3, 15, 24, 31]. Most of the are based on the so called source term upwinding, which discretizes the source term in exactly the same way as the corresponding numerical scheme for the hyperbolic part of the equation. The advantage is, that for steady states the variations in the source term can be directly balanced by the spatial variations in the flux function, such that no unphysical waves are generated.

As numerical scheme we develop an augmented Riemann Solver similar to the one presented in [15]. Based on common techniques like Roe-, HLLE- or Flux-Difference-Splitting schemes, this approach unifies several desired features. Beside some basic properties, it has an improved treatment of large rarefaction waves and it is well-balanced with respect to the bottom slope and variations in the cross section of the flow profile. The former property is important for the accurate
capturing of wet/dry fronts. For the special choice of the pressure law \( p(A) = \frac{g}{2} A^2 \), the Saint Venant equations reduce to the shallow water equations. In this case the solver coincides with the one used in the GEOCLAW framework [22], for which the stability and capacity to large scale applications is intensively tested.

In the next section we recall the Saint Venant equations and describe the different possible source terms. We put a special focus on the pressure law and the analytical properties required for the steady state solutions. In the numerical section the augmented Riemann Solver for the Saint Venant equations is developed and the well-balancing of the source terms is established. Finally we present different numerical test cases confirming the desired properties.

2 The Saint Venant Equations

The Saint Venant equations (StVE) describe the flow of a fluid in a spatially one dimensional structure, such as a river [1, 10, 20, 28] or tube partially filled with water [6, 7, 19, 25]. They are given by the following two hyperbolic equations

\[
\begin{align*}
\partial_t A + \partial_x Q &= 0 \\
\partial_t Q + \partial_x \left( \frac{Q^2}{A} + p(x, A) \right) &= -gA \partial_x z + S_w - S_f. 
\end{align*}
\] (2.1)  

Model: eq:StVE

Here \( A(t,x) \) denotes the wetted cross sectional area and \( Q(t,x) \) is the flow of water into \( x \)-direction at a given time \( t \in \mathbb{R}^+ \) and location \( x \in \mathbb{R} \). The function \( p(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^+ \) is the hydrostatic pressure law, which highly depends on the geometry of the flow profile. \( g \) is the gravitational acceleration and \( z(x) \) is the bottom elevation above a given datum. The first equation guarantees the conservation of mass, the second one is a balance law on the momentum. The source term \( S_w \) is balancing spatial variations of the flow profile and \( S_f \) introduces the friction into model.

The pressure law \( p \) can be derived under a hydrostatic assumption [1, 10] and has the general form

\[
p(x, A) = g \int_0^A \left( h(x, A) - h(x, a) \right) da. \] (2.2)  

Model: eq:pressure

The function \( h(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^+ \) denotes the relative height of water to the corresponding wetted area. This relation is a direct property of the geometry of the flow profile. The source term \( S_w \) balances the spatial variations of the pressure law

\[
S_w = g \int_0^A \left( h(x, A) - h(x, a) \right) \frac{\partial_x w(x, a)}{w(x, a)} da, \] (2.3)  

Model: eq:spatial

where \( w(\cdot, \cdot) : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^+ \) denotes the width of the flow profile at the height of the actual water level.

For the friction term the formula of Manning can be used [10] \( S_f = gn_f^2 \frac{Q}{A^{4/3}} \).

\( n_f \) is the Manning friction coefficient and \( r_{hy} = \frac{A}{U} \) is the hydraulic radius, where \( U \) is the wetted perimeter corresponding to \( A \).

With suitable choices of the initial conditions, sufficiently bounded source terms and appropriate boundary conditions, the system (2.1) is well posed. A proof can be found in [17].

2.1 The Pressure Law

As pressure law we refer to the function \( p \) given by (2.2), although it has the units \( m^4 s^{-1} \), which is not a pressure in the classical sense. In fact it is the accumulated force acting on the area \( A \).
due to hydrostatic pressure of the water. The definition (2.2) involves the two functions \( h \) and \( w \), which are specified by the geometric shape of the flow profile. For any reasonable cross section the function for the height \( h \) is strictly monotone increasing in \( A \) and in most cases even differentiable. If \( h \) is not differentiable, the profile can be slightly smoothed, such that this property is restored. Similarly we will assume that \( w \) is at least piecewise continuous in \( A \). Some simple relations of \( p \), \( A, h \) and \( w \) generally hold, independently of the actual shape on the flow profile

\[
A(h) = \int_0^h w(\xi) \, d\xi , \quad \partial_A h = \frac{1}{w(A)} , \quad \partial_A p = g \frac{A}{w(A)} .
\]  

(2.4)

The first two are immediate geometric properties, while the third one directly follows from the definition of \( p \) (2.2). As above and in the following we will handle the role of \( A \) or \( h \) as argument of the functions freely, since their relation is bijective.

2.2 Steady states

Steady states are of special interest for the long term behavior of the equation. Furthermore they take a prominent position as benchmark solutions for numerical schemes, to test the accurate incorporation of the source term into the method. In the following we will refer to steady states, when considering the solutions of

\[
\partial_x Q = 0,
\]

\[
\partial_x \left( \frac{Q^2}{A} + p(x,A) \right) = -gA\partial_x z + S_w .
\]

(2.5)

We have dropped the friction term, since it admits only steady states on finite intervals and it can be incorporated numerically easily, e.g. by splitting techniques [22].

The above equations admit several possible steady states. In the case of continuous solutions, the second equation of (2.5) can be transformed into

\[
\left( \frac{1}{2} \frac{Q^2}{A^2} + gh(x,A) + gz \right)_x = 0 ,
\]

(2.6)  

by using the relations (2.4). This condition is known as Bernoulli equation [10].

The most intuitive steady state might be the so called lake at rest, i.e. \( Q(t,x) \equiv 0 \) and \( h(t,x) + z(x) \equiv const. \). In the absence of sources due to the bottom slope and changing cross-sections, it is easy to prove, that only trivial solutions \( Q(t,x) \equiv const. \) and \( h(t,x) \equiv const. \) can arise, e.g. see [14]. In presence of variations in \( z(x) \), more involved examples can be constructed. In the context of the shallow water equations, the flow over a local bump with different velocities is a relevant benchmarking example [2, 15, 24]. For the situation including changes in the cross-sectional profile similar test cases can be considered [31].

It is important to note, that for the balancing of the source terms the relations (2.4) have to hold, since they establish the necessary relations between \( p, A, h \) and \( w \). In case of the shallow water equations these are trivially fulfilled, since \( w \) is constant and \( A \) is a linear function of \( h \).

3 Numerical method

Before we start with the actual numerical solver for the StVE, we will take a closer look on the discretization of the pressure law \( p \).

3.1 Discretizing the pressure law

As mentioned above, the function \( p \) is strictly monotone and differentiable. Although it is analytically of such a friendly type, it might cause some difficulties in practical realizations. For example
pipes with circular cross sections are the main element of sewer systems [7, 19, 25]. It is easy to
derive, that in this case the area $A$ has the following representation in terms of $h$

$$A(h) = r^2 \arccos \left( \frac{r - h}{r} \right) - (r - h) \sqrt{r^2 - (r - h)^2}.$$  

Unfortunately this equation can not be inverted globally towards $h$. Thus we can not derive an
explicit formula for the pressure law $p$.

This drawback can be overcome in different ways. A first choice might be to solve the above
expression numerically. In general this is not desirable, since it results in a computationally costly
method. A much more practical approach is to approximate the resulting function $p$ by a Taylor
series of sufficient high order. This ansatz is e.g. used in [18] with very satisfying results.

Nevertheless both methods have problems in balancing the source terms on the right hand
side. It is easy to see, that the relations (2.4) do not hold for general approximations of $p$. To
address this problem, we approximate $w$ as a function of $h$, by some piecewise linear function, i.e.
we approximate the cross-sectional profile instead of $p$. Based on this approach the functions $A(h)$
and $p(A)$ can be defined by the equations (2.4).

$$w(h) = \begin{cases} 
\frac{2}{2r^2} h & 0 \leq h \leq h_1 \\
\frac{2}{2r^2} (2r - h) & h_1 < h \leq h_2 \\
\frac{1}{4c^2} h^3 + \frac{2c}{3r} h^2 + \left( c + 2(\pi - 2) - 2 \frac{\pi^2}{12c} \right) r^2 h \\
\frac{1}{3c^2} h^3 + c^2 + \frac{\pi^2}{12c} & h_2 < h \leq 2r 
\end{cases}$$

$$h(A) = \begin{cases} 
\sqrt{\frac{(2 - \frac{\pi}{2}) A}{2r^2}} & 0 \leq A \leq A_1 \\
\frac{2r^2}{2r^2 - \pi r^2} A + (1 - \frac{\pi}{2}) r & A_1 < A \leq A_2 \\
2r - \sqrt{\frac{2 - \frac{\pi}{2} \left( \pi r^2 - A \right)}{r^2}} & A_2 < A \leq \pi r^2 
\end{cases}$$

$$p(h) = g \begin{cases} 
\frac{1}{3c^2} h^3 + (c + 2(\pi - 2) - 2 \frac{\pi^2}{12c}) r^2 h & 0 \leq h \leq h_1 \\
\frac{1}{3c^2} h^3 + c^2 + \frac{\pi^2}{12c} & h_1 < h \leq h_2 \\
\frac{1}{3c^2} h^3 + (c + 2(\pi - 2) - 2 \frac{\pi^2}{12c}) r^2 h & h_2 < h \leq 2r 
\end{cases},$$  

where $c = 2 - \frac{\pi}{2}$. We see that the pressure law of the hexagonal profile agrees quite good with the
circular one. With more discretization points, even better approximations can be obtained.

In Figure 1 the approximation to a circular tube and the resulting pressure law are shown.

In Figure 2 the solution to a Riemann Problem with different cross-sectional profiles is plotted.

Figure 1: Approximating a circular tube by piecewise linear functions: the shape (left) and the
pressure law $p$ (right) of a circular (blue), a diamond (red) and a hexagonal (green) tube.

Numerics:fig:CrossectionOfDifferentTubes

Figure 2: The solution to a Riemann Problem with different cross-sectional profiles is plotted.

In Figure 2, the structure of the waves remains unchanged and the intermediate state is
almost identical in all three cases. The solution of the diamond shaped profile produces wrong
wave speeds, whereas the hexagonal profile already matches the solution of the circular profile quite accurately. A proof of the continuous dependence on variations in the flux function can be found in [5].

So we can conclude, that is very easy to obtain accurate approximations of the physical solutions with the help of simple approximations of the flow profile. This approach has the advantage that the relations (2.4) also hold for the discretized functions.

3.2 An Augmented Riemann Solver for the StVE

In this section we construct a solver for the Saint Venant equations with source terms due to bottom topography and variations in the cross-sectional profile. Since the StVE are a classical conservation law, many solvers for hyperbolic systems can be applied. The choice of a numerical scheme should reflect most of the underlying properties of the considered system. This includes, besides an accurate approximation to the exact solution, features like the conservation of mass and stability especially in regions of shocks. In contrast to these rather popular properties, the well balancing of the bottom slope and of the change of the width of the cross-section or the accurate capturing of wet dry interfaces are only inherited by particular solvers. In the following the main focus is given to these special properties.

First the general concept of a Riemann Solver is introduced and some basic methods like the Roe-scheme, the wave propagation method or the flux-difference splitting approach are shortly reviewed. Based on these, the general structure of an augmented solver similar to [15] is presented. This scheme includes an additional wave, which can be used to improve the capturing of large rarefaction waves and the resolutions in situations near dry states. A special focus is given to the source terms and their balancing with the flux at steady states.

In the following, the spatial interval $[0, L]$, is separated into subintervals $C_i$, $i = 1, \ldots, N$ of equal size $\Delta x$. The center of $C_i$ is denoted by $x_i$, the interface between $C_{i-1}$ and $C_i$ is at $x_{i-\frac{1}{2}}$. For the numerical section all variables are assumed to be discretized according to this particular grid. For simplicity of notation, the step-function keeps the name of original variables, but can be distinguished by the presence of an index. E.g. the discrete initial values or the cross sectional area are given by the average

$$ (A_0)_i = \frac{1}{\Delta x} \int_{C_i} A_0(x) dx \quad i = 1, \ldots, N. $$

The brackets are omitted if possible. At the ends of the tube the boundary conditions are applied using the ghost cell technique [21].
### 3.2.1 **Roe-Solver and flux-difference splitting**

Before starting with the development of the actual augmented Riemann Solver, we review briefly some common schemes, which we used in its construction. A very popular choice for p-systems is the so called Roe-Solver [25] [27]. The idea of the Roe Solver is to approximate the flux at the interface by a suitable average state \( \hat{u} = \bar{u}(u_L, u_R) \). The so-called Roe-average is defined by the following relation

\[
f(u_R) - f(u_L) = \nabla f(\hat{u})(u_R - u_L) ,
\]

where \( \nabla f \) denotes the Jacobian of the flux function \( f \) and \( u_L, u_R \) are the states on the left and right hand side of the considered interface. In the present case of the StVE it is easier not to define an average of \( A \) and \( Q \), but to approximate the velocity by \( \hat{v} \) and the derivative of the pressure law by \( \hat{p'} \). Thus the above equation can be written as

\[
\begin{pmatrix}
\frac{Q_R - Q_L}{A_R} + p(A_R) - \frac{Q_L}{A_L} - p(A_L) \\
\frac{Q_R - Q_L}{A_R} + p(A_R) - \frac{Q_L}{A_L} - p(A_L)
\end{pmatrix} = 
\begin{pmatrix}
0 & 1 \\
-\hat{v} + \hat{p'} & 2\hat{v}
\end{pmatrix}
\begin{pmatrix}
A_R - A_L \\
Q_R - Q_L
\end{pmatrix} .
\]

This equation is fulfilled for the following choices of

\[
\hat{v} = \sqrt{\frac{A_R v_R + \sqrt{A_L v_L}}{A_R + \sqrt{A_L}}}
\text{ and }
\hat{p'} = \frac{p(A_R) - p(A_L)}{A_R - A_L} .
\]

\( \hat{v} \) is the common choice out of the two possible solutions to a quadratic equation [20]. Corresponding to these averages the eigenvalues of \( \nabla f(\hat{u}) \), the so-called Roe-speeds \( \hat{\lambda}^1, \hat{\lambda}^2 \), and the eigenvectors \( \hat{\nu}^1, \hat{\nu}^2 \) are

\[
\hat{\lambda}^1(\hat{q}) = \hat{v} - \sqrt{\hat{p}'}, \\
\hat{\lambda}^2(\hat{q}) = \hat{v} + \sqrt{\hat{p}'} ,
\]

\[
\hat{\nu}^1 = \left( \frac{1}{\hat{\lambda}^1(\hat{q})} \right) , \\
\hat{\nu}^2 = \left( \frac{1}{\hat{\lambda}^2(\hat{q})} \right) .
\]

These values can be used in the context of the wave propagation methods [21] or the flux-difference splitting approach [3]. Therefore the jump in the states at the cell interface is considered and decomposed into the directions of the eigenvectors

\[
u_R - \nu_L = \sum_{k=1}^{n_w} \alpha^k \hat{z}^k = \sum_{k=1}^{n_w} \hat{W}^k .
\]

In the present case the number of waves considered is \( n_w = 2 \), as in the analytical solution. By applying equation (4.2), the flux is split into so called \( f \)-waves \( \hat{\nu}^k = \hat{\lambda}^k \hat{\nu}^k \) by

\[
f(u_R) - f(u_L) = \sum_{k=1}^{n_w} \alpha^k \hat{z}^k = \sum_{k=1}^{n_w} \sum_{k=1}^{n_w} \alpha^k z^k = \sum_{k=1}^{n_w} \hat{z}^k ,
\]

for some given \( z^k \), here \( z^k = \hat{\lambda}^k \hat{\nu}^k \). Now we sort these waves by their direction of propagation, which is indicated by the sign of the eigenvalues. The \( f \)-waves with the same direction are combined to the so called fluctuations

\[
A^- \Delta u_{i-1/2} = \sum_{k : \hat{\lambda}^k_{i-1/2} < 0} \hat{z}^k_{i-1/2} , \\
A^+ \Delta u_{i-1/2} = \sum_{k : \hat{\lambda}^k_{i-1/2} > 0} \hat{z}^k_{i-1/2} .
\]

Thus the state \( u^j_i \) is only updated by the forward-going \( f \)-waves of the interface at \( x_{i-1/2} \) and the backward traveling \( f \)-waves from \( x_{i+1/2} \). This leads to the following update formula in terms of the fluctuations

\[
u^j_{i+1} = \nu^j_i - \frac{\Delta t}{\Delta x} \left( A^+ \Delta u_{i+1/2} - A^- \Delta u_{i-1/2} \right) .
\]

At this point it should be emphasized that the wave propagation method yields a conservative scheme only if equation (3.2) holds. In contrast, the flux-difference splitting approach is always conservative, due to its construction.
3.2.2 Augmented Riemann Solver

In the following, the above concept of the flux-difference splitting and the wave propagation is extended to a so-called Augmented Riemann Solver \[8\] \[15\]. There is a close relation of relaxation schemes to augmented Riemann Solvers as shown in \[24\].

The main idea of these schemes is to enlarge the number of waves considered in the splittings \((3.3)\) or \((3.4)\). In the present case this is achieved by combining both equations \((3.3)\) and \((3.4)\). This leads to a system of \(n_w = 4\) waves

\[
\begin{bmatrix}
  u_R - u_L \\
  f(u_R) - f(u_L)
\end{bmatrix} = \sum_{k=1}^{4} \alpha^k \begin{bmatrix}
  \hat{f}^k \\
  z^k
\end{bmatrix}.
\]

In case of the StVE the second line of \(u_R - u_L\) and the first one of \(f(u_R) - f(u_L)\) coincide, thus the system reduces to \(n_w = 3\) equations

\[
\begin{bmatrix}
  A_R - A_L \\
  Q_R - Q_L \\
  \phi_R - \phi_L
\end{bmatrix} = \sum_{k=1}^{3} \beta^k w^k,
\]

for some linearly independent vectors \(w^k \in \mathbb{R}^3\). Above and in the following \(\phi = \frac{Q^2}{2} + p(A)\) represents the momentum flux. As a direct consequence of the construction, the left hand side of \((3.7)\) lies in a 2-dimensional subspace spanned by \((1, \lambda_1, (\lambda_1)^2)^T, k = 1, 2\), since the wave propagation method and the flux-difference splitting approach coincide for these choices of \(w^k\).

In the following we construct a new set of vectors \(w^k\), in order to improve the approximation to the exact solution. In general these vectors have to be linear independent, to guarantee a well-defined decomposition. We can use particular choices of \(w^k\) to avoid some shortcomings of the Roe linearization, but still maintain the strength of the above methods. We will exploit the ability to choose three instead of two waves, to improve the behavior in case of large rarefaction waves \[14\] \[15\]. In case of a shock wave the choice of the Roe linearization directly provides the correct shock speed, since the state is an eigenvector of \((3.2)\).

A well known defect of Roe’s method is the convergence to entropy violating shocks, in cases of transsonic rarefaction waves. This can be overcome by the entropy fix proposed in the HLLE method \[14\]. The Roe speeds are replaced by the so-called Einfeldt-speeds

\[
\lambda^{E,1} = \min \left( \lambda^1(u_L), \lambda^1(\hat{q}) \right), \quad \lambda^{E,2} = \max \left( \lambda^2(u_R), \lambda^2(\hat{q}) \right).
\]

A further advantage of these choices is, that the positivity of the wetted cross sectional area is maintained \[12\]. With the Einfeldt-speeds as speeds for the first and third wave, a simple choice for the speed of the second wave is \(\frac{1}{2}(\lambda^{E,1} + \lambda^{E,2})\). This leads to the following vectors \(w^k, k = 1, 2, 3\) and their corresponding speeds of propagation \(s^k\)

\[
\begin{align*}
  w^1 &= \begin{bmatrix} 1, s^1, (s^1)^2 \end{bmatrix}^T, & s^1 &= \lambda^{E,1}, \\
  w^2 &= \begin{bmatrix} 0, 0, 1 \end{bmatrix}^T, & s^2 &= \frac{1}{2}(\lambda^{E,1} + \lambda^{E,2}), \\
  w^3 &= \begin{bmatrix} 1, s^3, (s^3)^2 \end{bmatrix}^T, & s^3 &= \lambda^{E,2}.
\end{align*}
\]

A detailed discussion of the advantages of these choices in the context of the shallow water equations can be found in \[14\]. This set of linearly independent vectors can now be applied in the splitting \((3.7)\).

Before applying the final update formula \((3.6)\), we have to define the fluctuations. As the state variables are only two dimensional vectors, the fluctuations have to be of the same size. Choosing the first two components of the \(w^k\) leads to a wave propagation type algorithm. In the present case we choose the second and the third components of the \(w^k\), which corresponds to a
flux difference splitting approach. This choice automatically guarantees the conservation of mass. Thus the fluctuations needed in (3.6) are defined as

\[
A^- \Delta u_{i-1/2} = \sum_{k:k_{i-1/2} < 0} \beta^k \tilde{w}^k_{i-1/2} + \frac{1}{2} \sum_{k:k_{i-1/2} = 0} \beta^k \tilde{w}^k_{i-1/2}, \tag{3.9a}
\]

\[
A^+ \Delta u_{i-1/2} = \sum_{k:k_{i-1/2} > 0} \beta^k \tilde{w}^k_{i-1/2} + \frac{1}{2} \sum_{k:k_{i-1/2} = 0} \beta^k \tilde{w}^k_{i-1/2}, \tag{3.9b}
\]

where \(\tilde{w}^k_{i-1/2}\) is the two dimensional vector composed out of the second and third component of \(w^k_{i-1/2}\). The case of \(s^k = 0\) for some \(k\) was not included in (3.5), as stationary waves automatically have zero strength in the Roe scheme. In the above method the contribution of these particular waves is split up into both directions equally as recommended in [15]. This is of special interest in combination with source terms, as discontinuous bottom slope or discontinuous channel width might cause stationary jumps in the states.

### 3.2.3 Dry states and large rarefaction waves

The choice of the first and the third wave in (3.8) is mainly devoted to the shock waves, as the Roe averages guarantee their accurate capturing. Thus the intermediate wave \(w^2\) can be used for improvements in the case of large rarefaction waves.

In general the capturing of rarefaction waves can be enhanced by some additional information on the new middle state of the Riemann Problem \(u^*\). The value of \(u^*\) can be computed by an Godunov solver, similar to the one in [29]. In order not to slow down the computation, it is sufficient to approximate \(u^*\), e.g. by some very few steps of a Newton method solving the system of the Lax-Curves [9].

This additional information we use to define the so-called generalized Einfeldt speeds

\[
s^1 = \min \left( \lambda^1(u_L), \tilde{\lambda}^1(\tilde{q}), \lambda^2(u^*) \right) \tag{3.10a}
\]

\[
s^3 = \max \left( \lambda^2(u_R), \tilde{\lambda}^2(\tilde{q}), \lambda^1(u^*) \right) \tag{3.10b}
\]

In general these speeds are combined with the above choices of \(w^k\) and \(s^2 = 0.5(s^1 + s^3)\), as in (3.5). In case of a strong rarefaction wave, but away from dry states, the \(w^2\) vector can be explicitly used to rise the resolution of this particular wave. The information, which kind of waves occur at the considered Riemann Problem, can be taken from the computation of the middle state \(u^*\).

In case of a strong 1-rarefaction wave, we define the middle wave \(w^2\) and its speed \(s^2\) by

\[
w^2 = \left( 1, \lambda^1(u^*), \lambda^1(u^*) \right)^T, \quad s^2 = \lambda^1(u^*).
\]

The case of a strong 2-rarefaction wave we choose analogously

\[
w^2 = \left( 1, \lambda^2(u^*), \lambda^2(u^*) \right)^T, \quad s^2 = \lambda^2(u^*).
\]

Both vectors and speeds are meant to represent the tale of the rarefaction wave, which exactly moves with \(s^2\). The head of the wave is captured by the adapted choices of (3.10).

In the case of two strong rarefaction or near dry states, we return to the initial configuration

\[
w^2 = (0, 0, 1)^T, \quad s^2 = \frac{1}{2}(s^1 + s^3),
\]

since the depth positivity can only be assured for these choices. A detailed discussion of the strengths of these constructions can be found in [14] [15].
3.2.4 Source terms

Of special interest for our solver of the StVE is the incorporation of the source terms. In particular we are interested in the influences of the slope of the bottom and the changes in the shape of the cross section. For both the well-balancing of the numerical flux with the discretization of the source terms at steady states is desired. For the shallow water equations different approaches have been addressed to this problem [1, 13, 3, 24]. In the following we adapt the ideas of [14, 15] to the StVE.

As mentioned in section 2.2 the steady states are given by equation (2.5). In the case of continuous solutions this expression can be reformulated as (2.6). But in general we can not assume, that all considered quantities are continuous in \( x \)-direction. The solutions to hyperbolic equations can form discontinuous shock waves, as well as the bottom and the cross sectional shape are not necessarily smooth. The discontinuities in the states are governed by the Rankine Hugoniot condition

\[
\Delta \left( \frac{Q^2}{A} + p(x, A) \right) = \left( \frac{Q^2_R}{A_R} + p(x_R, A_R) \right) - \left( \frac{Q^2_L}{A_L} + p(x_L, A_L) \right)
\]

where \( \alpha, \beta \) represent the indices according to the states on the left and right hand side of the shock.

For stationary solutions the shock speed is \( s = 0 \) and thus the equality of the fluxes results

\[
f(x_R, u_R) = f(x_L, u_L).
\]  

(3.11)  

The first component of this relation guarantees the conservation of mass i.e. \( \Delta Q = 0 \). Here and in the following the operator \( \Delta \) is a shorthand for the jump at the discontinuity \( (\cdot)_R - (\cdot)_L \). The second line of (3.11) is given as

\[
\Delta \left( \frac{Q^2}{A} + p(x, A) \right) = \left( \frac{Q^2_R}{A_R} + p(x_R, A_R) \right) - \left( \frac{Q^2_L}{A_L} + p(x_L, A_L) \right) = \Delta A \left( - |v_R v_L| + \frac{p(x_R, A_R) - p(x_L, A_L)}{A_R - A_L} \right),
\]

where we used the equality of \( Q_R = Q_L \) and the velocities \( v_L = \frac{Q_L}{A_L} \) and \( v_R = \frac{Q_R}{A_R} \). By defining \( \tilde{v}^2 = |v_R v_L|, p' + \frac{p}{A} = \frac{p(x_R, A_R) - p(x_L, A_L)}{A_R - A_L} \), the jump in the momentum flux \( \phi \) is given by

\[
\Delta \phi = \left( -\tilde{v}^2 + p' + \frac{p}{A} \right) \Delta A.
\]  

(3.12)  

The notation on the right hand side is motivated by the following approximation of the continuous case

\[
\frac{p(x_R, A_R) - p(x_L, A_L)}{A_R - A_L} \approx \partial_p + \frac{\partial_p}{\partial A} A.
\]

Starting at the other end, we aim that the relation (2.6) should hold for discontinuous steady states as well

\[
\left( \frac{1}{2} \tilde{v}^2 + gh(x, A) + g z \right)_R = \left( \frac{1}{2} \tilde{v}^2 + gh(x, A) + g z \right)_L.
\]
Multiplying this equation by $\overline{A} = \frac{1}{2}(A_R + A_L)$ gives

$$-g\overline{A}\Delta z - g\overline{A}\Delta h = \left(\frac{A_R + A_L}{2}\right)\left(\frac{u_R^2 - v_L^2}{2}\right)$$

$$= \left(\frac{1}{2}\right)^2 \left(A_R v_R^2 + A_L v_L^2 - A_R v_R^2 - A_L v_L^2\right)$$

$$= \left(\frac{1}{2}\right)^2 \left(2A_R v_R^2 - A_R v_R^2 + A_L v_L^2 - A_R v_R^2\right)$$

$$= \left(\frac{1}{2}\right)^2 \left(A_L v_R^2 + 2A_R v_R v_L + A_L v_L^2\right) - \left(A_R v_R^2 + 2A_R v_R v_L + A_R v_L^2\right)$$

$$= \left(\frac{1}{2}\right)^2 \left(A_L (v_R + v_L)^2 - A_R (v_R + v_L)^2\right)$$

$$= -\left(\frac{v_R + v_L}{2}\right)^2 \Delta A .$$

Rearranging the terms and defining the arithmetic average $\overline{\tau} = \frac{1}{2}(v_R + v_L)$ leads to

$$\left(-\overline{\tau}^2 + g\overline{A} \frac{\Delta h}{\Delta A}\right) \Delta A = -g\overline{A}\Delta z , \quad (3.13)$$

where the term $-g\overline{A}\Delta z$ is a direct approximation to the source term representing the elevation of the bottom. In order to obtain also an approximation to $S_w$ of (2.1), we define

$$SW = p(x_R, \overline{A}) - p(x_L, \overline{A}) \approx \int_{x_L}^{x_R} S_w(x, \overline{A}) dx .$$

Adding this to (3.13) gives

$$\left(-\overline{\tau}^2 + g\overline{A} \frac{\Delta h}{\Delta A} + \frac{SW}{\Delta A}\right) \Delta A = -g\overline{A}\Delta z + SW .$$

By defining $\overline{p} = g\overline{A} \frac{\Delta h}{\Delta A}$ and $\frac{\overline{p}}{\overline{A}} = \frac{SW}{\Delta A}$ a similar structure as in (3.12) is obtained

$$\left(-\overline{\tau}^2 + \overline{p} + \frac{\overline{p}}{\overline{A}}\right) \Delta A = -g\overline{A}\Delta z + SW . \quad (3.14)$$

Here $\overline{p} + \frac{\overline{p}}{\overline{A}}$ can be interpreted as an approximation to $p' = \partial p + \frac{\partial p}{\partial A}$ via the relations (2.4), while $p' + \frac{\overline{p}}{\overline{A}}$ uses a direct finite difference approximation. By solving (3.14) w.r.t. $\Delta A$ and inserting it into (3.12), we obtain some modified representation of the source term

$$\Delta \phi = \left(-\overline{\tau}^2 + p' + \frac{\overline{p}}{\overline{A}}\right) \left(-g\overline{A}\Delta z + SW\right) = : S \phi . \quad (3.15)$$

As the notation indicates, this expression is an approximation to the original source term. In the case, where the denominator of the correction factor is close to zero, the factor is set to 1 and thus the naive approximation $-g\overline{A}\Delta z + SW$ is used. In some special situations, e.g. for supersonic flow, it might occur that the correction factor becomes negative. Nevertheless this does not influence its validity, as discussed in (3.14).

In order to incorporate this source term into the Augmented Riemann Solver in a well-balanced way, we subtract it directly from the flux difference before the splitting [3].
Therefore the source terms have to be adapted to the augmented states. The first line of (3.7) is considering the jump in $A$. For this equation the relation (3.15) can be combined with (3.12) to obtain the correction for the first component

$$SA := \left(\begin{array}{c}
-\frac{gA\Delta z + SW}{-\nu^2 + \nu + \frac{\nu}{A}} \\
\end{array}\right).$$

(3.16)

The second line of (3.7) remains unchanged due to the conservation of mass $\Delta Q = 0$. For the difference in the momentum fluxes we can use (3.15) directly. Thus the augmented splitting of the flux difference with source term correction is given by

$$\left(\begin{array}{c}
AR - AL \\
QR - QL \\
\phi_R - \phi_L \\
\end{array}\right) = \left(\begin{array}{c}
SA \\
0 \\
S\phi \\
\end{array}\right) = \sum_{k=1}^{3} \beta^k w^k.$$  

(3.17)

Subtracting the source term at this point can be regarded as a simple way of an upwinding of the right hand side $[3, 4]$. In case of an steady state the variations in $\Delta u$ are directly balanced by the source term and thus no fluctuations can arise.

3.2.5 Some properties of the solver

**Depth Positivity:** The positivity of the new arising middle state is an important feature of solvers for the StVE. In the present case this is guaranteed by the choice of the Einfeld speeds. The update formula of interest is the first equation of the wave splitting (3.7). In absence of a source term only the speeds $s^1$, $s^3$ affect the area of the middle state $A^*_E$, since $w^2$ is zero in its first entry. Solving this equation for the middle state we obtain

$$A^*_E = \frac{QL - QR + s^3AR - s^1AL}{s^3 - s^1}.$$

The following easy computation shows that $A^*_E$ is always greater equal zero,

$$\left(s^3 - s^1\right)A^*_E = QL - QR + s^3AR - s^1AL$$

$$\geq ALv_L - ARv_R + (v_R + \sqrt{\partial_{AP}(AR)})AR - (v_L - \sqrt{\partial_{AP}(AL)})AL$$

$$\geq AR\sqrt{\partial_{AP}(AR)} + AL\sqrt{\partial_{AP}(AL)} \geq 0,$$

since $s^3 - s^1 > 0$. This estimate holds as long as the speeds $s^1$, $s^3$ satisfy $s^1 \leq v_L - \sqrt{\partial_{AP}(AL)}$ and $s^3 \geq v_R + \sqrt{\partial_{AP}(AR)}$. Thus the generalized Einfeldt speeds (3.10) guarantee a positive middle state, if no source term is present.

This situation slightly changes in the presence of source terms. Since subtracting the source terms can yield two different middle states, $A^*_L$ and $A^*_R$, we have to bound $SA$ such that both remain positive for all times. Nevertheless the mass in the splitted state can not differ from the one of the classical situation, as it remains unchanged by the sources. Thus in the case of subsonic flow, $s^1_L < 0 < s^3_L$, we can express the new states in terms of $A^*_E$

$$s^3_L(A_R - A^*_E) + s^1_L(A^*_E - AL) = s^3_L(A_R - A^*_L) + s^1_L(A^*_L - AL).$$

By either rising $A^*_R$ and decreasing $A^*_L$, or vice versa, we obtain the following bounds as maximal cross sectional areas

$$A^*_R = \frac{s^3_L - s^1_L}{s^3_L}A^*_E \quad \text{ if } A^*_L = 0,$$
\[ A_L^* = \frac{(s_3^E - s_1^E)}{-s_1^E} A_E^* \quad \text{if} \ A_R^* = 0. \]

Combining both gives upper and lower bounds for \( A_R^* - A_L^* \),

\[ \frac{(s_3^E - s_1^E)}{-s_1^E} A_E^* \leq A_R^* - A_L^* \leq \frac{(s_3^E - s_1^E)}{s_1^E} A_E^*. \]

Consequently these are the bounds, which may not be exceeded by the jump due to the source term, i.e.

\[ \frac{(s_3^E - s_1^E)}{-s_1^E} A_E^* \leq S \leq \frac{(s_3^E - s_1^E)}{s_1^E} A_E^*. \]

For supersonic flow no actual limits are needed, since if a negative middle state arises, this is directly refilled by the other wave traveling in the same direction \[14\]. Additionally we provide some special treatment of dry cells in the presence of a given bottom elevation. Without loss of generality the right state is assumed to have zero depth, i.e. \( A_R = 0 \).

Before starting the computation we verify whether the right cell can be wetted or if the slope of the bottom impedes the propagation of the flow. This is easily achieved by the help of the estimated middle state \( u^* \), which is used for the computation of the generalized Einfeldt speeds. If the height \( h(A^*) \) exceeds the jump in the bottom profile, the usual procedure is started. In the case, where the new height is not able to reach the bottom of the neighboring cell, a reflection on a solid wall is assumed. This is simply achieved by mirroring the velocity components \[16\].

**Exact C-Property:** The balancing of steady states can be directly deduced from the incorporation of the source terms. The exact C-Property refers to an exact balancing of the source terms and the approximation of the flux in the situation of flow at rest \[4, 31\]. In this particular case, i.e. \( Q_L = Q_R = 0 \) and \( h(x_R, A_R) + z_R = h(x_L, A_L) + z_L \), the expression (3.15) reduces to

\[ S \phi = \tilde{p}_x + \frac{\tilde{p}}{\Delta A} (-g \bar{A} \Delta z + SW) \]

\[ = \frac{p(x_R, A_R) - p(x_L, A_L)}{g \bar{A} h(x_R, A_R) - h(x_L, A_L)} + SW (-g \bar{A} (z_R - z_L) + SW) = \Delta p. \]

For the flow at rest, this represents exactly the jump in the momentum flux and thus no wave is generated by the third component of (3.17). The second line of (3.17) just gives the conservation of mass, which is trivial in this case \( Q_L = Q_R = 0 \). Thus it is only left to show that \( S \Delta A = \Delta A \).

Analyzing the formula (3.16) for this special situation gives

\[ S \Delta A = -g \bar{A} \Delta z + SW \frac{\tilde{p}}{\Delta A} = \frac{\tilde{p}}{\Delta A} \frac{g \bar{A} \Delta h + SW}{\Delta A} = \Delta A, \]

since \( \Delta h = -\Delta z \). As all jumps in the augmented state are balanced by the source terms, no waves will be generated at this interface and the neighboring states will remain unchanged.

### 3.2.6 Summary of the solver

First we compute approximately the middle state \( u^* \), with a procedure similar to the one in \[29\]. At this point it can already be checked, whether the slope in the bottom topography hinders a possible wetting of a neighboring dry state. In the next step, we compute the wave speeds and the splitting vectors according to (3.8) and (3.10). The influences of the source terms we handle with (3.15) and (3.16). Based on these the strength of the numerical waves is determined by
Finally we determine the fluctuations according to (3.9). In order to obtain a second order accurate method, the final update is performed with

\[ u_{i+1}^j = u_i^j + \frac{\Delta t}{\Delta x} \left( A^- \Delta u_{i+1/2} + A^+ \Delta u_{i-1/2} \right) + \frac{\Delta t}{\Delta x} \left( \tilde{F}_{i+1/2} - \tilde{F}_{i-1/2} \right). \]

Here \( \tilde{F}_{i+1/2} \) and \( \tilde{F}_{i-1/2} \) are second order corrections of the fluxes. These are computed from the first order waves with help of an appropriate limiter function. A detailed description can be found in e.g. [3, 21]. For situations including dry states, we have to assure that the correction will not violate the depth positivity of the solver [15].

4 Numerical test cases

In this section we present some numerical test cases, focusing on the well-balancing of the source terms. The general properties of the scheme are equivalent to those of [15] and not repeated here. All computations are run on a grid of 100 points and the time step is chosen adaptively according to a CFL number 0.95. The gravitational acceleration in (2.2) is taken as \( g = 9.81 \).

4.1 The source terms - Well balanced

For the well-balancing of source terms various benchmarks have been proposed. In the context of the shallow water equations examples can be found in [2, 3] and for channels of variable rectangular cross-section in e.g. [13, 31]. In [27] variable cross-sections of the special form \( w(x, A) = \sigma(x) h(A)^k \) for fixed \( kN \) and some function \( \sigma: \mathbb{R} \rightarrow \mathbb{R} \) are considered.

The majority of test cases consider a steady state for a particular flow regime. Among these the most popular one is the 'lake at rest' example. For \( Q \equiv 0 \) the even surface of the water has to remain planar for all times. As proven in section 3.2.5, the actual solver preserves these states up to the machine precision. Figures showing the results of such computations are omitted, since they do not give any additional insights.

4.1.1 Rectangular profile

More challenging are the test cases presented in [13]. For a rectangular channel of variable breadth and bottom subsonic, transsonic and supersonic flows are examined. All these scenarios use the same channel profile, but reach different stationary states, according to the different boundary conditions. The flow is considered in an open rectangular channel along the interval \([0, 3]\). At its center the bottom \( z \) has a smooth elevation

\[ z(x) = \begin{cases} 0.1 \cos^2(\pi(x - 1.5)) & \text{if } |x - 1.5| < 0.5 \\ 0 & \text{else} \end{cases} \]

Within the same region the breadth \( w(x) \) is contracted in a similar way

\[ w(x) = \begin{cases} 1 - 0.1 \cos^2(\pi(x - 1.5)) & \text{if } |x - 1.5| < 0.5 \\ 1 & \text{else} \end{cases}. \]

The initial states \((A_0, Q_0)^T\) are chosen to satisfy \( h(A_0) + z(x) \equiv 1 \) and \( Q_0 \equiv const \). The actual value for \( Q_0 \) given by the boundary conditions of the actual example. All figures show the results at \( t = 30 \), since no major changes can be observed later on. All results are compared to the solutions obtained using the solver developed in [13], which gives results close to the exact solutions. In the figures the reference scheme is plotted by a red and dashed line, while the augmented solver is indicated by the blue line.

The first example is dedicated to purely subsonic flow. The boundary conditions are given by \( Q(t, 0) = 0.5 \sqrt{g} \) at the left and \( h(A(t, 3)) = 1 \) at the right end. In Figure 3 the surface elevation
and the flow are plotted at time $t = 30$. The flow is constant along the channel, as it is necessary for any steady state. The surface of the water has a downward dent at the bottleneck. Both quantities show a very good agreement with the benchmark solution.

The second test considers transonic flow. The setup is similar to the subsonic case, but the speed of the flow is slightly increased. At the boundaries the following values are imposed, $Q(t, 0) = 0.6 \sqrt{g}$ on the left and $h(A(t, 3)) = 1$ on the right hand side. This leads to a stationary transonic shock at the end of the bump, as shown in Figure 4. In the flow, a peak at the position of the shock is observed. This is due to the jump in $A$, but has no negative effect on the general behavior. This artifact is as well present in the reference solution. A very small additional peak of the augmented solver can be observed, which might be caused by the switching to a naive discretization of the source terms for states close to sonic flow. Nevertheless a good agreement with the reference scheme is obtained.

The last test with this profile is dedicated to supersonic flow. Here only boundary conditions on the left side are given, i.e. $Q(t, 0) = 1.7 \sqrt{g}$ and $h(A(t, 3)) = 1$. As depicted in Figure 5, the flow is constant on the whole domain. For the depth of the water a symmetrical dent is formed, pointing upwards. Again no significant difference to the result of the scheme of 13 can be observed.

### 4.1.2 Hexagonal profile

For general non-rectangular profiles, no common benchmarks exist. Thus we adopt the above ones onto the flow in the hexagonal profile (3.1). We shrink the radius $r(x)$ of the tube at the center.
of the interval $[0, 3]$

$$r(x) = \begin{cases} 
0.5 - 0.05 \cos^2(\pi(x - 1.5)) & \text{if } |x - 1.5| < 0.5 \\
0.5 & \text{else}.
\end{cases}$$

The elevation of the bottom is kept as before

$$z(x) = \begin{cases} 
0.1 \cos^2(\pi(x - 1.5)) & \text{if } |x - 1.5| < 0.5 \\
0 & \text{else}.
\end{cases}$$

As boundary condition, $h(A(t, 0)) = 0.5$ on the right and $Q(t, 3) = 0.5\sqrt{\frac{g}{2}(\frac{5}{2}0.25)^{\frac{3}{2}}}$ on the left side are fixed. The corresponding initial conditions are $h(A_0(x)) + z(x) \equiv 0.5$ and $Q_0(x) = Q(0, 3)$.

In Figure 6, the surface of the water $h + z$ and the flow $Q$ at $t = 30$ are plotted for a computation using 150 cells. Again a downward dent evolves above the hump, while the flow is constant over the whole domain. As intended, the results look very similar to the case of the rectangular channel. Here we recall that although the flow is situated in the region, where the hexagonal profile has vertical walls, the balancing involves also nonlinear parts of the cross-section.

### 4.2 A small perturbation

So far we have only considered steady state solutions. For the well-balancing it is also important, that the original qualities of the solver remain unaffected by the corrections due to the source term.
Therefore we perform the test proposed in [3], considering a small perturbation of a given steady state. As this test case was developed for the shallow water equations, we reduce our scheme to this classical case. Consider the the lake at rest on the interval $[0, 1]$, but add a small additional amount of water at one end. The bottom slope is similar to the above situations

$$z(x) = \begin{cases} 0.25 \cos(\pi (x - 0.5)) + 1 & \text{if } |x - 0.5| < 0.1 \\ 0 & \text{else}, \end{cases}$$

while the width is constant, $w \equiv 1$. Furthermore the gravitational acceleration is set to $g = 1$. Within this setup two scenarios are compared, which differ only by the initial states

$$h(A_0(0, x)) + z(x) = \begin{cases} 1 + \epsilon & \text{if } 0.1 < x < 0.2 \\ 1 & \text{else}, \end{cases}$$

where $\epsilon$ is either $\epsilon = 0.2$ or $\epsilon = 0.01$. In both cases is $Q_0 \equiv 0$. At the boundaries free outflow conditions are imposed. From the initial disturbance of the even surface two small waves are generated. One wave leaves the computational domain at the left boundary. The other wave travels to the right and passes the hump in the bottom. There the wave is slightly deformed and two small waves are send back, while the main part moves further to the right.

![Figure 7: A small perturbation passing a hump. The water level $h + z$ for $\epsilon = 0.2$ (left) and $\epsilon = 0.01$ (right) at $t = 0.7$. Compared are the augmented solver (solid, blue) and the method of [13] (red, dashed).](smallperturbation)

As depicted in Figure 7, the general behavior is equal for both initial conditions, but scaled in their size. Here the water level $h + z$ of a computation with 150 cells is shown. Compared are the augmented solver (solid, blue) and the method of [13] (red, dashed). Both schemes show very close results. For the smaller disturbance the rarefaction wave traveling to the right does not evolve sharply anymore. This is not only caused by numerical diffusion, but is a consequence of the reduced difference in the speeds of propagation of the front and the rear end of the rarefaction wave.

## 5 Conclusion

In this paper we developed a well-balanced solver for the Saint Venant equations with variable cross-section. An important ingredient for this method is an exact representation of the relations between the pressure law and the depth of the water. These are used in the construction of the well-balanced source term approximations. The predicted properties of the solver are confirmed by several numerical test cases.
References


