ON THE COMPUTATION OF ROLL WAVES*

SHI JIN¹ AND YONG JUNG KIM²

Abstract. The phenomenon of roll waves occurs in a uniform open-channel flow down an incline, when the Froude number is above two. The goal of this paper is to analyze the behavior of numerical approximations to a model roll wave equation $u_t + uu_x = u$, $u(x,0) = u_0(x)$, which arises as a weakly nonlinear approximation of the shallow water equations. The main difficulty associated with the numerical approximation of this problem is its linear instability. Numerical round-off error can easily overtake the numerical solution and yields false roll wave solution at the steady state. In this paper, we first study the analytic behavior of the solution to the above model. We then discuss the numerical difficulty, and introduce a numerical method that predicts precisely the evolution and steady state of its solution. Various numerical experiments are performed to illustrate the numerical difficulty and the effectiveness of the proposed numerical method.

Mathematics Subject Classification. 35L65, 65M06, 76B15.

Received: September 9, 2000. Revised: January 29, 2001.

1. Introduction

The phenomenon of roll waves occurs in a uniform open-channel flow down an incline, when the Froude number is above two. Such a physical problem is described by the river equations—the shallow water equations with source terms modeling the balance between the slope and the friction of the river bottom. The two variables are the depth h and the mean velocity v, solving the system

$$h_t + (hv)_x = 0,$$

$$(hv)_t + \left(hv^2 + \frac{1}{2}gh^2\right)_x = ghS - C_f v^2,$$
(1.1)

where g is the gravitational constant, S is the slope of the river bottom and C_f is the friction coefficient. It is long known [26] that the initial value problem of the above system is linearly stable if

$$S < 4C_f, \tag{1.2}$$

namely, when the model describes flow in a river bed where the slope S is usually small compared to the friction coefficient C_f . When the system (1.1) is written in dimensionless form, the above inequality is equivalent to that

Keywords and phrases. Roll wave, conservation laws with source term, round-off error, shock capturing methods.

^{*} Research was supported in part by NSF grant DMS-9704957 and DMS-0072374.

Department of Mathematics, University of Wisconsin, Madison, WI 53706, USA. e-mail: jinmath.wisc.edu.

 $^{^2}$ Institute for Mathematics and its Applications, University of Minnesota, Minneapolis, MN 55455-0436, USA. e-mail: yjkimima.umn.edu

the Froude number is smaller than 2. This condition may be violated for steep channels, where the uniform flow eventually evolves to a series of breaking waves or bores that are separated by sections of gradually varying flow, in a staircase pattern. Under suitable conditions these formations constitute *discontinuous* periodic traveling waves, and are known as "roll waves" [25]. The phenomenon of roll waves has been actually observed (and photographed!) in inclined spill ways and river beds, see Cornish [4].

The mechanism leading to the instability was revealed by Jeffreys [12] in an early study of roll waves. Dressler constructed nonlinear periodic solutions that consist of piecewise smooth profiles separated by discontinuities (shocks) using the river equations [6]. Novik [23] proposed to model the roll waves by a Burgers equation with a source term. Continuous roll waves in the unstable regime were constructed by Needham and Merkin using the river equations with a diffusive term in the momentum equation [22]. It was realized by Whitham [26], and more recently explicitly formulated by Kranenburg [16], that one can obtain a weakly nonlinear asymptotic approximation for (1.1), which turns out to be a Burger's equation with a source term. Specifically, if one uses the ansatz

$$h = \overline{h} + \epsilon h_1(\xi, t), v = \overline{v} + \epsilon v_1(\xi, t), \qquad (1.3)$$

where $\overline{h}, \overline{v} = \sqrt{sg\overline{h}/C_f}$ are the equilibrium constants, $\xi = (x - ct)/\epsilon$ with $c_{\pm} = \overline{v} \pm \sqrt{g\overline{h}}$ being the equilibrium characteristic speeds, then the leading order approximation of (1.1) that corresponding to the wave speed c_{+} becomes (the readers are referred to the Appendix for details of this asymptotic result):

$$\partial_t h_1 + \frac{3}{2} \sqrt{\frac{g}{\overline{h}}} h_1 \partial_x h_1 = \sqrt{\frac{g}{\overline{h}}} s \left(1 - \sqrt{\frac{4C_f}{s}} \right) h_1. \tag{1.4}$$

The sign on the right hand side shows that, when the stability condition (1.2) is not satisfied, it induces exponential growth, thus the linear instability of (1.4), just like the original system (1.1). However, the damping mechanism of the nonlinear convection (known as the N-wave decay [20]) in the Burger's equation provides a competing force with the growth term, and under suitable conditions the solution does not tend to infinity. For a periodic problem with zero total mass, it converges to the steady state "roll wave" solution which is a bounded solution [16].

The model (1.4) arises not only in the river equations. It is the asymptotic approximation to general hyperbolic systems with relaxations when some characteristic stability condition is violated. See [14] for details. Theoretical justification of this asymptotic limit, with the addition of artificial viscosity, was made in [14].

Model (1.4), and its asymptotic origin, the river equations (1.1), are examples of hyperbolic systems with geometrical source terms. In recent years there have been increasing interests in developing shock capturing numerical methods for related problems that are able to capture the steady state solution with a better accuracy, since it is known that a scheme that ignores the steady state structure of the equations will not produce accurate steady state numerical solutions [10]. Several successful numerical approaches, which take into consideration of the balance between the flux and the source term, exist in the literature [1–3,7,9,10,13,19,24]. All these methods capture the steady state solutions either exactly or approximately with a formally second order accuracy.

Due to the instability of the roll wave, steady state preserving seems to be a necessary but not sufficient condition for the computation of steady state roll waves using evolutional models such as (1.4) and (1.1), as shown in Section 3. For such a physically unstable problem, the round-off errors, due to the finite precision arithmetic, can easily build up and get amplified, which eventually destroys the long time and the steady state structures of the roll wave solution. We have not found any report on the success of numerical computations of the phenomena to be studied in this paper.

The goal of this paper is to analyze the behavior of numerical approximations to the model problem

$$u_t + uu_x = u$$

 $u(x,0) = u_0(x),$ (1.5)

where $x \in \mathbb{R}, t \in \mathbb{R}^+, u_0 \in L^1$ and $u = u(x,t) \in \mathbb{R}$. We first study the analytic behavior of its solution in Section 2. In Section 3 we discuss the numerical difficulty and introduce a numerical method that predicts precisely the evolution and steady state of its solution. The key for the success of this new method is that it preserves the zero mass over each period of the wave if odd number of cells is used over each period. More numerical experiments are conducted in Section 4 and a few remarks are made to conclude the paper in Section 5.

2. Analytic behavior of the solution

2.1. General asymptotic behavior

We introduce the change of variables

$$s = e^t - 1, v = e^{-t}u,$$
 (2.1)

then (1.5) is transformed to the inviscid Burgers equation,

$$v_s + vv_x = 0,$$

 $v(x, 0) = u_0(x).$ (2.2)

The asymptotic behavior of inviscid problem has been studied in various contexts. It is well-known that a classical solution does not exist in general, and hence weak solutions are considered in a distribution sense. Since this kind of weak solutions are not unique, it is always considered with entropy condition which implies the physically meaningful solution, [17].

Asymptotic behavior of the solution v(x,s) of (2.2) is determined by two invariant variables

$$p = -\inf_{x} \int_{-\infty}^{x} u_0(y) dy, \qquad q = \sup_{x} \int_{x}^{\infty} u_0(y) dy, \tag{2.3}$$

and an N-wave

$$\mathcal{N}_{p,q}(x,s) = \begin{cases} x/s, & -\sqrt{2ps} < x < \sqrt{2qs}, \\ 0, & \text{otherwise.} \end{cases}$$
 (2.4)

Note that if the initial total mass $M = \int_{-\infty}^{\infty} u_0(y) dy$ and one of the two invariant constants are given, the other one is obtained by

$$q = \sup_{x} \int_{x}^{\infty} u_0(y) dy = \sup_{x} (M - \int_{-\infty}^{x} u_0(y) dy) = M - p.$$
 (2.5)

We have an optimal estimate for the solutions of Burgers, (2.2), in Appendix B. If we translate it under the change of variable (2.1) we get the following lemma.

Lemma 2.1. Let u(x,t) be the solution of (1.5). Suppose that the data have compact support supp $(u_0) \subset [A,B]$. Then, for t > 0,

$$u(x,t) = 0, \quad x < A - \sqrt{2p(e^t - 1)}, \ x > B + \sqrt{2q(e^t - 1)},$$
 (2.6)

and for $B - \sqrt{2p(e^t - 1)} \le x \le A + \sqrt{2q(e^t - 1)}$

$$A \le x - (1 - e^{-t})u(x, t) \le B. \tag{2.7}$$

In this lemma we can see that support of the solution of (1.5) expands exponentially and the solution has an N-wave structure which can also be obtained from (2.4) after using the transformation (2.1). This gives

$$e^{t} \mathcal{N}_{p,q}(x,s(t)) = \begin{cases} x \frac{e^{t}}{e^{t}-1}, & -\sqrt{2p(e^{t}-1)} < x < \sqrt{2q(e^{t}-1)}, \\ 0, & \text{otherwise.} \end{cases}$$

After neglecting the small terms in long time we get the following N-wave for (1.5)

$$N_{p,q}(x,t) = \begin{cases} x, & -\sqrt{2p} e^{t/2} < x < \sqrt{2q} e^{t/2}, \\ 0, & \text{otherwise,} \end{cases}$$
 (2.8)

which grows exponentially in time. Here we can clearly see that the invariance variables p, q for the Burgers equation also play the key role in the asymptotic behavior of (1.5).

Remark 2.1. We are particularly interested in the case p = q = 0. In the case the N-wave (2.8) does not give any information and hence we need more detailed analysis for the evolution of the solution. Nevertheless, the lemma shows that the support of the solution is the same as the initial data. This implies that the solution has no influence outside the initial support.

2.2. Emerging roll waves

In this section we prove that the solution of (1.5) converges to roll waves under certain conditions on the initial data $u_0 \in L^1_{loc}(\mathbb{R})$. The structure of emerging roll waves is decided by the structure of the initial data and we can predict it from the initial data.

From the observation in Remark 2.1 it is natural to try to divide the domain of the initial data into sub-domains on which the solution evolves independently. Here we introduce a natural concept for it. An interval (a_k, b_k) is called *self-contained* if initial data satisfy

$$\inf_{a_k < x < b_k} \int_{a_k}^x u_0(y) dy = 0, \quad \sup_{a_k < x < b_k} \int_x^{b_k} u_0(y) dy = 0.$$
 (H)

In the rest of the article we only consider initial data $u_0 \in L^1_{loc}(\mathbb{R})$ which have a partition $\mathbb{R} = \bigcup_k (a_k, b_k]$ that consists of *self-contained* finite intervals. This condition is clearly satisfied by mean zero periodic functions.

Lemma 2.2. Let $u_0 \in L^1_{loc}(\mathbb{R})$ be L-periodic and have mean zero, i.e.,

$$u_0(x+L) = u_0(x), \qquad \int_x^{x+L} u_0(y) dy = 0, \quad x \in \mathbb{R}.$$

Then there exists a partition $\mathbb{R} = \bigcup_k (a_k, b_k]$ satisfying (H) for all k.

Proof. Let $U_0(x) = \int_0^x u_0(y) dy$. Since u_0 has mean zero, U_0 is also L-periodic. Since U_0 is continuous, there exists $a \in [0, L]$ such that U_0 attains its minimum at the point, x = a. Hence we get

$$\inf_{a < x < a + L} \int_{a}^{x} u_0(y) dy = 0, \quad \sup_{a < x < a + L} \int_{x}^{a + L} u_0(y) dy = 0.$$

So we take the partition as $a_k = b_{k-1} = a + kL$.

Suppose the initial data $u_0 \in L^1_{loc}(\mathbb{R})$ have a partition satisfying (H). Let $\psi_k(x,t)$ be the solution of (1.5) with the initial data over an interval $[a_k, b_k]$, *i.e.*,

$$\psi_k(x,0) = \begin{cases} u_0(x), & a_k < x < b_k \\ 0, & \text{otherwise.} \end{cases}$$
 (2.9)

Then one can easily check that $u(x,t) = \sum \psi_k(x,t)$ is the solution with initial data u_0 . This implies that it suffices to consider the structure of the solution over each self-contained interval to determine the global structure.

Convergence to roll waves has been studied in [21] when initial data are periodic and have mean zero and the flux is strictly convex. We summarize some of the results we need in our context in the following lemma.

Lemma 2.3. Let u(x,t) be the solution of (1.5) with zero mean periodic initial data $u_0 \in C^0(\mathbb{R})$ and let $\mathcal{Z}(t) = \{x \in \mathbb{R} : u(x+,t) = u(x,t) = 0\}$. Then $\mathcal{Z}(t') \subset \mathcal{Z}(t)$ for all 0 < t < t'. If $a,b \in \mathcal{Z}(t)$ for all t > 0 and there is no such a point between a and b, then the uniform convergence holds on [a,b],

$$u(x,t) \to R_{a,b}(x) = \begin{cases} x - a, & a < x < \frac{a+b}{2}, \\ x - b, & \frac{a+b}{2} < x < b, \\ 0, & otherwise, \end{cases}$$
 as $t \to \infty$. (2.10)

The function $R_{a,b}(x)$ is a steady state of (1.5). In fact all the functions given by $\sum R_{a_i,b_i}(x)$ with ... $\leq a_i < b_i \leq a_{i+1} < ...$ are the steady states. Now we apply this result together with the previous observations to see how the structure of the emerging roll waves is determined from its initial data.

Proposition 2.4. Let u(x,t) be the solution of (1.5) with continuous initial data $u_0 \in L^1_{loc}(\mathbb{R})$ that has a partition $\mathbb{R} = \bigcup_k (a_k, b_k]$ consisting of self-contained intervals. Suppose that $\int_{a_k}^x u_0(y) dy > 0$ for all $x \in (a_k, b_k)$ and there are only finitely many number of zero points of u_0 in (a_k, b_k) . Then for all k, $u(x, t) \to R_{a_k, b_k}(x)$ as $t \to \infty$ for $x \in [a_k, b_k]$.

Proof. Let $\psi_k(x,t)$ be the solution of (1.5) with initial data (2.9). Since $\operatorname{supp}\psi_k\subset [a_k,b_k]$ for all k, the solution u(x,t) over the interval (a_k,b_k) is same as the solution $\bar{u}(x,t)$ of (1.5) with periodic initial data,

$$\bar{u}(x,0) = u_0(x), \quad a_k < x < b_k, \quad \bar{u}(x+L,0) = \bar{u}(x,0),$$

where $L = b_k - a_k$. The generalized characteristics through the end points a_k, b_k are strait lines $x = a_k, x = b_k$ and they do not intersect any of the shock since p = q = 0. So $a_k, b_k \in \mathcal{Z}(t)$ for all t > 0.

Suppose that there exists $c_k \in (a_k, b_k)$ such that $c_k \in \mathcal{Z}(t)$ for all t > 0 and there is no such a point between a_k and c_k . Then Lemma 2.3 implies that $u(x,t) \to R_{a_k,c_k}(x)$ for all $x \in (a_k, c_k)$ and hence $\int_{a_k}^{c_k} u(x,t) dx \to 0$. On the other hand we have

$$\frac{d}{dt} \int_{a_k}^{c_k} u(x,t) dx = \int_{a_k}^{c_k} u_t(x,t) dx = \int_{a_k}^{c_k} u(x,t) dx.$$

So $\int_{a_k}^{c_k} u(x,t) dx = e^t \int_{a_k}^{c_k} u_0(x) dx$, which contradicts the fact $\int_{a_k}^{c_k} u(x,t) dx \to 0$. So there is no such a constant c_k and hence we get the convergence $u(x,t) \to R_{a_k,b_k}(x)$ for $x \in [a_k,b_k]$.

With this proposition, we can determine the emerging roll waves from the initial data. In the following example we consider a periodic initial data with mean zero.

Example 2.5. Consider 1-periodic initial data,

$$u_0(x) = \begin{cases} 0.01\sin(8\pi x) + 0.001\sin(4\pi x), & 0 < x < 0.5, \\ 0.01\sin(8\pi x), & 0.5 < x < 1. \end{cases}$$
 (2.11)

We can easily check that $\int_0^x u_0(y) dy$ has minimum at $x = 0, \frac{1}{2}, \frac{3}{4}, 1$ and hence $(0, \frac{1}{2}] \cup (\frac{1}{2}, \frac{3}{4}] \cup (\frac{3}{4}, 1]$ is a partition of interval (0, 1] consists of self-contained intervals that satisfy the condition of Proposition 2.4. So for 0 < x < 1, the solution converges to roll waves given by

$$u(x,t) \to R_{0,\frac{1}{2}}(x) + R_{\frac{1}{2},\frac{3}{4}}(x) + R_{\frac{3}{4},1}(x), \quad as \quad t \to \infty.$$
 (2.12)

2.3. Instability of roll waves

Variables

$$p(t) = -\inf_{x} \int_{-\infty}^{x} u(y, t) dy, \quad q(t) = \sup_{x} \int_{x}^{\infty} u(y, t) dy, \tag{2.13}$$

are constants for the L^1 solution u(x,t) of homogeneous conservation laws with convex flux, for example the inviscid Burgers equation. So M(t) = q(t) - p(t) is also a constant. But for inhomogeneous problems like (1.5) these are functions in time. One can easily check that the total mass M(t) satisfies

$$M'(t) = \int_{-\infty}^{\infty} u_t(x, t) dx = \int_{-\infty}^{\infty} u(x, t) dx = M(t),$$

and hence it grows exponentially. Suppose $\int_{-\infty}^{x} u(x,t)dx$ has infimum along a differentiable curve x=g(t), then u(g(t),t)=0, and hence

$$p'(t) = g'(t)u(g(t), t) + \int_{-\infty}^{g(t)} u_t(x, t) dx = p(t).$$

So we conclude that M, p, q are given by

$$M(t) = M(0)e^{t}, \quad p(t) = p(0)e^{t}, \quad q(t) = q(0)e^{t}.$$
 (2.14)

From (2.14) we can see that if p, q, M are initially zero,

$$p(0) = q(0) = M(0) = 0, (2.15)$$

these variables will remain zero all the time. However, if a small perturbation to the initial data makes (2.15) be violated, then (2.14) shows exponential growth of this initial perturbation. This phenomenon contributes to the numerical difficulty when solving the physically unstable problem, as will be shown in next section.

We can easily check the same kind of instability from periodic solutions. Suppose there is an partition $\cup(a_k,b_k]$ of initial data u_0 such that all the intervals are self-contained except one of them, say (a_0,b_0) . Let $\psi_k(x,t)$ be the solution of (1.5) with initial data (2.9). Then we have $\operatorname{supp}\psi_k \subset [a_k,b_k]$ except ψ_0 . According to (2.14), $\psi_0(x,t)$ will grow exponentially, overtaking all of $\psi_k(x,t)$ and diverging.

Remark 2.6. As we have seen in Lemma 2.2 condition $\int_x^{x+L} u_0(y) dy = 0$ is enough for periodic initial data. Then we can always find a partition consisting of self-contained intervals by taking the endpoints with infimum points. In the rest of the article we always consider periodic initial data with the mean zero condition.

3. Behavior of numerical solutions

In this section we consider the behavior of numerical schemes for the initial-boundary value problem of (1.5) with initial data

$$u(x,0) = -\sin(\pi x), \qquad -1 < x < 1, \tag{3.1}$$

and the periodic boundary condition. Since the initial data are periodic with zero mass, the solution should remain periodic with zero mass for all time. In particular, the exact steady state solution to this problem, as discussed in the previous section, is

$$U_{\text{steady state}}(x) = x - \text{sign}(x), \qquad -1 < x < 1.$$
 (3.2)

Our interest is to find a numerical method for such a problem that is able to capture the steady state roll wave solution. The linear instability of the problem poses tremendous numerical challenges, and very careful numerical computations are needed to obtain the correct solutions. As will be see later, in most of the cases, numerical schemes do not converge to the steady state due to the numerical instability that is induced by the physical instability. Even if they converge, they may converge to the wrong steady state solution.

We first introduce the notations. Consider a uniform mesh with mesh points $x_{j+1/2}$ and width h, and a uniform time step k, where $j \in \mathbb{Z}, n \in \mathbb{Z}^+$. Let $U_j(t)$ be the cell average approximation of u over the domain $[x_{j-1/2}, x_{j+1/2}]$, namely,

$$U_j(t) = \frac{1}{h} \int_{x_{j-1/2}}^{x_{j+1/2}} u(x,t) dx,$$
(3.3)

and $U_j^n = U_j(t^n)$. We also need $U_{j+1/2}(t)$, approximation of u at $x_{j+1/2}$:

$$U_{j+1/2}^n \sim u(x_{j+1/2}, t), \quad t^n \le t < t^{n+1}.$$
 (3.4)

Here U_j and U_j^n will be called the cell average of u, while $U_{j+1/2}$ and $U_{j+1/2}^n$ will be called interface value of u. To derive a numerical discretization for (1.5), as in standard finite volume method, we integrating (1.5) over the domain $[x_{j-1/2}, x_{j+1/2}]$ to get

$$\partial_t U_j + \frac{1}{h} ((U_{j+1/2})^2 - (U_{j-1/2})^2) = U_j. \tag{3.5}$$

When the time discretization is the forward Euler method, (3.5) is the Godunov method for the homogeneous part, while the right hand side is the cell average of u. The fully discrete scheme takes the following form:

$$U_j^{n+1} - U_j^n + \frac{k}{h} ((U_{j+1/2}^n)^2 - (U_{j-1/2}^n)^2) = kU_j^n.$$
(3.6)

This scheme will be called the *cell-average method*, since it uses the cell-average value of u in the source term. Once the initial cell-average of u is given, (3.6) can be used after specifying the numerical flux $U_{j+1/2}$. Different definition of the flux gives different shock capturing method, and we leave it unspecified until later on in this section.

470 S. JIN AND Y.J. KIM

Let us consider the evolution of the total mass

$$M^n = \Sigma_j U_j^n \tag{3.7}$$

in the above method. Summing over j on (3.6) one gets

$$M^{n} = (1+k)M^{n-1} = (1+k)^{n}M^{0}.$$
(3.8)

So if the initial total mass M^0 of the numerical solution is not identically zero, the mass M^n will grow algebraically and eventually becomes unbounded. Thus the success of a method for this problem depends on whether the method *preserves the zero mass*.

We also consider another scheme, introduced in [13] for numerical capturing of the steady state of hyperbolic systems with geometrical source terms:

$$U_j^{n+1} - U_j^n + \frac{k}{h}((U_{j+1/2}^n)^2 - (U_{j-1/2}^n)^2) = \frac{k}{2}(U_{j+1/2}^n + U_{j-1/2}^n).$$
(3.9)

This scheme has the property of steady state preserving, namely, when the solution tends to steady state, (3.9) becomes

$$\frac{1}{h}((U_{j+1/2}^n)^2 - (U_{j-1/2}^n)^2) = \frac{1}{2}(U_{j+1/2}^n + U_{j-1/2}^n),$$

or

$$U_{j+1/2}^n - U_{j-1/2}^n = 0. (3.10)$$

(3.10) is clearly a second order approximation of the steady state solution (3.2) at cell $[x_{j-1/2}, x_{j+1/2}]$. This method will be referred to as the *interface method*.

In order to define the numerical flux, we use the Roe method [18]:

$$U_{j+1/2}^{n} = \begin{cases} U_{j}^{n} &, & \text{if } U_{j}^{n} + U_{j+1}^{n} \ge 0, \\ U_{j+1}^{n} &, & \text{if } U_{j}^{n} + U_{j+1}^{n} < 0, \end{cases}$$
(3.11a)

with the entropy fix

$$U_{j+1/2}^n = 0$$
, if $U_j^n < 0$, $U_{j+1}^n > 0$. (3.11b)

We solve (1.5) with initial data being the cell average of (3.1):

$$u_i^0 = \frac{1}{\pi \Delta x} \left[\cos(x_i + \frac{\Delta x}{2}) - \cos(x_i - \frac{\Delta x}{2}) \right]$$
 (3.12)

and use one ghost cell on each side of the domain. Periodic condition with period 2 is used to define the cell-average values at the ghost cells. All computations are made using *double precision arithmetic*. We conduct the following comparisons between the two schemes introduced above using different (even or odd) number of cells.

In the first experiment, we use odd number (51) of cells (so in the middle cell of the domain $u^0 = 0$) and $\Delta t = 0.02$. In this set up, it is rather easy to shown that the interface method (3.9) along with the Godunov flux (3.11) preserves the zero mass exactly, namely,

$$M^n = M^{n-1} = M^0. (3.13)$$

TABLE 3.1. Comparison of the interface method with the cell-average method using 51 cells.

Interface method			Cell-average method	
Time t	T Mass $M(t^n)$	L^1 error	$Mass M(t^n)$	L^1 error
0	-2.242130D-16	7.663521D-01	-2.242130D-16	7.663521D-01
5	-4.226220D-16	7.613592D-02	-4.876090D-14	2.111572D-01
10	-4.720616D-16	6.712015D-02	-6.869538D-12	2.129453D-01
15	-5.323432D-16	6.647604D-02	-9.704359D-10	2.129582D-01
20	-4.130810D-16	6.642417D-02	-1.370912D-07	2.129584D-01
25	-4.237062D-16	6.641996D-02	-1.936657D-05	2.129777D-01
30	-4.026726D-16	6.641962D-02	-2.735871D-03	2.156942D-01
35	-4.440892D-16	6.641959D-02	-0.386490D+00	5.643326D-01
36	-4.421376D-16	6.641959D-02	-1.040272D+00	1.154394D+00
37	-4.174178D-16	6.641959D-02	-2.799985D+00	2.827958D+00
37.4	-4.297777D-16	6.641959D-02	-4.160631D+00	6.270405D+00
37.44	-4.371503D-16	6.641959D-02	-4.328721D+00	3.454927D + 03
37.50	-4.169841D-16	6.641959D-02	-8.579398D+12	1.428020D + 30

However, due to the finite machine precision, the mass M^n calculated from (3.12) is not exactly zero. Rather it is a none zero number in the order of machine precision. This is not a problem for the interface method since for this method (3.13) holds, thus the mass M^n will remain at the level of machine precision. As a result, the solution of the interface method converges (as $t \to \infty$) to the correct steady state. However, for the the cell-average method, (3.8) indicates that the initial mass M^0 , which is a non-zero number in the order of machine precision, will be amplified by a factor of 1 + k after each time step. Consequently, the solution of the cell-average method eventually becomes unbounded and can never converge to the steady state solution.

In Table 3.1 the numerical behaviors of the two methods are compared at various times. The interface method preserves the zero mass up to machine precision, and converges to the steady state solution with an error in L^1 norm in the order of the mesh size. The cell-average method, however, fails to preserve the zero mass in the long time, due to the amplification and accumulation of the round-off errors. The numerical solution eventually becomes unbounded. Although for 5 < t < 30 the solution stays fairly close to the exact steady state, but this error eventually increases and the solution never converges to a steady state. Figure 3.1 shows the differences between the exact steady state and the two methods at t = 20. The solution of the interface method is a converged steady state, while that of the cell-average method is only a transient solution and are much less accurate than the interface method at this time.

We then study the convergence of the interface method by reducing the mesh size. We use 51, 101, 201, 401, 801 and 1601 cells respectively, and the corresponding time steps are $\Delta t = 0.02 * 50/(\text{number of cells} - 1)$. In all these meshes $u^0 = 0$ at the middle cell of the domain. The results are displayed in Table 3.2, which shows the first order convergence in L^1 norm of the method as the mesh size goes to zero.

We now compare the results using 50 cells, while other parameters in the numerical set up remain the same. In this case, u=0 is no longer in any cell, and the source operator of the interface method breaks the symmetry at both sides of x=0, yields an increase in the mass. Thus the zero mass condition cannot be preserved at the machine precision level after just one time step, and very soon (around t=5.86) the numerical solution becomes unbounded. The behavior of the cell-average method is pretty much the same as in the previous case, the amplification and accumulation of the round-off errors eventually destroy the zero mass condition. The results of both method are documented in Table 3.3. The result of the cell-average method at t=5.5, as which time the error between the numerical solution and the exact steady state is minimum, is shown in Figure 3.2.

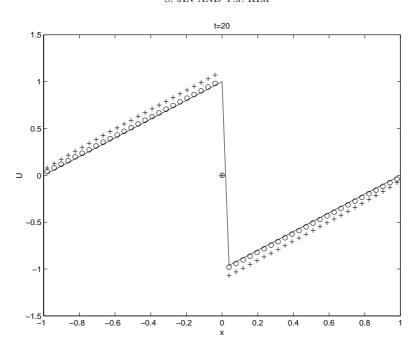


FIGURE 3.1. Exact steady state solution (solid line) versus the numerical solutions by the interface method (*) and by the cell-average method (+) at time t = 20. There are 51 cells in the domain [-1, 1], so in the middle cell $u^0 = 0$.

Table 3.2. Convergence study of the interface method.

Number of cells	L^1 error		
51	6.641959D-02		
101	2.148659D-02		
201	1.173174D-02		
401	1.500430D-02		
801	2.571934D-03		
1 601	1.305071D-03		

One should be informed that this is not a converged steady state solution. The figure only shows how close the numerical solution can be to the exact steady state.

In summary, we are able to capture the steady state roll wave solution with the interface method (3.8) along with the Godunov flux (3.10), under the condition that there are odd number of spatial cells over the period of the solution (so u=0 is a cell in the middle of the domain). One important criterion is that the method preserves the zero mass. This method fails when there are even number of cells in the period, so do the cell-average method for any number of cells, since they do not preserve the zero mass.

Table 3.3. Comparison of the interface method with the cell-average method using 50 cells.

Interface method			Cell-average method	
Time t	Mass $M(t^n)$	L^1 error	Mass $M(t^n)$	L^1 error
0	-1.444157D-16	7.747934D-01	-1.444157D-16	7.747934D-01
0.02	5.01993D-05	7.693004D-01	-9.454242D-17	7.716586D-01
1	4.219994D-02	3.758393D-01	-2.675810D-16	4.157055D-01
2	9.463334D-02	2.196706D-01	-9.376180D-16	2.288907D-01
3	2.633120D-01	3.267345D-01	-2.224349D-15	1.917004D-01
4	7.157849D-01	7.505731D-01	-5.689893D-15	1.887776D-01
5	1.936042D+00	1.936042D+00	-1.554485D-14	1.905798D-01
5.5	3.176304D+00	3.176304D+00	-2.564875D-14	1.912405D-01
5.86	-8.541924D+00	4.467215D+14	-3.669980D-14	1.924447D-01
10			-2.287877D-12	1.924447D-01
20			-4.747947D-08	1.924588D-01
30			-9.951297D-04	1.914637D-01
39			-7.711095D+00	7.711095D+00
39.14			-8.863705D+00	6.004777D+06

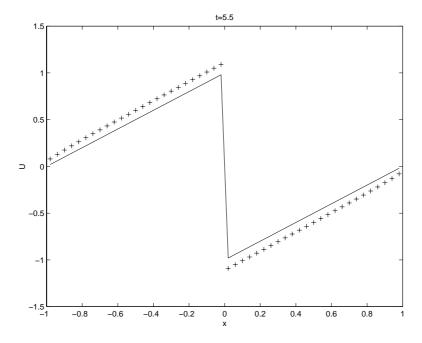


FIGURE 3.2. Exact steady state solution (solid line) versus the numerical solution of the cell-average method (+) at time t = 5.5. There are 50 cells over the domain [-1, 1].

204 cells

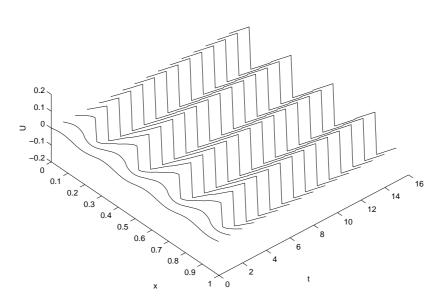


FIGURE 4.1. Numerical solution of (1.5) by the interface method with Godunov flux. The initial condition is given by (4.1). There are 204 cells over [0,1].

4. More numerical experiments

We now solve the problems studied in [16]. This is again a periodic problem over the domain [0,1], with the initial condition

$$u(x,0) = 0.01\sin(8\pi x). \tag{4.1}$$

The initial condition has four sine wave lengths over the domain [0,1]. According to the analysis in Section 2, the exact solution should have four roll waves over [0,1] with slope one, peak value 0.125 with jumps at x = 1/8, 3/8, 5/8 and 7/8. Therefore, in order to obtain the correct steady state, we need odd number of cells in *each* of the wave length. Namely, the zero mass condition should be preserved at *every* period of the sine wave. In this section we always use the interface method with the Godunov flux.

In the first run, we use 204 cells over [0,1], thus each period of the sine wave is resolved by 51 cells. We take $\Delta t = 0.005$ and display the numerical solution between $0 \le t \le 30$ in Figure 4.1. One can see that in this case the exact solution has been captured and the steady state is already reached around t = 5.

Next we use 202 cells. In this case there are odd number of cells in [0,0.5] and [0.5,1] respectively, however, each sine wave length has *even* number of cells. The numerical method we use can preserve the zero mass over [0,0.5] and [0.5,1], but not on every sine period. Thus the analysis and numerical experiment in the previous sections indicate that the numerical steady state should be two roll waves. This is indeed the case and the numerical results are shown in Figure 4.2. The steady state given in this run is clearly false.

Next we use 201 cells. In this case there are no odd number of cells on any of the period of the sine wave except in the whole domain [0, 1]. Thus the only preserved zero mass is the total mass, and one should expect only one roll wave at the steady state. This is confirmed numerically. See Figure 4.3. In this case, one again obtains a false steady state solution.

If 200 cells are used, then neither the submass nor the total mass will be preserved so eventually the numerical solution becomes unbounded and will never reach a steady state, just as studied in Section 3. We have observed this in our numerical experiment but will not report the result here.

202 cells

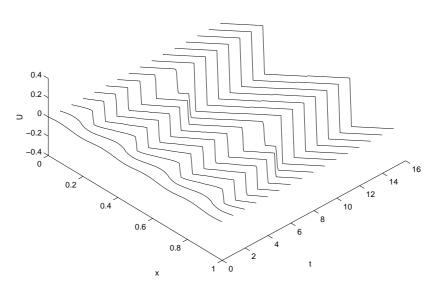


FIGURE 4.2. Numerical solution of (1.5) by the interface method with Godunov flux. The initial condition is given by (4.1). There are 202 cells over [0,1].

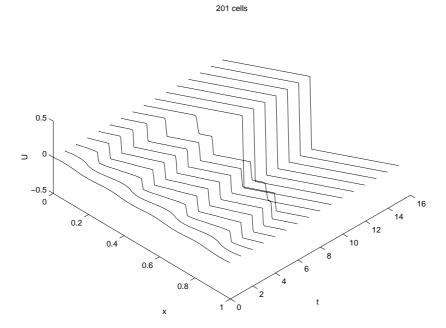


FIGURE 4.3. Numerical solution of (1.5) by the interface method with Godunov flux. The initial condition is given by (4.1). There are 201 cells over [0,1].

201 cells

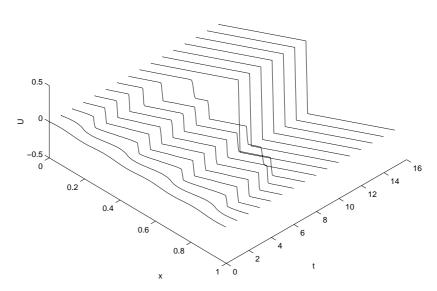


FIGURE 4.4. Numerical solution of (1.5) by the interface method with Godunov flux. The initial condition is given by (4.2). There are 201 cells over [0,1].

The last experiment we conduct is the so-called *biharmonic perturbation*. We perturb the initial data with a subharmonic disturbance

$$u(x,0) = 0.01\sin(8\pi x) + 0.001\sin(2\pi x). \tag{4.2}$$

Since the only zero mass is the total mass, we need to use odd number of cells over the domain. Figure 4.4 shows the numerical solutions with 201 cells and $\Delta t = 0.005$. One can see that the small perturbation basically overtakes the original profile and the steady state solution is the same as if one starts the initial profile with the perturbation. This can be justified using the analysis in Section 2.

5. Conclusions

In this paper, we studied the analytic and numerical issues associated with a model roll wave equation, which arises as the weakly nonlinear asymptotic approximation of the shallow water equations when the Froude number is bigger than two. The main difficulty for numerical approximation to this problem is its linear instability. The numerical round-off error, due to the finite precision arithmetic, can easily build up and eventually destroys the long time behavior of the roll wave solution. We utilize a numerical method introduced in [13], and show that it preserves the zero mass condition if odd number of cells are used over each period of the roll wave. This important property guarantees that the steady state roll wave will be captured numerically. Several numerical examples are used to illustrate the numerical subtlety of the problem and the performance of the proposed method

Although the numerical study is on a simpler model problem, the study here sheds some lights on the river equations, which will be the subject of future research by the authors. Increasing difficulty may arise there when one wants to compute a moving periodic roll waves constructed by Dressler, since the preservation of a constant mass, when the roll wave is moving, is a highly challenging numerical task.

Appendix A

The weakly nonlinear approximation of the shallow water equations

For completeness, in this appendix, we derive the asymptotic approximation (1.4) from the shallow water equations (1.1). This asymptotic was performed in [21] for the dimensionless form of the shallow water equations. Below it is carried out for the dimensional form of the equations (1.1). Let

$$h = \overline{h} + \epsilon h_1(\xi, t), v = \overline{v} + \epsilon v_1(\xi, t), \tag{A.1}$$

where $\overline{h}, \overline{v} = \sqrt{sg\overline{h}/C_f}$ are the equilibrium constants, $\xi = (x - ct)/\epsilon$ with $c_{\pm} = \overline{v} \pm \sqrt{g\overline{h}}$ being the equilibrium characteristic speeds. Applying the ansatz (A.1) into (1.1), one obtains the equations for h_1 and v_1 as

$$(\overline{v} - c)\partial_{\xi}h_{1} + \overline{h}\partial_{\xi}v_{1} = -\varepsilon\partial_{t}h_{1} - \varepsilon\partial_{\xi}(h_{1}v_{1}),$$

$$(-c\overline{v} + (\overline{v})^{2} + g\overline{h})\partial_{\xi}h_{1} + \overline{h}(2\overline{v} - c)\partial_{\xi}v_{1} + \varepsilon(2\overline{v} - c)\partial_{\xi}(h_{1}v_{1}) + \varepsilon\overline{h}\partial_{\xi}(v_{1}^{2})$$

$$+ \varepsilon^{2}\partial_{\xi}(h_{1}v_{1}^{2}) + \frac{\varepsilon}{2}g\partial_{\xi}(h_{1}^{2}) + \varepsilon\overline{h}\partial_{t}h_{1} + \varepsilon\overline{h}\partial_{t}v_{1} + \varepsilon^{2}\partial_{t}(h_{1}v_{1}) = \varepsilon gsh_{1} - 2\varepsilon C_{f}\overline{v}v_{1} - \varepsilon^{2}C_{f}v_{1}^{2}.$$
(A.2)

The leading order approximation of (A.2) is

$$(\overline{v} - c)\partial_{\xi}h_{1} + \overline{h}\partial_{\xi}v_{1} = 0$$

$$(-c\overline{v} + (\overline{v})^{2} + g\overline{h})\partial_{\xi}h_{1} + \overline{h}(2\overline{v} - c)\partial_{\xi}v_{1} = 0.$$
(A.3)

Multiplying (A.3a) by $2\overline{v} - c$, and then subtracting it by (A.3b), one gets an equation for c:

$$c^2 - 2\overline{(v)}c + (\overline{v})^2 - q\overline{h} = 0 \tag{A.4}$$

which gives

$$c_{\pm} = \overline{v} \pm \sqrt{g\overline{h}} \,. \tag{A.5}$$

With this c, (A.3) gives

$$v_1 = \frac{c - \overline{v}}{\overline{h}} h_1. \tag{A.6}$$

Now, multiplying (A.2a) by $2\overline{v} - c$, and then subtracting (A.2b), after ignoring the $O(\varepsilon^2)$ terms and some manipulations, one gets

$$2(c-\overline{v})\partial_t h_1 + \left[\frac{(c-\overline{v})^2}{\overline{h}} + \frac{1}{2}g\right]\partial_\xi(h_1^2) = \left[gs - \frac{2C_f\overline{v}}{\overline{h}}(c-\overline{v})\right]h_1. \tag{A.7}$$

For $c = c_+$, this gives

$$\partial_t h_1 + \frac{3}{2} \sqrt{\frac{g}{\overline{h}}} h_1 \partial_x h_1 = \sqrt{\frac{g}{\overline{h}}} s \left(1 - \sqrt{\frac{4C_f}{s}} \right) h_1. \tag{A.8}$$

Appendix B

AN OPTIMAL ESTIMATE OF INVISCID BURGERS

We consider piecewise continuous BV-solutions of the initial-value problem of the inviscid Burgers equation

$$v_s + \frac{1}{2}vv_x = 0, \quad v(x,0) = v_0(x), \qquad x \in \mathbb{R}, s > 0.$$
 (B.1)

The Oleinik estimate $\partial_x v(x,s) \leq 1/s$ implies that

$$v(x,s) - \frac{x}{s}$$
 is decreasing in x for any $s > 0$. (B.2)

It is well known that following quantities are invariant,

$$p = -\inf_{x} \int_{-\infty}^{x} v(y, s) dy, \quad q = \sup_{x} \int_{x}^{\infty} v(y, s) dy, \tag{B.3}$$

and that they play a key role in the evolution of the solution. Let $v(x_0, s) = v_0 \ge 0$. Then

$$0 \le \int_{x_0 - sv_0}^{x_0} v(y, s) - (\frac{y - x_0}{s} + v_0) dy = \int_{x_0 - sv_0}^{x_0} v(y, s) dy - \frac{s}{2} v_0^2.$$

So v_0 is bounded by

$$v_0 \le \sqrt{\frac{2}{s} \int_{x_0 - sv_0}^{x_0} v(y, s) dy}.$$
 (B.4)

Note that for any $a, b \in \mathbb{R}$ we get

$$\int_a^b v(y,s) dy = \int_{-\infty}^b v dy - \int_{-\infty}^a v dy \le \sup_{\eta} \int_{-\infty}^{\eta} v dy + p.$$

Since the supremum of the right hand side is decreasing in time (maximum principle), (B.4) gives a uniform upper bound. Similar arguments hold for $v_0 \le 0$ and we get a uniform estimate,

$$|v(x,s)| \le \sqrt{\frac{2}{s} \left(\sup_{\eta} \int_{-\infty}^{\eta} v(y,0) dy + p\right)}$$
.

We apply the method of generalized characteristics. For a detailed theory we refer [5, Ch 11]. Let $x = \xi(s), 0 < s < \infty$ be a generalized characteristic. We use the notation $v_{\pm}(s) = v(\xi(s) \pm, s)$ for simplification. A characteristic $\xi(\cdot)$ on [a,b] is called genuine if $v_{-}(s) = v_{+}(s)$ for almost all $s \in [a,b]$ or shock if $v_{-}(s) > v_{+}(s)$ for all $s \in (a,b)$. A genuine characteristic $\xi(s), 0 \le s \le b$ is given by

$$\xi(s) = \xi(0) + v(0)s, \quad 0 \le s \le b$$
 (B.5)

and a shock characteristic $\xi(s)$ on [a,b), a>0 is given by

$$\xi(s) = \xi(a) + \int_{a}^{t} \frac{1}{2} (v_{+}(s) + v_{-}(s)) ds.$$
(B.6)

Theorem B.1. Suppose that $v_0(x)$ has a compact support supp $(v_0) \subset [A, B]$. Then the solution v(x, s) of (B.1) satisfies

$$v(x,s) = 0,$$
 $x < A - \sqrt{2ps}, \ x > B + \sqrt{2qs},$ (B.7)

$$A \le x - sv(x, s) \le B, \qquad B - \sqrt{2ps} \le x \le A + \sqrt{2qs}.$$
 (B.8)

Proof. Let $x = \xi_2(s), 0 \le s \le \infty$ be the generalized characteristic with $\xi_2(0) = B$. Thus $v(\xi_2(s)+,s) = 0$. Let $v_2(s) = v(\xi_2(s)-,s)$. Then, since v(x,s) = 0 for all $x > \xi_2$,

$$\int_{\xi_2 - \nu_2}^{\xi_2} v(x, s) \mathrm{d}x = \int_{\xi_2 - \nu_2}^{\infty} v(x, s) \mathrm{d}x \le q,$$

and hence (B.4) implies that

$$v(\xi(s)-,s) \le \sqrt{\frac{2q}{s}}$$

Now $\xi_2(s)$ is easily estimated from (B.6),

$$\xi_2(s) \le B + \int_0^s \frac{1}{2} \sqrt{\frac{2q}{\tau}} d\tau = B + \sqrt{2qs}.$$

The characteristic $x = \xi_1(s)$ with $\xi_1(0) = A$ is similarly estimated and we get

$$A - \sqrt{2ps} \le \xi_1(s), \qquad \xi_2(s) \le B + \sqrt{2qs}.$$
 (B.9)

For a point ξ_0 satisfying $p = -\int_{-\infty}^{\xi_0} v(y,s) dy$, one can easily check that $v(\xi_0 \pm, s) = 0$ and

$$p = -\int_{-\infty}^{\xi_0} v(y, s) dy, \quad q = \int_{\xi_0}^{\infty} v(y, s) dy.$$

Since $v(y,s) - (y-\xi_0)/s$ is decreasing in y, (B.2), with a zero point $y = \xi_0 \le \xi_2$,

$$\int_{\xi_0}^{\xi_2} v(y,s) - \frac{y - \xi_0}{s} dy = q - (\xi_2 - \xi_0)^2 / 2s \le 0.$$

So $\xi_2 - \xi_0 \ge \sqrt{2qs}$. Similarly $\xi_0 - \xi_1 \ge \sqrt{2ps}$ and, hence,

$$\sqrt{2ps} + \sqrt{2qs} \le \xi_2(s) - \xi_1(s).$$
 (B.10)

Using (B.9),(B.10) the estimate for the characteristics ξ_1, ξ_2 is completed by

$$A - \sqrt{2ps} \le \xi_1(s) \le B - \sqrt{2ps},$$

$$A + \sqrt{2qs} \le \xi_2(s) \le B + \sqrt{2qs}.$$
(B.11)

It is clear that v(x,s)=0 for $x\not\in [\xi_1(s),\xi_2(s)],$ (B.7). It is enough to consider continuity points (x,s) to show (B.8). From (B.11) the condition $B-\sqrt{2ps}\le x\le A+\sqrt{2qs}$ simply implies $\xi_1(s)\le x\le \xi_2(s)$. If u is continuous

480 S. JIN AND Y.J. KIM

at the point (x, s), there exists a genuine characteristic $\xi(\tau)$, $0 < \tau < s$ such that $\xi(s) = x$. From (B.5) we get

$$\xi(s) - sv(\xi(s), s) = \xi(0).$$

Since
$$A = \xi_1(0) \le \xi(0) \le \xi_2(0) = B$$
, we get (B.8).

If the initial data is given by a delta function, $v_0(x) = q\delta(x)$, then all the corresponding constants A, B, p are zero. In the case estimates of the theorem, (B.7-8), give the exact structure of the solution. In the sense the estimates are optimal.

References

- [1] A. Bernudez and M.E. Vazquez, Upwind methods for hyperbolic conservation laws with source terms. *Comput. Fluids* **23** (1994) 1049–1071.
- [2] R. Botchorishvili, B. Perthame and A. Vasseur, Equilibrium schemes for scalar conservation laws with stiff sources. *Math. Comp.* (to appear).
- [3] A. Chinnayya and A.Y. Le Roux, A new general Riemann solver for the shallow-water equations with friction and topography. Preprint (1999).
- [4] V. Cornish, Ocean waves and kindred geophysical phenomena. Cambridge University Press, London (1934).
- [5] C.M. Dafermos, Hyperbolic conservation laws in continuum physics. Grundlehren der Mathematischen Wissenschaften 325, Springer-Verlag, Berlin (2000) xvi+443 pp.
- [6] R.F. Dressler, Mathematical solution of the problem of roll-waves in inclined open channels. Comm. Pure Appl. Math. 2 (1949)
- [7] T. Gallouët, J.-M. Hérard and N. Seguin, Some approximate Godunov schemes to compute shallow-water equations with topography. AIAA-2001 (to appear).
- [8] J. Goodman, Stability of the Kuramoto-Sivashinsky and related systems. Comm. Pure Appl. Math. 47 (1994) 293–306.
- [9] L. Gosse, A well-balanced flux-vector splitting scheme desinged for hyperbolic systems of conservation laws with source terms. Comp. Math. Appl. 39 (2000) 135–159.
- [10] J.M. Greenberg and A.-Y. Le Roux, A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. SIAM J. Numer. Anal. 33 (1996) 1–16.
- [11] J.K. Hunter, Asymptotic equations for nonlinear hyperbolic waves, in *Surveys in Appl. Math.* Vol. 2, J.B. Keller, G. Papanicolaou, D.W. McLaughlin, Eds. (1993).
- [12] H. Jeffreys, The flow of water in an inclined channel of rectangular section. Phil. Mag. 49 (1925) 793-807.
- [13] S. Jin, A steady-state capturing method for hyperbolic systems with source terms. ESAIM: M2AN (to appear).
- [14] S. Jin and M. Katsoulakis, Hyperbolic systems with supercharacteristic relaxations and roll waves. SIAM J. Appl. Math. 61 (2000) 271–292 (electronic).
- [15] Y.J. Kim and A.E. Tzavaras, Diffusive N-waves and metastability in Burgers equation. Preprint.
- [16] C. Kranenburg, On the evolution of roll waves. J. Fluid Mech. 245 (1992) 249–261.
- [17] P.D. Lax, Hyperbolic systems of conservation laws and the mathematical theory of shock waves. CBMS-NSF Regional Conference Series Appl. Math. 11, Philadelphia (1973).
- [18] R. LeVeque, Numerical methods for conservation laws. Lect. Math., ETH Zurich, Birkhauser (1992).
- [19] R.J. LeVeque, Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. J. Comp. Phys. 146 (1998) 346–365.
- [20] T.P. Liu, Nonlinear stability of shock waves for viscous conservation laws. Memoirs of the AMS 56 (1985).
- [21] A.N. Lyberopoulos, Asymptotic oscillations of solutions of scalar conservation laws with convexity under the action of a linear excitation. Quart. Appl. Math. XLVIII (1990) 755–765.
- [22] D.J. Needham and J.H. Merkin, On roll waves down an open inclined channel. Proc. Roy. Soc. Lond. A 394 (1984) 259–278.
- [23] O.B. Novik, Model description of roll-waves. J. Appl. Math. Mech. 35 (1971) 938–951.
- [24] P.L. Roe, Upwind differenced schemes for hyperbolic conservation laws with source terms. Lect. Notes Math. 1270, Springer, New York (1986) 41–51.
- [25] J.J. Stoker, Water Waves. John Wiley and Sons, New York (1958).
- [26] J. Whitham, Linear and nonlinear waves. Wiley, New York (1974).