INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2006; **65**:834–862
Published online 6 October 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme.1471

The generalized Riemann problem method for the shallow water equations with bottom topography

Jiequan Li[1,*,†] and Guoxian Chen[1,2]

[1]*Department of Mathematics, Capital Normal University, Beijing, 100037, People's Republic of China*
[2]*LMAM, School of Mathematical Sciences, Peking University, Beijing, 100871, People's Republic of China*

SUMMARY

This paper extends the generalized Riemann problem method (GRP) to the system of shallow water equations with bottom topography. The main contribution is that the generalized Riemann problem method (*J. Comput. Phys.* 1984; **55**(1):1–32) is used to evaluate the midpoint values of solutions at each cell interface so that the bottom topography effect is included in numerical fluxes, and at the same step the source term is discretized with an interface method in which only mid-point values are plugged in. This scheme is well balanced between the flux gradient and bottom topography when incorporating the surface gradient method (SGM) (*J. Comput. Phys.* 2001; **168**(1):1–25) into data reconstruction step, and it is also suitable for both steady and unsteady flow simulations. We illustrate the accuracy of this scheme by several 1-D and 2-D numerical experiments. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: the shallow water equations; the generalized Riemann problem method; the well-balanced property; the bottom topography; the surface gradient method; characteristic co-ordinates

1. INTRODUCTION

The system of shallow water equations (sometimes referred to as the Saint-Venant system) forms the basis of many mathematical models such as the tidal flows in estuary, lakes and coastal water regions, the flood routing in nature and man-made streams, hydraulic jump,

*Correspondence to: Jiequan Li, Department of Mathematics, Capital Normal University, Beijing, 100037, People's Republic of China.
†E-mail: jiequan@mail.cnu.edu.cn

Contract/grant sponsor: Zheng Ge Ru foundation
Contract/grant sponsor: Alexander von Humboldt Foundation
Contract/grant sponsor: NNSF; contract/grant number: 10301022
Contract/grant sponsor: Natural Science Foundation
Contract/grant sponsor: Fok Ying Tong Education Foundation
Contract/grant sponsor: Beijing Educational Commission; contract/grant number: KZ200510028018

Received 20 May 2005
Accepted 22 July 2005

open-channel flow and so on [1, 2]. In the case of two spatial dimensions it takes the form

$$\begin{aligned} h_t + (hu)_x + (hv)_y &= 0 \\ (hu)_t + (hu^2 + \tfrac{1}{2}gh^2)_x + (huv)_y &= -ghB_x \\ (hv)_t + (huv)_x + (hv^2 + \tfrac{1}{2}gh^2)_y &= -ghB_y \end{aligned} \tag{1}$$

where at a point (x, y) at time t, $h(x, y, t)$ is the height of fluid surface above the (fixed) elevation $B(x, y)$ of the bottom. The fluid is contained between $B(x, y)$ and $h(x, y, t)+B(x, y)$. The vector $(u, v)(x, y, t)$ is the velocity of fluid, and g is the gravitational constant taken to be $9.8\,\mathrm{m/s^2}$, unless otherwise stated. The source term results from the dynamical effect on the motion of fluid exerted by the gravitational force and the variable river bottom.

In recent years the numerical treatment of (1) has attracted much attention, leading to a variety of schemes, see References [3–18] and references therein. Special attention has been focused on the treatment of the source term (bottom topography). Among them the simplest is a fractional step method [19], which, however, provides relatively poor resolution for quasi-steady and steady flows [9]. The upwind discretization of the source term can improve significantly the accuracy of a numerical method, but it tends to be more complicated [6]. In Reference [10] the solution at the intermediate time level is computed by an approximate evolution operator in the predictor step, and the cell interface approximation of the source term is used in the corrector operator. Alternative approach in Reference [4] is used to construct numerical fluxes by incorporating the source term effects. Since the source term affects the motion of fluid, the physical flux function should include this influence, as explained in References [12, 14]. Therefore an efficient numerical scheme, in addition to being robust, should comply with the following basic principles: (i) The scheme is well-balanced between the flux gradient and the source term; (ii) Numerical fluxes reflect the source term influence; and (iii) Numerical solutions converge to steady states quickly.

The present paper provides the construction of numerical fluxes and the discretization of source term in the framework of finite volume schemes of the Godunov-type based on the above principles. In this context, it is essential to evaluate the mid-point values of solutions on each cell interface and to discretize the source term. We adopt the generalized Riemann problem method (GRP) for the mid-point values, and then use these values in order to discretize the source term (using the trapezoidal rule in space and the mid-point rule in time). Thus the central ingredient proposed here is the computation of the mid-point values, which conforms with the spirit of (ii) above. As a further ingredient, the surface gradient method (SGM) [15] is incorporated into the data reconstruction step, which makes the scheme well-balanced for quasi-steady flows.

We recall that the generalized Riemann problem method was originally devised for gas dynamical flows [20] and extended extensively [18, 21, 22]. When it is used for the system of shallow water equations, the auxiliary trick of Lagrangian co-ordinates used in Reference [20] is not needed, thanks to the existence of Riemann invariants for system (1). The process of evaluation of the midpoint values is thus straightforward. The idea of the interface method was proposed in Reference [8]. Here we use it in the discretization of the source term in an explicit scheme, where use is made of the mid-point values already obtained in the step of the construction of numerical fluxes. The resulting scheme is simple and direct in this regard. Since the mid-point value is upwind evaluated, so is the discretization of the source term.

This scheme is described in detail in Section 2 and the evaluation of mid-point values is given in Section 3. The well-balanced property to state the preservation of stationary flows is shown in Section 4, and both one-dimensional and two-dimensional numerical experiments are provided to display the accuracy of the scheme in Section 5.

2. THE STATEMENT OF THE SCHEME

In this section we propose our scheme for (1) on a finite volume Cartesian grid. We rewrite (1) in the form

$$W_t + F(W)_x + G(W)_y = \Psi(x, y, W) \tag{2}$$

where $W = (h, hu, hv)^\top$, $F(W) = (hu, hu^2 + \frac{1}{2}gh^2, huv)^\top$, $G(W) = (hv, huv, hv^2 + \frac{1}{2}gh^2)^\top$ and $\Psi(x, y, W) = (0, -gB_x h, -gB_y h)^\top$. Denote by W_{ij} the cell average of W over cell Ω_{ij}. With the Strang splitting method, we split this system into two subsystems in the x- and y-directions,

$$W_t + F(W)_x = \Psi_1(x, y, W) \tag{3}$$

and

$$W_t + G(W)_y = \Psi_2(x, y, W) \tag{4}$$

with $\Psi_1(x, y, t) = (0, -ghB_x(x, y), 0)^\top$ and $\Psi_2(x, y, t) = (0, 0, -ghB_y(x, y))^\top$. For further details about the Strang splitting, we refer to Reference [21, Chapter 7.3]. Then our scheme proceeds by solving the generalized Riemann problem normal to each interface between grid cells for (3) and (4). It clearly suffices to consider (3), which is now written in a one-dimensional form (and for simplicity, B is taken as $B(x)$)

$$U_t + F(U)_x = \Phi(x, U) \tag{5}$$

where $U = (h, hu, hv)^\top$, $F(U) = (hu, hu^2 + \frac{1}{2}gh^2, huv)^\top$, and $\Phi(x, U) = (0, -ghB'(x), 0)^\top$. Note that the first two equations of (5) form the basic system of one-dimensional shallow water equations decoupled from the third, which is an advection equation for v at propagation speed u. System (5) is called the augmented shallow water equations in the present paper (as compared to the basic one-dimensional system).

As customary, we denote by $C_j = [x_{j-1/2}, x_{j+1/2}]$ ($\Delta x = x_{j+1/2} - x_{j-1/2}$) the computational cell numbered j, and by $\{t_n\}_{n=0}^{\infty}$ the sequence of discretized time levels ($\Delta t = t_{n+1} - t_n$). Then the equivalent integral form of (5) over the control volume $[x_{j-1/2}, x_{j+1/2}] \times [t_n, t_{n+1}]$ is

$$\int_{x_{j-1/2}}^{x_{j+1/2}} U(x, t_{n+1})\,\mathrm{d}x - \int_{x_{j-1/2}}^{x_{j+1/2}} U(x, t_n)\,\mathrm{d}x + \int_{t_n}^{t_{n+1}} F(U(x_{j+1/2}, t))\,\mathrm{d}t$$

$$- \int_{t_n}^{t_{n+1}} F(U(x_{j-1/2}, t))\,\mathrm{d}t = \int_{t_n}^{t_{n+1}} \int_{x_{j-1/2}}^{x_{j+1/2}} \Phi(x, U(x, t))\,\mathrm{d}x\,\mathrm{d}t \tag{6}$$

Then this scheme consists of two steps.

Step 1. The evolution of solutions: Given a piecewise linear distribution for U at $t = t_n$

$$U_n(x) = U_j^n + \sigma_j^n \cdot (x - x_j) \quad x \in (x_{j-1/2}, x_{j+1/2}) \tag{7}$$

system (5) is evolved by approximating (6). Denote the midpoint value of U on the cell interface $x = x_{j+1/2}$ by

$$U_{j+1/2}^{n+1/2} = U\left(x_{j+1/2}, t_n + \frac{\Delta t}{2}\right) \tag{8}$$

Then we approximate the flux term with the mid-point rule

$$\frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} F(U(x_{j+1/2}, t))\,\mathrm{d}t \cong F(U_{j+1/2}^{n+1/2}) \tag{9}$$

and discretize the source term with the trapezoidal rule in space and the midpoint rule in time

$$\int_{t_n}^{t_{n+1}}\int_{x_{j-1/2}}^{x_{j+1/2}} \Phi(x, U(x,t))\,\mathrm{d}x\,\mathrm{d}t \cong \frac{\Delta x \Delta t}{2}(\Phi(x_{j-1/2}, U_{j-1/2}^{n+1/2}) + \Phi(x_{j+1/2}, U_{j+1/2}^{n+1/2})) \tag{10}$$

These approximations are formally second-order accurate. Thus (6) is discretized as

$$\begin{aligned} U_j^{n+1} - U_j^n = &-\frac{\Delta t}{\Delta x}(F(U_{j+1/2}^{n+1/2}) - F(U_{j-1/2}^{n+1/2})) \\ &+\frac{\Delta t}{2}(\Phi(x_{j-1/2}, U_{j-1/2}^{n+1/2}) + \Phi(x_{j+1/2}, U_{j+1/2}^{n+1/2})) \end{aligned} \tag{11}$$

where U_j^n is the cell average of U over cell C_j at time $t = t_n$

$$U_j^n \cong \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} U(x, t_n)\,\mathrm{d}x \tag{12}$$

In our case the source term has a special form $\Phi(x, U) = (0, -ghB'(x), 0)^\top$, so the second component of (11) reads

$$-\int_{t_n}^{t_{n+1}}\int_{x_{j-1/2}}^{x_{j+1/2}} ghB'(x)\,\mathrm{d}x\,\mathrm{d}t = -\frac{g\Delta x\Delta t}{2}(B'(x_{j+1/2})h_{j+1/2}^{n+1/2} + B'(x_{j-1/2})h_{j-1/2}^{n+1/2}) \tag{13}$$

In practice, we replace this by

$$-\int_{t_n}^{t_{n+1}}\int_{x_{j-1/2}}^{x_{j+1/2}} ghB'(x)\,\mathrm{d}x\,\mathrm{d}t = -g\frac{B(x_{j+1/2}) - B(x_{j-1/2})}{2} \cdot (h_{j-1/2}^{n+1/2} + h_{j+1/2}^{n+1/2})\Delta t \tag{14}$$

with only very minor differences (within second-order accuracy).

Our discretization of the source term in (10) is an explicit variant of the interface method proposed in a semi-discrete scheme [8]. The advantage of this method here is that we only exploit the midpoint value of the solution at each cell interface, which is already obtained in the construction of the numerical fluxes in (9). No more values are needed in the process.

Note that the midpoint values, as given in Section 3, are upwind evaluated. The treatment of source term is thus upwind consistent with the numerical fluxes.

Step 2. The data reconstruction: Given the central values $\{U_j^n\}$, we need to construct the associated piecewise linear data (7), namely we need to find σ_j^n. The method we adopt is called the surface gradient method (SGM) proposed in Reference [15], which is essential to balance well between the flux gradient and the source term for quasi-steady flows. We include below some details of the method for the convenience of readers. We refer to Reference [15] for more details.

The SGM does not treat the primitive variables (h, u, v) directly, but instead deals with $(h + B, u, v)$, where $h + B$ is the total elevation of the fluid surface over the zero reference level. Let

$$\eta(x, t) = h(x, t) + B(x) \tag{15}$$

Then we derive the slope $\delta\eta_j^n$ for $\eta(x, t)$ from $\{\eta_j^n\}$, $\eta_j^n = h_j^n + B(x_j)$, with any monotone slope limiter operator such as the minmod Limiter and the van Leer Limiter [19]

$$\delta\eta_j^n = \frac{1}{\Delta x}\,\text{Limiter}(\eta_{j-1}^n, \eta_j^n, \eta_{j+1}^n) \tag{16}$$

Thus the slope δh_j^n of $h_n(x)$ in $[x_{j-1/2}, x_{j+1/2}]$ is recovered as

$$\delta h_j^n = \delta\eta_j^n - B'(x_j) \tag{17}$$

This is the first component of σ_j^n in (7). For the slopes of u and v, we just use the usual monotone algorithm, as for η in (16),

$$\begin{aligned} \delta u_j^n &= \frac{1}{\Delta x}\,\text{Limiter}(u_{j-1}^n, u_j^n, u_{j+1}^n) \\ \delta v_j^n &= \frac{1}{\Delta x}\,\text{Limiter}(v_{j-1}^n, v_j^n, v_{j+1}^n) \end{aligned} \tag{18}$$

Steps 1 and 2 are closely interrelated and are both indispensable. In Section 4 we will see the role that each step plays in the well-balanced property of this scheme.

Remark 2.1

In our computation, we take the bottom topography $B(x)$ as a piecewise linear profile within cell C_j by

$$B(x) = B(x_j) + B'(x_j)(x - x_j) \tag{19}$$

with

$$B(x_j) \cong \frac{B(x_{j+1/2}) + B(x_{j-1/2})}{2}, \quad B'(x_j) \cong \frac{B(x_{j+1/2}) - B(x_{j-1/2})}{x_{j+1/2} - x_{j-1/2}} \tag{20}$$

This simplification is consistent with (14) but does not reduce the accuracy.

3. MIDPOINT VALUES ON CELL INTERFACES

There are several methods to calculate midpoint values $U_{j+1/2}^{n+1/2}$ such as the MUSCL-Hancock finite-volume method adopted in Reference [15]. In this study we use the generalized Riemann problem method, which was originally devised for the gas dynamical system [20], in order to evaluate $U_{j+1/2}^{n+1/2}$ in an explicit way. An alternative asymptotic approach for the generalized Riemann problem can be found in Reference [23, 24].

Assume that the solution is smooth along the cell interface $x = x_{j+1/2}$ for $t_n < t < t_{n+1}$, i.e. instantaneously following the resolution of the singularity at $(x_{j+1/2}, t_n)$. Then the standard Taylor series expansion gives

$$U_{j+1/2}^{n+1/2} = U(x_{j+1/2}, t_n + 0) + \frac{\Delta t}{2} \cdot \frac{\partial U}{\partial t}(x_{j+1/2}, t_n + 0) + O((\Delta t^2)) \tag{21}$$

Denote

$$U_{j+1/2}^{n} := U(x_{j+1/2}, t_n + 0), \quad \left(\frac{\partial U}{\partial t}\right)_{j+1/2}^{n} := \frac{\partial U}{\partial t}(x_{j+1/2}, t_n + 0) \tag{22}$$

In fact, $U_{j+1/2}^{n} = R(0; U_{j+1/2,-}^{n}, U_{j+1/2,+}^{n})$, where $R(x/t; U_{j+1/2,-}^{n}, U_{j+1/2,+}^{n})$ is the Riemann solution of the following associated Riemann problem:

$$U_t + F(U)_x = 0, \quad x \in \mathbb{R}, \quad t \in [0, \Delta t)$$

$$U(x, t_n) = \begin{cases} U_{j+1/2,-}^{n} := U_j^n + \sigma_j^n \dfrac{\Delta x}{2}, & x < 0 \\ U_{j+1/2,+}^{n} := U_{j+1}^n - \sigma_{j+1}^n \dfrac{\Delta x}{2}, & x > 0 \end{cases} \tag{23}$$

While $(\frac{\partial U}{\partial t})_{j+1/2}^{n}$ is the instantaneous value of time derivative of U at $(x_{j+1/2}, t_n + 0)$, which is evaluated by solving the generalized Riemann problem for (5) subject to the piecewise linear initial data

$$U(x, t_n) = \begin{cases} U_j^n + \sigma_j^n (x - x_j), & x < x_{j+1/2} \\ U_{j+1}^n + \sigma_j^n (x - x_{j+1}), & x > x_{j+1/2} \end{cases} \tag{24}$$

Once the values $U_{j+1/2}^{n}$ and $(\frac{\partial U}{\partial t})_{j+1/2}^{n}$ are defined, we can evaluate the midpoint value within second-order accuracy by ignoring $O(\Delta t^2)$ term in (21), and still denote from now on

$$U_{j+1/2}^{n+1/2} = U_{j+1/2}^{n} + \frac{\Delta t}{2} \left(\frac{\partial U}{\partial t}\right)_{j+1/2}^{n} \tag{25}$$

Remark 3.1

We see that the value $U_{j+1/2}^{n}$ is taken from the associated Riemann problem (23) for the homogeneous conservation laws and does not reflect the source term effect. This is why the

Riemann problem needs to be modified by taking the source term into account for the first-order Godunov-type scheme [11] that is proved to be quite complicated. However, the source term effect is reflected in the time derivative term $(\frac{\partial U}{\partial t})^n_{j+1/2}$. This can be seen heuristically from the following advection equation:

$$q_t + aq_x = k(x, q) \tag{26}$$

from which we have

$$q_t = -aq_x + k(x, q) \tag{27}$$

When this is plugged into (25), the midpoint value not only reflects the spatial variation of variables but includes the contribution of the source term as well. This demonstrates the necessity to introduce high resolution schemes for hyperbolic systems with source terms like (5).

Since there are many good Riemann solvers to get $U^n_{j+1/2}$, see Reference [19] and Appendix C in Reference [21], the remaining task is just to compute the value $(\frac{\partial U}{\partial t})^n_{j+1/2}$. As mentioned at the beginning of this section, the first two equations of (5) are decoupled from the third, an advection equation for v, which propagates with speed u. Therefore we first investigate the one-dimensional shallow water equations to derive the time derivatives of h, u at $(x_{j+1/2}, t_n + 0)$. Then we continue to find the time derivative of v. These are formulated to solve the generalized Riemann problems (5) and (24).

3.1. The time derivatives of h, u on the interface—The generalized Riemann problem for the one-dimensional shallow water equations

We consider the generalized Riemann problem for the one-dimensional shallow water equations with bottom topography

$$V_t + H(V)_x = \Gamma(x, V) \tag{28}$$

subject to the initial data

$$V(x, 0) = \begin{cases} V_- + \sigma_- x, & x < 0 \\ V_+ + \sigma_+ x, & x > 0 \end{cases} \tag{29}$$

where $V = (h, hu)^\top$, $H(V) = (hu, hu^2 + \frac{1}{2} gh^2)^\top$, $\Gamma(x, V) = (0, -ghB'(x))$. In (29), $V_\pm$ and $\sigma_\pm$ are all constant vectors. Note that, within second-order accuracy, not only h and hu, but also $u(x, 0)$ can be taken to be piecewise linear

$$u(x, 0) = \begin{cases} u_- + \delta u_- \cdot x, & x < 0 \\ u_+ + \delta u_+ \cdot x, & x > 0 \end{cases} \tag{30}$$

where $u_\pm$ and $\delta u_\pm$ are all constants. In the sequel, we use the notation δq to denote the slope of the quantity q.

We want to calculate

$$\left(\frac{\partial V}{\partial t}\right)_0 := \lim_{t \to 0+} \frac{\partial V}{\partial t}(0, t) \tag{31}$$

This computation depends on the local wave patterns associated with the Riemann problem (23) and characteristic co-ordinates play an important role. For this purpose, we write (28) for smooth flows as

$$\frac{\partial Q}{\partial t} + A \cdot \frac{\partial Q}{\partial x} = \Pi(x) \tag{32}$$

where Q, Π and A are

$$Q := (h, u)^{\top}, \quad \Pi(x) := (0, -gB'(x))^{\top}, \quad A(h, u) := \begin{pmatrix} u & h \\ g & u \end{pmatrix} \tag{33}$$

Denote $\psi(x) = -gB'(x)$. Then (32) can be diagonalized to be

$$\begin{aligned} \frac{\partial r}{\partial t} + \lambda \frac{\partial r}{\partial x} &= \psi \\ \frac{\partial s}{\partial t} + \mu \frac{\partial s}{\partial x} &= \psi \end{aligned} \tag{34}$$

where $\lambda = u - c$ and $\mu = u + c$ are the two eigenvalues of A, $r = u - 2c$ and $s = u + 2c$ are the two Riemann invariants, $c = \sqrt{gh}$. The quantity $Fr = u/c$ is the Froude number. The flow is subcritical if $Fr < 1$; and supercritical if $Fr > 1$.

The characteristic co-ordinates $\beta(x, t) = C_1$, $\alpha(x, t) = C_2$, C_1 and C_2 are arbitrary constants, are defined as two families of integral curves of the system,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \lambda(r(x, t), s(x, t)) \tag{35}$$

and

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \mu(r(x, t), s(x, t)) \tag{36}$$

In a region where $(x, t) \to (\alpha, \beta)$ is one-to-one, we can consider all variables (including x, t) as functions of (α, β). Then, in terms of the characteristic co-ordinates (α, β), (34) is written equivalently to be

$$\begin{aligned} \frac{\partial x}{\partial \alpha} &= \lambda \frac{\partial t}{\partial \alpha}, \quad \frac{\partial r}{\partial \alpha} = \psi \frac{\partial t}{\partial \alpha} \\ \frac{\partial x}{\partial \beta} &= \mu \frac{\partial t}{\partial \beta}, \quad \frac{\partial s}{\partial \beta} = \psi \frac{\partial t}{\partial \beta} \end{aligned} \tag{37}$$

It follows that $t = t(\alpha, \beta)$, as function of α and β, satisfies

$$(\lambda - \mu) \frac{\partial^2 t}{\partial \alpha \partial \beta} + \frac{\partial \lambda}{\partial \beta} \frac{\partial t}{\partial \alpha} - \frac{\partial \mu}{\partial \alpha} \frac{\partial t}{\partial \beta} = 0 \tag{38}$$

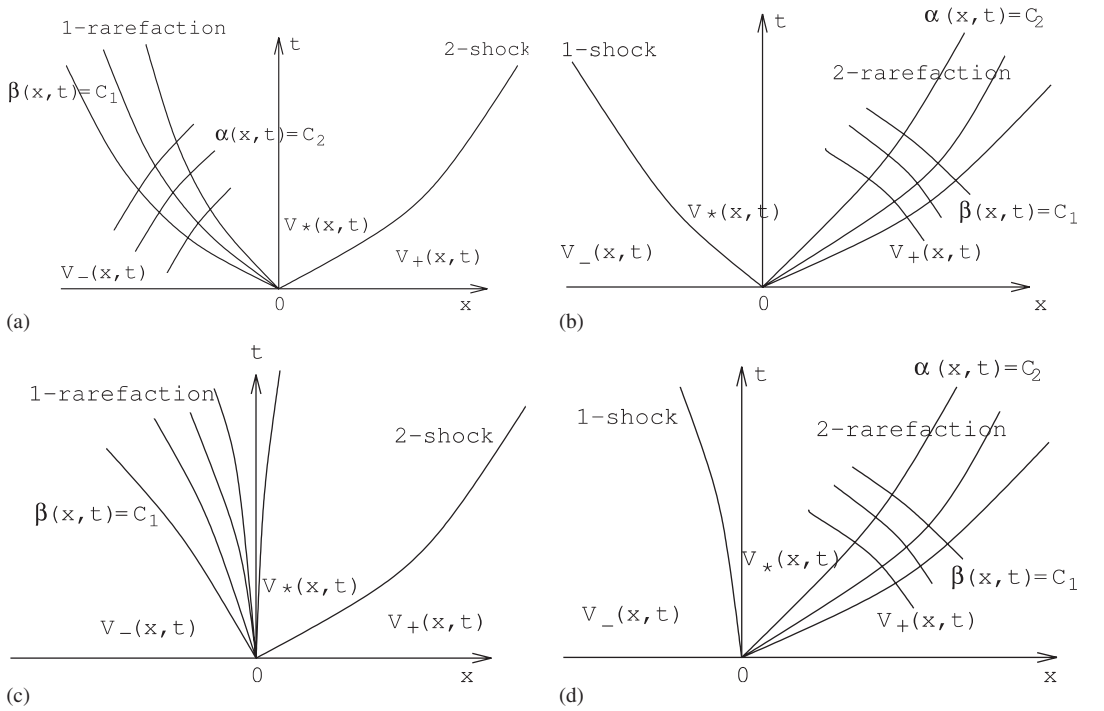

Figure 1. Typical wave patterns for one-dimensional shallow water equations: (a) the 1-rarefaction wave moves to the left and the 2-shock to the right; (b) the 2-shock moves to the left and the 1-rarefaction wave to the right; (c) the t-axis is inside the 1-rarefaction wave; and (d) the 1-shock is tangent to the t-axis.

and $s = s(\alpha, \beta)$ satisfies

$$\begin{aligned}\frac{\partial^2 s}{\partial\alpha\partial\beta} &= \frac{\partial^2 t}{\partial\alpha\partial\beta}\cdot\psi(x) + \frac{\partial t}{\partial\beta}\frac{\partial\psi}{\partial\alpha}\\ &= -\frac{\psi}{\lambda-\mu}\cdot\frac{\partial\lambda}{\partial\beta}\cdot\frac{\partial t}{\partial\alpha} + \frac{\partial t}{\partial\beta}\left(\frac{\psi}{\lambda-\mu}\frac{\partial\mu}{\partial\alpha} + \frac{\partial\psi}{\partial\alpha}\right)\end{aligned} \tag{39}$$

Now we consider the typical wave patterns as shown in Figure 1. When the t-axis is located inside the intermediate region between two waves, we denote by $V_*(x,t) = (h_*(x,t), u_*(x,t))$ the intermediate state. Then we have

Proposition 3.2

The instantaneous value $(\frac{\partial V}{\partial t})_0$, see (31), is obtained by solving the following system of linear algebraic equations:

$$\begin{aligned} a_-\left(\frac{\partial h}{\partial t}\right)_0 + b_-\left(\frac{\partial u}{\partial t}\right)_0 &= d_-\\ a_+\left(\frac{\partial h}{\partial t}\right)_0 + b_+\left(\frac{\partial u}{\partial t}\right)_0 &= d_+\end{aligned} \tag{40}$$

where a_+, b_+, and d_+ (resp. a_-, b_- and d_-) are constants depending only on the right initial data V_+, σ_+, (resp. the left initial data V_-, σ_-) and the associated Riemann solution $V_* = R(0; V_-, V_+)$.

Proof

The proof is divided into four cases. The first two deal with the wave patterns in Figure 1(a) and (b). The transcritical case (Figure 1(c)) is discussed as Case 3 and the stationary shock case (Figure 1(d)) is treated as Case 4.

Case 1. The resolution of centred rarefaction waves: Consider the standard centred rarefaction wave associated with the eigenvalue λ (recall that this is a solution for the associated Riemann problem so that $u_\pm$, $c_\pm$, u_* and c_* are all constants),

$$\begin{aligned} x/t &= u - c \\ u + 2c &= u_- + 2c_- := s_- \end{aligned} \tag{41}$$

The characteristic curves inside this rarefaction wave are chosen to be

$$\begin{aligned} &\beta = x/t, \quad \beta_- < \beta < \beta_*, \quad \beta_- = u_- - c_-, \quad \beta_* = u_* - c_* \\ &\alpha^{2/3} = s_- t^{2/3} - x t^{-1/3}, \quad \alpha > 0 \end{aligned} \tag{42}$$

That is

$$x = \beta t, \quad t = \frac{\alpha}{(s_- - \beta)^{3/2}} \tag{43}$$

so that (x, t) are given throughout the rarefaction wave by

$$t_{\text{ass}}(\alpha, \beta) = \frac{\alpha}{(s_- - \beta)^{3/2}}, \quad x_{\text{ass}}(\alpha, \beta) = \frac{\beta\alpha}{(s_- - \beta)^{3/2}} \tag{44}$$

The characteristic curves in the curved rarefaction wave R_λ associated with the generalized Riemann problem (28), (29) is a perturbation of (44) of second-order near the origin $(0, 0)$

$$x(\alpha, \beta) = x_{\text{ass}}(\alpha, \beta) + O(\alpha^2), \quad t(\alpha, \beta) = t_{\text{ass}}(\alpha, \beta) + O(\alpha^2) \tag{45}$$

as $\alpha \to 0$. It follows that

$$\frac{\partial t}{\partial \alpha}(0, \beta) = \frac{1}{(s_- - \beta)^{3/2}}, \quad \frac{\partial t}{\partial \beta}(0, \beta) = 0, \quad \beta_- \leqslant \beta \leqslant \beta_* \tag{46}$$

Then with (39), we can obtain $\frac{\partial s}{\partial \alpha}(0, \beta)$ by solving the following problem:

$$\begin{aligned} \frac{\partial}{\partial \beta}\left(\frac{\partial s}{\partial \alpha}(0, \beta)\right) &= \frac{-\psi(0)}{2(s_- - \beta)/3} \cdot \frac{1}{(s_- - \beta)^{3/2}} \\ \frac{\partial s}{\partial \alpha}(0, \beta_-) &= -\frac{1}{(3c_-)^{3/2}}(-\psi(0) + 2c_- \cdot \delta s_-) \end{aligned} \tag{47}$$

In (47), the initial data $\frac{\partial s}{\partial \alpha}(0, \beta_-)$ is obtained from (34)

$$\begin{aligned}\frac{\partial s}{\partial \alpha} &= \frac{\partial t}{\partial \alpha} \cdot \left(\frac{\partial s}{\partial t} + \lambda \frac{\partial s}{\partial x} \right) \\ &= \frac{\partial t}{\partial \alpha} \cdot \left(\frac{\partial s}{\partial t} + \mu \frac{\partial s}{\partial x} + (\lambda - \mu) \frac{\partial s}{\partial x} \right) \\ &= \frac{\partial t}{\partial \alpha} \cdot \left(\psi + (\lambda - \mu) \frac{\partial s}{\partial x} \right) \end{aligned} \tag{48}$$

Setting $(\alpha, \beta) = (0, \beta_-)$ (so that $\lambda = \lambda_- = u_- - c_-$, etc.), we get the initial data in (47). Consequently, we have

$$\frac{\partial s}{\partial \alpha}(0, \beta) = \frac{\partial s}{\partial \alpha}(0, \beta_-) - \psi(0)((s_- - \beta)^{-3/2} - (s_- - \beta_-)^{-3/2}) \tag{49}$$

Recall that

$$u(0, \beta) + 2c(0, \beta) = s_-, \quad \beta = u(0, \beta) - c(0, \beta) \tag{50}$$

Then we get

$$s_- - \beta = 3c(0, \beta), \quad s_- - \beta_- = 3c_- \tag{51}$$

Therefore, (49) gives

$$\frac{\partial s}{\partial \alpha}(0, \beta_*) = \frac{\partial s}{\partial \alpha}(0, \beta_-) - \psi(0)((3c_*)^{-3/2} - (3c_-)^{-3/2}) \tag{52}$$

Returning to the primitive variables $Q = (h, u)$, we calculate

$$\begin{aligned}\frac{\partial s}{\partial \alpha}(0, \beta) &= \frac{\partial t}{\partial \alpha}(0, \beta) \left[\frac{\partial s}{\partial t}(0, \beta) + \lambda \frac{\partial s}{\partial x}(0, \beta) \right] \\ &= \frac{\partial t}{\partial \alpha}(0, \beta) \left[\nabla_{h,u} s \cdot \frac{\partial Q}{\partial t}(0, \beta) + \lambda \nabla_{h,u} s \cdot \frac{\partial Q}{\partial x}(0, \beta) \right] \\ &= \frac{\partial t}{\partial \alpha}(0, \beta) \left[\nabla_{h,u} s \cdot \frac{\partial Q}{\partial t}(0, \beta) + \lambda \nabla_{h,u} s A^{-1} \cdot \left(\Pi(0) - \frac{\partial Q}{\partial t}(0, \beta) \right) \right] \\ &= \frac{\partial t}{\partial \alpha}(0, \beta) \left[\nabla_{h,u} s (A - \lambda I) A^{-1} \cdot \frac{\partial Q}{\partial t}(0, \beta) + \lambda \nabla_{h,u} s A^{-1} \cdot \Pi(0) \right] \\ &= \frac{\partial t}{\partial \alpha}(0, \beta) \left[(\mu - \lambda) \nabla_{h,u} s A^{-1} \cdot \frac{\partial Q}{\partial t}(0, \beta) + \lambda \nabla_{h,u} s A^{-1} \cdot \Pi(0) \right] \end{aligned} \tag{53}$$

where we use the fact that $\nabla_{h,u}s$ is the left eigenvector of A

$$A=\begin{pmatrix} u(0,\beta) & h(0,\beta) \\ g & u(0,\beta) \end{pmatrix} \tag{54}$$

associated with μ. Then we get

$$(\mu-\lambda)\nabla_{h,u}sA^{-1}\cdot\frac{\partial Q}{\partial t}(0,\beta)=\frac{\partial s}{\partial\alpha}(0,\beta)\left(\frac{\partial t}{\partial\alpha}\right)^{-1}(0,\beta)-\lambda\nabla_{h,u}sA^{-1}\cdot\Pi(0) \tag{55}$$

In particular, setting $\beta=\beta_*$, we arrive at

$$a_-\left(\frac{\partial h}{\partial t}\right)_0+b_-\left(\frac{\partial u}{\partial t}\right)_0=d_- \tag{56}$$

where

$$\begin{aligned}(a_-,b_-)&=(\mu_*-\lambda_*)\nabla_{h,u}s_*(A_*)^{-1}\\ d_-&=\left(\frac{\partial t}{\partial\alpha}(0,\beta_*)\right)^{-1}\frac{\partial s}{\partial\alpha}(0,\beta_*)-\lambda_*\nabla_{h,u}s_*(A_*)^{-1}\cdot\Pi(0)\end{aligned} \tag{57}$$

Case 2. The resolution of shocks: We first recall the basic shock relation for the system (28). Let $x=x(t)$ be the shock trajectory and denote by $\bar{Q}=(\bar{h},\bar{u})$ and $Q=(h,u)$ the limiting values ahead (resp. behind) the shock front. Then the Rankine–Hugoniot jump relation reads

$$\begin{aligned}x'(t)&=\frac{\bar{u}+u}{2}+\frac{\bar{h}+h}{2}\cdot\frac{\bar{u}-u}{\bar{h}-h}\\ (\bar{u}-u)^2&=\frac{g(\bar{h}+h)}{2\bar{h}h}(\bar{h}-h)^2\end{aligned} \tag{58}$$

which results in two possibilities,

$$\begin{cases} x'(t)=\dfrac{\bar{u}+u}{2}+\dfrac{\bar{h}+h}{2}\cdot\dfrac{\bar{u}-u}{\bar{h}-h}, \\ \dfrac{\bar{u}-u}{\bar{h}-h}=-\sqrt{\dfrac{g(\bar{h}+h)}{2\bar{h}h}}, \end{cases} \quad \text{or} \quad \begin{cases} x'(t)=\dfrac{\bar{u}+u}{2}+\dfrac{\bar{h}+h}{2}\cdot\dfrac{\bar{u}-u}{\bar{h}-h} \\ \dfrac{\bar{u}-u}{\bar{h}-h}=\sqrt{\dfrac{g(\bar{h}+h)}{2\bar{h}h}} \end{cases} \tag{59}$$

The former corresponds to the 1-shock associated with λ, the latter to the 2-shock associated with μ. Denote

$$G(h,u;\bar{h},\bar{u})=h\bar{h}(u-\bar{u})^2-\frac{g}{2}(h+\bar{h})(h-\bar{h})^2 \tag{60}$$

Then the Rankine–Hugoniot condition (58) implies

$$G(h,u;\bar{h},\bar{u})=0 \tag{61}$$

This is differentiated along the shock curve $x=x(t)$ to yield

$$\nabla_{h,u}G\left(\frac{\partial Q}{\partial t}+\gamma(t)\frac{\partial Q}{\partial x}\right)+\nabla_{\bar{h},\bar{u}}G\left(\frac{\partial \bar{Q}}{\partial t}+\gamma(t)\frac{\partial \bar{Q}}{\partial x}\right)=0 \tag{62}$$

where $\gamma(t)=x'(t)$ is the shock speed and the two terms in the left-hand side of (62) correspond to the differentiation on the two sides of the shock $x=x(t)$. Using (32), we substitute $\frac{\partial Q}{\partial x}$ in terms of $\frac{\partial Q}{\partial t}$ and $\frac{\partial \bar{Q}}{\partial t}$ in terms of $\frac{\partial \bar{Q}}{\partial x}$, as in Case 1, to get

$$\nabla_{h,u}G\left(\frac{\partial Q}{\partial t}+\gamma(t)A^{-1}\left(\Pi-\frac{\partial Q}{\partial t}\right)\right)+\nabla_{\bar{h},\bar{u}}G\left(\Pi-\bar{A}\frac{\partial \bar{Q}}{\partial x}+\gamma(t)\frac{\partial \bar{Q}}{\partial x}\right)=0 \tag{63}$$

Consider now the case of a 2-shock in Figure 1(a). Then in (63), $(\bar{h},\bar{u})\to(h_+,u_+)$ and $(h,u)\to(h_*,u_*)$ as $t\to 0+$. Note that $A(h_*,u_*)$ is not singular by the entropy condition (i.e. the flow is subcritical in the intermediate region). Thus we take the limit as $t\to 0+$ to obtain a_+, b_+ and d_+ in (40)

$$\begin{aligned}(a_+,b_+)&=\nabla_{h,u}G(h_*,u_*)\cdot(A(h_*,u_*)-\gamma_* I)A^{-1}(h_*,u_*)\\ d_+&=\nabla_{h_+,u_+}G\cdot(A(h_+,u_+)-\gamma_* I)\cdot(\delta h_+,\delta u_+)\\ &\quad -\gamma_*\nabla_{h,u}G(h_*,u_*)\cdot A^{-1}(h_*,u_*)\Pi(0)-\nabla_{\bar{h},\bar{u}}G\cdot\Pi(0)\end{aligned} \tag{64}$$

where $\gamma_*=\gamma(0)$, the initial shock speed (obtained from the solution to the associated Riemann problem). □

Remark 3.3
For the convenience of the reader, we include the following routine calculation:

$$\begin{aligned}\nabla_{h,u}G&=(h-\bar{h})\left(-\frac{g}{2h}(2h^2+\bar{h}^2+h\bar{h}),2h\bar{h}\frac{u-\bar{u}}{h-\bar{h}}\right)\\ \nabla_{\bar{h},\bar{u}}G&=(\bar{h}-h)\left(-\frac{g}{2\bar{h}}(2\bar{h}^2+h^2+h\bar{h}),2h\bar{h}\frac{u-\bar{u}}{h-\bar{h}}\right)\end{aligned} \tag{65}$$

Note that

$$u-\gamma(t)=-\bar{h}\frac{u-\bar{u}}{h-\bar{h}},\quad \text{or}\quad \bar{u}-\gamma(t)=-h\frac{u-\bar{u}}{h-\bar{h}} \tag{66}$$

Then we get

$$A(h,u)-\gamma(t)I=\begin{pmatrix} -\bar{h}\dfrac{u-\bar{u}}{h-\bar{h}} & h \\ g & -\bar{h}\dfrac{u-\bar{u}}{h-\bar{h}} \end{pmatrix}=\begin{pmatrix} \mp\sqrt{\dfrac{g(\bar{h}+h)\bar{h}}{2h}} & h \\ g & \mp\sqrt{\dfrac{g(\bar{h}+h)\bar{h}}{2h}} \end{pmatrix}$$

$$A(\bar{h},\bar{u})-\gamma(t)I=\begin{pmatrix} -h\dfrac{u-\bar{u}}{h-\bar{h}} & h \\ g & -h\dfrac{u-\bar{u}}{h-\bar{h}} \end{pmatrix}=\begin{pmatrix} \mp\sqrt{\dfrac{g(\bar{h}+h)h}{2\bar{h}}} & h \\ g & \mp\sqrt{\dfrac{g(\bar{h}+h)h}{2\bar{h}}} \end{pmatrix} \tag{67}$$

where '+' corresponds to the 1-shock, and '−' to the 2-shock. In particular, γ_* is not an eigenvalue of $A(h_*,u_*)$ so that in (64) necessarily $a_+,\ b_+\neq 0$.

Remark 3.4

In the other cases where the t-axis is located in the intermediate state region, we apply a similar method. Thus, when the right wave is a 2-rarefaction wave R_μ, we follow the analysis carried out in Case 1, and obtain, in analogy with (57)

$$\begin{aligned}(a_+,b_+)&=(\lambda_*-\mu_*)\nabla_{h,u}r(A_*)^{-1}\\ d_+&=\left(\frac{\partial t}{\partial\beta}(\alpha_*,0)\right)^{-1}\frac{\partial r}{\partial\beta}(\alpha_*,0)-\mu_*\nabla_{h,u}r\cdot A_*^{-1}\cdot(0,\psi(0))\end{aligned} \tag{68}$$

where

$$\begin{aligned}\frac{\partial r}{\partial\beta}(\alpha_*,0)&=\frac{\partial r}{\partial\beta}(\alpha_+,0)+\psi(0)((3c_*)^{-3/2}-(3c_+)^{-3/2})\\ \frac{\partial r}{\partial\beta}(\alpha_+,0)&=\frac{1}{(3c_+)^{3/2}}(\psi(0)+2c_+\delta r_+)\\ \nabla_{h,u}r&=(-\sqrt{g/h_*},1),\quad \alpha_*=\mu_*,\quad \alpha_+=\mu_+=u_++c_+\end{aligned} \tag{69}$$

When the 1-shock $x=x(t)$ associated with λ emanates from $(0,0)$ and propagates to the left, we can use the expressions in (64) for (a_-,b_-,d_-) by replacing all subscripts '+' with '−', and γ_* is naturally taken to be the one associated with the 1-shock.

Case 3. Critical cases: As far as the critical case is concerned that the t-axis is located inside a rarefaction wave, the above arguments become simple. Indeed, as the t-axis is located inside the rarefaction wave R_λ, one of characteristic curve is tangent to the t-axis. The speed of this characteristic curve is $\beta_0=\lambda_0=0$, correspondingly $(h_0,u_0)=(h,u)(0,0_+)$. Then we

derive the first equation in (40) as in Case 1. Indeed, since $\lambda_0 = 0$, we have

$$\frac{\partial s}{\partial \alpha}(0, \beta_0) = \frac{\partial t}{\partial \alpha}(0, \beta_0)\left[\frac{\partial s}{\partial t}(0, \beta_0) + \lambda_0 \frac{\partial s}{\partial x}(0, \beta_0)\right] = \frac{\partial t}{\partial \alpha}(0, \beta_0)\frac{\partial s}{\partial t}(0, \beta_0) \tag{70}$$

This yields

$$\frac{\partial s}{\partial t}(0, \beta_0) = \left(\frac{\partial t}{\partial \alpha}(0, \beta_0)\right)^{-1} \frac{\partial s}{\partial \alpha}(0, \beta_0) \tag{71}$$

Therefore

$$(a_-, b_-) = \nabla_{h,u} s(h_0, u_0) = (\sqrt{g/h_0}, 1), \quad d_- = \left(\frac{\partial t}{\partial \alpha}(0, \beta_0)\right)^{-1} \frac{\partial s}{\partial \alpha}(0, \beta_0) \tag{72}$$

For the second equation in (40), we use the fact that $\lambda(h_0, u_0) = 0$, and

$$\left(\frac{\partial r}{\partial t}\right)_0 + \lambda(h_0, u_0)\left(\frac{\partial r}{\partial x}\right)_0 = \left(\frac{\partial r}{\partial t}\right)_0 = -gB'(0) \tag{73}$$

Then we have

$$(a_+, b_+) = \nabla_{h,u} r(h_0, u_0) = \left(-\sqrt{\frac{g}{h_0}}, 1\right), \quad d_+ = -gB'(0) \tag{74}$$

Likewise, as $x = 0$ is located inside the rarefaction wave R_μ, we choose (a_+, b_+, d_+) as in Remark 3.4 and (a_-, b_-, d_-) as

$$(a_-, b_-) = \nabla_{h,u}(h_0, u_0) = \left(\sqrt{\frac{g}{h_0}}, 1\right), \quad d_- = -gB'(0) \tag{75}$$

Case 4. Stationary shocks: When the Riemann problem associated with (28) and (29) yields a stationary shock, the detection of the curvature of the shock $x = x(t)$ at $(x, t) = (0, 0)$ is necessary. That is, we need to detect the sign of $x''(0)$. This is done in the following.

The Rankine–Hugoniot condition for the shock is

$$\begin{aligned} x'(t)(h - \bar{h}) &= (hu - \bar{h}\bar{u}) \\ x'(t)(hu - \bar{h}\bar{u}) &= (hu^2 + \tfrac{1}{2} gh^2 - \bar{h}\bar{u}^2 - \tfrac{1}{2} g\bar{h}^2) \end{aligned} \tag{76}$$

Differentiating the first equation along $x = x(t)$ and taking into account $x'(0) = 0$, one arrives at

$$x''(0) = \frac{1}{h_* - \bar{h}}\left[\left(\frac{\partial (hu)}{\partial t}\right)_0 - \left(\frac{\partial (\bar{h}\bar{u})}{\partial t}\right)_0\right] \tag{77}$$

where the subscripts represent the values at $(0, 0)$. With this formula, we easily conclude the sign of $x''(0)$. As $x''(0) > 0$, this shock is regarded as a right-going wave; as $x''(0) < 0$, this

wave is a left-going wave. For the critical case that $x''(0)=0$, we can deduce

$$\begin{gathered}\left(\frac{\partial(hu)}{\partial t}\right)_0-\left(\frac{\partial(\bar{h}\bar{u})}{\partial t}\right)_0=0\\ \left(\frac{\partial(hu^2+\frac{1}{2}gh^2)}{\partial t}\right)_0-\left(\frac{\partial(\bar{h}\bar{u}^2+\frac{1}{2}g\bar{h}^2)}{\partial t}\right)_0=0\end{gathered}\tag{78}$$

Therefore, we can regard it as either a right-going or a left-going wave.

3.2. The time derivative of the velocity v for the augmented one-dimensional shallow water equations

Note that the time derivatives of (h,u) has been obtained in Section 3.1. To compute the time derivative of v, we need to introduce auxiliary characteristic co-ordinate $\xi(x,t)=C$ that is defined as the integral curve of the equation

$$\frac{\mathrm{d}x}{\mathrm{d}t}=u\tag{79}$$

Consider the typical case that a rarefaction wave moves to the left and the shock to the right as shown in Figure 2. The difference from Figure 1 is that a contact discontinuity for v is located in the intermediate region and propagates with speed u. Then we have

Proposition 3.5
Assume that a rarefaction wave moves to the left and a shock wave moves to the right, the t-axis is located in the intermediate region, as shown in Figure 2. Then

(i) if $u_*>0$, the time derivative of v is computed from the rarefaction wave side and

$$\left(\frac{\partial v}{\partial t}\right)_0=-\left(\frac{c_*}{c_-}\right)^3 Fr_* c_-\left(\frac{\partial v}{\partial x}\right)_-\tag{80}$$

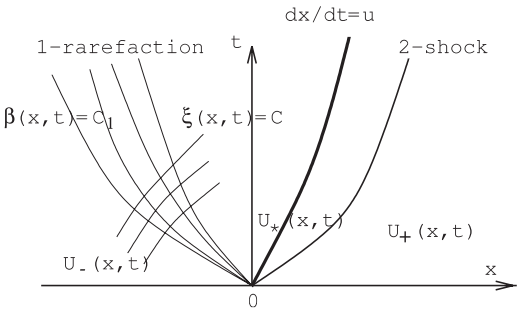

Figure 2. Typical wave pattern for the augmented system.

(ii) if $u_*<0$, the time derivative of v is computed from the shock wave side and

$$\left(\frac{\partial v}{\partial t}\right)_0 = -\frac{u_*(\gamma_+ - u_+)}{\gamma_+ - u_*}\left(\frac{\partial v}{\partial x}\right)_+ \tag{81}$$

Proof

First we consider (i). In analogy with the proof of Case 1 of Proposition 3.2, we now map the rarefaction wave by the co-ordinates (β, ξ). To obtain the expressions for (x, t), in the case of the associated Riemann problem, we solve $\mathrm{d}x/\mathrm{d}t = u = s_- - 2c$ and $x/t = u - c = s_- - c$ and get, along a streamline, $(s_- - \beta)t^{1/3} = \text{const}$. Taking this constant as $\xi^{1/3}$, we obtain

$$t^{\text{ass}}(\beta, \xi) = \frac{\xi}{(s_- - \beta)^3}, \quad x^{\text{ass}}(\beta, \xi) = \frac{\xi\beta}{(s_- - \beta)^3} \tag{82}$$

Then on the left side of the t-axis, the characteristic co-ordinate transform $(\beta, \xi) \to (x, t)$ at the singularity point $(0, 0)$ is

$$t(\beta, \xi) = t^{\text{ass}}(\beta, \xi) + O(\xi^2), \quad x(\beta, \xi) = x^{\text{ass}}(\beta, \xi) + O(\xi^2) \tag{83}$$

as $\xi \to 0$. It follows that

$$\frac{\partial t}{\partial \xi}(0, \beta) = \frac{1}{(s_- - \beta)^3} = \frac{1}{(3c)^3} \tag{84}$$

Consider the equations

$$\begin{aligned} &\frac{\partial r}{\partial t} + \lambda \frac{\partial r}{\partial x} = \psi \\ &\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} = 0 \end{aligned} \tag{85}$$

In terms of characteristic co-ordinates (β, ξ), (85) becomes

$$\frac{\partial r}{\partial \xi} = \psi \frac{\partial t}{\partial \xi}, \quad \frac{\partial v}{\partial \beta} = 0 \tag{86}$$

Then we have

$$\frac{\partial^2 v}{\partial \beta \partial \xi} = 0 \tag{87}$$

which gives

$$\frac{\partial v}{\partial \xi}(0, \beta) = \frac{\partial v}{\partial \xi}(0, \beta_-) \tag{88}$$

Note that using the advection equation for v in (85) gives

$$\frac{\partial v}{\partial \xi} = \frac{\partial t}{\partial \xi}\left(\frac{\partial v}{\partial t} + \lambda \frac{\partial v}{\partial x}\right) = \frac{\partial t}{\partial \xi}(\lambda - u)\frac{\partial v}{\partial x} = -c \frac{\partial t}{\partial \xi} \cdot \frac{\partial v}{\partial x} \tag{89}$$

So we arrive at

$$\frac{\partial v}{\partial \xi}(0,\beta_-) = -c_- \cdot \frac{\partial t}{\partial \xi}(0,\beta_-) \cdot \frac{\partial v}{\partial x}(0,\beta_-) = -c_-(3c_-)^{(-3)} \left(\frac{\partial v}{\partial x}\right)_- \tag{90}$$

The advection equation for v in (85) also gives

$$\frac{\partial v}{\partial \xi} = \frac{\partial t}{\partial \xi}\left(\frac{\partial v}{\partial t} + \lambda \frac{\partial v}{\partial x}\right) = \frac{\partial t}{\partial \xi}\left(\frac{\partial v}{\partial t} - \frac{\lambda}{u}\frac{\partial v}{\partial t}\right) = \frac{1}{Fr}\frac{\partial t}{\partial \xi}\frac{\partial v}{\partial t} \tag{91}$$

Therefore,

$$\begin{aligned}\left(\frac{\partial v}{\partial t}\right)_0 &= Fr_* \left(\frac{\partial t}{\partial \xi}\right)^{-1}(0,\beta_*) \cdot \frac{\partial v}{\partial \xi}(0,\beta_*) = Fr_*(3c_*)^3 \frac{\partial v}{\partial \xi}(0,\beta_-) \\ &= -c_- Fr_* \left(\frac{c_*}{c_-}\right)^3 \left(\frac{\partial v}{\partial x}\right)_-\end{aligned} \tag{92}$$

This is (80).

For (81) in (ii), we need to compute from the shock wave side. Note that the quantity v is continuous across the shock, we have

$$v = v_+ \tag{93}$$

where v and v_+ are the limiting states on the two sides of the shock. Differentiate this equality along the shock wave $x = x(t)$, $\gamma(t) = x'(t)$. Then we get

$$\frac{\partial v}{\partial t} + \gamma(t)\frac{\partial v}{\partial x} = \left(\frac{\partial v}{\partial t}\right)_+ + \gamma(t)\left(\frac{\partial v}{\partial x}\right)_+ \tag{94}$$

Using the advection equation for v in (85) yields

$$\left(1 - \frac{\gamma(t)}{u}\right)\frac{\partial v}{\partial t} = -u_+ \left(\frac{\partial v}{\partial x}\right)_+ + \gamma(t)\left(\frac{\partial v}{\partial x}\right)_+ \tag{95}$$

Taking the limit as $t \to 0$, we get (81). □

4. THE WELL-BALANCED PROPERTY

One of important criteria to justify a numerical scheme is to check if it can replicate the exact solution to the stationary flow problem,

$$h + B \equiv \text{constant}, \quad (u, v) \equiv 0 \tag{96}$$

This is the Z-property in Reference [15] and references therein. When a centred discretization is used for source terms, the scheme is said to satisfy the C-property.

Proposition 4.1 (Z-*Property*)
The numerical scheme in Sections 2 and 3 satisfies the Z-property. That is, let the initial state $(h_0(x, y), u_0(x, y), v_0(x, y))$ satisfy

$$(u_0, v_0)(x, y) = 0, \quad h_0(x, y) + B(x, y) = C \tag{97}$$

where C is constant. Then

$$u_{ij}^n \equiv 0, \quad v_{ij}^n \equiv 0, \quad h_{ij}^n + B(x_i, y_j) \equiv C \tag{98}$$

for all $n = 0, 1, 2, \ldots$.

Proof
Since we use the Strang splitting method, it suffices to show this proposition by considering (5) and use a single subscript for each component. Assume that $u_j^n \equiv 0$, $v_j^n \equiv 0$, and $h_j^n + B(x_j) \equiv C$. Then we want to prove

$$u_j^{n+1} = 0, \quad v_j^{n+1} = 0, \quad h_j^{n+1} + B(x_j) = C \tag{99}$$

Note that the data reconstruction step in Section 2 yields

$$\delta\eta_j^n = 0, \quad \delta h_j^n + B'(x_j) = 0, \quad \delta u_j^n = 0, \quad \delta v_j^n = 0 \tag{100}$$

The flow is subcritical and the waves from each grid interface are rarefaction waves. With the results in Section 3, we can get

$$\left(\frac{\partial h}{\partial t}\right)_{j+1/2}^n = 0, \quad \left(\frac{\partial u}{\partial t}\right)_{j+1/2}^n = 0, \quad \left(\frac{\partial v}{\partial t}\right)_{j+1/2}^n = 0 \tag{101}$$

Thus from (25) we conclude

$$h_{j+1/2}^{n+1/2} = h_{j+1/2}^n, \quad (hu)_{j+1/2}^{n+1/2} = (hu)_{j+1/2}^n = 0, \quad (hv)_{j+1/2}^{n+1/2} = (hv)_{j+1/2}^n = 0 \tag{102}$$

Therefore,

$$h_j^{n+1} + B(x_j) = h_j^n - \frac{\Delta t}{\Delta x}((hu)_{j+1/2}^{n+1/2} - (hu)_{j-1/2}^{n+1/2}) + B(x_j) = h_j^n + B(x_j) \tag{103}$$

and

$$\begin{aligned}(hu)_j^{n+1} = (hu)_j^n &- \frac{\Delta t}{\Delta x}\left(\left(hu^2 + \frac{1}{2}gh^2\right)_{j+1/2}^{n+1/2} - \left(hu^2 + \frac{1}{2}gh^2\right)_{j-1/2}^{n+1/2}\right) \\ &- \frac{g\Delta t}{2\Delta x}(h_{j+1/2}^{n+1/2} + h_{j-1/2}^{n+1/2})(B(x_{j+1/2}) - B(x_{j-1/2})) \\ &= -\frac{g\Delta t}{2\Delta x}(h_{j+1/2}^{n+1/2} + h_{j-1/2}^{n+1/2})(h_{j+1/2}^{n+1/2} + B(x_{j+1/2}) - h_{j-1/2}^{n+1/2} - B(x_{j-1/2})) \\ &= 0\end{aligned} \tag{104}$$

This implies

$$u_j^{n+1} = 0 \tag{105}$$

Moreover, we have for v,

$$\begin{aligned}(hv)_j^{n+1} &= (hv)_j^n - \frac{\Delta t}{\Delta x}((huv)_{j+1/2}^{n+1/2} - (huv)_{j-1/2}^{n+1/2}) \\ &= (hv)_j^n = 0\end{aligned} \tag{106}$$

which yields,

$$v_j^{n+1} = 0 \tag{107}$$

□

Remark 4.2
One of the important issue for the well-balanced property is to verify the non-stationary case, which becomes quite tedious. We will show the property numerically in the next section.

5. NUMERICAL EXPERIMENTS

In this section we demonstrate the scheme proposed in Sections 2 and 3 by testing several quasisteady and unsteady problems for the shallow water equations (1).

5.1. Quasi-stationary flows

The first case is proposed by LeVeque [9] to demonstrate the capability of the proposed scheme for computations involving small perturbations of the water surface. The bed topography is

$$B(x) = \begin{cases} 0.25[\cos(\pi(x - 0.5)/0.1) + 1] & \text{for } |x - 0.5| < 0.1 \\ 0 & \text{otherwise} \end{cases} \tag{108}$$

in a channel having a length 1.0, and $g \equiv 1$ for the present case. The data is initially stationary ($u = 0$) with a small perturbation for the river height

$$h(x) = \begin{cases} 1.0 + \varepsilon & \text{for } 0.1 < x < 0.2 \\ 1.0 - B(x) & \text{otherwise} \end{cases} \tag{109}$$

As pointed out in References [9, 14, 15], many numerical methods have difficulty with the calculation involving such small perturbations of the water surface. We show here our solutions in Figure 3, and reflection boundary conditions are used at both ends $x = 0$ and 1,

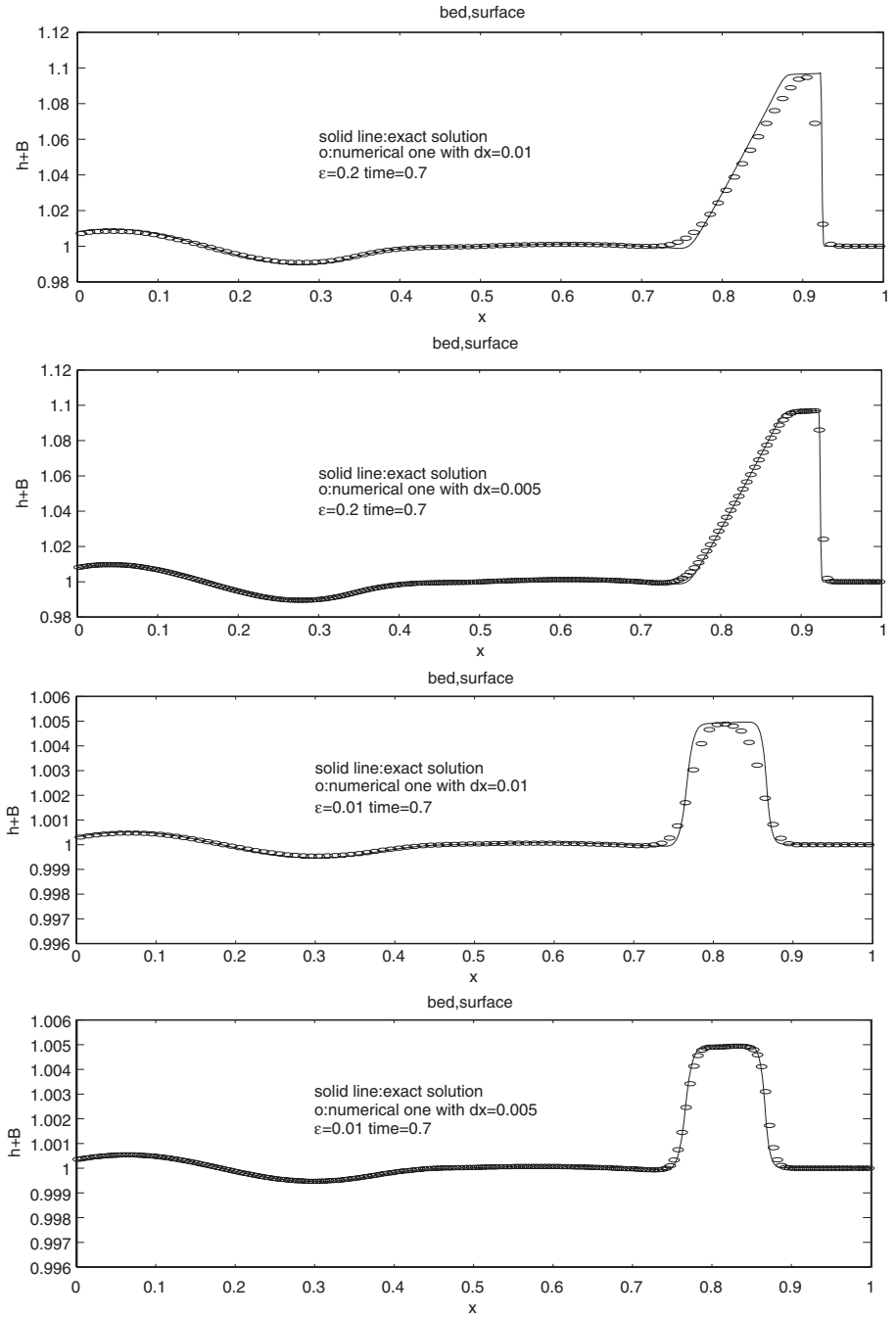

Figure 3. Quasi-stationary water equations with perturbation: Comparison of water elevation surface at output time $t = 0.7$. The top two figures are about the perturbation $\varepsilon = 0.2$ where calculations are made with 100 and 200 grid points, respectively. The bottom two are about the perturbation $\varepsilon = 0.01$ with 100 and 200 grid points, respectively.

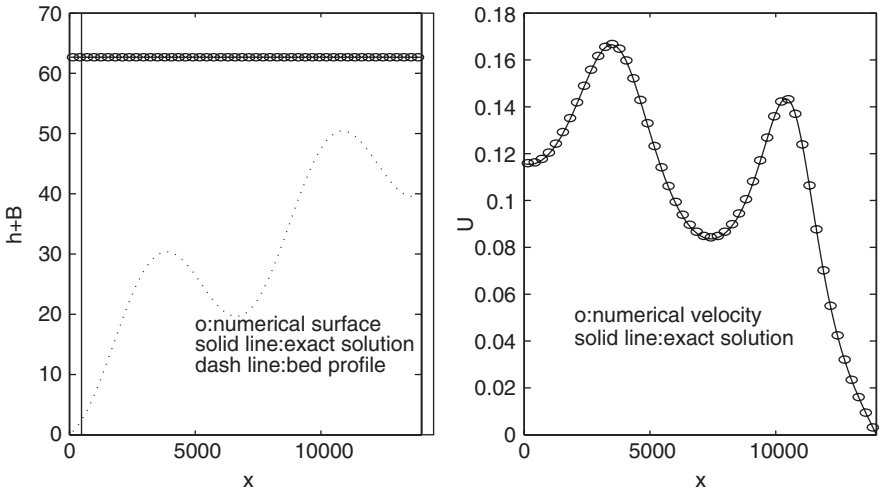

Figure 4. An unsteady test problem: Comparison of surface $\eta = h + B$ (the left) and velocity u (the right). The numerical solution is calculated with 50 grid points.

see Reference [9]. The numerical solutions are compared with the corresponding accurate solutions that are obtained with 1000 grid points. The agreement between the numerical and exact solutions is quite satisfactory.

The second quasi-stationary test problem is proposed by Bermudez and Vazquez [25]. The bottom topography is

$$B(x) = 10 + \frac{40x}{L} + 10 \sin\left[\pi\left(\frac{4x}{L} - \frac{1}{2}\right)\right] \tag{110}$$

where $L = 14,000\,\text{m}$ is the channel length. The initial and boundary conditions are

$$h(x, 0) = 60.5 - B(x), \quad u(x, 0) = 0 \tag{111}$$

and

$$h(0, t) = 64.5 - 4.0 \sin\left[\pi\left(\frac{4t}{86,400} + \frac{1}{2}\right)\right], \quad u(L, t) = 0.0 \tag{112}$$

With the above conditions, there is an approximate solution based on the asymptotic analysis given in [20]

$$h(x, t) = 64.5 - B(x) - 4.0 \sin\left[\pi\left(\frac{4t}{86,400} + \frac{1}{2}\right)\right] \tag{113}$$

and

$$u(x, t) = \frac{(x - L)\pi}{5400h(x, t)} \cos\left[\pi\left(\frac{4t}{86,400} + \frac{1}{2}\right)\right] \tag{114}$$

This is an unsteady flow. The numerical solutions $h + B$ and u are shown in Figure 4, in comparison with the asymptotic solutions above at time $t = 7552.13\,\text{s}$.

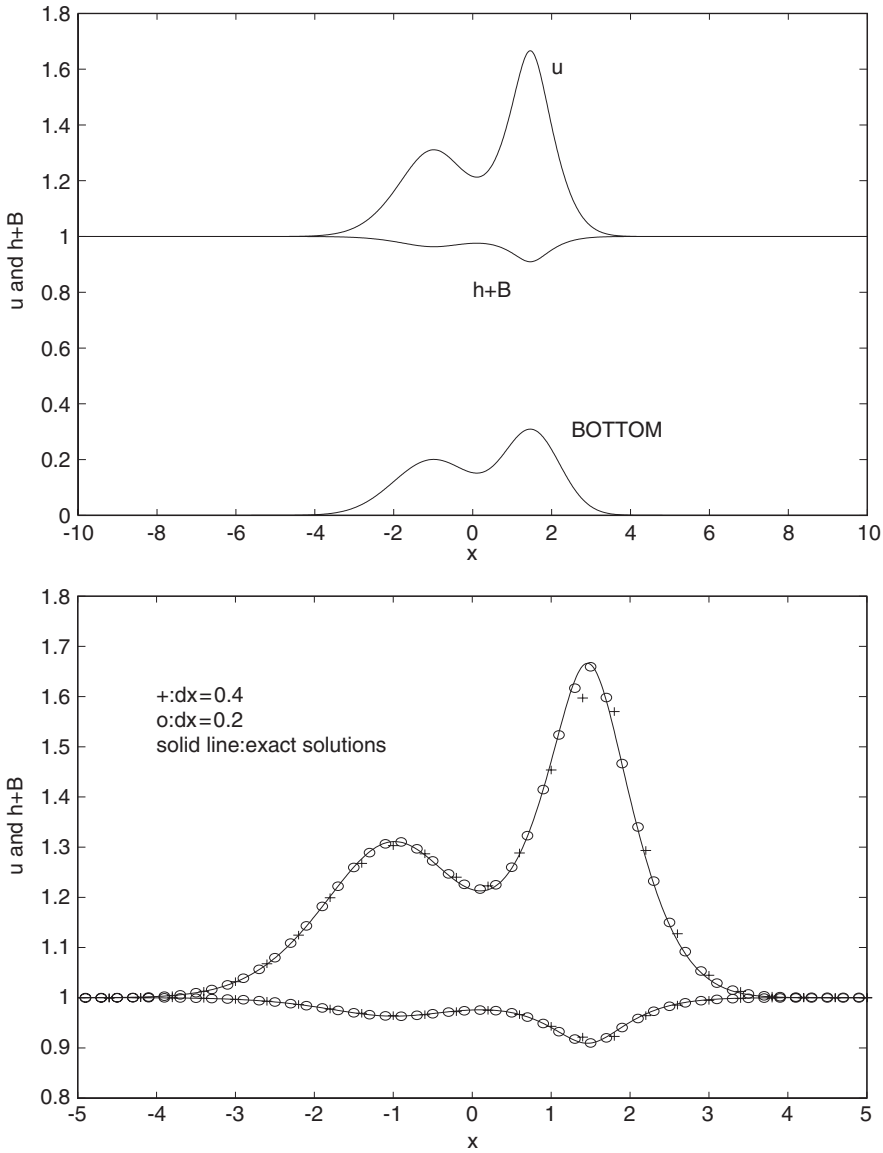

Figure 5. Steady state solutions of shallow water equations: The top figure shows the exact solution, the bottom figure shows the comparison of surface and velocity with the exact solutions.

5.2. *Steady flows*

This case was used in Reference [14] to check the dissipative and dispersive errors in the kinetic schemes. The bottom topography $B(x)$ is

$$B(x) = 0.2\mathrm{e}^{-(x+1)^2/2} + 0.3\mathrm{e}^{-(x-1.5)^2} \tag{115}$$

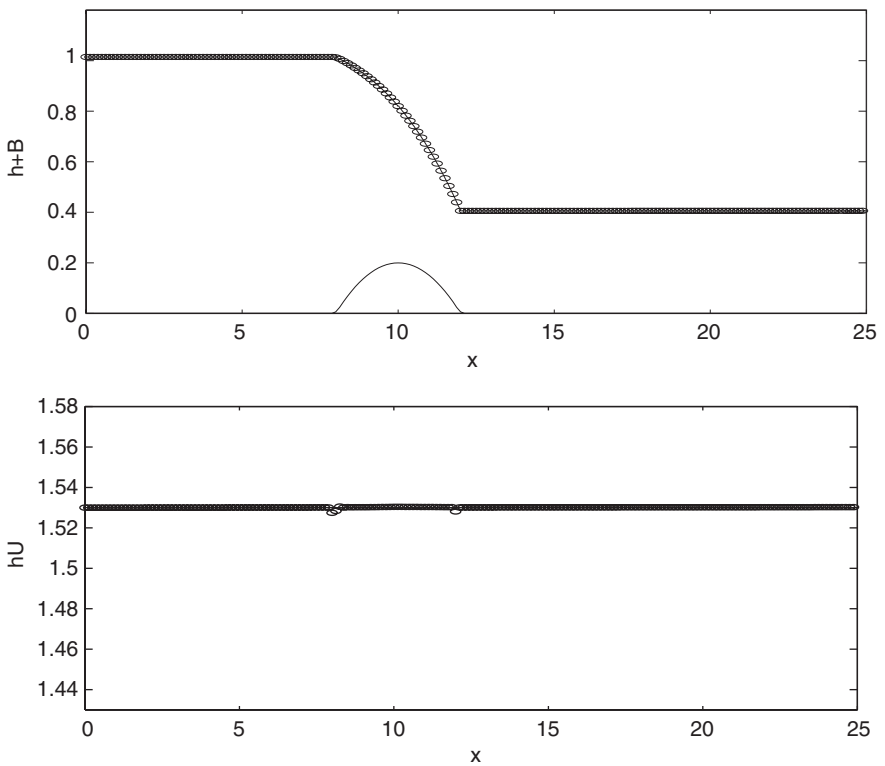

Figure 6. Steady transcritical flow over a bump without a shock: The surface is shown in the top figure, the discharge is shown in the bottom figure.

The equilibrium state over a computation domain $[-10, 10]$ is obtained by assuming the left boundary condition to be $h_- = 1.0$ and $u_- = 1.0$. The flow is steady and subcritical. Here we compare the numerical results with the analytic solutions and find that the current method has a good accuracy, see Figure 5.

5.3. *Transcritical flows*

We use the classical example to test our algorithm for transcritical flows, see References [14, 15]. The 1D simulation channel has a length of 25 m. The bed topography is

$$B(x) = \begin{cases} 0.2 - 0.05(x - 10)^2 & \text{for } 8 < x < 12 \\ 0 & \text{otherwise} \end{cases} \tag{116}$$

Transcritical flow without a shock: A discharge per unit width of $hu = 1.53\,\mathrm{m}^2/\mathrm{s}$ was imposed as the upstream boundary condition $x = 0$. No boundary condition is needed at the downstream end of the channel $x = 25$ if the flow is supercritical there, and the water level $h = 0.66$ is imposed otherwise. The numerical result is shown in Figure 6 at time $t = 250$.

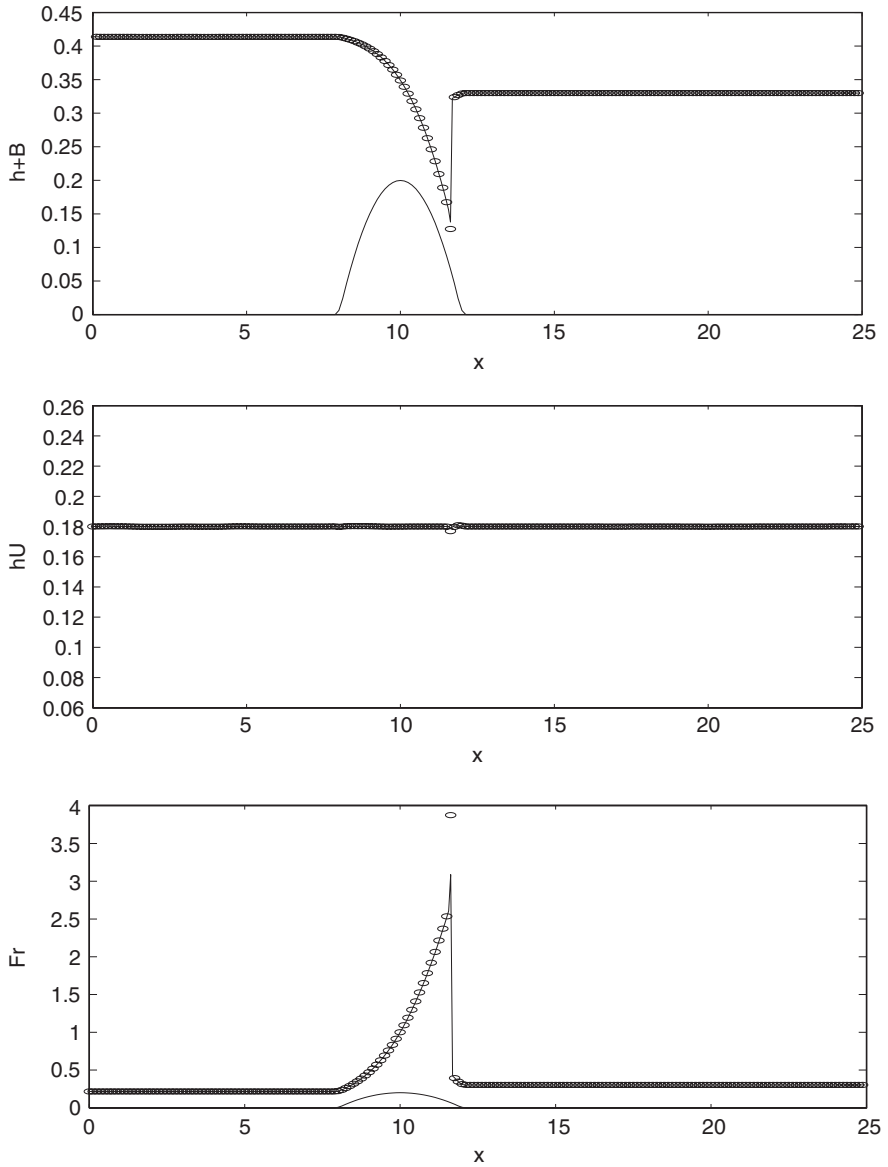

Figure 7. Steady transcritical flow over a bump with a shock: Water surface elevation on the top, discharge in the middle and Froude number at the bottom. The solid lines are the numerical results calculated with refined grids.

Transcritical flow with a shock: The discharge hu has a value $hu = 0.18\,\mathrm{m}^2/\mathrm{s}$ at the upstream boundary $x = 0$, and the water level $h = 0.33\,\mathrm{m}$ is taken to be the downstream boundary condition at $x = 25$. The numerical solution is shown in Figure 7.

Table I. Deviation from flatness in water surface at time $t = 0.1$ for 2D shallow water equations in the stationary case, on $N \times N$ grids.

N	$t = 1.7$ (GRP method)	$t = 0.1$ (GRP method)	$t = 0.1$ (Quasi-steady method)	$t = 0.1$ (Strang method)
50	3.178×10^{-8}	3.162×10^{-8}	1.2×10^{-3}	1.4×10^{-3}
100	4.7517×10^{-8}	4.7775×10^{-8}	2.5×10^{-4}	5.5×10^{-4}
200	4.87306×10^{-8}	6.6639×10^{-8}	6.3×10^{-5}	1.7×10^{-4}

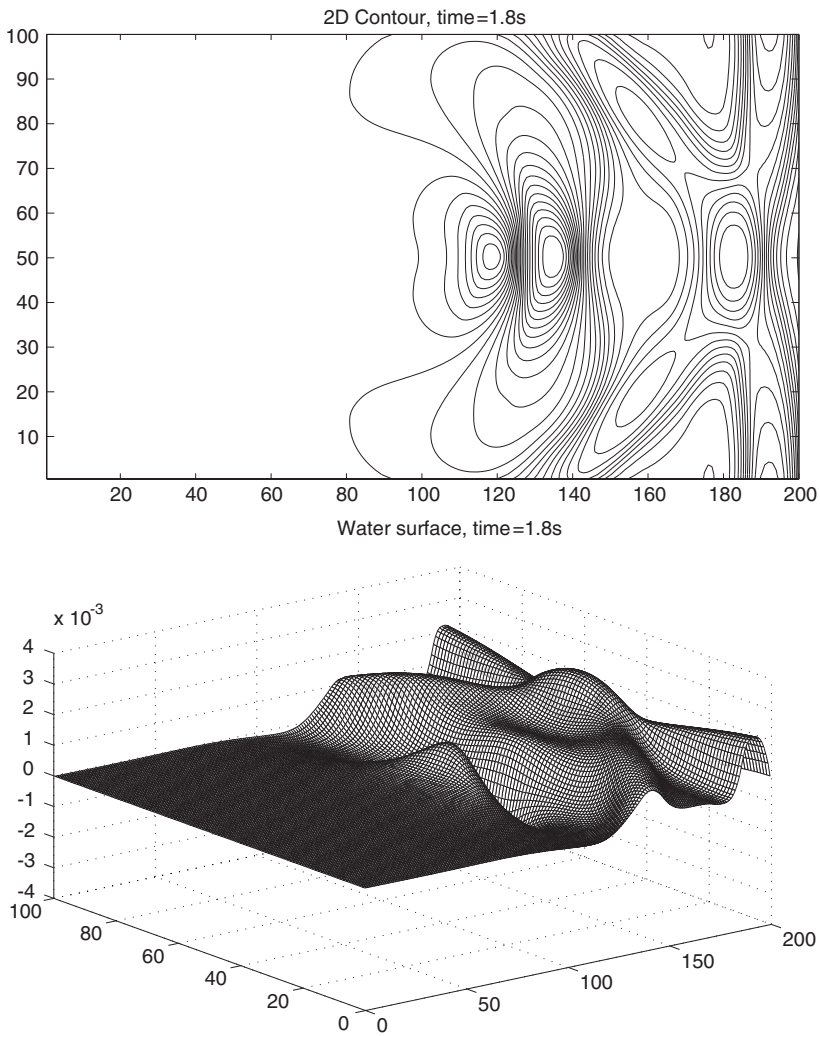

Figure 8. Quasi-stationary flow in 2-D. 200×100 grids are used in the calculation to show the water surface with the initial perturbation $\varepsilon = 0.01$.

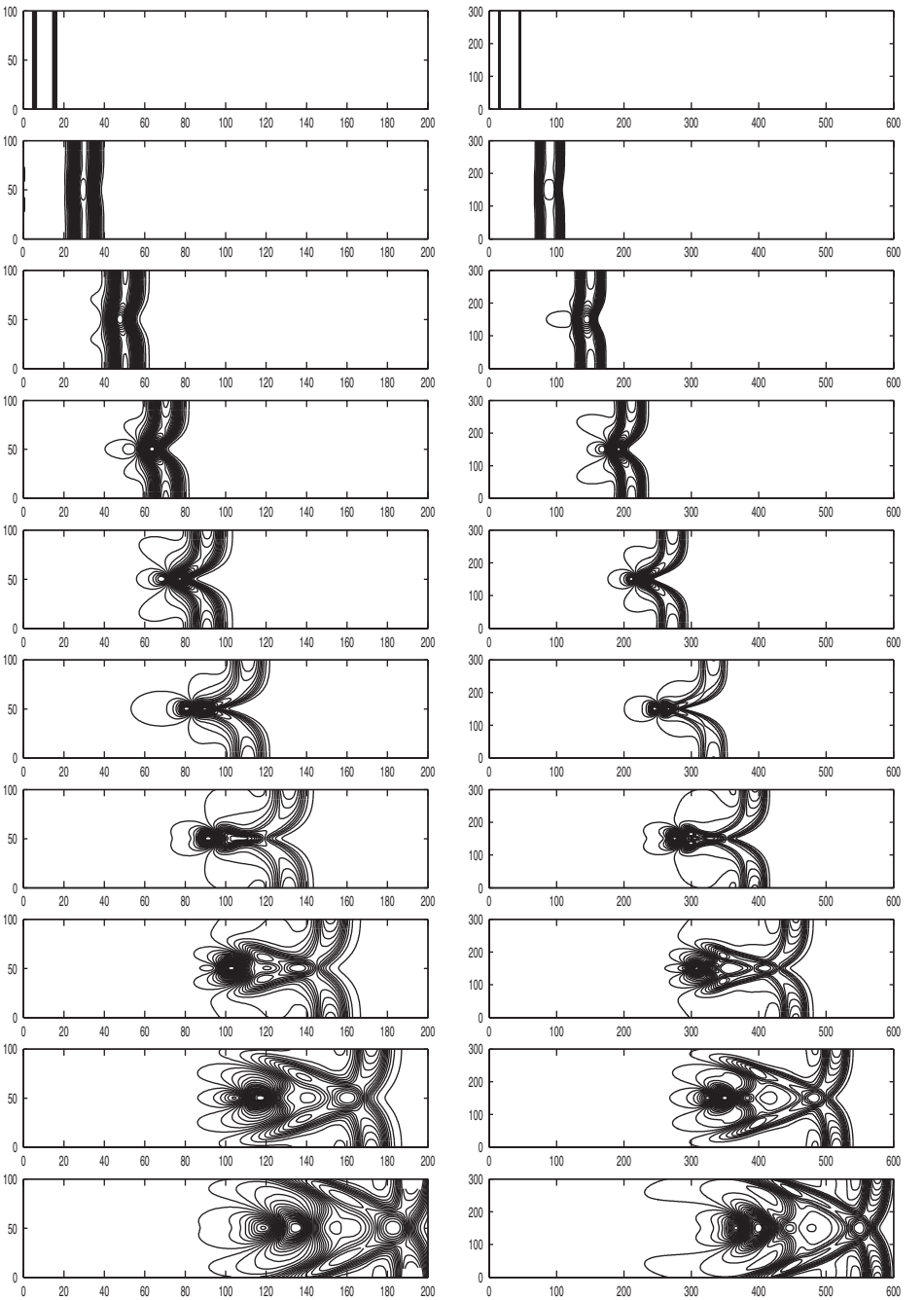

Figure 9. The propagation of water surface of example is computed.
Left column: 200×100, Right column: 600×300.

5.4. Two-dimensional numerical experiments

We show two two-dimensional numerical experiments, which can be found in References [9, 14], to show the accuracy of our scheme. The bottom topography is so discretized to keep the continuity. Otherwise, the accuracy is greatly lowered.

The first case is about the preservation of stationary flows. The bottom topography is

$$B(x, y) = 0.8 \exp(-50((x - 0.5)^2 + (y - 0.5)^2)), \quad (x, y) \in [0, 1] \times [0, 1] \tag{117}$$

The initial state is set to be $h(x, y, 0) = 1 - B(x, y)$, and $(u, v)(x, y, 0) \equiv 0$. Table I is the comparison of our algorithm with other two schemes, Quasi-steady method and Strang splitting method, for the max-norm of $h - 1 + B$ at time $t = 0.1$ and 1.7. The results in Table I show the quick convergence of our scheme.

The second example is similar to the case in one dimension with an elliptic bottom topography

$$B(x, y) = 0.8 \exp(-5(x - 0.9)^2 - 50(y - 0.5)^2), \quad (x, y) \in [0, 2] \times [0, 1] \tag{118}$$

The initial surface of water is $h(x, y, 0) = 1 - B(x, y)$ except for $x \in [0.05, 0.15]$ where the surface is perturbed with $\varepsilon = 0.01$. Our numerical result is shown in Figures 8 and 5.3. It can be seen that our algorithm is comparable to other schemes, such as the quasi-steady method in Reference [9], etc.

ACKNOWLEDGEMENTS

We thank M. Ben-Artzi and Zhouping Xin for their stimulating discussions, which essentially improve the present paper. This work was partially done when the second author was visiting the Institute of Mathematical Sciences, the Chinese University of Hongkong, and revised during his visit to IAN, Magdeburg University. The supports from the Zheng Ge Ru foundation and Alexander von Humboldt Foundation are gratefully acknowledged. This research is also supported in part under the grant No. 10301022 of NNSF of China, the Natural Science Foundation from Beijing Municipality and Fok Ying Tong Education Foundation.

REFERENCES

1. Gerbeau JF, Perthame B. Derivation of viscous Saint-Venant system for laminar shallow water, numerical validation. *Discrete and Continuous Dynamical Systems*, *Series B* 2001; **1**:89–102.
2. Keller JB. Shallow water theory for arbitrary slopes of the bottom. *Journal of Fluid Mechanics* 2003; **489**: 345–348.
3. Alcrudo F, Benkhaldoun F. Exact solutions to the Riemann problem of the shallow water equations with a bottom step. *Computers and Fluids* 2001; **30**(6):643–671.
4. Audusse E, Bouchut F, Bristeau M, Klein R, Perthame B. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM Journal on Scientific Computing* 2004; **25**(6): 2050–2065.
5. Audusse E, Bristeau M-O, Perthame B. Kinetic schemes for Saint-Venant equations with source terms on unstructured grids. *INRIA Report*, *RR 3989*, 2000.
6. Bermudez A, Dervieux A, Desideri J, Vazquez M. Upwind schemes for the two-dimensional shallow water equations with variable depth using unstructured meshes. *Computer Methods in Applied Mechanics and Engineering* 1998; **155**(49):49–72.
7. Chinnayya A, Le Roux A-Y. A new general Riemann solver for the shallow water equations with friction and topography. http://www.math.ntnu.no/conservation/1999

8. Jin S. A steady-state capturing method for hyperbolic systems with geometrical source terms. *Mathematical Modeling in Numerical Analysis* 2001; **35**(4):631–645.
9. LeVeque RJ. Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. *Journal of Computational Physics* 1998; **146**(1):346–365.
10. Lukacova M, Vlk V. Well-balanced finite volume evolution Galerkin methods for the shallow water equations with source terms. *International Journal for Numerical Methods in Fluids* 2005; **47**(10–11):1165–1171.
11. Noussair A. Riemann problem with nonlinear resonance effects and well-balanced Godunov scheme for shallow fluid flow past an obstacle. *SIAM Journal on Numerical Analysis* 2001; **39**(1):52–72.
12. Perthame B, Simeoni C. A kinetic scheme for the Saint-Venant system with a source term. *Calcolo* 2001; **38**(4):201–231.
13. Vazquez-Cendon ME. Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry. *Journal of Computational Physics* 1999; **148**(2):497–526.
14. Xu K. A well-balanced gas-kinetic scheme for the shallow-water equations with source terms. *Journal of Computational Physics* 2002; **178**(2):533–562.
15. Zhou JG, Causon DM, Mingham CG, Ingram DM. The surface gradient method for the treatment of source terms in the shallow-water equations. *Journal of Computational Physics* 2001; **168**(1):1–25.
16. Donea J. A Taylor–Galerkin method for convection transport problems. *International Journal for Numerical Methods in Engineering* 1984; **20**:101.
17. Greenberg JM, Leroux AY. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis* 1996; **33**(1):1–16.
18. Greenberg JM, Leroux AY, Baraille R, Noussair A. Analysis and approximation of conservation laws with source terms. *SIAM Journal on Numerical Analysis* 1997; **34**(5):1980–2007.
19. Toro E. *Riemann Solvers and Numerical Methods for Fluid Dynamics. A Practical Introduction*. Springer: Berlin, 1997.
20. Ben-Artzi M, Falcovitz J. A second-order Godunov-type scheme for compressible fluid dynamics. *Journal of Computational Physics* 1984; **55**(1):1–32.
21. Ben-Artzi M, Falcovitz J. *Generalized Riemann Problems in Computational Gas Dynamics*. Cambridge University Press: Cambridge, MA, 2003.
22. Ben-Artzi M, Falcovitz J. An upwind second-order scheme for compressible duct flows. *SIAM Journal on Scientific and Statistical Computing* 1986; **7**(3):744–768.
23. LeFloch Ph, Raviart P-A. An asymptotic expansion for the solution of the generalized Riemann problem. I. General theory. *Annales de l'Institut Henri Poincaré. Analyse Non Linéaire* 1988; **5**(2):179–207.
24. Bourgeade A, LeFloch Ph, Raviart P-A. An asymptotic expansion for the solution of the generalized Riemann problem. II. Application to the equations of gas dynamics. *Annales de l'Institut Henri Poincaré. Analyse Non Linéaire* 1989; **6**(6):437–480.
25. Bermudez A, Vazquez M. Upwind methods for hyperbolic conservation laws with source terms. *Computers and Fluids* 1994; **23**(8):1049–1071.

