NUMERICAL METHODS FOR RIVER FLOW MODELLING*

Marek Brandner, Jiří Egermaier, Hana Kopincová

Abstract

In this paper we propose a new numerical scheme to simulate the river flow in the presence of a variable bottom surface. We use the finite volume methods, our approach is based on the technique described by D. L. George for shallow water equations [1]. The main goal is to construct the scheme, which is well balanced, i.e. maintains not only some special steady states but all steady states which can occur. Furthermore this should preserve non–negativity of some quantities, which are essentially non–negative from their physical fundamental, for example cross section or depth. Our scheme can be extended to the second order accuracy. We also describe connections between the central and central-upwind schemes and the approximate Riemann solvers.

1. Introduction

We are interested in solving the problem describing the fluid flow through the channel with the general cross-section area

$$\begin{aligned} a_t + q_x &= 0, \\ q_t + \left(\frac{q^2}{a} + gI_1\right)_x &= -gaB_x + gI_2, \end{aligned} \tag{1}$$

where $a = a(x,t)$ is the unknown cross-section area, $q = q(x,t)$ is the unknown discharge, $B = B(x)$ is the function of elevation of the bottom, g is the gravitational constant and

$$I_1 = \int_0^{h(x)} [h(x) - \eta]\sigma(x,\eta)d\eta, \tag{2}$$

$$I_2 = \int_0^{h(x)} (h - \eta)\left[\frac{\partial \sigma}{\partial x}\right] d\eta, \tag{3}$$

where η is the depth integration variable, h is the water depth and $\sigma(x,\eta)$ is the width of the cross-section at the depth η.

The special cases are the equations reflecting the fluid flow through the varying rectangular channel

$$\begin{aligned} a_t + q_x &= 0, \\ q_t + \left(\frac{q^2}{a} + \frac{ga^2}{2l}\right)_x &= \frac{ga^2}{2l^2}l_x - gab_x, \end{aligned} \tag{4}$$

*This work was supported by the Research Plan MSM 4977751301 and by Moravian-Silesian region.

where $l = l(x)$ is a function describing the width of the channel, and the system for the constant rectangular channel (the shallow water equations)

$$\begin{aligned} h_t + (hu)_x &= 0, \\ (hu)_t + \left(hu^2 + \frac{1}{2}gh^2\right)_x &= -ghB_x, \end{aligned} \tag{5}$$

where $h(x,t)$ is the water depth and $u(x,t)$ is the horizontal velocity.

All of the presented systems can be briefly written in the matrix form

$$\mathbf{q}_t + [\mathbf{f}(\mathbf{q},x)]_x = \boldsymbol{\psi}(\mathbf{q},x), \tag{6}$$

where $\mathbf{q}(x,t)$ is the vector of conserved quantities, $\mathbf{f}(\mathbf{q},x)$ is the flux function and $\boldsymbol{\psi}(\mathbf{q},x)$ is the source term. We note that this relation represents the balance laws.

There are many numerical schemes for solving (6) with different properties and possibilities of failing. For example the central, upwind and central-upwind schemes. The main requirements on the numerical schemes are the consistency (in the finite volume meaning, i.e. consistency with the flux function), the conservativity (if there is possibility to rewrite the problem to the conservative form it is required to have conservative numerical scheme), positive semidefiniteness, i.e. the schemes preserve nonnegativity of some quantities, which are essentially nonnegative from their physical fundamental, and the well-balancing, i.e. the schemes maintain some or all steady states which can occur. The next properties are the order of the schemes and stability.

2. Augmented formulations

There are several ways how to formulate the fluid flow problems. Homogeneous, autonomous, conservative formulation usually used for standard cases like Euler equations or fluid flow through the channel with constant cross-section and flat bottom, has the form

$$\begin{aligned} \mathbf{q}_t + [\mathbf{f}(\mathbf{q})]_x &= \mathbf{0},\ x \in \mathbf{R},\ t \in (0,T), \\ \mathbf{q}(x,0) &= \mathbf{q}_0(x),\ x \in \mathbf{R}, \end{aligned} \tag{7}$$

where $\mathbf{q} = \mathbf{q}(x,t) :\ \mathbf{R} \times \langle 0,T) \to \mathbf{R}^m$ is an unknown function, $\mathbf{q}_0 = \mathbf{q}_0(x) :\ \mathbf{R} \to \mathbf{R}^m$ is an initial function, $\mathbf{f} = \mathbf{f}(\mathbf{q}) :\ \mathbf{R}^m \to \mathbf{R}^m$ is a given flux function. This formulation corresponds to (6) with zero right hand side and it represents the conservation law in the differential form. It is necessary to note that the local differential form is obtained from more fundamental integral form under special assumptions.

The homogeneous, nonautonomous, conservative case has the form

$$\begin{aligned} \mathbf{q}_t + [\mathbf{f}(\mathbf{q}, \mathbf{w}(x))]_x &= \mathbf{0}, \ x \in \mathbf{R}, \ t \in (0, T), \\ \mathbf{q}(x, 0) &= \mathbf{q}_0(x), \ x \in \mathbf{R}, \end{aligned} \tag{8}$$

where $\mathbf{w} = \mathbf{w}(x) : \mathbf{R} \to \mathbf{R}^s$ is a given function.

The system (8) can be rewritten to the homogeneous, autonomous, conservative formulation (we add the equation $\tilde{\mathbf{w}}_t = \mathbf{0}$, where $\tilde{\mathbf{w}}(x,t) = \mathbf{w}(x)$)

$$\begin{aligned} \tilde{\mathbf{q}}_t + [\tilde{\mathbf{f}}(\tilde{\mathbf{q}})]_x &= \mathbf{0}, \ x \in \mathbf{R}, \ t \in (0, T), \\ \tilde{\mathbf{q}}(x, 0) &= \tilde{\mathbf{q}}_0(x), \ x \in \mathbf{R}, \end{aligned} \tag{9}$$

where $\tilde{\mathbf{q}} = [\mathbf{q}, \tilde{\mathbf{w}}]^T$, $\tilde{\mathbf{f}}(\tilde{\mathbf{q}}) = [\mathbf{f}(\mathbf{q}, \tilde{\mathbf{w}}), \mathbf{0}]^T$ and $\tilde{\mathbf{q}}_0(x) = [\mathbf{q}_0(x), \mathbf{w}(x)]^T$.

Now we consider the system in the form (nonhomogeneous, nonautonomous case)

$$\begin{aligned} \mathbf{q}_t + [\mathbf{f}(\mathbf{q}, \mathbf{w}(x))]_x &= \mathbf{B}(\mathbf{q}, \mathbf{w}(x))\mathbf{w}_x, \ x \in \mathbf{R}, \ t \in (0, T), \\ \mathbf{q}(x, 0) &= \mathbf{q}_0(x), \ x \in \mathbf{R}, \end{aligned} \tag{10}$$

where $\mathbf{B} = \mathbf{B}(\mathbf{q}, \mathbf{w})$ is the matrix function of the type $m \times s$.
In the case of the river flow (4) this augmented formulation has the form

$$\begin{gathered} \mathbf{q} = [a, q]^T, \quad \mathbf{w}(x) = [l(x), b(x)]^T, \\ \mathbf{f}(\mathbf{q}, \mathbf{w}) = [q, \tfrac{q^2}{a} + \tfrac{ga^2}{2l}]^T, \\ \mathbf{B}(\mathbf{q}, \mathbf{w}(x)) = \begin{bmatrix} 0 & 0 \\ \frac{ga^2}{2l^2} & -ga \end{bmatrix}. \end{gathered}$$

We can rewrite the previous system to the augmented, homogeneous, autonomous, quasilinear formulation

$$\begin{aligned} \tilde{\mathbf{q}}_t + \mathbf{C}(\tilde{\mathbf{q}})\tilde{\mathbf{q}}_x &= \mathbf{0}, \ x \in \mathbf{R}, \ t \in (0, T), \\ \tilde{\mathbf{q}}(x, 0) &= \tilde{\mathbf{q}}_0(x), \ x \in \mathbf{R}, \end{aligned} \tag{11}$$

where

$$\mathbf{C}(\tilde{\mathbf{q}}) = \begin{bmatrix} \mathbf{f}_\mathbf{q} & \mathbf{f}_\mathbf{w} - \mathbf{B}(\mathbf{q}, \mathbf{w}) \\ \mathbf{0} & \mathbf{0} \end{bmatrix},$$

and the following relation holds $\mathbf{f}_x = \mathbf{f}_\mathbf{q}\mathbf{q}_x + \mathbf{f}_\mathbf{w}\mathbf{w}_x$.

The next extension can be done by adding another equation in the form

$$[\mathbf{f}(\mathbf{q})]_t + \mathbf{f}_\mathbf{q}[\mathbf{f}(\mathbf{q}, \mathbf{w}(x))]_x - \mathbf{f}_\mathbf{q}\mathbf{B}(\mathbf{q}, \mathbf{w}(x))\mathbf{w}_x = \mathbf{0}.$$

The previous relations provide some theoretical insight into how the model behaves.

The overdetermined system has the form

$$\begin{aligned} \hat{\mathbf{q}}_t + \hat{\mathbf{D}}(\hat{\mathbf{q}})\hat{\mathbf{q}}_x &= \mathbf{0},\ x \in \mathbf{R},\ t \in (0,T), \\ \hat{\mathbf{q}}(x,0) &= \hat{\mathbf{q}}_0(x),\ x \in \mathbf{R}, \end{aligned} \tag{12}$$

where $\hat{\mathbf{q}} = [\mathbf{q}, \tilde{\mathbf{w}}, \hat{\mathbf{f}}]^T$, $\hat{\mathbf{f}}(\mathbf{q}, \mathbf{w}, t) = \mathbf{f}(\mathbf{q}, \mathbf{w}(x))$,

$$\hat{\mathbf{D}}(\hat{\mathbf{q}}) = \begin{bmatrix} \mathbf{f_q} & \mathbf{f_w} - \mathbf{B}(\mathbf{q}, \tilde{\mathbf{w}}) & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -\mathbf{f_q}\mathbf{B}(\mathbf{q}, \tilde{\mathbf{w}}) & \mathbf{f_q} \end{bmatrix},$$

and $\hat{\mathbf{q}}_0(x) = [\mathbf{q}_0(x), \mathbf{w}(x), \mathbf{f}(\mathbf{q}_0(x), \mathbf{w}(x))]^T$. The advantage of this formulation is in the conversion of the nonhomogeneous systems to the homogeneous one. For our model of the river flow the matrix has the form

$$\hat{\mathbf{D}}(\hat{\mathbf{q}}) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{-q^2}{a^2} + \frac{ga}{l} & \frac{2q}{a} & \frac{-ga^2}{l^2} & ga & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{-ga^2}{2l^2} & ga & 0 & 1 \\ 0 & 0 & \frac{-gqa}{l^2} & 2gq & \frac{-q^2}{a^2} + \frac{ga}{l} & \frac{2q}{a} \end{bmatrix},$$

where $\hat{\mathbf{q}} = [a, q, l, b, q, \frac{q^2}{a} + \frac{ga^2}{2l}]$. The second and fifth row represents equations with the same unknown quantity, so the fifth equation can be rejected.

Therefore we can formulate problem in the form

$$\begin{aligned} \check{\mathbf{q}}_t + \check{\mathbf{D}}(\check{\mathbf{q}})\check{\mathbf{q}}_x &= \mathbf{0},\ x \in \mathbf{R},\ t \in (0,T), \\ \check{\mathbf{q}}(x,0) &= \check{\mathbf{q}}_0(x),\ x \in \mathbf{R}, \end{aligned} \tag{13}$$

where for the model of the river flow the matrix has the form

$$\check{\mathbf{D}}(\check{\mathbf{q}}) = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{-q^2}{a^2} + \frac{ga}{l} & \frac{2q}{a} & \frac{-ga^2}{l^2} & ga & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{-q^2}{a^2} + \frac{ga}{l} & \frac{-gqa}{l^2} & 2gg & \frac{2q}{a} \end{bmatrix}, \tag{14}$$

and $\check{\mathbf{q}} = [a, q, l, b, \frac{q^2}{a} + \frac{ga^2}{2l}]^T$.

The mentioned augmented formulations for general, i.e. nonconservative and semiconservative, systems of the nonlinear partial differential equations can be used for derivation of efficient numerical methods.

3. Finite volume methods

The finite volume methods are suitable for conservation law problems, because the numerical solution is modified only by the intercell fluxes. These methods are based on the integral formulation

$$\int_{x_1}^{x_2} \mathbf{q}(x,t^2)\ dx - \int_{x_1}^{x_2} \mathbf{q}(x,t^1)\ dx + \int_{t^1}^{t^2} \mathbf{f}(\mathbf{q}(x_2,t)) - \int_{t^1}^{t^2} \mathbf{f}(\mathbf{q}(x_1,t)) = \mathbf{0}, \quad \forall (x_1,x_2)\times(t^1,t^2) \subset \mathbf{R}\times(0,T), \tag{15}$$

They use approximations of the integral averages of the unknown functions instead of the approximations of the unknown functions. The consistency of these methods is related to the flux function.

We define the following discretisation

$$x_j = j\Delta x, \quad j \in \mathbf{Z}, \quad \Delta x > 0, \qquad t^n = n\Delta t, \quad n \in \mathbf{N}_0, \quad \Delta t > 0,$$
$$x_{j+1/2} = x_j + \Delta x/2, \qquad t^{n+1/2} = t^n + \Delta t/2.$$

We denote the conserved quantities at time t^n and point x_j: $\mathbf{q}_j^n = \mathbf{q}(x_j,t^n)$ and $\mathbf{q}_j(t) = \mathbf{q}(x_j,t)$ and its approximations: $\mathbf{Q}_j^n = \mathbf{Q}(x_j,t^n) \approx \mathbf{q}_j^n$ and $\mathbf{Q}_j(t) = \mathbf{Q}(x_j,t) \approx \mathbf{q}_j(t)$. The finite volumes mean the sets $(x_{j-1/2},x_{j+1/2})\times(t^n,t^{n+1})$.

We denote the integral averages of the conserved quantities over the finite volume

$$\bar{\mathbf{Q}}_j^n \approx \bar{\mathbf{q}}_j^n = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{q}(x,t^n)\ dx, \tag{16}$$

and the average flux along $x = x_{j+1/2}$

$$\bar{\mathbf{F}}_{j+1/2}^{n+1/2} \approx \bar{\mathbf{f}}_{j+1/2}^{n+1/2} = \frac{1}{\Delta t}\int_{t^n}^{t^{n+1}} \mathbf{f}(\mathbf{q}(x_{j+1/2},t))\ dt. \tag{17}$$

Fully discrete conservative method can be written as relation between approximations of the flux averages and approximations of the integral averages of the conserved quantities

$$\bar{\mathbf{Q}}_j^{n+1} = \bar{\mathbf{Q}}_j^n - \frac{\Delta t}{\Delta x}(\bar{\mathbf{F}}_{j+1/2}^{n+1/2} - \bar{\mathbf{F}}_{j-1/2}^{n+1/2}). \tag{18}$$

Sometimes it is useful to consider the discretisation in two steps. First step is discretisation only in the space (here the finite volume means intervals $(x_{j-1/2},x_{j+1/2})$)

$$\bar{\mathbf{Q}}_j = \bar{\mathbf{Q}}_j(t) \approx \bar{\mathbf{q}}_j = \bar{\mathbf{q}}_j(t) = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}} \mathbf{q}(x,t)\ dx. \tag{19}$$

This leads to the system of the ordinary differential equations in the time

$$\frac{d}{dt}\bar{\mathbf{Q}}_j = -\frac{1}{\Delta x}[\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}]. \tag{20}$$

4. Steady states

The steady states mean that the unknown quantities do not change in the time, i.e. $\mathbf{q_t} = \mathbf{0}$ and the flux function must balance the right hand side, i.e. $[\mathbf{f}(\mathbf{q})]_x = \boldsymbol{\psi}(\mathbf{q}, x)$.

Some schemes are constructed to preserve some special steady states like so called rest at lake, i.e. there is no motion and the free surface height is constant:

$$q(x,t) = 0, \quad h(x,t) + b(x) = \text{const.} \tag{21}$$

For our model (4) this special steady state has the following form

$$q(x,t) = 0, \ \left(\frac{q^2}{a} + \frac{ga^2}{2l}\right)_x - \frac{ga^2}{2l^2} l_x + gab_x = 0. \tag{22}$$

The second relation in (22) can be rewritten into the form (under the assumption $a = hl$)

$$ghl(h+b)_x = 0.$$

The discrete analogue to this smooth steady state can have the form

- semidiscrete approach and piecewise constant reconstruction

$$\bar{Q}_j = 0, \ \bar{H}_j + \bar{B}_j = \text{const.} \ \forall j \in \mathbf{Z}$$

- general reconstruction

$$Q^+_{j-1/2} = Q^-_{j+1/2} = 0, \ H^+_{j+1/2} + B^+_{j+1/2} = H^-_{j+1/2} + B^-_{j+1/2} = \text{const.} \ \forall j \in \mathbf{Z},$$

  where $(\cdot)^{\pm}_{j\pm 1/2}$ are left and right values at the central points obtained by polynomial reconstruction from values in adjacent cells.

For general steady states the following equalities hold

$$q_x = 0, \ (\frac{q^2}{a} + \frac{ga^2}{2l})_x = \frac{ga^2}{2l^2} l_x - gab_x. \tag{23}$$

The left term in the second equality we can rewrite as

$$\left(\frac{q^2}{a} + \frac{ga^2}{2l}\right)_x = \left(-u^2 + \frac{ga}{l}\right) a_x - \frac{ga^2}{2l^2} l_x, \tag{24}$$

and together we have

$$\left(-u^2 + \frac{ga}{l}\right) a_x = \frac{ga^2}{l} l_x - gab_x. \tag{25}$$

From (25) we obtain the following relation for general steady states (the Bernoulli equation)

$$\left(\frac{1}{2}u^2 + gb + \frac{ga}{l}\right)_x = 0. \tag{26}$$

For numerical methods it is important to choose such approximation which conserved these steady states. The equation (26) means that the term $\frac{1}{2}u^2 + gb + \frac{ga}{l}$ is constant for differentiable steady states. Therefore following property has to be satisfied

$$\left(\frac{1}{2}u^2 + gb + \frac{ga}{l}\right)_j = \left(\frac{1}{2}u^2 + gb + \frac{ga}{l}\right)_{j+1}. \tag{27}$$

We consider the piecewise constant reconstruction. We rearrange (27) and we can express the discrete relation analogous to the smooth one $\phi_x = \left(-u^2 + \frac{ga}{l}\right)a_x - \frac{ga^2}{2l^2}l_x$

$$\Delta\Phi = \left(-|U_L U_R| + g\frac{\bar{A}\bar{L}}{L_L L_R}\right)\Delta A - \frac{g}{2}\frac{\tilde{A}^2}{L_L L_R}\Delta L, \tag{28}$$

where $(\cdot)_L = (\bar{\cdot})_j$, $(\cdot)_R = (\bar{\cdot})_{j+1}$, $\Delta(\cdot) = (\cdot)_R - (\cdot)_L$, $\bar{L} = (L_L + L_R)/2$, $\bar{A} = (A_L + A_R)/2$, $\tilde{A}^2 = (A_L^2 + A_R^2)/2$. The details can be found in [1].

The discrete analogue of the equation (27) can be written as

$$\frac{1}{2}U_L^2 + gB_L + \frac{gA_L}{L_L} = \frac{1}{2}U_R^2 + gB_R + \frac{gA_R}{L_R}. \tag{29}$$

This equality we multiply by $\bar{A}$ and after algebraic manipulation we get

$$\left(-\bar{U}^2 + g\frac{\bar{A}}{\bar{L}}L_L L_R\right)\Delta A = -\bar{A}g\Delta B + \frac{g\bar{A}^2}{L_L L_R}\Delta L. \tag{30}$$

Using (30) together with (28) we get

$$\Delta\Phi = \frac{-|U_L U_R| + g\frac{\bar{A}\bar{L}}{L_L L_R}}{-\bar{U}^2 + g\frac{\bar{A}}{\bar{L}}L_L L_R}\left(-\bar{A}g\Delta B + \frac{g\bar{A}^2}{L_L L_R}\Delta L\right) - \frac{g}{2}\frac{\tilde{A}^2}{L_L L_R}\Delta L, \tag{31}$$

This approach can be viewed as a construction of the approximation of the generalized Rankine–Hugoniot condition for nonhomogeneous problem (4).

In the case of general reconstruction the following condition has to be satisfied

$$\frac{1}{2}(U_{j+1/2}^-)^2 + gB_{j+1/2}^- + \frac{gA_{j+1/2}^-}{L_{j+1/2}^-} = \frac{1}{2}(U_{j+1/2}^+)^2 + gB_{j+1/2}^+ + \frac{gA_{j+1/2}^+}{L_{j+1/2}^+},$$

$$\frac{1}{2}(U_{j+1/2}^-)^2 + gB_{j+1/2}^- + \frac{gA_{j+1/2}^-}{L_{j+1/2}^-} = \frac{1}{2}(U_{j-1/2}^+)^2 + gB_{j-1/2}^+ + \frac{gA_{j-1/2}^+}{L_{j-1/2}^+},$$

$$Q^-_{j+1/2} = Q^+_{j+1/2}, \quad Q^-_{j+1/2} = Q^+_{j-1/2}.$$

For the augmented systems steady state means that, for formulation (12), $\hat{\mathbf{D}}(\hat{\mathbf{q}})\hat{\mathbf{q}}_x = \mathbf{0}$. It follows that the steady state is defined by the vector $\hat{\mathbf{q}}$ satisfying either

$$\hat{\mathbf{q}}_x = \mathbf{0}$$

or

$$\hat{\mathbf{q}}_x = \sum_{p=1}^{k} \alpha_p \mathbf{r}_p \ \forall \alpha_p \in \mathbf{R},$$

where $\mathbf{r}_p$, $p = 1, \dots, k$ are eigenvectors corresponding with eigenvalues $\lambda_p = 0$.

5. Central methods

The central methods are universal schemes for solving hyperbolic partial differential equations. In these schemes there is not necessary to construct the characteristic decomposition of the flux f nor to compute the approximation of the Jacobian matrix. These schemes are Riemann problem free. They are robust but they are characterized by large numerical diffusion.

One simple example is the first-order Lax-Friedrichs scheme

$$\bar{\mathbf{Q}}^{n+1}_j = \frac{1}{2}(\bar{\mathbf{Q}}^n_{j-1} + \bar{\mathbf{Q}}^n_{j+1}) - \frac{\Delta t}{2\Delta x}[\mathbf{f}(\bar{\mathbf{Q}}^n_{j+1}) - \mathbf{f}(\bar{\mathbf{Q}}^n_{j-1})], \tag{32}$$

where the flux function for the conservative form can be written in the form

$$\mathbf{F}^{n+1/2}_{j+1/2} = \frac{1}{2}[\mathbf{f}(\bar{\mathbf{Q}}^n_j) + \mathbf{f}(\bar{\mathbf{Q}}^n_{j+1})] - \frac{\Delta x}{2\Delta t}(\bar{\mathbf{Q}}^n_{j+1} - \bar{\mathbf{Q}}^n_j). \tag{33}$$

In the case of the model describing fluid flow through the constant rectangular channel

$$\begin{aligned} h_t + q_x &= 0, \\ q_t + \left(\tfrac{q^2}{h} + \tfrac{1}{2}gh^2\right)_x &= -ghb_x, \end{aligned}$$

we substitute $y = h + b$ and then we can write

$$\begin{aligned} y_t + q_x &= 0, \\ q_t + \left(\tfrac{q^2}{y-b} + \tfrac{1}{2}g(y-b)^2\right)_x &= -g(y-b)b_x. \end{aligned} \tag{34}$$

The special steady state "rest at lake" means $y(x,t) = \text{const}$ and $q(x,t) = 0$.

Discrete approximations of the flux function and the right hand side are in the form

$$\begin{array}{rcl}
F^{n,1}_{j+1/2} & = & \frac{1}{2}(\bar{Q}^n_j + \bar{Q}^n_{j+1}) - \frac{\Delta x}{2\Delta t}(\bar{Y}^n_{j+1} - \bar{Y}^n_j), \\
F^{n,2}_{j+1/2} & = & \frac{1}{2}\left[\frac{(\bar{Q}^n_j)^2}{\bar{Y}^n_j - \bar{B}_j} + \frac{(\bar{Q}^n_{j+1})^2}{\bar{Y}^n_{j+1} - \bar{B}_{j+1}} + \frac{1}{2}g(\bar{Y}^n_j - \bar{B}_j) + \right. \\
 & + & \left. \frac{1}{2}g(\bar{Y}^n_{j+1} - \bar{B}_{j+1})\right] - \frac{\Delta x}{2\Delta t}(\bar{Q}^n_{j+1} - \bar{Q}^n_j), \\
S^{1,n}_j & = & 0, \\
S^{n,2}_j & = & -\frac{g}{4\Delta x}(\bar{B}_{j+1} - \bar{B}_j) \\
 & & (\bar{Y}^n_{j+1} - \bar{B}_{j+1} + \bar{Y}^n_j - \bar{B}_j + \bar{Y}^n_j - \bar{B}_j + \bar{Y}^n_{j-1} - \bar{B}_{j-1}).
\end{array}$$

This scheme preserves only special steady state "rest at lake". But in general these methods are not suitable for computation steady states [8]. One of their big disadvantages is the relatively large numerical dissipation.

The next type of the central method is for example the Rusanov scheme in semidiscrete form

$$\begin{array}{c}
\frac{d}{dt}\bar{\mathbf{Q}}_j = -\frac{1}{2\Delta x}[\mathbf{f}(\bar{\mathbf{Q}}_{j+1}) - \mathbf{f}(\bar{\mathbf{Q}}_{j-1})] + \frac{1}{2\Delta x}[\hat{a}_{j+1/2}(\bar{\mathbf{Q}}_{j+1} - \bar{\mathbf{Q}}_j) - \\
-\hat{a}_{j-1/2}(\bar{\mathbf{Q}}_j - \bar{\mathbf{Q}}_{j-1})],
\end{array} \tag{35}$$

where

$$\hat{a}_{j+1/2} = \max_p\{\max\{\lambda^p_j, \lambda^p_{j+1}\}\}.$$

This scheme can be written in the conservative form (20) where the numerical fluxes have the form

$$\mathbf{F}_{j+1/2} = \frac{1}{2}[\mathbf{f}(\bar{\mathbf{Q}}_j) + \mathbf{f}(\bar{\mathbf{Q}}_{j+1})] - \frac{1}{2}|\hat{a}_{j+1/2}|(\bar{\mathbf{Q}}_{j+1} - \bar{\mathbf{Q}}_j).$$

And as will be mentioned in the next section, this scheme can be rewritten in the fluctuation form.

The Rusanov scheme applied to our model (34) has the following form

$$\begin{array}{rcl}
\frac{d}{dt}\bar{\mathbf{Q}}_j & = & -\frac{1}{\Delta x}(\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}) + \mathbf{S}_j, \\
F^1_{j+1/2} & = & \frac{1}{2}(\bar{Q}_j + \bar{Q}_{j+1}) - \frac{1}{2}\hat{a}_{j+1/2}(\bar{Y}_{j+1} - \bar{Y}_j), \\
F^2_{j+1/2} & = & \frac{1}{2}\left[\frac{\bar{Q}^2_j}{\bar{Y}_j - \bar{B}_j} + \frac{\bar{Q}^2_{j+1}}{\bar{Y}_{j+1} - \bar{B}_{j+1}} + \frac{g}{2}(\bar{Y}_{j+1} + \bar{Y}_j - \bar{B}_{j+1} - \bar{B}_j)^2\right] - \\
 & - & \frac{1}{2}\hat{a}_{j+1/2}(\bar{Q}_{j+1} - \bar{Q}_j) \\
S^1_j & = & 0 \\
S^2_j & = & \frac{-g}{4\Delta x}(\bar{B}_{j+1} - \bar{B}_{j-1}) \\
 & & (\bar{Y}_{j+1} + \bar{Y}_j - \bar{B}_{j+1} - \bar{B}_j + \bar{Y}_j + \bar{Y}_{j_1} - \bar{B}_j - \bar{B}_{j-1})
\end{array} \tag{36}$$

The suitable choice of the CFL condition preserves positive semidefiniteness of some components of the solution. As in the previous case this scheme can preserve only special steady state "rest at lake".

6. Upwind methods

6.1. Scalar case

In this subsection we consider the equation

$$\begin{aligned} q_t + aq_x &= 0,\ x \in \mathbf{R},\ t \in (0,T),\ a \in \mathbf{R}, \\ q(x,0) &= q_0(x),\ x \in \mathbf{R}. \end{aligned} \tag{37}$$

This advection equation has known solution $q(x,t) = q_0(x - at)$. Usually the REA algorithm (reconstruct-evolve-average) is used to solve this problem. This algorithm is based on the piecewise polynomial reconstruction of the solution from the values $\bar{Q}_j(t)$. This reconstruction we denote $\hat{Q}_j(x,t)$ for $x \in (x_{j-1/2}, x_{j+1/2})$ and it is considered to be the initially condition for solving sets of the Riemann problems (in this case we can use the form of the solution).

The semidiscrete scheme (20) has the numerical flux in the form

$$F_{j+1/2} = \frac{1}{2}a(Q^-_{j+1/2} + Q^+_{j+1/2}) - \frac{1}{2}|a|(Q^+_{j+1/2} - Q^-_{j+1/2}), \tag{38}$$

where $Q^+_{j+1/2} = \hat{Q}_{j+1}(x_{(j+1/2)^+}, t)$, $Q^-_{j+1/2} = \hat{Q}_j(x_{(j+1/2)^-}, t)$ are one–sided limits. This scheme can be rewritten into so called fluctuation form

$$\frac{d\bar{Q}_j}{dt} = \frac{-1}{\Delta x}(a^-\Delta Q_{j+1/2} + a\Delta Q_j + a^+\Delta Q_{j-1/2}), \tag{39}$$

where fluctuations are defined

$$\begin{aligned} a\Delta Q_j &= a(Q^-_{j+1/2} - Q^+_{j-1/2}), \\ a^-\Delta Q_{j+1/2} &= a^-(Q^+_{j+1/2} - Q^-_{j+1/2}), \\ a^+\Delta Q_{j-1/2} &= a^+(Q^+_{j-1/2} - Q^-_{j-1/2}), \end{aligned}$$

where $a^+ = \max\{a, 0\}$, $a^- = \min\{a, 0\}$.

For simple piecewise constant reconstruction $Q^+_{j+1/2} = \bar{Q}_{j+1}$, $Q^-_{j+1/2} = \bar{Q}_j$ we obtain for $a > 0$

$$\frac{d}{dt}\bar{Q}_j = -\frac{a}{\Delta x}(\bar{Q}_j - \bar{Q}_{j-1}), \tag{40}$$

and for $a < 0$

$$\frac{d}{dt}\bar{Q}_j = -\frac{a}{\Delta x}(\bar{Q}_{j+1} - \bar{Q}_j). \tag{41}$$

6.2. Linear systems

Now we consider a linear system

$$\begin{aligned} \mathbf{q}_t + \mathbf{A}\mathbf{q}_x &= \mathbf{0},\ x \in \mathbf{R},\ t \in (0,T), \\ \mathbf{q}(x,0) &= \mathbf{q}_0(x),\ x \in \mathbf{R}, \end{aligned} \tag{42}$$

where $\mathbf{A}$ is a real matrix $m \times m$. We suppose that the matrix $\mathbf{A}$ has distinct real eigenvalues and is diagonalisable, i.e. there exists regular matrix $\mathbf{R}$ such that $\mathbf{\Lambda} = \mathbf{R}^{-1}\mathbf{A}\mathbf{R}$, where $\mathbf{\Lambda}$ is diagonal matrix. Thus we can rewrite (42) to the form

$$\boldsymbol{\gamma}_t + \mathbf{\Lambda}\boldsymbol{\gamma}_x = \mathbf{0}, \tag{43}$$

where $\boldsymbol{\gamma}(x,t) = \mathbf{R}^{-1}\mathbf{q}(x,t)$. The system (43) represents m advection equations which can be solved analogous to the scalar case.

After rewriting the system (42) to the conservation form, where $\mathbf{f}(\mathbf{q}) = \mathbf{A}\mathbf{q}$, and solving sets of the generalized Riemann problems we get the numerical fluxes in the form

$$\mathbf{F}_{j+1/2} = \frac{1}{2}\mathbf{A}(\mathbf{Q}^-_{j+1/2} + \mathbf{Q}^+_{j+1/2}) - \frac{1}{2}|\mathbf{A}|(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^-_{j-1/2}), \tag{44}$$

where $\mathbf{Q}^+_{j+1/2} = \hat{\mathbf{Q}}_{j+1}(x_{(j+1/2)+}, t)$, $\mathbf{Q}^-_{j+1/2} = \hat{\mathbf{Q}}_j(x_{(j+1/2)-}, t)$, $|\mathbf{A}| = \mathbf{R}|\mathbf{\Lambda}|\mathbf{R}^{-1}$ and $|\mathbf{\Lambda}| = \mathrm{diag}(|\lambda^p|)$. In analogy to the previous section we can rewrite the conservative scheme to the fluctuation form

$$\frac{d\bar{\mathbf{Q}}_j}{dt} = \frac{-1}{\Delta x}(\mathbf{A}^-\Delta\mathbf{Q}_{j+1/2} + \mathbf{A}\Delta\mathbf{Q}_j + \mathbf{A}^+\Delta\mathbf{Q}_{j-1/2}), \tag{45}$$

where

$$\begin{aligned}
\mathbf{A}\Delta\mathbf{Q}_j &= \mathbf{A}(\mathbf{Q}^-_{j+1/2} - \mathbf{Q}^+_{j-1/2}),\\
\mathbf{A}^-\Delta\mathbf{Q}_{j+1/2} &= \sum_{p=1}^{m} \lambda^{-,p}\Delta\gamma^p_{j+1/2}\mathbf{r}^p,\\
\mathbf{A}^+\Delta\mathbf{Q}_{j-1/2} &= \sum_{p=1}^{m} \lambda^{+,p}\Delta\gamma^p_{j-1/2}\mathbf{r}^p,\\
\Delta\mathbf{Q}_{j+1/2} &= \sum_{p=1}^{m} \Delta\gamma^p_{j+1/2}\mathbf{r}^p,\\
\Delta\mathbf{Q}_j &= \mathbf{Q}^-_{j+1/2} - \mathbf{Q}^+_{j-1/2},
\end{aligned}$$

$\mathbf{A}^+ = \mathbf{R}\mathbf{\Lambda}^+\mathbf{R}^{-1}$, $\mathbf{A}^- = \mathbf{R}\mathbf{\Lambda}^-\mathbf{R}^{-1}$, $\mathbf{\Lambda}^+ = \mathrm{diag}(\max\{\lambda^p, 0\})$, $\mathbf{\Lambda}^- = \mathrm{diag}(\min\{\lambda^p, 0\})$, $\Delta\boldsymbol{\gamma}_{j+1/2} = \mathbf{R}^{-1}\Delta\mathbf{Q}_{j+1/2}$.

6.3. Nonlinear systems

Now we consider nonlinear system

$$\begin{aligned}
\mathbf{q}_t + [\mathbf{f}(\mathbf{q})]_x &= \mathbf{0},\ x \in \mathbf{R},\ t \in (0,T),\\
\mathbf{q}(x,0) &= \mathbf{q}_0(x),\ x \in \mathbf{R},
\end{aligned} \tag{46}$$

The fluctuation form of the conservative scheme is as follows

$$\frac{d\bar{\mathbf{Q}}_j}{dt} = \frac{-1}{\Delta x}[\mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) + \mathbf{A}(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) + \mathbf{A}^+(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2})], \tag{47}$$

$$\begin{array}{rcl}\mathbf{A}(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2}) & = & \mathbf{f}(\mathbf{Q}^-_{j+1/2})-\mathbf{f}(\mathbf{Q}^+_{j-1/2}),\\ \mathbf{A}^-(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2}) & = & \mathbf{F}^-_{j+1/2}-\mathbf{f}(\mathbf{Q}^-_{j+1/2}),\\ \mathbf{A}^+(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2}) & = & \mathbf{f}(\mathbf{Q}^+_{j-1/2})-\mathbf{F}^+_{j-1/2}.\end{array} \tag{48}$$

This scheme can be rewritten to the form

$$\frac{d}{dt}\bar{\mathbf{Q}}_j = -\frac{1}{\Delta x}[\mathbf{F}^-_{j+1/2}-\mathbf{F}^+_{j-1/2}]. \tag{49}$$

If the fluctuations have the following property

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2})-\mathbf{f}(\mathbf{Q}^-_{j+1/2}) = \mathbf{A}^+(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2})+\mathbf{A}^-(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2}), \tag{50}$$

then $\mathbf{F}^-_{j+1/2}=\mathbf{F}^+_{j+1/2}$ $\forall j\in\mathbf{Z}$, i.e. the scheme is conservative.

It is difficult to solve nonlinear Riemann problems to take exact solution. It is efficient to use some approximate Riemann solvers such as HLL or Roe's solvers.

6.3.1. Roe's solver

This approximate Riemann solver is based on the approximation of the nonlinear system $\mathbf{q}_t+[\mathbf{f}(\mathbf{q})]_x\equiv\mathbf{q}_t+\mathbf{A}(\mathbf{q})\mathbf{q}_x=0$, where $\mathbf{A}(\mathbf{q})$ is the Jacobian matrix, by the linear system $\mathbf{q}_t+\mathbf{A}_{j+1/2}\mathbf{q}_x=0$, where $\mathbf{A}_{j+1/2}$ is the Roe-averaged Jacobian matrix, which is defined by suitable combination of $\mathbf{A}(\mathbf{Q}_j)$ and $\mathbf{A}(\mathbf{Q}_{j+1})$.

We define intercell numerical fluxes

$$\mathbf{F}_{j+1/2}=\frac{1}{2}[\mathbf{f}(\mathbf{Q}^-_{j+1/2})+\mathbf{f}(\mathbf{Q}^+_{j+1/2})]-\frac{1}{2}|\mathbf{A}_{j+1/2}|(\mathbf{Q}^+_{j+1/2}-\mathbf{Q}^-_{j+1/2}), \tag{51}$$

and intercell fluctuations in the scheme (47) by

$$\begin{array}{rcl}\mathbf{A}^-(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2}) & = & \sum\limits_{p=1}^{m}\lambda^{-,p}_{j+1/2}\mathbf{r}^p_{j+1/2}\Delta\gamma^p_{j+1/2},\\ \mathbf{A}^+(\mathbf{Q}^-_{j+1/2},\mathbf{Q}^+_{j+1/2}) & = & \sum\limits_{p=1}^{m}\lambda^{+,p}_{j+1/2}\mathbf{r}^p_{j+1/2}\Delta\gamma^p_{j+1/2},\end{array} \tag{52}$$

where $\mathbf{r}^p_{j+1/2}$ are eigenvectors of the Roe matrix $\mathbf{A}_{j+1/2}$, $\lambda^p_{j+1/2}$ are eigenvalues called Roe's speeds and $\Delta\boldsymbol{\gamma}_{j+1/2}=\mathbf{R}^{-1}_{j+1/2}\Delta\mathbf{Q}_{j+1/2}$, $\Delta\mathbf{Q}_{j+1/2}=\mathbf{Q}^+_{j+1/2}-\mathbf{Q}^-_{j+1/2}$.

6.3.2. HLL solver

This solver does not use the explicit linearization of the Jacobian matrix, but the solution is constructed by the consideration of two discontinuities, propagating at speeds s^1 and s^2. The middle state $\mathbf{Q}^*_{j+1/2}$ is determined by conservation law

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2})-\mathbf{f}(\mathbf{Q}^-_{j+1/2})=s^2_{j+1/2}(\mathbf{Q}^+_{j+1/2}-\mathbf{Q}^*_{j+1/2})+s^1_{j+1/2}(\mathbf{Q}^*_{j+1/2}-\mathbf{Q}^-_{j+1/2}), \tag{53}$$

$$\mathbf{Q}^*_{j+1/2}=\frac{\mathbf{f}(\mathbf{Q}^+_{j+1/2})-\mathbf{f}(\mathbf{Q}^-_{j+1/2})-s^2_{j+1/2}\mathbf{Q}^+_{j+1/2}+s^1_{j+1/2}\mathbf{Q}^-_{j+1/2}}{s^1_{j+1/2}-s^2_{j+1/2}}. \tag{54}$$

When the special choice of the characteristic speeds called Einfeldt speeds is used, the solver is called HLLE. The Einfeldt speeds are defined by

$$s^1_{j+1/2} = \min_p \{\min\{\lambda^p_j, \lambda^p_{j+1/2}, 0\}\}, \qquad s^2_{j+1/2} = \max_p \{\max\{\lambda^p_{j+1}, \lambda^p_{j+1/2}, 0\}\}, \tag{55}$$

where λ^p_j are eigenvalues of the matrix $\mathbf{A}_j = \mathbf{f}'(\mathbf{Q}^-_{j+1/2})$.

6.4. Wave–fluctuation scheme for augmented systems

Consider the model for river flow through the varying rectangular channel (4) as was presented in 1. Section and its augmented formulation (13) and (14) presented in 2. Section. The eigencomponents for the matrix $\check{\mathbf{D}}$ are

$$\lambda_1 = 0,\ \lambda_2 = 0,\ \lambda_3 = 2u,\ \lambda_4 = u + \sqrt{\frac{ga}{l}},\ \lambda_5 = u - \sqrt{\frac{ga}{l}},$$

and

$$\begin{aligned} \mathbf{r}_1 &= [-\tfrac{ga}{\lambda_4\lambda_5}, 0, 0, -1, ga]^T, & \mathbf{r}_2 &= [-\tfrac{ga^2}{l^2\lambda_4\lambda_5}, 0, 1, 0, \tfrac{ga^2}{2l^2}]^T, \\ \mathbf{r}_3 &= [0, 0, 0, 0, 1]^T, & \mathbf{r}_4 &= [1, \lambda_4, 0, 0, \lambda_4^2]^T, \\ \mathbf{r}_5 &= [1, \lambda_5, 0, 0, \lambda_5^2]^T. \end{aligned}$$

We realize the decomposition for the augmented quasilinear formulation i.e. for the system of five equations with Einfeldt speeds

$$s_1 = 0, \quad s_2 = 0, \quad s_3 = s_4 + s_5,$$

$$s_4 = \min_p \{\min\{\lambda^p_L, \lambda^p_{LR}\}\}, \qquad s_5 = \max_p \{\max\{\lambda^p_R, \lambda^p_{LR}\}\},$$

and approximation of the eigenvectors of the matrix $\check{\mathbf{D}}$

$$\begin{aligned} \mathbf{r}_1 &\approx [\tfrac{g\bar{A}}{\widetilde{s_4 s_5}}, 0, 0, -1, \tfrac{g\bar{A}\widehat{s_4 s_5}}{\widetilde{s_4 s_5}}]^T, \\ \mathbf{r}_2 &\approx [\tfrac{g\tilde{A}^2}{L_L L_R \widetilde{s_4 s_5}}, 0, 1, 0, \tfrac{g\tilde{A}^2 \widehat{s_4 s_5}}{L_L L_R \widetilde{s_4 s_5}} - \tfrac{g\tilde{A}^2}{2L_L L_R}]^T, \\ \mathbf{r}_3 &\approx [0, 0, 0, 0, 1]^T, \\ \mathbf{r}_4 &\approx [1, s_4, 0, 0, s_4^2]^T, \\ \mathbf{r}_5 &\approx [1, s_5, 0, 0, s_5^2]^T, \end{aligned}$$

where $\widetilde{s_4 s_5} = -\bar{U}^2 + \frac{g\bar{A}\bar{L}}{L_L L_R}$, $\widehat{s_4 s_5} = -|U_L U_R| + \frac{g\bar{A}\bar{L}}{L_L L_R}$, λ^p_L and λ^p_R are eigenvalues of the Jacobi matrix for the left end right values and λ^p_{LR} are eigenvalues of the Roe's matrix.

The decomposition of the augmented system has the following form

$$\begin{bmatrix} \Delta A \\ \Delta Q \\ \Delta L \\ \Delta B \\ \Delta \Phi \end{bmatrix} = \sum_{p=1}^{5} \gamma_p \mathbf{r}^p.$$

We have five linearly independent eigenvectors. The approximation is chosen to be able to prove the consistency and provide the stability of the algorithm. In some special cases this scheme is conservative and we can obtain the positive semidefiniteness, but only under the additional assumptions.

The basic version of the numerical scheme for piecewise constant reconstruction is in the form

$$\frac{d\bar{\mathbf{Q}}_j}{dt} = -\frac{1}{\Delta x}[\mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) + \mathbf{A}^+(\mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2})], \tag{56}$$

where fluctuations are defined by

$$\begin{aligned} \mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1 \\ s^{p,n}_{j+1/2}<0}}^{m} \gamma^p_{j+1/2}\mathbf{r}^p_{j+1/2}, \\ \mathbf{A}^+(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1 \\ s^{p,n}_{j+1/2}>0}}^{m} \gamma^p_{j+1/2}\mathbf{r}^p_{j+1/2}. \end{aligned}$$

7. Central-upwind method

Now we introduce so called central-upwind schemes. These schemes combine advantages of the upwind schemes i.e. lower numerical diffusion and usability for the steady states with advantages of the central schemes i.e. positive semidefiniteness. These schemes are Riemann solver free.

One simple method in the conservative form (20) has the numerical flux in the form

$$\mathbf{F}_{j+1/2} = \frac{a^+_{j+1/2}\mathbf{f}(\mathbf{Q}_j) - a^-_{j+1/2}\mathbf{f}(\mathbf{Q}_{j+1})}{a^+_{j+1/2} - a^-_{j+1/2}} + \frac{a^+_{j+1/2}a^-_{j+1/2}}{a^+_{j+1/2} - a^-_{j+1/2}}\left[\mathbf{Q}_{j+1} - \mathbf{Q}_j\right], \tag{57}$$

where

$$\begin{aligned} a^+_{j+1/2} &= \max\left\{\lambda_N\left(\mathbf{f}'(\mathbf{Q}_j)\right), \lambda_N\left(\mathbf{f}'(\mathbf{Q}_{j+1})\right), 0\right\}, \\ a^-_{j+1/2} &= \min\left\{\lambda_1\left(\mathbf{f}'(\mathbf{Q}_j)\right), \lambda_1\left(\mathbf{f}'(\mathbf{Q}_{j+1})\right), 0\right\}, \end{aligned}$$

represent maximal speeds of the propagation of the waves at the points $x_{j+1/2}$ and we suppose that $\lambda_1 < \lambda_2, \ \ldots, \ \lambda_N$.

For system (34) this scheme has the following form

$$\frac{d\bar{\mathbf{Q}}_j}{dt} = -\frac{1}{\Delta x}(\mathbf{F}_{j+1/2} - \mathbf{F}_{j-1/2}) + \mathbf{S}_j. \tag{58}$$

Numerical fluxes have the form

$$\begin{aligned} \mathbf{F}_{j+1/2} &= \frac{a^+_{j+1/2}\mathbf{f}(\mathbf{Q}^-_{j+1/2}) - a^-_{j+1/2}\mathbf{f}(\mathbf{Q}^+_{j+1/2})}{a^+_{j+1/2} - a^-_{j+1/2}} \\ &+ \frac{a^+_{j+1/2}a^-_{j+1/2}}{a^+_{j+1/2} - a^-_{j+1/2}}\left[\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^-_{j+1/2}\right], \end{aligned} \tag{59}$$

where $\mathbf{Q}^-_{j+1/2}$ and $\mathbf{Q}^+_{j+1/2}$ are left and right approximations of the unknown function at the points $x_{j+1/2}$, $a^+_{j+1/2}$ and $a^-_{j+1/2}$ represent maximal wave propagation speeds at points $x_{j+1/2}$.

$$\begin{aligned} a^+_{j+1/2} &= \max\left\{\lambda_N\left(\mathbf{f}'(\mathbf{Q}^-_{j+1/2})\right), \lambda_N\left(\mathbf{f}'(\mathbf{Q}^+_{j+1/2})\right), 0\right\}, \\ a^-_{j+1/2} &= \min\left\{\lambda_1\left(\mathbf{f}'(\mathbf{Q}^-_{j+1/2})\right), \lambda_1\left(\mathbf{f}'(\mathbf{Q}^+_{j+1/2})\right), 0\right\}, \end{aligned} \tag{60}$$

where $\lambda_1 < \lambda_2, \ \ldots, \ \lambda_N$. The source term is discretized by the following way

$$\begin{aligned} S^1_j &= 0 \\ S^2_j &= -g\tfrac{B_{j+1/2}-B_{j-1/2}}{\Delta x}\cdot \\ &\quad\cdot\tfrac{(Y^-_{j+1/2}+Y^+_{j-1/2}-B_{j+1/2}-B_{j-1/2})}{2}. \end{aligned} \tag{61}$$

These schemes preserve only steady state "rest at lake" such as central methods.

8. Decomposition of the flux function

All described schemes can be represented and understood by the same way. The amount of data about the structure of the solution of the Riemann problem included into schemes causes the differences between schemes. This information is employed in decomposition of the difference of the flux function.

Central schemes, for example Lax-Friedrichs scheme, are based on the following decomposition

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - \mathbf{f}(\mathbf{Q}^-_{j+1/2}) = s(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}) - s(\mathbf{Q}^*_{j+1/2} - \mathbf{Q}^-_{j+1/2}) = \sum_{p=1}^{2}\mathbf{Z}^p_{j+1/2}, \tag{62}$$

where $s = \frac{\Delta x}{\Delta t}$ and

$$\begin{aligned} \mathbf{Z}^2_{j+1/2} &= s(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}), \\ \mathbf{Z}^1_{j+1/2} &= -s(\mathbf{Q}^*_{j+1/2} - \mathbf{Q}^-_{j+1/2}). \end{aligned} \tag{63}$$

The only used information is upper estimation of the maximum speed of the discontinuities propagation. This estimate is used in the CFL condition. These methods are global because no information about discontinuities propagation in the neighbourhood of the point is used. From (62) we can express $\mathbf{Q}^*_{j+1/2}$ and we get

$$\mathbf{Q}^*_{j+1/2} = \frac{\Delta t}{2\Delta x}[\mathbf{f}(\mathbf{Q}^-_{j+1/2}) - \mathbf{f}(\mathbf{Q}^+_{j+1/2})] + \frac{1}{2}(\mathbf{Q}^-_{j+1/2} + \mathbf{Q}^+_{j+1/2}). \tag{64}$$

Next we define fluctuations

$$\begin{aligned}
\mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1 \\ s^p<0}}^{2} \mathbf{Z}^p_{j+1/2}, \\
\mathbf{A}^+(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1 \\ s^p>0}}^{2} \mathbf{Z}^p_{j+1/2}.
\end{aligned} \tag{65}$$

For conservative form of the scheme we have to evaluate $\mathbf{F}_{j+1/2} = \mathbf{f}(\mathbf{Q}^*_{j+1/2})$. We use the Rankine–Hugoniot jump condition in the form

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = s(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}) \tag{66}$$

and together with (64) we get

$$\mathbf{F}_{j+1/2} = \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = \frac{1}{2}[\mathbf{f}(\mathbf{Q}^-_{j+1/2} + \mathbf{f}(\mathbf{Q}^+_{j+1/2}] - \frac{\Delta t}{2\Delta x}(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^-_{j+1/2}). \tag{67}$$

This scheme can be understood by another way. We can derive them from fully discrete form (18) where we use the following partition of the x–axis

$$\langle x_{j+1/2,L}, x_{j+1/2,R}\rangle,$$

where $x_{j+1/2,L} = x_{j+1/2} - s\Delta t$, $x_{j+1/2,R} = x_{j+1/2} + s\Delta t$ and $\langle x_{j+1/2,L}, x_{j+1/2,R}\rangle = \langle x_{j-1/2}, x_{j+3/2}\rangle$ hold. On each of these intervals we use the balance in integral form (15). The points where the solution is discontinuous lie inside these intervals and this scheme is Riemann problem free.

We can use the relation (47) and we can derive the scheme in the conservative form. These schemes are not suitable for the semidiscrete formulation because of the infinite speed ($\Delta t \to 0$) of the propagating discontinuities which is typical for the parabolic type of the equations.

The semidiscrete central schemes use estimate of the upper bound of maximal local speed of the propagating discontinuities. They are based on the following decomposition

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2})-\mathbf{f}(\mathbf{Q}^-_{j+1/2}) = s_{j+1/2}(\mathbf{Q}^+_{j+1/2}-\mathbf{Q}^*_{j+1/2})-s_{j+1/2}(\mathbf{Q}^*_{j+1/2}-\mathbf{Q}^-_{j+1/2}) = \sum_{p=1}^{2} \mathbf{Z}^p_{j+1/2}, \tag{68}$$

where

$$s_{j+1/2} = \max_p\{\max\{|\lambda^p(\mathbf{Q}^-_{j+1/2})|, |\lambda^p(\mathbf{Q}^+_{j+1/2})|\}\},$$

and

$$\begin{aligned} \mathbf{Z}^2_{j+1/2} &= s_{j+1/2}(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}), \\ \mathbf{Z}^1_{j+1/2} &= -s_{j+1/2}(\mathbf{Q}^*_{j+1/2} - \mathbf{Q}^-_{j+1/2}). \end{aligned} \tag{69}$$

We can express

$$\mathbf{Q}^*_{j+1/2} = \frac{1}{2s_{j+1/2}}[\mathbf{f}(\mathbf{Q}^-_{j+1/2}) - \mathbf{f}(\mathbf{Q}^+_{j+1/2})] + \frac{1}{2}(\mathbf{Q}^-_{j+1/2} + \mathbf{Q}^+_{j+1/2}), \tag{70}$$

and we define

$$\begin{aligned} \mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1 \\ s^p_{j+1/2}<0}}^{2} \mathbf{Z}^p_{j+1/2}, \\ \mathbf{A}^+(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1 \\ s^p_{j+1/2}>0}}^{2} \mathbf{Z}^p_{j+1/2}. \end{aligned} \tag{71}$$

For evaluating $\mathbf{F}_{j+1/2} = \mathbf{f}(\mathbf{Q}^*_{j+1/2})$ we use the Rankine–Hugoniot jump condition in the form

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = s_{j+1/2}(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}), \tag{72}$$

and together with (70) we get

$$\mathbf{F}_{j+1/2} = \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = \frac{1}{2}[\mathbf{f}(\mathbf{Q}^-_{j+1/2} + \mathbf{f}(\mathbf{Q}^+_{j+1/2}] - \frac{1}{2}s_{j+1/2}(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^-_{j+1/2}). \tag{73}$$

This scheme we can derive from fully discrete form (18) where the x-axis is partitioned to subintervals of the following types

$$\langle x_{j-1/2,R}, x_{j+1/2,L}\rangle \quad \text{and} \quad \langle x_{j+1/2,L}, x_{j+1/2,R}\rangle,$$

where $x_{j+1/2,L} = x_{j+1/2}-s_{j+1/2}\Delta t$, $x_{j+1/2,R} = x_{j+1/2}+s_{j+1/2}\Delta t$. On these intervals we use the integral balance law (15). The points where the solution is discontinuous lie inside these intervals and as in the previous case, this method is Riemann solver free.

The central-upwind methods can be identified with HLL solver. The decomposition has the form

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2})-\mathbf{f}(\mathbf{Q}^-_{j+1/2}) = s^2_{j+1/2}(\mathbf{Q}^+_{j+1/2}-\mathbf{Q}^*_{j+1/2})-s^1_{j+1/2}(\mathbf{Q}^*_{j+1/2}-\mathbf{Q}^-_{j+1/2}) = \sum_{p=1}^{2}\mathbf{Z}^p_{j+1/2}, \tag{74}$$

where $s^1_{j+1/2} = a^-_{j+1/2}$, $s^2_{j+1/2} = a^+_{j+1/2}$ and

$$\begin{array}{rcl} \mathbf{Z}^2_{j+1/2} & = & s^2_{j+1/2}(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}), \\ \mathbf{Z}^1_{j+1/2} & = & -s^1_{j+1/2}(\mathbf{Q}^*_{j+1/2} - \mathbf{Q}^-_{j+1/2}). \end{array} \tag{75}$$

As in the previous cases we can express $\mathbf{Q}^*_{j+1/2}$

$$\mathbf{Q}^*_{j+1/2} = \frac{\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - \mathbf{f}(\mathbf{Q}^-_{j+1/2})}{s^1_{j+1/2} - s^2_{j+1/2}} + \frac{s^1_{j+1/2}\mathbf{Q}^-_{j+1/2} - s^2_{j+1/2}\mathbf{Q}^+_{j+1/2}}{s^1_{j+1/2} - s^2_{j+1/2}}. \tag{76}$$

We define

$$\begin{array}{rcl} \mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) & = & \sum\limits_{\substack{p=1 \\ s^p_{j+1/2}<0}}^{2} \mathbf{Z}^p_{j+1/2}, \\ \mathbf{A}^+(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) & = & \sum\limits_{\substack{p=1 \\ s^p_{j+1/2}>0}}^{2} \mathbf{Z}^p_{j+1/2}. \end{array} \tag{77}$$

Relation (76) with Rankine–Hugoniot jump condition in the form

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = s^2_{j+1/2}(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^*_{j+1/2}) \tag{78}$$

give us the following

$$\mathbf{F}_{j+1/2} = \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = \frac{s^1_{j+1/2}\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - s^2_{j+1/2}\mathbf{f}(\mathbf{Q}^-_{j+1/2})}{s^1_{j+1/2} - s^2_{j+1/2}} + \frac{s^1_{j+1/2}s^2_{j+1/2}}{s^1_{j+1/2} - s^2_{j+1/2}}(\mathbf{Q}^-_{j+1/2}-\mathbf{Q}^+_{j+1/2}). \tag{79}$$

We can derive these schemes from fully discrete method (18) by limiting process ($\Delta t \to 0$). The x–axis is partitioned to subintervals of following types

$$\langle x_{j-1/2,R}, x_{j+1/2,L}\rangle \quad \text{and} \quad \langle x_{j+1/2,L}, x_{j+1/2,R}\rangle,$$

where $x_{j+1/2,L} = x_{j+1/2} - s^1_{j+1/2}\Delta t$, $x_{j+1/2,R} = x_{j+1/2} + s^2_{j+1/2}\Delta t$. In analogy to the previous cases we formulate the integral balance law (15) on each of defined intervals. The solution is discontinuous at the points lying inside of these intervals and no Riemann problem we need to solve.

The previous schemes contain only one middle state $\mathbf{Q}^*_{j+1/2}$ between states $\mathbf{Q}^-_{j+1/2}$ and $\mathbf{Q}^+_{j+1/2}$. It is possible derive schemes with two or more middle states. For example, the Roe solver is based on the decomposition with $(m-1)$ middle states

$$\mathbf{f}(\mathbf{Q}^+_{j+1/2}) - \mathbf{f}(\mathbf{Q}^-_{j+1/2}) = \sum_{p=1}^{m} s^p_{j+1/2}\mathbf{W}^p_{j+1/2}, \tag{80}$$

where $s^p_{j+1/2} = \lambda^p_{j+1/2}$ are eigenvalues and $\mathbf{r}^p_{j+1/2}$ are eigenvectors of the Roe matrix (see paragraph 6.3.1.), $s^1_{j+1/2} < s^2_{j+1/2} < \cdots < s^m_{j+1/2}$, $\mathbf{W}^p_{j+1/2} = \gamma^p_{j+1/2}\mathbf{r}^p_{j+1/2}$ and $\gamma^p_{j+1/2} = \mathbf{R}^{-1}_{j+1/2}\Delta\mathbf{Q}_{j+1/2}$.

The middle states can be express in the following form

$$\mathbf{Q}^{p,*}_{j+1/2} = \mathbf{Q}^-_{j+1/2} + \sum_{k=1}^{p}\mathbf{W}^k_{j+1/2}. \tag{81}$$

Next we define

$$\mathbf{Z}^p_{j+1/2} = s^p_{j+1/2}\mathbf{W}^p_{j+1/2} \tag{82}$$

and than the following holds

$$\begin{aligned}\mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1\\ s^p_{j+1/2}<0}}^{m} \mathbf{Z}^p_{j+1/2},\\ \mathbf{A}^+(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}) &= \sum_{\substack{p=1\\ s^p_{j+1/2}>0}}^{m} \mathbf{Z}^p_{j+1/2}.\end{aligned} \tag{83}$$

From (48) and from the conservativity we get the following results

$$\begin{aligned}\mathbf{F}_{j+1/2} &= \mathbf{f}(\mathbf{Q}^-_{j+1/2}) + \mathbf{A}^-(\mathbf{Q}^-_{j+1/2}, \mathbf{Q}^+_{j+1/2}),\\ \mathbf{F}_{j-1/2} &= \mathbf{f}(\mathbf{Q}^+_{j-1/2}) - \mathbf{A}^+(\mathbf{Q}^-_{j-1/2}, \mathbf{Q}^+_{j-1/2}).\end{aligned} \tag{84}$$

The numerical flux function can be express in the form

$$\mathbf{F}_{j+1/2} = \mathbf{f}(\mathbf{Q}^*_{j+1/2}) = \frac{1}{2}[f(\mathbf{Q}^-_{j+1/2}) + f(\mathbf{Q}^+_{j+1/2})] - \frac{1}{2}|\mathbf{A}_{j+1/2}|(\mathbf{Q}^+_{j+1/2} - \mathbf{Q}^-_{j+1/2}) \tag{85}$$

This scheme can be derived in the same way as the previous ones. We define the partition of the x–axis

$$\langle x_{j-1/2,m}, x_{j+1/2,1}\rangle, \langle x_{j+1/2,1}, x_{j+1/2,2}\rangle, \ldots, \langle x_{j+1/2,m-1}, x_{j+1/2,m}\rangle,$$

where $x_{j+1/2,p} = x_{j+1/2} + s^p_{j+1/2}\Delta t$. The speeds $s^p_{j+1/2}$ was getting from linearized problem and it cannot be said that the discontinuities lie inside of the intervals. It is not possible to interprete this scheme as a scheme without Riemann solver.

9. Conclusion

We presented various numerical schemes for solving fluid flow problems with various properties. All described schemes can be understood in the same way. We note here that in the case of the methods the source term is not included in the decompositions and if we include them the central and central-upwind methods do not preserve general steady states anyway.

References

[1] D. L. George: *Finite Volume Methods and Adaptive Refinement for Tsunami Propagation and Innundation*, University of Washington, Ph.D. Thesis, 2006

[2] A. Kurganov and G. Petrova: *A Second-Order Well-Balanced Positivity Preserving Central-Upwind Scheme for the Saint-Senant System*, Communications in Mathematical Sciences, 5(2007), pp. 133–160

[3] R. J. LeVeque and M. Pelanti: *A Class of Approximate Riemann Solvers and their Relation to Relaxation Schemes*, Journal of Computational Physics, 172(2001),pp. 572–591

[4] P. L. Roe: *Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes*, Journal of Computational Physics,135(1997), pp. 250–258,

[5] R. J. LeVeque: *Simplified Multi-Dimensional Flux Limiter Methods*, Numerical Methods for Fluid Dynamics,4(1993), pp. 175–190

[6] A. Kurganov and G. Petrova: *Central Schemes and Contact Discontinuities*, Mathematical Modelling and Numerical Analysis, 34(2000), pp. 1259–1275

[7] R. J. LeVeque: *Finite volume methods for hyperbolic problems*, Cambridge University Press, 2002

[8] A. Kurganov and E. Tadmor: *New High-resolution central schemes for nonlinear conservation laws and convection-diffusion equations*, Journal of Computational Physics, 160(2000), pp. 241–282