TU München
Zentrum Mathematik
Lehrstuhl M16
E. Sonnendrücker
E. Franck

Numerical methods for hyperbolic systems

Exercise sheet 4: Nonlinear scalar equations

Exercice 1

Firstly we consider the Burgers equation on the non-conservative form

$$\partial_t u + u\partial_x u = 0, \quad \forall x \in \mathbb{R}, \quad t > 0, \tag{1}$$

We propose to approximate (1) with the finite volumes scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a_j^n}{\Delta x}(u_j^n - u_{j-1}^n) = 0, \tag{2}$$

where the discrete velocity is given by $a_j^n = u_j^n$, $a_j^n = u_{j-1}^n$ or $a_j^n = \frac{u_j^n + u_{j-1}^n}{2}$.
1. Discussing the conservativity of the scheme for the different discrete velocities.

A scheme is conservative if $\sum_j u_j^{n+1} = \sum_j u_j^n$ which is equivalent to $\sum_j a_j^n(u_j^n - u_{j-1}^n) = 0$. If $a_j^n = u_j^n$ we have

$$\sum_j = u_j^n(u_j^n - u_{j-1}^n) = u_1^{2,n} - u_0^n u_1^n + u_2^{2,n} - u_1^n u_2^n + ... + u_N^{2,n} - u_{N-1}^n u_N^n.$$

This sum is not necessary equal to zero if the solution is not constant, consequently the scheme is not conservative. The result is the same for $a_j^n = u_{j-1}^n$. For $a_j^n = \frac{u_j^n + u_{j-1}^n}{2}$ the sum can be rewritten on the following form

$$\sum_j = u_j^n(u_j^n - u_{j-1}^n) = \frac{1}{2}\sum_j u_j^{2,n} - u_{j-1}^{2,n} = u_1^n - u_N^n.$$

Since the boundary conditions are periodic $u_1 = u_N$ the sum is equal to zero. Consequently the scheme is conservative.

Now we consider a nonlinear scalar equation

$$\begin{cases} \partial_t u + \partial_x f(u) = 0, \quad \forall x \in \mathbb{R}, \quad t > 0, \\ u(t=0,x) = u^0(x), \end{cases} \tag{3}$$

and the following scheme

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{f_{j+\frac{1}{2}}^n - f_{j-\frac{1}{2}}^n}{\Delta x} = 0, \tag{4}$$

with $f^n_{j+\frac{1}{2}} = \frac{1}{2}(f(u^n_{j+1}) + f(u^n_j)) + \frac{c}{2}(u^n_j - u^n_{j+1})$, $m = \min_x u_0(x)$, $M = \max_x u_0(x)$ and $\max_{m \le x \le M} |f^{'}(x)| \le c$.

2. Prove that the scheme satisfy the maximum principle under a CFL condition.

Using the definition of the fluxes the scheme can be rewrite on the following form

$$u^{n+1}_j = u^n_j + \frac{c\Delta t}{2\Delta x}(u^n_{j+1} - 2u^n_j + u^n_{j-1}) - \frac{\Delta t}{2\Delta x}(f(u^n_{j+1}) - f(u^n_{j-1})),$$

equivalent to

$$u^{n+1}_j = u^n_j + \frac{c\Delta t}{2\Delta x}(u^n_{j+1} - 2u^n_j + u^n_{j-1}) - \frac{\Delta t}{2\Delta x}a^n_j(u^n_{j+1} - u^n_{j-1}),$$

with $a^n_j = \dfrac{f(u^n_{j+1}) - f(u^n_{j-1})}{(u^n_{j+1} - u^n_{j-1})}$.

We remark that $a^n_j = f^{'}(z^n_j)$ with $\min(u^n_{j-1}, u^n_{j+1}) \le z^n_j \le \max(u^n_{j-1}, u^n_{j+1})$ (Mean Value theorem corollary of Rolle's theorem).

Now we rewrite the scheme to obtain

$$u^{n+1}_j = \left(1 - \frac{c\Delta t}{\Delta x}\right) u^n_j + \frac{\Delta t}{2\Delta x}\left(c - a^n_j\right) u^n_{j+1} + \frac{\Delta t}{2\Delta x}\left(c + a^n_j\right) u^n_{j-1}.$$

The sum of the coefficients associates to u^n_j, u^n_{j-1}, and u^n_{j+1} is equal to one. By definition of c the coefficient $c - a^n_j > 0$ consequently all the coefficients are positive on the CFL condition $\frac{c\Delta t}{\Delta x} \le 1$. We obtain a convex combination thus the scheme preserve the maximum principle.

Additional questions

Now we propose to prove that the scheme is entropic which correspond to satisfy

$$\frac{\eta(u^{n+1}_j) - \eta(u^n_j)}{\Delta t} + \frac{\xi^n_{j+\frac{1}{2}} - \xi^n_{j-\frac{1}{2}}}{\Delta x} \le 0,$$

with $(\eta(u), \xi(u))$ a couple entropy-entropic flux and $\xi^n_{j+\frac{1}{2}}$ the numerical entropic flux

$$\xi^n_{j+\frac{1}{2}} = \frac{\xi(u^n_{j+1}) + \xi(u^n_j)}{2} + \frac{c}{2}(\eta(u^n_j) - \eta(u^n_{j+1})).$$

3. Prove that

$$\frac{\eta(u^{n+1}_j) - \eta(u^n_j)}{\Delta t} + \frac{\xi^n_{j+\frac{1}{2}} - \xi^n_{j-\frac{1}{2}}}{\Delta x} \le \frac{1}{2}(\phi(u^n_{j+1}) + \psi(u^n_{j-1})),$$

with

$\phi(z) = \nu\left(u^n_j + \frac{\Delta t}{\Delta x}c(z - u^n_j) - \frac{\Delta t}{\Delta x}(f(z) - f(u^n_j))\right) - \eta(u^n_j) - \frac{\Delta t}{\Delta x}c(\eta(z) - \eta(u^n_j)) + \frac{\Delta t}{\Delta x}(\xi(z) - \xi(u^n_j))$,

and

$\psi(z) = \nu\left(u_j^n + \frac{\Delta t}{\Delta x}c(-u_j^n + z) - \frac{\Delta t}{\Delta x}(f(u_j^n) - f(z))\right) - \eta(u_j^n) - \frac{\Delta t}{\Delta x}c(-\eta(u_j^n) + \eta(z)) + \frac{\Delta t}{\Delta x}(\xi(u_j^n) - \xi(z))$.

We write the scheme of the following form

$$u_j^{n+1} = \frac{1}{2}\left(u_j^n + \frac{c\Delta t}{\Delta x}(-u_j^n + u_{j-1}^n) - \frac{c\Delta t}{\Delta x}(f(u_j^n) - f(u_{j-1}^n))\right)$$
$$+\frac{1}{2}\left(u_j^n + \frac{c\Delta t}{\Delta x}(u_{j+1}^n + u_j^n) - \frac{c\Delta t}{\Delta x}(f(u_{j+1}^n) - f(u_j^n))\right).$$

Since η is a convex function we obtain

$$\eta(u_j^{n+1}) \leq \frac{1}{2}\eta\left(u_j^n + \frac{c\Delta t}{\Delta x}(-u_j^n + u_{j-1}^n) - \frac{c\Delta t}{\Delta x}(f(u_j^n) - f(u_{j-1}^n))\right)$$
$$+\frac{1}{2}\eta\left(u_j^n + \frac{c\Delta t}{\Delta x}(u_{j+1}^n - u_j^n) - \frac{c\Delta t}{\Delta x}(f(u_{j+1}^n) - f(u_j^n))\right),$$

$$\eta(u_j^{n+1}) - \eta(u_j^n) \leq \frac{1}{2}\eta(u_j^n) + \frac{1}{2}\eta\left(u_j^n + \frac{c\Delta t}{\Delta x}(-u_j^n + u_{j-1}^n) - \frac{c\Delta t}{\Delta x}(f(u_j^n) - f(u_{j-1}^n))\right).$$
$$+\frac{1}{2}\eta(u_j^n) + \frac{1}{2}\eta\left(u_j^n + \frac{c\Delta t}{\Delta x}(u_{j+1}^n - u_j^n) - \frac{c\Delta t}{\Delta x}(f(u_{j+1}^n) - f(u_j^n))\right).$$

By definition of $\xi_{j+\frac{1}{2}}$, ϕ and ψ

$$\frac{\Delta t}{\Delta x}(\xi_{j+\frac{1}{2}} - \xi_{j-\frac{1}{2}}) = \frac{1}{2}(\phi(u_{j+1}^n) + \psi(u_{j-1}^n)) + \eta(u_j^n)$$
$$-\frac{1}{2}\eta\left(u_j^n + \frac{c\Delta t}{\Delta x}(-u_j^n + u_{j-1}^n) - \frac{c\Delta t}{\Delta x}(f(u_j^n) - f(u_{j-1}^n))\right)$$
$$-\frac{1}{2}\eta\left(u_j^n + \frac{c\Delta t}{\Delta x}(u_{j+1}^n - u_j^n) - \frac{c\Delta t}{\Delta x}(f(u_{j+1}^n) - f(u_j^n))\right),$$

Consequently

$$\frac{\eta(u_j^{n+1}) - \eta(u_j^n)}{\Delta t} + \frac{\xi_{j+\frac{1}{2}}^n - \xi_{j-\frac{1}{2}}^n}{\Delta x} \leq \frac{1}{2}(\phi(u_{j+1}^n) + \psi(u_{j-1}^n)). \tag{5}$$

4. Prove that $\psi(z) \leq 0$, $\phi(z) \leq 0$ under the CFL $\frac{c\Delta t}{2\Delta x} < 1$ and conclude.

We consider $\phi^{'}(w) = \nu(c - f^{'}(w))\left(\eta^{'}(u_j^n + \nu c(w - u_j^n) - \nu(f(w) - f(u_j^n))) - \eta^{'}(w)\right)$ with $\nu = \frac{\Delta t}{\Delta x}$.

By definition of c the term $k_1 = \nu(c - f^{'}(w))$ is positive. We obtain

$$\phi^{'}(w) = k_1\left(\eta^{'}(u_j^n + \nu c(w - u_j^n) - \nu(f(w) - f(u_j^n))) - \eta^{'}(w)\right). \tag{6}$$

If a function f is convex we have $(f^{'}(y) - f^{'}(x))(y - x) > 0$. Therefore we can define $k_2 > 0$ with $f^{'}(y) - f^{'}(x) = k_2(y - x)$. Consequently when we apply this property for (6) we obtain

$$\phi^{'}(w) = k_1 k_2\left(u_j^n + \nu c(w - u_j^n) - \nu(f(w) - f(u_j^n))) - w\right). \tag{7}$$

The equation (7) is equivalent

$$\phi^{'}(w) = k_1 k_2\left(1 - \nu c + \nu\frac{(f(u_j^n) - f(w))}{u_j^n - w}\right)(u_j^n - w). \tag{8}$$

Using the Mean value theorem we obtain that $\frac{f(u_j^n)-f(w)}{u_j^n-w} = f^{'}(z)$ with $z \in \left[w, u_j^n\right]$ with $|f^{'}(z)| \leq c$. Under the CFL $\frac{c\Delta t}{2\Delta x} < 1$ the term

$$\left(1 - \nu c + \nu\frac{(f(u_j^n) - f(w))}{u_j^n - w}\right) \geq 0.$$

Consequently $\phi^{'}(w) = K(w)(u_j^n - w)$ with $K(w) > 0$. Indeed ϕ is convex consequently

$$\phi(u_j^n) > \phi(w) + \phi^{'}(w)(u_j^n - w).$$

Since $\phi(u_j^n) = 0$ and $\phi^{'}(w) = K(w)(u_j^n - w)$ then $\phi(w) \leq 0$.

Now we study the second term

$$\psi^{'}(w) = \nu(c + f^{'}(w))\left(\eta^{'}(u_j^n + \nu c(w - u_j^n) - \nu(f(u_j^n) - f(w))) - \eta^{'}(w)\right).$$

Using the same arguments that ϕ we obtain $\psi^{'}(w) = K(w)(u_j^n - w)$ with $K(w) > 0$. Since ψ is convex and $\psi(u_j^n) = 0$ we obtain $\psi(w) \leq 0$.

To conclude $\psi(w) \leq 0$, $\phi(w) \leq 0$ and (5) imply that the scheme is entropic.

Additional exercise We consider a linear hyperbolic system with stiff nonlinear source term.

$$\begin{cases} \partial_t p + \partial_x u = 0, \\ \partial_t u + a\partial_x p = \dfrac{1}{\varepsilon}(f(p) - u), \end{cases} \tag{9}$$

with $\sqrt{a} \geq |f^{'}(p)|$.

1. Formally prove that when ε tends to zero, the system (9) tends to $\partial_t p + \partial_x f(p) = 0$.

Idea : Try to obtain $\partial_t u + \partial_x f(u) = \varepsilon \partial_x \left[(a - f^{'}(p)^2)\partial_x p \right] + o(\varepsilon^2)$.

The second equation of (9) gives

$$u = f(p) - \varepsilon(\partial_t u + a\partial_x p). \tag{10}$$

Taking the derivate of (10) we obtain $\partial_t u = \partial_t f(p) - \varepsilon(\partial_{tt} u + a\partial_{t,x} p)$. Now we plug the last relation in (10) to obtain

$$\begin{aligned} u &= f(p) - \varepsilon(\partial_t f(p) + a\partial_x p) + O(\varepsilon^2), \\ &= f(p) - \varepsilon(f^{'}(p)\partial_t p + a\partial_x p) + O(\varepsilon^2), \\ &= f(p) - \varepsilon(-f^{'}(p)\partial_t v + a\partial_x p) + O(\varepsilon^2). \end{aligned}$$

The last relation is obtained using the first equation on (10). Now we use $u = f(p) + O(\varepsilon)$ to obtain

$$\begin{aligned} u &= f(p) - \varepsilon(-f^{'}(p)\partial_t f(p) + a\partial_x p) + O(\varepsilon^2), \\ &= f(p) - \varepsilon((a - f^{'}(p)^2)\partial_x p) + O(\varepsilon^2). \end{aligned}$$

Consequently we obtain

$$\partial_t p + \partial_x f(p) = \varepsilon \partial_x ((a - f^{'}(p)^2)\partial_x p) \tag{11}$$

This result show that the system (9) tends to $\partial_t p + \partial_x f(p)$ when ε tend to zero. The condition $\sqrt{a} \geq |f^{'}(p)|$ is a condition to obtain a dissipative equation.

2. We propose the splitting scheme (12)-(13)

$$\begin{cases} \dfrac{p_j^{n+\frac{1}{2}} - p_j^n}{\Delta t} = 0, \\ \\ \dfrac{u_j^{n+\frac{1}{2}} - u_j^n}{\Delta t} = \dfrac{1}{\varepsilon}(f(p_j^n) - u_j^n). \end{cases} \tag{12}$$

$$\begin{cases} \dfrac{p_j^{n+1} - p_j^{n+\frac{1}{2}}}{\Delta t} + \dfrac{u_{j+1}^{n+\frac{1}{2}} - u_{j-1}^{n+\frac{1}{2}}}{2\Delta x} - \dfrac{\sqrt{a}\Delta x}{2} \dfrac{p_{j+1}^{n+\frac{1}{2}} - 2p_j^{n+\frac{1}{2}} + p_{j-1}^{n+\frac{1}{2}}}{\Delta x^2} = 0, \\ \\ \dfrac{u_j^{n+1} - u_j^{n+\frac{1}{2}}}{\Delta t} + \dfrac{p_{j+1}^{n+\frac{1}{2}} - p_{j-1}^{n+\frac{1}{2}}}{2\Delta x} - \dfrac{\sqrt{a}\Delta x}{2} \dfrac{u_{j+1}^{n+\frac{1}{2}} - 2u_j^{n+\frac{1}{2}} + u_{j-1}^{n+\frac{1}{2}}}{\Delta x^2} = 0. \end{cases} \tag{13}$$

Assuming that $u_j^0 = f(p_j^0) + O(\varepsilon)$ (the initial data are close to the equilibrium). Explain why this scheme is not adapted to treat the system (9) with big time step.

We assume that this equality $u_j^0 = f(p_j^0) + O(\varepsilon)$ is propagated in time consequently $u_j^n = f(p_j^n) + O(\varepsilon)$. Now we propose to study the step $n+1$. Using (12) we have

$$\begin{aligned} u_j^{n+\frac{1}{2}} &= u_j^n + \tfrac{\Delta t}{\varepsilon} f(p_j^n) - \tfrac{\Delta t}{\varepsilon} u_j^n, \\ &= (1 - \tfrac{\Delta t}{\varepsilon}) u_j^n + f(p_j^n) + (1 - \tfrac{\Delta t}{\varepsilon}) f(p_j^n), \\ &= f(p_j^n) + 2(1 - \tfrac{\Delta t}{\varepsilon}) f(p_j^n) + O(\varepsilon + \Delta t). \end{aligned}$$

The last relation come from to $u_j^n = f(p_j^n) + O(\varepsilon)$. Plugging the last relation in the first equation of (13) we obtain

$$\frac{p_j^{n+1} - p_j^{n+\frac{1}{2}}}{\Delta t} + \frac{f(p_{j+1}^n) - f(p_{j-1}^n)}{2\Delta x} - \frac{\sqrt{a}\Delta x}{2} \frac{p_{j+1}^n - 2p_j^n + p_{j-1}^n}{\Delta x^2}$$
$$+2\left(1 - \frac{\Delta t}{\varepsilon}\right) \frac{f(p_{j+1}^n) - f(p_{j-1}^n)}{2\Delta x} + O(\varepsilon + \Delta t).$$

For example we choose $\Delta t = \frac{\sqrt{\varepsilon}}{2}$. In this case we obtain

$$\frac{p_j^{n+1} - p_j^{n+\frac{1}{2}}}{\Delta t} + \left(3 - \frac{1}{\sqrt{\varepsilon}}\right) \frac{f(p_{j+1}^n) - f(p_{j-1}^n)}{2\Delta x} - \frac{\sqrt{a}\Delta x}{2} \frac{p_{j+1}^n - 2p_j^n + p_{j-1}^n}{\Delta x^2}.$$

If ε tends to zero the coefficient between the term discretizing $\partial_x f(u)$ is very large. Consequently the limit scheme is not a good approximation of the limit equation $\partial_t u + \partial_x f'(u) = 0$ when ε tends to zero. We obtain a good approximation if $\Delta t << \varepsilon$.

3. Propose a modification of the previous scheme to obtain a better accuracy for big time step and justify your modification.

To solve the problem we propose the following scheme

$$\begin{cases} \dfrac{p_j^{n+\frac{1}{2}} - p_j^n}{\Delta t} = 0, \\[2ex] \dfrac{u_j^{n+\frac{1}{2}} - u_j^n}{\Delta t} = \dfrac{1}{\varepsilon}(f(p_j^{n+\frac{1}{2}}) - u_j^{n+\frac{1}{2}}). \end{cases} \tag{14}$$

$$\begin{cases} \dfrac{p_j^{n+1} - p_j^{n+\frac{1}{2}}}{\Delta t} + \dfrac{u_{j+1}^{n+\frac{1}{2}} - u_{j-1}^{n+\frac{1}{2}}}{2\Delta x} - \dfrac{\sqrt{a}\Delta x}{2} \dfrac{p_{j+1}^{n+\frac{1}{2}} - 2p_j^{n+\frac{1}{2}} + p_{j-1}^{n+\frac{1}{2}}}{\Delta x^2} = 0, \\[2ex] \dfrac{u_j^{n+1} - u_j^{n+\frac{1}{2}}}{\Delta t} + \dfrac{p_{j+1}^{n+\frac{1}{2}} - p_{j-1}^{n+\frac{1}{2}}}{2\Delta x} - \dfrac{\sqrt{a}\Delta x}{2} \dfrac{u_{j+1}^{n+\frac{1}{2}} - 2u_j^{n+\frac{1}{2}} + u_{j-1}^{n+\frac{1}{2}}}{\Delta x^2} = 0. \end{cases} \tag{15}$$

We assume that this equality $u_j^0 = f(p_j^0) + O(\varepsilon)$ is propagated in time, thus $u_j^n = f(p_j^n) + O(\varepsilon)$. Now we propose to study the step $n+1$. Using (14) we have

$$u_j^{n+\frac{1}{2}} = u_j^n + \frac{\Delta t}{\varepsilon} f(p_j^{n+\frac{1}{2}}) - \frac{\Delta t}{\varepsilon} u_j^{n+\frac{1}{2}},$$

$$\left(1 + \frac{\Delta t}{\varepsilon}\right) u_j^{n+\frac{1}{2}} = u_j^n + \frac{\Delta t}{\varepsilon} f(p_j^n).$$

Simplifying the last estimation we obtain

$$u_j^{n+\frac{1}{2}} = \frac{\varepsilon}{\varepsilon + \Delta t} u_j^n + \frac{\Delta t}{\Delta t + \varepsilon} f(p_j^n).$$

Using $u_j^n = f(p_j^n) + O(\varepsilon)$ we obtain $u_j^{n+\frac{1}{2}} = f(p_j^n) + O(\varepsilon)$. Indeed $(\frac{\varepsilon}{\varepsilon+\Delta t}) = O(\varepsilon)$ because $\frac{\varepsilon}{\varepsilon+\Delta t} \leq 1$.
Plugging $u_j^{n+\frac{1}{2}} = f(p_j^n) + O(\varepsilon)$ in the scheme (15) we obtain

$$\frac{p_j^{n+1} - p_j^{n+\frac{1}{2}}}{\Delta t} + \frac{f(p_{j+1}^n) - f(p_{j-1}^n)}{2\Delta x} - \frac{\sqrt{a}\Delta x}{2} \frac{p_{j+1}^n - 2p_j^n + p_{j-1}^n}{\Delta x^2} + O(\varepsilon).$$

We obtain a good approximation of the limit equation $\partial_t u + \partial_x f(u) = 0$ when ε tends to zero for all values of Δt.