Méthodes numériques pour les EDP instationnaires

Solution TD 3: jeudi 07.10.2021
Transport equation with variable coefficients

Solution

1 Transport equation with variable velocity

Take a function $a \in C^1(\mathbb{R})$ such that there exists $A, A_1, A_2 \in [0, +\infty)$ with $|a(x)| \le A$, $|a'(x)| \le A_1$ and $|a''(x)| \le A_2$ for all $x \in \mathbb{R}$. The nonconservative transport equation is

$$\begin{aligned} \partial_t \bar{u}(x,t) + a(x)\, \partial_x \bar{u}(x,t) &= 0, && \forall (x,t) \in \mathbb{R} \times \mathbb{R}_*^+, \\ \bar{u}(x,0) &= u_0(x), && \forall x \in \mathbb{R}. \end{aligned} \tag{1}$$

The conservative transport equation is

$$\begin{aligned} \partial_t \hat{u}(x,t) + \partial_x (a(x) \hat{u}(x,t)) &= 0, && \forall (x,t) \in \mathbb{R} \times \mathbb{R}_*^+, \\ \hat{u}(x,0) &= u_0(x), && \forall x \in \mathbb{R}. \end{aligned} \tag{2}$$

We assume that $u_0 \in C^2(\mathbb{R})$ with bounded derivatives. We introduce a discretization of the domain using a regular mesh:

$$(x_j, t_n) = (j\Delta x, n\Delta t), \qquad \forall j \in \mathbb{Z},\ \forall n \in \mathbb{N},$$

where Δx, respectively Δt, denotes the space step, respectively the time step.
We denote $a_j = a(x_j)$, $a_j^+ = \max(a_j, 0)$, $a_j^- = \max(-a_j, 0)$. Note the relation $a_j = a_j^+ - a_j^-$.

1.1 Scheme for equation (1)

The scheme is

$$\Delta x \frac{u_j^{n+1} - u_j^n}{\Delta t} + a_j^- (u_j^n - u_{j+1}^n) - a_{j-1}^+ (u_{j-1}^n - u_j^n) = 0\,. \tag{3}$$

1. Define the discrete iteration operator $J_{h,\Delta t}$.

We define

$$\nu_j = \frac{a_j \Delta t}{\Delta x}\,, \qquad \nu = \frac{A\Delta t}{\Delta x}\,.$$

The scheme can be rewritten as

$$\begin{aligned} u_j^{n+1} &= u_j^n - \nu_j^-(u_j^n - u_{j+1}^n) - \nu_{j-1}^+(u_{j-1}^n - u_j^n)\,, \\ &= (1 - \nu_j^- - \nu_{j-1}^+) u_j^n + \nu_j^- u_{j+1}^n + \nu_{j-1}^+ u_{j-1}^n\,, \end{aligned}$$

so the discrete iteration operator is

$$J_{h,\Delta t} = (1 - \nu_j^- - \nu_{j-1}^+) + A_h\,,$$

where

$$(A_h u)_j = \nu_j^- u_{j+1} + \nu_{j-1}^+ u_{j-1}\,.$$

2. Check that under a CFL condition, the scheme satisfies the discrete maximum principle and thus deduce the L^∞ stability of the scheme.

Since

$$u_j^{n+1} = (1 - \nu_j^- - \nu_{j-1}^+) u_j^n + \nu_j^- u_{j+1}^n + \nu_{j-1}^+ u_{j-1}^n\,,$$

the scheme is L^∞-stable if and only if $1 - \nu_j^- - \nu_{j-1}^+ \ge 0$ for all $j \in \mathbb{Z}$. This is in particular the case if $\nu \le \frac{1}{2}$.

3. We assume that the solution $\bar{u}$ of (1) satisfies $\bar{u} \in C^2(\mathbb{R} \times \mathbb{R}_+)$, where the absolute value of all first and second order derivatives of $\bar{u}$ are bounded by some value C_2, and we consider the truncation error

$$T_j^n = \frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t} + a_j^- \frac{\bar{u}_j^n - \bar{u}_{j+1}^n}{\Delta x} - a_{j-1}^+ \frac{\bar{u}_{j-1}^n - \bar{u}_j^n}{\Delta x}\,,$$

where $\bar{u}_j^n = \bar{u}(x_j, t_n)$. Prove that $|T_j^n| \le C_3(\Delta t + \Delta x)$.

We have

$$\begin{aligned} \frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t} &= \partial_t \bar{u}(x_j, t_n) + O(\Delta t)\,, \\ \frac{\bar{u}_j^n - \bar{u}_{j+1}^n}{\Delta x} &= -\partial_x \bar{u}(x_j, t_n) + O(\Delta x)\,, \\ \frac{\bar{u}_{j-1}^n - \bar{u}_j^n}{\Delta x} &= \partial_x \bar{u}(x_j, t_n) + O(\Delta x)\,, \end{aligned}$$

so

$$T_j^n = \partial_t \bar{u}(x_j, t_n) - \left[a(x_j)^- - a(x_{j-1})^+\right] \partial_x \bar{u}(x_j, t_n) + O(\Delta t + \Delta x)\,.$$

One can show that for all $a, b \in \mathbb{R}$,

$$a^+ - |b| \le (a+b)^+ \le a^+ + |b|\,,$$

so

$$\left|(a+b)^+ - a^+\right| \le |b|\,.$$

Since

$$a(x_{j-1}) = a(x_j) + O(\Delta x)\,,$$

using this last inequality, one has

$$\left|a(x_{j-1})^+ - a(x_j)^+\right| \le O(\Delta x)\,.$$

Therefore

$$T_j^n = \partial_t \bar{u}(x_j, t_n) + a(x_j) \partial_x \bar{u}(x_j, t_n) + O(\Delta t + \Delta x) = O(\Delta t + \Delta x)\,,$$

since

$$a(x_j)^- - a(x_{j-1})^+ = -a(x_j) + a(x_j)^+ - a(x_{j-1})^+\,.$$

We have shown that $|T_j^n| \le C_3 (\Delta t + \Delta x)$ where C_3 depends on C_2 and A_1.

4. Prove the convergence in $L^\infty(\mathbb{R})$ for $u_0 \in W^{2,\infty}(\mathbb{R})$.

We have stability and consistence in L^∞, so by Lax theorem, the scheme is convergent in L^∞ for $\nu \le \frac{1}{2}$.

1.2 Scheme for equation (2)

The scheme is

$$\Delta x \frac{u_j^{n+1} - u_j^n}{\Delta t} + (a_j^+ u_j^n - a_j^- u_{j+1}^n) - (a_{j-1}^+ u_{j-1}^n - a_{j-1}^- u_j^n) = 0\,. \tag{4}$$

1. Define the discrete iteration operator $J_{h,\Delta t}$.

We have

$$u_j^{n+1} = (1 - \nu_j^+ - \nu_{j-1}^-) u_j^n + \nu_j^- u_{j+1}^n + \nu_{j-1}^+ u_{j-1}^n\,,$$

so

$$J_{h,\Delta t} = (1 - \nu_j^+ - \nu_{j-1}^-) + A_h\,,$$

where

$$(A_h u)_j = \nu_j^- u_{j+1} + \nu_{j-1}^+ u_{j-1}\,.$$

2. Assume $u_j^0 \geqslant 0$ for all $j \in \mathbb{Z}$. Find a CFL condition such that $u_j^n \geqslant 0$ for all $j \in \mathbb{Z}$ and all $n \in \mathbb{N}$.

Since the scheme is a convex combination when $1 - \nu_j^+ - \nu_{j-1}^- \ge 0$. This is in particular the case when $\nu \le \frac{1}{2}$, so in this case the positivity of the initial data is preserved.

3. Assume $\sum_{j\in\mathbb{Z}} |u_j^0| < \infty$. Under the same CFL condition, prove that $\sum_{j\in\mathbb{Z}} |u_j^{n+1}| < \infty$ for all $n \in \mathbb{N}$.

Under the CFL condition $\nu \le \frac{1}{2}$, one has

$$\begin{aligned} \sum_{j\in\mathbb{Z}} |u_j^{n+1}| &\le \sum_{j\in\mathbb{Z}} (1 - \nu_j^+ - \nu_{j-1}^-)|u_j^n| + \sum_{j\in\mathbb{Z}} \nu_j^- |u_{j+1}^n| + \sum_{j\in\mathbb{Z}} \nu_{j-1}^+ |u_{j-1}^n| \\ &\le \sum_{j\in\mathbb{Z}} (1 - \nu_j^+ - \nu_{j-1}^-)|u_j^n| + \sum_{j\in\mathbb{Z}} \nu_{j-1}^- |u_j^n| + \sum_{j\in\mathbb{Z}} \nu_j^+ |u_j^n| \le \sum_{j\in\mathbb{Z}} |u_j^n|\,, \end{aligned}$$

which proves the claim.

4. Under the previous assumptions, prove that the scheme preserves the discrete mass, *i.e* for a given $n \in \mathbb{N}$,

$$\sum_{j\in\mathbb{Z}} \Delta x u_j^{n+1} = \sum_{j\in\mathbb{Z}} \Delta x u_j^n\,.$$

One has

$$\begin{aligned} \sum_{j\in\mathbb{Z}} u_j^{n+1} &= \sum_{j\in\mathbb{Z}} (1 - \nu_j^+ - \nu_{j-1}^-) u_j^n + \sum_{j\in\mathbb{Z}} \nu_j^- u_{j+1}^n + \sum_{j\in\mathbb{Z}} \nu_{j-1}^+ u_{j-1}^n \\ &= \sum_{j\in\mathbb{Z}} (1 - \nu_j^+ - \nu_{j-1}^-) u_j^n + \sum_{j\in\mathbb{Z}} \nu_{j-1}^- u_j^n + \sum_{j\in\mathbb{Z}} \nu_j^+ u_j^n \\ &= \sum_{j\in\mathbb{Z}} u_j^n\,. \end{aligned}$$

5. Study the L^1 stability of the scheme.

We have shown that $\|J_{h,\Delta t}\|_{\mathcal{L}(L^1)} \le 1 + 4\nu$, so $\|J_{h,\Delta t}^n\|_{\mathcal{L}(L^1)} \le (1+4\nu)^n \le e^{4\nu n}$ and the scheme is L^1-stable. Under the CFL condition $\nu \le \frac{1}{2}$ and for positive initial data, the L^1-norm is preserved by $J_{h,\Delta t}$, so $\|J_{h,\Delta t}\|_{\mathcal{L}(L^1)} = 1$.

6. Prove the convergence in $L^1(\mathbb{R})$ for compactly supported initial data.

The first step is to prove the consistency in L^∞. One has

$$\begin{aligned} \frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t} &= \partial_t \bar{u}(x_j, t_n) + O(\Delta t)\,, \\ a_j^+ \bar{u}_j^n - a_j^- \bar{u}_{j+1}^n &= a_j^+ \bar{u}_j^n - a_j^- \left(\bar{u}_j^n + \partial_x \bar{u}(x_j, t_n)\Delta x + O(\Delta x^2)\right) \\ &= a_j \bar{u}_j^n - a_j^- \partial_x \bar{u}(x_j, t_n)\Delta x + O(\Delta x^2)\,, \\ a_{j-1}^+ \bar{u}_{j-1}^n - a_{j-1}^- \bar{u}_j^n &= a_{j-1}^+ \bar{u}_{j-1}^n - a_{j-1}^- \left(\bar{u}_{j-1}^n + \partial_x \bar{u}(x_j, t_n)\Delta x + O(\Delta x^2)\right) \\ &= a_{j-1} \bar{u}_{j-1}^n - a_{j-1}^- \partial_x \bar{u}(x_j, t_n)\Delta x + O(\Delta x^2)\,, \end{aligned}$$

so

$$\begin{aligned} T_j^n &= \partial_t \bar{u}(x_j, t_n) + \partial_x (a\bar{u})(x_j, t_n) + \left[a_{j-1}^- - a_j^-\right] \partial_x \bar{u}(x_j, t_n) + O(\Delta t + \Delta x) \\ &= O(\Delta t + \Delta x)\,. \end{aligned}$$

We have shown that $|T_j^n| \le C_3 (\Delta t + \Delta x)$ where C_3 depends on C_2, $A_1 = \sup_x |a'(x)|$ and $A_2 = \sup_x |a''(x)|$.
Since u_0 has compact support the exact solution $\bar{u}$ has compact support in space, so

$$\sum_j |T_j^n| \Delta x \le R |T_j^n| \le R C_3 (\Delta t + \Delta x)\,,$$

where R is linked to the support of the solution at time t.
Finally, one has stability and consistency in L^1 so convergence in L^1 for $\nu \le \frac{1}{2}$.

2 Mixed cases

Mixed schemes are schemes which are stable in some norm (typically L^2) but not in L^∞.
The Crank-Nicholson (second order in time) discretization of the heat equation is

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{u_{j+1}^{n+\frac12} - 2u_j^{n+\frac12} + u_{j-1}^{n+\frac12}}{\Delta x^2}\,, \qquad u_j^{n+\frac12} = \frac{1}{2}(u_j^n + u_j^{n+1})\,.$$

1. Show it is unconditionally stable in L^2.

Defining $\nu = \frac{\Delta t}{\Delta t^2}$, one has

$$\begin{aligned} u_j^{n+1} &= u_j^n + \nu \left(u_{j+1}^{n+\frac12} - 2u_j^{n+\frac12} + u_{j-1}^{n+\frac12}\right) \\ &= u_j^n + \frac{\nu}{2}\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right) + \frac{\nu}{2}\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right)\,, \end{aligned}$$

so the scheme can be rewritten as

$$U^{n+1} = U^n + \frac{\nu}{2} A U^n + \frac{\nu}{2} A U^{n+1}\,,$$

where $U^n \in \mathbb{R}^{\mathbb{Z}}$ and $M \in \mathbb{R}^{\mathbb{Z}\times\mathbb{Z}}$ are given by

$$U^n = \begin{pmatrix} \vdots \\ u_{-1}^n \\ u_0^n \\ u_1^n \\ u_2^n \\ \vdots \end{pmatrix}, \qquad A = \mathrm{diag}\,(1, -2, 1) = \begin{pmatrix} \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \ddots & -2 & 1 & 0 & 0 & \ddots \\ \ddots & 1 & -2 & \ddots & 0 & \ddots \\ \ddots & 0 & \ddots & \ddots & 1 & \ddots \\ \ddots & 0 & 0 & 1 & -2 & \ddots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

The symbol of A is

$$\lambda(\theta) = a + b e^{i\theta} + b^{-i\theta} = -2 + 2\cos\theta\,,$$

so the spectrum satisfy $\sigma(A) \subset [-4, 0]$.
The scheme can also be written as

$$\left(1 - \frac{\nu}{2} A\right) U^{n+1} = \left(1 + \frac{\nu}{2} A\right) U^n\,,$$

and since $\left(1 - \frac{\nu}{2} A\right)$ is invertible

$$U^{n+1} = J U^n\,, \qquad \text{where} \qquad J = \left(1 - \frac{\nu}{2} A\right)^{-1} \left(1 + \frac{\nu}{2} A\right).$$

The symbol of J is therefore

$$\frac{1 + \frac{\nu}{2}\lambda(\theta)}{1 - \frac{\nu}{2}\lambda(\theta)}\,,$$

so is bounded by one and $\rho(J) = 1$. This proves that the scheme is L^2-stable for any value of ν. We remark this is the case as soon as $\sigma(A) \subset (-\infty, 0]$.

2. Show it is conditionally unitary stable in L^∞.

One has

$$(1+\nu)\, u_j^{n+1} = (1-\nu)\, u_j^n + \frac{\nu}{2} u_{j+1}^n + \frac{\nu}{2} u_{j-1}^n + \frac{\nu}{2} u_{j+1}^{n+1} + \frac{\nu}{2} u_{j-1}^{n+1}\,,$$

so for $0 \le \nu \le 1$,

$$\begin{aligned} (1+\nu) \sup_j |u_j^{n+1}| &\le (1-\nu) \sup_j |u_j^n| + \frac{\nu}{2} \sup_j |u_{j+1}^n| + \frac{\nu}{2} \sup_j |u_{j-1}^n| + \frac{\nu}{2} \sup_j |u_{j+1}^{n+1}| + \frac{\nu}{2} \sup_j |u_{j-1}^n| \\ &\le \sup_j |u_j^n| + \nu \sup_j |u_j^{n+1}|\,, \end{aligned}$$

and

$$\sup_j |u_j^{n+1}| \le \sup_j |u_j^n|\,.$$

We note that it is possible to prove that

$$\|J_{h,\Delta t}\|_{\mathcal{L}(L^\infty)} = \begin{cases} 1, & 0 < \nu \le \frac{3}{2}, \\ 3 - \frac{4}{\sqrt{1+2\nu}}, & \nu \ge \frac{3}{2}. \end{cases}$$

See for example Faragóa & Palenciab, *Sharpening the estimate of the stability constant in the maximum-norm of the Crank–Nicolson scheme for the one-dimensional heat equation*, Applied Numerical Mathematics, 2002.

3 Transport on nonuniform meshes

Consider a nonuniform mesh with mesh sizes $0 < \alpha h < \Delta x_j \le h$. The Finite Volume scheme for advection $a > 0$ writes

$$\Delta x_j \frac{u_j^{n+1} - u_j^n}{\Delta t} + a u_j^n - a u_{j-1}^n = 0, \qquad j \in \mathbb{Z}, \quad n \in \mathbb{N}\,.$$

1. Write the scheme under the form $\frac{u_h^{n+1} - u_h^n}{\Delta t} = A_h u_h^n$.

One has

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = a \frac{u_{j-1}^n - u_j^n}{\Delta x_j}\,,$$

so

$$\frac{u_h^{n+1} - u_h^n}{\Delta t} = A_h u_h^n\,,$$

where

$$(A_h u)_j = \frac{a}{\Delta x j}\left(u_{j-1} - u_j\right)\,.$$

2. Study the stability in L^∞ of the iteration operator $J_{h,\Delta t} = I_h + \Delta t A_h$.

Denoting $\nu_j = \frac{a\Delta t}{\Delta x_j}$, one has

$$u_j^{n+1} = u_j^n + \frac{a\Delta t}{\Delta x_j} u_{j-1}^n - \frac{a\Delta t}{\Delta x_j} u_j^n = (1 - \nu_j)\, u_j^n + \nu_j u_{j-1}^n\,,$$

so the scheme is L^∞-stable if and only if $1 - \nu_j \ge 0$ for all $j \in \mathbb{Z}$. This is in particular the case if $1 - \frac{a\Delta t}{\alpha h} \ge 0$, *i.e.* $\frac{a\Delta t}{\alpha h} \le 1$.

3. For $x_j = \frac{1}{2}(x_{j+\frac12} + x_{j-\frac12})$ the middle of cell number j, define the interpolation of the exact solution as $v_h^n = (u(x_j, t_n))_{j\in\mathbb{Z}}$. Show this approach does not yield consistency.

We have

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\Delta t} &= \partial_t u(t_n, x_j) + O(\Delta t)\,, \\ \frac{v_j^n - v_{j-1}^n}{\Delta x_j} &= \partial_x u(t_n, x_j) \frac{x_j - x_{j-1}}{\Delta x_j} + O(\Delta x)\,, \end{aligned}$$

so the scheme is consistent if and only if $x_j - x_{j-1} = \Delta x_j$. Since $x_j - x_{j-1} = \frac{\Delta x_j}{2} + \frac{\Delta x_{j-1}}{2}$, this last condition is equivalent to $\Delta x_j = \Delta x_{j-1}$, *i.e.* $\Delta x_j = h$ for all $j \in \mathbb{Z}$. The conclusion is that this scheme is consistent if and only if the grid is uniform.

4. Define another interpolation of the exact solution $w_h^n = \left(u(x_{j+\frac12}, t_n)\right)_{j\in\mathbb{Z}}$. Show consistency and convergence.

Since

$$\begin{aligned} \frac{w_j^{n+1} - w_j^n}{\Delta t} &= \partial_t u(t_n, x_{j+\frac12}) + O(\Delta t)\,, \\ \frac{w_j^n - w_{j-1}^n}{\Delta x_j} &= \frac{u(t_n, x_{j+\frac12}) - u(t_n, x_{j-\frac12})}{\Delta x_j} = \partial_x u(t_n, x_{j+\frac12}) \frac{x_{j+\frac12} - x_{j-\frac12}}{\Delta x_j} + O(\Delta x)\,, \end{aligned}$$

and $x_{j+\frac12} - x_{j-\frac12} = \Delta x_j$, this proves the consistency of this interpolation. Since one has consistency and stability in L^∞, this scheme is convergent under the condition $\frac{a\Delta t}{\alpha h} \le 1$.

5. Show stability in all l^p equipped with the norm $||v_h||_p = \left(\sum_j \Delta x_j |v_j|^p\right)^{\frac1p}$.

Since

$$u_j^{n+1} = (1 - \nu_j)\, u_j^n + \nu_j u_{j-1}^n\,,$$

if $0 \le \nu_j \le 1$ one has

$$|u_j^{n+1}|^p \le (1 - \nu_j)\, |u_j^n|^p + \nu_j |u_{j-1}^n|^p$$

by convexity. Therefore

$$||u_h^{n+1}||_p^p = \sum_j \Delta x_j |u_j^{n+1}|^p \le \sum_j (\Delta x_j - a\Delta t)\, |u_j^n|^p + \sum_j a\Delta |u_{j-1}^n|^p = \sum_j \Delta x_j |u_j^n|^p = ||u_h^n||_p^p\,.$$

4 Compactness techniques

Compactness techniques, when applied to numerical analysis, often provide a different strategy to prove convergence, however without an explicit calculation of the rate of convergence. It can be generalized to nonlinear equations and numerical schemes as well (a main asset, but not detailed in this course). Usually the proof has three steps: (a) an estimate of the discrete derivative; (b) a compactness result; (c) convergence to a solution.

Here we focus on BV (Bounded Variation) techniques. The characterization of the BV norm that we consider is

$$BV(\mathbb{R}^d) = \left\{ u \in L^1_{\mathrm{loc}}(\mathbb{R}^d),\ |u|_{BV} := \sup_{\varphi \in W^{1,\infty}_{b,0}(\mathbb{R}^d)} - \int_{\mathbb{R}^d} u(x) \nabla \cdot \varphi(x) \mathrm{d}x < \infty \right\},$$

where

$$W^{1,\infty}_{b,0}(\mathbb{R}^d) = \left\{ \varphi \in \left(W^{1,\infty}_0(\mathbb{R}^d)\right)^d : \|\varphi\|_{L^\infty(\mathbb{R}^d)} \le 1 \right\}$$

is the space of compactly supported vector fields in $W^{1,\infty}(\mathbb{R}^d)$ bounded by one.

One has the Helly's selection theorem: if $(u_n)_n$ is a bounded sequence in $L^1(\mathbb{R}) \cap BV(\mathbb{R})$, then there exists $u \in L^1(\mathbb{R}) \cap BV(\mathbb{R})$ such that, up to the extraction of a subsequence, $\lim_{n\to\infty} \|u_n - u\|_{L^1_{\mathrm{loc}}(\mathbb{R})} = 0$.
In the context of numerical methods, see Godlewski-Raviart, *Hyperbolic systems of conservations laws*, Ellipse, 1991, page 53.
As similar compactness result holds for functions in $L^1(\Omega) \cap BV(\Omega)$ provided $\Omega \subset \mathbb{R}^d$ is bounded with Lipschitz boundary (Giusti, *Minimal surfaces and functions with bounded variations*, 1984).

(0) Properties of $BV(\mathbb{R})$

1. For $u \in W^{1,1}(\mathbb{R})$, show that $|u|_{BV(\mathbb{R})} = \|u'\|_{L^1(\mathbb{R})}$.

Since $u \in W^{1,1}(\mathbb{R})$, we can integrate by parts

$$-\int_{\mathbb{R}} u(x)\varphi'(x)\mathrm{d}x = \int_{\mathbb{R}} u'(x)\varphi(x)\mathrm{d}x \le \|u'\|_{L^1(\mathbb{R})} \|\varphi\|_{L^\infty(\mathbb{R})} \le \|u'\|_{L^1(\mathbb{R})},$$

so $|u|_{BV(\mathbb{R})} \le \|u'\|_{L^1(\mathbb{R})}$. Taking φ as a mollification of $\mathrm{sign}(u')$, we can saturate the inequality, hence $|u|_{BV(\mathbb{R})} = \|u'\|_{L^1(\mathbb{R})}$.

2. Let $u(x) = 1$ for $-1 < x < 1$ and $u(x) = 0$ otherwise. Show that $|u|_{BV(\mathbb{R})} = 2$.

We have

$$-\int_{\mathbb{R}} u(x)\varphi'(x)\mathrm{d}x = -\int_{-1}^{+1} \varphi'(x)\mathrm{d}x = -\varphi(1) + \varphi(-1) < 2.$$

Taking φ as a smooth function such that $\varphi(1) = -1$ and $\varphi(-1) = 1$, we obtain $|u|_{BV(\mathbb{R})} = 2$.

3. Let $u_h = (u_j)_{j\in\mathbb{Z}} \in L^1(\mathbb{R})$ be a numerical profile, that is

$$u_h(x) = u_j \text{ for } (j - \frac{1}{2})h < x < (j + \frac{1}{2})h.$$

Show that $|u|_{BV(\mathbb{R})} = \sum_{j\in\mathbb{Z}} |u_j - u_{j-1}|$.

In the same spirit as before

$$\begin{aligned} -\int_{\mathbb{R}} u(x)\varphi'(x)\mathrm{d}x &= -\sum_{j\in\mathbb{Z}} u_j \int_{(j-\frac12)h}^{(j+\frac12)h} \varphi'(x)\mathrm{d}x \\ &= -\sum_{j\in\mathbb{Z}} u_j \left(\varphi((j+\tfrac12)h) - \varphi(j - \tfrac12)h)\right) \\ &= \sum_{j\in\mathbb{Z}} (u_j - u_{j-1})\, \varphi((j - \tfrac12)h) \\ &\le \sum_{j\in\mathbb{Z}} |u_j - u_{j-1}|\,, \end{aligned}$$

and by a good choice of φ the inequality is saturated.

(a) Discrete estimates Now we apply this material to the numerical scheme.

1. Take the scheme (3). Show that, under CFL, it can be recast under the Harten form

$$u_j^{n+1} = (1 - C_j - D_j)\, u_j^n + C_j u_{j-1}^n + D_j u_{j+1}^n$$

with $0 \le C_j, D_j$ and $C_j + D_j \le 1$.

The scheme is defined by

$$u_j^{n+1} = u_j - \nu_j^-\left(u_j^n - u_{j+1}^n\right) + \nu_{j-1}^+\left(u_{j-1}^n - u_j^n\right) = \left(1 - \nu_j^- - \nu_{j-1}^+\right) u_j^n + \nu_j^- u_{j+1}^n + \nu_{j-1}^+ u_{j-1}^n\,,$$

where $\nu_j = \frac{a_j\Delta t}{\Delta x}$. Therefore if $\nu = \frac{A\Delta t}{\Delta x} \le \frac{1}{2}$, we have $\nu_j^- + \nu_{j-1}^+ \le 1$.

2. Show that, under CFL, $\sum_{j\in\mathbb{Z}} |u_j^{n+1} - u_{j-1}^{n+1}| \le \sum_{j\in\mathbb{Z}} |u_j^n - u_{j-1}^n|$.

First, we arrange the terms,

$$\begin{aligned} u_j^{n+1} - u_{j-1}^{n+1} &= \left(1 - \nu_j^- - \nu_{j-1}^+\right) u_j^n + \nu_j^- u_{j+1}^n + \nu_{j-1}^+ u_{j-1}^n \\ &\quad - \left(1 - \nu_{j-1}^- - \nu_{j-2}^+\right) u_{j-1}^n - \nu_{j-1}^- u_j^n - \nu_{j-2}^+ u_{j-2}^n \\ &= \nu_j^-\left(u_{j+1}^n - u_j^n\right) + \nu_{j-2}^+\left(u_{j-1}^n - u_{j-2}^n\right) + \left(1 - \nu_{j-1}^+ - \nu_{j-1}^-\right)\left(u_j^n - u_{j-1}^n\right)\,. \end{aligned}$$

Therefore, for $\nu \le \frac{1}{2}$,

$$\begin{aligned} \sum_{j\in\mathbb{Z}} \left|u_j^{n+1} - u_{j-1}^{n+1}\right| &\le \sum_{j\in\mathbb{Z}} \nu_j^- \left|u_{j+1}^n - u_j^n\right| + \sum_{j\in\mathbb{Z}} \nu_{j-2}^+ \left|u_{j-1}^n - u_{j-2}^n\right| + \sum_{j\in\mathbb{Z}} \left(1 - \nu_{j-1}^+ - \nu_{j-1}^-\right) \left|u_j^n - u_{j-1}^n\right| \\ &\le \sum_{j\in\mathbb{Z}} \nu_{j-1}^- \left|u_j^n - u_{j-1}^n\right| + \sum_{j\in\mathbb{Z}} \nu_{j-1}^+ \left|u_j^n - u_{j-1}^n\right| + \sum_{j\in\mathbb{Z}} \left(1 - \nu_{j-1}^+ - \nu_{j-1}^-\right) \left|u_j^n - u_{j-1}^n\right| \\ &\le \sum_{j\in\mathbb{Z}} \left|u_j^n - u_{j-1}^n\right|\,. \end{aligned}$$

3. Show that, under CFL, a 2D discrete BV inequality holds

$$\sum_{0\le n\le T/\Delta t} \left(\sum_{j\in\mathbb{Z}} \frac{\left|u_j^n - u_{j-1}^n\right|}{\Delta x} + \sum_{j\in\mathbb{Z}} \frac{\left|u_j^{n+1} - u_j^n\right|}{\Delta t} \right) \Delta x \Delta t \le C |u|_{BV(\mathbb{R})}. \tag{5}$$

Using the definition of the scheme, we directly get

$$\begin{aligned} \sum_{j\in\mathbb{Z}} \Delta x \left|u_j^{n+1} - u_j^n\right| &\le \sum_{j\in\mathbb{Z}} \Delta x \nu_j^- \left|u_{j+1}^n - u_j^n\right| + \sum_{j\in\mathbb{Z}} \Delta x \nu_{j-1}^+ \left|u_j^n - u_{j-1}^n\right| \\ &\le 2\Delta x \nu \sum_{j\in\mathbb{Z}} \left|u_j^n - u_{j-1}^n\right| \le 2A\Delta t \sum_{j\in\mathbb{Z}} \left|u_j^n - u_{j-1}^n\right|\,. \end{aligned}$$

By the previous step,

$$\sum_{j\in\mathbb{Z}} \left|u_j^n - u_{j-1}^n\right| \le \sum_{j\in\mathbb{Z}} \left|u_j^0 - u_{j-1}^0\right| = |u_0|_{BV(\mathbb{R})}$$

hence, we obtain

$$\begin{aligned} \sum_{0\le n\le T/\Delta t} \left(\sum_{j\in\mathbb{Z}} \Delta t \left|u_j^n - u_{j-1}^n\right| + \sum_{j\in\mathbb{Z}} \Delta x \left|u_j^{n+1} - u_j^n\right| \right) &\le \sum_{0\le n\le T/\Delta t} (2A+1)\Delta t |u_0|_{BV(\mathbb{R})} \\ &\le (2A+1) T |u_0|_{BV(\mathbb{R})}. \end{aligned}$$

4. Defining $u_h(t,x) = u_j^n$ for $(x,t) \in \Omega_j^n$, where

$$\Omega_j^n := \{(x,t) \mid (j - 1/2)h < x < (j+1/2)h\} \text{ and } n\Delta t < t < (n+1)\Delta t,$$

the left hand side of (5) is equivalent to $|u_h|_{BV([0,T]\times\mathbb{R})}$, so we have $|u_h|_{BV([0,T]\times\mathbb{R})} \le C|u|_{BV(\mathbb{R})}$.

(b) Compactness result By compactness, there exists a subsequence converging to u in $L^1_{\mathrm{loc}}([0,T]\times\mathbb{R})$ for $h = \Delta x \to 0$ with Δt fixed such that the CFL is satisfied.
Note that slightly different but perhaps more intuitive compactness result can be used: Leveque, *Numerical Methods for Conservation Laws*, 1990, page 162.

(c) Convergence to a solution The aim is now to prove that u is a solution of the original equation. To simplify the notations, we consider that $a(x) = a > 0$ is a constant.

1. Show the following identity for all $\varphi \in C_0^1([0,\infty) \times \mathbb{R})$:

$$\sum_{0\le n\le T/\Delta t} \sum_{j\in\mathbb{Z}} \Delta x \Delta t \left(\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{a(u_j^n - u_{j-1}^n)}{\Delta x} \right) \varphi(x_j, t_n) = 0.$$

By definition of the scheme,

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

so there is nothing to do.

2. Recast as

$$\sum_{1\le n\le T/\Delta t} \sum_{j\in\mathbb{Z}} \Delta x \Delta t u_j^n \left(\frac{\varphi(x_j, t_{n-1}) - \varphi(x_j, t_n)}{\Delta t} \right) - \sum_{j\in\mathbb{Z}} \Delta x u_j^0 \varphi(x_j, 0) + \sum_{0\le n\le T/\Delta t} \sum_{j\in\mathbb{Z}} \Delta x \Delta t u_j^n \left(\frac{a\left(\varphi(x_j, t_n) - \varphi(x_{j+1}, t_n)\right)}{\Delta x} \right) = 0.$$

The idea is to have u_h evaluated at the same points,

$$\sum_{j\in\mathbb{Z}} \left(u_j^n - u_{j-1}^n\right) \varphi(x_j, t_n) = \sum_{j\in\mathbb{Z}} u_j^n \left(\varphi(x_j, t_n) - \varphi(x_{j+1}, t_n)\right)\,,$$

and

$$\sum_{0\le n\le T/\Delta t} \left(u_j^{n+1} - u_j^n\right) \varphi(x_j, t_n) = \sum_{1\le n\le T/\Delta t} u_j^n \left(\varphi(x_j, t_{n-1}) - \varphi(x_j, t_n)\right) - u_j^0 \varphi(x_j, 0),$$

so we obtain the claimed formula.

3. Show that each part admits a limit (up to subsequence extraction), and that the limit satisfies

$$-\int_0^T \int_{\mathbb{R}} u(x,t) \left(\partial_t \varphi(x,t) + a\partial_x \varphi(x,t)\right) \mathrm{d}x\mathrm{d}t - \int_{\mathbb{R}} u_0(x) \varphi(x, 0) \mathrm{d}x = 0 \tag{6}$$

for all $\varphi \in C_0^1([0,\infty) \times \mathbb{R})$.

The equality found at previous step can be written as

$$\int_{\Delta t}^T \int_{\mathbb{R}} u_h(x,t) \left(\frac{\varphi(x_j, t_{n-1}) - \varphi(x_j, t_n)}{\Delta t} \right) \mathrm{d}x\mathrm{d}t - \int_{\mathbb{R}} u_0(x_j) \varphi(x_j, 0) \mathrm{d}x + \int_0^T \int_{\mathbb{R}} u_h(x,t) \left(\frac{a\left(\varphi(x_j, t_n) - \varphi(x_{j+1}, t_n)\right)}{\Delta x} \right) \mathrm{d}x\mathrm{d}t = 0, \tag{7}$$

where x_j and t_n are now interpreted as step functions on Ω_j^n. We have

$$\begin{aligned} \frac{\varphi(x_j, t_{n-1}) - \varphi(x_j, t_n)}{\Delta t} &\to -\partial_t \varphi(x,t) && \text{in } L^\infty([0,T]\times\mathbb{R}), \\ \frac{\varphi(x_j, t_n) - \varphi(x_{j+1}, t_n)}{\Delta x} &\to -\partial_x \varphi(x,t) && \text{in } L^\infty([0,T]\times\mathbb{R}), \end{aligned}$$

as $h \to 0$. In general $v_n \to v$ in $L_1(\Omega)$ and $w_n \to w$ in $L_\infty(\Omega)$, we have

$$\int_\Omega v_n w_n \to \int_\Omega v w\,,$$

since

$$\begin{aligned} \left| \int_\Omega v_n w_n - \int_\Omega v w \right| &\le \left| \int_\Omega v_n (w_n - w) \right| + \left| \int_\Omega (v_n - v)\, w \right| \\ &\le \|v_n\|_{L_1(\Omega)} \|w_n - w\|_{L_\infty(\Omega)} + \|v_n - v\|_{L_1(\Omega)} \|w_n\|_{L_\infty(\Omega)}\,. \end{aligned}$$

Therefore, we can pass to the limit in (7) to get

$$-\int_0^T \int_{\mathbb{R}} u(x,t) \left(\partial_t \varphi(x,t) + a\partial_x \varphi(x,t)\right) \mathrm{d}x\mathrm{d}t - \int_{\mathbb{R}} u_0(x) \varphi(x, 0) \mathrm{d}x = 0,$$

for any $\varphi \in C_0^1([0,\infty) \times \mathbb{R})$.

4. Assume some extra regularity, for example $u \in C^1([0,T] \times \mathbb{R})$. Show that u is the unique solution to $\partial_t u + a\partial_x u = 0$ plus initial condition.

If $u \in C^1([0,T] \times \mathbb{R})$, then we can integrate by parts,

$$\begin{aligned} -\int_0^T \int_{\mathbb{R}} u(x,t) \partial_t \varphi(x,t) \mathrm{d}x\mathrm{d}t &= \int_0^T \int_{\mathbb{R}} \partial_t u(x,t) \varphi(x,t) + \int_{\mathbb{R}} u(x,0) \varphi(x,0) \mathrm{d}x, \\ -\int_0^T \int_{\mathbb{R}} u(x,t) \partial_x \varphi(x,t) \mathrm{d}x\mathrm{d}t &= \int_0^T \int_{\mathbb{R}} \partial_x u(x,t) \varphi(x,t), \end{aligned}$$

and therefore

$$\int_0^T \int_{\mathbb{R}} \left(\partial_t u(x,t) + a\partial_x u(x,t)\right) \varphi(x,t) \mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}} \left(u(x,0) - u_0(x)\right) \mathrm{d}x.$$

Since this is valid for any $\varphi \in C_0^1([0,\infty) \times \mathbb{R})$, u has to solve

$$\partial_t u(x,t) + a\partial_x u(x,t) = 0\,, \qquad u(x,0) = u_0(x).$$

5. Assume much less regularity $u \in L^1_{\mathrm{loc}}([0,T] \times \mathbb{R})$. Show that (6) yields that $u(x,t) = u_0(x - at)$ for almost all (x,t).

We make the change of variable $z = x - at$, and define

$$\psi(z,t) = \varphi(z + at, t),$$

Since $\partial_t \psi = \partial_t \varphi + a\partial_x \varphi$, we have

$$\int_0^T \int_{\mathbb{R}} u(x,t) \left(\partial_t \varphi(x,t) + a\partial_x \varphi(x,t)\right) \mathrm{d}x\mathrm{d}t = \int_0^T \int_{\mathbb{R}} u(z + at, t) \left(\partial_t \psi(z,t)\right) \mathrm{d}z\mathrm{d}t.$$

Since

$$\int_{\mathbb{R}} u_0(x) \varphi(x,0) \mathrm{d}x = \int_{\mathbb{R}} u_0(x) \psi(x,0) \mathrm{d}x = -\int_0^\infty \int_{\mathbb{R}} u_0(z) \partial_t \psi(z,t) \mathrm{d}z\mathrm{d}t,$$

the weak formulation (6) is equivalent to

$$\int_0^T \int_{\mathbb{R}} \left(u(z + at, t) - u_0(z)\right) \left(\partial_t \psi(z,t)\right) \mathrm{d}z\mathrm{d}t = 0.$$

Therefore $u(z + at, t) = u_0(z)$ *i.e.* $u(x,t) = u_0(x - at)$.