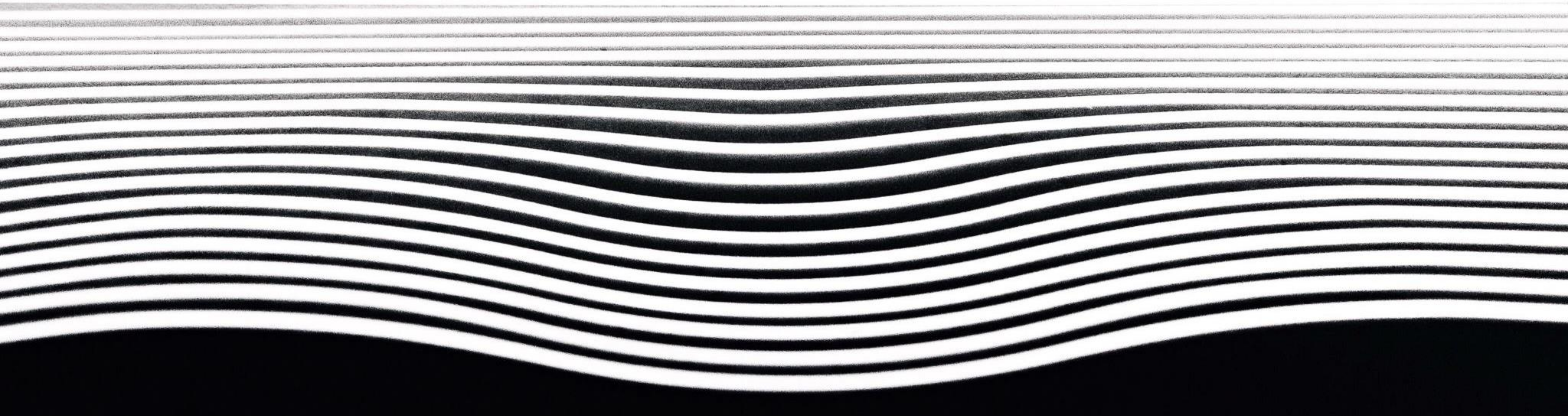
Numerical methods for conservation laws
8: Monotone Schemes



We have

- introduced finite difference schemes
- discussed conservative schemes
- given examples for numerical fluxes
- proven that limits of such schemes are weak solutions

We now discuss monotone schemes as an important special class.

Up to now, we have numerical schemes that produce a sequence of states, indexed over the time indices, starting with the initial values:

$$U^0, U^1, U^2, \dots, U^{n-1}, U^n, U^{n+1}, \dots$$

Every state is constructed from its predecessor. We write this as

$$U^{n+1} = G(U^n)$$

We call the numerical scheme G monotone if

$$U \leq V \quad \text{implies} \quad G(U) \leq G(V)$$

Why does monotonicity matter? (1/3)

1) For many conservation laws, $u_0 \leq v_0$ at the initial data implies that $u(t) \leq v(t)$ for all subsequent times.

If the scheme is monotone, then $\mathrm{U}^0 \leq V^0$ implies

$$U^1 \leq V^1, \qquad U^2 \leq V^2, \qquad \dots, \qquad U^n \leq V^n, \qquad \dots$$

In other words, monotone schemes reproduce a maximum principle at the discrete level.

Why does monotonicity matter? (2/3)

2) Suppose we have a monotone consistent conservative scheme

$$U_j^{n+1} = U_j^n - \frac{k}{h}\left(F^n_{j+\frac{1}{2}} - F^n_{j+\frac{1}{2}}\right)$$

Any state vector with contant values $V_j^n = U^{\max}$ is preserved under this scheme.

If $U^0 \leq U^{max}$, then $U^n \leq U^{max}$ for all subsequent time steps.

Similarly for lower bounds of U…

Why does monotonicity matter? (3/3)

3) Lastly, if a scheme is not monotone, then we may expect uncontrolled growth of the solution.

Important observation

A numerical scheme G is monotone if and only if $G(U)$ is non-decreasing in each variable U_j.

Example (Transport Equation)

The FTBS scheme is

$$U_j^{n+1} = U_j^n - c\frac{k}{h}\left(U_j^n - U_{j-1}^n\right) = \left(1 - c\frac{k}{h}\right)U_j^n + c\frac{k}{h}U_{j-1}^n$$

For the scheme to be monotone we need the right-hand side is non-decreasing in U_j^n and U_{j-1}^n. That is,

1. $c \geq 0$ (all waves travel to the right, upwinding)
2. $c \leq \frac{h}{k}$ (time step can't be too large)

Example (Burger's Equation)

The FTBS scheme for Burgers' equation is

$$U_j^{n+1} = U_j^n - \frac{k}{h}\left((U_j^n)^2 \ - (U_{j-1}^n)^2\right) = \left(U_j^n - \frac{k}{h}(U_j^n)^2\right) + \frac{k}{h}(U_{j-1}^n)^2$$

Provided that

1. U_j is positive
2. $2U_j \leq \frac{h}{k}$

the scheme "behaves" monotone. In other words, all waves travel to the right (upwinding) and not faster than the mesh velocity.

Let us focus on the case of a finite difference scheme in where the numerical fluxes only depend on the immediate neighbors.

$$
\begin{aligned}
U_j^{n+1} &= U_j^n - \frac{k}{h}\left(F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n \right) \\
&= U_j^n - \frac{k}{h}\left(F(U_j^n, U_{j+1}^n) - F(U_{j-1}^n, U_j^n) \right) \\
&=: G(U_{j-1}^n, U_j^n, U_{j+1}^n)
\end{aligned}
$$

We also write with some minor abuse of notation:

$$
G(U^n)_j = G(U_{j-1}^n, U_j^n, U_{j+1}^n)
$$

We call the scheme monotone if G is non-decreasing in each argument.
Let's express that in terms of the differentials.

$$0 \le \frac{\partial G}{\partial U^n_{j-1}} = \frac{k}{h}\, \partial_1 F(U^n_{j-1}, U^n_j)$$

$$0 \le \frac{\partial G}{\partial U^n_j} = 1 - \frac{k}{h}\left(\partial_1 F(U^n_j, U^n_{j+1}) - \partial_2 F(U^n_{j-1}, U^n_j) \right)$$

$$0 \le \frac{\partial G}{\partial U^n_{j+1}} = -\frac{k}{h}\, \partial_2 F(U^n_j, U^n_{j+1})$$

The first and the third inequality are equivalent to F non-decreasing in its first argument and non-increasing in its second argument.

The second inequality requires that the quotient k/h must be sufficiently small, in particular, the time step must be smaller than the space step.

We verify those three conditions with the Lax-Friedrich flux

$$F_{LF}(u, v) = \frac{1}{2}(f(u) + f(v)) - \frac{\alpha}{2}(v - u)$$

where α is the maximum norm of f'. The first two conditions hold:

$$\partial_u F(u, v) = \frac{1}{2} f'(u) + \frac{\alpha}{2} \geq 0$$

$$\partial_v F(u, v) = \frac{1}{2} f'(v) - \frac{\alpha}{2} \leq 0$$

We check the third condition:

$$0 \leq 1 - \frac{k}{2h}(f'(u) - \alpha - f'(u) - \alpha) = 1 - \frac{k}{h}\alpha$$

Apparently, the third condition says

$$0 \leq 1 - \frac{k}{h}\alpha$$

In other words, the time step must be small enough such that

$$k\alpha \leq h$$

The higher the first derivative of the flux, the smaller the time step necessary to ensure the scheme is monotone. The "grid speed" h/k must be at least as large as the physical speed.

Such conditions are known as **Courant–Friedrichs–Lewy (CFL) conditions.**

What about the Lax-Wendroff flux?

$$F_{LW}(u,v) = \frac{1}{2}(f(u) + f(v)) - \frac{k}{2h} f'\left(\frac{u+v}{2}\right)(f(v) - f(u))$$

This is generally not monotone. For example, when $f(u) = u^2$, then

$$F(u,v) = \frac{u^2 + v^2}{2} - \frac{k}{2h} \cdot (u+v)(v^2 - u^2)$$

With some calculations, we find that

$$\partial_u F(u,v) = u - \frac{k}{2h}(-3u^2 - 2uv + v^2)$$

For example, when $|u| \ll |v|$ and v is negative, then this derivative is negative.
We conclude that the Lax-Wendroff flux is generally not monotone.

We call a numerical scheme G **L1 contractive** if

$$\sum_j \left|G(U)_j - G(V)_j\right| \leq \sum_j \left|U_j - V_j\right|$$

Theorem

Suppose that G is a conservative monotone scheme.

Then G is L1 contractive.

Proof

We write

$$a^+ = \max(a, 0), \qquad a^- = \min(a, 0)$$

Given U and V, we define W by

$$W_j = \max(U_j, V_j) = V_j + (U_j - V_j)^+$$

By construction and monotonicity

$$G(U)_j \, , \quad G(V)_j \leq G(W)_j$$

Furthermore, by monotonicity

$$\left(G(U)_j - G(V)_j \right)^+ \leq G(W)_j - G(V)_j$$

Hence, using conservation

$$\sum \left[G(U)_j - G(V)_j \right]^+ \leq \sum_j G(W)_j - G(V)_j$$

$$= \sum W_j - V_j$$

$$= \sum V_j + (U_j - V_j)^+ - V_j$$

$$= \sum (U_j - V_j)^+$$

It follows that

$$\begin{aligned}
&\sum |G(U)_j - G(V)_j| \\
&= \sum \left[G(U)_j - G(V)_j\right]^+ + \sum \left\{G(V)_j - G(U)_j\right\}^+ \\
&\le \sum (U_j - V_j)^+ + \sum (V_j - U_j)^+ \\
&= \sum |U_j - V_j|
\end{aligned}$$

$\square$

We call a numerical scheme G **total variation diminishing (TVD)** if

$$\sum_j \left|G(U)_j - G(U)_{j+1}\right| \le \sum_j \left|U_j - U_{j+1}\right|$$

Theorem

Suppose that G is L1 contractive. Then G is TVD.

Proof

Given U, we define V by $V_j = U_{j+1}$. Then the L1 contraction applied to U and V implies

$$\sum_j \left|G(U)_j - G(U)_{j+1}\right| \le \sum_j \left|U_j - U_{j+1}\right|$$

We have shown that monotone schemes replicate important properties of the original physical system:

- maximum principles
- L1-contraction
- TV-diminishment.

We now study the connection to entropy solutions.

We write

$$u \wedge c = \min(u, c), \qquad u \vee c = \max(u, c)$$

and similarly for vectors, entrywise.

Recall (and rewrite) the Kruzkov entropy pairs

$$\eta(u) = |u - c| = u \vee c - u \wedge c$$

$$\psi(u) = \operatorname{sgn}(u-c)\left[f(u) - f(c)\right] = f(u \vee c) - f(u \wedge c)$$

We define a discrete entropy pair.

$$\eta(U^n)_j = \eta(U_j^n) = U_j^n \vee c - U_j^n \wedge c$$

$$\Psi^n_{j+\frac{1}{2}} = F(U_j^n \vee c) - F(U_j^n \wedge c)$$

We want to show a discrete entropy condition

$$\frac{\eta(U_j^{n+1}) - \eta(U_j^n)}{k} + \frac{\Psi^n_{j+\frac{1}{2}} - \Psi^n_{j-\frac{1}{2}}}{h} \leq 0$$

Consider, using definitions

$$G(U^n \vee c)_j = (U^n \vee c)_j - \frac{k}{h}\left[F(U^n_j \vee c) - F(U^n_{j-1} \vee c) \right]$$

$$G(U^n \wedge c)_j = (U^n \wedge c)_j - \frac{k}{h}\left[F(U^n_j \wedge c) - F(U^n_{j-1} \wedge c) \right]$$

some abuse of notation

Subtracting gives

$$G(U^n \vee c)_j - G(U^n \wedge c)_j = |U^n_j - c| - \frac{k}{h}\left[\psi^n_{j+\frac{1}{2}} - \psi^n_{j-\frac{1}{2}} \right]$$

Via consistency and monotonicity

$$c \vee U_j^{n+1} = G(c) \vee G(U^n)_j \leq G(U^n \vee c)_j \vee G(U^n \vee c)_j = G(U^n \vee c)_j$$

$$c \wedge U_j^{n+1} = G(c) \wedge G(U^n)_j \geq G(U^n \wedge c)_j \wedge G(U^n \wedge c)_j = G(U^n \wedge c)_j$$

Hence

$$c \vee U_j^{n+1} \leq G(U^n \vee c)_j \,, \qquad c \wedge U_j^{n+1} \geq G(U^n \wedge c)_j$$

We get

$$|U_j^{n+1} - c| = U_j^{n+1} \vee c - U_j^{n+1} \wedge c \leq G(U^n \vee c)_j - G(U^n \wedge c)_j$$

Finally $\quad \eta(U_j^{n+1}) = |U_j^{n+1} - c|$

$$\leq G(U^n \vee c)_j - G(U^n \wedge c)_j$$

$$= \underbrace{|U_j^n - c|}_{=\eta(U_j^n)} - \frac{k}{h}\left[\Psi^n_{j+\frac{1}{2}} - \Psi^n_{j-\frac{1}{2}}\right]$$

This shows the discrete entropy condition

$$\frac{\eta(U_j^{n+1}) - \eta(U_j^n)}{k} + \frac{\Psi^n_{j+\frac{1}{2}} - \Psi^n_{j-\frac{1}{2}}}{h} \leq 0 \qquad \square$$

We call a numerical scheme G **monotonicity preserving (MP)** if it maps monotonely increasing sequences to monotonely increasing sequences.

More explicitly,

$$\forall j: \; U_j \le U_{j+1} \quad \Rightarrow \quad \forall j: \; G(U)_j \le G(U)_{j+1}$$

Intuitively, this prevents the emergence of spurious oscillations.

Recall that schemes (in practice) are defined via local stencils.

$$G(U)_j = G(U_{j-p}, \dots, U_j, \dots, U_{j+q})$$

If G is MP, then the following is true for all indices $i < j$:

If $U_{i-p}, \dots, U_{j+q}$ increases monotonely, then so does $G(U)_i, \dots, G(U)_j$.

Theorem

Suppose that a scheme G is TVD. Then G is MP.

Proof: Suppose that G is TVD. Consider the sequence

$$U_j^n = \begin{cases} U^- & \text{if } j \le J^- \\ \text{monotonely} & \text{if } J^- \le j \le J^+ \\ U^+ & \text{if } J^+ \le j \end{cases}$$

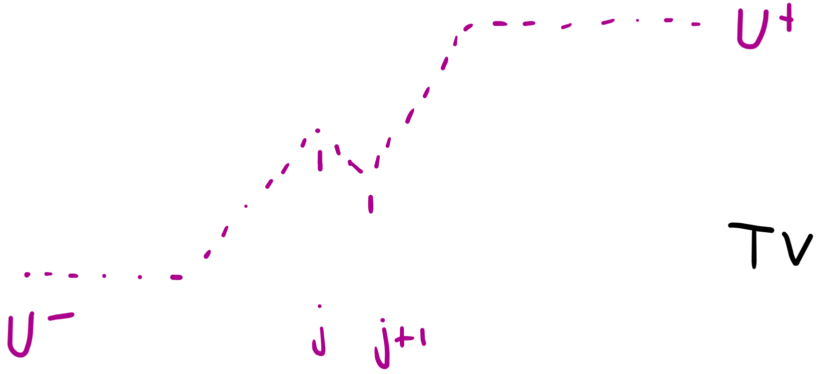
The entire sequence increases monotonely. Its total variation is:

$$TV(U^n) = U^+ - U^-$$

Now suppose this TVD scheme is not MP.

For some choice of U as above, $G(U)$ is not monotonely increasing.

- There exists j such that $G(U)_j > G(U)_{j+1}$.
- Since TVD schemes are consistent, $G(U)$ takes the values U^- and U^+ to the left and the right.



Hence

$$TV(G(U)) \geq |U^- - G(U)_j| + |G(U)_j - G(U)_{j+1}| + |U^+ - G(U)_{j+1}|$$

$$> |U^- - U^+|$$

This contradicts G being TVD. ↯ $\square$

In simple terms:

Monotone schemes preserve the order of solutions in terms of amplitudes.

L1 contractive schemes decrease the differences between solutions over time.

TVD schemes do not accentuate existing extrema or create to many new ones.

Monotonicity-preserving schemes do not produce new undershoots and overshoots in the solution.

Summary of important properties:

Monotone $\Rightarrow$ L^1 contraction $\Rightarrow$ TVD $\Rightarrow$ MP

Example: Linear schemes

$$U_j^{n+1} = G(U^n)_j = \sum_{\ell=-p}^{q} c_\ell \, U_{j+\ell}^n$$

Notice that

$$\frac{\partial U_j^{n+1}}{\partial U_{j+\ell}^n} = c_\ell$$

The scheme is monotone if and only if all $c_\ell \geq 0$.

What about monotonicity-preservation?

Consider a monotonely increasing sequence U^n and

$$U^{n+1}_{j+1} - U^{n+1}_{j} = \sum_{\ell=-p}^{q} c_\ell \Big(\underbrace{U^n_{j+\ell+1} - U^n_{j+\ell}}_{\geqslant 0} \Big)$$

For the scheme to be MP, it suffices that all $c_\ell \geq 0$.
This condition is also necessary. To see this, let

$$U^n_j = \begin{cases} 0 & j \leq 0 \\ 1 & j > 0 \end{cases} \Rightarrow U^n_{j+1} - U^n_j = \begin{cases} 1 & j = 0 \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow U^{n+1}_{j+1} - U^{n+1}_{j} = \begin{cases} c_{-j} & j \in [0-p, 0+q] \\ 0 & \end{cases}$$

Thus, G is MP if and ony if all $c_\ell \geq 0$.

Lemma: for a linear scheme G, the following are equivalent:

1. G is monotone
2. G is monotonicity-preserving
3. G is has only non-negative coefficients.

$$\begin{tikzcd} \text{Monotone} \arrow[r, Rightarrow] & L^1 \text{ contraction} \arrow[r, Rightarrow] & \text{TVD} \arrow[r, Rightarrow] & \text{MP} \arrow[lll, bend left, "\text{Linear}"] \end{tikzcd}$$

We have studied important qualitative properties of finite-difference schemes:

- Monotonicity
- L1 contraction
- TV-diminishing
- MP

And the implications between those.

However, there are intrinsic limitations of monotone schemes. To understand those, we need to address the error analysis.