MA4110/MA5020: Computational Methods for Fluid Flow

Lecture 9: Conservative Finite Difference Schemes, the Lax-Wendroff Theorem, and the Method of Lines

Instructor: Biswarup Biswas

1 Introduction

In Lectures 6 through 8, we examined the Finite Difference Method (FDM) applied to linear hyperbolic advection $u_t + au_x = 0$. We studied truncation error, Von Neumann stability, and modified equations, culminating in Godunov's Barrier Theorem: linear difference schemes cannot be higher than first-order accurate without generating spurious oscillations near sharp gradients.

When dealing with **nonlinear hyperbolic conservation laws**:

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0 \tag{1}$$

the computational challenges multiply. As we established in Lecture 5 via the method of characteristics, solutions to nonlinear equations inevitably break down and develop discontinuous shock waves, even from infinitely smooth initial data.

At a shock discontinuity, classical derivatives do not exist. In this lecture, we answer the foundational question of computational shock physics within the finite difference framework:

- **Why must a finite difference scheme be written in conservative (flux) form?**
- **Why do non-conservative schemes compute shock waves that propagate at the wrong physical speed, even in the limit as grid spacing $\Delta x \to 0$?**
- **What is the Lax-Wendroff Theorem, and how does it guarantee correct shock speeds?**
- **How does the Method of Lines (semi-discretization) decouple spatial differencing from time integration?**

2 The Breakdown of Differential Forms at Shocks: Review from Lecture 5

In Lecture 5, we analyzed the nonlinear inviscid Burgers' equation and established that whenever faster states overtake slower states ($u_L > u_R$), classical derivatives blow up ($\frac{\partial u}{\partial x} \to -\infty$) at the breaking time t_s, forming a discontinuous shock wave.

For smooth solutions, the chain rule allows us to freely interchange between two differential forms:

$$\text{Conservative (Divergence) Form: } \frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = \frac{\partial u}{\partial t} + \frac{\partial}{\partial x}\left(\frac{u^2}{2}\right) = 0 \tag{2}$$

$$\text{Non-Conservative (Advective) Form: } \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} = 0 \tag{3}$$

Analytically, equations (2) and (3) are identical as long as $u(x,t)$ remains differentiable.

However, across a shock discontinuity, $\frac{\partial u}{\partial x}$ becomes a singular Dirac delta distribution. In Lecture 5, we resolved this breakdown by going back to the macroscopic integral conservation law:

$$\frac{d}{dt}\int_{x_1}^{x_2} u(x,t)\,dx = f(u(x_1,t)) - f(u(x_2,t)) \tag{4}$$

which yielded the fundamental **Rankine-Hugoniot (RH) jump condition**:

$$\boxed{s = \frac{f(u_L) - f(u_R)}{u_L - u_R} = \frac{[f]}{[u]}} \tag{5}$$

For Burgers' equation ($f(u) = u^2/2$), the exact physical shock propagation speed is:

$$s = \frac{\frac{1}{2}u_L^2 - \frac{1}{2}u_R^2}{u_L - u_R} = \frac{u_L + u_R}{2} \tag{6}$$

2.1 The Central Numerical Question

In this lecture, we transition from continuous mathematics to numerical computation on a discrete grid with finite differences. Both forms can be discretized directly:

$$\text{Discretizing Conservative Form (2):} \quad \frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{F_{j+1/2}^n - F_{j-1/2}^n}{\Delta x} = 0 \tag{7}$$

$$\text{Discretizing Non-Conservative Form (3):} \quad \frac{u_j^{n+1} - u_j^n}{\Delta t} + u_j^n \left(\frac{u_j^n - u_{j-1}^n}{\Delta x} \right) = 0 \tag{8}$$

Because both discrete schemes are consistent with the PDE for smooth flows ($\mathcal{O}(\Delta t, \Delta x)$), one might naively expect both to converge to the exact shock solution as the mesh is refined ($\Delta x, \Delta t \to 0$).

They do not. As we will prove in Section 4, the non-conservative scheme computes a shock that moves at a completely false speed ($s = u_R$), and refining the grid will *never* correct the error! Only conservative difference schemes are guaranteed to capture the true Rankine-Hugoniot shock speed.

3 Conservative Finite Difference Schemes

3.1 Definition of a Conservative Scheme

A finite difference scheme for point values $u_j^n \approx u(x_j, t^n)$ is said to be in **conservative (flux-difference) form** if it can be written as:

$$\boxed{\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{F_{j+1/2}^n - F_{j-1/2}^n}{\Delta x} = 0} \tag{9}$$

Here, the spatial derivative $\frac{\partial f(u)}{\partial x}$ at grid point x_j is approximated by a difference of numerical fluxes evaluated at the intermediate locations $x_{j\pm1/2}$:

$$F_{j+1/2}^n = \mathcal{F}(u_{j-p+1}^n, u_{j-p+2}^n, \ldots, u_{j+q}^n) \tag{10}$$

where $\mathcal{F}$ is a $(p+q)$-point **numerical flux function**.

3.2 Properties of the Numerical Flux Function

For the discrete scheme to represent the continuous conservation law, the function $\mathcal{F}$ must satisfy:

1. **Consistency:** When all state arguments are equal, the numerical flux must return the exact physical flux:
$$\mathcal{F}(u, u, \ldots, u) = f(u) \tag{11}$$
2. **Lipschitz Continuity:** There exists a constant K such that:
$$|\mathcal{F}(v_1, \ldots, v_{p+q}) - \mathcal{F}(w_1, \ldots, w_{p+q})| \leq K \sum_{m=1}^{p+q} |v_m - w_m| \tag{12}$$

3.3 The Telescoping Property and Exact Discrete Conservation

Consider a computational domain covering nodes $j = 1, 2, \ldots, N$. Multiply equation (9) by Δx and sum over all grid points:

$$\sum_{j=1}^{N} u_j^{n+1} \Delta x = \sum_{j=1}^{N} u_j^n \Delta x - \Delta t \sum_{j=1}^{N} \left(F_{j+1/2}^n - F_{j-1/2}^n \right) \tag{13}$$

Expanding the flux summation explicitly:

$$\begin{aligned} \sum_{j=1}^{N} \left(F_{j+1/2}^n - F_{j-1/2}^n \right) &= (F_{3/2}^n - F_{1/2}^n) + (F_{5/2}^n - F_{3/2}^n) + \cdots + (F_{N+1/2}^n - F_{N-1/2}^n) \\ &= \boxed{F_{N+1/2}^n - F_{1/2}^n} \end{aligned} \tag{14}$$

Every interior numerical flux $F_{j+1/2}^n$ enters cell j with a positive sign and cell $j+1$ with an identical negative sign, **canceling out telescopically throughout the interior of the domain**!

Discrete Global Conservation Law

The total discrete mass inside the domain changes exclusively due to fluxes crossing the external physical boundaries:

$$\sum_{j=1}^{N} u_j^{n+1} \Delta x = \sum_{j=1}^{N} u_j^n \Delta x - \Delta t \left(F_{N+1/2}^n - F_{1/2}^n \right)$$

No artificial mass, momentum, or energy is created or destroyed in the interior grid, regardless of the presence of sharp gradients or discontinuities.

4 Why Non-Conservative Schemes Fail: Shock Speed Error

To see why non-conservative differencing is disastrous for hyperbolic problems, consider discretizing the non-conservative Burgers' equation (3) using standard first-order upwinding:

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x} u_j^n \left(u_j^n - u_{j-1}^n \right) \tag{15}$$

4.1 The Hidden Numerical Source Term

Algebraically rewrite the non-conservative product $u_j(u_j - u_{j-1})$:

$$u_j^n(u_j^n - u_{j-1}^n) = \frac{1}{2} \left((u_j^n)^2 - (u_{j-1}^n)^2 \right) + \frac{1}{2}(u_j^n - u_{j-1}^n)^2 \tag{16}$$

Substitute this identity directly back into equation (15):

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} + \frac{\frac{1}{2}(u_j^n)^2 - \frac{1}{2}(u_{j-1}^n)^2}{\Delta x} = \underbrace{-\frac{1}{2\Delta x}(u_j^n - u_{j-1}^n)^2}_{\text{Spurious Numerical Mass Sink } S_j^n} \tag{17}$$

Equation (17) reveals what the non-conservative scheme is *actually* computing: it is solving the true conservation law perturbed by a spurious numerical source term S_j^n!

4.2 Behavior as Mesh is Refined ($\Delta x \to 0$)

Now analyze the magnitude of the spurious source S_j^n in two distinct flow regimes:

- **In Smooth Regions:** By Taylor series, $u_j^n - u_{j-1}^n = \Delta x \frac{\partial u}{\partial x} + \mathcal{O}(\Delta x^2)$. Hence:

$$S_j^n = -\frac{1}{2\Delta x}\left(\Delta x \frac{\partial u}{\partial x}\right)^2 = -\frac{\Delta x}{2}\left(\frac{\partial u}{\partial x}\right)^2 \xrightarrow{\Delta x \to 0} 0 \tag{18}$$

  In smooth regions, the spurious source term vanishes linearly as $\Delta x \to 0$, and the non-conservative scheme converges to the correct smooth solution.

- **Across a Shock Wave:** Across a shock, the states jump by a finite non-zero amount: $u_j^n - u_{j-1}^n \approx u_R - u_L = [u] \neq 0$. Therefore:

$$S_j^n = -\frac{1}{2\Delta x}(u_R - u_L)^2 \xrightarrow{\Delta x \to 0} -\infty \tag{19}$$

  The point-wise source term blows up as $\Delta x \to 0$!

Now compute the **total spurious mass destroyed per unit time** by integrating across the shock transition layer ($\sum_j S_j^n \Delta x$):

$$\text{Total Integrated Mass Loss} = \sum_{\text{shock}} S_j^n \Delta x = -\frac{1}{2}(u_R - u_L)^2 \neq 0 \tag{20}$$

The grid spacing Δx cancels out completely! Even if we refine the mesh to $\Delta x = 10^{-6}$ or take the theoretical limit $\Delta x \to 0$, the non-conservative scheme continues to destroy a finite, non-zero amount of mass at the shock every second.

4.3 Analytical Derivation of the Wrong Shock Speed

Because the non-conservative scheme destroys mass at the rate $-\frac{1}{2}(u_R - u_L)^2$, the global Rankine-Hugoniot balance is corrupted:

$$\begin{aligned} s_{\text{num}}\,[u] &= [f] + \text{Mass Loss} \\ s_{\text{num}}\,(u_L - u_R) &= \frac{1}{2}(u_L^2 - u_R^2) - \frac{1}{2}(u_L - u_R)^2 \\ s_{\text{num}}\,(u_L - u_R) &= u_R(u_L - u_R) \end{aligned} \tag{21}$$

Dividing by $(u_L - u_R)$:

$$\boxed{s_{\text{num}} = u_R \qquad \left(\text{compared to the exact physical speed } s_{\text{exact}} = \frac{u_L + u_R}{2}\right)} \tag{22}$$

Catastrophic Consequence for CFD

Consider a Riemann problem with $u_L = 1.0$ and $u_R = 0.0$:

- **Exact Physical Speed:** $s_{\text{exact}} = \frac{1.0+0.0}{2} = \mathbf{0.5}$. The shock moves steadily forward.
- **Non-Conservative Scheme Speed:** $s_{\text{num}} = u_R = \mathbf{0.0}$. The shock remains **completely frozen at the origin forever**!

Refining the mesh ($\Delta x \to 0$) does not fix this error: the non-conservative numerical solution converges cleanly to a sharp jump that stands completely stationary, delivering an entirely incorrect physical prediction (see Figure 1).

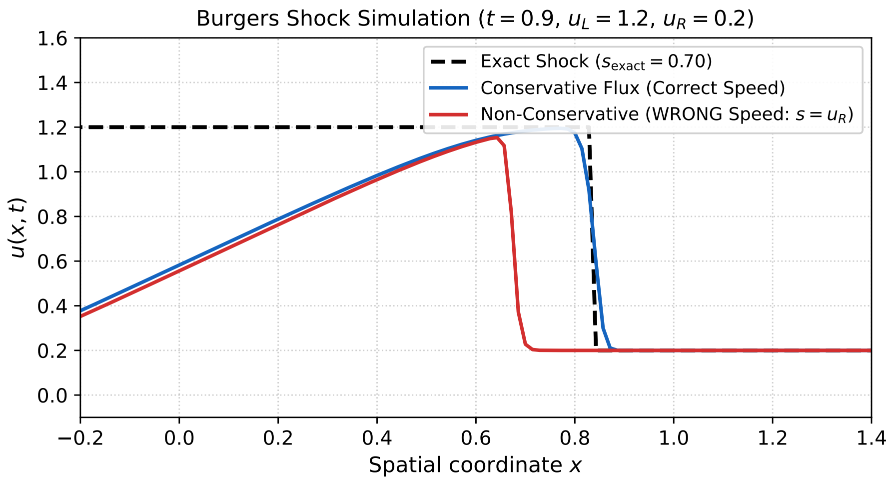

Figure 1: Shock propagation for Burgers' equation ($u_L = 1.2, u_R = 0.2$). The conservative finite difference scheme tracks the exact physical shock speed $s_{\text{exact}} = 0.70$, while the non-conservative scheme violates Rankine-Hugoniot and lags severely behind ($s_{\text{num}} = u_R = 0.2$). Generated with Julia.

5 The Lax-Wendroff Theorem

The mathematical foundation ensuring correct shock speeds is the celebrated theorem established by Peter Lax and Burton Wendroff in 1960:

The Lax-Wendroff Theorem (1960)

Let u_j^n be the solution generated by a consistent finite difference scheme in **conservative form** (9) with numerical flux $\mathcal{F}$. Suppose that as the grid is refined ($\Delta x, \Delta t \to 0$), the sequence of grid functions $u_{\Delta x}(x, t)$ converges boundedly almost everywhere to a function $u(x, t)$:

$$u_{\Delta x}(x, t) \longrightarrow u(x, t) \quad \text{a.e.}$$

Then $u(x, t)$ is guaranteed to be a **weak solution** of the original conservation law:

$$\frac{\partial u}{\partial t} + \frac{\partial f(u)}{\partial x} = 0$$

and in particular, all discontinuities satisfy the Rankine-Hugoniot jump condition $s = \frac{[f]}{[u]}$.

5.1 Critical Caveats of the Theorem

1. **Convergence is an Assumption:** The theorem states that *if* the scheme converges, it converges to a weak solution. It does not prove that the scheme converges (that requires stability and compactness/TVD arguments).

2. **Entropy Condition:** A weak solution is not necessarily the unique physical entropy solution. A conservative scheme can converge to an unphysical expansion shock unless an **entropy fix** is incorporated.

3. **Non-Conservative Exclusion:** The theorem holds *strictly* for conservative schemes. Non-conservative schemes can converge boundedly, but their limit is generally not a weak solution.

6 Semi-Discretization: The Method of Lines (MOL)

In classical schemes like Lax-Wendroff (Lecture 7), spatial differencing and temporal advancing are mixed together in a single formula. In modern CFD, it is far more flexible to decouple space and time via the **Method of Lines (MOL)**.

6.1 The Semi-Discrete ODE System

In the semi-discrete finite difference framework, we discretize the spatial flux derivative at each node x_j while keeping time t continuous. Let $u_j(t) \approx u(x_j, t)$:

$$\boxed{\frac{du_j(t)}{dt} = \mathcal{L}_j(\mathbf{u}) = -\frac{F_{j+1/2}(t) - F_{j-1/2}(t)}{\Delta x}} \tag{23}$$

Equation (23) defines a system of N coupled ordinary differential equations:

$$\frac{d\mathbf{u}}{dt} = \mathbf{L}(\mathbf{u}), \qquad \mathbf{u}(t) \in \mathbb{R}^N \tag{24}$$

Any suitable ODE time-stepper can now be applied to integrate $\mathbf{u}(t)$ forward in time.

6.2 Classical Semi-Discrete Numerical Fluxes

For linear advection $f(u) = au$ $(a > 0)$:

1. **First-Order Upwind Flux:**

$$F_{j+1/2} = a\,u_j \implies \frac{du_j}{dt} = -\frac{a(u_j - u_{j-1})}{\Delta x} \tag{25}$$

2. **Central Difference with Scalar Artificial Dissipation (Rusanov / Local Lax-Friedrichs Flux):**

$$F_{j+1/2} = \frac{1}{2}\left(f(u_j) + f(u_{j+1})\right) - \frac{s_{\max}}{2}\left(u_{j+1} - u_j\right) \tag{26}$$

where $s_{\max} = \max(|f'(u_j)|, |f'(u_{j+1})|)$. For linear advection ($s_{\max} = a$):

$$F_{j+1/2} = \frac{a}{2}(u_j + u_{j+1}) - \frac{a}{2}(u_{j+1} - u_j) = a\,u_j \tag{27}$$

which naturally recovers the upwind flux!

6.3 Fourier Spectral Analysis of the Spatial Operator

Let us examine the eigenvalues λ of the semi-discrete operator $\frac{d\mathbf{u}}{dt} = \mathbf{L}\mathbf{u}$ for linear advection $u_t + au_x = 0$ by substituting a harmonic Fourier mode $u_j(t) = \hat{u}(t)e^{\mathrm{i}j\xi}$:

$$\begin{aligned} \textbf{Central Differencing:} \quad & \frac{d\hat{u}}{dt} = -a\frac{e^{\mathrm{i}\xi} - e^{-\mathrm{i}\xi}}{2\Delta x}\,\hat{u} = \lambda_{\text{central}}(\xi)\,\hat{u} \\ \implies & \lambda_{\text{central}}(\xi) = -\mathrm{i}\frac{a}{\Delta x}\sin\xi \qquad \xi \in [-\pi, \pi] \end{aligned} \tag{28}$$

Remarkable Result: The spatial eigenvalues of central differencing are **purely imaginary**, lying entirely on the segment $[-\mathrm{i}a/\Delta x, +\mathrm{i}a/\Delta x]$ on the imaginary axis!

$$\begin{aligned} \textbf{Upwind Differencing:} \quad & \frac{d\hat{u}}{dt} = -a\frac{1 - e^{-\mathrm{i}\xi}}{\Delta x}\,\hat{u} = \lambda_{\text{upwind}}(\xi)\,\hat{u} \\ \implies & \lambda_{\text{upwind}}(\xi) = -\frac{a}{\Delta x}(1 - \cos\xi) - \mathrm{i}\frac{a}{\Delta x}\sin\xi \end{aligned} \tag{29}$$

The upwind eigenvalues have negative real parts $\mathrm{Re}(\lambda) \le 0$, tracing an ellipse in the stable left-half complex plane.

6.4 Time Integration via Multi-Stage Runge-Kutta Methods

When integrating $\frac{d\mathbf{u}}{dt} = \mathbf{L}(\mathbf{u})$ with an explicit ODE method, the product $\lambda\Delta t$ must fall inside the linear stability region of the time-stepper:

- **Forward Euler ($|1+z| \leq 1$):** The stability region touches the imaginary axis **only at the single point** $z = 0$. Because central differencing eigenvalues are purely imaginary ($\lambda\Delta t = -i\nu \sin\xi$), they lie outside the stability domain for all $\nu > 0$. This provides a deep ODE-spectral explanation for why FTCS is unconditionally unstable!
- **Classical Fourth-Order Runge-Kutta (RK4):** The stability region includes the imaginary axis up to $|y| \leq 2.828$. Thus, central spatial differencing combined with RK4 is **conditionally stable** up to CFL $\nu \leq 2.828$!

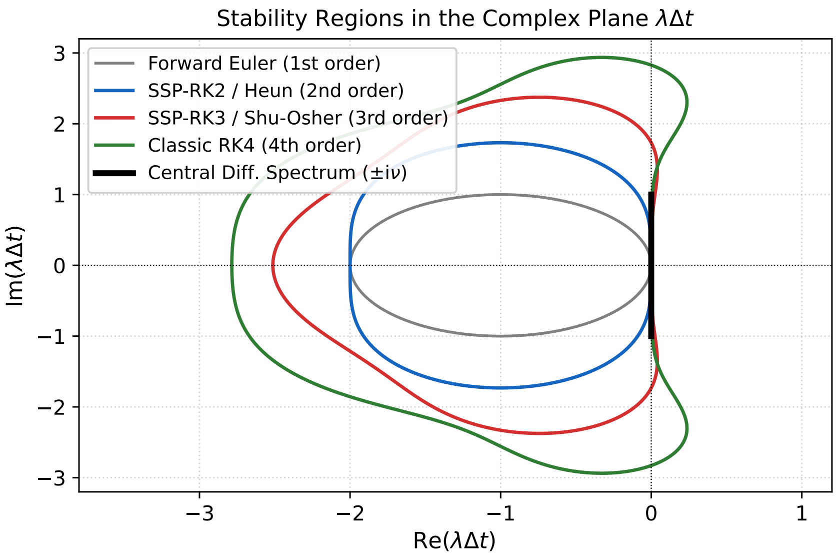

Figure 2: Linear stability regions in the complex plane $\lambda\Delta t$ for explicit time integrators. The spatial eigenvalues of central differencing lie on the imaginary axis segment $[-i\nu, +i\nu]$. Forward Euler is unstable because its stability boundary touches the imaginary axis only at $z = 0$. Multi-stage Runge-Kutta methods (SSP-RK3, RK4) enclose a finite portion of the imaginary axis, achieving conditional stability.

7 Summary: Finite Difference Schemes in Flux Form

Concept	Mathematical Formulation	Physical Significance
Conservative Difference	$\frac{u_j^{n+1}-u_j^n}{\Delta t} + \frac{F_{j+1/2}-F_{j-1/2}}{\Delta x} = 0$	Telescoping interior cancellation guarantees global conservation
Non-Conservative Defect	$S_j = -\frac{1}{2\Delta x}(u_j - u_{j-1})^2$	Creates a finite mass sink at shocks; yields wrong shock speed $s = u_R$
Lax-Wendroff Theorem	$u_{\Delta x} \to u$ as $\Delta x \to 0$	Guarantees limit is a weak solution satisfying Rankine-Hugoniot
Method of Lines (MOL)	$\frac{du_j}{dt} = -\frac{1}{\Delta x}(F_{j+1/2} - F_{j-1/2})$	Decouples spatial differencing from multi-stage ODE integration

8 Exercises

Exercises for Lecture 9

E1. Conservative Form of Lax-Wendroff: Recall the second-order Lax-Wendroff scheme for linear advection $u_t + au_x = 0$:

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

Show that this scheme can be written in explicit conservative flux form with numerical flux:

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

Verify that this numerical flux is consistent with the physical flux $f(u) = au$.

E2. Mass Loss Calculation on a Linear Profile: Suppose a steady non-conservative upwind calculation has a numerical shock profile spreading over 3 grid cells: $u_0 = 1.0$, $u_1 = 0.6$, $u_2 = 0.2$, $u_3 = 0.0$. Compute the sum of the discrete source terms $\sum_{j=1}^{3} S_j \Delta x = -\frac{1}{2}\sum_{j=1}^{3}(u_j - u_{j-1})^2$, and compare it to $-\frac{1}{2}(u_R - u_L)^2$.

E3. Imaginary Stability Bound for RK3: For the semi-discrete central difference operator $\lambda = -\mathrm{i}\frac{a}{\Delta x}\sin\xi$, show that when integrated using standard 3rd-order Runge-Kutta with amplification factor $R(z) = 1 + z + \frac{z^2}{2} + \frac{z^3}{6}$, the maximum stable CFL number is $\nu_{\max} = \frac{a\Delta t}{\Delta x} = \sqrt{3} \approx 1.732$.