Numerical methods for hyperbolic system of conservation laws

Praveen Chandrashekar
Centre for Applicable Mathematics
Tata Institute of Fundamental Research
Bangalore-560065
http://cpraveen.github.io

August 13, 2019

Contents

Chapter 1

Conservation law

Let us consider a system of coupled equations of the form

$$\frac{\partial U}{\partial t} + \sum_{j=1}^{d} \frac{\partial}{\partial x_j} F_j(U) = 0 \tag{1.1}$$

where U is called the set of *conserved variables* and F_j are the *flux vectors*

$$U = \begin{bmatrix} U_1 \\ U_2 \\ . \\ . \\ U_p \end{bmatrix} \in \mathcal{U}_{ad} \subset \mathbb{R}^p, \qquad F_j = \begin{bmatrix} F_{1j} \\ F_{2j} \\ . \\ . \\ F_{pj} \end{bmatrix} \in \mathbb{R}^p, \qquad 1 \leq j \leq d$$

Here $\mathcal{U}_{ad}$ is the set of physically admissible states and depends on the particular problem we are dealing with.

For any spatial domain $D \subset \mathbb{R}^d$ with outward unit normal vector $n = (n_1, \ldots, n_d)$ to ∂D

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_D U \mathrm{d}x + \sum_{j=1}^{d} \int_{\partial D} F_j(U) n_j \mathrm{d}S = 0$$

The above equations tell us that the total amount of U inside any domain D changes only due to the fluxes across the domain boundary. Due to this property, we say that we have a system of conservation laws.

Define the flux jacobian

$$A_j(U) = F_j'(U) = \left[\frac{\partial}{\partial U_k} F_{ij}(U) \right]_{1 \leq i,k \leq p} \in \mathbb{R}^{p \times p}$$

Definition 1.1 (Hyperbolicity)**.** *The system of conservation laws (1.1) is said to be* hyperbolic *if for every $U \in \mathcal{U}_{ad}$ and for every $\omega = (\omega_1, \ldots, \omega_d) \in \mathbb{R}^d$, the matrix*

$$A(U, \omega) = \sum_{j=1}^{d} A_j(U) \omega_j$$

1. *has p real eigenvalues $\lambda_1(U, \omega) \leq \lambda_2(U, \omega) \leq \ldots \leq \lambda_p(U, \omega)$*
2. *and p linearly independent eigenvectors $r_1(U, \omega), \ldots, r_p(U, \omega)$, i.e.,*

$$A(U, \omega) r_j(U, \omega) = \lambda_j(U, \omega) r_j(U, \omega) \qquad 1 \leq j \leq p$$

Moreover, if the eigenvalues are all distinct, then it is said to be strictly hyperbolic*. In this case, condition (2) is automatically satisfied.*

Definition 1.2 (Cauchy problem or Initial Value Problem (IVP)). *Find a function* $U : (x,t) \in \mathbb{R}^d \times [0,\infty) \to U(x,t) \in \mathcal{U}_{ad}$ *which is a solution of (1.1) and satisfies the initial condition*

$$U(x,0) = U_0(x) \qquad x \in \mathbb{R}^d$$

Example 1.3 (Riemann problem (1-D)). *This corresponds to the IVP with initial condition*

$$U(x,0) = U_0(x) = \begin{cases} U_l & x < 0 \\ U_r & x > 0 \end{cases}$$

1.1 Notion of weak solution

The solution of hyperbolic PDE can develop discontinuities even when the initial condition and other data are very smooth. We would like to allow discontinuous solutions since they are observed in many physical phenomena. The classical notion of solution which requires all derivatives appearing in the PDE to exist is not valid in this case. Instead, let us multiply the conservation law by a smooth test function with compact support in space and time

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}^d} (\partial_t U + \partial_j F_j(U)) \cdot \Phi \mathrm{d}x\mathrm{d}t = 0, \qquad \Phi \in C_c^1(\mathbb{R}^d \times \mathbb{R}^+; \mathbb{R}^p)$$

and do integration by parts to transfer derivatives onto the test function

$$\int_{\mathbb{R}^+} \int_{\mathbb{R}^d} (U \cdot \partial_t \Phi + F_j(U) \cdot \partial_j \Phi) \mathrm{d}x\mathrm{d}t + \int_{\mathbb{R}^d} U(x,0)\Phi(x,0)\mathrm{d}x = 0, \qquad \forall \Phi \in C_c^1(\mathbb{R}^d \times \mathbb{R}^+; \mathbb{R}^p)$$

This equation makes sense even if the function U is not smooth, as long as the integrals exist. If a weak solution is smooth, then it is also a classical solution.

1.2 Jump conditions

Consider a discontinuity in the solution across the surface

$$x = X(t)$$

The solution on either side of the discontinuity surface cannot be arbitrary but must satisfy the Rankine-Hugoniot jump conditions

$$[\![F_j n_j]\!] = S\,[\![U]\!]$$

where

$$S = \dot{X}_j n_j$$

is normal speed of the discontinuity surface. This can be proved starting from the definition of weak solution. Moreover, a piecewise smooth solution which satisfies the jump condition at points of discontinuity is a weak solution.

1.3 Entropy condition

The price we pay for adopting the notion of weak solutions is that we lose uniqueness. There can be multiple or an infinite number of solutions that satisfy the definition of weak solution. In order to select a unique solution among the set of all weak solutions, we need an additional principle, which is usually called an *entropy condition*. For physical problems, the laws of thermodynamics must hold and the second law says that the entropy of an isolated system cannot decrease with time. This notion can be introduced for a general system of conservation laws in the following way.

Assume that we have a strictly convex function $Q = Q(U)$ called the *entropy function* and associated *entropy fluxes* G_j such that

$$Q'(U)F_j'(U) = G_j'(U), \qquad 1 \le j \le d$$

The set of functions $(Q(U), G_j(U))$ is called an *entropy pair*. Multiply the conservation law by $Q'(U)$

$$\begin{aligned} Q'(U)\partial_t U + Q'(U)\partial_j F_j(U) &= 0 \\ \partial_t Q(U) + Q'(U)F_j'(U)\partial_j U &= 0 \\ \partial_t Q(U) + \partial_j G_j(U) &= 0 \end{aligned}$$

Hence smooth solutions of the conservation law satisfy the additional entropy conservation law. However, we cannot expect that the entropy equation holds for discontinuous solutions, since the entropy jump conditions may not be consistent with the jump conditions of conservation law.

Usually, it happens that hyperbolic equation is a simplification of a more realistic model where certain terms of small order have been dropped. An important realistic model involves parabolic terms

$$\partial_t U + \partial_j F_j(U) = \epsilon \Delta U, \qquad \epsilon > 0$$

Because $\epsilon \ll 1$ in many physical situations, we might have dropped this term from our model. However, we lose some important information, the entropy condition, when we throw away the Laplacian term. Multiplying throughout by $Q'(U)$

$$\begin{aligned} Q'(U)\partial_t U + Q'(U)F_j'(U)\partial_j U &= \epsilon Q'(U)\partial_j^2 U \\ &= \epsilon \partial_j [Q'(U)\partial_j U] - \underbrace{\epsilon [\partial_j U]^\top Q''(U)\partial_j U}_{\geq 0} \end{aligned}$$

As $Q(U)$ is strictly convex, the hessian $Q''(U)$ is symmetric, positive definite, and we obtain the inequality

$$\partial_t Q(U) + \partial_j G_j(U) \leq \epsilon \partial_j [Q'(U)\partial_j U]$$

In the limit of $\epsilon \to 0$, we obtain the entropy inequality

$$\partial_t Q(U) + \partial_j G_j(U) \leq 0$$

We will demand that this inequality must be satisfied in the weak sense. Across a discontinuity moving with speed S,

$$[\![G_j n_j]\!] \leq S [\![Q]\!]$$

must be satisfied.

1.4 Euler equations

The Euler equations model the flow of ideal gas in which there is no frictional effects. In such a fluid, the stress field is isotropic and given by a scalar pressure field p. The equations are mathematical statements of conservation laws of mass, momentum and energy. We can derive these equations by considering a fixed control volume V and writing down the conservation law. Let

ρ	mass density, i.e., mass per unit volume, Kg/m^3
$v = (v_1, v_2, v_3)$	velocity vector, m/s
p	pressure, N/m^2

The quantities (ρ, v, p) completely specify the state of the system at any particular time.

Mass conservation The mass inside V changes due to flow of fluid across the boundary of V, which we denote by ∂V. Net mass flow out of ∂V happens only due to the normal component of velocity. Hence

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_V \rho \mathrm{d}x = - \oint_{\partial V} \rho (v \cdot n) \mathrm{d}s = - \int_V \nabla \cdot (\rho v) \mathrm{d}x$$

where we used Gauss divergence theorem. Since this must hold for every control volume, we deduce the condition

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) = 0$$

or in index notation

$$\partial_t \rho + \partial_j (\rho v_j) = 0$$

Momentum equation This is just Newton's law applied to fluids. Let f be some external force field per unit volume, e.g., gravity, electromagnetic force, etc. Then Newton's law is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho v \mathrm{d}x = -\oint_{\partial V}(\rho v)(v\cdot n)\mathrm{d}s + \oint_{\partial V} pn\mathrm{d}s + \int_V \rho f \mathrm{d}x = -\int_V (\rho v\otimes v + pI)\mathrm{d}x + \int_V \rho f \mathrm{d}x$$

and hence we obtain the equation

$$\frac{\partial}{\partial t}(\rho v) + \nabla\cdot(\rho v\otimes v + pI) = \rho f$$

or in index notation

$$\partial_t(\rho v_i) + \partial_j(\rho v_i v_j + p\delta_{ij}) = \rho f_i, \qquad i = 1,2,3$$

Remark 1.4. *If gravity is acting in the negative x_3 direction, then $f = (0,0,-g)$ where $g = 9.81\ m^2/s$ is the acceleration due to gravity.*

Energy conservation The first law of thermodynamics says that the total energy of an isolated system is conserved. For a system that interacts with its environment, the total energy changes due to flow of energy into/out of V and the work done on the system. In an ideal fluid, energy flows only due to convection by the flow, i.e., we ignore effects like conduction of heat. Let

$$E = \text{ total energy per unit volume}, J/m^3$$

Then the energy conservation law is

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_V E\mathrm{d}x = -\oint_{\partial V} E(v\cdot n)\mathrm{d}x - \oint_{\partial V} p(v\cdot n)\mathrm{d}s + \int_V \rho f\cdot v\mathrm{d}x$$

Using divergence theorem, we obtain

$$\frac{\partial E}{\partial t} + \nabla\cdot(Ev + pv) = \rho f\cdot v$$

or in index notation

$$\partial_t E + \partial_j[(E+p)v_j] = \rho f_j v_j$$

The total energy is made up of internal energy, kinetic energy, gravitational energy, etc. Let us assume that only internal and kinetic energy are relevant in our problem. Then

$$E = \rho e + \frac{1}{2}\rho|v|^2$$

where $e = e(\rho, p)$ is the internal energy per unit mass. For a *calorically perfect gas*

$$e = c_v T, \qquad c_v = \text{constant}$$

combined with *ideal gas* law

$$p = \rho R T$$

yields

$$c_v = \frac{R}{\gamma - 1} \quad \Longrightarrow \quad e = \frac{p}{(\gamma-1)\rho}$$

where

$$\gamma = \frac{c_p}{c_v} > 1$$

is the ratio of specific heats at constant pressure and constant volume, respectively. For air which is mostly composed of nitrogen, $\gamma = 1.4$. The total energy is given by

$$E = \frac{p}{\gamma-1} + \frac{1}{2}\rho|v|^2 \quad \Longrightarrow \quad p = (\gamma-1)\left[E - \frac{1}{2}\rho|v|^2\right]$$

Summary of equations All of the conservation laws we have derived have a common structure

$$\frac{\partial}{\partial t}(\text{some density}) + \nabla \cdot (\text{corresponding flux}) = \text{source term}$$

We can hence write the Euler equations as

$$\partial_t U + \partial_j F_j(U) = S(U)$$

where

$$U = \begin{bmatrix} \rho \\ \rho v_1 \\ \rho v_2 \\ \rho v_3 \\ E \end{bmatrix}, \quad F_1 = \begin{bmatrix} \rho v_1 \\ p + \rho v_1^2 \\ \rho v_1 v_2 \\ \rho v_1 v_3 \\ (E+p)v_1 \end{bmatrix}, \quad F_2 = \begin{bmatrix} \rho v_2 \\ \rho v_2 v_1 \\ p + \rho v_2^2 \\ \rho v_2 v_3 \\ (E+p)v_2 \end{bmatrix}, \quad F_3 = \begin{bmatrix} \rho v_3 \\ \rho v_3 v_1 \\ \rho v_3 v_2 \\ p + \rho v_3^2 \\ (E+p)v_3 \end{bmatrix}$$

$$S = \begin{bmatrix} 0 \\ \rho f_1 \\ \rho f_2 \\ \rho f_3 \\ \rho(f_1 v_1 + f_2 v_2 + f_3 v_3) \end{bmatrix}, \qquad p = (\gamma - 1)\left[E - \frac{1}{2}\rho |v|^2\right]$$

This is a hyperbolic system for which the matrix $A(U, \omega)$ has real eigenvalues

$$v \cdot \omega - a, \quad v \cdot \omega, \quad v \cdot \omega, \quad v \cdot \omega, \quad v \cdot \omega + a$$

where

$$a = \sqrt{\frac{\gamma p}{\rho}}$$

is the speed of sound. Though there are repeated eigenvalues, we can find a full set of eigenvectors.

State space Physically admissible states must have strictly positive values of density and pressure, i.e.,

$$\rho > 0, \quad p > 0 \quad \Longrightarrow \quad \mathcal{U}_{ad} = \left\{ U \in \mathbb{R}^5 : U_1 > 0,\ U_5 - \frac{1}{2U_1}(U_2^2 + U_3^2 + U_4^2) > 0 \right\}$$

This is a convex subset of $\mathbb{R}^5$. It is very important that the numerical scheme should yield positive solutions, since otherwise, the computations will break down.

1.4.1 Non-conservative form

It is some times useful to write the Euler equations in non-conservative form. They are given by

$$\begin{aligned} \frac{\partial \rho}{\partial t} + v \cdot \nabla \rho + \rho \nabla \cdot v &= 0 \\ \frac{\partial v}{\partial t} + v \cdot \nabla v + \frac{1}{\rho} \nabla p &= 0 \\ \frac{\partial p}{\partial t} + v \cdot \nabla p + \gamma p \nabla \cdot v &= 0 \end{aligned}$$

1.4.2 Entropy equation

The thermodynamic entropy can be taken as

$$s = \frac{p}{\rho^\gamma}$$

and using the Euler equations, we can derive the entropy equation

$$\frac{\partial s}{\partial t} + v \cdot \nabla s = 0$$

This implies that entropy of a fluid element is constant. Using the continuity equation we can rewrite this in conservation form

$$\frac{\partial}{\partial t}(\rho s) + \nabla \cdot (\rho s v) = 0$$

In fact a more general equation holds. Define

$$Q = \rho H(s), \qquad G_j = \rho v_j H(s)$$

Then

$$\partial_t Q + \partial_j G_j = 0$$

holds for smooth solutions. Moreover, if $H'(s) < 0$ then $Q(U)$ is a strictly convex function[1]. For Navier-Stokes equations with Fourier law of heat conduction, the correct entropy function is $H(s) = -\ln(s)$ so that

$$Q = -\rho \ln s, \qquad G_j = -\rho v_j \ln s$$

1.5 Isothermal/isentropic gas

If the gas system behaves in such a way that

$$p = p(\rho)$$

then the energy equation is not required and we have just mass and momentum equation

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho v) &= 0 \\ \frac{\partial (\rho v)}{\partial t} + \nabla \cdot (\rho v \otimes v + p(\rho) I) &= 0 \end{aligned}$$

If the gas temperature is constant, then ideal gas assumption implies that

$$p = C\rho \quad \Longrightarrow \quad \text{Isothermal Euler equations}$$

If the gas has constant entropy, then

$$p = C\rho^\gamma \quad \Longrightarrow \quad \text{Isentropic Euler equations}$$

for some $\gamma > 1$. In both cases, we have a hyperbolic system of equations. For these models, the total energy E plays the role of a convex entropy function.

1.6 p-system

Model for one-dimensional isentropic gas dynamics in Lagrangian coordinates

$$\begin{aligned} \frac{\partial v}{\partial t} - \frac{\partial u}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} p(v) &= 0 \end{aligned}$$

v = specific volume = $\frac{1}{\rho}$
u = velocity
p = pressure = $Cv^{-\gamma}$, $\gamma \geq 1$ "Conserved" variables and flux vector

$$U = \begin{bmatrix} v \\ u \end{bmatrix}, \qquad F = \begin{bmatrix} -u \\ p(v) \end{bmatrix}, \qquad \mathcal{U}_{ad} = \{(v,u) \in \mathbb{R}^2 : v > 0\}$$

Flux jacobian

$$A = F'(U) = \begin{bmatrix} 0 & -1 \\ p'(v) & 0 \end{bmatrix}$$

Eigenvalues are real and distinct provided $p'(v) < 0$

$$\lambda_1 = -\sqrt{-p'(v)}, \qquad \lambda_2 = \sqrt{-p'(v)}$$

[1] See Bouchut

1.7 Linearized Euler equations

In some problems, the flow may be constant with only small perturbations around a constant state (ρ_0, v_0, p_0). In this case, the flow may be treated as isentropic

$$p = p(\rho, s) = C(s)\rho^\gamma$$

and we can ignore the energy equation. Then

$$\rho = \rho_0 + \rho', \qquad v = v_0 + v', \qquad p = p(\rho_0) + p' = p_0 + (\gamma p_0/\rho_0)\rho'$$

where

$$|\rho'| \ll \rho_0, \qquad |v'| \ll |v_0|$$

Ignoring terms quadratic in the perturbations, the mass and momentum equations take the form

$$\begin{aligned} \frac{\partial \rho'}{\partial t} + v_0 \cdot \nabla \rho' + \rho' \nabla \cdot v_0 &= 0 \\ \frac{\partial v'}{\partial t} + v_0 \cdot \nabla v' + v' \cdot \nabla v_0 + \frac{a_0^2}{\rho_0} \nabla \rho' &= 0, \quad a_0 = \sqrt{\gamma p_0/\rho_0} \end{aligned}$$

This is a hyperbolic system.

We can also take the linearized pressure equation

$$\frac{\partial p'}{\partial t} + v_0 \cdot \nabla p' + \gamma p_0 \nabla \cdot v' = 0$$

instead of the density equation.

Sound waves Let us consider the background state to be stationary, $v_0 = 0$. Then

$$\frac{\partial v'}{\partial t} + \frac{1}{\rho_0} \nabla p' = 0, \qquad \frac{\partial p'}{\partial t} + \gamma p_0 \nabla \cdot v' = 0$$

Differentiating pressure equation wrt time

$$\partial_{tt} p' + \gamma p_0 \nabla \cdot \partial_t v' = 0 \quad \implies \quad \partial_{tt} p' = a_0^2 \Delta p'$$

The pressure perturbations are governed by the wave equations and the perturbations propagate with speed a_0 which is the sound speed.

1.8 Shallow water equations

1.9 Maxwell's equations

Chapter 2

FVM in 1-D

The finite volume method is based on the integral form of the conservation laws and gives an approximation to the weak solution. Consider a 1-D conservation law system

$$U_t + F(U)_x = 0$$

with some initial condition

$$U(x,0) = U_0(x)$$

We partition the domain into disjoint cells

$$I_{j+\frac{1}{2}} = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}], \qquad \Delta x = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}, \qquad x_j = \frac{1}{2}(x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})$$

Integrate the conservation law over one cell

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{I_{j+\frac{1}{2}}} U(x,t)\mathrm{d}x + F(x_{j+\frac{1}{2}},t) - F(x_{j-\frac{1}{2}},t) = 0$$

Define the cell average value

$$U_j(t) = \frac{1}{\Delta x}\int_{I_{j+\frac{1}{2}}} U(x,t)\mathrm{d}x$$

We have to make some approximation to estimate the fluxes. Suppose we have some method to do this

$$F(x_{j+\frac{1}{2}},t) \approx F_{j+\frac{1}{2}}(t) = F(\ldots, U_j(t), U_{j+1}(t), \ldots)$$

and $F_{j+\frac{1}{2}}$ is called a *numerical flux function*. We obtain the *semi-discrete finite volume scheme*

$$\Delta x\frac{\mathrm{d}U_j}{\mathrm{d}t} + F_{j+\frac{1}{2}} - F_{j-\frac{1}{2}} = 0$$

2.1 Basic scheme

The numerical solution is defined by the cell averages, which give a piecewise constant approximation. Suppose

$$F_{j+\frac{1}{2}} = F(U_j, U_{j+1})$$

where the numerical flux is constant in the sense that

$$F(U,U) = F(U) \quad \forall U \in \mathcal{U}_{ad}$$

We partition the time axis into intervals Δt; we will compute the numerical solution at the time levels

$$t_n = n\Delta t, \qquad n = 0, 1, 2, \ldots$$

and denote the cell average by

$$U_j^n \approx U(x, t_n) \qquad x \in I_{j+\frac{1}{2}}$$

The superscript denotes the time level and is not a power. The time derivative can be approximated by a forward difference formula in time, also called *forward Euler scheme*

$$\frac{\mathrm{d}U_j}{\mathrm{d}t}(t_n) \approx \frac{U_j^{n+1} - U_j^n}{\Delta t}$$

Then the fully discrete scheme is

$$\Delta x \frac{U_j^{n+1} - U_j^n}{\Delta t} + F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n = 0$$

Since the solution is known at time level n, we put all these quantities on the right hand side to obtain the *finite volume update equation*

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}[F_{j+\frac{1}{2}}^n - F_{j-\frac{1}{2}}^n], \qquad n = 0, 1, 2, \ldots$$

Given the initial condition

$$U_j^0 = U_0(x_j) \qquad \text{or} \qquad U_j^0 = \frac{1}{\Delta x}\int_{I_{j+\frac{1}{2}}} U_0(x)\mathrm{d}x$$

we repeatedly apply the update equation to generate the solution at future times.

Remark 2.1. *How to compute the numerical flux ? This is the key question in the finite volume method. The natural thing to try is to perform some interpolation, for example*

$$F_{j+\frac{1}{2}} = \frac{1}{2}[F_j + F_{j+1}]$$

but this leads to a central difference scheme

$$\Delta x \frac{dU_j}{dt} + \frac{1}{2}(F_{j+1} - F_{j-1}) = 0$$

which is unstable for non-linear hyperbolic problems. In the remaining chapters, we will see better ways to approximate the flux. The piece-wise constant solution representation creates a Riemann problem at each cell face. We can solve the Riemann problem exactly or approximately to compute the flux.

2.2 Local truncation error

The local truncation error is useful to make some conclusions on the order of accuracy that can be expected from the scheme. Let us assume that we have a consistent numerical flux function and also that it is a smooth function of its two arguments. The local truncation error is obtained by substituting a smooth exact solution into the numerical scheme

$$T_h = \frac{U(x, t + \Delta t) - U(x, t)}{\Delta t} + \frac{F(U(x,t), U(x + \Delta x, t)) - F(U(x - \Delta x, t), U(x, t))}{\Delta x}$$

We perform Taylor expansion around (x, t). The first term is

$$\frac{U(x, t + \Delta t) - U(x, t)}{\Delta t} = \partial_t U(x, t) + \mathcal{O}(\Delta t)$$

Define shorthand notation

$$U = U(x, t), \qquad \partial_x U = \partial_x U(x, t)$$

Let us call the two arguments of $F(\cdot,\cdot)$ are X, Y respectively. Then, differentiating the flux consistency condition, we get

$$\frac{\partial}{\partial X}F(U,U) + \frac{\partial}{\partial Y}F(U,U) = \frac{\partial}{\partial U}F(U)$$

Now

$$\begin{aligned} F(U(x,t),U(x+\Delta x,t)) &= F(U, U+\Delta x \partial_x U + \mathcal{O}\left(\Delta x^2\right)) \\ &= F(U,U) + \Delta x \frac{\partial}{\partial Y}F(U,U)\cdot \partial_x U + \mathcal{O}\left(\Delta x^2\right) \\ &= F(U) + \Delta x \frac{\partial}{\partial Y}F(U,U)\cdot \partial_x U + \mathcal{O}\left(\Delta x^2\right) \end{aligned}$$

and similarly

$$F(U(x-\Delta x,t),U(x,t)) = F(U) - \Delta x \frac{\partial}{\partial X}F(U,U)\cdot \partial_x U + \mathcal{O}\left(\Delta x^2\right)$$

The flux difference term becomes

$$\begin{aligned} & \frac{F(U(x,t),U(x+\Delta x,t)) - F(U(x-\Delta x,t),U(x,t))}{\Delta x} \\ = & \left[\frac{\partial}{\partial X}F(U,U) + \frac{\partial}{\partial Y}F(U,U)\right]\partial_x U + \mathcal{O}\left(\Delta x\right) \\ = & \frac{\partial}{\partial U}F(U)\cdot\partial_x U + \mathcal{O}\left(\Delta x\right) \\ = & \partial_x F + \mathcal{O}\left(\Delta x\right) \end{aligned}$$

Hence the truncation error is

$$T_h = \partial_t U + \partial_x F + \mathcal{O}\left(\Delta t\right) + \mathcal{O}\left(\Delta x\right) = \mathcal{O}\left(\Delta t\right) + \mathcal{O}\left(\Delta x\right)$$

We get first order accuracy just from a smooth, consistent flux.

2.3 Implementation of scheme

Let us write the scheme in residual form

$$\Delta x \frac{\mathrm{d}U_j}{\mathrm{d}t} + R_j = 0, \qquad R_j = F_{j+\frac{1}{2}} - F_{j-\frac{1}{2}}$$

The update equation is

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} R_j^n$$

Note that the flux $F_{j+\frac{1}{2}}$ appears in R_j as shown above and also in R_{j+1}

$$R_{j+1} = F_{j+\frac{3}{2}} - F_{j+\frac{1}{2}}$$

but with opposite sign. Since the flux computation can be expensive, we will compute each flux only once and add/subtract it from the two residuals. This requires that we should loop over the faces. The algorithm is given in (1). To implement this method, we need arrays to store the solution and the residual. Here is a Fortran-type pseudo-code. It uses Neumann boundary conditions at both ends of the domain.

Algorithm 1: First order finite volume scheme

Allocate memory for all variables;
Set initial condition;
Set time counter $t = 0$;
while $t < T$ **do**
  Compute time step Δt;
  Set residual to zero;
  for *each face* **do**
    Compute flux;
    Add flux to left residual;
    Subtract flux from right residual;
  end
  Update solution to next time level;
  $t = t + \Delta t$;
end

Listing 2.1: First order FVM

```
integer :: nvar=3, nx=100, j
real    :: U(nvar,nx), R(nvar,nx), flux(nvar), t, dt, T=1.0, &
           cfl=0.9
call set_initial_condition(U) ! Set initial condition in U
t = 0.0
do while(t < T)
   dt = compute_dt(cfl, U) ! Compute time step using CFL condition
   R = 0
   do j=0,nx
      if(j == 0)then ! first face
         call num_flux(U(:,1), U(:,1), flux)
         R(:,1) = R(:,1) - flux ! subtract from right cell
      else if(j == nx)then ! last face
         call num_flux(U(:,nx), U(:,nx), flux)
         R(:,nx) = R(:,nx) + flux ! add to left cell
      else ! interior faces
         call num_flux(U(:,j), U(:,j+1), flux)
         R(:,j  ) = R(:,j  ) + flux ! add to left cell
         R(:,j+1) = R(:,j+1) - flux ! subtract from right cell
      endif
   enddo
   U = U - dt * R
enddo
```

Chapter 3

Linear hyperbolic system in 1-D

Let us consider a system of m conservation laws

$$U_t + F(U)_x = 0$$

where the flux is linear

$$F(U) = AU, \qquad A \in \mathbb{R}^{m\times m} \text{ constant}$$

We are interested in the IVP

$$U_t + AU_x = 0, \qquad U(x,0) = U^0(x)$$

The system $U_t + AU_x = 0$ is said to be *hyperbolic* provided

- A has m real eigenvalues
$$\lambda_1 < \lambda_2 < \ldots < \lambda_m$$
- The eigenvectors form a basis for $\mathbb{R}^m$.

It is not necessary that the eigenvalues should be distinct. But we need linearly independent eigenvectors.

3.1 Where are the waves ?

By definition of eigenvalues and eigenvectors

$$Ar_k = \lambda_k r_k, \qquad r_k \in \mathbb{R}^m$$

We can stack the equations side by side

$$A[r_1, \ldots, r_m] = [r_1, \ldots, r_m]\text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m)$$

Define the matrices

$$R = [r_1, r_2, \ldots, r_m] \in \mathbb{R}^{m\times m}, \qquad \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^{m\times m}$$

we get

$$AR = R\Lambda, \qquad A = R\Lambda R^{-1}$$

The inverse of R exists since its columns are linearly independent and so it has full rank. Now we can modify the conservation law as

$$\frac{\partial U}{\partial t} + R\Lambda R^{-1}\frac{\partial U}{\partial x} = 0 \qquad \Longrightarrow \qquad R^{-1}\frac{\partial U}{\partial t} + \Lambda R^{-1}\frac{\partial U}{\partial x} = 0$$

Define the **characteristic variables** by

$$W = R^{-1}U \qquad \Longrightarrow \qquad \frac{\partial W}{\partial t} + \Lambda\frac{\partial W}{\partial x} = 0$$

These equations become decoupled

$$\frac{\partial W_i}{\partial t} + \lambda_i \frac{\partial W_i}{\partial x} = 0, \qquad i = 1, 2, \ldots, m$$

We get m linear advection equations !!! This also shows the importance of eigenvalues; they represent the wave speeds !!!

3.2 General solution

The initial condition for W is

$$W(x, 0) = W^0(x) = R^{-1} U^0(x)$$

and the solution is given by

$$W_i(x, t) = W_i^0(x - \lambda_i t) = [R^{-1} U^0(x - \lambda_i t)]_i$$

Transforming, we get the solution U

$$\begin{aligned} U(x,t) &= RW(x,t) \\ &= \sum_{i=1}^{m} W_i(x,t) r_i \qquad \text{(linear combination of eigenvectors)} \\ &= \sum_{i=1}^{m} W_i^0(x - \lambda_i t) r_i \\ &= \sum_{i=1}^{m} [R^{-1} U^0(x - \lambda_i t)]_i r_i \end{aligned}$$

Geometrical interpretation We have m characteristic curves $\frac{\mathrm{d}x}{\mathrm{d}t} = \lambda_i$. Through any point (x, t) draw all the m characteristic curves until they hit the line $t = 0$ on which we know the initial condition. The characteristic with slope λ_i hits the initial line at $x_i = x - \lambda_i t$ and we take the value $W_i(x_i, 0)$ from this point. Then we know the p values

$$W_1(x, t) = W_1(x_1, 0), \quad \ldots \quad , W_p(x, t) = W_p(x_p, 0)$$

and we now convert back to conserved variables by multiplying with R.

3.3 Riemann problem

Consider initial condition

$$U^0(x) = \begin{cases} U_l & x < 0 \\ U_r & x > 0 \end{cases}$$

Initial condition for W

$$W^0(x) = R^{-1} U^0(x) = \begin{cases} R^{-1} U_l & x < 0 \\ R^{-1} U_r & x > 0 \end{cases} =: \begin{cases} W_l & x < 0 \\ W_r & x > 0 \end{cases}$$

Solution for W_i, $i = 1, 2, \ldots, m$

$$W_i(x, t) = W_i^0(x - \lambda_i t) = \begin{cases} W_{l,i} & x/t < \lambda_i \\ W_{r,i} & x/t > \lambda_i \end{cases}$$

Solution U

$$\begin{aligned} U(x,t) &= \sum_i W_i^0(x - \lambda_i t) r_i \\ &= \sum_{i : x/t < \lambda_i} W_{l,i} r_i + \sum_{i : x/t > \lambda_i} W_{r,i} r_i \end{aligned}$$

Since we ordered the eigenvalues, there is an $n = n(x/t)$ such that

$$\lambda_n < \frac{x}{t} < \lambda_{n+1}$$

and the above solution can also be written as

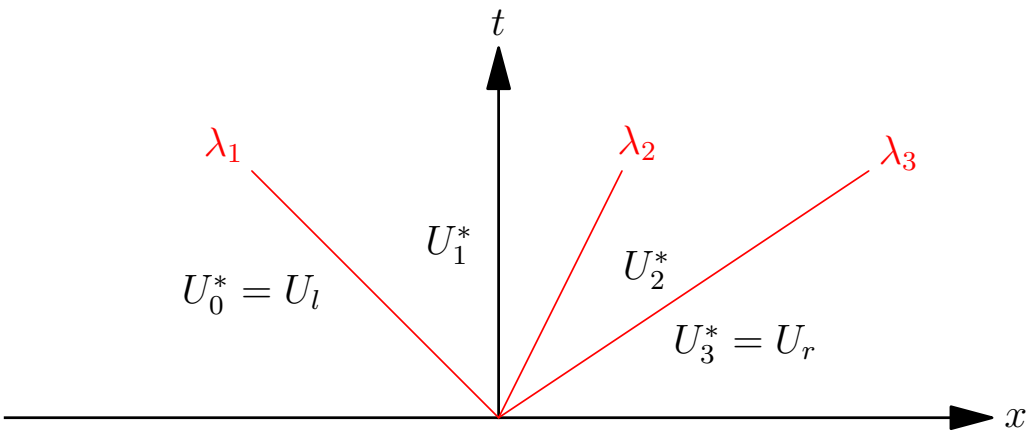
$$U(x,t) = \sum_{i=1}^{n(x/t)} W_{r,i} r_i + \sum_{i=n(x/t)+1}^{m} W_{l,i} r_i$$

We see that the solution of the Riemann problem is self-similar in the sense that it depends only on the ratio x/t, i.e.,

$$U(x,t) = U_R(x/t)$$

Example 3.1. *System of 3 equations* ($m = 3$)

$$U(x,t) = \begin{cases} U_l & x/t < \lambda_1 \\ U_1^* & \lambda_1 < x/t < \lambda_2 \\ U_2^* & \lambda_2 < x/t < \lambda_3 \\ U_r & x/t > \lambda_3 \end{cases}$$



The intermediate states are given by

$$\begin{aligned} U_1^* &= W_{r,1} r_1 + W_{l,2} r_2 + W_{l,3} r_3 \\ U_2^* &= W_{r,1} r_1 + W_{r,2} r_2 + W_{l,3} r_3 \end{aligned}$$

Remark 3.2. *Show that the the jump in the intermediate states satisfies*

$$U_i^* - U_{i-1}^* = (W_{r,i} - W_{l,i}) r_i, \qquad i = 1, 2, \ldots, m$$

with the convention that $U_0^* = U_l$ and $U_m^* = U_r$. Hence the initial discontinuity breaks into m discontinuity waves which propagate at speeds λ_i, $i = 1, \ldots, m$. The i'th wave will appear in the solution if the corresponding amplitude $|W_{r,i} - W_{l,i}| > 0$. Moreover, across each wave, the jump condition

$$F(U_i^*) - F(U_{i-1}^*) = A(U_i^* - U_{i-1}^*) = (W_{r,i} - W_{l,i}) A r_i = (W_{r,i} - W_{l,i}) \lambda_i r_i = \lambda_i (U_i^* - U_{i-1}^*)$$

is satisfied.

Solution on $x/t = 0$: For future use in finite volume method, we compute the solution along $x/t = 0$. It is given by

$$U_R(0) = \sum_{i:\lambda_i>0} W_{l,i} r_i + \sum_{i:\lambda_i<0} W_{r,i} r_i$$

and the corresponding flux is

$$F(U_R(0)) = AU_R(0) = \sum_{i:\lambda_i>0} \lambda_i W_{l,i} r_i + \sum_{i:\lambda_i<0} \lambda_i W_{r,i} r_i$$

This can be re-written as

$$F(U_R(0)) = \sum_i \lambda_i^+ W_{l,i} r_i + \sum_i \lambda_i^- W_{r,i} r_i$$

where we have defined

$$\begin{aligned} \lambda^+ &= \max(0,\lambda) = \frac{1}{2}(\lambda + |\lambda|) \geq 0 \\ \lambda^- &= \min(0,\lambda) = \frac{1}{2}(\lambda - |\lambda|) \leq 0 \end{aligned}$$

Define diagonal matrix

$$\Lambda^\pm = \text{diag}(\lambda_1^\pm, \ldots, \lambda_n^\pm)$$

The above flux can also be written as

$$\begin{aligned} F(U_R(0)) &= R\Lambda^+ W_l + R\Lambda^- W_r \\ &= R\Lambda^+ R^{-1} U_l + R\Lambda^- R^{-1} U_r \\ &= A^+ U_l + A^- U_r \end{aligned}$$

where

$$A^\pm = R\Lambda^\pm R^{-1}$$

Another formula is obtained using the second definition of $\lambda^\pm$;

$$\begin{aligned} F(U_R(0)) &= \sum_i \lambda_i^+ W_{l,i} r_i + \sum_i \lambda_i^- W_{r,i} r_i \\ &= \sum_i \frac{1}{2}(\lambda_i + |\lambda_i|) W_{l,i} r_i + \sum_i \frac{1}{2}(\lambda_i - |\lambda_i|) W_{r,i} r_i \\ &= \frac{1}{2}(F_l + F_r) - \frac{1}{2}\sum_i |\lambda_i| (W_{r,i} - W_{l,i}) r_i \\ &= \frac{1}{2}(F_l + F_r) - \frac{1}{2} R|\Lambda|(W_r - W_l) \\ &= \frac{1}{2}(F_l + F_r) - \frac{1}{2}|A|(U_r - U_l), \qquad |A| = R|\Lambda|R^{-1} \end{aligned}$$

Flux difference form Let s be such that

$$\lambda_s < 0 < \lambda_{s+1}$$

Then solution on $x = 0$ is given by

$$U_R(0) = U_s^* = U_l + \sum_{i=1}^{s} (U_i^* - U_{i-1}^*) = U_r - \sum_{i=s+1}^{m} (U_i^* - U_{i-1}^*)$$

so that the flux on $x = 0$ is

$$F(U_R(0)) = AU_R(0) = F(U_l) + \sum_{i=1}^{s} \lambda_i (U_i^* - U_{i-1}^*) = F(U_r) - \sum_{i=s+1}^{m} \lambda_i (U_i^* - U_{i-1}^*)$$

This can also be written as

$$F(U_R(0)) = F(U_l) + \underbrace{\sum_{i=1}^{m} \lambda_i^-(U_i^* - U_{i-1}^*)}_{(\Delta F)^-} = F(U_r) - \underbrace{\sum_{i=1}^{m} \lambda_i^+(U_i^* - U_{i-1}^*)}_{(\Delta F)^+}$$

The flux difference

$$\Delta F = F_r - F_l = \sum_{i=1}^{m} \lambda_i (U_i^* - U_{i-1}^*) = (\Delta F)^- + (\Delta F)^+$$

is split into two parts, $(\Delta F)^-$ due to left moving waves and $(\Delta F)^+$ due to right moving waves.

3.4 Upwind scheme

The system of conservation laws can be transformed to a set of decoupled linear advection equations

$$\frac{\partial W_i}{\partial t} + \lambda_i \frac{\partial W_i}{\partial x} = 0, \quad 1 \le i \le n$$

which represent waves moving with velocity λ_i. We can try to build a scheme for the system of conservation laws by applying the upwind scheme to the above advection equations. For the grid point j we have

$$\frac{W_{i,j}^{n+1} - W_{i,j}^n}{\Delta t} + \lambda_i^+ \frac{W_{i,j}^n - W_{i,j-1}^n}{\Delta x} + \lambda_i^- \frac{W_{i,j+1}^n - W_{i,j}^n}{\Delta x} = 0, \quad i = 1, 2, \ldots, m$$

or using matrix-vector notation,

$$\frac{W_j^{n+1} - W_j^n}{\Delta t} + \Lambda^+ \frac{W_j^n - W_{j-1}^n}{\Delta x} + \Lambda^- \frac{W_{j+1}^n - W_j^n}{\Delta x} = 0$$

Multiplying by R from the left, we transform back to the conserved variables U, the above scheme becomes

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + A^+ \frac{U_j^n - U_{j-1}^n}{\Delta x} + A^- \frac{U_{j+1}^n - U_j^n}{\Delta x} = 0$$

CIR splitting: We could have obtained this scheme using the CIR splitting technique; separating the Jacobian A into positive and negative parts

$$A = A^+ + A^-, \qquad A^\pm = R\Lambda^\pm R^{-1}, \qquad \frac{\partial U}{\partial t} + A^+ \frac{\partial U}{\partial x} + A^- \frac{\partial U}{\partial x} = 0$$

and using backward and forward differencing for the A^+ and A^- terms respectively,

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + A^+ \frac{U_j^n - U_{j-1}^n}{h} + A^- \frac{U_{j+1}^n - U_j^n}{h} = 0$$

we obtain exactly the upwind scheme.

Flux splitting scheme: Another way to arrive at this scheme is to start with flux splitting. The eigenvalue splitting leads to the flux splitting

$$F = A^+ U + A^- U = F^+ + F^-$$

so that conservation law can be written as

$$\frac{\partial U}{\partial t} + \frac{\partial F^+}{\partial x} + \frac{\partial F^-}{\partial x} = 0$$

Since

$$\frac{\partial F^+}{\partial U} = A^+ \geq 0, \qquad \frac{\partial F^-}{\partial U} = A^- \leq 0$$

we use backward and forward differencing for the F^+ and F^- terms respectively.

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{F_j^+ - F_{j-1}^+}{h} + \frac{F_{j+1}^- - F_j^-}{h} = 0$$

We can write this as a finite volume scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{(F_j^+ + F_{j+1}^-) - (F_{j-1}^+ + F_j^-)}{h} = 0$$

with the numerical flux

$$F_{j+\frac{1}{2}} = F_j^+ + F_{j+1}^- = A^+ U_j + A^- U_{j+1} = \frac{1}{2}(F_j + F_{j+1}) - \frac{1}{2}|A|(U_{j+1} - U_j)$$

We can compare this flux to the upwind flux for linear advection equation; the factor $|a|$ has been replaced by the matrix $|A|$.

Upwind property This scheme ***upwind property*** in the following sense: If all eigenvalues are positive, i.e., all the waves are moving to the right, then

$$F_j^+ = F_j, \qquad F_{j+1}^- = 0 \qquad \Longrightarrow \qquad F_{j+\frac{1}{2}} = F_j$$

The flux is entirely determined from the left state U_j which is physically meaningful. Conversely if all eigenvalues are negative, then

$$F_j^+ = 0, \qquad F_{j+1}^- = F_{j+1} \qquad \Longrightarrow \qquad F_{j+\frac{1}{2}} = F_{j+1}$$

the flux is now entirely determined from the right state U_{j+1}.

Chapter 4

Euler equations in 1-D

The Euler equations in 1-D are given by

$$U_t + F(U)_x = 0$$

where

$$U = \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix}, \quad F(U) = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ (E+p)u \end{bmatrix}$$

Here

$$\rho = \text{density}, \quad u = \text{velocity}, \quad p = \text{pressure}$$

$$E = \text{total energy per unit volume} = \rho e + \frac{1}{2}\rho u^2$$

$$\rho e = \text{internal energy per unit volume}$$

$$e = \text{internal energy per unit mass}$$

The pressure p is related to the internal energy e by the caloric equation of state $p = p(\rho, e)$; for a calorically ideal gas, $p = (\gamma - 1)\rho e$, so that

$$p = (\gamma - 1)\left[E - \frac{1}{2}\rho u^2\right]$$

4.1 Flux Jacobian

The flux jacobian $A \in \mathbb{R}^{3\times 3}$ is defined as

$$A(U) := F'(U) = \frac{\partial F}{\partial U}$$

The jacobian can be computed by first expressing the flux vector in terms of the conserved variables. The pressure is given by

$$p = (\gamma - 1)\left[E - \frac{(\rho u)^2}{2\rho}\right] = (\gamma - 1)\left[U_3 - \frac{U_2^2}{2U_1}\right]$$

Then the flux can be written as

$$F(U) = \begin{bmatrix} U_2 \\ p(U) + U_2^2/U_1 \\ (U_3 + p(U))U_2/U_1 \end{bmatrix} = \begin{bmatrix} U_2 \\ \frac{1}{2}(3-\gamma)\frac{U_2^2}{U_1} + (\gamma-1)U_3 \\ \gamma\frac{U_2 U_3}{U_1} - \frac{1}{2}(\gamma-1)\frac{U_2^3}{U_1^2} \end{bmatrix}$$

The jacobian components are then given by

$$A_{ij} = \frac{\partial F_i}{\partial U_j}, \quad 1 \le i, j \le 3$$

and for the Euler equations we obtain

$$A(U) = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{2}(3-\gamma)\left(\frac{U_2}{U_1}\right)^2 & (3-\gamma)\frac{U_2}{U_1} & \gamma-1 \\ -\gamma\frac{U_2U_3}{U_1^2} + (\gamma-1)\left(\frac{U_2}{U_1}\right)^3 & \gamma\frac{U_3}{U_1} - \frac{3}{2}(\gamma-1)\left(\frac{U_2}{U_1}\right)^2 & \gamma\frac{U_2}{U_1} \end{bmatrix}$$

Defining the *total specific enthalpy* H

$$H = (E+p)/\rho = \frac{a^2}{\gamma-1} + \frac{1}{2}u^2, \qquad a = \sqrt{\frac{\gamma p}{\rho}} = \text{sound speed}$$

the jacobian matrix can be written as

$$A(U) = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3)u^2 & (3-\gamma)u & \gamma-1 \\ u[\frac{1}{2}(\gamma-1)u^2 - H] & H-(\gamma-1)u^2 & \gamma u \end{bmatrix}$$

4.2 Hyperbolicity

The flux Jacobian A has eigenvalues

$$\lambda_1 = u-a, \qquad \lambda_2 = u, \qquad \lambda_3 = u+a$$

The corresponding right eigenvectors are

$$r_1 = \begin{bmatrix} 1 \\ u-a \\ H-ua \end{bmatrix}, \quad r_2 = \begin{bmatrix} 1 \\ u \\ \frac{1}{2}u^2 \end{bmatrix}, \quad r_3 = \begin{bmatrix} 1 \\ u+a \\ H+ua \end{bmatrix}$$

which are linearly independent. Thus the time dependent Euler equations are hyperbolic. The flux Jacobian can be expressed in terms of the eigenvalues and eigenvectors by the following diagonal decomposition

$$A(U) = R(U)\Lambda(U)R^{-1}(U)$$

where the matrix R has the eigenvectors on its columns

$$R = [r_1, r_2, r_3] \qquad \text{and} \qquad \Lambda = \text{diag}(\lambda_1, \lambda_2, \lambda_3)$$

The rows of R^{-1} are the left eigenvectors of A; the left and right eigenvectors are mutually orthogonal. In fact, since

$$R^{-1} = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix} = \begin{bmatrix} \frac{\gamma-1}{4}\frac{u^2}{a^2} + \frac{u}{2a} & -\frac{\gamma-1}{2}\frac{u}{a^2} - \frac{1}{2a} & \frac{\gamma-1}{2a^2} \\ 1 - \frac{\gamma-1}{2}\frac{u^2}{a^2} & (\gamma-1)\frac{u}{a^2} & -\frac{\gamma-1}{a^2} \\ \frac{\gamma-1}{4}\frac{u^2}{a^2} - \frac{u}{2a} & -\frac{\gamma-1}{2}\frac{u}{a^2} + \frac{1}{2a} & \frac{\gamma-1}{2a^2} \end{bmatrix}$$

we have $l_i r_j = \delta_{ij}$.

4.3 Homogeneity property

If the equation of state $p = p(\rho, e)$ satisfies

$$p(\alpha\rho, e) = \alpha p(\rho, e) \qquad \text{for every} \quad \alpha > 0$$

then it is easy to check[1] that the flux vector satisfies[2]

$$F(\alpha U) = \alpha F(U) \quad \text{for every} \quad \alpha > 0$$

[1] See [**?**]

[2] We say that F is homogeneous of degree one.

Differentiating wrt α

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\alpha}F(\alpha U) =& \frac{\mathrm{d}}{\mathrm{d}\alpha}[\alpha F(U)]\\ F'(\alpha U)\frac{\mathrm{d}}{\mathrm{d}\alpha}(\alpha U) =& F(U)\\ A(\alpha U)U =& F(U)\end{aligned}$$

and setting $\alpha = 1$ we get

$$F(U) = F'(U)U = A(U)U$$

which is called the **homogeneity property**. It can also be directly checked by computing the product $A(U)U$. This special property of the Euler equations is used in the Steger-Warming flux splitting scheme and in the Beam-Warming scheme.

4.4 Primitive form

The primitive variables are

$$V = [\rho,\ u,\ p]^\top$$

The transformation between U and V is given by

$$\begin{array}{l|l} U_1 = \rho & \rho = U_1 \\ U_2 = \rho u & u = U_2/U_1 \\ U_3 = p/(\gamma-1) + \rho u^2/2 & p = (\gamma-1)(U_3 - U_2^2/(2U_1)) \end{array}$$

Defining the jacobian $M := U'(V)$, the Euler equations can be transformed to the primitive form

$$\begin{aligned}\frac{\partial U}{\partial V}\frac{\partial V}{\partial t} + \frac{\partial F}{\partial U}\frac{\partial U}{\partial V}\frac{\partial V}{\partial x} &= 0\\ M\frac{\partial V}{\partial t} + AM\frac{\partial V}{\partial x} &= \\ \frac{\partial V}{\partial t} + \tilde{A}\frac{\partial V}{\partial x} &= 0, \quad \tilde{A} = M^{-1}AM\end{aligned}$$

The Jacobian of the transformation is

$$M = \begin{bmatrix} 1 & 0 & 0 \\ u & \rho & 0 \\ \frac{u^2}{2} & \rho u & \frac{1}{\gamma-1} \end{bmatrix}$$

This matrix is invertible since $det(M) = \rho/(\gamma-1) > 0$. The matrix $\tilde{A}$ can be computed as

$$\tilde{A} = \begin{bmatrix} u & \rho & 0 \\ 0 & u & \frac{1}{\rho} \\ 0 & \rho a^2 & u \end{bmatrix}$$

whose eigenvalues are again $u-a$, u and $u+a$. This is obvious since A and $\tilde{A}$ are related by a similarity transformation. It is easier to compute the eigenvalues/vectors of $\tilde{A}$ since it has a simpler structure. The eigenvectors of A can be obtained as follows.

$$\begin{aligned}\tilde{A}\tilde{r} &= \lambda\tilde{r}\\ M^{-1}AM\tilde{r} &= \lambda\tilde{r}\\ A(M\tilde{r}) &= \lambda(M\tilde{r})\end{aligned}$$

Hence $r = M\tilde{r}$ is the eigenvector of A corresponding to eigenvalue λ.

The primitive form can also be derived by manipulating the conservation form in the following way. The continuity equation gives

$$\boxed{\frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + \rho\frac{\partial u}{\partial x} = 0}$$

which is in the primitive form. The momentum equation can be written as

$$\rho\frac{\partial u}{\partial t} + u\frac{\partial \rho}{\partial t} + u^2\frac{\partial \rho}{\partial x} + 2\rho u\frac{\partial u}{\partial x} + \frac{\partial p}{\partial x} = 0$$

Using the continuity equation to eliminate the time derivative of ρ we have

$$\boxed{\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho}\frac{\partial p}{\partial x} = 0}$$

Similarly, from the energy equation and eliminating ρ_t and u_t, we obtain

$$\boxed{\frac{\partial p}{\partial t} + \rho a^2\frac{\partial u}{\partial x} + u\frac{\partial p}{\partial x} = 0}$$

Writing the three equations as a system, we have

$$\frac{\partial}{\partial t}\begin{bmatrix}\rho\\ u\\ p\end{bmatrix} + \begin{bmatrix}u & \rho & 0\\ 0 & u & \frac{1}{\rho}\\ 0 & \rho a^2 & u\end{bmatrix}\frac{\partial}{\partial x}\begin{bmatrix}\rho\\ u\\ p\end{bmatrix} = 0$$

which immediately gives us the matrix $\tilde{A}$.

4.5 Entropy equation

Consider the quantity $s = p/\rho^\gamma$. Using the primitive form of the Euler equations, we can show that

$$\begin{aligned}\frac{\partial s}{\partial t} &= \frac{1}{\rho^\gamma}\left(\frac{\partial p}{\partial t} - a^2\frac{\partial \rho}{\partial t}\right)\\ &= -u\frac{1}{\rho^\gamma}\left(\frac{\partial p}{\partial x} - a^2\frac{\partial \rho}{\partial x}\right)\\ &= -u\frac{\partial s}{\partial x}\end{aligned}$$

which gives us an additional equation, atleast for smooth solutions

$$\frac{\partial s}{\partial t} + u\frac{\partial s}{\partial x} = 0$$

This equation tells us that the quantity s which is the **entropy**, is convected along with the fluid; the entropy of a fluid element remains constant. For smooth solutions, the entropy equation implies that $p = \text{const.}\rho^\gamma$ along a particle path. If the initial condition has constant entropy and if the inflow has same entropy, then the entropy is constant everywhere inside the domain at future times also.

This is however not always true, e.g. when shocks are present. Using the continuity equation this can also be written in conservation form

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

As discussed before, we can only demand an inequality in general

$$\frac{\partial}{\partial t}(-\rho s) + \frac{\partial}{\partial x}(-\rho s u) \le 0$$

4.6 Characteristic form

We can put the Euler equations in the form

$$\frac{\partial \phi}{\partial t} + \lambda \frac{\partial \phi}{\partial x} = 0$$

which leads to the characteristic equation

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = 0 \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = \lambda$$

The entropy equation is already in this form, i.e.,

$$\frac{\mathrm{d}s}{\mathrm{d}t} = 0 \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u$$

Combining the primitive form of the momentum and pressure equations, we have

$$\left(\frac{\partial p}{\partial t} + u\frac{\partial p}{\partial x} + \rho a^2 \frac{\partial u}{\partial x}\right) + a\left(\rho \frac{\partial u}{\partial t} + \rho u \frac{\partial u}{\partial x} + \frac{\partial p}{\partial x}\right) = 0$$

or

$$\frac{\partial p}{\partial t} + (u+a)\frac{\partial p}{\partial x} + \rho a\left[\frac{\partial u}{\partial t} + (u+a)\frac{\partial u}{\partial x}\right] = 0$$

which implies that

$$\frac{1}{\rho a}\frac{\mathrm{d}p}{\mathrm{d}t} + \frac{\mathrm{d}u}{\mathrm{d}t} = 0 \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u + a$$

Integrating this equation we have

$$\int \left(\frac{\mathrm{d}p}{\rho a} + \mathrm{d}u\right) = C \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u + a$$

If we assume that the entropy is constant in the whole domain, then ρ, a can be written as functions of pressure so that the first integral can be evaluated to

$$\frac{a}{\gamma - 1} + \frac{u}{2} = \text{const.}, \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u + a$$

Similarly we get

$$\frac{a}{\gamma - 1} - \frac{u}{2} = \text{const.}, \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u - a$$

We see that the quantities that are constant along the characteristics are the following **Riemann invariants**

$$\begin{aligned}
R^- &= \frac{a}{\gamma - 1} - \frac{u}{2} \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u - a \\
R^0 &= s \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u \\
R^+ &= \frac{a}{\gamma - 1} + \frac{u}{2} \quad \text{along} \quad \frac{\mathrm{d}x}{\mathrm{d}t} = u + a
\end{aligned}$$

4.7 Jump conditions

Let us consider a shock which is perpendicular to the x-axis. The Euler equations have the form

$$\partial_t U + \partial_x F(U) = 0$$

where

$$U = \begin{bmatrix} \rho \\ \rho u \\ \rho v \\ \rho w \\ E \end{bmatrix}, \qquad F = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ \rho uv \\ \rho uw \\ (E+p)u \end{bmatrix}, \qquad E = \frac{p}{\gamma - 1} + \frac{1}{2}\rho(u^2 + v^2 + w^2)$$

The eigenvalues of the flux jacobian are

$$\lambda_1 = u - a, \qquad \lambda_2 = \lambda_3 = \lambda_4 = u, \qquad \lambda_5 = u + a$$

The characteristic fields associated to λ_1, λ_5 are genuinely non-linear

$$\lambda_1'(U) r_1 \neq 0, \qquad \lambda_5'(U) r_5 \neq 0, \qquad \forall\ U \in \mathcal{U}_{ad}$$

while those associated to $\lambda_2, \lambda_3, \lambda_4$ are linearly degenerate

$$\lambda_i'(U) r_i = 0, \qquad i = 2, 3, 4, \qquad \forall\ U \in \mathcal{U}_{ad}$$

Define the jump operator

$$[\![\cdot]\!] = (\cdot)_r - (\cdot)_l$$

Across a discontinuity moving with speed S, the jump condition

$$[\![F]\!] = S\,[\![U]\!]$$

must be satisfied.

4.7.1 Contact/shear wave

This is associated with the eigenvalue u and the eigenvectors are linearly degenerate. The eigenvalue is a Riemann invariant and has same value across the wave, $u_l = u_r = u$. The contact wave moves with speed u. Fluid particles do not cross a contact wave. Now we write down the jump conditions.

$$\begin{aligned}
\rho_r u - \rho_l u &= u(\rho_r - \rho_l) \quad \checkmark \\
(p_r + \rho_r u^2) - (p_l + \rho_l u^2) &= u(\rho_r u - \rho_l u) \quad \Longrightarrow \quad p_r = p_l \\
\rho_r u v_r - \rho_l u v_l &= u(\rho_r v_r - \rho_l v_l) \quad \checkmark \\
\rho_r u w_r - \rho_l u w_l &= u(\rho_r w_r - \rho_l w_l) \quad \checkmark \\
(E_r + p_r)u - (E_l + p_l)u &= u(E_r - E_l) \quad \checkmark
\end{aligned}$$

If $\rho_l \neq \rho_r$ then it is a *material wave* or *contact wave*. If $v_l \neq v_r$ and/or $w_l \neq w_r$ then it is called a *shear wave*. Any two states with same values of u, p are admissible for a contact/shear wave. (FIGURE)

4.7.2 Shock wave

This is associated with the eigenvalue $u - a$ and/or $u + a$, which are genuinely non-linear. Let us perform a change of coordinate system in which the shock is stationary. Then the jump conditions become $F_r = F_l$, i.e.,

$$\begin{aligned}
\rho_r u_r &= \rho_l u_l \\
p_r + \rho_r u_r^2 &= p_l + \rho_l u_l^2 \\
\rho_r u_r v_r &= \rho_l u_l v_l \\
\rho_r u_r w_r &= \rho_l u_l w_l \\
(E_r + p_r)u_r &= (E_l + p_l)u_l
\end{aligned}$$

Note that $u_l \neq 0$, $u_r \neq 0$ and the y, z momentum jumps show that $v_l = v_r$ and $w_r = w_l$. Thus the tangential velocity components are continuous across a shock and only the normal component has a jump. The conditions can be reduced to

$$\begin{aligned} \rho_r u_r &= \rho_l u_l \\ p_r + \rho_r u_r^2 &= p_l + \rho_l u_l^2 \\ H_r &= H_l \end{aligned}$$

Then we can derive relations between the right and left states

$$\frac{\rho_r}{\rho_l} = \frac{u_l}{u_r} = \frac{(\gamma+1)M_l^2}{2+(\gamma-1)M_l^2}, \qquad \frac{p_r}{p_l} = 1 + \frac{2\gamma}{\gamma+1}(M_l^2 - 1)$$

where M is the mach number, $M = u/a$.

Entropy condition I Now we consider the entropy condition. Let us assume that $u_l > 0$ (and hence $u_r > 0$); then flow is from left to right and l is the *pre-shock* state and r is the *post-shock* state. An examination of the entropy condition shows that

$$(-\rho_r u_r s_r) - (-\rho_l u_l s_l) < 0 \quad \Longrightarrow \quad s_r > s_l \quad \Leftrightarrow \quad M_l > 1$$

i.e., the pre-shock flow must be supersonic relative to the shock. Under this condition we can deduce that

$$\rho_r > \rho_l, \qquad u_r < u_l, \qquad p_r > p_l$$

The density and pressure of a fluid element increases as it crosses a shock wave, and the velocity decreases. (FIGURE)

Entropy condition II An alternate way to check the entropy condition is by using Lax characterization; the characteristics must enter the shock curve. For the stationary 1-shock, the Lax condition is

$$u_l - a_l > 0 > u_r - a_r$$

This gives two inequalities

$$u_l > a_l > 0, \qquad u_r < a_r$$

The pre-shock state is supersonic and post-shock state is subsonic. From the first inequality and using mass jump condition

$$\rho_l u_l > \rho_l a_l \quad \Longrightarrow \quad \rho_r u_r > \rho_l a_l$$

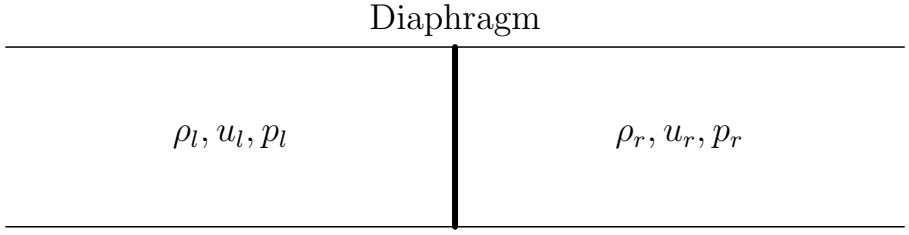
Combining this with second Lax inequality yields

$$\rho_l a_l < \rho_r u_r < \rho_r a_r \quad \Longrightarrow \quad \frac{\rho_r p_r}{\rho_l p_l} > 1$$

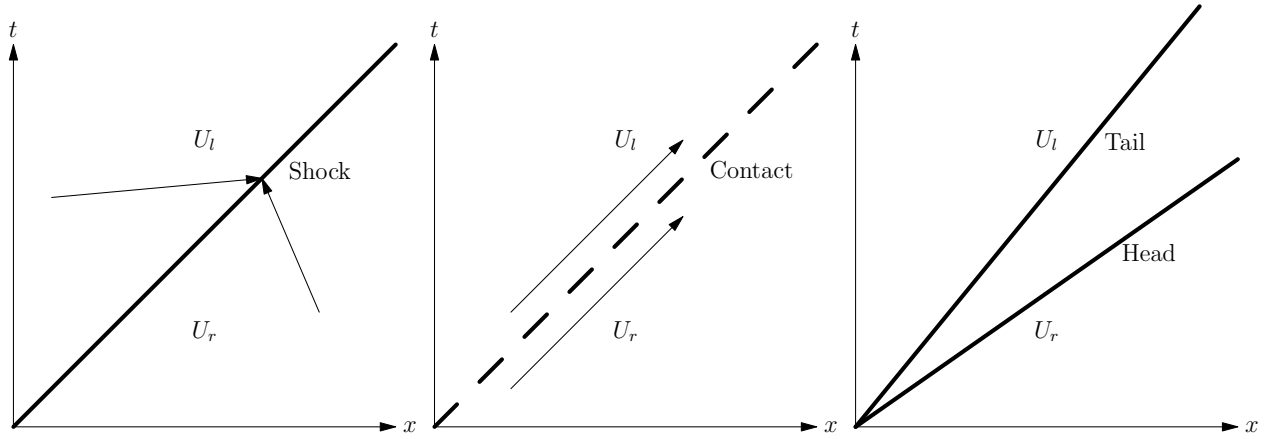
We know the ratio of density and pressure in terms of M_l and this tells us that $M_l > 1$ which again implies that $\rho_r > \rho_l$, $u_r < u_l$ and $p_r > p_l$.

4.8 Riemann problem (Shock tube problem)

The gas on the two sides of the diaphragm are at different states; when the diaphragm is ruptured, a pattern of waves is set up in the tube which may travel along the length of the tube.



We have seen that for a linear system of m hyperbolic PDEs, the Riemann problem consists of m discontinuity waves propagating with speeds given by the eigenvalues. A non-linear system also gives rise to m waves. For the 1-D Euler equations, the Riemann problem has in general three waves known as *shock*, *contact* and *expansion* wave. What type of waves are actually present in the solution will depend on the initial conditions of the Riemann problem.



- A shock is a discontinuity across which all the flow variables density, velocity, pressure, are discontinuous. A shock is associated with the characteristic fields corresponding to the eigenvalues $\lambda_1 = u - a$ and $\lambda_3 = u + a$. The characteristics on either side of the shock intersect into the shock. Fluid particles can cross the shock; when this happens, their velocity decreases, and, density and pressure increase.

- A contact is a discontinuity across which density is discontinuous but pressure and velocity are continuous. It is associated with the characteristic field corresponding to the eigenvalue $\lambda_2 = u$. The characteristics on either side of the contact are parallel to the contact line. Fluid particles do not cross a contact discontinuity.

- A rarefaction or expansion wave is a continuous wave which consists of a *head* and a *tail*; all the flow quantities vary continuously through the wave and the entropy is constant. This wave is associated with the characteristic fields corresponding to the eigenvalues $\lambda_1 = u - a$ and $\lambda_3 = u + a$. The density of a fluid particle decreases as it crosses this wave, and hence the name rarefaction wave.

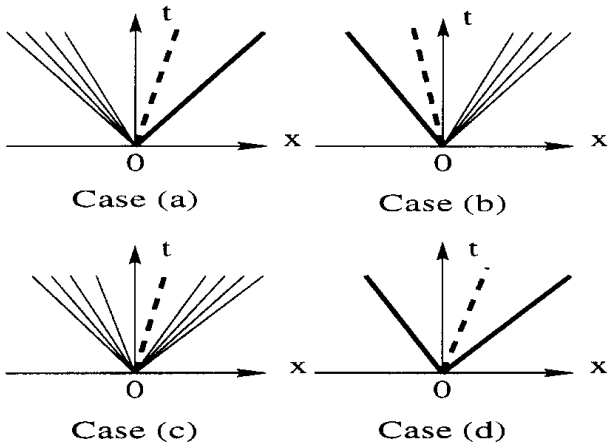


Fig. 4.2. Possible wave patterns in the solution of the Riemann problem: (a) left rarefaction, contact, right shock (b) left shock, contact, right rarefaction (c) left rarefaction, contact, right rarefaction (d) left shock, contact, right shock

Test	ρ_L	u_L	p_L	ρ_R	u_R	p_R
1	1.0	0.0	1.0	0.125	0.0	0.1
2	1.0	-2.0	0.4	1.0	2.0	0.4
3	1.0	0.0	1000.0	1.0	0.0	0.01
4	1.0	0.0	0.01	1.0	0.0	100.0
5	5.99924	19.5975	460.894	5.99242	-6.19633	46.0950

Table 4.1. Data for five Riemann problem tests

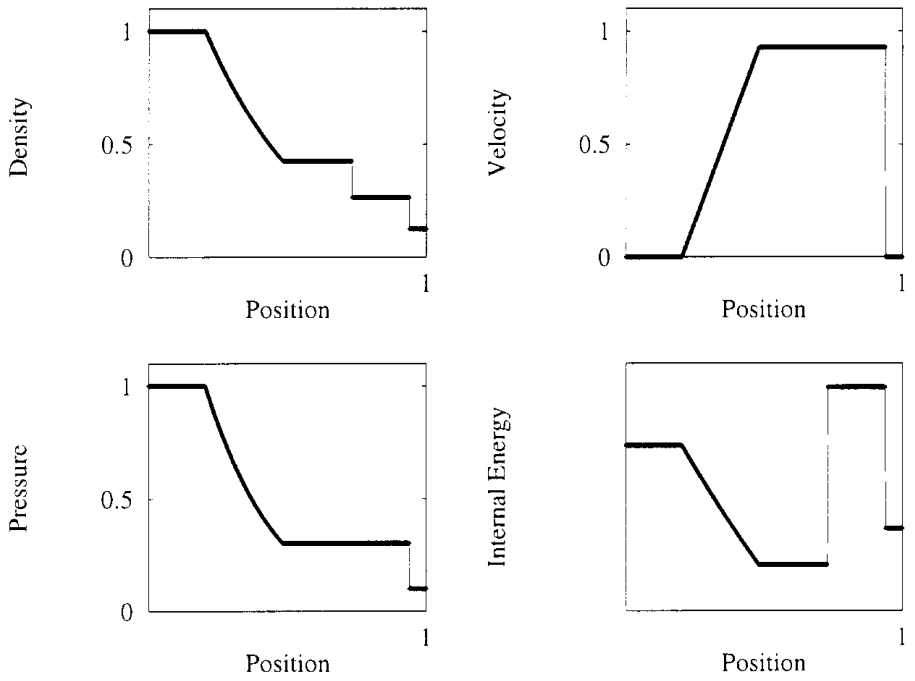


Fig. 4.7. Test 1: Exact solution for density, velocity, pressure and specific internal energy at time t = 0.25 units

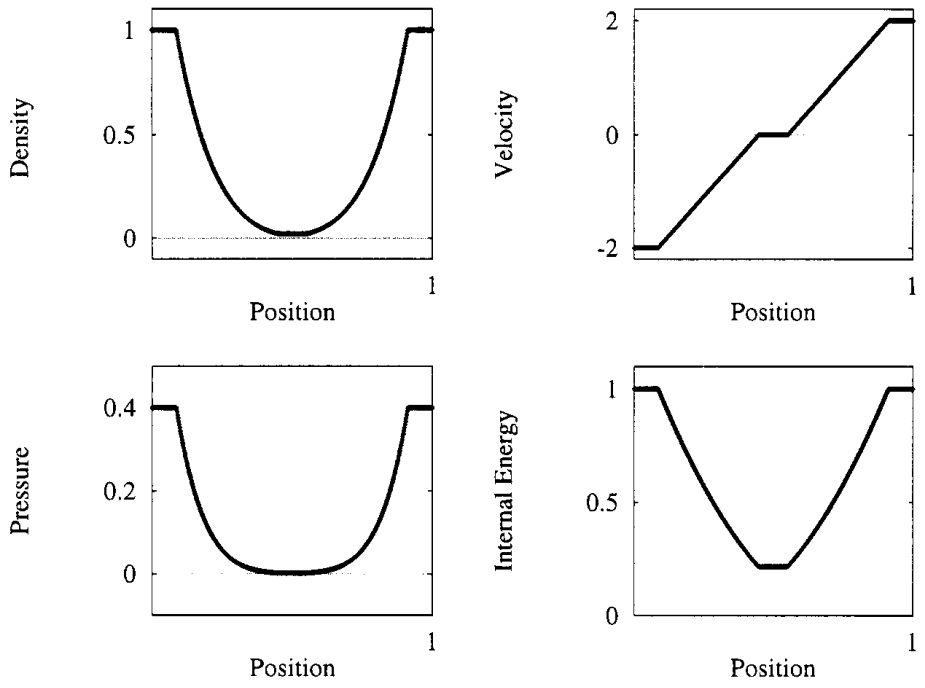


Fig. 4.8. Test 2: Exact solution for density, velocity, pressure and specific internal energy at time t = 0.15 units

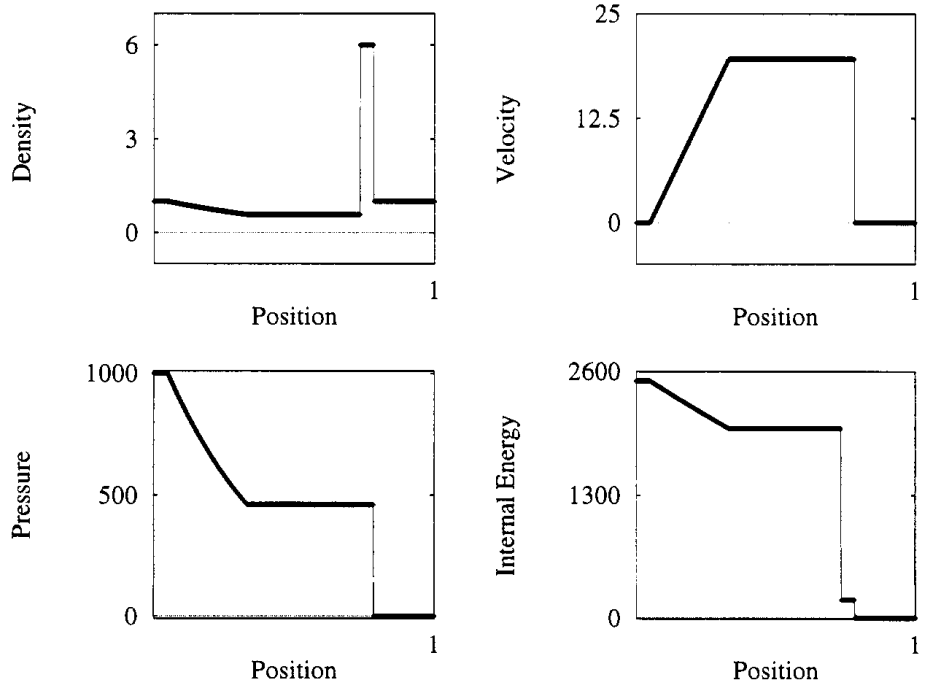


Fig. 4.9. Test 3: Exact solution for density, velocity, pressure and specific internal energy at time t = 0.012 units

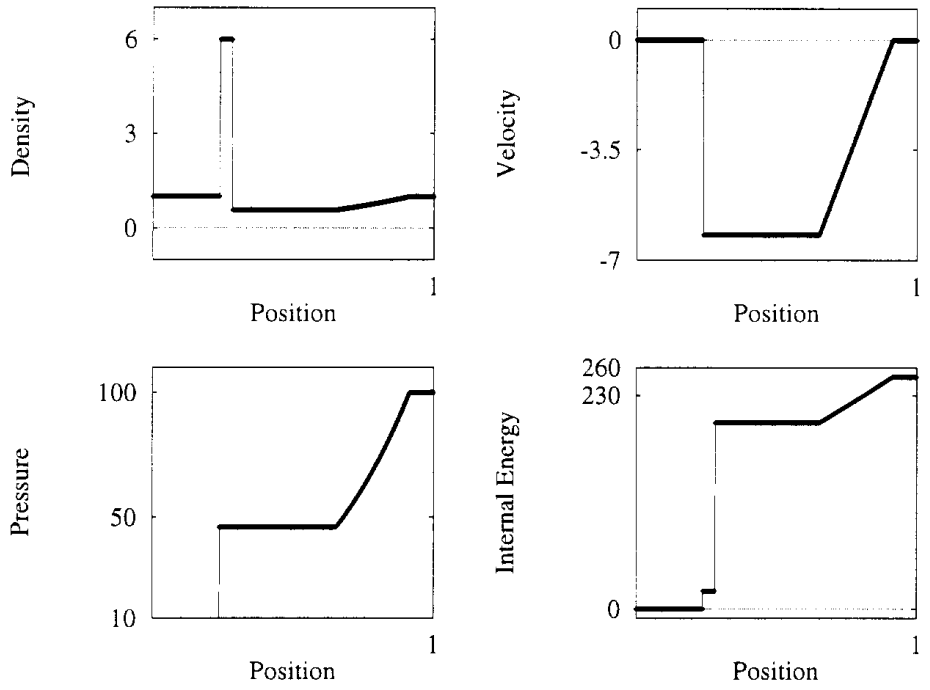


Fig. 4.10. Test 4: Exact solution for density, velocity, pressure and specific internal energy at time t = 0.035 units

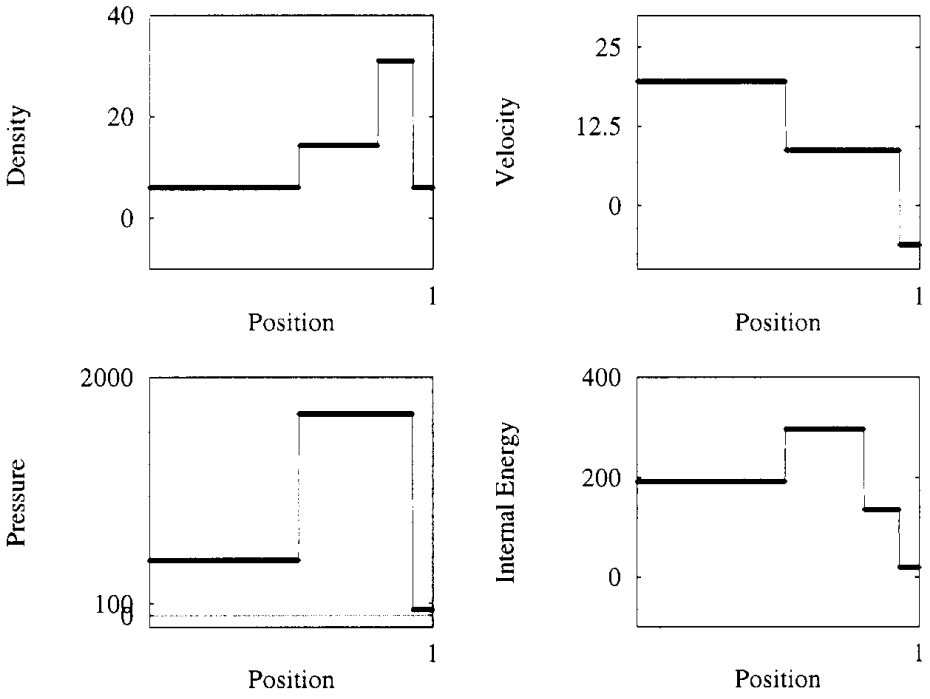


Fig. 4.11. Test 5: Exact solution for density, velocity, pressure and specific internal energy at time t = 0.035 units

Chapter 5

Lax-Friedrich flux

Let us consider a 1-D system of conservation laws

$$U_t + F(U)_x = 0$$

The central difference scheme

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + \frac{F_{j+1}^n - F_j^n}{2\Delta x} = 0$$

is unstable for hyperbolic problems. A stable version can be obtained by a small modification

$$\frac{U_j^{n+1} - \frac{1}{2}(U_{j-1}^n + U_{j+1}^n)}{\Delta t} + \frac{F_{j+1}^n - F_j^n}{\Delta x} = 0$$

which can be written in finite volume form with numerical flux

$$F_{j+\frac{1}{2}} = \frac{1}{2}(F_j + F_{j+1}) - \frac{1}{2}\frac{\Delta x}{\Delta t}(U_{j+1} - U_j)$$

This has the usual structure

$$F_{j+\frac{1}{2}} = \frac{1}{2}(F_j + F_{j+1}) - \frac{1}{2}\lambda(U_{j+1} - U_j), \qquad \lambda = \frac{\Delta x}{\Delta t} = \text{speed}$$

Since the time step must satisfy a CFL condition (show this by Fourier analysis)

$$\Delta t = \text{CFL}\frac{\Delta x}{\max_j \sigma(F'(U_j^n))}, \qquad \text{CFL} \le 1$$

where σ is the spectral radius. We see that

$$\lambda \approx \max_j \sigma(F'(U_j^n))$$

The parameter λ is related to the maximum wave speed in the whole computational domain. For this reason, this scheme is also sometimes called *global Lax-Friedrich* scheme. It is very simple in the sense that we do not need to know anything about the eigenvectors but only need to know the largest eigenvalue. It is a very robust scheme and seldom fails to give an answer, but the results will be very diffusive.

5.1 Rusanov or local Lax-Friedrich scheme

A simple way to improve the scheme is to use a local estimate of λ so that the flux has the form

$$F_{j+\frac{1}{2}} = \frac{1}{2}(F_j + F_{j+1}) - \frac{1}{2}\lambda_{j+\frac{1}{2}}(U_{j+1} - U_j)$$

where $\lambda_{j+\frac{1}{2}}$ is an estimate of the maximum wave speed arising in the Riemann problem at the face $j+\frac{1}{2}$. For convex fluxes, a simple choice is

$$\lambda_{j+\frac{1}{2}} = \max\{\sigma(F'(U_j^n)), \sigma(F'(U_{j+1}^n))\}$$

This scheme was proposed by Rusanov [4] and is also called *local Lax-Friedrich* scheme. For the Euler equations

$$\lambda_{j+\frac{1}{2}} = \max\{|u_j| + a_j, |u_{j+1}| + a_{j+1}\}$$

is a good choice.

5.2 Positivity property

$$\begin{aligned} U_j^{n+1} &= \left[1 - \frac{\Delta t}{2\Delta x}(\lambda_{j-\frac{1}{2}} + \lambda_{j+\frac{1}{2}})\right] U_j^n \\ &\quad + \frac{\lambda_{j-\frac{1}{2}}\Delta t}{2\Delta x}\left[U_{j-1} + \frac{1}{\lambda_{j-\frac{1}{2}}}F_{j-1}\right] + \frac{\lambda_{j+\frac{1}{2}}\Delta t}{2\Delta x}\left[U_{j+1} - \frac{1}{\lambda_{j+\frac{1}{2}}}F_{j+1}\right] \\ &= \left[1 - \frac{\Delta t}{2\Delta x}(\lambda_{j-\frac{1}{2}} + \lambda_{j+\frac{1}{2}})\right] U_j^n \\ &\quad + \frac{\lambda_{j-\frac{1}{2}}\Delta t}{2\Delta x}U_{j-1}^+ + \frac{\lambda_{j+\frac{1}{2}}\Delta t}{2\Delta x}U_{j+1}^- \end{aligned}$$

Theorem 5.1. *If $U \in \mathcal{U}_{ad}$ and $\lambda \geq |u| + a$ then*

$$U^{\pm} := U \pm \frac{1}{\lambda}F(U) \in \mathcal{U}_{ad}$$

Proof. The first component is

$$\begin{aligned} \rho \pm \frac{1}{\lambda}\rho u &= \frac{1}{\lambda}\rho(\lambda \pm u) \\ &\geq \frac{1}{\lambda}(|u| \pm u + a) \\ &\geq \frac{1}{\lambda}a \qquad \text{since } |u| \pm u \geq 0 \\ &> 0 \end{aligned}$$

The pressure component is

$$= 2\left(\rho \pm \frac{1}{\lambda}\rho u\right)\left(E \pm \frac{1}{\lambda}(E+p)u\right) - \left(\rho u \pm \frac{1}{\lambda}(p + \rho u^2)\right)^2$$

□

Theorem 5.2. *The FV scheme with Rusanov flux and $\lambda_{j+\frac{1}{2}}$ as defined above is positive under the condition*

$$\frac{\Delta t}{2\Delta x}(\lambda_{j-\frac{1}{2}} + \lambda_{j+\frac{1}{2}}) \leq 1$$

Chapter 6

Flux Vector Splitting schemes

The main idea is to split the flux F into two parts

$$F(U) = F^+(U) + F^-(U)$$

such that

$$\begin{aligned} F^+ &= \text{due to waves moving to right, positive speed} \\ F^- &= \text{due to waves moving to left, negative speed} \end{aligned}$$

Since eigenvalues determine the wave speeds, we should ensure that

$$\frac{\partial F^+}{\partial U} \text{ has positive eigenvalues}, \quad \frac{\partial F^-}{\partial U} \text{ has negative eigenvalues}$$

Then the numerical flux function is given by

$$F_{i+1/2} = F(U_i, U_{i+1}) = F^+(U_i) + F^-(U_{i+1})$$

This is a consistent flux; if $U_i = U_{i+1} = U$

$$F_{i+\frac{1}{2}} = F(U, U) = F^+(U) + F^-(U) = F(U)$$

6.1 Steger-Warming scheme

The Steger-Warming flux [6] is based on the homogeneity property of the Euler flux

$$F(U) = A(U)U, \quad A(U) = \frac{\partial F}{\partial U}$$

Split the flux jacobian matrix using eigenvalue splitting

$$A(U) = A^+(U) + A^-(U), \quad A^\pm(U) = R(U)\Lambda^\pm(U)R^{-1}(U)$$

Then the flux can be split based on eigenvalue splitting

$$F(U) = A^+(U)U + A^-(U)U = F^+(U) + F^-(U) \quad \Longrightarrow \quad F^\pm(U) = A^\pm(U)U$$

The Steger-Warming flux is given by

$$F_{i+1/2} = F^+(U_i) + F^-(U_{i+1}) = A^+(U_i)U_i + A^-(U_{i+1})U_{i+1}$$

Some properties The flux has the upwind property; e.g., if both U_i, U_{i+1} are supersonic to the right, then

$$A^+(U_i) = A(U_i), \qquad A^-(U_{i+1}) = 0 \quad \Longrightarrow \quad F(U_i, U_{i+1}) = A(U_i)U_i = F(U_i)$$

Moreover, if $\gamma \in [1, 5/3]$

$$\frac{\partial F^+}{\partial U} \geq 0, \quad \frac{\partial F^-}{\partial U} \leq 0$$

It is found to add excessive numerical dissipation and leads to poor resolution of contact waves. Moreover, $F^\pm$ are not differentiable at sonic and stagnation points. Laney recommends smoothing the eigenvalues

$$\tilde{\lambda}_i^\pm = \frac{1}{2}\left(\lambda_i \pm \sqrt{\lambda_i^2 + \delta^2}\right), \qquad i = 1, 2, 3$$

Implementation A direct implementation can be costly since it involves many matrix-matrix products. First express the conserved vector U in terms of the eigenvectors of A

$$U = \alpha_1 r_1 + \alpha_2 r_2 + \alpha_3 r_3, \qquad \alpha_1 = \alpha_3 = \frac{\rho}{2\gamma}, \qquad \alpha_2 = \rho\frac{\gamma - 1}{\gamma}$$

or

$$U = R\alpha, \qquad \alpha = [\alpha_1,\ \alpha_2,\ \alpha_3]^\top$$

The split fluxes are given by

$$F^\pm = A^\pm U = R\Lambda^\pm R^{-1}(R\alpha) = R\Lambda^\pm \alpha = \alpha_1\lambda_1^\pm r_1 + \alpha_2\lambda_2^\pm r_2 + \alpha_3\lambda_3^\pm r_3$$

Remark 6.1. *A variant of this flux was proposed by Vijayasundaram [9]*

$$F_{i+1/2} = A^+(U_{i+\frac{1}{2}})U_i + A^-(U_{i+\frac{1}{2}})U_{i+1}, \qquad U_{i+\frac{1}{2}} = \frac{1}{2}(U_i + U_{i+1})$$

6.2 van Leer splitting

- The Mach number: $M = \frac{u}{a}$ tells us about sign of eigenvalues.
- If $M > 1$: all eigenvalues are positive

  $$M > 1 \qquad \Longrightarrow \qquad u - a > 0, \quad u > 0, \quad u + a > 0$$

  and if $M < -1$, all eigenvalues are negative. We can use the Mach number to determine the upwind direction.
- Euler flux: polynomials in M

  $$F = \begin{bmatrix} \rho a M \\ \frac{\rho a^2}{\gamma}(\gamma M^2 + 1) \\ \rho a^3 M\left(\frac{1}{2}M^2 + \frac{1}{\gamma - 1}\right) \end{bmatrix}$$

- Each component flux of the form

  $$F = G(\rho, a)H(M)$$

  Split the flux as

  $$F = F^+ + F^-, \qquad F^\pm = G(\rho, a)H^\pm(M)$$

- Split polynomial $H(M)$ such that we have upwinding and smoothness

  1. $H(M) = H^+(M) + H^-(M)$
  2. $H^+(M) = 0$ for $M \leq -1$ ($\Longrightarrow$ $H^-(M) = H(M)$)
  3. $H^+(M) = H(M)$ for $M \geq 1$ ($\Longrightarrow$ $H^-(M) = 0$)
  4. $\frac{\mathrm{d}}{\mathrm{d}M}H^+(-1) = 0$, $\frac{\mathrm{d}}{\mathrm{d}M}H^+(1) = \frac{\mathrm{d}}{\mathrm{d}M}H(1)$.

6.2.1 van Leer: Mass flux

$$H(M) = M = M^+(M) + M^-(M)$$

- To satisfy all conditions, $M^\pm$ must be quadratic polynomial in M

$$M^+ = \begin{cases} 0 & M \le -1 \\ \left(\frac{M+1}{2}\right)^2 & -1 < M < 1 \\ M & M \ge 1 \end{cases}$$

$$M^- = \begin{cases} M & M \le -1 \\ -\left(\frac{M-1}{2}\right)^2 & -1 < M < 1 \\ 0 & M \ge 1 \end{cases}$$

6.2.2 van Leer: momentum flux

Split $\gamma M^2 + 1$ using cubic polynomials in M

$$(\gamma M^2 + 1) = (\gamma M^2 + 1)^+ + (\gamma M^2 + 1)^-$$

$$(\gamma M^2 + 1)^+ = \begin{cases} 0 & M \le -1 \\ \left(\frac{M+1}{2}\right)^2 [(\gamma - 1)M + 2] & -1 < M < +1 \\ \gamma M^2 + 1 & M \ge 1 \end{cases}$$

$$(\gamma M^2 + 1)^- = \begin{cases} \gamma M^2 + 1 & M \le -1 \\ -\left(\frac{M-1}{2}\right)^2 [(\gamma - 1)M - 2] & -1 < M < +1 \\ 0 & M \ge 1 \end{cases}$$

6.2.3 van Leer: Energy flux

Split energy flux using quartic polynomials in M

$$F_3^+ = \begin{cases} 0 & M \le -1 \\ \frac{[(\gamma-1)u+2a]^2 F_1^+}{2(\gamma+1)(\gamma-1)} & -1 < M < 1 \\ F_3 & M > 1 \end{cases}$$

$$F_3^- = \begin{cases} F_3 & M \le -1 \\ \frac{[(\gamma-1)u-2a]^2 F_1^-}{2(\gamma+1)(\gamma-1)} & -1 < M < 1 \\ 0 & M > 1 \end{cases}$$

6.2.4 van Leer flux

- Final flux formulae

$$F^\pm = \pm\frac{1}{4}\rho a (M \pm 1)^2 \begin{bmatrix} 1 \\ \frac{(\gamma-1)u \pm 2a}{\gamma} \\ \frac{[(\gamma-1)u \pm 2a]^2}{2(\gamma+1)(\gamma-1)} \end{bmatrix}$$

- Has upwind property of split flux jacobians

$$\frac{\partial F^+}{\partial U} \ge 0, \quad \frac{\partial F^-}{\partial U} \le 0$$

- Adds excessive dissipation for contact discontinuity

6.3 Liou and Steffen (1993)

Separate flux into convective and pressure parts

$$F = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ \rho H u \end{bmatrix} = \begin{bmatrix} \rho u \\ \rho u^2 \\ \rho H u \end{bmatrix} + \begin{bmatrix} 0 \\ p \\ 0 \end{bmatrix} = M \begin{bmatrix} \rho a \\ \rho u a \\ \rho H a \end{bmatrix} + \begin{bmatrix} 0 \\ p \\ 0 \end{bmatrix} = M F_c + F_p$$

Flux splitting

$$F^{\pm} = M^{\pm} F_c + F_p^{\pm}, \qquad F_p^{\pm} = \begin{bmatrix} 0 \\ p^{\pm} \\ 0 \end{bmatrix}$$

$M^{\pm}$ is same as in van Leer scheme. Pressure $p = p^+ + p^-$

$$p^+ = p \begin{cases} 0 & M \le -1 \\ \frac{1}{2}(1+M) & -1 < M < 1 \\ 1 & M \ge 1 \end{cases}, \qquad p^- = p \begin{cases} 1 & M \le -1 \\ \frac{1}{2}(1-M) & -1 < M < 1 \\ 0 & M \ge 1 \end{cases}$$

Remark: See the papers on AUSM family of schemes.

6.4 Zha-Bilgen flux vector splitting (1993)

$$F = \begin{bmatrix} \rho u \\ p + \rho u^2 \\ (E+p)u \end{bmatrix} = \begin{bmatrix} \rho u \\ \rho u^2 \\ E u \end{bmatrix} + \begin{bmatrix} 0 \\ p \\ p u \end{bmatrix} = u \begin{bmatrix} \rho \\ \rho u \\ E \end{bmatrix} + \begin{bmatrix} 0 \\ p \\ p u \end{bmatrix}$$

Flux vector splitting

$$F^{\pm} = u^{\pm} U + \begin{bmatrix} 0 \\ p^{\pm} \\ (pu)^{\pm} \end{bmatrix}, \qquad u^{\pm} = \frac{1}{2}(u \pm |u|)$$

$p^{\pm}$ is same as in Liou-Steffen scheme.

$$(pu)^+ = p \begin{cases} 0 & M \le -1 \\ \frac{1}{2}(u+a) & -1 < M < 1 \\ u & M \ge 1 \end{cases}, \qquad (pu)^- = p \begin{cases} u & M \le -1 \\ \frac{1}{2}(u-a) & -1 < M < 1 \\ 0 & M \ge 1 \end{cases}$$

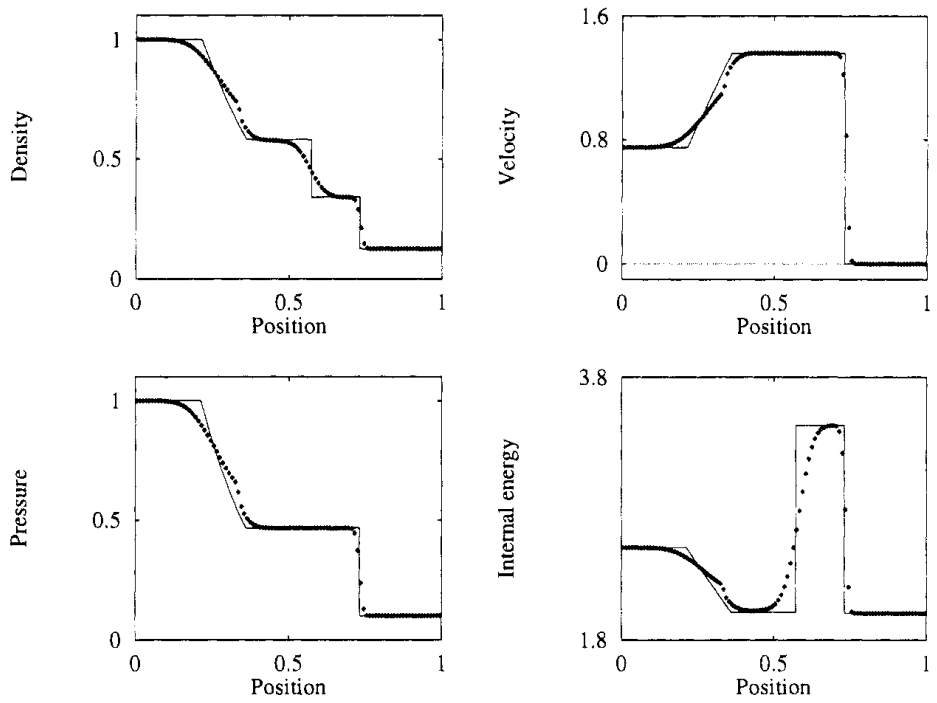


Fig. 8.3. Steger and Warming FVS scheme applied to Test 1, with $x_0 = 0.3$. Numerical (symbol) and exact (line) solutions are compared at time 0.2 units

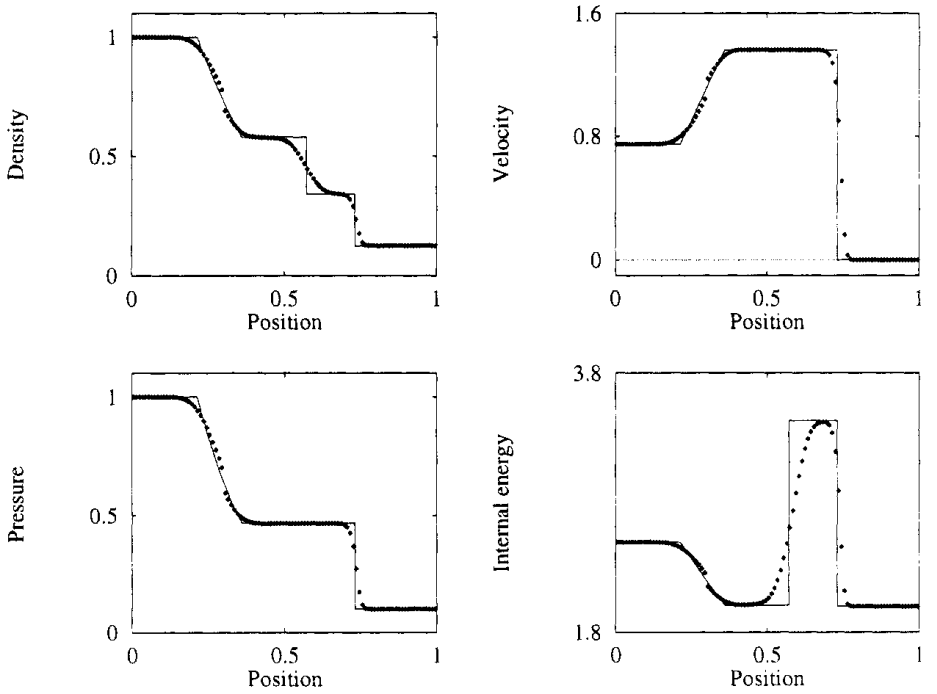


Fig. 8.4. Van Leer FVS scheme applied to Test 1, with $x_0 = 0.3$. Numerical (symbol) and exact (line) solutions are compared at time 0.2 units

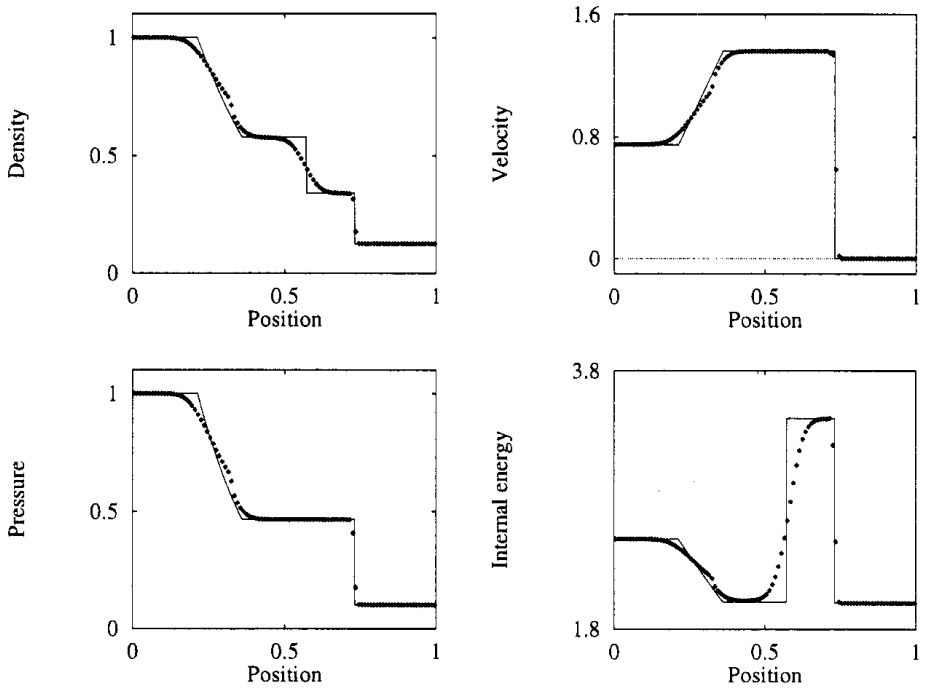


Fig. 8.5. Liou and Steffen scheme applied to Test 1, with $x_0 = 0.3$. Numerical (symbol) and exact (line) solutions are compared at time 0.2 units

Chapter 7

Godunov scheme

- At any time t^n, FV solution is constant in each cell

$$U(x,t^n) = U_i^n, \quad x_{i-1/2} < x < x_{i+1/2}$$

- Riemann problem at every cell interface
- Godunov's idea
  1. Solve Riemann problem for U^n at every cell interface exactly
  2. Evolve the Riemann solution upto next time level $t^{n+1} = t^n + \Delta t$
  3. Average the solution at t^{n+1} to get cell average values U^{n+1}
- Solution of Riemann problem at $x_{i+1/2}$

$$U_{i+1/2}\left(\frac{x - x_{i+1/2}}{t - t^n}\right)$$

  waves moving to left and right
- Waves from successive Riemann problems must not intersect

$$\Delta t < \frac{h}{2S^n}$$

- S^n = maximum wave speed of all Riemann problems
- Average solution at new time level

$$U_i^{n+1} = \frac{1}{h}\left[\int_0^{\frac{1}{2}h} U_{i-1/2}\left(\frac{\xi}{\Delta t}\right)\mathrm{d}\xi + \int_{-\frac{1}{2}h}^{0} U_{i+1/2}\left(\frac{\xi}{\Delta t}\right)\mathrm{d}\xi\right]$$

- Difficult to implement numerically when expansion waves are present
- CFL condition is more restrictive
- Exact solution of Riemann problem

$$\tilde{U}(x,t) = U_{i+1/2}\left(\frac{x - x_{i+1/2}}{t - t^n}\right), \quad x_i \le x \le x_{i+1}, \quad t^n \le t \le t^{n+1}$$

- Satisfies integral conservation law

$$\begin{aligned}\int_{x_{i-1/2}}^{x_{i+1/2}} \tilde{U}(x,t^{n+1})\mathrm{d}x &= \int_{x_{i-1/2}}^{x_{i+1/2}} \tilde{U}(x,t^n)\mathrm{d}x + \int_0^{\Delta t} F[\tilde{U}(x_{i-1/2},t)]\mathrm{d}t \\ &\quad - \int_0^{\Delta t} F[\tilde{U}(x_{i+1/2},t)]\mathrm{d}t\end{aligned}$$

But

$$\int_0^{\Delta t} F[\tilde{U}(x_{i+1/2}, t)]\mathrm{d}t = \int_0^{\Delta t} F[U_{i+1/2}(0)]\mathrm{d}t = F[U_{i+1/2}(0)]\Delta t$$

etc., so that we finally have

$$U_i^{n+1} = U_i^n - \frac{\Delta t}{h}\left[F(U_{i+1/2}(0)) - F(U_{i-1/2}(0))\right]$$

- Godunov flux

$$F_{i+1/2} = F(U_i, U_{i+1}) = F(U_{i+1/2}(0))$$

- right moving waves from $U_{i-1/2}$ should not reach $x_{i+1/2}$ and vice versa: CFL condition

$$\Delta t < \frac{h}{S^n}$$

- Accurate but expensive - not used for practical computations
- Recall: Linear system of equations $F = AU$, A constant matrix

$$F(U_{i+1/2}(0)) = A^+ U_i + A^- U_{i+1}$$

Here, Godunov scheme is identical to upwind scheme

Chapter 8

Roe scheme

The exact solution of Riemann problem is a costly process and in some systems we may not have an exact solution available. Since we are anyway computing an approximate solution to the PDE, there is not much advantage in solving the Riemann problem exactly. P. L. Roe decided to solve the Riemann problem approximately. His approach was to linearize the non-linear problem

$$\frac{\partial U}{\partial t} + A(U)\frac{\partial U}{\partial x} = 0 \quad \Longrightarrow \quad \frac{\partial U}{\partial t} + \bar{A}\frac{\partial U}{\partial x} = 0$$

and then solve Riemann problem exactly for the linear problem. The matrix $\bar{A} = \bar{A}(U_l, U_r)$ in the linear problem must satisfy certain consistency conditions.

1. Consistency: $\bar{A}(U, U) = A(U)$
2. Hyperbolicity: $\bar{A}$ has all real eigenvalues and linearly independent eigenvectors.
3. Conservation: $F(U_r) - F(U_l) = \bar{A}(U_l, U_r)(U_r - U_l)$

This is known as a ***Roe-type linearization*** of the non-linear hyperbolic PDE. The main task is to construct such a matrix and this was achieved by Roe for the Euler equations in his famous paper [3].

Let us introduce the parameter vector

$$Z = \begin{bmatrix} Z_1 \\ Z_2 \\ Z_3 \end{bmatrix} = \sqrt{\rho} \begin{bmatrix} 1 \\ u \\ H \end{bmatrix}$$

Define the straight line path connecting Z_l and Z_r

$$Z(\alpha) = Z_l + \alpha(Z_r - Z_l), \qquad Z(0) = Z_l, \qquad Z(1) = Z_r$$

Now U and F are homogeneous functions of degree two in the parameter vector Z. Then

$$D(Z) = U'(Z), \qquad C(Z) = F'(Z)$$

are linear in Z. We can express the jump in U as

$$\begin{aligned}
\Delta U &= U_r - U_l \\
&= U(Z(1)) - U(Z(0)) \\
&= \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\alpha} U(Z(\alpha)) \mathrm{d}\alpha \\
&= \int_0^1 \frac{\partial}{\partial Z} U(Z(\alpha)) \cdot \frac{\partial}{\partial \alpha} Z(\alpha) \mathrm{d}\alpha \\
&= \left[\int_0^1 \underbrace{D(Z(\alpha))}_{\text{linear in } \alpha} \mathrm{d}\alpha \right] (Z_r - Z_l) \\
&= D(\bar{Z})\Delta Z, \qquad \bar{Z} = \frac{1}{2}(Z_l + Z_r)
\end{aligned}$$

Similarly, we can show for the flux difference

$$\Delta F = C(\bar{Z})\Delta Z$$

The matrices C, D are given by

$$D(Z) = \begin{bmatrix} 2Z_1 & 0 & 0 \\ Z_2 & Z_1 & 0 \\ \frac{1}{\gamma}Z_3 & \frac{\gamma-1}{\gamma}Z_2 & \frac{1}{\gamma}Z_1 \end{bmatrix}, \quad C(Z) = \begin{bmatrix} Z_2 & Z_1 & 0 \\ \frac{\gamma-1}{\gamma}Z_3 & \frac{\gamma+1}{\gamma}Z_2 & \frac{\gamma-1}{\gamma}Z_1 \\ 0 & Z_3 & Z_2 \end{bmatrix}$$

Hence we have

$$\Delta F = \hat{A}\Delta U, \quad \hat{A} = C(\bar{Z})D^{-1}(\bar{Z})$$

An explicit computation gives

$$\hat{A} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{\gamma-3}{\gamma}\left(\frac{\bar{Z}_2}{\bar{Z}_1}\right)^2 & (3-\gamma)\frac{\bar{Z}_2}{\bar{Z}_1} & \gamma-1 \\ \frac{\gamma-1}{2}\left(\frac{\bar{Z}_2}{\bar{Z}_1}\right)^3 - \frac{\bar{Z}_2\bar{Z}_3}{\bar{Z}_1^2} & \frac{\bar{Z}_3}{\bar{Z}_1} - (\gamma-1)\left(\frac{\bar{Z}_2}{\bar{Z}_1}\right)^2 & \gamma\frac{\bar{Z}_2}{\bar{Z}_1} \end{bmatrix}$$

Define *Roe averages*

$$\bar{u} = \frac{\bar{Z}_2}{\bar{Z}_1} = \frac{u_l\sqrt{\rho}_l + u_r\sqrt{\rho}_r}{\sqrt{\rho}_l + \sqrt{\rho}_r}, \quad \bar{H} = \frac{\bar{Z}_3}{\bar{Z}_1} = \frac{H_l\sqrt{\rho}_l + H_r\sqrt{\rho}_r}{\sqrt{\rho}_l + \sqrt{\rho}_r}$$

In terms of these average quantities, the matrix $\hat{A}$ is

$$\hat{A} = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{2}(\gamma-3)\bar{u}^2 & (3-\gamma)\bar{u} & \gamma-1 \\ \bar{u}[\frac{1}{2}(\gamma-1)\bar{u}^2 - \bar{H}] & \bar{H} - (\gamma-1)\bar{u}^2 & \gamma\bar{u} \end{bmatrix}$$

Define the average density

$$\bar{\rho} = \sqrt{\rho_l\rho_r}$$

Define $\bar{U}$ = conserved vector corresponding to $(\bar{\rho}, \bar{u}, \bar{H})$. Then

$$\hat{A} = A(\bar{U})$$

Hence if we take $\bar{A} = \hat{A}$, all three conditions satisfied. The eigenvalues of $\bar{A}$ are given by

$$\bar{u} - \bar{a}, \qquad \bar{u}, \qquad \bar{u} + \bar{a}$$

where

$$\bar{a} = \sqrt{(\gamma-1)\left[\bar{H} - \frac{1}{2}\bar{u}^2\right]}$$

Let $U_R(x/t)$ be the solution of the Riemann problem for the Roe-linearized equation. The flux is given by $F(U_R(0))$ and we know from linear systems

$$F(U_R(0)) = \frac{1}{2}(F_l + F_r) - \frac{1}{2}|\bar{A}(U_l, U_r)|(U_r - U_l)$$

Moreover, this flux has the upwind property.

Remark 8.1. *Show that if the two states U_l, U_r are positive, then the Roe average sound speed $\bar{a}$ is well defined.*

8.1 Roe scheme: entropy violation

- Roe scheme – derived from linearized problem – has only contact discontinuities – no expansion wave
- sonic expansion wave – Roe scheme can give rise to entropy violating shocks
- Eigenvalue of $\bar{A}$ becomes zero – loss of numerical dissipation

$$\bar{\lambda}_1 = \bar{u} - \bar{a}, \quad \bar{\lambda}_2 = \bar{u}, \qquad \bar{\lambda}_3 = \bar{u} + \bar{a}$$

- Entropy fix – do not allow eigenvalue to become zero

$$|\hat{\lambda}_i| = \begin{cases} \frac{\bar{\lambda}_i^2}{4\epsilon\bar{a}} + \epsilon\bar{a} & \text{if } |\bar{\lambda}_i| < 2\epsilon\bar{a} \\ |\bar{\lambda}_i| & \text{otherwise} \end{cases}$$

  This fix is applied to $\lambda_1 = u - a$ and $\lambda_3 = u + a$.

8.2 Roe scheme formulae

The Roe flux is given by

$$F(U_l, U_r) = \frac{1}{2}(F_l + F_r) - \frac{1}{2}|\bar{A}(U_l, U_r)|(U_r - U_l) \qquad \bar{A} = R|\Lambda|R^{-1}$$

The eigenvectors of $\bar{A}$

$$r_1 = \begin{bmatrix} 1 \\ \bar{u} - \bar{a} \\ \bar{H} - \bar{u}\bar{a} \end{bmatrix}, \quad r_2 = \begin{bmatrix} 1 \\ \bar{u} \\ \frac{1}{2}\bar{u}^2 \end{bmatrix}, \quad r_3 = \begin{bmatrix} 1 \\ \bar{u} + \bar{a} \\ \bar{H} + \bar{u}\bar{a} \end{bmatrix}$$

A direct computation requires several for loops to compute the matrix-matrix and matrix-vector products which is not an efficient way to write the code.

We note that

$$|\bar{A}(U_l, U_r)|(U_r - U_l) = R|\Lambda| \underbrace{R^{-1}(U_r - U_l)}_{\alpha}$$

Write jump in terms of eigenvectors

$$\Delta U = U_r - U_l = \alpha_1 r_1 + \alpha_2 r_2 + \alpha_3 r_3 = R\alpha, \quad \alpha \in \mathbb{R}^3$$

or

$$\begin{aligned} \alpha_1 + \alpha_2 + \alpha_3 &= \Delta U_1 \\ \alpha_1(\bar{u} - \bar{a}) + \alpha_2\bar{u} + \alpha_3(\bar{u} + \bar{a}) &= \Delta U_2 \\ \alpha_1(\bar{H} - \bar{u}\bar{a}) + \alpha_2\frac{\bar{u}^2}{2} + \alpha_3(\bar{H} + \bar{u}\bar{a}) &= \Delta U_3 \end{aligned}$$

The solution is

$$\begin{aligned} \alpha_2 &= \frac{\gamma - 1}{\bar{a}^2}[(\bar{H} - \bar{u}^2)\Delta U_1 + \bar{u}\Delta U_2 - \Delta U_3] \\ \alpha_1 &= \frac{1}{2\bar{a}}[(\bar{u} + \bar{a})\Delta U_1 - \Delta U_2 - \bar{a}\alpha_2] \\ \alpha_3 &= \Delta U_1 - \alpha_1 - \alpha_2 \end{aligned}$$

Then the flux is

$$F(U_l, U_r) = \frac{1}{2}(F_l + F_r) - \frac{1}{2}\sum_{j=1}^{3} \alpha_j |\bar{\lambda}_j| r_j$$

8.3 Roe scheme for general system

For a general system of conservation laws

$$\frac{\partial U}{\partial t} + \frac{\partial F}{\partial x} = 0, \quad U, F \in \mathbb{R}^n$$

assume that there exists an entropy-entropy flux pair (η, θ), with η strictly convex such that

$$\theta'(U) = \eta'(U)F'(U)$$

Theorem 8.2 (Harten and Lax)**.** *If the hyperbolic system has an entropy-entropy flux pair, then it admits a Roe-type linearization.*

Proof. Define the entropy variables

$$Q(U) = \eta'(U)^\top \in \mathbb{R}^n$$

and define the linear path

$$Q(\alpha) = Q_l + \alpha(Q_r - Q_l), \qquad \alpha \in [0, 1]$$

Then

$$\Delta F = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\alpha} F(Q(\alpha))\mathrm{d}\alpha = \left(\int_0^1 F'(Q(\alpha))\mathrm{d}\alpha\right) \Delta Q = B(Q_l, Q_r)\Delta Q$$

and B is a symmetric matrix. Using the linear path

$$U(\alpha) = U_l + \alpha(U_r - U_l), \qquad \alpha \in [0, 1]$$

we get

$$\Delta Q = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}\alpha} Q(U(\alpha))\mathrm{d}\alpha = \left(\int_0^1 Q'(U(\alpha))\mathrm{d}\alpha\right) \Delta U = C(U_l, U_r)\Delta U$$

and C is symmetric, positive definite matrix. If we take $\bar{A} = BC$, then we see that

$$C^{\frac{1}{2}} \bar{A} C^{-\frac{1}{2}} = C^{\frac{1}{2}} B C^{\frac{1}{2}}$$

so that $\bar{A}$ is similar to a symmetric matrix and hence has real eigenvalues. □

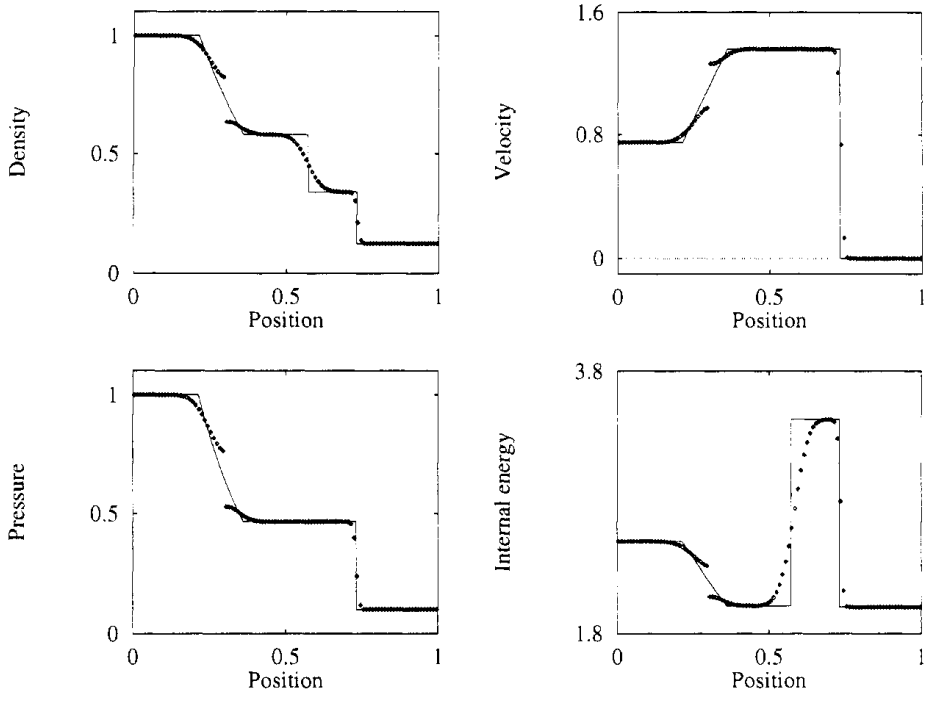


Fig. 11.4. Godunov's method with Roe's Riemann solver (no entropy fix) for Test 1, $x_0 = 0.3$. Numerical (symbol) and exact (line) solutions compared at time 0.2

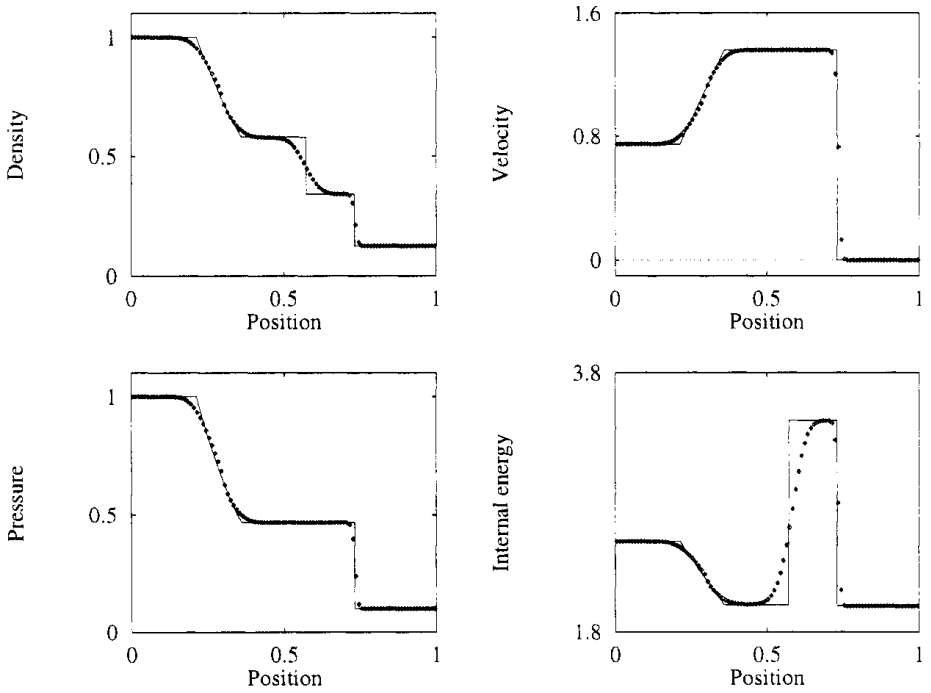


Fig. 11.5. Godunov's method with Roe's Riemann solver applied to Test 1, with $x_0 = 0.3$. Numerical (symbol) and exact (line) solutions are compared at time 0.2

Chapter 9

HLL and HLLC Riemann solvers

The Riemann solution for a system gives rise to as many waves as the number of independent variables in the system. One way to simplify the solution of the Riemann problem is to use a simpler wave model where only some waves are included. Clearly, we must include the slowest and fastest waves since they define the extent of the Riemann solution. We use only simple waves in our model, i.e., shocks or contact discontinuities.

9.1 HLL Riemann solver

The simplest model is to include only the slowest and fastest waves in the solution [2]. Let us assume that we have an estimate of these speeds denoted by S_l, S_r with $S_l < S_r$, and let us consider the subsonic case

$$S_l < 0 < S_r$$

There is one intermediate state which we denote by U_*. (FIGURE) To compute the intermediate state, let us integrate the conservation law over the space-time region $(S_l\Delta t, S_r\Delta t) \times (0, \Delta t)$.

$$\int_0^{\Delta t} \int_{S_l\Delta t}^{S_r\Delta t} (U_t + F_x)\mathrm{d}x\mathrm{d}t = 0$$

Integrate in t in the first term and x in the second term

$$\int_{S_l\Delta t}^{S_r\Delta t} U(x, \Delta t)\mathrm{d}x - \int_{S_l\Delta t}^{S_r\Delta t} U(x, 0)\mathrm{d}x + \int_0^{\Delta t} [F(S_r\Delta t, t) - F(S_l\Delta t, t)]\mathrm{d}t = 0$$

$$\int_{S_l\Delta t}^{S_r\Delta t} U_*\mathrm{d}x - \int_{S_l\Delta t}^{0} U_l\mathrm{d}x - \int_0^{S_r\Delta t} U_r\mathrm{d}x + \int_0^{\Delta t} [F_r - F_l]\mathrm{d}t = 0$$

$$U_*(S_r - S_l)\Delta t - (0 - S_l\Delta t)U_l - (S_r\Delta t - 0)U_r + (F_r - F_l)\Delta t = 0$$

Canceling Δt throughout and solving for U_* yields

$$U_* = \frac{S_rU_r - S_lU_l - (F_r - F_l)}{S_r - S_l}$$

What we require for the finite volume scheme is the flux across $x = 0$ line. Integrate the conservation law over the space-time region $(S_l\Delta t, 0) \times (0, \Delta t)$

$$\int_0^{\Delta t} \int_{S_l\Delta t}^{0} (U_t + F_x)\mathrm{d}x\mathrm{d}t = 0$$

$$\int_{S_l\Delta t}^{0} U_*\mathrm{d}x - \int_{S_l\Delta t}^{0} U_l\mathrm{d}x + \int_0^{\Delta t} [F_* - F_l]\mathrm{d}t = 0$$

$$F_* = F_l + S_l(U_* - U_l)$$

Alternately, we can integrate over the region $(0, S_r\Delta t) \times (0, \Delta t)$ to obtain

$$F_* = F_r + S_r(U_* - U_r)$$

However the two expressions for F_* yield the same value as can be checked by substituting the expression for U_*

$$F_* = \frac{S_r F_l - S_l F_r + S_l S_r (U_r - U_l)}{S_r - S_l}$$

Note that $F_* \neq F(U_*)$ since we have a non-linear conservation law.

Alternate derivation We can write the jump conditions across the two waves S_l, S_r

$$F_* - F_l = S_l(U_* - U_l), \qquad F_r - F_* = S_r(U_r - U_*)$$

Adding the two equations yields the formula for U_* and eliminating U_* from any one of the jump conditions yields the formula for the flux. This shows that the HLL solver satisfies the jump conditions across the two waves.

Numerical flux We have derived the flux for the subsonic case. If $0 < S_l < S_r$, then the solution along $x = 0$ is equal to the left state, and conversely if $S_l < S_r < 0$, then the solution along $x = 0$ is equal to the right state. Hence, the numerical flux is given by

$$F(U_l, U_r) = \begin{cases} F_l & 0 < S_l < S_r \\ F_r & S_l < S_r < 0 \\ F_* & \text{otherwise} \end{cases}$$

This is a consistent flux and has the upwind property.

9.1.1 Estimation of wave speeds, entropy condition

The only missing information is how to estimate the wave speeds S_l, S_r. We do not know the exact values of these speeds S_l^{true}, S_r^{true} since that would require exact solution of the Riemann problem. Einfeldt et al. suggest the following estimate

$$S_l = \min\{\lambda_1(U_l), \lambda_1(\bar{U})\}, \qquad S_r = \max\{\lambda_1(U_r), \lambda_m(\bar{U})\}$$

where $\bar{U}$ is the Roe-average state. For the Euler equations, these are given by

$$S_l = \min\{u_l - a_l, \bar{u} - \bar{a}\}, \qquad S_r = \max\{u_r + a_r, \bar{u} + \bar{a}\}$$

Theorem 9.1. *If the speed estimates satisfy*

$$S_l \leq S_l^{true}, \qquad S_r \geq S_r^{true}$$

i.e., the numerical Riemann fan completely encloses the true Riemann fan, then the entropy condition is satisfied by the numerical solution.

Remark 9.2. *From the above description, we see that the HLL solver can be applied to any system of conservation laws as long as we can estimate the wave speeds S_l, S_r. No knowledge of the eigenvectors is required in this solver.*

Remark 9.3. *If we choose*

$$S_l = -\lambda, \qquad S_r = \lambda$$

where $\lambda > 0$ is the maximum absolute speed in the Riemann problem, then the HLL flux becomes

$$F_* = \frac{1}{2}(F_l + F_r) - \frac{1}{2}\lambda(U_r - U_l)$$

which is the Rusanov flux. Thus the Rusanov scheme can be interpreted as an approximate Riemann solver of HLL-type with $S_r = -S_l = \lambda$. However, this scheme is not upwind since the Riemann fan is always subsonic.

9.1.2 Positivity of intermediate state

9.2 HLLC Riemann solver

The HLL solver does not include the contact wave which is linearly degenerate. Such waves will be dissipated by upwind schemes and they cannot steepen themselves unlike non-linear waves. Hence including contact waves in the approximate Riemann solver will enhance the accuracy of solutions. The HLLC solver [8] uses a wave model with three waves with speeds $S_l < S_* < S_r$ and two intermediate states $U_{*l}.U_{*r}$. (FIGURE) The tangential velocity does not change across shocks so that

$$v_{*l} = v_l, \qquad w_{*l} = w_l$$

$$v_{*r} = v_r, \qquad w_{*r} = w_r$$

The pressure and normal velocity are continuous across the contact wave

$$p_{*l} = p_{*r} = p_*, \qquad u_{*l} = u_{*r} = u_*$$

and the speed of the contact wave coincides with the intermediate velocity $S_* = u_*$.

The jump conditions across the S_l, S_r waves are

$$F_{*\alpha} - F_\alpha = S_\alpha(U_{*\alpha} - U_\alpha), \qquad \alpha = l, r$$

or

$$F_{*\alpha} - S_\alpha U_{*\alpha} = F_\alpha - S_\alpha U_\alpha$$

or in full

$$\begin{bmatrix} \rho_{*\alpha}u_* \\ p_* + \rho_{*\alpha}u_*^2 \\ \rho_{*\alpha}u_*v_\alpha \\ \rho_{*\alpha}u_*w_\alpha \\ (E_{*\alpha} + p_*)u_* \end{bmatrix} - S_\alpha \begin{bmatrix} \rho_{*\alpha} \\ \rho_{*\alpha}u_* \\ \rho_{*\alpha}v_\alpha \\ \rho_{*\alpha}w_\alpha \\ E_{*\alpha} \end{bmatrix} = \begin{bmatrix} \rho_\alpha u_\alpha \\ p_\alpha + \rho_\alpha u_\alpha^2 \\ \rho_\alpha u_\alpha v_\alpha \\ \rho_\alpha u_\alpha w_\alpha \\ (E_\alpha + p_\alpha)u_\alpha \end{bmatrix} - S_\alpha \begin{bmatrix} \rho_\alpha \\ \rho_\alpha u_\alpha \\ \rho_\alpha v_\alpha \\ \rho_\alpha w_\alpha \\ E_\alpha \end{bmatrix}$$

The first jump condition yields

$$\rho_{*\alpha} = \left(\frac{S_\alpha - u_\alpha}{S_\alpha - u_*}\right)\rho_\alpha \tag{9.1}$$

With this solution, the third and fourth conditions are satisfied. From the second jump condition

$$\begin{aligned} p_* &= p_\alpha + \rho_\alpha u_\alpha(u_\alpha - S_\alpha) + \rho_{*\alpha}u_*(S_\alpha - u_*) \\ &= p_\alpha + \rho_\alpha u_\alpha(u_\alpha - S_\alpha) + \rho_\alpha u_*(S_\alpha - u_\alpha) \quad \text{using (9.1)} \\ &= p_\alpha + \rho_\alpha(S_\alpha - u_\alpha)(u_* - u_\alpha) \end{aligned}$$

We get two estimates of pressure p_* from the l, r states and equating the two

$$p_l + \rho_l(S_l - u_l)(u_* - u_l) = p_r + \rho_r(S_r - u_r)(u_* - u_r)$$

gives us a formula for u_*

$$u_* = \frac{\rho_r u_r(S_r - u_r) - \rho_l u_l(S_l - u_l) - (p_r - p_l)}{\rho_r(S_r - u_r) - \rho_l(S_l - u_l)}$$

Once u_* is known, p_* can be computed using either $\alpha = l$ or $\alpha = r$. Finally, from the last jump condition we obtain

$$E_{*\alpha} = \frac{1}{S_\alpha - u_*}[(S_\alpha - u_\alpha)E_\alpha + p_*u_* - p_\alpha u_\alpha]$$

This finishes the determination of all the intermediate states and the flux is given by

$$F(U_l, U_r) = \begin{cases} F_l & 0 < S_l \\ F_r & S_r < 0 \\ F_{*l} = F_l + S_l(U_{*l} - U_l) & S_l < 0 < u_* \\ F_{*r} = F_r + S_r(U_{*r} - U_r) & u_* < 0 < S_r \end{cases}$$

Remark 9.4. *To simplify some expressions, we can define*

$$m_l = \rho_l(u_l - S_l), \qquad m_r = \rho_r(u_r - S_r)$$

Then

$$u_* = \frac{m_r u_r - m_l u_l - (p_r - p_l)}{m_r - m_l}, \qquad p_* = \frac{m_r p_l - m_l p_r + m_l m_r (u_r - u_l)}{m_r - m_l}$$

Chapter 10

FVM in 1-D: high order schemes

The cell average value is the basic unknown in the finite volume method. If we assume that the solution is given by piece-wise constant approximation based on cell-averages, then we obtain a first order accurate scheme. To obtain higher order accuracy, we can use a more accurate representation of the solution. Let us start with a piece-wise linear approximation in each cell

$$U_i(x,t) = U_i(t) + \frac{1}{\Delta x}(x - x_i)S_i(t), \qquad x \in I_{i+\frac{1}{2}}$$

where S_i is an approximation to the derivative

$$\frac{1}{\Delta x}S_i(t) \approx \frac{\partial U}{\partial x}(x_i, t)$$

Using this we can evaluate the state at each face

$$U^L_{i+\frac{1}{2}} = U_i + \frac{1}{2}S_i, \qquad U^R_{i+\frac{1}{2}} = U_{i+1} - \frac{1}{2}S_{i+1}$$

The semi-discrete finite volume scheme is given by

$$\Delta x \frac{\mathrm{d}U_i}{\mathrm{d}t} + F_{i+\frac{1}{2}}(t) - F_{i-\frac{1}{2}}(t) = 0$$

where

$$F_{i+\frac{1}{2}}(t) = F(U^L_{i+\frac{1}{2}}(t), U^R_{i+\frac{1}{2}}(t))$$

We have a higher order approximation in space and we combine this with a high order time integration scheme. This is known as *method of lines* approach.

10.0.1 Second order SSPRK

Stage 1

$$U^{(1)} = U^n - \frac{\Delta t}{\Delta x}R(U^n)$$

Stage 2

$$U^{n+1} = U^{(2)} = \frac{1}{2}U^n + \frac{1}{2}\left[U^{(1)} - \frac{\Delta t}{\Delta x}R(U^{(1)})\right]$$

10.0.2 Third order SSPRK

Stage 1

$$U^{(1)} = U^n - \frac{\Delta t}{\Delta x}R(U^n)$$

Stage 2

$$U^{(2)} = \frac{3}{4}U^n + \frac{1}{4}\left[U^{(1)} - \frac{\Delta t}{\Delta x}R(U^{(1)})\right]$$

Stage 3

$$U^{n+1} = U^{(3)} = \frac{1}{3}U^n + \frac{2}{3}\left[U^{(2)} - \frac{\Delta t}{\Delta x}R(U^{(2)})\right]$$

10.1 Estimate of reconstruction slope

We have several choices for this based on finite difference formulae

$$S_i^b = U_i - U_{i-1}, \qquad S_i^f = U_{i+1} - U_i$$

The reconstructed function must not oscillate and to achieve this, we can choose the smallest of the two estimates

$$S_i = \text{minmod}(S_i^b, S_i^f)$$

where

$$\text{minmod}(a,b) = \begin{cases} s\min(|a|,|b|) & s = \text{sign}(a) = \text{sign}(b) \\ 0 & \text{otherwise} \end{cases}$$

The total variation of the piece-wise linear reconstruction will not be larger than that of the piece-wise constant function. This leads to a very robust method, but it be can rather dissipative at shocks and smooth extrema.

A small improvement can be made using the central difference estimate

$$S_i^c = \frac{1}{2}(U_{i+1} - U_{i-1})$$

and

$$S_i = \text{minmod}(\beta S_i^b, S_i^c, \beta S_i^f), \qquad \beta \in [1,2]$$

where

$$\text{minmod}(a,b,c) = \begin{cases} s\min(|a|,|b|,|c|) & s = \text{sign}(a) = \text{sign}(b) = \text{sign}(c) \\ 0 & \text{otherwise} \end{cases}$$

If $\beta = 1$ we obtain the previous scheme. A value of $\beta = 2$ is likely to pick the central difference approximation and hence gives a more accurate reconstruction, though it might also generate some small oscillations.

Remark 10.1. *There is huge variety of other schemes to perform the solution reconstruction.*

10.2 Local truncation error

Let us study the accuracy in space by substituting a smooth exact solution into the finite volume scheme. If $U^L_{i+\frac{1}{2}}, U^R_{i+\frac{1}{2}}$ are obtained from reconstruction using cell averages of the exact solution and define

$$\Delta U^L_{i+\frac{1}{2}} = U(x_{i+\frac{1}{2}}) - U^L_{i+\frac{1}{2}}, \qquad \Delta U^R_{i+\frac{1}{2}} = U(x_{i+\frac{1}{2}}) - U^R_{i+\frac{1}{2}}$$

Let us assume that the reconstruction is exact for degree k polynomials so that

$$\Delta U^L_{i+\frac{1}{2}} = \mathcal{O}\left(\Delta x^{k+1}\right), \qquad \Delta U^R_{i+\frac{1}{2}} = \mathcal{O}\left(\Delta x^{k+1}\right)$$

Then

$$\begin{aligned} F(U^L_{i+\frac{1}{2}}, U^R_{i+\frac{1}{2}}) &= F(U(x_{i+\frac{1}{2}}) + \Delta U^L_{i+\frac{1}{2}}, U(x_{i+\frac{1}{2}}) + \Delta U^R_{i+\frac{1}{2}}) \\ &= F(U(x_{i+\frac{1}{2}}), U(x_{i+\frac{1}{2}})) + \frac{\partial}{\partial U_l} F(U(x_{i+\frac{1}{2}}), U(x_{i+\frac{1}{2}}))\Delta U^L_{i+\frac{1}{2}} \\ &\quad + \frac{\partial}{\partial U_r} F(U(x_{i+\frac{1}{2}}), U(x_{i+\frac{1}{2}}))\Delta U^R_{i+\frac{1}{2}} + \mathcal{O}\left(\Delta U\right)^2 \\ &= F(U(x_{i+\frac{1}{2}})) + \mathcal{O}\left(\Delta x^{k+1}\right) + \mathcal{O}\left(\Delta x^{2k+2}\right) \end{aligned}$$

The flux integral term in the scheme has the error

$$F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}} = F(U(x_{i+\frac{1}{2}})) - F(U(x_{i-\frac{1}{2}})) + \mathcal{O}\left(\Delta x^{k+1}\right)$$

Chapter 11

2-D finite volume method

Consider a 2-D conservation law of the form

$$U_t + F(U)_x + G(U)_y = 0$$

To approximate this by a numerical scheme, we first partition the domain $\Omega = [x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$ by a Cartesian mesh with cells of size $\Delta x, \Delta y$ along the x and y directions respectively. The finite volume method is however more general and allows the use of non-uniform grids. Let us index the cells by (i, j) and the faces by $(i + \frac{1}{2}, j)$, $(i, j + \frac{1}{2})$, etc., and denote this cell by

$$I_{i,j} = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$$

and its cell center is (x_i, y_j). Integrating the conservation law over one cell

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\int_{I_{i,j}} U \mathrm{d}x\mathrm{d}y + \int_{y_j-\frac{1}{2}\Delta y}^{y_j+\frac{1}{2}\Delta y} [F(x_{i+\frac{1}{2}}, y) - F(x_{i-\frac{1}{2}}, y)]\mathrm{d}y \\ + \int_{x_i-\frac{1}{2}\Delta x}^{x_i+\frac{1}{2}\Delta x} [G(x, y_{j+\frac{1}{2}}) - G(x, y_{j-\frac{1}{2}})]\mathrm{d}x = 0\end{aligned}$$

Define the cell average

$$U_{i,j} = \frac{1}{\Delta x \Delta y}\int_{I_{i,j}} U \mathrm{d}x\mathrm{d}y$$

and the average fluxes

$$F_{i+\frac{1}{2},j} = \frac{1}{\Delta y}\int_{y_j-\frac{1}{2}\Delta y}^{y_j+\frac{1}{2}\Delta y} F(x_{i+\frac{1}{2}}, y)\mathrm{d}y, \qquad G_{i,j+\frac{1}{2}} = \frac{1}{\Delta x}\int_{x_i-\frac{1}{2}\Delta x}^{x_i+\frac{1}{2}\Delta x} G(x, y_{j+\frac{1}{2}})\mathrm{d}x$$

The semi-discrete takes the form

$$\Delta x \Delta y \frac{\mathrm{d}U_{i,j}}{\mathrm{d}\Delta t} + [F_{i+\frac{1}{2},j} - F_{i-\frac{1}{2},j}]\Delta y + [G_{i,j+\frac{1}{2}} - G_{i,j-\frac{1}{2}}]\Delta x = 0$$

We have to still specify how to compute the fluxes which determines the order of accuracy of the scheme.

11.1 First order scheme

Let us consider the solution to be piece-wise constant given by the cell average values. There is a 1-D Riemann problem defined at each point of the vertical faces

$$U_t + F(U)_x = 0, \qquad U(x, y, t_n) = \begin{cases} U^n_{i,j} & x < x_{i+\frac{1}{2}} \\ U^n_{i+1,j} & x > x_{i+\frac{1}{2}} \end{cases}$$

from which an approximation to the flux can be obtained.

$$F_{i+\frac{1}{2},j} = \frac{1}{\Delta y}\int_{y_{j-\frac{1}{2}}}^{y_{j+\frac{1}{2}}} F(U_{i,j}, U_{i+1,j})\mathrm{d}y = F(U_{i,j}, U_{i+1,j})$$

Similarly, at any horizontal face, there is a 1-D Riemann problem

$$U_t + G(U)_y = 0, \qquad U(x, y, t_n) = \begin{cases} U^n_{i,j} & y < y_{j+\frac{1}{2}} \\ U^n_{i,j+1} & y > y_{j+\frac{1}{2}} \end{cases}$$

from which an approximation to the flux can be obtained

$$G_{i,j+\frac{1}{2}} = G(U_{i,j}, U_{i,j+1})$$

We are using 1-D numerical fluxes along x and y directions. Using forward Euler scheme in time, we get the first order finite volume method as

$$U^{n+1}_{i,j} = U^n_{i,j} - \frac{\Delta t}{\Delta x \Delta y} R^n_{i,j}$$

where

$$R^n_{i,j} = [F^n_{i+\frac{1}{2},j} - F^n_{i-\frac{1}{2},j}]\Delta y + [G^n_{i,j+\frac{1}{2}} - G^n_{i,j-\frac{1}{2}}]\Delta x$$

CFL condition A Fourier stability analysis yields the CFL condition

$$\Delta t \leq \frac{1}{\frac{\lambda_x}{\Delta x} + \frac{\lambda_y}{\Delta y}}$$

where λ_x, λ_y are maximum wave speeds along the x, y directions in the whole mesh

$$\lambda_x = \max_{i,j} \sigma(F'(U_{i,j}), \qquad \lambda_y = \max_{i,j} \sigma(G'(U_{i,j})$$

where the maximum is over all the cells in the grid. If $\Delta x = \Delta y = h$, $\lambda_x = \lambda_y = \lambda$ then

$$\Delta t \leq \frac{1}{2}\frac{h}{\lambda}$$

and the time step is half the value in 1-D.

Remark 11.1. *For Euler equations with velocity components (u, v) and sound speed a, the CFL condition is given by*

$$\Delta t \leq \frac{1}{\max_{i,j}\left[\frac{|u_{i,j}|+a_{i,j}}{\Delta x} + \frac{|v_{i,j}|+a_{i,j}}{\Delta y}\right]}$$

11.2 Implementation of scheme

11.3 Second order scheme

A second order scheme must have a local truncation error which is $O(h^2)$. Another way to characterize a second order scheme is if it is exact for a linear solution. Hence we have to use a piece-wise linear approximation of the solution inside each cell. The cell average value is a second order approximation to the cell-center value

$$U_{i,j} - U(x_i, y_j) = O(h^2)$$

so we take it as the point value for the purpose of developing the second order scheme.

Let us approximate the flux integral by the mid-point rule which is exact for linear functions,

$$F_{i+\frac{1}{2},j} = F(U^L_{i+\frac{1}{2},j}, U^R_{i+\frac{1}{2},j}), \qquad G_{i,j+\frac{1}{2}} = G(U^L_{i,j+\frac{1}{2}}, U^R_{i,j+\frac{1}{2}})$$

where $U^L_{i+\frac{1}{2},j}$ etc. are the trace values at the face mid-points obtained by some reconstruction process. For example

$$U^L_{i+\frac{1}{2},j} = U_{i,j} + \frac{1}{2}\Delta_x U_{i,j}, \qquad U^R_{i+\frac{1}{2},j} = U_{i+1,j} - \frac{1}{2}\Delta_x U_{i+1,j}$$

where $\Delta_x U_{i,j}$ is an estimate of the derivative of U at (x_i, y_j). The simplest choice is to take

$$\Delta_x U_{i,j} = \frac{1}{2}[U_{i+1,j} - U_{i-1,j}]$$

which is second order accurate for smooth functions. But this can lead to oscillations if we are near a shock or steep solution variation. Then some form of non-linear limiter must be used to perform the reconstruction. A simple choice is the minmod reconstruction given by

$$\Delta_x U_{i,j} = \text{minmod}\,[U_{i,j} - U_{i-1,j}, U_{i+1,j} - U_{i,j}]$$

For scalar problems, this leads to a TVD scheme which eliminates oscillations but the solutions can be rather diffused at shocks and smooth extrema since the method becomes first order in those regions. A slightly better version of this scheme is the MC limiter of van Leer given by

$$\Delta_x U_{i,j} = \text{minmod}\left[\beta(U_{i,j} - U_{i-1,j}), \frac{1}{2}(U_{i+1,j} - U_{i-1,j}), \beta(U_{i+1,j} - U_{i,j})\right], \quad \beta \in [1,2]$$

A larger value of β gives more sharp resolution of shocks, though there may be small oscillations.

A similar procedure is used along the y direction. The overall method can be viewed as first constructing a linear polynomial approximation

$$R_{i,j}(x,y) = U_{i,j} + \frac{x - x_i}{\Delta x}\Delta_x U_{i,j} + \frac{y - y_j}{\Delta y}\Delta_y U_{i,j}, \qquad (x,y) \in I_{i,j}$$

and using this to get the solution at face mid-points

$$U^L_{i+\frac{1}{2},j} = R_{i,j}(x_{i+\frac{1}{2}}, y_j), \qquad U^R_{i+\frac{1}{2},j} = R_{i+1,j}(x_{i+\frac{1}{2}}, y_j)$$

$$U^L_{i,j+\frac{1}{2}} = R_{i,j}(x_i, y_{j+\frac{1}{2}}), \qquad U^R_{i,j+\frac{1}{2}} = R_{i,j+1}(x_i, y_{j+\frac{1}{2}})$$

The high order spatial scheme must be coupled with a high order time integration scheme. For the second order scheme, a second or third order SSPRK scheme is a good choice.

Remark 11.2. *For second order schemes, it is not necessary to reconstruct conserved variables and we can use some other set of variables. For Euler equations, the primitive variables are a good choice as this ensures better control of oscillations. In particular, reconstructing pressure rather than total energy can help to ensure positivity of pressure.*

11.4 Higher order scheme

To achieve $(k+1)$'th order accuracy, i.e., $\|U_{exact} - U_{num}\| = \mathcal{O}\left(\Delta x^{k+1}\right)$, we must ensure three properties are satisfied.

1. Reconstruct the solution inside each cell to $(k+1)$'th order accuracy. This is satisfied if the reconstruction uses degree k polynomials.

2. Calculate the flux integral using a quadrature rule which is exact for degree k polynomials.

3. Use a $(k+1)$'th order accurate time integration scheme.

For second order accuracy ($k = 1$), we used linear reconstruction and mid-point quadrature, both of which are exact for a linear solution.

11.4.1 Solution reconstruction

Our solution is known in terms of cell averages using which we must reconstruct a polynomial inside each cell.

$$U(x,y) \approx R_{i,j}(x,y), \qquad (x,y) \in I_{i,j}$$

We would like the polynomial $R_{i,j}$ to have the correct cell average value

$$\frac{1}{\Delta x \Delta y} \int_{I_{i,j}} R_{i,j}(x,y) \mathrm{d}x\mathrm{d}y = U_{i,j} \tag{11.1}$$

We need more equations to determine the polynomial. E.g., for third order accuracy, we need a quadratic polynomial which has six coefficients. We use an interpolation or least-squares approach to find the polynomial. For interpolation, we need as many equations as the number of unknowns. For least-squares, we use more equations than unknowns. For example, if we use the 3×3 cells, we can solve the following least-squares problem

$$\min \sum_{I \in S} \left[\frac{1}{|I|} \int_I R_{i,j} \mathrm{d}x\mathrm{d}y - U_I \right]^2 \qquad \text{such that (11.1) is satisfied}$$

where

$$S = \{I_{i-1,j}, I_{i+1,j}, I_{i,j-1}, I_{i,j+1}, I_{i-1,j-1}, I_{i+1,j-1}, I_{i+1,j+1}, I_{i-1,j+1}\}$$

where the constrained minimization is performed wrt the six coefficients in the quadratic polynomial.

Of course, we need to implement some form of limiter in case of discontinuous solutions !!!

11.4.2 Flux quadrature

Once the solution is reconstructed in all cells, we can compute the flux integral on cell faces using suitable quadrature rule.

$$F_{i+\frac{1}{2},j} = \sum_q F(U^L(y_q), U^R(y_q)) \omega_q$$

where y_q are the location of quadrature points on the face, ω_q are the quadrature weights, and U^L, U^R denote some reconstructions from the left and right cells, respectively.

$$U^L(y_q) = R_{i,j}(x_{i+\frac{1}{2}}, y_q), \qquad U^R(y_q) = R_{i+1,j}(x_{i+\frac{1}{2}}, y_q)$$

For third order accuracy, the quadrature must be exact for quadratic polynomials and we could use Simpson rule which has three quadrature points. More efficient methods are based on Gauss-Legendre quadrature; for third order accuracy, 2-point GL quadrature is enough. In generate, the N-point GL quadrature is exact for polynomials of degree upto $2N-1$.

Chapter 12

ENO and WENO schemes

12.1 1-D grid

Given a grid in the domain $[a, b]$

$$a = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \ldots < x_{N-\frac{1}{2}} < x_{N+\frac{1}{2}} = b$$

define the cells, cell centers and cell sizes by

$$I_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}], \qquad x_i = \frac{1}{2}(x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}}), \qquad \Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}, \qquad i = 1, 2, \ldots, N$$

We denote the maximum cell size by

$$\Delta x = \max_{1 \le j \le N} \Delta x_j$$

12.2 First order finite volume scheme

For a general conservation law

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

the semi-discrete finite volume scheme is given by

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + \frac{1}{\Delta x}[\hat{f}_{i+\frac{1}{2}} - \hat{f}_{i-\frac{1}{2}}] = 0$$

where the numerical flux depends on

$$\hat{f}_{i+\frac{1}{2}} = \hat{f}(u_i, u_{i+1})$$

The numerical solution is piecewise constant in each cell, with possible jump at cell interfaces.

12.3 Higher order finite volume scheme

- We know only cell averages, information has been lost in the averaging process.
- **Reconstruct** solution in each cell by a linear polynomial

  $$p_i(x) = u_i + (x - x_i)s_i$$

  This preserves cell average value

  $$\frac{1}{\Delta x}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} p_i(x)\mathrm{d}x = u_i$$

- How to determine the slope s_i ?

$$s_i^- = \frac{u_i - u_{i-1}}{\Delta x}, \qquad s_i^+ = \frac{u_{i+1} - u_i}{\Delta x}$$

  Two candidate reconstructions corresponding to above two choices.

- Choose the one with smallest slope (to avoid discontinuities)

$$s_i = \text{minmod}(s_i^-, s_i^+)$$

  where

$$\text{minmod}(a,b) = \begin{cases} s\min(|a|,|b|) & \text{if } s = sign(a) = sign(b) \\ 0 & \text{otherwise} \end{cases}$$

- High order in space, semi-discrete scheme

$$\frac{\mathrm{d}u_i}{\mathrm{d}t} + \frac{1}{\Delta x}[\hat{f}_{i+\frac{1}{2}} - \hat{f}_{i-\frac{1}{2}}] = 0$$

  where

$$\hat{f}_{i+\frac{1}{2}} = \hat{f}(u^-_{i+\frac{1}{2}}, u^+_{i+\frac{1}{2}})$$

  and

$$u^-_{i+\frac{1}{2}} = p_i(x_{i+\frac{1}{2}}), \qquad u^+_{i+\frac{1}{2}} = p_{i+1}(x_{i+\frac{1}{2}})$$

- Forward Euler time discretization

$$u_i^{n+1} = u_i^n - \frac{\Delta t}{\Delta x}[\hat{f}^n_{i+\frac{1}{2}} - \hat{f}^n_{i-\frac{1}{2}}]$$

- Total variation

$$\text{TV}(u_h) = \sum_i |u_i - u_{i-1}|$$

- Total variation dimishing property

$$\text{TV}(u_h^{n+1}) \le \text{TV}(u_h^n)$$

12.4 ENO scheme

- Stencil for reconstruction is chosen adaptively, based on data smoothness
- Extend this idea to higher orders, beyond linear polynomial.
- We need a way to reconstruct solution from cell averages with a high degree polynomial
- This polynomial must be constructed using the smoothest data (stencil)

12.4.1 Polynomial reconstruction

For some smooth function $v(x)$, we are given its cell average values

$$v_i = \frac{1}{\Delta x_i}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} v(x)\mathrm{d}x$$

We want to find a polynomial $p_i(x)$ of degree at most $k-1$ such that it is a k-th order accurate approximation to $v(x)$, i.e.,

$$p_i(x) = v(x) + \mathcal{O}\left(\Delta x^k\right), \qquad x \in I_i, \qquad i = 1, 2, \ldots, N$$

We can evaluate this polynomial at the cell boundaries

$$v^+_{i-\frac{1}{2}} = p_i(x_{i-\frac{1}{2}}), \qquad v^-_{i+\frac{1}{2}} = p_i(x_{i+\frac{1}{2}})$$

which are required to compute the flux in a finite volume method. We want these values to be also k-th order accurate, i.e.,

$$v^+_{i-\frac{1}{2}} = v(x_{i-\frac{1}{2}}) + \mathcal{O}\left(\Delta x^k\right), \qquad v^-_{i+\frac{1}{2}} = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^k\right)$$

12.4.2 Primitive function

Define the primitive function

$$V(x) = \int_{-\infty}^{x} v(\xi)\mathrm{d}\xi, \qquad v(x) = \frac{\mathrm{d}}{\mathrm{d}x}V(x)$$

The lower bound in the above integral can be arbitrarily chosen. Now we know the values of the primitive function at the cell faces $x_{i+\frac{1}{2}}$ since

$$V(x_{i+\frac{1}{2}}) = \sum_{j=-\infty}^{i} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} v(\xi)\mathrm{d}\xi = \sum_{j=-\infty}^{i} v_j \Delta x_j$$

Using these known values of $V(x)$, we can construct a polynomial approximation to $V(x)$ of degree at most k, which we denote by $P_i(x)$, i.e.,

$$P_i(x) = V(x) + \mathcal{O}\left(\Delta x^{k+1}\right), \qquad x \in I_i$$

Then our desired polynomial p_i is given by

$$p_i(x) = \frac{\mathrm{d}}{\mathrm{d}x}P_i(x)$$

12.4.3 Construction of $P_j(x)$

To construct degree k polynomial, we need $k+1$ data points. First note that

$$V(x_{i+\frac{1}{2}}) - V(x_{i-\frac{1}{2}}) = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} v(x) = v_i \Delta x_i$$

Hence in the construction of $P_i(x)$ we must make use of the values $V_{i-\frac{1}{2}}$ and $V_{i+\frac{1}{2}}$. Let us take the data at the following $k+1$ points

$$x_{i-r-\frac{1}{2}}, \ldots, x_{i+s+\frac{1}{2}} \qquad \text{with} \qquad r+s+1 = k$$

The polynomial which interpolates this data can be obtained by Lagrange interpolation

$$P_i(x) = \sum_{m=0}^{k} V(x_{i-r+m-\frac{1}{2}}) \prod_{l=0, l\neq m}^{k} \frac{x - x_{i-r+l-\frac{1}{2}}}{x_{i-r+m-\frac{1}{2}} - x_{i-r+l-\frac{1}{2}}}$$

12.4.4 Newton form of reconstruction

Define the 0-th divided difference

$$V[x_{i-\frac{1}{2}}] := V(x_{i-\frac{1}{2}})$$

Then the j-th divided difference, for $j \geq 1$ are defined recursively by

$$V[\underbrace{x_{i-\frac{1}{2}}, \ldots, x_{i+j-\frac{1}{2}}}_{j+1 \text{ terms}}] := \frac{V[\overbrace{x_{i+\frac{1}{2}}, \ldots, x_{i+j-\frac{1}{2}}}^{j \text{ terms}}] - V[\overbrace{x_{i-\frac{1}{2}}, \ldots, x_{i+j-\frac{3}{2}}}^{j \text{ terms}}]}{x_{i+j-\frac{1}{2}} - x_{i-\frac{1}{2}}}$$

In particular, first divided difference is

$$V[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] = \frac{V(x_{i+\frac{1}{2}}) - V(x_{i-\frac{1}{2}})}{x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}} = v_i$$

Let us also define the divided differences of the cell average values by

$$v[x_i] := v_i, \qquad v[x_i, \ldots, x_{i+j}] = \frac{v[x_{i+1}, \ldots, x_{i+j}] - v[x_i, \ldots, x_{i+j-1}]}{x_{i+j} - x_i}$$

The polynomial P_i interpolating the primitive function V can be written in terms of the Newton divided differences of V

$$P_i(x) = \sum_{j=0}^{k} V[x_{i-r-\frac{1}{2}}, \ldots, x_{i-r+j-\frac{1}{2}}] \prod_{m=0}^{j-1} (x - x_{i-r+m-\frac{1}{2}})$$

Then we obtain $p_i(x) = P_i'(x)$ as

$$p_i(x) = \sum_{j=1}^{k} V[x_{i-r-\frac{1}{2}}, \ldots, x_{i-r+j-\frac{1}{2}}] \sum_{m=0}^{j-1} \prod_{l=0, l\neq m}^{j-1} (x - x_{i-r+l-\frac{1}{2}})$$

Note that only first and higher degree divided differences of V appear in the above expression. These can be written in terms of the divided differences of the cell averages.

12.4.5 Smoothness indicator

How to choose the left shift r and right shift s which determines the stencil for reconstructing the polynomial ? Since we have the freedom to choose the stencil, we would like to choose it so that the data used in the reconstruction is as smooth as possible. In particular, we would like to avoid choosing data across a shock.

To measure the smoothness of the data, we make use of Newton divided differences. The j-th divided difference has the property that

$$V[x_{i-\frac{1}{2}}, \ldots, x_{i+j-\frac{1}{2}}] = \frac{V^{(j)}(\xi)}{j!}, \qquad x_{i-\frac{1}{2}} < \xi < x_{i+j-\frac{1}{2}}$$

provided the function $V(x)$ is smooth in this stencil. If $V(x)$ is discontinuous at some point inside the stencil, then it is easy to verify that

$$V[x_{i-\frac{1}{2}}, \ldots, x_{i+j-\frac{1}{2}}] = \mathcal{O}\left(\frac{1}{\Delta x^j}\right)$$

Thus the divided difference is a measurement of the smoothness of the function inside the stencil.

12.5 ENO reconstruction

Our job is to find a stencil of $k+1$ consecutive points, which must include $x_{i-\frac{1}{2}}$ and $x_{i+\frac{1}{2}}$, such that $V(x)$ is *the smoothest* in this stencil, compared with other possible stencils.

First choose the two point stencil

$$\tilde{S}_2(i) = \{x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\}$$

The linear interpolation on this stencil is

$$P_i^1(x) = V[x_{i-\frac{1}{2}}] + V[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}](x - x_{i-\frac{1}{2}})$$

Next, we have two choices to increase the stencil. If we add the left value $x_{i-\frac{3}{2}}$, then we get

$$P_i^{2,l}(x) = P_j^1(x) + V[x_{i-\frac{3}{2}}, x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}](x - x_{i-\frac{1}{2}})(x - x_{i+\frac{1}{2}})$$

but if we add the right value $x_{i+\frac{3}{2}}$ we get

$$P_i^{2,r}(x) = P_j^1(x) + V[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}, x_{i+\frac{3}{2}}](x - x_{i-\frac{1}{2}})(x - x_{i+\frac{1}{2}})$$

Among these two quadratic polynomials, we choose the one which has the smallest divided difference. Hence we enhance the stencil as follows

$$\tilde{S}_3(i) = \begin{cases} \tilde{S}_2(i) \cup \{x_{i-\frac{3}{2}}\}, & |V[x_{i-\frac{3}{2}}, x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]| < |V[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}, x_{i+\frac{3}{2}}]| \\ \tilde{S}_2(i) \cup \{x_{i+\frac{3}{2}}\}, & \text{otherwise} \end{cases}$$

This procedure is repeated until the stencil contains $k+1$ points. Using $\tilde{S}_{k+1}(i)$ we obtain the polynomial $p_i(x) = P_i'(x)$ which is used to compute

$$v_{i-\frac{1}{2}}^+ = p_i(x_{i-\frac{1}{2}}), \qquad v_{i+\frac{1}{2}}^- = p_i(x_{i+\frac{1}{2}})$$

Finally the flux can be obtained as $g_{i+\frac{1}{2}} = g(v_{i+\frac{1}{2}}^-, v_{i+\frac{1}{2}}^+)$ which can be used in a finite volume scheme.

Remark 12.1. *Corresponding to each stencil $\tilde{S}(i)$ for primitive function V we have stencil $S(i)$ for conserved variable v. It is enough to build the stencil $S(i)$ to obtain the reconstruction polynomial $p_i(x)$.*

Remark 12.2. *For a uniform grid, it is enough to take undivided differences, which may be useful to avoid roundoff errors.*

12.6 TVB property of ENO reconstruction

For a piecewise smooth function $V(x)$, ENO reconstruction starting with the two point stencil $\tilde{S}_2(i) = \{x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\}$ has following properties:

- The accuracy condition

  $$P_i(x) = V(x) + \mathcal{O}\left(\Delta x^{k+1}\right), \qquad x \in I_i$$

  is valid for any cell I_i that does not contain a discontinuity. Hence we recover high order accuracy right upto the discontinuity.

- The reconstruction is *total variation bounded*. That is, there exists a function $z(x)$ satisfying

  $$z(x) = P_i(x) + \mathcal{O}\left(\Delta x^{k+1}\right), \qquad x \in I_i$$

  for any cell I_i such that

  $$TV(z) \leq TV(V)$$

  This implies that

  $$TV(p_i) \leq TV(v) + \mathcal{O}\left(\Delta x^k\right)$$

12.7 Sign property of ENO scheme

Theorem 12.3 (Fjordholm, Mishra, Tadmor)**.** *The ENO reconstruction $p_i(x)$ which is a polynomial of degree at most $k-1$ satisfies the following sign property*

$$\begin{aligned} &\text{if} \quad v_{i+1} - v_i \geq 0 \qquad \text{then} \quad v^+_{i+\frac{1}{2}} - v^-_{i+\frac{1}{2}} \geq 0 \\ &\text{if} \quad v_{i+1} - v_i \leq 0 \qquad \text{then} \quad v^+_{i+\frac{1}{2}} - v^-_{i+\frac{1}{2}} \leq 0 \end{aligned}$$

In particular, if $v_i = v_{i+1}$ then the ENO reconstruction is continuous, i.e., $v^-_{i+\frac{1}{2}} = v^+_{i+\frac{1}{2}}$.

12.8 WENO scheme

ENO scheme has some deficiencies.

- The stencil might change due to roundoff errors since we are making a comparison of divided differences and choosing the one with the smallest absolute value. The scheme would be non-smooth due to randomly changing stencil from one cell to its neighbour, especially in regions where the solution is very smooth or nearly constant.
- Obtaining steady state solutions can be difficult.
- In the stencil selection process, k candidate stencils are considered, covering $2k-1$ cells, but only one of the stencils is actually used.
- If all the $2k-1$ cells in the potential stencils are used, one could get $(2k-1)$-th order accuracy in smooth regions.

12.8.1 Third order approximation for $v^L_{i+\frac{1}{2}}$: $k = 2$

The candidate stencils are $S_0 = \{i, i+1\}$ and $S_1 = \{i-1, i\}$. (FIGURE) Using S_0, find linear polynomial p_0 such that

$$\int_{I_i} p_0(x)\mathrm{d}x = v_i \Delta x, \qquad \int_{I_{i+1}} p_0(x)\mathrm{d}x = v_{i+1}\Delta x$$

If we take p_0 of the form

$$p_0(x) = v_i + (x - x_i)s_0$$

then the first condition is already satisfied and we determine s_0 from second condition leading to

$$p_0(x) = v_i + (x - x_i)\frac{v_{i+1} - v_i}{\Delta x}, \qquad v^{(0)}_{i+\frac{1}{2}} = p_0(x_{i+\frac{1}{2}}) = \frac{1}{2}(v_i + v_{i+1}) = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^2\right)$$

Similarly using S_1, find p_1 such that

$$\int_{I_i} p_1(x)\mathrm{d}x = v_i \Delta x, \qquad \int_{I_{i-1}} p_1(x)\mathrm{d}x = v_{i-1}\Delta x$$

Hence

$$p_1(x) = v_i + (x - x_i)\frac{v_i - v_{i-1}}{\Delta x}, \qquad v^{(1)}_{i+\frac{1}{2}} = p_1(x_{i+\frac{1}{2}}) = \frac{3}{2}v_i - \frac{1}{2}v_{i-1} = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^2\right)$$

Using the stencil $\{i-1, i, i+1\}$ which is union of stencils S_0, S_1, construct a quadratic polynomial $p(x)$ such that

$$\frac{1}{\Delta x}\int_{I_{i-1}} p(x)\mathrm{d}x = v_{i-1}, \qquad \frac{1}{\Delta x}\int_{I_i} p(x)\mathrm{d}x = v_i, \qquad \frac{1}{\Delta x}\int_{I_{i+1}} p(x)\mathrm{d}x = v_{i+1}$$

which yields

$$p(x) = v_i + (x - x_i)\frac{v_{i+1} - v_{i-1}}{2\Delta x} + \left[(x - x_i)^2 - \frac{\Delta x^2}{12}\right]\frac{v_{i-1} - 2v_i + v_{i+1}}{2\Delta x^2}$$

and

$$v_{i+\frac{1}{2}} = p(x_{i+\frac{1}{2}}) = -\frac{1}{6}v_{i-1} + \frac{5}{6}v_i + \frac{1}{3}v_{i+1} = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^3\right)$$

We now try to write the quadratic approximation as a linear combination of the two linear approximations

$$d_0 p_0(x_{i+\frac{1}{2}}) + d_1 p_1(x_{i+\frac{1}{2}}) = p(x_{i+\frac{1}{2}})$$

$$-\frac{d_1}{2}v_{i-1} + \frac{d_0 + 3d_1}{2}v_i + \frac{d_0}{2}v_{i+1} = -\frac{1}{6}v_{i-1} + \frac{5}{6}v_i + \frac{1}{3}v_{i+1}$$

which has the unique solution

$$d_0 = \frac{2}{3}, \qquad d_1 = \frac{1}{3}$$

12.8.2 Fifth order approximation for $v^L_{i+\frac{1}{2}}$: $k = 3$

Let us take the candidate stencils as $S_0 = \{i, i+1, i+2\}$, $S_1 = \{i-1, i, i+1\}$ and $S_2 = \{i-2, i-1, i\}$ and construct the quadratic polynomial using each stencil. (FIGURE) We evaluate these polynomials at $x_{i+\frac{1}{2}}$ to obtain

$$\begin{aligned}
v^{(0)}_{i+\frac{1}{2}} &= \frac{1}{3}v_i + \frac{5}{6}v_{i+1} - \frac{1}{6}v_{i+2} = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^3\right) \\
v^{(1)}_{i+\frac{1}{2}} &= -\frac{1}{6}v_{i-1} + \frac{5}{6}v_i + \frac{1}{3}v_{i+1} = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^3\right) \\
v^{(2)}_{i+\frac{1}{2}} &= \frac{1}{3}v_{i-2} - \frac{7}{6}v_{i-1} + \frac{11}{6}v_i = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^3\right)
\end{aligned}$$

Now construct a degree 4 polynomial using $\{i-2, i-1, i, i+1, i+2\} = S_0 \cup S_1 \cup S_2$ and evaluate it at $x_{i+\frac{1}{2}}$

$$v_{i+\frac{1}{2}} = \frac{1}{30}v_{i-2} - \frac{13}{60}v_{i-1} + \frac{47}{60}v_i + \frac{9}{20}v_{i+1} - \frac{1}{20}v_{i+2}$$

We want to choose d_0, d_1, d_2 so that

$$\begin{aligned}
v_{i+\frac{1}{2}} &= d_0 v^{(0)}_{i+\frac{1}{2}} + d_1 v^{(1)}_{i+\frac{1}{2}} + d_2 v^{(2)}_{i+\frac{1}{2}} \\
&= \frac{d_2}{3}v_{i-2} - \frac{7d_2 + d_1}{6}v_{i-1} + \frac{2d_0 + 5d_1 + 11d_2}{6}v_i + \frac{2d_1 + 5d_0}{6}v_{i+1} - \frac{d_0}{6}v_{i+2}
\end{aligned}$$

The coefficients of $v_{i\pm 2}$ give

$$d_0 = \frac{3}{10}, \qquad d_2 = \frac{1}{10}$$

We can find d_1 from any other coefficient, say v_{i-1}

$$\frac{7d_2 + d_1}{6} = \frac{13}{60} \quad \Longrightarrow \quad d_1 = \frac{3}{5}$$

We can verify that the above solution matches the coefficients of v_i and v_{i+1} also. Thus we have a unique solution.

12.8.3 Non-linear blending

We have seen that each of the small stencils with k cells give an approximation with error $\mathcal{O}\left(\Delta x^k\right)$)

$$v^{(r)}_{i+\frac{1}{2}} = p_r(x_{i+\frac{1}{2}}) = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^k\right), \qquad p_r \in \mathbb{P}_{k-1}$$

while the union of all the stencils with $2k-1$ cells gives an approximation of $\mathcal{O}\left(\Delta x^{2k-1}\right)$

$$v_{i+\frac{1}{2}} = \sum_r d_r v^{(r)}_{i+\frac{1}{2}} = v(x_{i+\frac{1}{2}}) + \mathcal{O}\left(\Delta x^{2k-1}\right)$$

The coefficients d_r are called *linear weights*; they are positive and they sum to unity

$$\sum_r d_r = 1$$

We would like to use the high order approximation if the solution in smooth in the larger stencil. But if there is a shock somewhere in the large stencil, this approximation would not be optimal. In that case, it is better to use one of the smaller stencils which does not contain the shock. To achieve this let us modify the linear weights and write

$$v_{i+\frac{1}{2}} = \sum_r \omega_r v^{(r)}_{i+\frac{1}{2}}, \qquad \sum_r \omega_r = 1$$

If the solution is smooth in all the stencils then we want $\omega_r \approx d_r$ so that we get $(2k-1)$ accuracy. What should be the condition on ω_r to achieve this ?

$$\begin{aligned}
\sum_r \omega_r v^{(r)}_{i+\frac{1}{2}} - \sum_r d_r v^{(r)}_{i+\frac{1}{2}} &= \sum_r (\omega_r - d_r) v^{(r)}_{i+\frac{1}{2}} - \sum_r (\omega_r - d_r) v(x_{i+\frac{1}{2}}) \\
&= \sum_r (\omega_r - d_r)(v^{(r)}_{i+\frac{1}{2}} - v(x_{i+\frac{1}{2}})) \\
&= \sum_r (\omega_r - d_r)\mathcal{O}\left(\Delta x^k\right) = \mathcal{O}\left(\Delta x^{2k-1}\right)
\end{aligned}$$

A sufficient condition for the above to hold is

$$\omega_r = d_r + \mathcal{O}\left(\Delta x^{k-1}\right)$$

If there is a shock in the r'th stencil, then we want $\omega_r \approx 0$. Due to this requirement, the weights ω_r must depend on the function v which makes the approximation to be non-linear; due to this reason the ω_r are usually called *non-linear weights*.

12.8.4 WENO-JS scheme

To measure the smoothness of the polynomials, Jiang and Shu [**?**] proposed to use a Sobolev norm defined as

$$\beta_r = \sum_{l=1}^{k-1} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \Delta x^{2l-1} \left(\frac{\partial^l p_r(x)}{\partial x^l}\right)^2 \mathrm{d}x$$

and the β_r are called *smoothness indicators*. For third order scheme, the smoothness indicators are given by

$$\beta_0 = (v_{i+1} - v_i)^2, \qquad \beta_1 = (v_i - v_{i-1})^2$$

while for the fifth order scheme

$$\begin{aligned}
\beta_0 &= \frac{13}{12}(v_i - 2v_{i+1} + v_{i+2})^2 + \frac{1}{4}(3v_i - 4v_{i+1} + v_{i+2})^2 \\
\beta_1 &= \frac{13}{12}(v_{i-1} - 2v_i + v_{i+1})^2 + \frac{1}{4}(v_{i-1} - v_{i+1})^2 \\
\beta_2 &= \frac{13}{12}(v_{i-2} - 2v_{i-1} + v_i)^2 + \frac{1}{4}(v_{i-2} - 4v_{i-1} + 3v_i)^2
\end{aligned}$$

For smooth solution,

$$\frac{\partial^l p_r(x)}{\partial x^l} = \mathcal{O}(1) \qquad \Longrightarrow \qquad \beta_r = \mathcal{O}\left(\Delta x^2\right) + \mathcal{O}\left(\Delta x^3\right) + \ldots = \mathcal{O}\left(\Delta x^2\right)$$

while for discontinuous solution

$$\frac{\partial^l p_r(x)}{\partial x^l} = \mathcal{O}\left(\Delta x^{-l}\right) \qquad \Longrightarrow \qquad \beta_r = \mathcal{O}(1) + \mathcal{O}(1) + \ldots = \mathcal{O}(1)$$

Jiang and Shu define the non-linear weights as

$$\omega_r = \frac{\alpha_r}{\sum_s \alpha_s}, \qquad \alpha_r = \frac{d_r}{(\epsilon + \beta_r)^2}, \qquad 0 < \epsilon \ll 1$$

The ϵ avoids division by zero. Let us consider the fifth order scheme. A Taylor expansion around $x = x_i$ shows that for smooth solutions

$$\beta_s = [v'(x_i)\Delta x]^2 + \mathcal{O}\left(\Delta x^4\right)$$

Solution smooth in all stencils In this case

$$\beta_s = [v'(x_i)\Delta x]^2[1 + \mathcal{O}\left(\Delta x^2\right)] \qquad \Longrightarrow \qquad \frac{1}{\beta_s^2} = \frac{1}{[v'(x_i)\Delta x]^4}[1 + \mathcal{O}\left(\Delta x^2\right)]$$

and since $\sum_s d_s = 1$

$$\sum_s \alpha_s = \sum_s \frac{d_s}{\beta_s^2} = \frac{1}{[v'(x_i)\Delta x]^4} + \mathcal{O}\left(\Delta x^{-2}\right) \quad \Longrightarrow \quad \frac{1}{\sum_s \alpha_s} = [v'(x_i)\Delta x]^4[1 + \mathcal{O}\left(\Delta x^2\right)]$$

Hence

$$\omega_r = \frac{d_r}{[v'(x_i)\Delta x]^4}[1 + \mathcal{O}\left(\Delta x^2\right)] \times [v'(x_i)\Delta x]^4[1 + \mathcal{O}\left(\Delta x^2\right)] = d_r[1 + \mathcal{O}\left(\Delta x^2\right)]$$

which is sufficient for fifth order accuracy.

Shock in r'th stencil Suppose there is shock in stencil S_0 between cell $i+1$ and $i+2$, so that case $\beta_0 = \mathcal{O}(1)$. Since solution is smooth in S_1, S_2, then

$$\sum_s \alpha_s = \mathcal{O}(1) + \mathcal{O}\left(\Delta x^{-4}\right) + \mathcal{O}\left(\Delta x^{-4}\right) \quad \Longrightarrow \quad \frac{1}{\sum_s \alpha_s} = \mathcal{O}\left(\Delta x^4\right)\left[\mathcal{O}(1) + \mathcal{O}\left(\Delta x^4\right)\right] = \mathcal{O}\left(\Delta x^4\right)$$

Hence

$$\omega_0 = \frac{d_0}{\beta_0^2} \times \frac{1}{\sum_s \alpha_s} = \mathcal{O}(1) \times \mathcal{O}\left(\Delta x^4\right) = \mathcal{O}\left(\Delta x^4\right) \approx 0$$

There is no shock in stencil S_2 so that

$$\beta_2 = \mathcal{O}\left(\Delta x^2\right), \qquad \omega_2 = \frac{d_2}{\beta_2^2} \times \mathcal{O}\left(\Delta x^4\right) = \mathcal{O}(1)$$

and similarly, $omega_1 = \mathcal{O}(1)$. Hence the WENO gives

$$v_{i+\frac{1}{2}} = \omega_0 p_0(x_{i+\frac{1}{2}}) + \omega_1 p_1(x_{i+\frac{1}{2}}) + \omega_2 p_2(x_{i+\frac{1}{2}}) = \omega_1 p_1(x_{i+\frac{1}{2}}) + \omega_2 p_2(x_{i+\frac{1}{2}})$$

and we get third order accuracy.

Remark 12.4. *We can now summarize the WENO algorithm as follows. The third order WENO reconstruction is given by*

$$v^L_{i+\frac{1}{2}} = WENO3(v_{i-1}, v_i, v_{i+1})$$

while the fifth order WENO reconstruction is given by

$$v^L_{i+\frac{1}{2}} = WENO5(v_{i-2}, v_{i-1}, v_i, v_{i+1}, v_{i+2})$$

Note that the stencil is biased to the left of $x_{i+\frac{1}{2}}$ in the sense that we have more points on the left of $x_{i+\frac{1}{2}}$ than to its right. To compute the values on the right of the face, we can use the same method as above but pass the correct values; e.g., the fifth order approximation is

$$v^R_{i+\frac{1}{2}} = WENO5(v_{i+3}, v_{i+2}, v_{i+1}, v_i, v_{i-1})$$

12.8.5 Characteristic variable reconstruction

For systems of equations, it is beneficial to apply the reconstruction on characteristic variables rather than conserved variables. Locally, characteristic variables are propagated by waves and hence obey some maximum principle, which is not true for conserved variables. For linear systems, a TVD property can be proved if the limiting is applied to characteristic variables. Let U denote the vector of conserved variables. Compute an average state at the face $i+\frac{1}{2}$, e.g.,

$$U_{i+\frac{1}{2}} = \frac{1}{2}(U_i + U_{i+1})$$

Compute the matrix of right ($R_{i+\frac{1}{2}}$) and left ($L_{i+\frac{1}{2}} = R^{-1}_{i+\frac{1}{2}}$) eigenvectors of the flux Jacobian at the average state and they satisfy

$$R_{i+\frac{1}{2}} L_{i+\frac{1}{2}} = I$$

Let us explain how to compute $U^L_{i+\frac{1}{2}}$ using the fifth order WENO scheme.

1. Convert conserved to characteristic variables

$$W_j = W_j^{(i+\frac{1}{2})} = L_{i+\frac{1}{2}} U_j, \qquad j \in \{i-2, i-1, i, i+1, i+2, i+3\}$$

2. Compute WENO reconstruction of each of the characteristic variables[1]

$$W^L_{i+\frac{1}{2}} = WENO5(W_{i-2}, W_{i-1}, W_i, W_{i+1}, W_{i+2})$$

$$W^R_{i+\frac{1}{2}} = WENO5(W_{i+3}, W_{i+2}, W_{i+1}, W_i, W_{i-1})$$

3. Convert to conserved variables

$$U^L_{i+\frac{1}{2}} = R_{i+\frac{1}{2}} W^L_{i+\frac{1}{2}}, \qquad U^R_{i+\frac{1}{2}} = R_{i+\frac{1}{2}} W^R_{i+\frac{1}{2}}$$

12.9 Example: error convergence

$$u_t + u_x = 0, \qquad x \in (-1, +1)$$
$$u(x, 0) = \sin(\pi x)$$

Accuracy on $u_t + u_x = 0$ *with* $u_0(x) = \sin(\pi x)$.

Method	N	L_∞ error	L_∞ order	L_1 error	L_1 order
WENO-5	10	2.98e-2	-	1.60e-2	-
	20	1.45e-3	4.36	7.41e-4	4.43
	40	4.58e-5	4.99	2.22e-5	5.06
	80	1.48e-6	4.95	6.91e-7	5.01
	160	4.41e-8	5.07	2.17e-8	4.99
	320	1.35e-9	5.03	6.79e-10	5.00

N = number of cells, error measured at time $t = 1$

[1] We do not show the superscript.

12.10 Example: Periodic advection of isentropic vortex

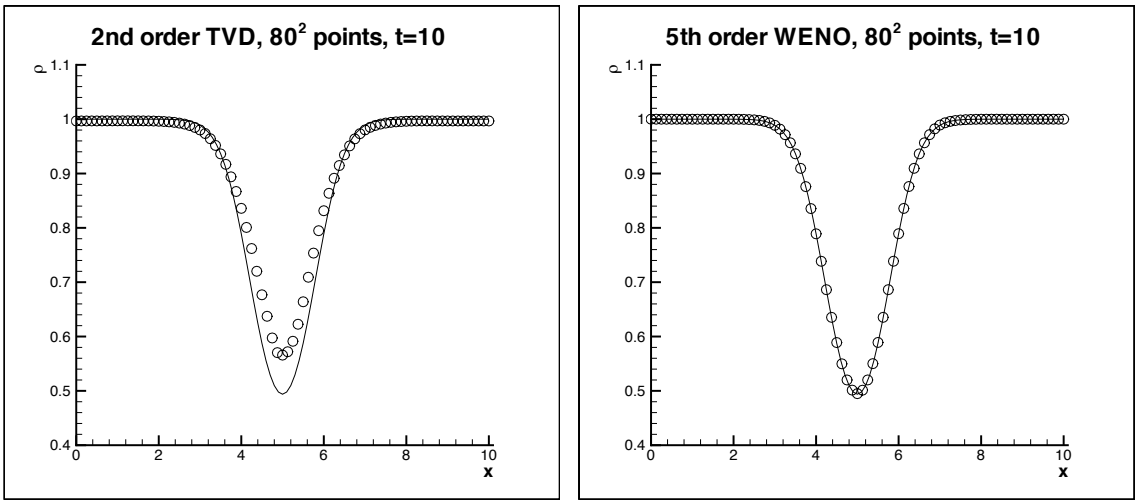


FIG. 1.1. *Vortex evolution. Cut at $x = 5$. Density ρ. 80^2 uniform mesh. $t = 10$ (after one time period). Solid: exact solution; circles: computed solution. Left: second order TVD scheme; right: fifth order WENO scheme.*

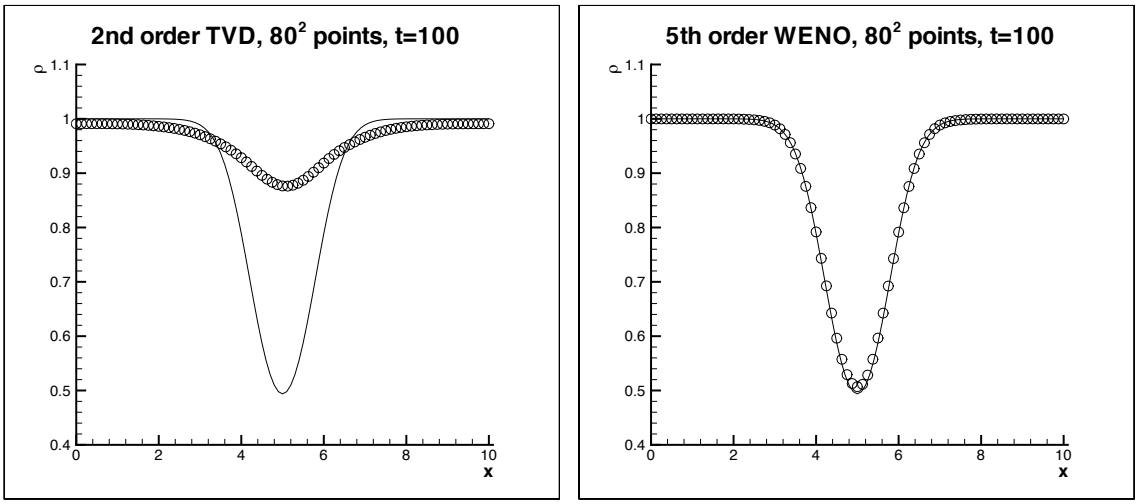


FIG. 1.2. *Vortex evolution. Cut at $x = 5$. Density ρ. 80^2 uniform mesh. $t = 100$ (after 10 time periods). Solid: exact solution; circles: computed solution. Left: second order TVD scheme; right: fifth order WENO scheme.*

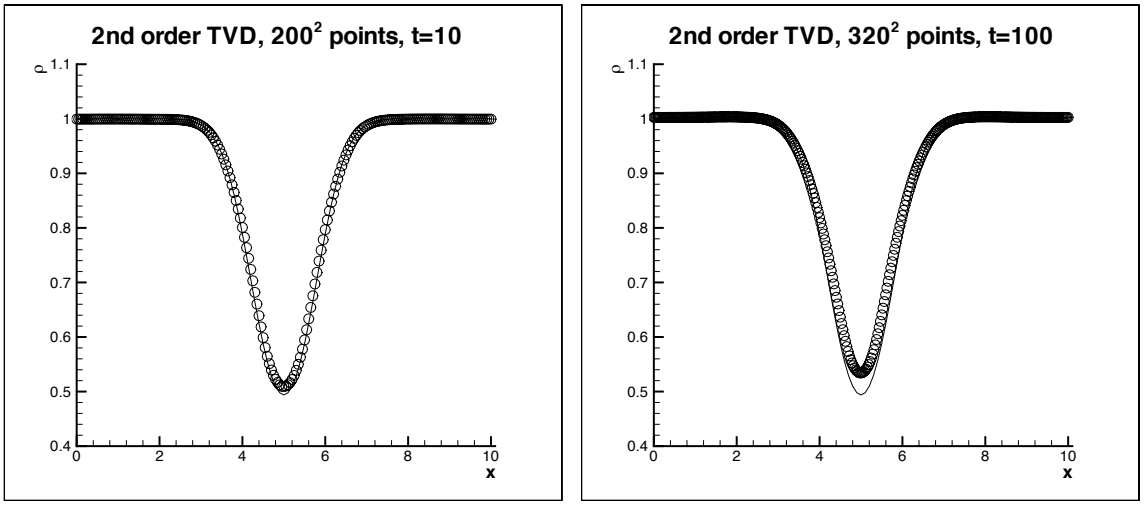


FIG. 1.3. *Vortex evolution. Cut at $x = 5$. Density ρ. Second order TVD scheme. Solid: exact solution; circles: computed solution. Left: 200^2 uniform mesh, $t = 10$ (after one time period); right: 320^2 uniform mesh, $t = 100$ (after 10 time periods).*

12.11 Finite volume WENO in 2-D

12.12 Finite difference WENO scheme

In the finite volume scheme, we reconstruct the solution at the cell face using the cell averages and use any numerical flux function. In the finite difference approach, we assume that we have the solution at the grid points and we aim to approximate the derivative of the flux at the grid points. This is achieved by reconstructing the flux rather than the solution. The semi-discrete form of the finite difference scheme is

$$\frac{\mathrm{d}v_i}{\mathrm{d}t} + \frac{1}{h}[f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}] = 0$$

The scheme has order of accuracy p if

$$\frac{1}{h}[f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}] = \frac{\partial f}{\partial x}(x_i) + \mathcal{O}\left(h^p\right)$$

If we had a function $F(x)$ such that

$$f(x) = \frac{1}{h}\int_{x-h/2}^{x+h/2} F(\xi)\mathrm{d}\xi$$

Then

$$\frac{\partial f}{\partial x}(x_i) = \frac{F(x_{i+\frac{1}{2}}) - F(x_{i-\frac{1}{2}})}{h}$$

and we can take

$$f_{i+\frac{1}{2}} = F(x_{i+\frac{1}{2}})$$

The point value of the flux is

$$f_i = f(x_i) = \frac{1}{h}\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} F(\xi)\mathrm{d}\xi$$

is the cell average value of the unknown function $F(x)$ on the i'th cell and this value is known to us. We can apply the WENO idea to construct an approximation to $F(x_{i+\frac{1}{2}})$. However, to obtain a stable scheme, we must ensure some form of upwinding or correct domain of dependence of the stencil. For this purpose let us split the flux into two parts

$$f = f^+ + f^-$$

where

$$\frac{\partial f^-}{\partial u} < 0, \qquad \frac{\partial f^+}{\partial u} > 0$$

The simplest splitting is given by Lax-Friedrich splitting

$$f^- = \frac{1}{2}[f(u) - \lambda u], \qquad f^+ = \frac{1}{2}[f(u) + \lambda u]$$

where

$$\lambda = \max_u |f'(u)|$$

with the maximum being taken over the range of values of u in our grid or locally at each face. The split fluxes $f^{\pm}_{i+\frac{1}{2}}$ are computed using a WENO scheme. Since f^+ represents waves moving with positive speed, we will use a stencil which is biased towards the left of $x_{i+\frac{1}{2}}$ to compute this flux, and similarly, for $f^-_{i+\frac{1}{2}}$ we use a stencil which is biased towards the right of $x_{i+\frac{1}{2}}$. Then the flux will be computed as

$$f_{i+\frac{1}{2}} = f^+_{i+\frac{1}{2}} + f^-_{i+\frac{1}{2}}$$

Using the fifth order WENO scheme, we can compute the split fluxes as

$$f^+_{i+\frac{1}{2}} = WENO5(f^+_{i-2}, f^+_{i-1}, f^+_i, f^+_{i+1}, f^+_{i+2})$$

$$f^-_{i+\frac{1}{2}} = WENO5(f^-_{i+3}, f^-_{i+2}, f^-_{i+1}, f^-_i, f^-_{i-1})$$

12.12.1 Analysis of WENO-JS scheme

In this section, we drop the superscript $\pm$. Recall that we have three approximations $f^r_{i\pm\frac{1}{2}}$ coming from quadratic polynomials which are third order accurate

$$f^r_{i\pm\frac{1}{2}} = F(x_{i\pm\frac{1}{2}}) + A_r \Delta x^3 + \mathcal{O}\left(\Delta x^4\right)$$

Their combination with linear weights is fifth order accurate

$$\sum_r d_r f^r_{i\pm\frac{1}{2}} = F(x_{i\pm\frac{1}{2}}) + B\Delta x^5 + \mathcal{O}\left(\Delta x^6\right)$$

WENO scheme combines these three values in a non-linear way

$$f_{i\pm\frac{1}{2}} = \sum_r \omega_r^{\pm} f^r_{i\pm\frac{1}{2}}$$

A Taylor expansion of the smoothness indicators yields

$$\begin{aligned}
\beta_0 &= (f_i')^2\Delta x^2 + \left(\frac{13}{12}(f_i'')^2 - \frac{2}{3}f_i' f_i'''\right)\Delta x^4 + \left(\frac{13}{6}f_i'' f_i''' - \frac{1}{2}f_i' f_i''''\right)\Delta x^5 + \mathcal{O}\left(\Delta x^6\right) \\
\beta_1 &= (f_i')^2\Delta x^2 + \left(\frac{13}{12}(f_i'')^2 + \frac{2}{3}f_i' f_i'''\right)\Delta x^4 + \mathcal{O}\left(\Delta x^6\right) \\
\beta_2 &= (f_i')^2\Delta x^2 + \left(\frac{13}{12}(f_i'')^2 - \frac{2}{3}f_i' f_i'''\right)\Delta x^4 - \left(\frac{13}{6}f_i'' f_i''' - \frac{1}{2}f_i' f_i''''\right)\Delta x^5 + \mathcal{O}\left(\Delta x^6\right)
\end{aligned}$$

Add and subtract the linear weights

$$\begin{aligned}
f_{i\pm\frac{1}{2}} &= \sum_r d_r f^r_{i\pm\frac{1}{2}} + \sum_r (\omega_r^{\pm} - d_r) f^r_{i\pm\frac{1}{2}} \\
&= \left[F(x_{i\pm\frac{1}{2}}) + B\Delta x^5 + \mathcal{O}\left(\Delta x^6\right)\right] + \sum_r (\omega_r^{\pm} - d_r) f^r_{i\pm\frac{1}{2}}
\end{aligned}$$

since the linear weights are chosen to give fifth order accuracy. The second term is

$$\begin{aligned}
\sum_r (\omega_r^{\pm} - d_r) f^r_{i\pm\frac{1}{2}} &= \sum_r (\omega_r^{\pm} - d_r)\left(F(x_{i\pm\frac{1}{2}}) + A_r\Delta x^3 + \mathcal{O}\left(\Delta x^4\right)\right) \\
&= F(x_{i\pm\frac{1}{2}}) \underbrace{\sum_r (\omega_r^{\pm} - d_r)}_{=0} + \Delta x^3 \sum_r A_r(\omega_r^{\pm} - d_r) + \sum_r (\omega_r^{\pm} - d_r)\mathcal{O}\left(\Delta x^4\right)
\end{aligned}$$

The first term is zero since the weights sum to unity. The finite difference approximation is

$$\begin{aligned}
\frac{f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}}{\Delta x} &= \frac{F(x_{i+\frac{1}{2}}) - F(x_{i-\frac{1}{2}})}{\Delta x} + \mathcal{O}\left(\Delta x^5\right) + \frac{\sum_r (\omega_r^+ - d_r) f^r_{i+\frac{1}{2}} - \sum_r (\omega_r^- - d_r) f^r_{i-\frac{1}{2}}}{\Delta x} \\
&= f'(x_i) + \mathcal{O}\left(\Delta x^5\right) + \Delta x^2 \sum_r A_r(\omega_r^+ - \omega_r^-) \\
&\quad + \sum_r (\omega_r^+ - d_r)\mathcal{O}\left(\Delta x^3\right) - \sum_r (\omega_r^- - d_r)\mathcal{O}\left(\Delta x^3\right)
\end{aligned}$$

In case of smooth function, the WENO-JS weights satisfy

$$\omega_r^{\pm} = d_r + \mathcal{O}\left(\Delta x^2\right)$$

then the last two terms are of $\mathcal{O}\left(\Delta x^5\right)$.Using Taylor expansions we can show that

$$\Delta x^2 \sum_r A_r(\omega_r^+ - \omega_r^-) = \mathcal{O}\left(\Delta x^5\right)$$

and hence we get fifth order accuracy

12.12.2 WENO-Z scheme

The discussion of the WENO-JS scheme shows that the conclusions are not valid if $v'(x_i) = 0$, i.e., we have a local extremum at $x = x_i$. Let us assume this is the case but also that $f_i'' \neq 0$. Then

$$\beta_0 = \beta_2 = \frac{13}{12}[f_i''\Delta x^2]^2[1 + \mathcal{O}(\Delta x)], \qquad \beta_1 = \frac{13}{12}[f_i''\Delta x^2]^2[1 + \mathcal{O}\left(\Delta x^2\right)]$$

This leads to

$$\omega_0 = d_0 + \mathcal{O}(\Delta x), \qquad \omega_2 = d_2 + \mathcal{O}(\Delta x)$$

so that

$$\Delta x^2 \sum_r A_r(\omega_r^+ - \omega_r^-) = \Delta x^2 \sum_r A_r(\omega_r^+ - d_r + d_r - \omega_r^-) = \Delta x^2 \mathcal{O}(\Delta x) = \mathcal{O}\left(\Delta x^3\right)$$

and we get only third order accuracy.

A modification of the non-linear weights was proposed in [1] to take care of such extrema. Define

$$\tau = |\beta_0 - \beta_2|$$

If solution is smooth in all stencils then

$$\tau = \frac{13}{3}|f_i'' f_i'''|\Delta x^5 + \mathcal{O}\left(\Delta x^6\right)$$

Define smoothness indicators as

$$\beta_r^z = \frac{\beta_r + \epsilon}{\beta_r + \tau + \epsilon}, \qquad 0 < \epsilon \ll 1$$

Then the nonlinear weights are given by

$$\omega_r = \frac{\alpha_r}{\sum_s \alpha_s}, \qquad \alpha_s = \frac{d_s}{\beta_s^z} = d_s\left(1 + \frac{\tau}{\beta_s + \epsilon}\right)$$

The WENO-Z scheme yields fifth order accuracy even at smooth extrema.

12.13 Central WENO (CWENO) schemes

Appendix A

Euler test cases

A.1 1-D: linear advection

If the initial velocity and pressure are constant, then they remain constant for future times. The density is advected at the constant velocity. So an initial condition

$$\rho(x,0) = f(x), \qquad u(x,0) = u_0, \qquad p(x,0) = p_0$$

has exact solution

$$\rho(x,t) = f(x - u_0 t), \qquad u(x,t) = u_0, \qquad p(x,t) = p_0$$

Take $f(x)$ to be a periodic function and use periodic boundary conditions. E.g.

$$f(x) = 1 + \frac{1}{2}\sin(2\pi x), \qquad x \in [0,1]$$

with $u_0 = p_0 = 1$ and $\gamma = 1.4$.

A.2 1-D: Sod test

This is a Riemann problem [5] given in Table (A.1) which develops a left rarefaction, a contact wave and a right shock. Solve this on a domain $[0,1]$ with initial jump at $x = 0.5$, $\gamma = 1.4$ and upto time $t = 0.2$ using Neumann boundary conditions.

A.3 1-D: Sod test with sonic rarefaction

This is a modification of Sod test case where the flow reaches sonic state inside the expansion fan and is a good test to check entropy condition satisfaction. The domain is $[0,1]$ and initial conditions are given in Table (A.2). The solution is computed up to time $t = 0.2$ with $\gamma = 1.4$.

A.4 1-D: Shu-Osher test case

An initial jump separates a constant state on the left with a smooth density profile on the right. The solution involves interaction of a shock with a smooth profile and hence is challenging to get accurate solutions. Solve this on a domain $[-5,+5]$ with initial jump at $x = -4$ upto time $t = 1.8$. Use $\gamma = 1.4$ and Neumann boundary conditions. The initial condition is given in Table (A.3).

	ρ	u	p
$x < 0.5$	1.0	0.0	1.0
$x > 0.5$	0.125	0.0	0.1

Table A.1: Initial conditions for Sod test case

	ρ	u	p
$x < 0.3$	1.0	0.75	1.0
$x > 0.3$	0.125	0.0	0.1

Table A.2: Initial conditions for modified Sod test case with sonic rarefaction

	ρ	u	p
$x < -4$	3.857143	2.699369	10.33333
$x > -4$	$1.0 + 0.2\sin(5x)$	0.0	1.0

Table A.3: Initial conditions for Shu-Osher test case

A.5 1-D: 123 problem

The initial condition is given by [7]

$$(\rho, u, p) = \begin{cases} (1.0,\ -2.0,\ 0.4) & x < 0.5 \\ (1.0,\ +2.0,\ 0.4) & x > 0.5 \end{cases}$$

The computational domain is $[0, 1]$ and the final time is $t = 0.15$. The density becomes very small in the middle of the domain and can be challenging test in terms of maintaining positive density.

A.6 1-D: Interaction of blast waves

The problems involves interaction of two shock waves and the initial condition is given by [10]

$$(\rho, u, p) = \begin{cases} (1.0,\ 0.0,\ 1000.0) & x < 0.1 \\ (1.0,\ 0.0,\ 0.01) & 0.1 < x < 0.9 \\ (1.0,\ 0.0,\ 100.0) & x > 0.9 \end{cases}$$

The domain can be taken as $[0, 1]$ and the computations performed upto the time $t = 0.038$. Solid wall boundary conditions are used on both end points of the domain.

A.7 2-D: Isentropic vortex

The flow consists of a vortex that moves with a constant speed [11]. The initial condition is given by

$$\begin{aligned} \rho &= \left[1 - \frac{(\gamma-1)\beta^2}{8\gamma\pi^2}\exp(1-r^2)\right]^{1/(\gamma-1)} \\ u &= M\cos\alpha - \frac{\beta}{2\pi}\exp(\tfrac{1}{2}(1-r^2))(y-y_0) \\ v &= M\sin\alpha + \frac{\beta}{2\pi}\exp(\tfrac{1}{2}(1-r^2))(x-x_0) \\ p &= \rho^\gamma \end{aligned}$$

where

$$r^2 = (x-x_0)^2 + (y-y_0)^2$$

Here (x_0, y_0) is the center of the vortex. The vortex moves at an angle α to the x axis at a constant speed of M. The flow has constant entropy everywhere and for all times.

Take a domain of $[-5, +5] \times [-5, +5]$ with these parameters

$$(x_0, y_0) = (0, 0), \quad \beta = 5, \quad M = 0.5, \quad \alpha = \pi/4$$

Use periodic boundary conditions and run up to a time of $t = 10\sqrt{2}/M$; the vortex would have come back to its initial position. Note that for high order methods, it is better to use a larger domain since the exponential functions may not have decayed at the boundary.

A.8 2-D: Shock reflection

This problem consist of a shock which hits the bottom wall and gets reflected and the flow eventually reaches a steady state. The domain is $[0, 4] \times [0, 1]$ and the initial condition can be taken as

$$\rho = 1, \qquad (u, v) = (2.9, 0.0), \qquad p = 1/\gamma$$

with $\gamma = 1.4$. The flow enters the domain from the left and top sides. The conditions at the left side are same as the above initial conditions while the conditions on the top are

$$\rho = 1.69997, \qquad (u, v) = (2.61934, -0.50632), \qquad p = 1.52819$$

The botton side is a solid wall and we can use Neumann conditions on the right side.

Bibliography

[1] R. Borges, M. Carmona, B. Costa, and W. S. Don, *An improved weighted essentially non-oscillatory scheme for hyperbolic conservation laws*, Journal of Computational Physics, 227 (2008), pp. 3191–3211.

[2] A. Harten, P. D. Lax, and B. van Leer, *On Upstream Differencing and Godunov-Type Schemes for Hyperbolic Conservation Laws*, SIAM Review, 25 (1983), pp. 35–61.

[3] P. L. Roe, *Approximate Riemann solvers, parameter vectors, and difference schemes*, Journal of Computational Physics, 43 (1981), pp. 357–372.

[4] V. Rusanov, *The calculation of the interaction of non-stationary shock waves and obstacles*, USSR Computational Mathematics and Mathematical Physics, 1 (1962), pp. 304–320.

[5] G. A. Sod, *A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws*, Journal of Computational Physics, 27 (1978), pp. 1–31.

[6] J. L. Steger and R. Warming, *Flux vector splitting of the inviscid gasdynamic equations with application to finite-difference methods*, Journal of Computational Physics, 40 (1981), pp. 263–293.

[7] E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[8] E. F. Toro, M. Spruce, and W. Speares, *Restoration of the contact surface in the HLL-Riemann solver*, Shock Waves, 4 (1994), pp. 25–34.

[9] G. Vijayasundaram, *Transonic flow simulations using an upstream centered scheme of Godunov in finite elements*, Journal of Computational Physics, 63 (1986), pp. 416–433.

[10] P. Woodward and P. Colella, *The numerical simulation of two-dimensional fluid flow with strong shocks*, Journal of Computational Physics, 54 (1984), pp. 115–173.

[11] H. Yee, N. Sandham, and M. Djomehri, *Low-Dissipative High-Order Shock-Capturing Methods Using Characteristic-Based Filters*, Journal of Computational Physics, 150 (1999), pp. 199–238.