# Finite volume methods: foundation and analysis

Timothy Barth[1] and Mario Ohlberger[2]

[1] *NASA Ames Research Center, Information Sciences Directorate, Moffett Field, California, 94035, USA*
[2] *Institute of Applied Mathematics, Freiburg University, Hermann-Herder-Str. 10, 79104 Freiburg, Germany and CSCAMM, University of Maryland, 4127 CSIC Building, College Park, Maryland, 20742-3289, USA*

ABSTRACT

Finite volume methods are a class of discretization schemes that have proven highly successful in approximating the solution of a wide variety of conservation law systems. They are extensively used in fluid mechanics, porous media flow, meteorology, electromagnetics, models of biological processes, semi-conductor device simulation and many other engineering areas governed by conservative systems that can be written in integral control volume form.

This article reviews elements of the foundation and analysis of modern finite volume methods. The primary advantages of these methods are numerical robustness through the obtention of discrete maximum (minimum) principles, applicability on very general unstructured meshes, and the intrinsic local conservation properties of the resulting schemes. Throughout this article, specific attention is given to scalar nonlinear hyperbolic conservation laws and the development of high order accurate schemes for discretizing them. A key tool in the design and analysis of finite volume schemes suitable for non-oscillatory discontinuity capturing is discrete maximum principle analysis. A number of building blocks used in the development of numerical schemes possessing local discrete maximum principles are reviewed in one and several space dimensions, e.g. monotone fluxes, E-fluxes, TVD discretization, non-oscillatory reconstruction, slope limiters, positive coefficient schemes, etc. When available, theoretical results concerning *a priori* and *a posteriori* error estimates are given. Further advanced topics are then considered such as high order time integration, discretization of diffusion terms and the extension to systems of nonlinear conservation laws.

KEY WORDS: finite volume methods, conservation laws, non-oscillatory approximation, discrete maximum principles, higher order schemes

Contents

**1 Introduction: Scalar nonlinear conservation laws** **2**
1.1 Characteristics of scalar conservation laws . . . . . 3
1.2 Weak solutions . . . . . 4
1.3 Entropy weak solutions and vanishing viscosity . . . . . 5
1.4 Measure-valued or entropy process solutions . . . . . 6

**2 Finite volume (FV) methods for nonlinear conservation laws** **7**
2.1 Godunov finite volume discretizations . . . . . 8
2.1.1 Monotone schemes. . . . . 10

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

2.1.2 E-flux schemes. . . . . 11
2.2 Stability, convergence and error estimates . . . . . 12
2.2.1 Discrete maximum principles and stability. . . . . 12
2.2.2 Convergence results. . . . . 13
2.2.3 Error estimates and convergence rates. . . . . 16
2.2.4 A posteriori error estimate. . . . . 17
2.2.5 A priori error estimate. . . . . 17
2.2.6 Convergence proofs via the streamline diffusion discontinuous Galerkin finite element method. . . . . 18

**3 Higher order accurate FV generalizations** **19**
3.1 Higher order accurate FV schemes in 1-D . . . . . 19
3.1.1 TVD schemes. . . . . 20
3.1.2 MUSCL schemes. . . . . 22
3.1.3 ENO/WENO schemes. . . . . 24
3.1.4 Reconstruction via primitive function. . . . . 25
3.1.5 ENO reconstruction. . . . . 25
3.1.6 WENO reconstruction. . . . . 26
3.2 Higher order accurate FV schemes in multi-dimensions. . . . . 27
3.2.1 Positive coefficient schemes on structured meshes. . . . . 28
3.2.2 FV schemes on unstructured meshes utilizing linear reconstruction. . . . 30
3.2.3 Linear reconstruction operators on simplicial control volumes. . . . . 35
3.2.4 Linear reconstruction operators on general control volumes shapes. . . . 36
3.2.5 General $p$-exact reconstruction operators on unstructured meshes. . . . 38
3.2.6 Positive coefficient schemes on unstructured meshes . . . . . 39

**4 Further Advanced Topics** **41**
4.1 Higher order time integration schemes . . . . . 41
4.1.1 Explicit SSP Runge-Kutta methods. . . . . 42
4.1.2 Optimal second and third order nonlinear SSP Runge-Kutta methods. . 43
4.2 Discretization of elliptic problems . . . . . 43
4.3 Conservation laws including diffusion terms . . . . . 45
4.3.1 Choices of the numerical diffusion flux $d_{jk}$. . . . . 45
4.3.2 Note on stability, convergence and error estimates. . . . . 46
4.4 Extension to systems of nonlinear conservation laws . . . . . 47
4.4.1 Numerical flux functions for systems of conservation laws. . . . . 48

**5 Concluding Remarks** **51**

## 1. Introduction: Scalar nonlinear conservation laws

Many problems from physics, chemistry, biology, mechanics, and gas dynamics lead to the study of nonlinear hyperbolic conservation laws. As a prototype conservation law, consider the

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Cauchy initial value problem

$$\begin{aligned} \partial_t u + \nabla \cdot f(u) &= 0 \qquad \text{in } \mathbb{R}^d \times \mathbb{R}^+ \ , & \text{(1a)} \\ u(x,0) &= u_0(x) \ \text{ in } \mathbb{R}^d. & \text{(1b)} \end{aligned}$$

Here $u(x,t) : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$ denotes the dependent solution variable, $f(u) \in C^1(\mathbb{R})$ denotes the flux function, and $u_0(x) : \mathbb{R}^d \to \mathbb{R}$ the initial data.

The function $u$ is a *classical solution* of the scalar initial value problem if $u \in C^1(\mathbb{R}^d \times \mathbb{R}^+)$ satisfies (1a-1b) pointwise. An essential feature of nonlinear conservation laws is that, in general, gradients of $u$ blow up in finite time, even when the initial data $u_0$ is arbitrarily smooth. Beyond some critical time $t_0$ classical solutions of (1a-1b) do not exist. This behavior will be demonstrated using the method of characteristics. By introducing the notion of weak solutions of (1a-1b) together with an entropy condition, it then becomes possible to define a class of solutions where existence and uniqueness is guaranteed for times greater than $t_0$. These are precisely the solutions that are numerically sought in the finite volume method.

### *1.1. Characteristics of scalar conservation laws*

Let $u$ be a classical solution of (1a) subject to initial data (1b). Further, define the vector

$$a(u) = f'(u) = (f_1'(u), \cdots, f_d'(u))^T.$$

A characteristic $\Gamma_y$ is a curve (x(t),t) such that

$$\begin{aligned} x'(t) &= a(u(x(t),t)) \quad \text{for } t > 0 \ , \\ x(0) &= y \ . \end{aligned}$$

Since $u$ is assumed to be a classical solution, it is readily verified that

$$\frac{d}{dt} u(x(t),t) = \partial_t u + x'(t) \nabla u = \partial_t u + a(u) \nabla u = \partial_t u + \nabla \cdot f(u) = 0 \ .$$

Therefore, $u$ is constant along a characteristic curve and $\Gamma_y$ is a straight line since

$$x'(t) = a(u(x(t),t)) = a(u(x(0),0)) = a(u(y,0)) = a(u_0(y)) = \text{const} \ .$$

In particular $x(t)$ is given by

$$x(t) = y + t a(u_0(y)) \ . \tag{2}$$

This important property may be used to construct classical solutions. If $x$ is fixed and $y$ determined as a solution of (2), then

$$u(x,t) = u_0(y) \ .$$

This procedure is the basis of the so-called method of characteristics. On the other hand, this construction shows that the intersection of any two straight characteristic lines leads to a contradiction in the definition of $u(x,t)$. Thus, classical solutions can only exist up to the first time $t_0$ at which any two characteristics intersect.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*1.2. Weak solutions*

Since, in general, classical solutions only exist for a finite time $t_0$, it is necessary to introduce the notion of weak solutions that are well-defined for times $t > t_0$.

**Definition 1.1 (Weak solution)** *Let $u_0 \in L^\infty(\mathbb{R}^d)$. Then, $u$ is a weak solution of (1a-1b) if $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}^+)$ and (1a-1b) hold in the distributional sense, i.e.*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^+} (u\partial_t\phi + f(u)\cdot\nabla\phi)\, dt\, dx + \int_{\mathbb{R}^d} u_0\phi(x,0)\, dx = 0 \quad \text{for all } \phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+)\ . \tag{3}$$

Note that classical solutions are weak solutions and weak solutions that lie in $C^1(\mathbb{R}^d \times \mathbb{R}^+)$ satisfy (1a-1b) in the classical sense.

It can be shown (see Kruzkov, 1970; Oleinik, 1963) that there always exists at least one weak solution to (1a-1b) if the flux function $f$ is at least Lipschitz continuous. Nevertheless, the class of weak solutions is too large to ensure uniqueness of solutions. An important class of solutions are piecewise classical solutions with discontinuities separating the smooth regions. The following lemma gives a necessary and sufficient condition imposed on these discontinuities such that the solution is a weak solution (see for example Godlewski and Raviart, 1991; Kröner, 1997). Later a simple example is given where infinitely many weak solutions exist.

**Lemma 1.2 (Rankine-Hugoniot jump condition)** *Assume that $\mathbb{R}^d \times R^+$ is separated by a smooth hypersurface $S$ into two parts $\Omega_l$ and $\Omega_r$. Furthermore, assume $u$ is a $C^1$-function on $\Omega_l$ and $\Omega_r$, respectively. Then, $u$ is a weak solution of (1a-1b) if and only if the following two conditions hold:*

*i) $u$ is a classical solution in $\Omega_l$ and $\Omega_r$, respectively.*

*ii) $u$ satisfies the Rankine-Hugoniot jump condition, i.e.*

$$[u]s = [f(u)]\cdot\nu \quad \text{on } S\ . \tag{4}$$

*Here, $(\nu, -s)^T$ denotes a unit normal vector for the hypersurface $S$ and $[w]$ denotes the jump in $w$ across the hypersurface $S$.*

In one space dimension, it may be assumed that $S$ is parameterized by $(\sigma(t), t)$ such that $s = \sigma'(t)$ and $\nu = 1$. The Rankine-Hugoniot jump condition then reduces to

$$s = \frac{[f(u)]}{[u]} \quad \text{on } S\ . \tag{5}$$

**Example 1.3 (Non-uniqueness of weak solutions)** *Consider the one-dimensional Burgers' equation, $f(u) = u^2/2$, with Riemann data: $u_0(x) = u_l$ for $x < 0$ and $u_0(x) = u_r$ for $x \geq 0$. Then, for any $a \geq max(u_l, -u_r)$ a function $u$ given by*

$$u(x,t) = \begin{cases} u_l, & x < s_1 t \\ -a, & s_1 t < x < 0 \\ a, & 0 < x < s_2 t \\ u_r, & s_2 t < x \end{cases} \tag{6}$$

*is a weak solution if $s_1 = (u_l - a)/2$ and $s_2 = (a + u_r)/2$. This is easily checked since $u$ is piecewise constant and satisfies the Rankine-Hugoniot jump condition. This elucidates a one-parameter family of weak solutions. In fact, there is also a classical solution whenever $u_l \leq u_r$.*

*Encyclopedia of Computational Mechanics*. Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*In this case, the characteristics do not intersect and the method of characteristics yields the classical solution*

$$u(x,t) = \begin{cases} u_l, & x < u_l t \\ x/t, & u_l t < x < u_r t \\ u_r, & u_r t < x \end{cases} . \tag{7}$$

This solution is the unique classical solution but not the unique weak solution. Consequently, additional conditions must be introduced in order to single out one solution within the class of weak solutions. These additional conditions give rise to the notion of a unique entropy weak solution.

### *1.3. Entropy weak solutions and vanishing viscosity*

In order to introduce the notion of entropy weak solutions, it is useful to first demonstrate that there is a class of additional conservation laws for any classical solution of (1a). Let $u$ be a classical solution and $\eta : \mathbb{R} \to \mathbb{R}$ a smooth function. Multiplying (1a) by $\eta'(u)$, one obtains

$$0 = \eta'(u)\partial_t u + \eta'(u)\nabla \cdot f(u) = \partial_t \eta(u) + \nabla \cdot F(u) \tag{8}$$

where $F$ is any primitive of $\eta' f'$. This reveals that for a classical solution $u$, the quantity $\eta(u)$, henceforth called an entropy function, is a conserved quantity.

**Definition 1.4 (Entropy - entropy flux pair)** *Let $\eta : \mathbb{R} \to \mathbb{R}$ be a smooth convex function and $F : \mathbb{R} \to \mathbb{R}$ a smooth function such that*

$$F' = \eta' f' \tag{9}$$

*in (8). Then $(\eta, F)$ is called an entropy - entropy flux pair or more simply an entropy pair for the equation (1a).*

**Note 1.5 (Kruzkov entropies)** *The family of smooth convex entropies $\eta$ may be equivalently replaced by the non-smooth family of so-called Kruzkov entropies, i.e. $\eta_\kappa(u) \equiv |u - \kappa|$ for all $\kappa \in \mathbb{R}$ (see Kröner, 1997).*

Unfortunately, the relation (8) can not be fulfilled for weak solutions in general, as it would lead to additional jump conditions which would contradict the Rankine-Hugoniot jump condition lemma. Rather, a weak solution may satisfy the relation (8) in the distributional sense with inequality. To see that this concept of entropy effectively selects a unique, physically relevant solution among all weak solutions, consider the viscosity perturbed equation

$$\partial_t u_\epsilon + \nabla \cdot f(u_\epsilon) = \epsilon \, \Delta u_\epsilon \tag{10}$$

with $\epsilon > 0$. For this parabolic problem, it may be assumed that a unique smooth solution $u_\epsilon$ exists. Multiplying by $\eta'$ and rearranging terms yields the additional equation

$$\partial_t \eta(u_\epsilon) + \nabla \cdot F(u_\epsilon) = \epsilon \Delta \eta(u_\epsilon) - \epsilon \eta''(u_\epsilon) |\nabla u|^2 \ .$$

Furthermore, since $\eta$ is assumed convex ($\eta'' \geq 0$), the following inequality is obtained

$$\partial_t \eta(u_\epsilon) + \nabla \cdot F(u_\epsilon) \leq \epsilon \Delta \eta(u_\epsilon) \ .$$

Taking the limit $\epsilon \to 0$ establishes (see Málek, Nečas, Rokyta and Růžička, 1996) that $u_\epsilon$ converges towards some $u$ a.e. in $\mathbb{R}^d \times \mathbb{R}^+$ where $u$ is a weak solution of (1a-1b) and satisfies the entropy condition

$$\partial_t \eta(u) + \nabla \cdot F(u) \leq 0 \tag{11}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

in the sense of distributions on $\mathbb{R}^d \times \mathbb{R}^+$.

By this procedure, a unique weak solution has been identified as the limit of the approximating sequence $u_\epsilon$. The obtained solution $u$ is called the vanishing viscosity weak solution of (1a-1b). Motivated by the entropy inequality (11) of the vanishing viscosity solution, it is now possible to introduce the notion of entropy weak solutions. This notion is weak enough for the existence and strong enough for the uniqueness of solutions to (1a-1b).

**Definition 1.6 (Entropy weak solution)** *Let $u$ be a weak solution of (1a-1b). Then, $u$ is called an entropy weak solution if $u$ satisfies for all entropy pairs $(\eta, F)$*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^+} (\eta(u)\partial_t\phi + F(u)\cdot\nabla\phi)\, dt\, dx + \int_{\mathbb{R}^d} \eta(u_0)\phi(x,0)\, dx \geq 0 \quad \text{for all } \phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+, \mathbb{R}^+)\,. \tag{12}$$

From the vanishing viscosity method, it is known that entropy weak solutions exist. The following $L^1$ contraction principle guarantees that entropy solutions are uniquely defined (see Kruzkov, 1970).

**Theorem 1.7 ($\mathbf{L^1}$ contraction principle)** *Let $u$ and $v$ be two entropy weak solutions of (1a-1b) with respect to initial data $u_0$ and $v_0$. Then, the following $L^1$ contraction principle holds*

$$||u(\cdot,t) - v(\cdot,t)||_{L^1(\mathbb{R}^d)} \leq ||u_0 - v_0||_{L^1(\mathbb{R}^d)} \tag{13}$$

*for almost every $t > 0$.*

This principle demonstrates a continuous dependence of the solution on the initial data and consequently the uniqueness of entropy weak solutions. Finally, note that an analog of the Rankine-Hugoniot condition exists (with inequality) in terms of the entropy pair for all entropy weak solutions

$$[\eta(u)]\, s \geq [F(u)]\cdot\nu \quad \text{on } S\ . \tag{14}$$

### *1.4. Measure-valued or entropy process solutions*

The numerical analysis of conservation laws requires an even weaker formulation of solutions to (1a-1b). For instance, the convergence analysis of finite volume schemes makes it necessary to introduce so called measure-valued or entropy process solutions (see DiPerna, 1985; Eymard, Gallouët and Herbin, 2000).

**Definition 1.8 (Entropy process solution)** *A function $\mu(x,t,\alpha) \in L^\infty(\mathbb{R}^d \times \mathbb{R}^+ \times (0,1))$ is called an entropy process solution of (1a-1b) if $u$ satisfies for all entropy pairs $(\eta, F)$*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^+}\int_0^1 \eta(\mu)\partial_t\left(\phi + F(\mu)\cdot\nabla\phi\right) d\alpha\, dt\, dx + \int_{\mathbb{R}^d} \eta(u_0)\phi(x,0)\, dx \geq 0 \ \text{ for all } \phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+, \mathbb{R}^+)\ .$$

The most important property of such entropy process solutions is the following uniqueness and regularity result (see Eymard, Gallouët and Herbin, 2000 [Theorem 6.3]).

**Theorem 1.9 (Uniqueness of entropy process solutions)** *Let $u_0 \in L^\infty(\mathbb{R}^d)$ and $f \in C^1(\mathbb{R})$. The entropy process solution $\mu$ of problem (1a-1b) is unique. Moreover, there exists a function $u \in L^\infty(\mathbb{R}^d \times \mathbb{R}^+)$ such that $u(x,t) = \mu(x,t,\alpha)$ a.e. for $(x,t,\alpha) \in \mathbb{R}^d \times \mathbb{R}^+ \times (0,1)$ and $u$ is the unique entropy weak solution of (1a-1b).*

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

## 2. Finite volume (FV) methods for nonlinear conservation laws

In the finite volume method, the computational domain, $\Omega \subset \mathbb{R}^d$, is first tessellated into a collection of non overlapping control volumes that completely cover the domain. Notationally, let $\mathcal{T}$ denote a tessellation of the domain $\Omega$ with control volumes $T \in \mathcal{T}$ such that $\cup_{T\in\mathcal{T}}\overline{T} = \overline{\Omega}$. Let $h_T$ denote a length scale associated with each control volume $T$, e.g. $h_T \equiv \text{diam(T)}$. For two distinct control volumes $T_i$ and $T_j$ in $\mathcal{T}$, the intersection is either an oriented edge (2-D) or face (3-D) $e_{ij}$ with oriented normal $\nu_{ij}$ or else a set of measure at most $d-2$. In each control volume, an *integral conservation law* statement is then imposed.

**Definition 2.1 (Integral conservation law)** *An integral conservation law asserts that the rate of change of the total amount of a substance with density $u$ in a fixed control volume $T$ is equal to the flux $f$ of the substance through the boundary $\partial T$*

$$\frac{\partial}{\partial t}\int_T u\,dx + \int_{\partial T} f(u)\cdot d\nu = 0\ . \tag{15}$$

This integral conservation law statement is readily obtained upon spatial integration of the divergence equation (1a) in the region $T$ and application of the divergence theorem. The choice of control volume tessellation is flexible in the finite volume method. For example, Fig.

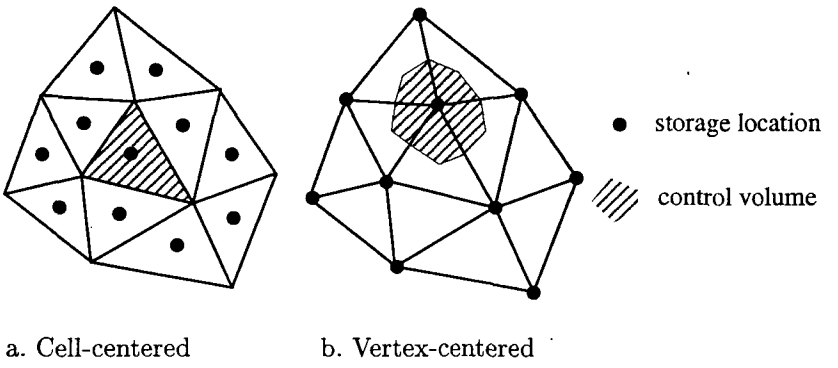



Figure 1. Control volume variants used in the finite volume method: (a) cell-centered and (b) vertex-centered control volume tessellation.

1 depicts a 2-D triangle complex and two typical control volume tessellations (among many others) used in the finite volume method. In the *cell-centered* finite volume method shown in Fig. 1a, the triangles themselves serve as control volumes with solution unknowns stored on a per triangle basis. In the *vertex-centered* finite volume method shown in Fig. 1b, control volumes are formed as a geometric dual to the triangle complex and solution unknowns stored on a per triangulation vertex basis.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

### 2.1. Godunov finite volume discretizations

Fundamental to finite volume methods is the introduction of the control volume cell average for each $T_j \in \mathcal{T}$

$$u_j \equiv \frac{1}{|T_j|} \int_{T_j} u \, dx \ . \tag{16}$$

For stationary meshes, the finite volume method can be interpreted as producing an evolution equation for cell averages

$$\frac{\partial}{\partial t} \int_{T_j} u \, dx = |T_j| \frac{\partial}{\partial t} u_j \ . \tag{17}$$

Godunov, 1959 pursued this interpretation in the discretization of the gas dynamic equations by assuming piecewise constant solution representations in each control volume with value equal to the cell average. However, the use of piecewise constant representations renders the numerical solution multivalued at control volume interfaces thereby making the calculation of a single solution flux at these interfaces ambiguous. The second aspect of Godunov's scheme and subsequent variants was the idea of supplanting the true flux at interfaces by a *numerical flux function*, $g(u, v) : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$, a Lipschitz continuous function of the two interface states $u$ and $v$. A single unique numerical flux was then calculated from an exact or approximate local solution of the Riemann problem in gas dynamics posed at these interfaces. Figure 2 depicts a representative 1-D solution profile in Godunov's method. For a given control volume $T_j = [x_{j-1/2}, x_{j+1/2}]$, Riemann problems are solved at each interface $x_{j\pm1/2}$. For example, at the interface $x_{j+1/2}$ the Riemann problem counterpart of (1a-1b)

$$\partial_\tau w_{j+1/2}(\xi, \tau) + \partial_\xi f(w_{j+1/2}(\xi, \tau)) = 0 \quad \text{in } \mathbb{R} \times \mathbb{R}^+$$

for $w_{j+1/2}(\xi, \tau) \in \mathbb{R}$ with initial data

$$w_{j+1/2}(\xi, 0) = \begin{cases} u_j & \text{if } \xi < 0 \\ u_{j+1} & \text{if } \xi > 0 \end{cases}$$

is solved either exactly or approximately. From this local solution, a single unique numerical flux at $x_{j+1/2}$ is computed from $g(u_j, u_{j+1}) = f(w_{j+1/2}(0, \mathbb{R}^+))$. This construction utilizes the fact that the solution of the Riemann problem at $\xi = 0$ is a constant for all time $\tau > 0$.

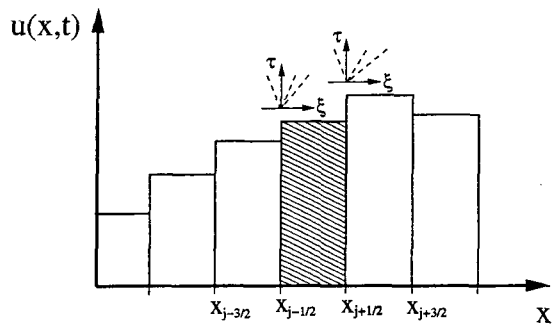



Figure 2. 1-D control volume, $T_j = [x_{j-1/2}, x_{j+1/2}]$, depicting Godunov's interface Riemann problems, $w_{j\pm1/2}(\xi, \tau)$, from piecewise constant interface states.

*Encyclopedia of Computational Mechanics*. Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

In higher space dimensions, the flux integral appearing in (15) is similarly approximated by

$$\int_{\partial T_j} f(u) \cdot d\nu \approx \sum_{\forall e_{jk} \in \partial T_j} g_{jk}(u_j, u_k) \tag{18}$$

where the numerical flux is assumed to satisfy the properties:

- (Conservation) This property insures that fluxes from adjacent control volumes sharing a mutual interface exactly cancel when summed. This is achieved if the numerical flux satisfies the identity
$$g_{jk}(u, v) = -g_{kj}(v, u) \ . \tag{19a}$$
- (Consistency) Consistency is obtained if the numerical flux with identical state arguments reduces to the true flux of that same state, i.e.
$$g_{jk}(u, u) = \int_{e_{jk}} f(u) \cdot d\nu \ . \tag{19b}$$

Combining (17) and (18) yields perhaps the simplest finite volume scheme in semi-discrete form. Let $V_h^0$ denote the space of piecewise constants, i.e.

$$V_h^0 = \{v \,|\, v|_T \in \chi(T), \ \forall T \in \mathcal{T}\} \tag{20}$$

with $\chi(T)$ a characteristic function in the control volume $T$.

**Definition 2.2 (Semi-discrete finite volume method)** *The semi-discrete finite volume approximation of (1a-1b) utilizing continuous in time solution representation, $t \in [0, \tau]$, and piecewise constant solution representation in space, $u_h(t) \in V_h^0$, such that*

$$u_j(t) = \frac{1}{|T_j|} \int_{T_j} u_h(x, t) \ dx$$

*with initial data*

$$u_j(0) = \frac{1}{|T_j|} \int_{T_j} u_0(x) \ dx$$

*and numerical flux function $g_{jk}(u_j, u_k)$ is given by the following system of ordinary differential equations*

$$\frac{d}{dt} u_j + \frac{1}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} g_{jk}(u_j, u_k) = 0 \ , \quad \forall T_j \in \mathcal{T} \ . \tag{21}$$

This system of ordinary differential equations can be marched forward using a variety of explicit and implicit time integration methods. In Sect. 4.1, time integration schemes that preserve properties of the spatial discretization are considered in more detail. Let $u_j^n$ denote a numerical approximation of the cell average solution in the control volume $T_j$ at time $t^n \equiv n\Delta t$. A particularly simple time integration method is the forward Euler scheme

$$\frac{d}{dt} u_j \approx \frac{u_j^{n+1} - u_j^n}{\Delta t}$$

thus producing the fully-discrete finite volume form.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

**Definition 2.3 (Fully-discrete finite volume method)** *The fully-discrete finite volume approximation of (1a-1b) for the time slab interval $[t^n, t^n + \Delta t]$ utilizing the piecewise constant solution representation in space, $u_h^n \in V_h^0$, such that*

$$u_j^n = \frac{1}{|T_j|} \int_{T_j} u_h^n(x)\ dx$$

*with initial data*

$$u_j^0 = \frac{1}{|T_j|} \int_{T_j} u_0(x)\ dx$$

*and numerical flux function $g_{jk}(u_j^n, u_k^n)$ is given by the following fully-discrete system*

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} g_{jk}(u_j^n, u_k^n) \ , \quad \forall T_j \in \mathcal{T} \ . \tag{22}$$

In subsequent sections, properties of the semi-discrete scheme (21) and fully-discrete scheme (22) will be examined in more detail.

*2.1.1. Monotone schemes.* Unfortunately, the numerical flux conditions (19a) and (19b) are insufficient to guarantee convergence to entropy satisfying weak solutions (12) and additional numerical flux restrictions are necessary. Two classes of numerical fluxes that guarantee such convergence for piecewise constant numerical solution data are *monotone fluxes* and *E-fluxes*. Specifically, Harten, Hyman and Lax, 1976 provide the following result concerning convergence of the fully-discrete one-dimensional scheme to weak solutions which was later generalized to (22) and irregular grids by Cockburn, Coquel and Lefloch, 1994.

**Theorem 2.4 (Monotone schemes and weak solutions)** *Consider a 1-D finite volume discretization of (1a-1b) with $2k+1$ stencil on a uniformly spaced mesh in both time and space with corresponding mesh spacing parameters $\Delta t$ and $\Delta x$*

$$\begin{aligned} u_j^{n+1} &= H_j(u_{j+k}, \ldots, u_j, \ldots, u_{j-k}) \\ &= u_j^n - \frac{\Delta t}{\Delta x}(g_{j+1/2} - g_{j-1/2}) \end{aligned} \tag{23}$$

*and consistent numerical flux of the form*

$$g_{j+1/2} = g(u_{j+k}, \ldots, u_{j+1}, u_j, \ldots, u_{j-k+1})$$

*that is monotone in the sense*

$$\frac{\partial H_j}{\partial u_{j+l}} \geq 0 \ , \quad \forall\ |l| \leq k \ . \tag{24}$$

*Then as $\Delta t$ and $\Delta x$ tend to zero with $\Delta t/\Delta x =$ constant, $u_j^n$ converges boundedly almost everywhere to $u(x,t)$, an entropy satisfying weak solution of (1a-1b).*

The monotonicity condition (24) motivates the introduction of Lipschitz continuous monotone fluxes satisfying

$$\frac{\partial g_{j+1/2}}{\partial u_l} \geq 0 \text{ if } \ l = j \tag{25a}$$

$$\frac{\partial g_{j+1/2}}{\partial u_l} \leq 0 \text{ if } \ l \neq j \tag{25b}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

together with a CFL (Courant-Friedrichs-Levy) like condition

$$1 - \frac{\Delta t}{\Delta x}\left(\frac{\partial g_{j+1/2}}{\partial u_j} - \frac{\partial g_{j-1/2}}{\partial u_j}\right) \geq 0$$

so that (24) is satisfied. Some examples of monotone fluxes for (1a) include

- (Godunov flux)

$$g^{\mathrm{G}}_{j+1/2} = \begin{cases} \min\limits_{u\in[u_j,u_{j+1}]} f(u) & \text{if } u_j < u_{j+1} \\ \max\limits_{u\in[u_j,u_{j+1}]} f(u) & \text{if } u_j > u_{j+1} \end{cases} \tag{26}$$

- (Lax-Friedrichs flux)

$$g^{\mathrm{LF}}_{j+1/2} = \frac{1}{2}\left(f(u_j) + f(u_{j+1})\right) - \frac{1}{2}\sup_{u\in[u_j,u_{j+1}]} |f'(u)|\,(u_{j+1} - u_j) \ . \tag{27}$$

*2.1.2. E-flux schemes.* A more general class of numerical fluxes was introduced and analyzed by Osher, 1984 that still guarantees convergence to weak entropy solutions when used in (22) or (23). These fluxes are called E-fluxes, $g_{j+1/2} = g^{\mathrm{E}}(u_{j+k}, \ldots, u_{j+1}, u_j, \ldots, u_{j-k+1})$, due to the relationship to Olienick's well-known E-condition which characterizes entropy satisfying discontinuities. E-fluxes satisfy the inequality

$$\frac{g^{\mathrm{E}}_{j+1/2} - f(u)}{u_{j+1} - u_j} \leq 0 \ , \quad \forall u \in [u_j, u_{j+1}] \ . \tag{28}$$

E-fluxes can be characterized by their relationship to Godunov's flux. Specifically, E-fluxes are precisely those fluxes such that

$$g^{\mathrm{E}}_{j+1/2} \leq g^{\mathrm{G}}_{j+1/2} \quad \text{if } u_{j+1} < u_j \tag{29a}$$

$$g^{\mathrm{E}}_{j+1/2} \geq g^{\mathrm{G}}_{j+1/2} \quad \text{if } u_{j+1} > u_j \ . \tag{29b}$$

Viewed another way, note that any numerical flux can be written in the form

$$g_{j+1/2} = \frac{1}{2}\left(f(u_j) + f(u_{j+1})\right) - \frac{1}{2}Q(u_{j+1} - u_j) \tag{30}$$

where $Q(\cdot)$ denotes a viscosity for the scheme. When written in this form, E-fluxes are those fluxes that contribute at least as much viscosity as Godunov's flux, i.e.

$$Q^{\mathrm{G}}_{j+1/2} \leq Q_{j+1/2} \ . \tag{31}$$

The most prominent E-flux is the Enquist-Osher flux

$$g^{\mathrm{EO}}_{j+1/2} = \frac{1}{2}\left(f(u_j) + f(u_{j+1})\right) - \frac{1}{2}\int_{u_j}^{u_{j+1}} |f'(s)|\,ds \ , \tag{32}$$

although other fluxes such as certain forms of Roe's flux with entropy fix fall into this category. From (29a-29b), the monotone fluxes of Godunov $g^{\mathrm{G}}_{j+1/2}$ and Lax-Friedrichs $g^{\mathrm{LF}}_{j+1/2}$ are also E-fluxes.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*2.2. Stability, convergence and error estimates*

Several stability results are provided here that originate from discrete maximum principle analysis and are straightforwardly stated in multi-dimensions and on general unstructured meshes. In presenting results concerning convergence and error estimates, a notable difference arises between one and several space dimensions. This is due to the lack of a BV bound on the approximate solution in multi-dimensions. Thus, before considering convergence and error estimates for finite volume methods, stability results are presented first together with some *a priori* bounds on the approximate solution.

*2.2.1. Discrete maximum principles and stability.* A compelling motivation for the use of monotone and E-fluxes in the finite volume schemes (21) and (22) is the obtention of discrete maximum principles in the resulting numerical solution of nonlinear conservation laws (1a). A standard analysis technique is to first construct local discrete maximum principles which can then be applied successively to obtain global maximum principles and stability results.

The first result concerns the boundedness of local extrema in time for semi-discrete finite volume schemes that can be written in nonnegative coefficient form.

**Theorem 2.5 (LED Property)** *The semi-discrete scheme for each* $T_j \in \mathcal{T}$

$$\frac{du_j}{dt} = \frac{1}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} C_{jk}(u_h)(u_k - u_j), \tag{33}$$

*is Local Extremum Diminishing (LED), i.e. local maxima are non-increasing and local minima are nondecreasing, if*

$$C_{jk}(u_h) \geq 0 \ , \ \forall e_{jk} \in \partial T_j \ . \tag{34}$$

Rewriting the semi-discrete finite volume scheme (21) in terms of monotone fluxes or E-fluxes

$$\begin{aligned} \frac{du_j}{\partial t} &= -\frac{1}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} \frac{g_{jk}(u_j, u_k) - f(u_j) \cdot \nu_{jk}}{u_k - u_j}(u_k - u_j) \\ &= -\frac{1}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} \frac{\partial g_{jk}}{\partial u_k}(u_j, \tilde{u}_{jk})(u_k - u_j) \end{aligned} \tag{35}$$

for appropriately chosen $\tilde{u}_{jk} \in [u_j, u_k]$ together with the monotone flux conditions (25a-25b) or the E-flux condition (28) reveals that monotone flux and E-flux finite volume schemes (21) are LED. In order to obtain local space-time maximum principle results for the fully-discrete discretization (22) requires the introduction of an additional CFL-like condition for non-negativity of coefficients in space-time.

**Theorem 2.6 (Local space-time discrete maximum principle)** *The fully-discrete scheme for the time slab increment* $[t^n, t^{n+1}]$ *and each* $T_j \in \mathcal{T}$

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} C_{jk}(u_h^n)(u_k^n - u_j^n) \tag{36}$$

*exhibits a local space-time discrete maximum principle*

$$\min_{\forall e_{jk} \in \partial T_j} (u_k^n, u_j^n) \leq u_j^{n+1} \leq \max_{e_{jk} \in \partial T_j} (u_k^n, u_j^n) \tag{37}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*if*

$$C_{jk}(u_h^n) \geq 0 \ , \quad \forall e_{jk} \in \partial T_j \tag{38}$$

*and satisfies the CFL-like condition*

$$1 - \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} C_{jk}(u_h^n) \geq 0 \ . \tag{39}$$

Again noting that the flux terms in the fully-discrete finite volume scheme (22) can be written in the form (35) reveals that the monotone flux conditions (25a-25b) or the E-flux condition (28) together with a local CFL-like condition obtained from (39) imply a local space-time discrete maximum principle. By successive application of Theorem 2.6, a global $L^\infty$-stability bound is obtained for the scalar initial value problem (1a-1b) in terms of initial data $u_0(x)$.

**Theorem 2.7 ($L^\infty$-stability)** *Assume a fully-discrete finite volume scheme (22) for the scalar initial value problem (1a-1b) utilizing monotone fluxes or E-fluxes that satisfy a local CFL-like condition as given in Theorem 2.6 for each time slab increment $[t^n, t^{n+1}]$. Under these conditions, the finite volume scheme is $L^\infty$-stable and the following estimate holds:*

$$\inf_{x \in \mathbf{R}^d} u_0(x) \leq u_j^n \leq \sup_{x \in \mathbf{R}^d} u_0(x), \quad \textit{for all } (T_j, t^n) \in \mathcal{T} \times [0, \tau]. \tag{40}$$

Consider now steady-state solutions, $u^{n+1} = u^n = u^*$, using monotone flux or E-flux schemes in the fully-discrete finite volume scheme (22). At steady state, non-negativity of the coefficients $C(u_h)$ in (36) implies a discrete maximum principle.

**Theorem 2.8 (Local discrete maximum principle in space)** *The fully-discrete scheme (36) exhibits a local discrete maximum principle at steady state, $u_h^*$, for each $T_j \in \mathcal{T}$*

$$\min_{\forall e_{jk} \in \partial T_j} u_k^* \leq u_j^* \leq \max_{\forall e_{jk} \in \partial T_j} u_k^* \tag{41}$$

*if*

$$C_{jk}(u_h^*) \geq 0 \ , \quad \forall e_{jk} \in \partial T_j \ .$$

Once again by virtue of (25a-25b) and (28), the conditions for a local discrete maximum principle at steady state are fulfilled by monotone flux and E-flux finite volume schemes (22). Global maximum principles for characteristic boundary valued problems are readily obtained by successive application of the local maximum principle result.

The local maximum principles given in (37) and (41) preclude the introduction of spurious extrema and $\mathcal{O}(1)$ Gibbs-like oscillations that occur near solution discontinuities computed using many numerical methods (even in the presence of grid refinement). For this reason, discrete maximum principles of this type are a highly sought after design principle in the development of numerical schemes for nonlinear conservation laws.

*2.2.2. Convergence results.* The $L^\infty$-stability bound (40) is an essential ingredient in the proof of convergence of the fully-discrete finite volume scheme (22). This bound permits the subtraction of a subsequence that converges against some limit in the $L^\infty$ weak-* sense. The primary task that then remains is to identify this limit with the unique solution of the problem. So although $L^\infty$-stability is enough to ascertain convergence of the scheme, stronger estimates are needed in order to derive convergence rates.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Let BV denote the space of functions with bounded variation, i.e.

$$\mathrm{BV} = \{g \in L^1(\mathbb{R}^d) \mid |g|_{BV} < \infty\} \text{ with } |g|_{BV} = \sup_{\substack{\varphi \in C_c^1(\mathbb{R}^d)^d \\ ||\varphi||_\infty \leq 1}} \int_{\mathbb{R}^d} g \, \nabla \cdot \varphi \, \mathrm{dx} \ .$$

From the theory of scalar conservation laws, it is known that, provided the initial data is in BV, the solution remains in BV for all times. Therefore, it is desirable to have an analog of this property for the approximate solution as well. Unfortunately, up to now, such an result is only rigorously proved in the one-dimensional case or in the case of tensor product cartesian meshes in multiple space dimensions. In the general multi-dimensional case, the approximate solution can only be shown to fulfill some weaker estimate which is thus called a *weak BV estimate* (see Vila, 1994; Cockburn, Coquel and Lefloch, 1994; Eymard, Gallouët, Ghilani and Herbin, 1998).

**Theorem 2.9 (Weak BV estimate)** *Let $\mathcal{T}$ be a regular triangulation, and let $J$ be a uniform partition of $[0,\tau]$, e.g. $\Delta t^n \equiv \Delta t$. Assume that there exists some $\alpha > 0$ such that $\alpha h^2 \leq |T_k|$, $\alpha\,|\partial T_k| \leq h$. For the time step $\Delta t^n$, assume the following CFL-like condition for a given $\xi \in (0,1)$*

$$\Delta t^n \leq \frac{(1-\xi)\alpha^2 h}{L_g}$$

*where $L_g$ is the Lipschitz constant of the numerical flux function. Furthermore, let $u_0 \in L^\infty(\mathbb{R}^d) \cap BV(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Then, the numerical solution of the fully-discrete discretization (22) fulfills the following estimate*

$$\sum_n \Delta t \sum_{jl} \chi_{jl}\ h\ |u_j^n - u_l^n| Q_{jl}(u_j^n, u_l^n) \quad \leq \quad K\sqrt{T|B_{R+h}(0)|}\sqrt{h} \ , \tag{42}$$

*where $K$ only depends on $\alpha$, $L_g$, $\xi$ and the initial function $u_0$. In this formula $Q_{jl}$ is defined as*

$$Q_{jl}(u,v) \equiv \frac{2g_{jl}(u,v) - g_{jl}(u,u) - g_{jl}(v,v)}{u-v}$$

*and $\chi_{jl}$ denotes the discrete cutoff function on $B_R(0) \subset \mathbb{R}^d$, i.e.*

$$\chi_{jl} = \begin{cases} 1, & \text{if } (T_j \cup T_l) \cap B_R(0) \neq \emptyset \\ 0, & \text{else} \end{cases} \ .$$

Note that in the case of a strong BV estimate, the right-hand side of (42) would be $\mathcal{O}(h)$ instead of $\mathcal{O}(\sqrt{h})$.

Another important property of monotone finite volume schemes is that they preserve the $L^1$–contraction property (see Theorem 1.7).

**Theorem 2.10 ($L^1$–contraction property and Lipschitz estimate in time)** *Let $u_h, v_h \in V_h^0$ be the approximate monotone finite volume solutions corresponding to initial data $u_0$, $v_0$ assuming that the CFL-like condition for stability has been fulfilled. Then the following discrete $L^1$–contraction property holds*

$$||u_h(\cdot, t+\tau) - v_h(\cdot, t+\tau)||_{L^1(\mathbb{R}^d)} \leq ||u_h(\cdot,t) - v_h(\cdot,t)||_{L^1(\mathbb{R}^d)} \ .$$

*Furthermore, a discrete Lipschitz estimate in time is obtained*

$$\sum_j |T_j||u_j^{n+1} - u_j^n| \leq L_g \Delta t^n \sum_j \sum_l |e_{jl}||u_j^0 - u_l^0| \ .$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

The principle ingredients of the convergence theory for scalar nonlinear conservation laws are compactness of the family of approximate solutions and the passage to the limit within the entropy inequality (12). In dealing with nonlinear equations, strong compactness is needed in order to pass to the limit in (12). In one space dimension, due to the BV estimate and the selection principle of Helly, strong compactness is ensured and the passage to the limit is summarized in the well known Lax-Wendroff theorem (see Lax and Wendroff, 1960).

**Theorem 2.11 ( Lax-Wendroff theorem)** *Let $(u_m)_{m\in\mathbb{N}}$ be a sequence of discrete solutions defined by the finite volume scheme in one space dimension with respect to initial data $u_0$. Assume that $(u_m)_{m\in\mathbb{N}}$ is uniformly bounded with respect to $m$ in $L^\infty$ and $u_m$ converges almost everywhere in $\mathbb{R}\times\mathbb{R}^+$ against some function $u$. Then $u$ is the uniquely defined entropy weak solution of (1a-1b).*

With the lack of a BV estimate for the approximate solution in multiple space dimensions, one cannot expect a passage to the limit of the nonlinear terms in the entropy inequality in the classical sense, i.e. the limit of $u_m$ will not in general be a weak solution. Nevertheless, the weak compactness obtained by the $L^\infty$-estimate is enough to obtain a measure-valued or entropy process solution in the limit.

The key theorem for this convergence result is the following compactness theorem of Tartar (see Tartar, 1983; Eymard, Galluoët and Herbin, 2000).

**Theorem 2.12 (Tartar's Theorem)** *Let $(u_m)_{m\in\mathbb{N}}$ be a family of bounded functions in $L^\infty(\mathbb{R}^n)$. Then, there exists a subsequence $(u_m)_{m\in\mathbb{N}}$, and a function $u\in L^\infty(\mathbb{R}^n\times(0,1))$ such that for all functions $g\in C(\mathbb{R})$ the weak-$\star$ limit of $g(u_m)$ exists and*

$$\lim_{m\to\infty}\int_{\mathbb{R}^n} g(u_m(x))\phi(x)dx = \int_0^1\int_{\mathbb{R}^n} g(u(x,\alpha))\phi(x)dx\,d\alpha,\ \text{for all } \phi\in L^1(\mathbb{R}^n)\ . \tag{43}$$

In order to prove the convergence of a finite volume method, it now remains to show that the residual of the entropy inequality (12) for the approximate solution $u_h$ tends to zero if $h$ and $\Delta t$ tend to zero. Before presenting this estimate for the finite volume approximation, a general convergence theorem is given which can be viewed as a generalization of the classical Lax-Wendroff result (see Eymard, Galluoët and Herbin, 2000).

**Theorem 2.13 (Sufficient condition for convergence)** *Let $u_0\in L^\infty(\mathbb{R}^d)$ and $f\in C^1(\mathbb{R})$. Further, let $(u_m)_{m\in\mathbb{N}}$ be any family of uniformly bounded functions in $L^\infty(\mathbb{R}^d\times\mathbb{R}^+)$ that satisfies the following estimate for the residual of the entropy inequality using the class of Kruzkov entropies $\eta_\kappa$ (see Note 1.5).*

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^+}(\eta_\kappa(u_m)\partial_t\phi + F_{\eta_\kappa}(u_m)\cdot\nabla\phi)\,dt\,dx + \int_{\mathbb{R}^d}\eta_\kappa(u_0)\phi(x,0)\,dx \geq -R(\kappa,u_m,\phi) \tag{44}$$

*for all $\kappa\in\mathbb{R}$ and $\phi\in C_0^1(\mathbb{R}^d\times\mathbb{R}^+,\mathbb{R}^+)$ where the residual $R(\kappa,u_m,\phi)$ tends to zero for $m\to\infty$ uniformly in $\kappa$. Then, $u_m$ converges strongly to the unique entropy weak solution of (1a-1b) in $L^p_{loc}(\mathbb{R}^d\times\mathbb{R}^+)$ for all $p\in[1,\infty)$ .*

**Theorem 2.14 (Estimate on the residual of the entropy inequality)** *Let $(u_m)_{m\in\mathbb{N}}$ be a sequence of monotone finite volume approximations satisfying a local CFL-like condition as given in (39) such that $h,\Delta t$ tend to zero for $m\to\infty$. Then, there exist measures*

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*$\mu_m \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^+)$ and $\nu_m \in \mathcal{M}(\mathbb{R}^d)$ such that the residual $R(\kappa, u_m, \phi)$ of the entropy inequality is estimated by*

$$R(\kappa, u_m, \phi) \leq \int_{\mathbb{R}^d} \int_{\mathbb{R}^+} (|\partial_t \phi(x,t)| + |\nabla \phi(x,t)|) \, d\mu_m(x,t) + \int_{\mathbb{R}^d} \phi(x,0) \, d\nu_m(x)$$

*for all $\kappa \in \mathbb{R}$ and $\phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+, \mathbb{R}^+)$. The measures $\mu_m$ and $\nu_m$ satisfy the following properties:*

1. *For all compact subsets $\Omega \subset\subset \mathbb{R}^d \times \mathbb{R}^+$, $\lim_{m \to \infty} \mu_m(\Omega) = 0$.*
2. *For all $g \in C_0(\mathbb{R}^d)$ the measure $\nu_m$ is given by $\langle \nu_m, g \rangle = \int_{\mathbb{R}^d} g(x) |u_0(x) - u_m(x,0)| \, dx$.*

These theorems are sufficient for establishing convergence of monotone finite volume schemes.

**Corollary 2.15 (Convergence theorem)** *Let $(u_m)_{m \in \mathbb{N}}$ be a sequence of monotone finite volume approximations satisfying the assumptions of Theorem 2.14. Then, $u_m$ converges strongly to the unique entropy weak solution of (1a-1b) in $L^p_{loc}(\mathbb{R}^d \times \mathbb{R}^+)$ for all $p \in [1, \infty)$.*

Convergence of higher order finite volume schemes can also be proven within the given framework as long as they are $L^\infty$-stable and allow for an estimate on the entropy residual in the sense of Theorem 2.14, for details see Kröner, Noelle and Rokyta, 1995; Chainais-Hillairet, 2000.

*2.2.3. Error estimates and convergence rates.* There are two primary approaches taken to obtain error estimates for approximations of scalar nonlinear conservation laws. One approach is based on the ideas of Oleinik and is applicable only in one space dimension (see Oleinik, 1963; Tadmor, 1991). The second approach which is widely used in the numerical analysis of conservation laws is based on the doubling of variables technique of Kruzkov (see Kruzkov, 1970; Kuznetsov, 1976). In essence, this technique enables one to estimate the error between the exact and approximate solution of a conservation law in terms of the entropy residual $R(\kappa, u_m, \Phi)$ introduced in (44). Thus, an *a posteriori* error estimate is obtained. Using *a priori* estimates of the approximate solution (see Section 2.2.1, and Theorems 2.9, 2.10), a convergence rate or an *a priori* error estimate is then obtained. The next theorem gives a fundamental error estimate for conservation laws independent of the particular finite volume scheme (see Eymard, Galluoët and Herbin, 2000; Chainais-Hillairet, 1999; Kröner and Ohlberger, 2000).

**Theorem 2.16 (Fundamental error estimate)** *Let $u_0 \in BV(\mathbb{R}^d)$ and let $u$ be an entropy weak solution of (1a-1b). Furthermore, let $v \in L^\infty(\mathbb{R}^d \times \mathbb{R}^+)$ be a solution of the following entropy inequalities with residual term $R$:*

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^+} \eta_\kappa(v) \partial_t \phi + F_{\eta_\kappa}(v) \cdot \nabla \phi + \int_{\mathbb{R}^d} \eta_\kappa(u_0) \phi(\cdot, 0) \geq -R(\kappa, v, \phi) \tag{45}$$

*for all $\kappa \in \mathbb{R}$ and $\phi \in C_0^1(\mathbb{R}^d \times \mathbb{R}^+, \mathbb{R}^+)$. Suppose that there exist measures $\mu_v \in \mathcal{M}(\mathbb{R}^d \times \mathbb{R}^+)$ and $\nu_v \in \mathcal{M}(\mathbb{R}^d)$ such that $R(\kappa, v, \phi)$ can be estimated independently of $\kappa$ by*

$$R(\kappa, v, \phi) \leq \langle |\partial_t \phi| + |\nabla \phi|, \mu_v \rangle + \langle |\phi(\cdot, 0)|, \nu_v \rangle. \tag{46}$$

*Let $K \subset\subset \mathbb{R}^d \times \mathbb{R}^+$, $\omega \equiv Lip(f)$ , and choose $T, R$ and $x_0$ such that $T \in ]0, \frac{R}{\omega}[$ and $K$ lies within its cone of dependence $D_0$, i.e. $K \subset D_0$ where $D_\delta$ is given as*

$$D_\delta := \bigcup_{0 \leq t \leq T} B_{R - \omega t + \delta}(x_0) \times \{t\}. \tag{47}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*Then, there exists a $\delta \geq 0$ and positive constants $C_1, C_2$ such that $u$, $v$ satisfy the following error estimate*

$$||u - v||_{L^1(K)} \leq T(\nu_v(B_{R+\delta}(x_0)) + C_1\mu_v(D_\delta) + C_2\sqrt{\mu_v(D_\delta)}). \tag{48}$$

This estimate can be used either as an *a posteriori* control of the error, as the right-hand side of the estimate (48) only depends on $v$, or it can be used as an *a priori* error bound if one is able to estimate further the measures $\mu_v$ and $\nu_v$ using some *a priori* bounds on $v$. Finally, note that comparable estimates to (48) are obtainable in an $L^\infty(0,T;L^1(\mathbb{R}^d))$-norm (see Cockburn and Gau, 1995; Bouchut and Perthame, 1998).

*2.2.4. A posteriori error estimate.*

**Theorem 2.17 (A posteriori error estimate)** *Assume the conditions and notations as in Theorem 2.16. Let $v = u_h$ be a numerical approximation to (1a-1b) obtained from a monotone finite volume scheme that satisfies a local CFL-like condition as given in (39). Then the following error estimate holds*

$$\int_K |u - u_h| \quad \leq \quad T(||u_0 - u_h(\cdot,0)||_{L^1(B_{R+h}(x_0))} + C_1\eta + C_2\sqrt{\eta}), \tag{49}$$

*where*

$$\eta \equiv \sum_{n\in I_0} \sum_{j\in M(t^n)} |u_j^{n+1} - u_j^n|\Delta t^n h_j^d + 2\sum_{n\in I_0} \sum_{(j,l)\in E(t^n)} \Delta t^n(\Delta t^n + h_{jl})Q_{jl}(u_j^n, u_l^n)|u_j^n - u_l^n| \ ,$$

$$Q_{jl}(u,v) \equiv \frac{2g_{jl}(u,v) - g_{jl}(u,u) - g_{jl}(v,v)}{u - v}$$

*with the index sets $I_0, M(t), E(t)$ given by*

$$\begin{aligned} I_0 &\equiv \{n \mid 0 \leq t^n \leq \min\{\frac{R+\delta}{\omega}, T\}\ , \\ M(t) &\equiv \{j \mid \text{there exists} \quad x \in T_j \quad \text{such that } (x,t) \in D_{R+\delta}\}\ , \\ E(t) &\equiv \{(j,l) \mid \text{there exists} \quad x \in T_j \cup T_l \quad \text{such that} \quad (x,t) \in D_{R+\delta}\}\ . \end{aligned}$$

*Furthermore, the constants $C_1, C_2$ only depend on $T$, $\omega$, $||u_0||_{BV}$ and $||u_0||_{L^\infty}$ (for details see Kröner and Ohlberger, 2000).*

Note that this *a posteriori error* estimate is local, since the error on a compact set $K$ is estimated by discrete quantities that are supported in the cone of dependence $D_{R+\delta}$.

*2.2.5. A priori error estimate.* Using the weak BV estimate (Theorem 2.9) and the Lipschitz estimate in time (Theorem 2.10), the right-hand side of the *a posteriori* error estimate (49) can be further estimated. This yields an *a priori* error estimate as stated in the following theorem (for details see Eymard, Gallouët and Herbin, 2000; Chainais-Hillairet, 1999).

**Theorem 2.18 (A priori error estimate)** *Assume the conditions and notations as in Theorem 2.16 and let $v = u_h$ be the approximation to (1a), (1b) given by a monotone finite volume scheme that satisfies a local CFL-like condition as given in (39). Then there exists a constant $C \geq 0$ such that*

$$\int_K |u - u_h| \quad \leq \quad Ch^{1/4}.$$

*Moreover, in the one-dimensional case, the optimal convergence rate of $h^{1/2}$ is obtained.*

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*2.2.6. Convergence proofs via the streamline diffusion discontinuous Galerkin finite element method.* It is straightforward to show that the fully-discrete finite volume scheme (22) can be viewed as a specific case of the more general streamline diffusion discontinuous Galerkin (SD-DG) finite element method which utilizes the mesh dependent broken space $V_h^p$ defined as

$$V_h^p = \{v \,|\, v|_T \in \mathcal{P}_p(T),\ \forall T \in \mathcal{T}\} \tag{50}$$

with $\mathcal{P}_p(T)$ the space of polynomials of degree $\leq p$ in the control volume $T$. By generalizing the notion of gradient and flux to include the time coordinate as well, the discontinuous Galerkin finite element method for a space-time tessellation $\mathcal{T}^n$ spanning the time slab increment $[t^n, t^{n+1}]$ is given compactly by the following variational statement.

*SD-DG(p) finite element method.* Find $u_h \in V_h^p$ such that $\forall\, v_h \in V_h^p$ and $n = 0, 1, \ldots$

$$\sum_{T \in \mathcal{T}^n} \left( \int_T (v_h + \delta(u_h) f'(u_h) \cdot \nabla v_h)\ \nabla \cdot f(u_h)\, dx + \int_T \widehat{\epsilon}(u_h)\, \nabla u_h \cdot \nabla v_h\, dx \right.$$
$$\left. + \int_{\partial T} v_{h-} \left( g(u_{h-}, u_{h+}) - f(u_{h-}) \cdot \nu \right) ds \right) = 0 \tag{51}$$

where in the integration over $\partial T$ it is understood for the portion $x \in \partial T \cap \partial T'$ that $u_{h-}, v_{h-}$ denotes the trace restriction of $u_h(T)$ and $v_h(T)$ onto $\partial T$ and $u_{h+}$ denotes the trace restriction of $u_h(T')$ onto $\partial T'$. Given this space-time formulation, convergence results for a scalar nonlinear conservation law in multi-dimensions and unstructured meshes are given in Jaffre, Johnson and Szepessy, 1995 for specific choices of the stabilization functions $\delta(u_h) : \mathbb{R} \mapsto \mathbb{R}^+$ and $\epsilon(u_h) : \mathbb{R} \mapsto \mathbb{R}^+$ together with a monotone numerical flux function $g(u_{h-}, u_{h+})$. Using their stabilization functions together with a monotone flux function, the following convergence result is obtained:

**Theorem 2.19 (SD-DG($p$) convergence)** *Suppose that components of $f'(u) \in C^d(\mathbb{R})$ are bounded and that $u_0 \in L_2(\mathbb{R}^d)$ has compact support. Then the solution $u_h$ of the SD-DG(p) method converges strongly in $L_p^{\text{loc}}(\mathbb{R}^d \times R^+)$, $1 \leq p \leq 2$, to the unique solution $u$ of the scalar nonlinear conservation law system (1a-1b) as $H \equiv \max(\|h\|_{L_\infty(\mathbb{R}^d)}, \Delta t)$ tends to zero.*

The proof of convergence to a unique entropy solution on general meshes for $p \geq 0$ is based on an extension by Szepessy, 1989 of a uniqueness result by DiPerna, 1985 by providing convergence for a sequence of approximations satisfying:

- a uniform $L_\infty$ bound in time and $L_2$ in space,
- entropy consistency and inequality for all Kruzkov entropies,
- consistency with initial data.

By choosing SD-DG(0), the dependence on the as yet unspecified stabilization functions $\delta(u_h)$ and $\epsilon(u_h)$ vanishes identically and the fully-discrete scheme (22) with monotone flux function is *exactly* reproduced, thus yielding a convergence proof for general scalar conservation laws for the finite volume method as well.

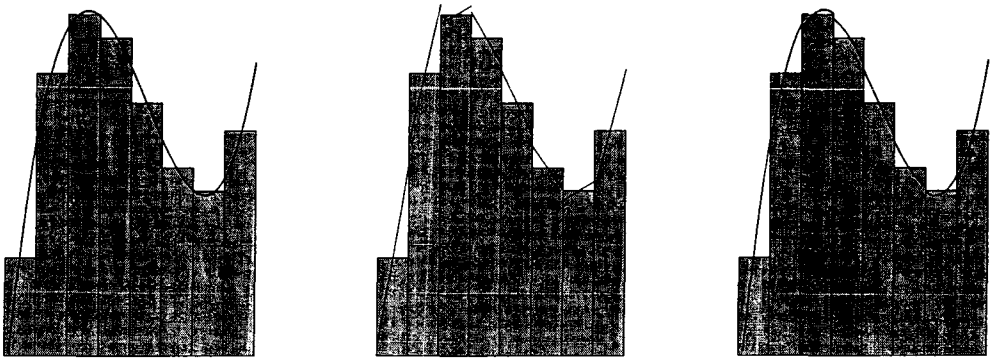
*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

## 3. Higher order accurate FV generalizations

Although an $\mathcal{O}(h^{1/2})$ $L_1$-norm error bound for the monotone and E-flux schemes of Sect. 2 is known to be sharp (Peterson, 1991), an $\mathcal{O}(h)$ solution error is routinely observed in numerical experiments with convex flux functions. Even so, first order accurate schemes are generally considered too inaccurate for most quantitative calculations unless the mesh spacing is made excessively small thus rendering the schemes inefficient. Godunov, 1959 has shown that all *linear* schemes that preserve solution monotonicity are at most first order accurate. The low order accuracy of these monotonicity preserving linear schemes has motivated the development of higher order accurate schemes with the important distinction that these new schemes utilize essential *nonlinearity* so that monotone resolution of discontinuities and high order accuracy away from discontinuities are simultaneously attained.

### *3.1. Higher order accurate FV schemes in 1-D*

A significant step forward in the generalization of Godunov's finite volume method to higher order accuracy is due to van Leer, 1979. In the context of Lagrangian hydrodynamics with Eulerian remapping, van Leer generalized Godunov's method by employing linear solution *reconstruction* in each cell (see Fig. 3b). Let $N$ denote the number of control volume cells in



a. Cell averaging of quartic data b. Linear reconstruction c. Quadratic reconstruction

Figure 3. Piecewise polynomial approximation used in the finite volume method: (a) cell averaging of analytic data, (b) piecewise linear reconstruction from cell averages and (c) piecewise quadratic reconstruction from cell averages.

space so that the $j$-th cell extends over the interval $T_j = [x_{j-1/2}, x_{j+1/2}]$ with length $\Delta x_j$ such that $\cup_{1\le j\le N} T_j = [0,1]$ with $T_i \cap T_j = \emptyset, i \neq j$. In a purely Eulerian setting, the higher order accurate schemes of van Leer are of the form

$$\frac{du_j}{dt} + \frac{1}{\Delta x_j}(g(u^-_{j+1/2}, u^+_{j+1/2}) - g(u^-_{j-1/2}, u^+_{j-1/2})) = 0$$

where $g(u,v)$ is a numerical flux function utilizing states $u^-_{j\pm1/2}$ and $u^+_{j\pm1/2}$ obtained from evaluation of the linear solution reconstructions from the left and right cells surrounding the

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

interfaces $x_{j\pm1/2}$, By altering the slope of the linear reconstruction in cells, non-oscillatory resolution of discontinuities can be obtained. Note that although obtaining the exact solution of the scalar nonlinear conservation law with linear initial data is a formidable task, the solution at each cell interface location for small enough time is the same as the solution of the Riemann problem with piecewise constant data equal to the linear solution approximation evaluated at the same interface location. Consequently, the numerical flux functions used in Sect. 2 can be once again used in the generalized schemes of van Leer. This single observation greatly simplifies the construction of higher order accurate generalizations of Godunov's method. The ideas of van Leer have been extended to quadratic approximations in each cell (see Fig. 3c) by Colella and Woodward, 1984. Although these generalizations of Godunov's method and further generalizations given later can be interpreted in 1-D as finite difference discretizations, concepts originally developed in 1-D such as solution monotonicity, positive coefficient discretization and discrete maximum principle analysis are often used in the design of finite volume methods in multiple space dimensions and on unstructured meshes where finite difference discretization is problematic.

*3.1.1. TVD schemes.* In considering the scalar nonlinear conservation law (1a-1b), Lax, 1973 made the following basic observation:

> *"the total increasing and decreasing variations of a differentiable solution between any pair of characteristics are conserved".*

Furthermore, in the presence of shock wave discontinuities, information is lost and the total variation *decreases*. For the 1-D nonlinear conservation law with compactly supported (or periodic) solution data $u(x,t)$, integrating along the constant time spatial coordinate at times $t_1$ and $t_2$ yields

$$\int_{-\infty}^{\infty} |du(x,t_2)| \leq \int_{-\infty}^{\infty} |du(x,t_1)|, \qquad t_2 \geq t_1 \ . \tag{52}$$

This motivated Harten, 1983 to consider the discrete total variation

$$\mathrm{TV}(u_h) \equiv \sum_j |\Delta_{j+1/2} u_h| \ , \quad \Delta_{j+1/2} u_h \equiv u_{j+1} - u_j$$

and the discrete total variation non-increasing (TVNI) bound counterpart to (52)

$$\mathrm{TV}(u_h^{n+1}) \leq \mathrm{TV}(u_h^n) \tag{53}$$

in the design of numerical discretizations for nonlinear conservation laws. A number of simple results relating TVNI schemes and monotone schemes follow from simple analysis.

**Theorem 3.1 (TVNI and monotone scheme properties, Harten, 1983)** *(i) Monotone schemes are TVNI. (ii) TVNI schemes are monotonicity preserving, i.e. the number of solution extrema is preserved in time.*

Property (i) follows from the $L_1$ contraction property of monotone schemes. Property (ii) is readily shown using a proof by contradiction by assuming a TVNI scheme with monotone initial data that produces new solution data at a later time with interior solution extrema present. Using the notion of discrete total variation, Harten, 1983 then constructed sufficient algebraic conditions for achieving the TVNI inequality (53).

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

**Theorem 3.2 (Harten's explicit TVD criteria)** *The fully discrete explicit 1-D scheme*

$$u_j^{n+1} = u_j^n + \Delta t \left( C_{j+1/2}(u_h^n)\, \Delta_{j+1/2} u_h^n + D_{j+1/2}(u_h^n)\, \Delta_{j-1/2} u_h^n \right), \quad j = 1, \ldots, N \tag{54}$$

*is total variation non-increasing if for each j*

$$C_{j+1/2} \geq 0 , \tag{55a}$$

$$D_{j+1/2} \leq 0 , \tag{55b}$$

$$1 - \Delta t \left( C_{j-1/2} - D_{j+1/2} \right) \geq 0 . \tag{55c}$$

Note that although the inequality constraints (55a-55c) in Theorem 3.2 insure that the total variation is non-increasing, these conditions are often referred to as total variation diminishing (TVD) conditions. Also note that inequality (55c) implies a CFL-like time step restriction that may be more restrictive than the time step required for stability of the numerical method. The TVD conditions are easily generalized to wider support stencils written in incremental form, see for example Jameson and Lax, 1986 and their corrected result in Jameson and Lax, 1987.

While this simple Euler explicit integration scheme may seem too crude for applications requiring true high order space-time accuracy, special attention and analysis is given to this fully-discrete form because it serves as a fundamental building block for an important class of high order accurate Runge-Kutta time integration techniques discussed in Sect. 4.1 that, by construction, inherit TVD (and later maximum principle) properties of the fully-discrete scheme.

**Theorem 3.3 (Generalized explicit TVD criteria)** *The fully discrete explicit 1-D scheme*

$$u_j^{n+1} = u_j^n + \Delta t \sum_{l=-k}^{k-1} C_{j+1/2}^{(l)}(u_h^n) \Delta_{j+l+1/2} u_h^n, \quad j = 1, \ldots, N \tag{56}$$

*with stencil width parameter k is total variation non-increasing if for each j*

$$C_{j+1/2}^{(k-1)} \geq 0 , \tag{57a}$$

$$C_{j+1/2}^{(-k)} \leq 0 , \tag{57b}$$

$$C_{j+1/2}^{(l-1)} - C_{j-1/2}^{(l)} \geq 0 , \quad -k+1 \leq l \leq k-1,\ l \neq 0 , \tag{57c}$$

$$1 - \Delta t \left( C_{j-1/2}^{(0)} - C_{j+1/2}^{(-1)} \right) \geq 0 . \tag{57d}$$

The extension to implicit methods follows immediately upon rewriting the implicit scheme in terms of the solution spatial increments $\Delta_{j+l+1/2} u_h$ and imposing sufficient algebraic conditions such that the implicit matrix acting on spatial increments has a nonnegative inverse.

**Theorem 3.4 (Generalized implicit TVD criteria)** *The fully discrete implicit 1-D scheme*

$$u_j^{n+1} - \Delta t \sum_{l=-k}^{k-1} C_{j+1/2}^{(l)}(u_h^{n+1}) \Delta_{j+l+1/2} u_h^{n+1} = u_j^n, \quad j = 1, \ldots, N \tag{58}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*with stencil width parameter $k$ is total variation non-increasing if for each $j$*

$$C_{j+1/2}^{(k-1)} \geq 0 , \tag{59a}$$

$$C_{j+1/2}^{(-k)} \leq 0 , \tag{59b}$$

$$C_{j+1/2}^{(l-1)} - C_{j-1/2}^{(l)} \geq 0 , \quad -k+1 \leq l \leq k-1, \; l \neq 0 . \tag{59c}$$

Theorems 3.3 and 3.4 provide sufficient conditions for non-increasing total variation of explicit (56) or implicit (58) numerical schemes written in incremental form. These incremental forms do not imply *discrete conservation* unless additional constraints are imposed on the discretizations. A sufficient condition for discrete conservation of the discretizations (56) and (58) is that these discretizations can be written in a finite volume flux balance form

$$g_{j+1/2} - g_{j-1/2} = \sum_{l=-k}^{k-1} C_{j+1/2}^{(l)}(u_h)\Delta_{j+l+1/2}u_h$$

where $g_{j\pm1/2}$ are the usual numerical flux functions. Section 3.1.2 provides an example of how the discrete TVD conditions and discrete conservation can be simultaneously achieved. A more comprehensive overview of finite volume numerical methods based on TVD constructions can be found the books by Godlewski and Raviart, 1991 and LeVeque, 2002.

*3.1.2. MUSCL schemes.* A general family of TVD discretizations with 5-point stencil is the Monotone Upstream-centered Scheme for Conservation Laws (MUSCL) discretization of van Leer, 1979; van Leer, 1985. MUSCL schemes utilize a $\kappa$-parameter family of interpolation formulas with *limiter function* $\Psi(R) : \mathbb{R} \mapsto \mathbb{R}$

$$\begin{aligned} u_{j+1/2}^{-} &= u_j + \frac{1+\kappa}{4}\Psi(1/R_i)\Delta_{j+1/2}u_h + \frac{1-\kappa}{4}\Psi(R_j)\Delta_{j-1/2}u_h \\ u_{j+1/2}^{+} &= u_j - \frac{1+\kappa}{4}\Psi(R_{j+1})\Delta_{j+1/2}u_h - \frac{1-\kappa}{4}\Psi(1/R_{j+1})\Delta_{j+3/2}u_h \end{aligned} \tag{60}$$

where $R_j$ is a ratio of successive solution increments

$$R_j \equiv \frac{\Delta_{j+1/2}u_h}{\Delta_{j-1/2}u_h} . \tag{61}$$

The technique of incorporating limiter functions to obtain non-oscillatory resolution of discontinuities and steep gradients dates back to Boris and Book, 1973. For convenience, the interpolation formulas (60) have been written for a uniformly spaced mesh although the extension to irregular mesh spacing is straightforward. The unlimited form of this interpolation is obtained by setting $\Psi(R) = 1$. In this unlimited case, the truncation error for the conservation law divergence in (1a) is given by

$$\text{Truncation Error} = -\frac{\left(\kappa - \frac{1}{3}\right)}{4}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u) .$$

This equation reveals that for $\kappa = 1/3$, the 1-D MUSCL formula yields an overall spatial discretization with $\mathcal{O}(\Delta x^3)$ truncation error. Using the MUSCL interpolation formulas given in (60), sufficient conditions for the discrete TVD property are easily obtained.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

**Theorem 3.5 (MUSCL TVD scheme)** *The fully discrete 1-D scheme*

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{\Delta x_j}\left(g_{j+1/2}^n - g_{j-1/2}^n\right) \ , \quad j = 1, \ldots, N$$

*with monotone Lipschitz continuous numerical flux function*

$$g_{j+1/2} = g(u_{j+1/2}^-, u_{j+1/2}^+)$$

*utilizing the $\kappa$-parameter family of MUSCL interpolation formulas (60) and (61) is total variation non-increasing if there exists a $\Psi(R)$ such that $\forall R \in \mathbb{R}$*

$$0 \le \Psi(R) \le \frac{3-\kappa}{1-\kappa} - (1+\alpha)\frac{1+\kappa}{1-\kappa} \tag{62a}$$

*and*

$$0 \le \frac{\Psi(R)}{R} \le 2 + \alpha \tag{62b}$$

*with $\alpha \in [-2, 2\,(1-\kappa)/(1+\kappa)]$ under the time step restriction*

$$1 - \frac{\Delta t}{\Delta x_j}\,\frac{2-(2+\alpha)\kappa}{1-\kappa}\left|\frac{\partial g}{\partial u}\right|_j^{\max} \ge 0$$

*where*

$$\left|\frac{\partial g}{\partial u}\right|_j^{\max} \equiv \sup_{\substack{\widetilde{u}\in[u_{j-1/2}^-,\,u_{j+1/2}^-] \\ \widetilde{\widetilde{u}}\in[u_{j-1/2}^+,\,u_{j+1/2}^+]}} \left(\frac{\partial g}{\partial u^-}(\widetilde{u}, u_{j+1/2}^+) - \frac{\partial g}{\partial u^+}(u_{j-1/2}^-, \widetilde{\widetilde{u}})\right) \ .$$

For accuracy considerations away from extrema, it is desirable that the unlimited form of the discretization is obtained. Consequently, the constraint $\Psi(1) = 1$ is also imposed upon the limiter function. This constraint together with the algebraic conditions (62a-b) are readily achieved using the well known *MinMod* limiter, $\Psi^{\mathrm{MM}}$, with compression parameter $\beta$ determined from the TVD analysis

$$\Psi^{\mathrm{MM}}(R) = \max(0, \min(R, \beta)) \ , \quad \beta \in [1, (3-\kappa)/(1-\kappa)] \ .$$

Table I. Members of the MUSCL TVD family of schemes.

| $\kappa$ | Unlimited Scheme | $\beta_{\max}$ | Truncation Error |
|---|---|---|---|
| 1/3 | Third-Order | 4 | 0 |
| -1 | Fully Upwind | 2 | $\frac{1}{3}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$ |
| 0 | Fromm's | 3 | $\frac{1}{12}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$ |
| 1/2 | Low Truncation Error | 5 | $-\frac{1}{24}(\Delta x)^2 \frac{\partial^3}{\partial x^3} f(u)$ |

Table I summarizes the MUSCL scheme and maximum compression parameter for a number of familiar discretizations. Another limiter due to van Leer that meets the technical conditions of Theorem 3.5 and also satisfies $\Psi(1) = 1$ is given by

$$\Psi^{\mathrm{VL}}(R) = \frac{R + |R|}{1 + |R|} \ .$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

This limiter exhibits differentiability away from $R = 0$ which improves the iterative convergence to steady state for many algorithms. Numerous other limiter functions are considered and analyzed in Sweby, 1984.

Unfortunately, TVD schemes locally degenerate to piecewise constant approximations at smooth extrema which locally degrades the accuracy. This is an unavoidable consequence of the strict TVD condition.

**Theorem 3.6 (TVD critical point accuracy, Osher, 1984)** *The TVD discretizations (54), (56) and (58) all reduce to at most first order accuracy at non-sonic critical points, i.e. points $u^*$ at which $f'(u^*) \neq 0$ and $u_x^* = 0$.*

*3.1.3. ENO/WENO schemes.* To circumvent the degradation in accuracy of TVD schemes at critical points, weaker constraints on the solution total variation were devised. To this end, Harten proposed the following abstract framework for generalized Godunov schemes in operator composition form (see Harten et al., 1986; Harten et al., 1987; Harten, 1989)

$$u_h^{n+1} = A \cdot E(\tau) \cdot R_p^0(\cdot; u_h^n) \ . \tag{63}$$

In this equation, $u_h^n \in V_h^0$ denotes the global space of piecewise constant cell averages as defined in (20), $R_p^0(x)$ is a reconstruction operator which produces a cell-wise discontinuous $p$-th order polynomial approximation from the given solution cell averages, $E(\tau)$ is the evolution operator for the PDE (including boundary conditions), and $A$ is the cell averaging operator. Since $A$ is a nonnegative operator and $E(\tau)$ represents exact evolution in the small, the control of solution oscillations and Gibbs-like phenomena is linked directly to oscillation properties of the reconstruction operator, $R_p^0(x)$. One has formally in one space dimension

$$\mathrm{TV}(u_h^{n+1}) = \mathrm{TV}(A \cdot E(\tau) \cdot R_p^0(\cdot; u_h^n)) \leq \mathrm{TV}(R_p^0(x; u_h^n))$$

so that the total variation depends entirely upon properties of the reconstruction operator $R_p^0(x; u_h^n)$. The requirements of high order accuracy for smooth solutions and discrete conservation give rise to the following additional design criterion for the reconstruction operator

- $R_p^0(x; u_h) = u(x) + e(x)\ \Delta x^{p+1} + O(\Delta x^{p+2})$ whenever $u$ is smooth (64a)
- $A|_{T_j} R_p^0(x; u_h) = u_h|_{T_j} = u_j, \quad j = 1, \ldots, N$ to insure discrete conservation (64b)
- $\mathrm{TV}(R(x; u_h^n)) \leq \mathrm{TV}(u_h^n) + O(\Delta x^{p+1})$ an essentially non-oscillatory reconstruction. (64c)

Note that $e(x)$ may not be Lipschitz continuous at certain points so that the cumulative error in the scheme is $O(\Delta x^p)$ in a maximum norm but remains $O(\Delta x^{p+1})$ in an $L_1$-norm. To achieve the requirements of (64a-64c), Harten and coworkers considered breaking the task into two parts

- Polynomial reconstruction from a given stencil of cell averages
- Construction of a "smoothest" polynomial approximation by a solution adaptive stencil selection algorithm.

In the next section, a commonly used reconstruction technique from cell averages is considered. This is then followed by a description of the solution adaptive stencil algorithm proposed by Harten et al., 1986.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*3.1.4. Reconstruction via primitive function.* Given cell averages $u_j$ of a piecewise smooth function $u(x)$, one can inexpensively evaluate pointwise values of the *primitive function* $U(x)$

$$U(x) = \int_{x_0}^{x} u(\xi)\, d\xi$$

by exploiting the relationship

$$\sum_{j=j_0}^{j} \Delta x_j\, u_j = U(x_{j+1/2}) \ .$$

Let $H_p(x;u)$ denote a $p$-th order piecewise polynomial interpolant of a function $u$. Since

$$u(x) \equiv \frac{d}{dx} U(x) \ ,$$

an interpolant of the primitive function given pointwise samples $U(x_{j+1/2})$ yields a reconstruction operator

$$R_p^0(x;u_h) = \frac{d}{dx} H_{p+1}(x;U) \ .$$

As a polynomial approximation problem, whenever $U(x)$ is smooth one obtains

$$\frac{d^k}{dx^k} H_p(x;U) = \frac{d^k}{dx^k} U(x) + O(\Delta x^{p+1-k}) \ , 0 \le k \le p$$

and consequently

$$\frac{d^l}{dx^l} R_p(x;u_h) = \frac{d^l}{dx^l} u(x) + O(\Delta x^{p+1-l}) \ .$$

By virtue of the use of the primitive function $U(x)$, it follows that

$$A|_{T_j} R_p^0(x;u_h) = u_j$$

and from the polynomial interpolation problem for smooth data

$$R_p^0(x;u_h) = u(x) + O(\Delta x^{p+1})$$

as desired.

*3.1.5. ENO reconstruction.* The reconstruction technique outlined in Section 3.1.4 does not satisfy the oscillation requirement given in (64c). This motivated Harten and coworkers to consider a new algorithm for essentially non-oscillatory (ENO) piecewise polynomial interpolation. When combined with the reconstruction technique of Section 3.1.4, the resulting reconstruction then satisfies (64a-c). Specifically, a new interpolant $H_p(x;u)$ is constructed so that when applied to piecewise smooth data $v(x)$ gives high order accuracy

$$\frac{d^k}{dx^k} H_p(x;v) = \frac{d^k}{dx^k} v(x) + O(\Delta x^{p+1-k}) \ , 0 \le k \le p$$

but avoids having Gibbs oscillations at discontinuities in the sense

$$TV(H_p(x;v)) \le TV(v) + \mathcal{O}(\Delta x^{p+1}) \ .$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

The strategy pursued by Harten and coworkers was to construct such an ENO polynomial $H_p(x;w)$ using the following steps. Define

$$H_p^{\mathrm{ENO}}(x;w) = P_{p,j+1/2}^{\mathrm{ENO}}(x;w) \quad \text{for } x_j \le x \le x_{j+1}, \quad j = 1,\ldots,N$$

where $P_{p,j+1/2}^{\mathrm{ENO}}$ is the $p$-th degree polynomial which interpolates $w(x)$ at the $p+1$ successive points $\{x_i\}$, $i_p(j) \le i \le i_p(j)+p$ that include $x_j$ and $x_{j+1}$, i.e.

$$P_{p,j+1/2}^{\mathrm{ENO}}(x_i;w) = w(x_i) \ , \quad i_p(j) \le i \le i_p(j)+p \ , \quad 1-p \le i_p(j)-j \le 0 \ . \tag{65}$$

Equation (65) describes $p$ possible polynomials depending on the choice of $i_p(j)$ for an interval $(x_j, x_{j+1})$. The ENO strategy selects the value $i_p(j)$ for each interval that produces the "smoothest" polynomial interpolant for a given input data. More precisely, information about smoothness of $w(x)$ is extracted from a table of divided differences of $w(x)$ defined recursively for $i = 1,\ldots,N$ by

$$\begin{aligned} w[x_i] &= w(x_i) \\ w[x_i, x_{i+1}] &= \frac{w[x_{i+1}] - w[x_i]}{x_{i+1} - x_i} \\ &\vdots \\ w[x_i,\ldots,x_{i+k}] &= \frac{w[x_{i+1},\ldots,x_{i+k}] - w[x_i,\ldots,x_{i+k-1}]}{x_{i+k} - x_i} \ . \end{aligned}$$

The stencil producing the smoothest interpolant is then chosen hierarchically by setting

$$i_1(j) = j$$

and for $1 \le k \le p-1$

$$i_{k+1}(j) = \begin{cases} i_k(j) - 1 & \text{if } |w[x_{i_k(j)-1},\ldots,w[x_{i_k(j)+k}]| < |w[x_{i_k(j)},\ldots,w[x_{i_k(j)+k+1}]| \\ i_k(j) & \text{otherwise} \ . \end{cases} \tag{66}$$

Harten et al., 1986 demonstrate the following properties of this ENO interpolation strategy

- The accuracy condition

$$P_{p,j+1/2}^{\mathrm{ENO}}(x) = w(x) + \mathcal{O}(\Delta x^{p+1}) \ , \quad x \in (x_j, x_{j+1}) \ .$$

- $P_p^{\mathrm{ENO}}(x)$ is monotone in any cell interval containing a discontinuity.
- There exists a function $z(x)$ nearby $P_p^{\mathrm{ENO}}(x)$ in the interval $(x_j, x_{j+1})$ in the sense

$$z(x) = P_{p,j+1/2}^{\mathrm{ENO}}(x) + \mathcal{O}(\Delta x^{p+1}) \ , \quad x \in (x_j, x_{j+1})$$

that is total variation bounded, i.e. the nearby function $z(x)$ satisfies

$$TV(z) \le TV(w) \ .$$

*3.1.6. WENO reconstruction.* The solution adaptive nature of the ENO stencil selection algorithm (66) yields non-differentiable fluxes that impede convergence to steady state. In addition, the stencil selection algorithm chooses only one of $p$ possible stencils and other slightly less smooth stencils may give similar accuracy. When $w(x)$ is smooth, using a linear

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

combination of *all* $p$ stencils with optimized weights yields a more accurate $\mathcal{O}(\Delta x^{2p-1})$ interpolant. More specifically, let $P^{(k)}_{p,j+1/2}$ denote the unique polynomial interpolating $p+1$ points with stencil $\{x_{j+1-p+k}, x_{j+1+k}\}$ then

$$P_{p,j+1/2}(w(x)) = \sum_{k=0}^{p-1} \omega_k P^{(k)}_{p,j+1/2}(x) + \mathcal{O}(\Delta x^{2p-1}) \ , \quad \sum_{k=0}^{p-1} \omega_k = 1 \ .$$

For example, optimized weights for $p = 1, 2, 3$ yielding $\mathcal{O}(\Delta x^{2p-1})$ accuracy are readily computed

$$\begin{aligned} p=1: &\quad \omega_0 = 1, \\ p=2: &\quad \omega_0 = \frac{2}{3},\ \omega_1 = \frac{1}{3}, \\ p=3: &\quad \omega_0 = \frac{3}{10},\ \omega_1 = \frac{3}{5},\ \omega_2 = \frac{1}{10} \ . \end{aligned}$$

In the WENO schemes of Jiang and Shu, 1996; Shu, 1999, approximate weights, $\widetilde{\omega}$, are devised such that for smooth solutions

$$\widetilde{\omega}_k = \omega_k + \mathcal{O}(\Delta x^{p-1})$$

so that the $\mathcal{O}(\Delta x^{2p-1})$ accuracy is still retained using these approximations

$$P_{p,j+1/2}(w(x)) = \sum_{k=0}^{p-1} \widetilde{\omega}_k P^{(k)}_{p,j+1/2}(x) + \mathcal{O}(\Delta x^{2p-1}) \ , \quad \sum_{k=0}^{p-1} \widetilde{\omega}_k = 1 \ .$$

The approximate weights are constructed using the *ad hoc* formulas

$$\alpha_k = \frac{\omega_k}{(\epsilon + \beta_k)^2} \ , \quad \widetilde{\omega}_k = \frac{\alpha_k}{\sum_{k=0}^{p-1} \alpha_k}$$

where $\epsilon$ is an approximation to the square root of the machine precision and $\beta_k$ is a smoothness indicator

$$\beta_k = \sum_{l=1}^{k-1} \int_{x_{j-1/2}}^{x_{j+1/2}} \Delta x_j^{2l-1} \left( \frac{d^l P_p^k(x)}{\partial^l x} \right)^2 dx \ .$$

For a sequence of smooth solutions with decreasing smoothness indicator $\beta_k$, these formulas approach the optimized weights, $\widetilde{\omega}_k \to \omega_k$. These formulas also yield vanishing weights $\widetilde{\omega}_k \to 0$ for stencils with large values of the smoothness indicator such as those encountered at discontinuities. In this way, the WENO construction retains some of the attributes of the original ENO formulation but with increased accuracy in smooth solution regions and improved differentiability often yielding superior robustness for steady state calculations.

### *3.2. Higher order accurate FV schemes in multi-dimensions.*

Although the one-dimensional TVD operators may be readily applied in multi-dimensions on a dimension-by-dimension basis, a result of Goodman and LeVeque, 1985 shows that TVD schemes in two or more space dimensions are only first order accurate.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

**Theorem 3.7 (Accuracy of TVD schemes in multi-dimensions)** *Any two-dimensional finite volume scheme of the form*

$$u_{i,j}^{n+1} = u_{i,j}^n - \frac{\Delta t}{|T|_{i,j}}(g_{i+1/2,j}^n - g_{i-1/2,j}^n) - \frac{\Delta t}{|T|_{i,j}}(h_{i,j+1/2}^n - h_{i,j-1/2}^n) \ , \quad 1 \le i \le M,\ 1 \le j \le N$$

*with Lipschitz continuous numerical fluxes for integers $p, q, r, s$*

$$\begin{aligned} g_{i+1/2,j} &= g(u_{i-p,j-q}, \ldots, u_{i+r,j+s}), \\ h_{i,j+1/2} &= h(u_{i-p,j-q}, \ldots, u_{i+r,j+s}), \end{aligned}$$

*that is total variation non-increasing in the sense*

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

*where*

$$TV(u) \equiv \sum_{i,j} \left[\Delta y_{i+1/2,j}\,|u_{i+1,j} - u_{i,j}| + \Delta x_{i,j+1/2}\,|u_{i,j+1} - u_{i,j}|\right]$$

*is at most first-order accurate.*

Motivated by the negative results of Goodman and LeVeque, weaker conditions yielding solution monotonicity preservation have been developed from discrete maximum principle analysis. These alternative constructions have the positive attribute that they extend to unstructured meshes as well.

*3.2.1. Positive coefficient schemes on structured meshes.* Theorem 2.6 considers schemes of the form

$$u_j^{n+1} = u_j^n + \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} C_{jk}(u_h^n)(u_k^n - u_j^n) \ , \quad \forall T_j \in \mathcal{T}$$

and provides a local space-time discrete maximum principle

$$\min_{\forall e_{jk} \in \partial T_j} (u_k^n, u_j^n) \le u_j^{n+1} \le \max_{\forall e_{jk} \in \partial T_j} (u_k^n, u_j^n)$$

$\forall T_j \in \mathcal{T}$ under a CFL-like condition on the time step parameter if all coefficients $C_{jk}$ are nonnegative. Schemes of this type are often called *positive coefficient schemes* or more simply *positive schemes*. To circumvent the negative result of Theorem 3.7, Spekreijse, 1987 developed a family of high order accurate positive coefficient schemes on two-dimensional structured $M \times N$ meshes. For purposes of positivity analysis, these schemes are written in incremental form on a $M \times N$ logically rectangular 2-D mesh

$$\begin{aligned} u_{i,j}^{n+1} = u_{i,j}^n + \Delta t \,( \quad & A_{i+1,j}^n\,(u_{i+1,j}^n - u_{i,j}^n) + B_{i,j+1}^n\,(u_{i,j+1}^n - u_{i,j}^n) \\ + \quad & C_{i-1,j}^n\,(u_{i-1,j}^n - u_{i,j}^n) + D_{i,j-1}^n\,(u_{i,j-1}^n - u_{i,j}^n) \quad ) \ , \quad 1 \le i \le M,\ 1 \le j \le N \end{aligned} \tag{67}$$

where the coefficients are nonlinear functions of the solution

$$\begin{aligned} A_{i+1,j}^n &= A(\ldots, u_{i-1,j}^n, u_{i,j}^n, u_{i+1,j}^n, \ldots) \\ B_{i,j+1}^n &= B(\ldots, u_{i,j-1}^n, u_{i,j}^n, u_{i,j+1}^n, \ldots) \\ C_{i-1,j}^n &= C(\ldots, u_{i-1,j}^n, u_{i,j}^n, u_{i+1,j}^n, \ldots) \\ D_{i,j-1}^n &= D(\ldots, u_{i,j-1}^n, u_{i,j}^n, u_{i,j+1}^n, \ldots) \ . \end{aligned}$$

Once written in incremental form, the following theorem follows from standard positive coefficient maximum principle analysis.

*Encyclopedia of Computational Mechanics*. Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

**Theorem 3.8 (Positive coefficient schemes in multi-dimensions)** *The discretization (67) is a positive coefficient scheme if for each $1 \le i \le M,\ 1 \le j \le N$ and time slab increment $[t^n, t^{n+1}]$*

$$A^n_{i+1,j} \ge 0,\ B^n_{i,j+1} \ge 0,\ C^n_{i-1,j} \ge 0,\ D^n_{i,j-1} \ge 0, \tag{68}$$

*and*

$$1 - \Delta t\ (A^n_{i+1,j} + B^n_{i,j+1} + C^n_{i-1,j} + D^n_{i,j-1}) \ge 0 \tag{69}$$

*with discrete space-time maximum principle*

$$\min(u^n_{i,j}, u^n_{i-1,j}, u^n_{i+1,j}, u^n_{i,j-1}, u^n_{i,j+1}) \le u^{n+1}_{i,j} \le \max(u^n_{i,j}, u^n_{i-1,j}, u^n_{i+1,j}, u^n_{i,j-1}, u^n_{i,j+1})$$

*and discrete maximum principle at steady state*

$$\min(u^*_{i-1,j}, u^*_{i+1,j}, u^*_{i,j-1}, u^*_{i,j+1}) \le u^*_{i,j} \le \max(u^*_{i-1,j}, u^*_{i+1,j}, u^*_{i,j-1}, u^*_{i,j+1})$$

*where $u^*$ denotes the numerical steady state.*

Using a procedure similar to that used in the development of MUSCL TVD schemes in 1-D, Spekreijse, 1987 developed a family of monotonicity preserving MUSCL approximations in multi-dimensoins from the positivity conditions of Theorem 3.8.

**Theorem 3.9 (MUSCL positive coefficient scheme)** *The fully discrete 2-D finite volume scheme*

$$u^{n+1}_{i,j} = u^n_{i,j} - \frac{\Delta t}{|T|_{i,j}}(g^n_{i+1/2,j} - g^n_{i-1/2,j}) - \frac{\Delta t}{|T|_{i,j}}(h^n_{i,j+1/2} - h^n_{i,j-1/2})\ ,\quad 1 \le i \le M,\ 1 \le j \le N$$

*utilizing monotone Lipschitz continuous numerical flux functions*

$$\begin{aligned} g_{i+1/2,j} &= g(u^-_{i+1/2,j}, u^+_{i+1/2,j}) \\ h_{i,j+1/2} &= h(u^-_{i,j+1/2}, u^+_{i,j+1/2}) \end{aligned}$$

*and MUSCL extrapolation formulas*

$$\begin{aligned} u^-_{i+1/2,j} &= u_{i,j} + \frac{1}{2}\Psi(R_{i,j})\,(u_{i,j} - u_{i-1,j}) \\ u^+_{i+1/2,j} &= u_{i,j} + \frac{1}{2}\Psi(1/R_{i,j})\,(u_{i,j} - u_{i+1,j}) \\ u^-_{i,j+1/2} &= u_{i,j} + \frac{1}{2}\Psi(S_{i,j})\,(u_{i,j} - u_{i,j-1}) \\ u^+_{i,j+1/2} &= u_{i,j} + \frac{1}{2}\Psi(1/S_{i,j})\,(u_{i,j} - u_{i,j+1}) \end{aligned}$$

*where*

$$R_{i,j} \equiv \frac{u_{i+1,j} - u_{i,j}}{u_{i,j} - u_{i-1,j}}\ ,\quad S_{i,j} \equiv \frac{u_{i,j+1} - u_{i,j}}{u_{i,j} - u_{i,j-1}}$$

*satisfies the local maximum principle properties of Lemma 3.8 and is second order accurate if the limiter $\Psi = \Psi(R)$ has the properties that there exist constants $\beta \in (0, \infty)$, $\alpha \in [-2, 0]$ such that $\forall R \in \mathbb{R}$*

$$\alpha \le \Psi(R) \le \beta\ ,\quad -\beta \le \frac{\Psi(R)}{R} \le 2 + \alpha \tag{70}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*with the constraint* $\Psi(1) = 1$ *and the smoothness condition* $\Psi(R) \in C^2$ *near* $R = 1$ *together with a time step restriction for stability*

$$1 - (1+\beta)\frac{\Delta t}{|T_{i,j}|}\left(\left|\frac{\partial g}{\partial u}\right|^{n,\max}_{i,j} + \left|\frac{\partial h}{\partial u}\right|^{n,\max}_{i,j}\right) \geq 0$$

*where*

$$\left|\frac{\partial g}{\partial u}\right|^{\max}_{i,j} \equiv \sup_{\substack{\widetilde{u}\in[u^-_{i-1/2,j},u^-_{i+1/2,j}]\\ \widetilde{\widetilde{u}}\in[u^+_{i-1/2,j},u^+_{i+1/2,j}]}} \left(\frac{\partial g}{\partial u^-}(\widetilde{u}, u^+_{i+1/2,j}) - \frac{\partial g}{\partial u^+}(u^-_{i-1/2,j}, \widetilde{\widetilde{u}})\right) \geq 0$$

$$\left|\frac{\partial h}{\partial u}\right|^{\max}_{i,j} \equiv \sup_{\substack{\widehat{u}\in[u^-_{i,j-1/2},u^-_{i,j+1/2}]\\ \widehat{\widehat{u}}\in[u^+_{i,j-1/2},u^+_{i,j+1/2}]}} \left(\frac{\partial h}{\partial u^-}(\widehat{u}, u^+_{i,j+1/2}) - \frac{\partial h}{\partial u^+}(u^-_{i,j-1/2}, \widehat{\widehat{u}})\right) \geq 0 \ .$$

Many limiter functions satisfy the technical conditions (70) of Theorem 3.9. Some examples include

- the van Leer limiter
$$\Psi^{\mathrm{VL}}(R) = \frac{R + |R|}{1 + |R|} \ ,$$
- the van Albada limiter
$$\Psi^{\mathrm{VA}}(R) = \frac{R + R^2}{1 + R^2} \ .$$

In addition, Koren, 1988 has constructed the limiter

$$\Psi^{\mathrm{K}}(R) = \frac{R + 2R^2}{2 - R + 2R^2}$$

which also satisfies the technical conditions (70) and corresponds for smooth solutions in 1-D to the most accurate $\kappa = 1/3$ MUSCL scheme of van Leer.

*3.2.2. FV schemes on unstructured meshes utilizing linear reconstruction.* Higher order finite volume extensions of Godunov discretization to unstructured meshes are of the general form

$$\frac{du_j}{dt} = -\frac{1}{|T_j|}\sum_{\forall e_{jk}\in\partial T_j} g_{jk}(u^-_{jk}, u^+_{jk}) \ , \quad \forall T_j \in \mathcal{T} \tag{71}$$

with the numerical flux $g_{jk}(u,v)$ given by the quadrature rule

$$g_{jk}(u^-_{jk}, u^+_{jk}) \equiv \sum_{q=1}^{Q} \omega_q \, g(\nu_{jk}(x_q); u^-_{jk}(x_q), u^+_{jk}(x_q)) \ , \tag{72}$$

where $\omega_q \in \mathbb{R}$ and $x_q \in e_{jk}$ represent quadrature weights and locations, $q = 1, \ldots, Q$. Given the global space of piecewise constant cell averages, $u_h \in V^0_h$, the extrapolated states $u^-_{jk}(x)$ and $u^+_{jk}(x)$ are evaluated using a $p$-th order polynomial reconstruction operator, $R^0_p : V^0_h \mapsto V^p_h$,

$$u^-_{jk}(x) \equiv \lim_{\epsilon\downarrow 0} R^0_p(x - \epsilon\,\nu_{jk}(x); u_h)$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

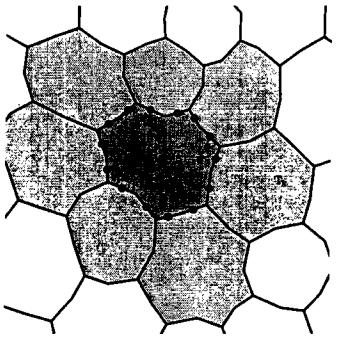


Figure 4. Polygonal control volume cell $T_j$ and perimeter quadrature points (solid circles).

$$u_{jk}^{+}(x) \equiv \lim_{\epsilon \downarrow 0} R_p^0(x + \epsilon\, \nu_{jk}(x); u_h)$$

for $x \in e_{jk}$. In addition, it is assumed that the reconstruction satisfies the property $\frac{1}{|T_j|}\int_{T_j} R_p^0(x; u_h)\, dx = u_j$ stated previously in (64b). In the general finite volume formulation, the control volume shapes need not be convex, see for example Fig. 4. Even so, the solution accuracy and maximum stable time step for explicit schemes may depend strongly on the shape of individual control volumes. In the special case of linear reconstruction, $R_1^0(x; u_h)$, the impact of control volume shape on stability of the scheme can be quantified more precisely. Specifically, the maximum principle analysis presented later for the scheme (71) reveals an explicit dependence on the geometrical shape parameter

$$\Gamma^{geom} = \sup_{0 \le \theta \le 2\pi} \alpha^{-1}(\theta) \tag{73}$$

where $0 < \alpha(\theta) < 1$ represents the smallest fractional perpendicular distance from the gravity center to one of two minimally separated parallel hyperplanes with orientation $\theta$ and hyperplane location such that all quadrature points in the control volume lie between or on the hyperplanes as shown in Fig. 5. Table II lists $\Gamma^{geom}$ values for various control volume shapes in $\mathbb{R}^1$, $\mathbb{R}^2$, $\mathbb{R}^3$, and $\mathbb{R}^d$. As might be expected, those geometries that have exact quadrature point symmetry with respect to the control volume gravity center have geometric shape parameters $\Gamma^{\mathrm{geom}}$ equal to 2 regardless of the number of space dimensions involved.

**Lemma 3.10 (Finite volume interval bounds on unstructured meshes, $\mathbf{R_1^0(x; u_h)}$)** *The fully discrete finite volume scheme*

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} g_{jk}(u_{jk}^{-,n}, u_{jk}^{+,n}) \quad , \quad \forall T_j \in \mathcal{T} \tag{74}$$

*with monotone Lipschitz continuous numerical flux function, nonnegative quadrature weights, and linear reconstructions*

$$\begin{aligned} u_{jk}^{-}(x) &\equiv \lim_{\epsilon \downarrow 0} R_1^0(x - \epsilon\, \nu_{jk}(x); u_h) \quad , \quad x \in e_{jk} \quad , \quad u_h \in V_h^0 \\ u_{jk}^{+}(x) &\equiv \lim_{\epsilon \downarrow 0} R_1^0(x + \epsilon\, \nu_{jk}(x); u_h) \quad , \quad x \in e_{jk} \quad , \quad u_h \in V_h^0 \quad , \end{aligned}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

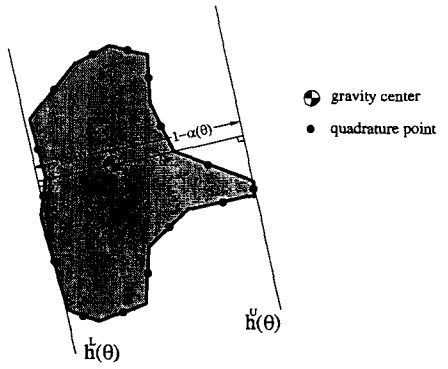



Figure 5. Minimally separated hyperplanes $h^L(\theta)$ and $h^U(\theta)$ and the fractional distance ratio $\alpha(\theta)$ for use in the calculation of $\Gamma^{\text{geom}}$.

Table II. Reconstruction geometry factors for various control volume shapes utilizing midpoint quadrature rule.

| control volume shape | space dimension | $\Gamma^{\text{geom}}$ |
|---|---|---|
| segment | 1 | 2 |
| triangle | 2 | 3 |
| parallelogram | 2 | 2 |
| regular $n$-gon | 2 | $n/\lceil\frac{n-1}{2}\rceil$ |
| tetrahedron | 3 | 4 |
| parallelepiped | 3 | 2 |
| simplex | d | d+1 |
| hyper-parallelepiped | d | 2 |
| polytope | d | Eqn. (73) |

*with extremal trace values at control volume quadrature points*

$$U_j^{\min} \equiv \min_{\substack{\forall e_{jk}\in\partial T_j \\ 1\le q\le Q}} u_{jk}^{\pm}(x_q) \ , \quad U_j^{\max} \equiv \max_{\substack{\forall e_{jk}\in\partial T_j \\ 1\le q\le Q}} u_{jk}^{\pm}(x_q) \ , \quad x_q \in e_{jk}$$

*exhibits the local interpolated interval bound*

$$\sigma_j U_j^{\min,n} + (1-\sigma_j)u_j^n \le u_j^{n+1} \le (1-\sigma_j)\, u_j^n + \sigma_j U_j^{\max,n} \tag{75}$$

*with the time step proportional interpolation parameter $\sigma_j$ defined by*

$$\sigma_j \equiv \frac{\Delta t}{|T_j|}\, \Gamma^{\text{geom}} \sum_{\substack{\forall e_{jk}\in\partial T_j \\ 1\le q\le Q}} \sup_{\substack{\widetilde{u}\in[U_j^{\min,n},U_j^{\max,n}] \\ \widetilde{\widetilde{u}}\in[U_j^{\min,n},U_j^{\max,n}]}} \left|\frac{\partial g}{\partial u^+}(\nu_{jk}(x_q));\widetilde{u},\widetilde{\widetilde{u}})\right| \tag{76}$$

*that depends on the shape parameter $\Gamma^{\text{geom}}$ defined in (73).*

*Encyclopedia of Computational Mechanics*. Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Given the two-sided bound of Lemma 3.10, a discrete maximum principle is obtained under a CFL-like time step restriction if the limits $U_j^{\max}$ and $U_j^{\min}$ can be bounded from above and below respectively by the neighboring cell averages. This idea is given more precisely in the following theorem.

**Theorem 3.11 (Finite volume maximum principle on unstructured meshes, $\mathbf{R}_1^0$)** *Let $u_j^{\min}$ and $u_j^{\max}$ denote the minimum and maximum value of solution cell averages for a given cell $T_j$ and corresponding adjacent cell neighbors, i.e.*

$$u_j^{\min} \equiv \min_{\forall e_{jk} \in \partial T_j} (u_j, u_k) \quad \text{and} \quad u_j^{\max} \equiv \max_{\forall e_{jk} \in \partial T_j} (u_j, u_k) \ . \tag{77}$$

*The fully discrete finite volume scheme*

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} g_{jk}(u_{jk}^{-,n}, u_{jk}^{+,n}) \ , \quad \forall T_j \in \mathcal{T} \tag{78}$$

*with monotone Lipschitz continuous numerical flux function, nonnegative quadrature weights, and linear reconstructions*

$$\begin{aligned} u_{jk}^-(x) &\equiv \lim_{\epsilon \downarrow 0} R_1^0(x - \epsilon \nu_{jk}(x); u_h) \ , \quad x \in e_{jk} \ , \quad u_h \in V_h^0 \\ u_{jk}^+(x) &\equiv \lim_{\epsilon \downarrow 0} R_1^0(x + \epsilon \nu_{jk}(x); u_h) \ , \quad x \in e_{jk} \ , \quad u_h \in V_h^0 \end{aligned} \tag{79}$$

*exhibits the local space-time maximum principle for each $T_j \in \mathcal{T}$*

$$\min_{\forall e_{jk} \in \partial T_j} (u_j^n, u_k^n) \le u_j^{n+1} \le \max_{\forall e_{jk} \in \partial T_j} (u_j^n, u_k^n)$$

*as well as the local spatial maximum principle at steady state ($u^{n+1} = u^n = u^*$)*

$$\min_{\forall e_{jk} \in \partial T_j} u_k^* \le u_j^* \le \max_{\forall e_{jk} \in \partial T_j} u_k^*$$

*if the linear reconstruction satisfies $\forall e_{jk} \in \partial T_j$ and $x_q \in e_{jk}, q = 1, \ldots, Q$*

$$\max(u_j^{\min,n}, u_k^{\min,n}) \le u_{jk}^{-,n}(x_q) \le \min(u_j^{\max,n}, u_k^{\max,n}) \tag{80}$$

*under the time step restriction*

$$1 - \frac{\Delta t}{|T_j|} \Gamma^{\text{geom}} \sum_{\substack{\forall e_{jk} \in \partial T_j \\ 1 \le q \le Q}} \sup_{\substack{\widetilde{u} \in [u_j^{\min,n}, u_j^{\max,n}] \\ \widetilde{\widetilde{u}} \in [u_j^{\min,n}, u_j^{\max,n}]}} \left| \frac{\partial g}{\partial u^+}(\nu_{jk}(x_q)); \widetilde{u}, \widetilde{\widetilde{u}}) \right| \ge 0$$

*with $\Gamma^{\text{geom}}$ defined in (73).*

Note that a variant of this theorem also holds if the definition of $u^{\max}$ and $u^{\min}$ are expanded to include more control volume neighbors. Two alternative definitions frequently used when the control volume shape is a simplex are given by

$$u_j^{\min} \equiv \min_{\substack{T_k \in \mathcal{T} \\ T_j \cap T_k \neq \emptyset}} u_k \quad \text{and} \quad u_j^{\max} \equiv \max_{\substack{T_k \in \mathcal{T} \\ T_j \cap T_k \neq \emptyset}} u_k \ . \tag{81}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

These expanded definitions include adjacent cells whose intersection with $T_j$ in $\mathbb{R}^d$ need only be a set of measure zero or greater.

*Slope limiters for linear reconstruction.* Given a linear reconstruction $R_1^0(x;u_h)$ that does not necessarily satisfy the requirements of Theorem 3.11, it is straightforward to modify the reconstruction so that the new modified reconstruction does satisfy the requirements of Theorem 3.11. For each control volume $T_j \in \mathcal{T}$ a modified reconstruction operator $\widetilde{R}_1^0(x;u_h)$ of the form

$$\widetilde{R}_1^0(x;u_h)|_{T_j} = u_j + \alpha_{T_j}(R_1^0(x;u_h)|_{T_j} - u_j)$$

is assumed for $\alpha_{T_j} \in [0,1]$. By construction, this modified reconstruction correctly reproduces the control volume cell average for all values of $\alpha_{T_j}$, i.e.

$$\frac{1}{|T_j|}\int_{T_j} \widetilde{R}_1^0(x;u_h)\,dx = u_j \ . \tag{82}$$

The most restrictive value of $\alpha_{T_j}$ for each control volume $T_j$ is then computed based on the Theorem 3.11 constraint (80), i.e.

$$\alpha_{T_j}^{\mathrm{MM}} = \min_{\substack{\forall e_{jk}\in\partial T_j \\ 1\le q\le Q}} \begin{cases} \frac{\min(u_j^{\max},u_k^{\max})-u_j}{R_1^0(x_q;u_h)|_{T_j}-u_j} & \text{if } R_1^0(x_q;u_h)|_{T_j} > \min(u_j^{\max},u_k^{\max}) \\ \frac{\max(u_j^{\min},u_k^{\min})-u_j}{R_1^0(x_q;u_h)|_{T_j}-u_j} & \text{if } R_1^0(x_q;u_h)|_{T_j} < \max(u_j^{\min},u_k^{\min}) \\ 1 & \text{otherwise} \end{cases} \tag{83}$$

where $u^{\max}$ and $u^{\min}$ are defined in (77). When the resulting modified reconstruction operator is used in the extrapolation formulas (79), the discrete maximum principle of Theorem 3.11 is attained under a CFL-like time step restriction. By utilizing the inequalities

$$\max(u_j,u_k) \le \min(u_j^{\max},u_k^{\max}) \quad \text{and} \quad \min(u_j,u_k) \ge \max(u_j^{\min},u_k^{\min})$$

it is straightforward to construct a simpler but more restrictive limiter function

$$\alpha_{T_j}^{\mathrm{LM}} = \min_{\substack{\forall e_{jk}\in\partial T_j \\ 1\le q\le Q}} \begin{cases} \frac{\max(u_j,u_k)-u_j}{R_1^0(x_q;u_h)|_{T_j}-u_j} & \text{if } R_1^0(x_q;u_h)|_{T_j} > \max(u_j,u_k) \\ \frac{\min(u_j,u_k)-u_j}{R_1^0(x_q;u_h)|_{T_j}-u_j} & \text{if } R_1^0(x_q;u_h)|_{T_j} < \min(u_j,u_k) \\ 1 & \text{otherwise} \end{cases} \tag{84}$$

that yields modified reconstructions satisfying the technical conditions of Theorem 3.11. This simplified limiter (84) introduces additional slope reduction when compared to (83). This can be detrimental to the overall accuracy of the discretization. The limiter strategy (84) and other variants for simplicial control volumes are discussed further in Liu, 1993; Wierse, 1994; Batten, Lambert and Causon, 1996.

In Barth and Jespersen, 1989, a variant of (83) was proposed

$$\alpha_{T_j}^{\mathrm{BJ}} = \min_{\substack{\forall e_{jk}\in\partial T_j \\ 1\le q\le Q}} \begin{cases} \frac{u_j^{\max}-u_j}{R_1^0(x_q;u_h)|_{T_j}-u_j} & \text{if } R_1^0(x_q;u_h)|_{T_j} > u_j^{\max} \\ \frac{u_j^{\min}-u_j}{R_1^0(x_q;u_h)|_{T_j}-u_j} & \text{if } R_1^0(x_q;u_h)|_{T_j} < u_j^{\min} \\ 1 & \text{otherwise} \end{cases} \ . \tag{85}$$

Although this limiter function does not produce modified reconstructions satisfying the requirements of Theorem 3.11, using Lemma 3.10 it can be shown that the Barth and Jespersen

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

limiter yields finite volume schemes (74) possessing a global extremum diminishing property, i.e. that the solution maximum is non-increasing and the solution minimum is nondecreasing between successive time levels. This limiter function produces the least amount of slope reduction when compared to the limiter functions (83) and (84). Note that in practical implementation, all three limiters (83), (84) and (85) require some modification to prevent near zero division for nearly constant solution data.

*3.2.3. Linear reconstruction operators on simplicial control volumes.* Linear reconstruction operators on simplicial control volumes that satisfy the cell averaging requirement (64b) often exploit the fact that the cell average is also a pointwise value of any valid linear reconstruction evaluated at the gravity center of a simplex. This reduces the reconstruction problem to that of gradient estimation given pointwise samples at the gravity centers. In this case, it is convenient to express the reconstruction in the form

$$R_1^0(x; u_h)|_{T_j} = u_j + (\nabla u_h)_{T_j} \cdot (x - x_j^{\circledcirc}) \tag{86}$$

where $x_j^{\circledcirc}$ denotes the gravity center for the simplex $T_j$ and $(\nabla u_h)_{T_j}$ is the gradient to be determined. Figure 6 depicts a 2-D simplex $\triangle_{123}$ and three adjacent neighboring simplices. Also shown are the corresponding four pointwise solution values $\{A, B, C, O\}$ located at gravity centers of each simplex. By selecting any three of the four pointwise solution values, a set of four possible gradients are uniquely determined, i.e. $\{\nabla(ABC), \nabla(ABO), \nabla(BCO), \nabla(CAO)\}$. Using the example of Fig. 6, a number of slope limited reconstruction techniques are possible

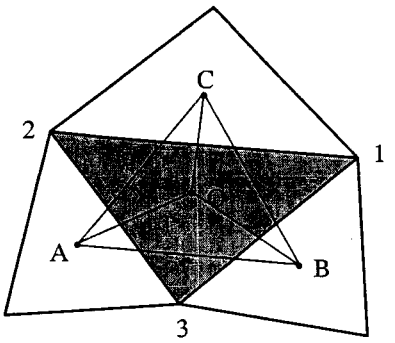



Figure 6. Triangle control volume $\triangle_{123}$ (shaded) with three adjacent cell neighbors.

for use in the finite volume scheme (78) that meet the technical conditions of Theorem 3.11.

1. Choose $(\nabla u_h)_{T_{123}} = \nabla(ABC)$ and limit the resulting reconstruction using (83) or (84). This technique is pursued in Barth and Jespersen, 1989 but using the limiter (85) instead.
2. Limit the reconstructions corresponding to gradients $\nabla(ABC), \nabla(ABO), \nabla(BCO)$ and $\nabla(CAO)$ using (83) or (84) and choose the limited reconstruction with largest gradient magnitude. This technique is a generalization of that described in Batten, Lambert and Causon, 1996 wherein limiter (84) is used.
3. Choose the unlimited reconstruction $\nabla(ABC), \nabla(ABO), \nabla(BCO)$ and $\nabla(CAO)$ with largest gradient magnitude that satisfies the maximum principle reconstruction bound inequality (80). If all reconstructions fail the bound inequality, the reconstruction gradient is set equal to zero, see Liu, 1993.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

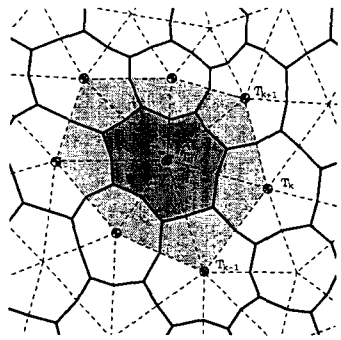


Figure 7. Triangulation of gravity center locations showing a typical control volume $T_0$ associated with the triangulation vertex $v_0$ with cyclically indexed graph neighbors $T_k, k = 1, \ldots, N_0$.

*3.2.4. Linear reconstruction operators on general control volumes shapes.* In the case of linear reconstruction on general volume shapes, significant simplification is possible when compared to the general $p$-exact reconstruction formulation given in Sect. 3.2.5. It is again convenient to express the reconstruction in the form

$$R_1^0(x; u_h)|_{T_j} = u_j + (\nabla u_h)_{T_j} \cdot (x - x_j^{\diamondsuit}) \tag{87}$$

where $x_j^{\diamondsuit}$ denotes the gravity center for the control volume $T_j$ and $(\nabla u_h)_{T_j}$ is the gradient to be determined. Two common techniques for simplified linear reconstruction include a simplified least squares technique and a Green-Gauss integration technique.

*Simplified least squares linear reconstruction.* As was exploited in the linear reconstruction techniques for simplicial control volumes, linear reconstructions satisfying (64b) on general control volume shapes are greatly simplified by exploiting the fact that the cell average value is also a pointwise value of all valid linear reconstructions evaluated at the gravity center of a general control volume shape. This again reduces the linear reconstruction problem to that of gradient estimation given pointwise values. In the simplified least squares reconstruction technique, a triangulation (2D) or tetrahedralization (3D) of gravity centers is first constructed as shown in Fig. 7. Referring to this figure, for each edge of the simplex mesh incident to the vertex $v_0$, an edge projected gradient constraint equation is constructed subject to a prespecified nonzero scaling $w_k$

$$w_k \, (\nabla u_h)_{T_0} \cdot (x_k^{\diamondsuit} - x_0^{\diamondsuit}) = w_k (u_k - u_0) \ .$$

The number of edges incident to a simplex mesh vertex in $\mathbb{R}^d$ is greater than or equal to $d$ thereby producing the following generally non-square matrix of constraint equations

$$\begin{bmatrix} w_1 \Delta x_1^{\diamondsuit} & w_1 \Delta y_1^{\diamondsuit} \\ \vdots & \vdots \\ w_{N_0} \Delta x_{N_0}^{\diamondsuit} & w_{N_0} \Delta y_{N_0}^{\diamondsuit} \end{bmatrix} (\nabla u_h)_{T_0} = \begin{pmatrix} w_1 (u_1 - u_0) \\ \vdots \\ w_{N_0} (u_{N_0} - u_0) \end{pmatrix}$$

or in abstract form

$$\begin{bmatrix} \vec{L}_1 & \vec{L}_2 \end{bmatrix} \nabla u = \vec{f} \ .$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

This abstract form can be symbolically solved in a least squares sense using an orthogonalization technique yielding the closed form solution

$$\nabla u = \frac{1}{l_{11}l_{22} - l_{12}^2} \begin{pmatrix} l_{22}(\vec{L}_1 \cdot \vec{f}) - l_{12}(\vec{L}_2 \cdot \vec{f}) \\ l_{11}(\vec{L}_2 \cdot \vec{f}) - l_{12}(\vec{L}_1 \cdot \vec{f}) \end{pmatrix} \tag{88}$$

with $l_{ij} = \vec{L}_i \cdot \vec{L}_j$. The form of this solution in terms of scalar dot products over incident edges suggests that the least squares linear reconstruction can be efficiently computed via an edge data structure without the need for storing a non-square matrix.

*Green-Gauss linear reconstruction.* This reconstruction technique specific to simplicial meshes assumes nodal solution values at vertices of the mesh which uniquely describes a $C^0$ linear interpolant, $u_h$. Gradients are then computed from the mean value approximation

$$|\Omega_0| \, (\nabla u_h)_{\Omega_0} \approx \int_{\Omega_0} \nabla u_h \, dx = \int_{\partial\Omega_0} u_h \, d\nu \ . \tag{89}$$

For linear interpolants, the right-hand side term can be written in the following equivalent

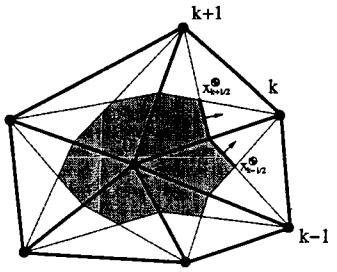



Figure 8. Median dual control volume $T_0$ demarcated by median segments of triangles incident to the vertex $v_0$ with cyclically indexed adjacent vertices $v_k, k = 1, \ldots N_0$.

form using the configuration depicted in Fig. 8

$$\int_{\Omega_0} \nabla u_h \, dx = \sum_{k=1}^{N_0} \frac{3}{2}(u_0 + u_k) \, \nu_{0k}$$

where $\nu_{0k}$ represents *any* path integrated normal connecting pairwise adjacent simplex gravity centers, i.e.

$$\nu_{0k} = \int_{x^{\bullet}_{k-1/2}}^{x^{\bullet}_{k+1/2}} d\nu \ . \tag{90}$$

A particularly convenient path is one that traces out portions of median segments as shown in Fig. 8. These segments demarcate the so-called *median dual* control volume. By construction, the median dual volume $|T_0|$ is precisely equal to $|\Omega_0|/3$ in 2-D. Consequently, a linear reconstruction operator on non-overlapping median dual control volumes is given by

$$|T_0|(\nabla u_h)_{T_0} \approx \sum_{k=1}^{N_0} \frac{1}{2}(u_0 + u_k) \, \nu_{0k} \ . \tag{91}$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

The gradient calculation is exact whenever the numerical solution varies linearly over the support of the reconstruction. Since mesh vertices are not located at the gravity centers of median dual control volumes, the cell averaging property (64b) and the bounds of Theorem 3.11 are only approximately satisfied using the Green-Gauss technique.

A number of slope limited linear reconstruction strategies for general control volume shapes are possible for use in the finite volume scheme (78) that satisfy the technical conditions of Theorem 3.11. Using the example depicted in Fig. 7, let $\nabla_{k+1/2}u_h$ denote the unique linear gradient calculated from the cell average set $\{u_0, u_k, u_{k+1}\}$. Three slope limiting strategies that are direct counterparts of the simplex control volume case are:

1. Compute $(\nabla u_h)_{T_0}$ using the least squares linear reconstruction or any other valid linear reconstruction technique and limit the resulting reconstruction using (83) or (84).
2. Limit the reconstructions corresponding to the gradients $\nabla_{k+1/2}u_h, k = 1, \ldots, N_0$ and $(\nabla u_h)_{T_0}$ using (83) or (84) and choose the limited reconstruction with largest gradient magnitude.
3. Choose the unlimited reconstruction from $\nabla_{k+1/2}u_h, k = 1, \ldots, N_0$ and $(\nabla u_h)_{T_0}$ with largest gradient magnitude that satisfies the maximum principle reconstruction bound inequality (80). If all reconstructions fail the bound inequality, the reconstruction gradient is set equal to zero.

*3.2.5. General p-exact reconstruction operators on unstructured meshes.* Abstractly, the reconstruction operator serves as a finite-dimensional pseudo inverse of the cell averaging operator $A$ whose $j$-th component $A_j$ computes the cell average of the solution in $T_j$

$$A_j u = \frac{1}{|T_j|} \int_{T_j} u \, dx \ .$$

The development of a general polynomial reconstruction operator, $R_p^0$, that reconstructs $p$-degree polynomials from cell averages on unstructured meshes follows from the application of a small number of simple properties.

1. (Conservation of the mean) Given solution cell averages $u_h$, the reconstruction $R_p^0 u_h$ is required to have the correct cell average, i.e.

   $$\text{if } v = R_p^0 u_h \text{ then } u_h = Av \ .$$

   More concisely,

   $$AR_p^0 = I$$

   so that $R_p^0$ is a right inverse of the averaging operator $A$.
2. ($p$-exactness) A reconstruction operator $R_p^0$ is $p$-exact if $R_p^0 A$ reconstructs polynomials of degree $p$ or less exactly. Denoting by $\mathcal{P}_p$ the space of all polynomials of degree $p$,

   $$\text{if } u \in \mathcal{P}_p \text{ and } v = Au \text{ then } R_p^0 v = u \ .$$

   This can be written succinctly as

   $$R_p^0 A|_{\mathcal{P}_p} = I$$

   so that $R_p^0$ is a left inverse of the averaging operator $A$ restricted to the space of polynomials of degree at most $p$.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

3. (Compact support) The reconstruction in a control volume $T_j$ should only depend of cell averages in a relatively small neighborhood surrounding $T_j$. Recall that a polynomial of degree $p$ in $\mathbb{R}^d$ contains $\binom{p+d}{d}$ degrees of freedom. The support set for $T_j$ is required to contain at least this number of neighbors. As the support set becomes even larger for fixed $p$, not only does the computational cost increase, but eventually the accuracy decreases as less valid data from further away is brought into the calculation.

Practical implementations of polynomial reconstruction operators fall into two classes:

- Fixed support stencil reconstructions. These methods choose a fixed support set as a preprocessing step. Various limiting strategies are then employed to obtain non-oscillatory approximation, see for example Barth and Frederickson, 1990; Delanaye, 1996 for further details.
- Adaptive support stencil reconstructions. These ENO-like methods dynamically choose reconstruction stencils based on solution smoothness criteria, see for example Harten and Chakravarthy, 1991; Vankeirsblick, 1993; Abgrall, 1994; Sonar, 1997; Sonar, 1998 for further details.

*3.2.6. Positive coefficient schemes on unstructured meshes* Several related positive coefficient schemes have been proposed on multi-dimensional simplicial meshes based on one-dimensional interpolation. The simplest example is the *upwind triangle scheme* as introduced by Billey et al., 1987; Desideri and Dervieux, 1988; Rostand and Stoufflet, 1988 with later improved variants given by Jameson, 1993; Cournède, Debiez and Dervieux, 1998. These schemes are not Godunov methods in the sense that a single multi-dimensional gradient is not obtained in each control volume. The basis for these methods originates from the gradient estimation formula (91) generalized to the calculation of flux divergence on a median dual tessellation. In deriving this flux divergence formula, the assumption has been made that flux components vary linearly within a simplex yielding the discretization formula

$$\int_{T_j} \operatorname{div}(f)\, dx = \int_{\partial T_j} f \cdot d\nu = \sum_{\forall e_{jk} \in \partial T_j} \frac{1}{2} \left(f(u_j) + f(u_k)\right) \cdot \nu_{jk}$$

where $\nu_{jk}$ is computed from a median dual tessellation using (90). This discretization is the unstructured mesh counterpart of central differencing on a structured mesh. Schemes using this discretization of flux divergence lack sufficient stability properties for computing solutions of general nonlinear conservation laws. This lack of stability can be overcome by adding suitable diffusion terms. One of the simplest modifications is motivated by upwind domain of dependence arguments yielding the numerical flux

$$g_{jk}(u_j, u_k) = \frac{1}{2} \left(f(u_j) + f(u_k)\right) \cdot \nu_{jk} - \frac{1}{2} |a|_{jk}\, \Delta_{jk} u \tag{92}$$

with $a_{jk}$ a mean value (a.k.a. Murman-Cole) linearization satisfying

$$\nu_{jk} \cdot \Delta_{jk} f = a_{jk}\, \Delta_{jk} u \ .$$

Away from sonic points where $f'(u^*) = 0$ for $u^* \in [u_j, u_{j+1}]$, this numerical flux is formally an E-flux satisfying (28). With suitable modifications of $a_{jk}$ near sonic points, it is then possible to produce a modified numerical flux that is an E-flux for all data, see Osher, 1984. Theorems

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

2.6, 2.7 and 2.8 show that schemes such as (22) using E-fluxes exhibit local discrete maximum principles and $L_\infty$ stability.

Unfortunately, schemes based on (92) are too dissipative for most practical calculations. The main idea in the upwind triangle scheme is to add anti-diffusion terms to the numerical flux function (92) such that the sum total of added diffusion and anti-diffusion terms in the numerical flux function vanish entirely whenever the numerical solution varies linearly over the support of the flux function. In all remaining situations, the precise amount of anti-diffusion is determined from maximum principle analysis.

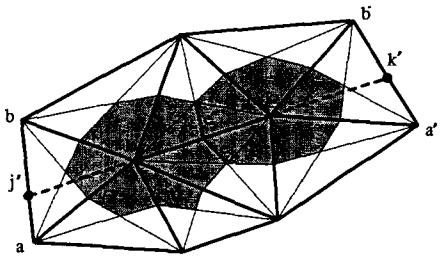



Figure 9. Triangle complex used in the upwind triangle schemes showing the linear extension of $e_{jk}$ into neighboring triangle for the determination of points $x_{j'}$ and $x_{k'}$.

**Theorem 3.12 (Maximum Principles for the Upwind Triangle Scheme)** *Let $\mathcal{T}$ denote the median dual tessellation of an underlying simplicial mesh. Also let $u_j$ denote the nodal solution value at a simplex vertex in one-to-one dual correspondence with the control volume $T_j \in \mathcal{T}$ such that a $C^0$ linear solution interpolant is uniquely specified on the simplicial mesh. Let $g_{jk}(u_{j'}, u_j, u_k, u_{k'})$ denote the numerical flux function with limiter function $\Psi(\cdot) : \mathbb{R} \mapsto \mathbb{R}$*

$$g_{jk}(u_{j'}, u_j, u_k, u_{k'}) \equiv \frac{1}{2}(f(u_j) + f(u_k)) \cdot \nu_{jk} \quad - \quad \frac{1}{2} a^+_{jk} \left(1 - \Psi\left(\frac{h_{jk}\ \Delta_{j'j} u}{h_{j'j}\ \Delta_{jk} u}\right)\right) \Delta_{jk} u$$
$$+ \quad \frac{1}{2} a^-_{jk} \left(1 - \Psi\left(\frac{h_{jk}\ \Delta_{kk'} u}{h_{kk'}\ \Delta_{jk} u}\right)\right) \Delta_{jk} u \ ,$$

*utilizing the mean value speed $a_{jk}$ satisfying*

$$\nu_{jk} \cdot \Delta_{jk} f = a_{jk}\, \Delta_{jk} u$$

*and variable spacing parameter $h_{jk} = |\Delta_{jk} x|$. The fully discrete finite volume scheme*

$$u_j^{n+1} = u_j^n - \frac{\Delta t}{|T_j|} \sum_{\forall e_{jk} \in \partial T_j} g_{jk}(u_{j'}^n, u_j^n, u_k^n, u_{k'}^n) \ , \quad \forall T_j \in \mathcal{T} \ ,$$

*with linearly interpolated values $u_{j'}$ and $u_{k'}$ as depicted in Fig. 9 exhibits the local space-time maximum principle*

$$\min_{\forall e_{jk} \in \partial T_j} (u_j^n, u_k^n) \le u_j^{n+1} \le \max_{\forall e_{jk} \in \partial T_j} (u_j^n, u_k^n)$$

*and the local spatial maximum principle at steady state ($u^{n+1} = u^n = u^*$)*

$$\min_{\forall e_{jk} \in \partial T_j} u_k^* \le u_j^* \le \max_{\forall e_{jk} \in \partial T_j} u_k^*$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*if the limiter* $\Psi(R)$ *satisfies* $\forall R \in \mathbb{R}$

$$0 \leq \Psi(R)/R \ , \quad 0 \leq \Psi(R) \leq 2 \ .$$

Some standard limiter functions that satisfy the requirements of Theorem 3.12 include

- the MinMod limiter with maximum compression parameter equal to 2

$$\Psi^{\mathrm{MM}}(R) = \max(0, \min(R, 2))$$

- the van Leer limiter

$$\Psi^{\mathrm{VL}}(R) = \frac{R + |R|}{1 + |R|} \ .$$

Other limiter formulations involving three successive one-dimensional slopes are given in Jameson, 1993; Cournède, Debiez and Dervieux, 1998.

## 4. Further Advanced Topics

The material presented in previous sections gives a brief overview of the derivation and analysis of finite volume methods. For simplicity and brevity of the presentation, exclusive attention has been devoted to scalar nonlinear conservation laws in divergence form. In this overview, special consideration has been given to the formulation and stability analysis of higher order accurate schemes since these developments have had the largest impact on development of industrial computer codes in use at the time of this writing. The remainder of this chapter will consider several extensions of the finite volume method. Section 4.1 considers a class of higher order accurate discretizations in time that still preserve the stability properties of the fully-discrete schemes using Euler time integration. This is followed by a discussion of generalizations of the finite volume method for problems including second order diffusion terms and the extension to systems of nonlinear conservation laws.

### *4.1. Higher order time integration schemes*

The derivation of finite volume schemes in Sect. 2 first yielded a semi-discrete formulation (21) that was later extended to a fully-discrete formulation (22) by the introduction of first order accurate forward Euler time integration. These latter schemes where then subsequently extended to higher order accuracy in space using a variety of techniques. For many computing problems of interest, first order accuracy in time is then no longer enough. To overcome this low order accuracy in time, a general class of higher order accurate time integration methods was developed that preserve stability properties of the fully-discrete scheme with forward Euler time integration. Following Gottlieb, Shu and Tadmor, 2001 and Shu, 2001, these methods will be referred to as *Strong Stability Preserving* (SSP) time integration methods.

Explicit SSP Runge-Kutta methods were originally developed by Shu, 1988; Shu and Osher, 1988 and Gottlieb and Shu, 1998 and called TVD Runge-Kutta time discretizations. In a slightly more general approach, total variation bounded (TVB) Runge-Kutta methods were considered by Cockburn and Shu, 1989; Cockburn, Lin and Shu, 1989; Cockburn, Hou and Shu, 1990; Cockburn and Shu, 1998 in combination with the discontinuous Galerkin discretization

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

in space. Küther, 2000 later gave error estimates for second order TVD Runge-Kutta finite volume approximations of hyperbolic conservation laws.

To present the general framework of SSP Runge-Kutta methods, consider writing the semi-discrete finite volume method in the following form

$$\frac{d}{dt}U(t) = L(U(t)) \tag{93}$$

where $U = U(t)$ denotes the solution vector of the semi-discrete finite volume method. Using this notation together with forward Euler time integration yields the fully-discrete form

$$U^{n+1} = U^n - \Delta t\ L(U^n)\ , \tag{94}$$

where $U^n$ is now an approximation of $U(t^n)$. As demonstrated in Section 2.2, the forward Euler time discretization is stable with respect to the $L^\infty$-norm, i.e.

$$||U^{n+1}||_\infty \leq ||U^n||_\infty\ , \tag{95}$$

subject to a CFL-like time step restriction

$$\Delta t \leq \Delta t_0\ . \tag{96}$$

With this assumption, a time integration method is said to be SSP (see Gottlieb, Shu and Tadmor, 2001) if it preserves the stability property (95), albeit with perhaps a slightly different restriction on the time step

$$\Delta t \leq c\ \Delta t_0 \tag{97}$$

where $c$ is called the CFL coefficient of the SSP method. In this framework, a general objective is to find SSP methods that are higher order accurate, have low computational cost and storage requirements, and have preferably a large CFL coefficient. Note that the TVB Runge-Kutta methods can be embedded into this class if the following relaxed notion of stability is assumed

$$||U^{n+1}||_\infty \leq (1 + \mathcal{O}(\Delta t))\ ||U^n||_\infty\ . \tag{98}$$

*4.1.1. Explicit SSP Runge-Kutta methods.* Following Shu and Osher, 1988 and the review articles by Gottlieb, Shu and Tadmor, 2001; Shu, 2001, a general $m$ stage Runge-Kutta method for integrating (93) in time can be algorithmically represented as

$$\begin{aligned} \tilde{U}^0 &:= U^n\ , \\ \tilde{U}^l &:= \sum_{k=0}^{l-1} \left(\alpha_{lk}\tilde{U}^k + \beta_{lk}\ \Delta t\ L(\tilde{U}^k)\right), \qquad \alpha_{lk} \geq 0, \qquad l = 1, ..., m\ , \\ U^{n+1} &:= \tilde{U}^m\ . \end{aligned} \tag{99}$$

To ensure consistency, the additional constraint $\sum_{k=0}^{l-1} \alpha_{lk} = 1$ is imposed. If, in addition, all $\beta_{lk}$ are assumed to be non-negative, it is straightforward to see that the method can be written as a convex (positive weighted) combination of simple forward Euler steps with $\Delta t$ replaced by $\frac{\beta_{lk}}{\alpha_{lk}}\Delta t$. From this property, Shu and Osher, 1988 concluded the following lemma:

**Lemma 4.1.** *If the forward Euler method (94) is $L^\infty$-stable subject to the CFL condition (96), then the Rung-Kutta method (99) with $\beta_{lk} \geq 0$ is SSP, i.e. the method is $L^\infty$-stable under the time step restriction (97) with CFL coefficient*

$$c = \min_{l,k} \frac{\beta_{lk}}{\alpha_{lk}}\ . \tag{100}$$

In the case of negative $\beta_{lk}$, a similar result can be proven, see Shu and Osher, 1988.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

*4.1.2. Optimal second and third order nonlinear SSP Runge-Kutta methods.* Gottlieb, Shu and Tadmor, 2001 [Proposition 3.1] show that the maximal CFL coefficient for any $m$-stage, $m$-th order accurate SSP Runge-Kutta methods is $c = 1$. Therefore, SSP Runge-Kutta methods that achieve $c = 1$ are termed "optimal". Note that this restriction is not true if the number of stages is higher than the order of accuracy, see Shu, 1988.

Optimal second and third order nonlinear SSP Runge-Kutta methods are given in Shu and Osher, 1988. The optimal second order, two-stage non-linear SSP Runge-Kutta method is given by

$$\begin{aligned} \tilde{U}^0 &:= U^n \ , \\ \tilde{U}^1 &:= \tilde{U}^0 + \Delta t \ L(\tilde{U}^0) \ , \\ U^{n+1} &:= \frac{1}{2}\tilde{U}^0 + \frac{1}{2}\tilde{U}^1 + \frac{1}{2}\Delta t \ L(\tilde{U}^1) \ . \end{aligned}$$

This method corresponds to the well known method of Heun. Similarly, the optimal third order, three-stage non-linear SSP Runge-Kutta method is given by

$$\begin{aligned} \tilde{U}^0 &:= U^n \ , \\ \tilde{U}^1 &:= \tilde{U}^0 + \Delta t \ L(\tilde{U}^0) \ , \\ \tilde{U}^2 &:= \frac{3}{4}\tilde{U}^0 + \frac{1}{4}\tilde{U}^1 + \frac{1}{4}\Delta t \ L(\tilde{U}^1) \ , \\ U^{n+1} &:= \frac{1}{3}\tilde{U}^0 + \frac{2}{3}\tilde{U}^2 + \frac{2}{3}\Delta t \ L(\tilde{U}^2) \ . \end{aligned}$$

Further methods addressing even higher order accuracy or lower storage requirements are given in the review articles of Gottlieb, Shu and Tadmor, 2001 and Shu, 2001 where SSP multi-step methods are also discussed.

### *4.2. Discretization of elliptic problems*

Finite volume methods for elliptic boundary value problems have been proposed and analyzed under a variety of different names: box methods, covolume methods, diamond cell methods, integral finite difference methods and finite volume element methods, see Bank and Rose, 1987; Cai, 1991; Süli, 1991; Lazarov, Michev and Vassilevsky, 1996; Viozat et al., 1998; Chatzipantelidis, 1999; Chou and Li, 2000; Hermeline, 2000; Eymard, Galluoët and Herbin, 2000; Ewing, Lin and Lin, 2002. These methods address the discretization of the following standard elliptic problem in a convex polygonal domain $\Omega \subset \mathbb{R}^2$

$$\begin{aligned} -\nabla \cdot A \nabla u &= f \ \text{ in } \Omega \ , \\ u(x) &= 0 \ \text{ on } \partial\Omega \end{aligned} \tag{101}$$

for $A \in \mathbb{R}^{2\times 2}$, a symmetric positive definite matrix (assumed constant). Provided $f \in H^\beta(\Omega)$ then a solution $u$ exists such that $u \in H_0^{\beta+2}(\Omega), -1 \leq \beta \leq 1, \beta \neq \pm 1/2$, where $H^s(\Omega)$ denotes the Sobolev space of order $s$ in $\Omega$.

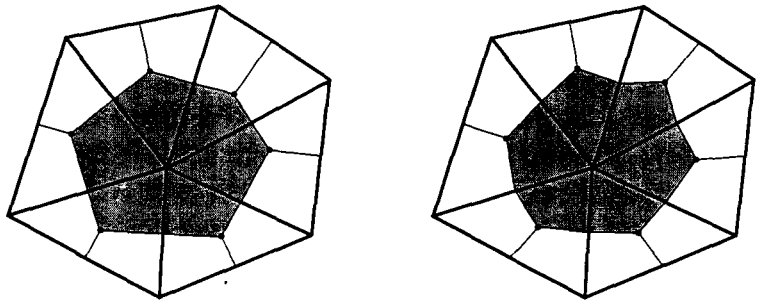
Nearly all the above mentioned methods can be recast in Petrov-Galerkin form using a piecewise constant test space together with a conforming trial space. A notable exception is given in Chatzipantelidis, 1999 wherein nonconforming Crouzeix-Raviart elements are utilized and analyzed. To formulate and analyze the Petrov-Galerkin representation, two tessellations

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

of $\Omega$ are considered: a triangulation $\mathcal{T}$ with simplicial elements $K \in \mathcal{T}$ and a dual tessellation $\mathcal{T}^*$ with control volumes $T \in \mathcal{T}^*$. In the class of conforming trial space methods such as the finite volume element (FVE) method, a globally continuous, piecewise $p$-th order polynomial trial space with zero trace value on the physical domain boundary is constructed

$$X_h = \{v \in C^0(\Omega) \,|\, v|_K \in \mathcal{P}_p(K),\ \forall K \in \mathcal{T} \text{ and } v|_{\partial\Omega} = 0\}$$

using nodal Lagrange elements on the simplicial mesh. A dual tessellation $\mathcal{T}^*$ of the Lagrange element is then constructed, see for example Fig. 10 which shows a linear Lagrange element with two dual tessellation possibilities. These dual tessellated regions form control volumes for the finite volume method. The tessellation technique extends to higher order Lagrange



a. Voronoi dual tessellation b. Median dual tessellation

Figure 10. Two control volume variants used in the finite volume discretization of second order derivative terms: (a) Voronoi dual where edges of the Voronoi dual are perpendicular to edges of the triangulation and (b) median dual formed from median dual segments in each triangle.

elements in a straightforward way. A piecewise constant test space is then constructed using $\mathcal{T}^*$

$$Y_h = \{v \,|\, v|_T \in \chi(T),\ \forall T \in \mathcal{T}^*\}$$

where $\chi(T)$ is a characteristic function in the control volume $T$. The finite volume element discretization of (101) then yields the following Petrov-Galerkin formulation: Find $u_h \in X_h$ such that

$$\sum_{\forall T \in \mathcal{T}^*} \left( \int_{\partial T} w_h\, A \nabla u_h \cdot d\nu + \int_T w_h\, f\, dx \right) = 0 \ , \quad \forall w_h \in Y_h \ . \tag{102}$$

The analysis of (102) by Ewing, Lin and Lin, 2002 using linear elements gives an *a priori* estimate in an $L^2$ norm that requires the least amount of solution regularity when compared to previous methods of analysis.

**Theorem 4.2 (FVE a priori error estimate, Ewing, Lin and Lin, 2002)** *Assume a 2-D quasi-uniform triangulation $\mathcal{T}$ with dual tessellation $\mathcal{T}^*$ such that $\exists\, C > 0$ satisfying*

$$C^{-1}h^2 \le |T| \le Ch^2,\ \forall T \in \mathcal{T}^* \ .$$

*Assume that $u$ and $u_h$ are solutions of (101) and (102) respectively with $u \in H^2(\Omega)$, $f \in H^\beta$, $(0 \le \beta \le 1)$. Then $\exists\, C' > 0$ such that the a priori estimate holds*

$$\|u - u_h\|_{L^2(\Omega)} \le C' \left( h^2 \|u\|_{H^2(\Omega)} + h^{1+\beta} \|f\|_{H^\beta(\Omega)} \right) \ . \tag{103}$$

*Encyclopedia of Computational Mechanics*. Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Unlike the finite element method, the error estimate (103) reveals that optimal order convergence is obtained only if $f \in H^\beta$ with $\beta \geq 1$. Moreover, numerical results show that the source term regularity can not be reduced without deteriorating the measured convergence rate. Optimal convergence rates are also shown for the nonconforming Crouzeix-Raviart element based finite volume method analyzed by Chatzipantelidis, 1999 for $u \in H^2(\Omega)$ and $f \in H^1(\Omega)$.

An extensive presentation and analysis of finite volume methods for elliptic equations without utilizing a Petrov-Galerkin formulation is given in Eymard, Galluoët and Herbin, 2000. In this work, general boundary conditions that include non-homogeneous Dirichlet, Neumann and Robin conditions are discussed. In addition, the analysis is extended to general elliptic problems in divergence form including convection, reaction and singular source terms.

### *4.3. Conservation laws including diffusion terms*

As demonstrated in Sect. 1, hyperbolic conservation laws are often approximations to physical problems with small or nearly vanishing viscosity. In other problems, the quantitative solution effects of these small viscosity terms are actually sought. Consequently, it is necessary in these problems to include viscosity terms into the conservation law formulation. As a model for these latter problems, a second order Laplacian term with small diffusion parameter is added to the first order Cauchy problem, i.e.

$$\partial_t u + \nabla \cdot f(u) - \varepsilon \Delta u = 0 \quad \text{in } \mathbb{R}^d \times \mathbb{R}^+, \tag{104a}$$

$$u(x,0) = u_0 \text{ in } \mathbb{R}^d \ . \tag{104b}$$

Here $u(x,t) : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$ denotes the dependent solution variable, $f \in C(\mathbb{R})$ the hyperbolic flux and $\varepsilon \geq 0$ a small diffusion coefficient. Application of the divergence and Gauss theorems to (104a) integrated in a region $T$ yields the following integral conservation law form

$$\frac{\partial}{\partial t} \int_T u \, dx + \int_{\partial T} f(u) \cdot d\nu - \int_{\partial T} \varepsilon \nabla u \cdot d\nu = 0 \ . \tag{105}$$

A first goal is to extend the fully-discrete form (22) of Sect. 2 to the integral conservation law (105) by the introduction of a numerical diffusion flux function $d_{jk}(u_h)$ for a control volume $T_j \in \mathcal{T}$ such that

$$\int_{\partial T_j} \varepsilon \nabla u \cdot d\nu \approx \sum_{e_{jk} \in \partial T_j} d_{jk}(u_h) \ .$$

When combined with the general finite volume formulation (22) for hyperbolic conservation laws, the following fully-discrete scheme is produced

$$u_j^{n+1} = u_j^n - \frac{\Delta t^n}{|T_j|} \sum_{e_{jk} \in \partial T_j} \left( g_{jk}(u_j^m, u_k^m) - d_{jk}(u_h^m) \right) \ , \quad \forall T_j \in \mathcal{T} \ . \tag{106}$$

In this equation, the index $m$ may be chosen either as $n$ or $n+1$, corresponding to an explicit or implicit discretization.

*4.3.1. Choices of the numerical diffusion flux $d_{jk}$.* The particular choice of the numerical diffusion flux function $d_{jk}$ depends on the type of control volume that is used. Since the approximate solution $u_h$ is assumed to be a piecewise constant function, the definition of $d_{jk}$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

involves a gradient reconstruction of $u_h$ in the normal direction to each cell interface $e_{jk}$. The reconstruction using piecewise constant gradients is relatively straightforward if the control volumes are vertex-centered, or if the cell interfaces are perpendicular to the straight lines connecting the storage locations (see Fig. 10).

*Vertex-centered finite volume schemes.* In the case of vertex-centered control volumes such as the median dual control volume, a globally continuous, piecewise linear approximate solution $\tilde{u}_h$ is first reconstructed on the primal mesh. $\nabla\tilde{u}_h$ is then continuous on the control volume interfaces and the numerical diffusion flux straightforwardly computed as

$$d_{jk}(u_h^m) \equiv \int_{e_{jk}} \nabla\tilde{u}_h^m \cdot d\nu_{jk} \ . \tag{107}$$

*Cell-centered finite volume schemes.* In the case of cell-centered finite volume schemes where an underlying primal-dual mesh relationship may not exist, a simple numerical diffusion flux can be constructed whenever cell interfaces are exactly or approximately perpendicular to the straight lines connecting the storage locations, e.g. Voronoi meshes, quadrilateral meshes, etc. In these cases, the reconstructed gradient of $u_h$ projected normal to the cell interface $e_{jk}$ can be represented by

$$\nabla u_h^m \cdot \nu_{jk} = \frac{u_k^m - u_j^m}{|x_k - x_j|}$$

where $x_i$ denotes the storage location of cell $T_i$. The numerical diffusion flux for this case is then given by

$$d_{jk}(u_h^m) \equiv \frac{|e_{jk}|}{|x_k - x_j|}\,(u_k^m - u_j^m) \ . \tag{108}$$

Further possible constructions and generalizations are given in Eymard, Gallouët and Herbin, 2001; Gallouët, Herbin and Vignal, 2000; Herbin and Ohlberger, 2002.

*4.3.2. Note on stability, convergence and error estimates.* Stability analysis reveals a CFL-like stability condition for the explicit scheme choice ($m = n$) in (106)

$$\Delta t^n \leq \frac{\alpha^3 (h_{min}^n)^2}{\alpha L_g h_{min}^n + \varepsilon}$$

where $L_g$ denotes the Lipschitz constant of the hyperbolic numerical flux, $\alpha$ is a positive mesh dependent parameter and $\epsilon$ is the diffusion coefficient. In constructing this bound, a certain form of shape regularity is assumed such that there exists an $\alpha > 0$ such that for all $j, k$ with $h_k \equiv \text{diam}\,(T_k)$

$$\alpha h_k^2 \leq |T_k|, \qquad \alpha\,|\partial T_k| \leq h_k, \qquad \alpha h_k \leq |x_k - x_l| \ . \tag{109}$$

Thus, $\Delta t^n$ is of the order $h^2$ for large $\varepsilon$ and of the order $h$ for $\varepsilon \leq h$. In cases where the diffusion coefficient is larger than the mesh size, it is therefore advisable to use an implicit scheme ($m = n + 1$). In this latter situation, no time step restriction has to be imposed (see Eymard et al., 2002; Ohlberger, 2001b).

In order to demonstrate the main difficulties when analyzing convection dominated problems, consider the following result from Feistauer et al., 1999 for a homogeneous diffusive boundary value problem. In this work, a mixed finite volume finite element method sharing similarities

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

with the methods described above is used to obtain a numerical approximation $u_h$ of the exact solution $u$. Using typical energy-based techniques, they prove the following *a priori* error bound.

**Theorem 4.3.** *For initial data* $u_0 \in L^\infty(\mathbb{R}^2) \cap W^{1,2}(\mathbb{R}^2)$ *and* $\tau > 0$ *there exist constants* $c_1, c_2 > 0$ *independent of* $\varepsilon$ *such that*

$$||u(\cdot, t^n) - u_h(\cdot, t^n)||_{L^2(\Omega)} \leq c_1 \, h \, e^{c_2 \tau/\epsilon} \ . \tag{110}$$

This estimate is fundamentally different from estimates for the purely hyperbolic problems of Sects. 2 and 3. Specifically, this result shows how the estimate strongly depends on the small parameter $\varepsilon$; ultimately becoming unbounded as $\varepsilon$ tends to zero.

In the context of convection dominated or degenerate parabolic equations, Kruzkov-techniques have been recently used by Carrillo, 1999; Karlsen and Risebro, 2000 in proving uniqueness and stability of solutions. Utilizing these techniques, convergence of finite volume schemes (uniform with respect to $\varepsilon \to 0$) was proven in Eymard et al., 2002 and *a priori* error estimates were obtained for viscuous approximations in Jakobsen and Karlsen, 2001 and Eymard, Gallouët and Herbin, 2002. Finally, in Ohlberger, 2001a; Ohlberger, 2001b uniform *a posteriori* error estimates suitable for adaptive meshing are given.

### *4.4. Extension to systems of nonlinear conservation laws*

A positive attribute of finite volume methods is the relative ease in which the numerical discretization schemes of Sects. 2 and 3 can be algorithmically extended to systems of nonlinear conservation laws of the form

$$\begin{aligned} \partial_t u + \nabla \cdot f(u) &= 0 \qquad \text{in } \mathbb{R}^d \times \mathbb{R}^+ \ , & \text{(111a)} \\ u(x,0) &= u_0(x) \quad \text{in } \mathbb{R}^d & \text{(111b)} \end{aligned}$$

where $u(x,t) : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}^m$ denotes the vector of dependent solution variables, $f(u) : \mathbb{R}^m \mapsto \mathbb{R}^{m \times d}$ denotes the flux vector, and $u_0(x) : \mathbb{R}^d \to \mathbb{R}^m$ denotes the initial data vector at time $t = 0$. It is assumed that this system is strictly hyperbolic, i.e. the eigenvalues of the flux jacobian $A(\nu; u) \boxminus \partial f / \partial u \cdot \nu$ are real and distinct for all bounded $\nu \in \mathbb{R}^d$.

The main task in extending finite volume methods to systems of nonlinear conservation laws is the construction of a suitable numerical flux function. To gain insight into this task, consider the one-dimensional *linear* Cauchy problem for $u(x,t) : \mathbb{R} \times \mathbb{R}^+ \mapsto R^m$ and $u_0(x) : R \mapsto \mathbb{R}^m$

$$\begin{aligned} \partial_t u + \partial_x (A\, u) &= 0 \qquad \text{in } \mathbb{R} \times \mathbb{R}^+ \ , \\ u(x,0) &= u_0(x) \quad \text{in } \mathbb{R} \end{aligned} \tag{112}$$

where $A \in \mathbb{R}^{m \times m}$ is a *constant* matrix. Assume the matrix $A$ has $m$ real and distinct eigenvalues, $\lambda_1 < \lambda_1 < \cdots < \lambda_m$, with corresponding right and left eigenvectors denoted by $r_k \in \mathbb{R}^m$ and $l_k \in \mathbb{R}^m$ respectively for $k = 1, \ldots, m$. Furthermore, let $X \in \mathbb{R}^{m \times m}$ denote the matrix of right eigenvectors, $X = [r_1, \ldots, r_m]$, and $\Lambda \in \mathbb{R}^{m \times m}$ the diagonal matrix of eigenvalues, $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_m)$ so that $A = X \Lambda X^{-1}$. The one-dimensional system (112) is readily decoupled into scalar equations via the transformation into characteristic variables $\alpha = X^{-1} u$

$$\partial_t \alpha + \partial_x (\Lambda\, \alpha) = 0 \qquad \text{in } \mathbb{R} \times \mathbb{R}^+ \ ,$$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

$$\alpha(x,0) \;=\; \alpha_0(x) \text{ in } \mathbb{R} \tag{113}$$

and component-wise solved exactly

$$\alpha^{(k)}(x,t) = \alpha_0^{(k)}(x-\lambda_k\, t) \;\;, \quad k=1,\ldots,m$$

or recombined in terms of the original variables

$$u(x,t) = \sum_{k=1}^{m} l_k \cdot u_0(x-\lambda_k\, t)\, r_k \;\;.$$

Using this solution, it is straightforward to solve exactly the associated Riemann problem for $w(\xi,\tau) \in R^m$

$$\partial_\tau w + \partial_\xi (A\, w) = 0 \quad \text{ in } \mathbb{R}\times\mathbb{R}^+$$

with initial data

$$w(\xi,0) = \begin{cases} u & \text{if } \xi < 0 \\ v & \text{if } \xi > 0 \end{cases}$$

thereby producing the following Godunov-like numerical flux function

$$\begin{aligned} g(u,v) \;&=\; Aw(0,\mathbb{R}^+) \\ &=\; \frac{1}{2}(Au+Av) - \frac{1}{2}|A|\,(v-u) \end{aligned} \tag{114}$$

with $|A| \equiv X|\Lambda|X^{-1}$. When used in one-dimensional discretization together with piecewise constant solution representation, the linear numerical flux (114) produces the well-known Courant-Isaacson-Rees (CIR) upwind scheme for linear systems of hyperbolic equations

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

where $A^\pm = X\Lambda^\pm X^{-1}$. Note that higher order accurate finite volume methods with slope limiting procedures formally extend to this linear system via component wise slope limiting of the characteristic components $\alpha^{(k)}, k=1,\ldots m$ for use in the numerical flux (114).

*4.4.1. Numerical flux functions for systems of conservation laws.* In Godunov's original work (see Godunov, 1959), exact solutions of the one-dimensional *nonlinear* Riemann problem of gas dynamics were used in the construction of a similar numerical flux function

$$g^{\mathrm{G}}(u,v) = f(w(0,\mathbb{R}^+))\cdot\nu \tag{115}$$

where $w(\xi,\tau) \in \mathbb{R}^m$ is now a solution of a nonlinear Riemann problem

$$\partial_\tau w + \partial_\xi f^{(\nu)}(w) = 0 \quad \text{ in } \mathbb{R}\times\mathbb{R}^+$$

with initial data

$$w(\xi,0) = \begin{cases} u & \text{if } \xi < 0 \\ v & \text{if } \xi > 0 \end{cases} \;.$$

Recall that solutions of the Riemann problem for gas dynamic systems are a composition of shock, contact and rarefaction wave family solutions. For the gas dynamic equations considered by Godunov, a unique solution of the Riemann problem exists for general states $u$ and $v$

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

except those states producing a vacuum. Even so, the solution of the Riemann problem is both mathematically and computationally nontrivial. Consequently, a number of alternative numerical fluxes have been proposed that are more computationally efficient. These alternative numerical fluxes can be sometimes interpreted as approximate Riemann solvers. A partial list of alternative numerical fluxes is given here. A more detailed treatment of this subject is given in Godlewski and Raviart, 1991, Kröner, 1997, and LeVeque, 2002.

- Osher-Solomon flux (Osher and Solomon, 1982). This numerical flux is a system generalization of the Enquist-Osher flux of Sect. 2. All wave families are approximated in state space as rarefaction or inverted rarefaction waves with Lipschitz continuous partial derivatives. The Osher-Solomon numerical flux is of the form

$$g^{\mathrm{OS}}(u,v) = \frac{1}{2}(f(u)+f(v))\cdot\nu - \frac{1}{2}\int_u^v |A(\nu;w)|\,dw$$

  where $|A|$ denotes the usual matrix absolute value. By integrating on $m$ rarefaction wave integral subpaths that are each parallel to a right eigenvector, a system decoupling occurs on each subpath integration. Furthermore, for the gas dynamic equations with ideal gas law, it is straightforward to construct $m-1$ Riemann invariants on each subpath thereby eliminating the need for path integration altogether. This reduces the numerical flux calculation to purely algebraic computations with special care taken at sonic points, see Osher and Solomon, 1982.
- Roe flux (Roe, 1981). Roe's numerical flux can be a interpreted as approximating all waves families as discontinuities. The numerical flux is of the form

$$g^{\mathrm{Roe}}(u,v) = \frac{1}{2}(f(u)+f(v))\cdot\nu - \frac{1}{2}|A(\nu;u,v)|\,(v-u)$$

  where $A(\nu;u,v)$ is the "Roe matrix" satisfying the matrix mean value identity

$$(f(v)-f(u))\cdot\nu = A(\nu;u,v)\,(v-u)$$

  with $A(\nu;u,u) = A(\nu;u)$. For the equations of gas dynamics with ideal gas law, the Roe matrix takes a particularly simple form. Steady discrete mesh-aligned shock profiles are resolved with one intermediate point. The Roe flux does not preclude the formation of entropy violating expansion shocks unless additional steps are taken near sonic points.
- Steger-Warming flux vector splitting (Steger and Warming, 1981). Steger and Warming considered a splitting of the flux vector for the gas dynamic equations with ideal gas law that exploited the fact that the flux vector is homogeneous of degree one in the conserved variables. From this homogeneity property, Euler's identity then yields that $f(u)\cdot\nu = A(\nu;u)\,u$. Steger and Warming then considered the matrix splitting

$$A = A^+ + A^- \ , \quad A^\pm \equiv X\Lambda^\pm X^{-1}$$

  where $\Lambda^\pm$ is computed component wise. From this matrix splitting, the final upwind numerical flux function was constructed as

$$g^{\mathrm{SW}}(u,v) = A^+(\nu;u)\,u + A^-(\nu;v)\,v \ .$$

  Although not part of their explicit construction, for the gas dynamic equations with ideal gas law, the jacobian matrix $\partial g^{\mathrm{SW}}/\partial u$ has eigenvalues that are all nonnegative and the

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

jacobian matrix $\partial g^{\mathrm{SW}}/\partial v$ has eigenvalues that are all nonpositive whenever the ratio of specific heats $\gamma$ lies in the interval $[1, 5/3]$. The matrix splitting leads to numerical fluxes that do not vary smoothly near sonic and stagnation points. Use of the Steger-Warming flux splitting in the schemes of Sect. 2 and 3 results in rather poor resolution of linearly degenerate contact waves and velocity slip surfaces due to the introduction of excessive artificial diffusion for these wave families.

- Van Leer flux vector splitting. Van Leer, 1982 provided an alternative flux splitting for the gas dynamic equations that produces a numerical flux of the form

$$g^{\mathrm{VL}}(u, v) = f^{-}(u) + f^{+}(v)$$

using special Mach number polynomials to construct fluxes that remain smooth near sonic and stagnation points. As part of the splitting construction, the jacobian matrix $\partial g^{\mathrm{SW}}/\partial u$ has eigenvalues that are all nonnegative and the matrix $\partial g^{\mathrm{SW}}/\partial v$ has eigenvalues that are all nonpositive. The resulting expressions for the flux splitting are somewhat simpler when compared to the Steger-Warming splitting. The van Leer splitting also introduces excessive diffusion in the resolution of linearly degenerate contact waves and velocity slip surfaces.

- System Lax-Friedrichs flux. This numerical flux is the system equation counterpart of the scalar Lax-Friedrichs flux (27). For systems of conservation laws the Lax-Friedrichs flux is given by

$$g^{\mathrm{LF}}(u, v) = \frac{1}{2}(f(u) + f(v)) \cdot \nu - \frac{1}{2}\alpha(\nu)\,(v - u)$$

where $\alpha(\nu)$ is given through the eigenvalues $\lambda_k(\nu; w)$ of $A(\nu; w)$

$$\alpha(\nu) = \max_{1 \le k \le m} \sup_{w \in [u,v]} |\lambda_k(\nu; w)| \ .$$

The system Lax-Friedrichs flux is usually not applied on the boundary of domains since it generally requires an over specification of boundary data. The system Lax-Friedrichs flux introduces a relatively large amount of artificial diffusion when used in the schemes of Sect. 2. Consequently, this numerical flux is typically only used together with relatively high order reconstruction schemes where the detrimental effects of excessive artificial diffusion are mitigated.

- Harten-Lax-van Leer flux (Harten, Lax and van Leer, 1983). The Harten-Lax-van Leer numerical flux originates from a simplified two wave model of more general $m$ wave systems such that waves associated with the smallest and largest characteristic speeds of the $m$ wave system are always accurately represented in the two wave model. The following numerical flux results from this simplified two wave model

$$g^{\mathrm{HLL}}(u, v) = \frac{1}{2}(f(u) + f(v)) \cdot \nu - \frac{1}{2}\frac{\alpha_{\max} + \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}}(f(v) - f(u)) \cdot \nu + \frac{\alpha_{\max}\alpha_{\min}}{\alpha_{\max} - \alpha_{\min}}\,(v - u)$$

where

$$\alpha_{\max}(\nu) = \max_{1 \le k \le m}\,(0, \sup_{w \in [u,v]} \lambda_k(\nu; w)) \ , \quad \alpha_{\min} = \min_{1 \le k \le m}\,(0, \inf_{w \in [u,v]} \lambda_k(\nu; w)) \ .$$

When compared to the Lax-Friedrichs flux, this flux can be considerably more accurate in flow situations where $0 < |(\alpha_{\max} + \alpha_{\min})/(\alpha_{\max} - \alpha_{\min})| < 1$. Using this flux, full upwinding is obtained for supersonic flow. Modifications of this flux are suggested in Einfeldt et al., 1998 to improve the resolution of intermediate waves as well.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Further examples of numerical fluxes (among others) include the kinetic flux vector splitting due to Deshpande, 1986, the advection upstream splitting flux (AUSM) of Liou and Steffen, 1993, and the convective upwind and split pressure (CUSP) flux of Jameson, 1993 and Tatsumi et al., 1994.

## 5. Concluding Remarks

The literature associated with the foundation and analysis of the finite volume methods is extensive. This article gives a very brief overview of finite volume methods with particular emphasis on theoretical results that have significantly impacted the design of finite volume methods in everyday use at the time of this writing. More extensive presentations and references on various topics in this article can be found in the books by Godlewski and Raviart, 1991, Kröner, 1997, Eymard, Galluoët and Herbin, 2000 and LeVeque, 2002.

## REFERENCES

Abgrall R. On essentially non-oscillatory schemes on unstructured meshes: analysis and implementation. *J. Comp. Phys.* 1994; **114**:45–58.

Angermann L, Knabner P and Thiele K. An error estimate for a finite volume discretization of density driven flow in porous media. *Appl. Numer. Math.* 1998; **26**:179–191.

Bank R and Rose DJ. Some error estimates for the box method. *SIAM J. Numer. Anal.* 1987; **24**:777–787.

Barth TJ and Jespersen DC. The design and application of upwind schemes on unstructured meshes. *American Institite for Aeronautics and Astronautics* 1989; Report 89-0366:1–12.

Barth TJ and Frederickson PO. Higher order solution of the Euler equations on unstructured grids using quadratic reconstruction. *American Institite for Aeronautics and Astronautics* 1990; Report AIAA-90-0013.

Batten P, Lambert, C and Causon DM. Positively Conservative high-resolution convection schemes for unstructured elements. *Int. J. Numer. Meth. Engrg* 1996; **39**:1821–1838.

Billey V, Périaux J, Perrier P and Stoufflet B. 2-D and 3-D Euler computations with finite element methods in aerodynamics. *Lecture Notes in Mathematics* (Vol. 1270), Springer-Verlag: Berlin, 1987.

Boris JP and Book DL. Flux corrected transport: SHASTA, a fluid transport algorithm that works. *J. Comp. Phys.* 1973; **11**:38–69.

Bouchut F and Perthame B. Kruzkov's estimates for scalar conservation laws revisited. *Trans. Am. Math. Soc.* 1998; **350**(7):2847–2870.

Carrillo J. Entropy solutions for nonlinear degenerate problems. *Arch. Ration. Mech. Anal.* 1999; **147**:269–361.

Cai Z. On the finite volume element method. *Numer. Math.* 1991; **58**:713–735.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Chatzipantelidis P. A finite volume method based on the Crouzeix-Raviart element for elliptic problems. *Numer. Math.* 1999; **82**:409–432.

Chou SH and Li Q. Error estimates in $L^2$, $H^1$ and $L^\infty$ in covolume methods for elliptic and parabolic problems: a unified approach. *Math. Comp.* 2000; **69**:103–120.

Cockburn B, Coquel F and Lefloch PG. An error estimate for finite volume methods for multidimensional conservation laws. *Math. Comput.* 1994; **63**:77–103.

Cockburn B and Gau H. A posteriori error estimates for general numerical methods for scalar conservation laws. *Comput. Appl. Math.* 1995; **14**:37–47.

Cockburn B, Hou S and Shu CW. The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. *Math. Comp.* 1990; **54**(190):545–581.

Cockburn B, Lin SY and Shu CW. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems. *J. Comput. Phys.* 1989; **84**(1):90–113.

Cockburn B and Shu CW. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. *Math. Comp.* 1989; **52**:411–435.

Cockburn B and Shu CW. The Runge-Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems. *J. Comput. Phys.* 1998; **141**(2):199–224.

Colella P and Woodward P. The piecewise parabolic methods for gas-dynamical simulations, *J. Comp. Phys.* 1984; **54**:174–201.

Chainais-Hillairet C. Finite volume schemes for a nonlinear hyperbolic equation: convergence towards the entropy solution and error estimates. *M2AN Math. Model. Numer. Anal.* 1999; **33**:129–156.

Chainais-Hillairet C. Second-order finite-volume schemes for a non-linear hyperbolic equation: error estimates. *Math. Methods Appl. Sci.* 2000; **23**(5):467–490.

Cournède P-H and Debiez C and Dervieux A. A positive MUSCL scheme for triangulations. Institut National De Recherche En Informatique Et En Automatique (INRIA), Report 3465, 1998.

DiPerna RJ. Measure-valued solutions to conservation laws. *Arch. Rational Mech. Anal.* 1985; **88**(3):223–270.

Delanaye M. Polynomial Reconstruction Finite Volume Schemes for the Compressible Euler and Navier-Stokes Equations on Unstructured Adaptive Grids. *Ph.D. Thesis* 1996, University of Liége, Belgium.

Deshpande, SM. On the Maxwellian distribution, symmetric form, and entropy conservation for the Euler equations. *NASA Langley, Hampton, Virginia* 1986; NASA Report TP-2583.

Desideri JA and Dervieux A. Compressible flow solvers using unstructured grids. *Von Karman Institute Lecture Notes* 1988-05; Von Karman Institute for Fluid Dynamics, Belgium.

Einfeldt B, Munz C, Roe P and Sjögreen B. On Godunov-type methods near low densities. *J. Comput. Phys.* 1992; **92**:272–295.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Ewing RE, Lin T and Lin Y. On the accuracy of the finite volume element method bases on piecwise linear polynomials. *SIAM J. Numer. Anal.* 2002; **39**(6):1865–1888.

Eymard R, Gallouët T, Ghilani M, and Herbin R. Error estimates for the approximate solution of a nonlinear hyperbolic equation given by finite volume schemes. *IMA Journal of Numerical Analysis* 1998; **18**:563–594.

Eymard R, Galluoët T and Herbin R. Finite volume methods. *Handbook of Numerical Analysis*, North Holland: Amsterdam, 2000; **7**:713–1020.

Eymard R, Gallouët T and Herbin R. Finite volume approximation of elliptic problems and convergence of an approximate gradient. *Appl. Numer. Math.* 2001; **37**(1-2):31–53.

Eymard R, Gallouët T and Herbin R. Error estimates for approximate solutions of a nonlinear convection-diffusion problem. *Adv. Differential Equations* 2002; **7**(4):419-440.

Eymard R, Gallouët T, Herbin R. and Michel A. Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.* 2002; **92**(1):41–82.

Feistauer M, Felcman J, Lukáčová-Medvid'ová M, and Warnecke G. Error estimates for a combined finite volume-finite element method for nonlinear convection-diffusion problems. *SIAM J. Numer. Anal.* 1999; **36**(5):1528–1548.

Gallouët T, Herbin R and Vignal MH. Error estimates on the approximate finite volume solution of convection diffusion equations with general boundary conditions. *SIAM J. Numer. Anal.* 2000; **37**(6):1935–1972.

Godunov SK. A finite difference method for the numerical computation of discontinuous solutions of the equations of fluid dynamics. *Mat. Sb.* 1959; **47**:271–290.

Godlewski E and Raviart P-A. Hyperbolic systems of conservation laws. *Mathematiques & Applications* Ellipses, Paris 1991.

Goodman JD and LeVeque RJ. On the accuracy of stable schemes for 2D conservation laws. *Math. Comp.* 1985; **45**(171):15–21.

Gottlieb S and Shu CW. Total variation diminishing Runge-Kutta schemes. *Math. Comput.* 1998; **67**(221):73–85.

Gottlieb S, Shu CW and Tadmor E. Strong stability-preserving high-order time discretization methods. *SIAM Rev.* 2001; **43**(1):89–112.

Harten A, Hyman JM, and Lax PD. On finite-difference approximations and entropy conditions for shocks. *Comm. Pure and Appl. Math.* 1976; **24**:297–322.

Harten A. High resolution schemes for hyperbolic conservation laws. *J. Comp. Phys.* 1983; **49**:357–393.

Harten A, Lax PD and van Leer, B. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Rev.* 1983; **25**:35–61.

Harten A, Osher S, Engquist B and Chakravarthy S. Some results on uniformly high order accurate essentially non-oscillatory schemes. *Appl. Num. Math.* 1986; **2**:347–377.

Harten A, Osher S, Engquist B and Chakravarthy S. Uniformly high-order accurate essentially nonoscillatory schemes III. *J. Comp. Phys.* 1987; **71**(2):231–303.

Harten A. ENO schemes with subcell resolution. *J. Comp. Phys.* 1989; **83**:148–184.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Harten A and Chakravarthy S. Multi-dimensional ENO schemes for general geometries. *Institite for Computer Applications in Science and Engineering* 1991; Report ICASE-91-76.

Herbin R and Ohlberger M. A posteriori error estimate for finite volume approximations of convection diffusion problems. In proceedings:*Finite volumes for complex applications - problems and perspectives, Porquerolles,* 753–760; Hermes Science Publications, Paris, 2002.

Hermeline F. A finite volume method for the approximation of diffusion operators on distorted meshes. *J. Comput. Phys.* 2000; **160**(2):481–499.

Jaffre J, Johnson C and Szepessy A. Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws. *Math. Models and Methods in Appl. Sci.* 1995; **5**(3):367–386.

Jakobsen ER and Karlsen KH. Continuous dependence estimates for viscosity solutions of fully nonlinear degenerate parabolic equations. *J. Differential Equations* 2002; **183**(2):497–525.

Jameson A and Lax PD. Conditions for the construction of multipoint variation diminishing difference schemes, *Appl. Numer. Math.* 1986; **2**(3-5):335-345.

Jameson A and Lax PD. Corrigendum: Conditions for the construction of multipoint variation diminishing difference schemes, *Appl. Numer. Math.* 1987; **3**(3):289.

Jameson A. Artificial Diffusion, Upwind biasing, limiters and their effect on accuracy and convergence in transonic and hypersonic flows. *American Institite for Aeronautics and Astronautics* 1993; Report AIAA-93-3359:1–28.

Jiang G and Shu CW. Efficient implementation of weighted ENO schemes. *J. Comp. Phys.* 1996; **126**:202–228.

Johnson C and Szepessy A. Adaptive finite element methods for conservation laws based on a posteriori error estimates. *Commun. Pure Appl. Math.* 1995; **48**:199–234.

Karlsen KH and Risebro NH. On the uniqueness and stability of entropy solutions of nonlinear degenerate parabolic equations with rough coefficients. Preprint 143, Department of Mathematics, University of Bergen, 2000.

Koren, B. Upwind schemes for the Navier-Stokes equations. *Proceedings of the Second International Conference on Hyperbolic Problems.* Vieweg: Braunschweig, 1988.

Kröner D. *Numerical Schemes for Conservation Laws.* Wiley–Teubner: Stuttgart, 1997.

Kröner D, Noelle S and Rokyta M. Convergence of higher order upwind finite volume schemes on unstructured grids for conservation laws in several space dimensions. *Numer. Math.* 1995; **71**:527–560.

Kröner D and Ohlberger M. A-posteriori error estimates for upwind finite volume schemes for nonlinear conservation laws in multidimensions. *Math. Comput.* 2000; **69**:25–39.

Kruzkov SN. First order quasilinear equations in several independent variables. *Math. USSR Sbornik* 1970; **10**:217–243.

Küther M. Error estimates for second order finite volume schemes using a TVD-Runge-Kutta time discretization for a nonlinear scalar hyperbolic conservation law. *East-West J. Numer. Math.* 2000; **8**(4):299–322.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes.

Kuznetsov NN. Accuracy of some approximate methods for computing the weak solutions of a first-order quasi-linear equation. *USSR, Comput. Math. and Math. Phys.* 1976; **16**(6):159–193.

Lax PD. *Hyperbolic Systems of Conservation Laws.* SIAM Pub.:Philadelphia, 1973.

Lax PD and Wendroff B. Systems of conservation laws. *Comm. Pure Appl. Math.* 1960; **13**:217–237.

Lazarov RD, Michev ID and Vassilevsky PS. Finite volume methods for convection-diffusion problems. *SIAM J. Numer. Anal.* 1996; **33**:31–35.

LeVeque R. High resolution finite volume methods on arbitrary grids via wave propagation. *J. Comp. Phys.* 1988; **78**:36–83.

LeVeque R. *Finite Volume Methods for Hyperbolic Problems.* Cambridge University Press: Cambridge, 2002.

Liou MS and Steffen CJ. A new flux-splitting scheme. *J. Comp. Phys.* 1993; **107**:23–39.

Liu X-D. A maximum principle satisfying modification of triangle based adaptive stencils for the solution of scalar hyperbolic conservation laws. *SIAM J. Numer. Anal.* 1993; **30**:701–716.

Málek J, Nečas J, Rokyta M, and Růžička M. Weak and measure-valued solutions to evolutionary PDEs. *Applied Mathematics and Mathematical Computation* (Vol 13). Chapman and Hall: London, 1968.

Ohlberger M. A posteriori error estimates for finite volume approximations to singularly perturbed nonlinear convection-diffusion equations. *Numer. Math.* 2001a; **87**(4):737–761.

Ohlberger M. A posteriori error estimates for vertex centered finite volume approximations of convection-diffusion-reaction equations. *M2AN Math. Model. Numer. Anal.* 2001b; **35**(2):355–387.

Oleinik OA. Discontinuous solutions of non-linear differential equations. *Amer. Math. Soc. Transl. (2)* 1963; **26**:95–172.

Osher S and Solomon F. Upwind Difference Schemes for Hyperbolic Systems of Conservation Laws. *Math. Comp.* 1982; **38**(158):339–374.

Osher S. Riemann solvers, the entropy condition, and difference approximations. *SIAM J. Numer. Anal.* 1984; **21**(2):217–235.

Osher S. Convergence of generalized MUSCL schemes. *SIAM J. Numer. Anal.* 1985; **22**(5):947–961.

Peterson T. A Note on the convergence of the discontinuous Galerkin method for a scalar hyperbolic equation. *SIAM J. Numer. Anal.* 1991; **28**(1):133–140.

Roe PL. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comp. Phys.* 1981; **43**:357–372.

Rostand P and Stoufflet B. TVD schemes to compute compressible viscous flows on unstructured meshes. *Proceedings of the Second International Conference on Hyperbolic Problems.* Vieweg: Braunschweig, 1988.

*Encyclopedia of Computational Mechanics.* Edited by Erwin Stein, Rene de Borst and Thomas J.R. Hughes.