High order methods for CFD

Remi Abgrall, Mario Ricchiuto

High order methods for CFD

R. Abgrall and M. Ricchiuto

September 16, 2016

Abstract

We provide a review of high order methods for CFD. Besides recalling some classical methods, we show a framework allowing on one hand to see and work with these methods under a different light, and on the other to provide a different path to construct numerical methods for flow equations. In particular, we focus on Residual Based techniques, and Residual Distribution methods, as a framework to construct schemes of arbitrary order. The somewhat classical second-order multidimensional upwind fluctuation-splitting/residual-distribution schemes are reviewed in the chapter by Deconinck and Ricchiuto.

Contents

1 Introduction

In this chapter we discuss with some extend high order methods for advection dominated problems. Typical examples are the advection diffusion equation (with large Peclet number), the Euler equations, the Navier Stokes equations, the Shallow water equations and many problems in geophysical flow, to mention just a few. The list is immense.

In this kind of problems, one has to deal with contransting constraints. First the solution must be accurate. Wherever the solution is smooth, the truncation error must scale as h^{r+1} where h is the typical size of the mesh elements, and r is an integer. However, it is also well known that in many case, the solution is not globally smooth, and that it may admit local very large gradients. For example, these may be shock waves (or slip lines) for the Euler equations, and boundary layers for the Navier-Stokes equations at very high Reynolds number. Other effects may come into play, as e.g. dispersion, as in some geophysical and environnemental applications (wave propagation, capillary flows).

In this chapter, we address these issues and give several examples of successful modern high order methods. The literature on this topic has exploded since the mid 90's, and it is not possible to give an exhaustive view of all what has been achieved, so we had to make choices, which, of course, are biased by our own work.

We will start by reviewing to some extent possibly the two most popular methods today: WENO finite volume schemes, and the Discontinuous Galerkin (DG) methods. We give some details, in particular indicate the main principles. However, we do not discuss all the issues related to these methods. In particular, set aside the choice of the approximation of the viscous terms: this can easily be found in the many papers that have appeared on the topics or in monographs such as [1, 2]. Our main focus is on a less successful approach known as the *residual distribution method*. As also discussed in the companion chapter by Deconinck and Ricchiuto, these schemes share a lot with continuous Finite Element Methods, such as the Streamline diffusion method, but also embed properties typical of the finite volume method: a lot of emphasis is put on L^∞ stability constraints, allowingto avoid spurious oscillations at discontinuities. In this contribution, however, we extend this analysis, showing how the Residual-Based philosophy underlying these schemes provides a framework which allows to embed most (or all) other arbitrary order methods, and work with them under a different light, thus providing mode insight in these methods, and perhaps new, alternative contructions. Of course it also provides a setting to construct different arbitrary order schemes, and we will review the main challenges encountered when doing that, as well as some of the solutions proposed so far to overcome these challenges.

This chapter is organized into 4 sections. The first one gives a review of the WENO and DG methods. The second section develop in details the Residual distribution method, both from an historical perspective and its most recent achievements. The third section provides several applications, both for compressible flows and aerodynamics, and for some geophysical flows. A conclusion follows. We hope that the bibliography is rich enough to cover and complete all the topics we have mentioned in the text.

2 A review of existing methods

We start by considering the following problem:

$$\frac{\partial \mathbf{w}}{\partial t} + \text{ div } \mathbf{f}(\mathbf{w}) - \text{div } \mathbf{f}_v(\mathbf{w}, \nabla \mathbf{w}) = 0 \tag{1a}$$

defined in $\Omega \subset \mathbb{R}^d$ with $d = 1, 2, 3$, and $\mathbf{w} : \Omega \times \mathbb{R}^+ \to \mathcal{D} \subset \mathbb{R}^m$. We need to set up an initial condition

$$\mathbf{w}(\mathbf{x}, 0) = u_0(\mathbf{x}), \qquad \mathbf{x} \in \Omega \tag{1b}$$

and boundary conditions on $\partial\Omega$. In (1a), the flux $\mathbf{f} = (\mathbf{f}_1, \ldots, \mathbf{f}_d)$ is assumed to be defined on $\mathcal{D}$ and to be smooth enough. The viscous flux $\mathbf{f}_v$ is assumed to satisfy similar hypothesis (however see below).

Here, we are mostly interested in fluid mechanics problems where the Navier Stokes equation is the canonical example. In that case, the state variable $\mathbf{w}$ needs to satisfy additional constraints: the density and the internal energy need to be positive; this is what is expressed in saying that $\mathbf{w} \in \mathcal{D}$.

In this section, we review some of the high order methods that have been developed in the recent years. The research activity in this field has been very intense over the last years, so that it is impossible to make an exhaustive review of what has been done. We had to make choices, this choices are biased.

The second choice we have made, at least for this section, is to deal with a simplified problem. Instead of (1a), we will deal with

$$\frac{\partial \mathbf{w}}{\partial t} + \text{ div } \mathbf{f}(\mathbf{w}) = 0 \tag{2}$$

with the initial condition (1b) and relevant boundary conditions. In the case of fluid mechanics, the simplest system of this kind is the Euler system. There, the state variable is

$$\mathbf{w} = (\rho, \rho\mathbf{u}, E)^T$$

where ρ is the density, $\mathbf{u}$ is the fluid velocity, and E is the total energy, i.e. the sum of the internal energy e and the quinetic energy. The flux is given by:

$$\mathbf{f} = (\rho\mathbf{u}, \rho\mathbf{u} \otimes \mathbf{u} + p \text{ Id}_d, (E+p)\mathbf{u})^T.$$

To close the system, we need to define the pressure $p = p(\rho, e)$. General assumptions on this function can be found in [3], a typical example is that of perfect gaz:

$$p = (\gamma - 1)e.$$

It is well known that only weak solutions of (2) can be considered because there is no hope to have regular solutions in general. Hence, we need to consider weak solutions, i.e. measurable functions in $L^\infty(\Omega \times \mathbb{R}^+)^m \cap L^1(\Omega \times \mathbb{R}^+)^m$ such that for any compactly supported regular test function $\varphi \in C_0^1(\Omega \times \mathbb{R}^+)^m$, we have

$$\int_{\Omega\times\mathbb{R}^+} \frac{\partial \varphi}{\partial t}(\mathbf{x},t) \cdot \mathbf{w}(\mathbf{x},t) d\mathbf{x}dt + \int_{\Omega\times\mathbb{R}^+} \nabla\varphi(\mathbf{x},t) \cdot \mathbf{f}(\mathbf{w}) d\mathbf{x}dt - \int_\Omega \varphi(\mathbf{x},0) \cdot \mathbf{w}_0(\mathbf{x}) d\mathbf{x} = 0. \tag{3}$$

Of course the initial condition needs to be also in $L^\infty(\Omega)^m \cap L^1(\Omega)^m$. Note we have not taken into account the boundary conditions. This is a complex problem (for a rigorous treatment and also from a practical point of view), we refer to [4] for systems.

It is also well known that the definition of weak solution is not enough. Even in the scalar case, this does not guaranty the uniqueness of solution and some selection mechanism is needed. For the system case, the solution is way more complex. To go towards this, one classical consider an entropy, i.e. a strictly convex function S defined on $\mathcal{D}$ such that there exists an entropy flux $\mathbf{G}$ such that:

$$\nabla_{\mathbf{u}}\mathbf{G} = \nabla_{\mathbf{w}} S \cdot \nabla_{\mathbf{w}} \mathbf{f}.$$

An entropy solution should satisfy, in the sens of distribution, the following inequality:

$$\frac{\partial S}{\partial t} + \text{ div } \mathbf{G} \leq 0. \tag{4}$$

In the case of the Euler equations, the (mathematical) entropy is given by $S = -\rho s$ where s is the (physical) entropy. The entropy flux is $\mathbf{G} = S\mathbf{u}$. The reader may consult [5, 6] for further considerations.

The form (3) is the origin of all possible forms of numerical approximation of the system (1). The first thing to do is to approximate the domain Ω. For the sake of simplicity, we assume here that Ω is polygonal. Then we discretize Ω using meshes. With the vocabulary of unstructured meshes, we consider a tessaletion $\mathcal{T}_h$. The domain Ω is

$$\Omega = \cup_{K\in\mathcal{T}_h} K.$$

As usual, we assume that the elements K are non overlapping. The elements K will be triangles or quadrangles in 2D, tetrahedrons, hexahedrons, pyramid, etc in 3D, or may have more complicated forms. All depends on the choices made for approximating the solution and the choices of test functions φ.

- Finite volume schemes: one considers that $\mathbf{w}$ is constant in each cell,

  $$\mathbf{w}_K(t) \approx \frac{1}{|K|} \int_K \mathbf{w}(\mathbf{x},t) d\mathbf{x},$$

  and the test functions are also constant. On can nevertheless get high order accuracy of the averaged value. This is the topic of section 2.1.1.

- Continuous finite elements. Here we assume a globally continuous approximation $\mathbf{w}_h$ of $\mathbf{w}$. Typicaly, for any element, $\mathbf{w}_{h|K}$ is a polynomial of degree k. Because of the continuity requirement, this imposes constraints on the mesh: the intersection of two elements is either empty, or reduced to a (complete) face or they are identical. The mesh is said to be conformal. The elements need also, in general, be simplices because of the polynomial approximation. These methods are sketched in 2.1.3.
- Discontinuous Galerkin methods. Here the continuity requirement is dropped. This enables a lot of freedom: the mesh needs not be conformal, the element can be general, so that mesh refinement becomes simple in principle. These methods are sketched in 2.1.2.

In the rest of this section, we first consider the spatial approximation (hence using a semi-discrete formulation), and then we discuss a bit the temporal approximation.

2.1 Space discretizations

2.1.1 ENO and WENO

Here, we consider the finite volume formulation of (3). The states are described by $\{\mathbf{w}_K(t)\}_{K\in\mathcal{T}_h}$. Starting from (3), we get

$$\frac{d}{dt}\int_K \mathbf{w}(\mathbf{x},t)d\mathbf{x} + \int_{\partial K} \mathbf{f}(\mathbf{w})\cdot\mathbf{n}d\partial K = 0$$

Here $\mathbf{n}$ is the outward unit normal to the boundary ∂K or K. This can be obtained from (3) by first regularizing via mollification φ_ε the characteristic function of K, and taking the limit when $\varepsilon \to 0$, we see that $\nabla\varphi_\varepsilon \to \mathbf{n}$.

Since K is polygonal, denoting by e a generic face/edge of K, we see that an approximation of (3) is

$$\frac{d}{dt}\mathbf{w}_K(t) + \frac{1}{|K|}\sum_{e\in\partial K}\int_e \mathbf{f}(\mathbf{w}_h)\cdot\mathbf{n}de = 0. \tag{5}$$

This relation has not yet a meaning since $\mathbf{u}_h$ is discontinuous accross edges. In the normal direction to e, we need to solve the following Riemann problem:

$$\frac{\partial\mathbf{w}}{\partial t} + \frac{\partial\mathbf{f}(\mathbf{w})\cdot\mathbf{n}}{\partial\mathbf{n}} = 0$$

with the initial condition:

$$\mathbf{w}(\mathbf{x},0) = \begin{cases} \mathbf{w}_K & \text{if } \mathbf{x}\cdot\mathbf{n} < 0 \\ \mathbf{w}_{K^+} & \text{else.} \end{cases}$$

Here, $\mathbf{u}_{K^+}$ is the state on the other side of e. This problem is solved either exactly or in an approximated way. A meaning of the edge integral is given thanks to the use of numerical flux $\hat{\mathbf{f}}(\mathbf{u}_K, \mathbf{u}_{K^+})$, see [3, 7, 8] for an extensive discussion about numerical flux and Riemann solvers. Hence, the finite volume method in its simplest form is:

$$\frac{d}{dt}\mathbf{w}_K(t) + \frac{1}{|K|}\sum_{e\in\partial K}\int_e \hat{\mathbf{f}}(\mathbf{w}_{h|K}, \mathbf{w}_{h|K^+}, \mathbf{n})de = 0. \tag{6}$$

Note that the edges integrals are evaluated via quadrature formula.

As this, only first order accuracy can be achieved. Formal high order accuracy can be obtained by using the MUSCL method due to van Leer [9]. The idea is to consider a polynomial reconstruction of degree p, $\mathcal{R}(u_h)$, within each cell K and to replace (6) by

$$\frac{d}{dt}\mathbf{w}_K(t) + \frac{1}{|K|}\sum_{e\in\partial K}\int_e \hat{\mathbf{f}}(\mathcal{R}(\mathbf{w}_h)_K, \mathcal{R}(\mathbf{w}_h)_{K^+}, \mathbf{n})de = 0. \tag{7}$$

and the quadrature formula need to be of sufficient order, typicaly exact for polynomials of degree p.

The design of a reconstruction operator is a research field by itself since one wants to avoid the Gibbs phenomena where the solution becomes steep or discontinuous. After the seminal work of van Leer, a very large literature has been devoted to this problem. A large part of it is about Total Variation Diminishing schemes (see Sweby's paper [10]), but it is rather difficult to reach higher than second order accuracy (see [11] for an attempt in 1D), and it can be shown that a TVD scheme in more than one dimension,

even in the scalar case, is at most first order accurate, see [12]. This negative result has motivated to look for criteria that are less strict than the TVD one, and the most successfull method is probably the Essentialy Non Oscillatory method, orginally due to Harten and co-workers [13, 14], then refined by Shu and co-workers [15, 16]. Extension to unstructured meshes can be found in [17]. Better stability properties are obtained by the so-called Weighted Essentially Non Oscillatory technique (WENO), see [18] for the original reference, and further refined by Shu and co-workers, see [19] for a review.

The principle can be explained assuming a regular mesh, in one dimension. Extension to 2D, and more general meshes can be found in [20, 21, 22] for example. Taking a mesh $\{x_j\}_{j\in\mathbf{z}}$, with $x_j = j\Delta x$, we first define four approximation of a smooth function at $x_{i+1/2} = \frac{x_i+x_{i+1}}{2}$ by:

- using the stencil $\mathcal{S}^{(1)} = \{x_{i-2}, x_{i-1}, x_i\}$, we have $u^{(1)}_{i+1/2} = \frac{3}{8}u_{i-2} - \frac{5}{4}u_{i-1} + \frac{15}{8}u_i = u(x_{i+1/2}) + O(\Delta x^3)$
- Using the stencil With $\mathcal{S}^{(2)} = \{x_{i-1}, x_i, x_{i+1}\}$, we get: $u^{(2)} = -\frac{1}{8}u_{i-1} + \frac{3}{4}u_i + \frac{3}{8}u_{i+1} = u(x_{i+1/2}) + O(\Delta x^3)$
- With the stencil, With $\mathcal{S}^{(3)} = \{x_i, x_{i+1}, x_{i+2}\}$, we have $u^{(3)}(x_{i+1/2}) = \frac{3}{8}u_i + \frac{3}{4}u_{i+1} - \frac{1}{8}u_{i+2} + O(\Delta x^3)$
- Last, with $\mathcal{S} = \{x_{i-2}, x_{i-1}, x_i, x_{i+1}, x_{i+2}\}$, we obtain

$$u^{(4)}_{i+1/2} = \frac{3}{128}u_{i-2} - \frac{5}{32}u_{i-1} + \frac{45}{64}u_i + \frac{15}{32}u_{i+1} - \frac{5}{128}u_{i+2} = u(x_{i+1/2}) + O(\Delta x^5).$$

Then we notice that $u^{(4)}_{i+1/2} = \gamma_1 u^{(1)}_{i+1/2} + \gamma_2 u^{(2)}_{i+1/2} + \gamma_3 u^{(3)}_{i+1/2}$ with $\gamma_1 = \frac{1}{16}$, $\gamma_2 = \frac{5}{8}$ and $\gamma_3 = \frac{5}{16}$. Note that $\gamma_1 + \gamma_2 + \gamma_3 = 1$. This enables to approximate $u(x_{i+1/2})$ by $u_{i+1/2} = w_1 u^{(1)}_{i+1/2} + w_2 u^{(2)}_{i+1/2} + w_3 u^{(3)}_{i+1/2}$ with $w_j \approx \gamma_j$ and $w_j \approx 0$ if a discontinuity exists in $\mathcal{S}^{(j)}$. In order to achieve this, we define $w_j = \frac{\tilde{w}_j}{\tilde{w}_1+\tilde{w}_2+\tilde{w}_3}$ with $\tilde{w}_j = \frac{\gamma_j}{(\varepsilon+\beta_j)^2}$ and the smoothner indicater β_j is:

$$\beta_j = \sum_{l=1}^{2} \Delta x^{2l-1} \int_{x_{i-1/2}}^{x_{i+1/2}} \left(\frac{d^l}{dx^l} p_j(x)\right)^2 dx.$$

Compared to the method we are going to discuss now, for a given formal accuracy they clearly need the lowest possible storage. The price to pay for this is that the computational stencil is quite large. In the case of an irregular mesh, many precautions need to be taken in order to effectively reach the formal accuracy. In the case of unstructured mesh, their extension is possible but very technical.

2.1.2 Discontinuous Galerkin methods.

Formulation and basic properties. This class of method has original been designed by W.H. Reed and T.R. Hill [23], the first analysis was done by Lesaint and Raviart [24] and further refined by [25]. The references [26] and mostly [27, 28] and their sequel paved the way to the success of DG methods for hyperbolic problems and the Navier Stokes equations (among many other applications). The reference [29] represents the state-of-the-art in the early 2000, the reference [2] is a more mathematical presentation of the theory, it also contains a lot of informations on how to approximate parabolic problems in that framework. It is impossible to give a complete survey of this field because the number of papers grows exponentially. Again, we will sketch the method for purely hyperbolic problems, and refer to the reference section of [2] for more informations on how to approximate the second order terms.

Again, we start from (3). We consider a tessellation of the computational domain Ω, like in the finite volume method, but here we look for solutions that are polynomial of degree $r \geq 0$ in each element. More precisely, we want to approximate the solution in V_h defined by:

$$V_h = \{\mathbf{w}_h \in (L^\infty(\Omega))^m \cap (L^1(\Omega))^m, \text{for any element } K, (\mathbf{w}_h)_{|K} \in (\mathbb{P}^r(K))^m\}$$

No continuity requirement is needed, and more over the degree r may depend on the element. Then we apply the weak formulation, taking as test function any $\varphi \in V_h$: for any element K,

$$\int_K \frac{\partial \varphi}{\partial t} \cdot \mathbf{w}_h(\mathbf{x}, t) d\mathbf{x} - \int_K \nabla\varphi \cdot \mathbf{f}(\mathbf{w}_h(\mathbf{x}, t)) d\mathbf{x} + \int_{\partial K} \varphi \cdot (\mathbf{f}(\mathbf{w}_h(\mathbf{x}, t)) \cdot \mathbf{n}) d\partial K = 0.$$

However, as for the Finite Volume method, this formulation is meaningless because $\mathbf{u}_h$ appearing in the boundary integral is in general multivalued, so the flux term cannot be given a meaning. The idea contained in [27, 28] is, as for the finite volume method, to introduce a numerical flux $\widehat{\mathbf{f}}$. The DG (semi discrete) formulation is thus: find $\mathbf{w}_h \in V_h$ such that for any K and any $\varphi_h \in V_h$,

$$\int_K \frac{\partial \varphi}{\partial t} \cdot \mathbf{w}_h(\mathbf{x},t) d\mathbf{x} - \int_K \nabla\varphi \cdot \mathbf{f}(\mathbf{w}_h(\mathbf{x},t)) d\mathbf{x} + \int_{\partial K} \varphi \cdot \widehat{\mathbf{f}}(\mathbf{w}_{h|K}, \mathbf{w}_{h|K^+}, \mathbf{n}) d\partial K = 0 \tag{8}$$

In any element K, $\mathbf{w}_h \in \mathbb{P}^r(K)$. This vector space is spanned by a finite set of polynomial functions:

$$\mathbf{w}_h = \sum_{j=0}^{R} \mathbf{w}^{(j)} \varphi_j,$$

so that we arrive at the following form of the semi-discrete scheme: for any K,

$$M_K \frac{d}{dt} \mathbf{W}_K + F(\mathbf{w}_{h|K}) = 0$$

where the mass matrix is

$$(M_K)_{ij} = \int_K \varphi_i \varphi_j d\mathbf{x}$$

is clearly invertible. This is a block diagonal matrix, hence its inversion (needed for time discretization) is rather straightforward. The vector F is defined by its components:

$$F_j = -\int_K \nabla\varphi_j \cdot \mathbf{f}(\mathbf{w}_h(\mathbf{x},t)) d\mathbf{x} + \int_{\partial K} \varphi_j \cdot \widehat{\mathbf{f}}(\mathbf{w}_{h|K}, \mathbf{w}_{h|K^+}, \mathbf{n}) d\partial K.$$

The choice of the degree of freedom, i.e. the choice of the basis function, is an issue by itself. The choices are made depending whether to favor a geometrical interpretation (Lagrange basis), to facilitate the change of polynomial degree within elements (in the case of degree adaptivity), or if the element shape is or not completely general (for example in the case of Lagrangian hydrodynamics, [30]), etc.

Non linear stability. One can show, in the scalar case, that a global entropy inequality can be easily derived, see [31]. In the scalar case, a natural entropy is $U(u) = \frac{u^2}{2}$: this is a convex function, and an entropy $g = (g_x, g_y)$ flux satisfies

$$u f'_x = g'_x, \qquad u\ f'_y = g'_y.$$

We see that $g_x = u f_x - \int^u f_x du$, $g_y = u f_y - \int^u f_y du$. In the following, we set $h_x = \int^u f_x du$ and $h_y = \int^u f_y du$. We wish to establish an inequality of the type: for any K,

$$\int_K \frac{\partial U(u_h)}{\partial t} dx + \int_{\partial K} \hat{g} \cdot \mathbf{n} dx \leq 0. \tag{9}$$

Here $\hat{g} \cdot \mathbf{n}$ is an entropy flux, i.e. a numerical flux consistant with g. This inequality simply states that we have a local L^2 energy bound.

For any v_h,

$$\int_K \frac{\partial u_h}{\partial t} v_h d\mathbf{x} - \int_K f(u_h) \nabla v_h x d\mathbf{x} + \int_{\partial K} v_h \hat{f}(u_{h|K}, u_{h|K^-}, \mathbf{n}) d\partial K = 0.$$

Then we choose $v_h = u_h$, so that

$$\frac{1}{2}\int_K \frac{\partial (u_h)^2}{\partial t} dx - \int_K f(u_h) \nabla u_h dx + \int_{\partial K} u_h \hat{f}(u_{h|K}, u_{h|K^-}, \mathbf{n}) d\partial K = 0 \tag{10}$$

We get

$$\frac{1}{2}\int_K \frac{\partial (u_h)^2}{\partial t} dx + \int_{\partial K} \hat{G}(u_{h|K}, u_{h|K^-}, \mathbf{n}) d\partial K + A_K = 0$$

with

$$\hat{G}_{j+1/2} = \frac{u_{h|K} + u_{h|K^-}}{2} \hat{\mathbf{f}}(u_{h|K}, u_{h|K^-}, \mathbf{n}) - \frac{1}{2}\big(g(u_{h|K}) + g(u_{h|K^-})\big) \cdot \mathbf{n}$$

This flux is consistant with $g \cdot \mathbf{n}$ and we also have set

$$A_K = \int_{[u_{h|K}, u_{h|K^-}]} \Big(\hat{\mathbf{f}}(u_{h|K}, u_{h|K^-}, \mathbf{n}) - \mathbf{f}(v) \cdot \mathbf{n} \Big) dv.$$

Using the mean value theorem, we see that

$$A_K = (u_{h|K} - u_{h|K^-}) \Big(\hat{\mathbf{f}}(u_{h|K}, u_{h|K^-}, \mathbf{n}) - \mathbf{f}(\xi) \cdot \mathbf{n} \Big)$$

for a suitable ξ between $u_{h|K}$ and $u_{h|K^-}$. If $\hat{\mathbf{f}}$ is an E-Scheme (see [32]), i.e. if for any ξ between u and v, $\big(\hat{\mathbf{f}}(u, v, \mathbf{n}) - \mathbf{f}(\xi) \cdot \mathbf{n}\big)\ (u - v) \leq 0$, we see that (9) holds true for any E-scheme. Typical example are the exact Godunov solver, the Rusanov scheme:

$$\hat{\mathbf{f}}(u, v, \mathbf{n}) = \frac{1}{2} \Big(\mathbf{f}(u) \cdot \mathbf{n} + \mathbf{f}(v) \cdot \mathbf{n} \Big) + \alpha(u - v)$$

for $\alpha \geq \max_{\xi \in [\min(u,v), \max(u,v)]} \big| \mathbf{f}(\xi) \cdot \mathbf{n} \big|$.

The case of systems is of course more complex. One possible solution could be to rewrite the system (1) in term of entropy variables: $\mathbf{v} = \nabla_{\mathbf{w}} S(\mathbf{w})$ where S is a (mathematical) entropy, in the following form:

$$\mathbf{w} \frac{\partial \mathbf{w}}{\partial t} + \text{ div } \mathbf{h}(\mathbf{v}) = 0$$

The change of variables $\mathbf{v} \mapsto \mathbf{w}$ is one-to-one and does not affect the Rankine-Hugoniot relations. Instead of approximating the state variable $\mathbf{w}$ by piecewise polynomials, we can approximate the entropy with polynomials of degree $r + 1$ and define the approximation state as

$$V_h = \{ \mathbf{v} \in (L^1(\Omega))^m \cap (L^\infty(\Omega))^m, \mathbf{v} \in \mathbb{P}^{r+1}(K) \}$$

and look for $\mathbf{w}(\mathbf{v})$, with $\mathbf{v} \in V_h$ such that for any $\varphi \in V_h$, for any K

$$\int_K \varphi \frac{\partial \mathbf{w}(\mathbf{v})}{\partial t} - \int_K \nabla \varphi \cdot \mathbf{h}(\mathbf{v}) d\mathbf{x} + \int_{\partial K} \varphi \widehat{\mathbf{h}}(\mathbf{v}_K, \mathbf{v}_{K^-}, \mathbf{n}) d\partial K = 0$$

Clearly, if $\widehat{\mathbf{h}}$ is an E-scheme, we have the entropy inequality:

$$\int_K \frac{\partial S(\mathbf{v})}{\partial t} + \int_{\partial K} \widehat{\mathbf{G}}(\mathbf{v}_K, \mathbf{v}_{K^-}, \mathbf{n}) d\partial K \leq 0$$

for a suitable consistant entropy flux $\widehat{\mathbf{G}}$.

Controlling spurious oscillations. Another and somewhat related issue is to control the Gibbs phenomena: when the numerical solution develops steep gradients, either because the mesh resolution is not sufficient or because discontinuities are appearing, spurious oscillations will appear. One of the fundamental question is how to control them as automatically as possible. One of the solution is to get inspired by what has been done for finite volume schemes, taking into account the negative result of Goodman and LeVeque [12]. In order to describe what has been achieved, let us turn back to the 1D case for the simplification of exposure. In that case, the scheme reduces to:

$$\int_{K_j} \varphi \frac{du_h}{dt} - \int_{K_j} \varphi' f(u_h) dx + \varphi(x_{j+1/2}) \hat{f}(u^-_{h,j+1/2}, u^+_{h,j+1/2}) - \varphi(x_{j-1/2}) \hat{f}(u^+_{h,j-1/2}, u^-_{h,j-1/2}) = 0 \quad (11)$$

where $K_j = (x_{j-1/2}, x_{j+1/2})$, $x_j = \frac{x_{j-1/2}, x_{j+1/2}}{2}$, $\Delta_j = x_{j+1/2} - x_{j-1/2}$ and $u^\pm_{h,j+1/2} = u_h(x_j \pm \frac{\delta_j u}{2}) = u_h(x_j) \pm \delta^\pm_j u$.

One of the main remark that enables to understand the behavior of methods is what is called Harten's lemma. Instead of considering the semi-discrete case, let us use full discretized form, we will come back to this in section 2.2 Assume we have a sequence $\{u^n_j\}_{j \in \mathbf{z}, n \in \mathbb{N}}$ that satisfies ($\lambda > 0$):

$$u^{n+1}_j = u^n_j - \lambda \Big(C_{j+1/2}(u^n_{j+1} - u^n_j) - D_{j-1/2}(u^n_j - u^n_{j-1}) \Big). \quad (12)$$

If for any $j \in \mathbb{Z}$, we have

$$\begin{array}{l} C_{j+1/2} \geq 0, D_{j+1/2} \geq 0, \\ \lambda\big(C_{j+1/2} + D_{j+1/2}\big) \leq 1 \end{array} \tag{13}$$

then the sequence satisfies a L^∞, a L^1 and a TVD bound, where the total variation of $u = \{u_j\}_{j\in\mathbf{z}}$ is:

$$TV(u) = \sum_{j\in\mathbf{z}} |u_{j+1} - u_j|.$$

There is no reason why the arguments $u_j^\pm$ are such that the sequence defined by (12) are such that the

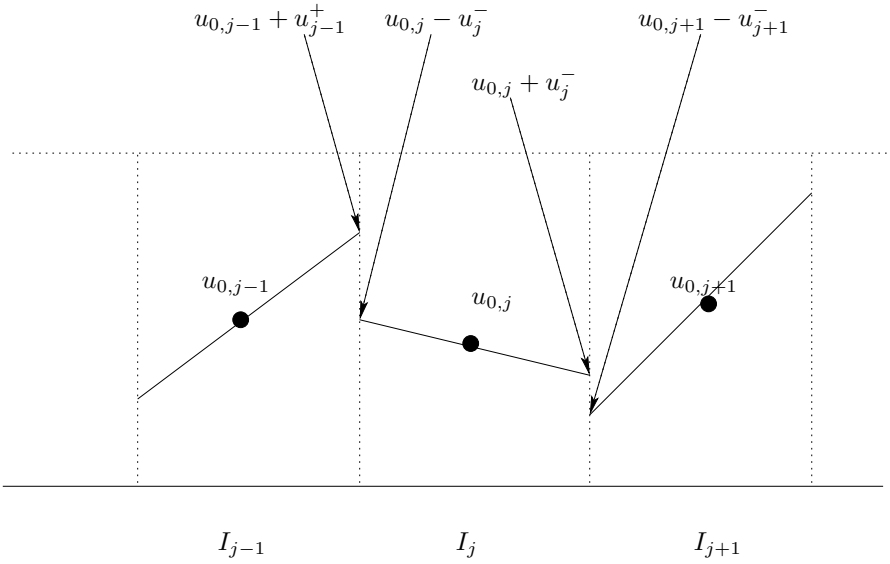


Figure 1: Geometrical representation of a $\mathbb{P}^1$ approximation. In the element K_j, $u_h = u_{0,j} + \delta u_j(x - x_j)$. x_j is the centroid of K_j, δ_j is its lengh. We have set $\delta u_j^\pm = u_h(x_j \pm \frac{\Delta_j}{2}) - u_{0,j}$.

conditions (13) are true. To do so, one technique is to introduce a limiter. The simplest is the generalized minmod limiter:

$$m(a_1, a_2, \ldots, a_m) = \left\{ \begin{array}{ll} s \min(|a_1|, |a_2|, \ldots, |a_m|) & \text{if } s = \text{sign}(a_1) = \ldots \text{sign}(a_m) \\ 0 & \text{else} \end{array} \right. \tag{14}$$

The arguments in the flux $\hat{f}$ (11) are modified as follows. We replace $\delta_j^\pm$ by

$$\big(\delta_j^\pm\big)^{mod} = m(\delta_j^\pm, \Delta_+ u_{0,j}, \Delta_- u_{0,j}).$$

There is a huge litterature on this topic. One may quote among others: [33, 34, 35, 36, 37, 38]. Other kind of polynomial representation can also be used, such as Hermite approximation, see [39, 40] and their relation to limiting. Another approach is to combine WENO limiters and Discontinuous Galerkin method, see [41, 22].

2.1.3 Stabilized Continous FEM

Another approach for spatial approximation is to consider the following trial space:

$$V_h = \{\mathbf{w}_h \in (L^\infty(\Omega))^m \cap (L^1(\Omega))^m \cap (C^0(\Omega))^m, \text{for any element } K, (\mathbf{w}_h)_{|K} \in (\mathbb{P}^r(K))^m\}.$$

The fundamental difference is that we now require continuity. If one use elements of V_h as test function, i.e. look for solutions $\mathbf{w}_h$ such that: for any $\varphi \in V_h$,

$$\int_\Omega \varphi \cdot \frac{\partial \mathbf{w}_h}{\partial t} d\mathbf{x} + a(\mathbf{w}_h, \varphi) = 0 \tag{15a}$$

where

$$a(\mathbf{w}, v) = -\int_\Omega \nabla v \cdot \mathbf{f}(\mathbf{w}_h) d\mathbf{x} + BC(\mathbf{w}, \varphi). \tag{15b}$$

The problem amounts to solving (with clear notations)

$$M \frac{d\mathbf{w}_h}{dt} + A(\mathbf{w}_h) = 0, \qquad M_{ij} = \int_\Omega \varphi_i \varphi_j d\mathbf{x}.$$

The mass matrix M is also invertible. It is a sparse matrix but it is not block diagonal, contrarily to the Discontinuous Galerkin method. In (15b), the operator BC describes the approximation of the weakly enforced boundary conditions. We do not describe it because it depends on the nature of the boundary conditions: it is problem dependent.

The method (15) is known have stability difficulties, so it is better to add to the Galerkin variational form a stabilization operator. There are several forms of this stabilization operator: the stream line operator [6, 25] and a jump operator [42]. Instead of solving (15), we solve:

$$\int_\Omega \varphi \cdot \frac{\partial \mathbf{w}_h}{\partial t} + a(\mathbf{w}_h, \varphi) + a_S(\mathbf{w}_h, \varphi) = 0 \tag{16a}$$

where a_S is a stabilisation operator.

In the case of the streamline operator, the choice is [6]:

$$a_S(\mathbf{w}_h, \varphi) = \sum_K h_K \int_K \big(\nabla_{\mathbf{w}} \mathbf{f}(\mathbf{w}_h) \cdot \nabla \varphi) \mathcal{T}_K \big(\frac{\partial \mathbf{w}_h}{\partial t} + \nabla_u \mathbf{f}(\mathbf{w}_h) \cdot \nabla \mathbf{w}_h\big) d\mathbf{x}\Big) = 0, \tag{16b}$$

where h_K represents the diameter of K and $\mathcal{T}_K \geq 0$ is a stabilisation parameter (or matrix).

In the case of the jump operator, we take [42]

$$a_S(\mathbf{w}_h, \varphi) = \sum_{\text{internal edges}} \gamma_e h_e^2 \int_e \big|\big|\nabla_{\mathbf{w}} \mathbf{f}(\mathbf{w}_h)\big|\big| [\nabla \mathbf{w}_h][\nabla \varphi_h] \tag{16c}$$

where $\gamma_e \geq 0$, h_e is the measure of the edge e. The choice of the stabilisation operator is done so that if the exact solution also satisfies (16). Note that the structure of the mass matrix is affected in the case (16b), hence its invertibility is less obvious. In the case of (16c), the mass matrix is not changed, but the compactness of the computational stencil is slightly affected. Since these methods share a lot of similarities with the residual distribution methods (indeed, they can be seen as a particular case), we postpone the discussion later in the text.

2.2 Temporal discretizations

There are several standard ways of approximating in time, depending on how we look at time with respect to space. Either they are two unrelated parameters, so that one first approximate in space and then in time thanks to the method of lines. Or we consider the equation

$$\frac{\partial \mathbf{w}}{\partial t} + \text{ div } \mathbf{f}(\mathbf{w}) = 0$$

as a space-time divergence applied to the flux $(\mathbf{w}, \mathbf{f}(\mathbf{w}))$. Here, we focus on the explicit method of lines.

A typical example is the well knows method of line. After having discretized in space, we have to discretize a problem of the form:

$$\frac{\partial \mathbf{w}}{\partial t} = L(\mathbf{w}). \tag{17}$$

Depending on the stiffness of the problem, more general speaking, of the properties we are looking for, one may consider explicit or implicit scheme. A very popular method is the so called Strong Stability

Preserving technique [43]: If the Euler operator $\mathbf{w} \mapsto \mathbf{v} = \mathbf{w} - \Delta t L(\mathbf{w})$ preserves the L^∞ norm or the L^1 norm or the TVD semi norm for $\Delta t \leq \Delta t_0$, then the SSP Runge-Kutta method it is built on will also have the same property under a condition of the type $\Delta t \leq C\Delta t_0$. In the cases we are interested in, Δt_0 is defined by mean of a CFL-type condition, and is approximation dependent.

The explicit SSP RK schemes are written as Runge Kutta schemes

$$\begin{array}{l} u^{(0)} = u^n \\ u^{(i)} = \sum_{k=0}^{i-1} \left(\alpha_{i,k} u^{(k)} + \Delta t \beta_{i,k} L(u^{(k)})\right), \qquad i = 1, \ldots, m \\ u^{n+1} = u^{(m)} \end{array}$$

where the $\alpha_{i,k}$ and $\beta_{i,k}$ are all *positive*. By consistency, $\sum_{k=1}^{i-1} \alpha_{i,k} = 1$, s that the intermediate stages can be written as convex combination of the Euler operator. The integer m is the number of stages. Examples of such SSP RK methods are the following:

- Second order in time and $C = 1$ (no degradation of the time step)

$$\begin{array}{l} u^{(1)} = u^n + \Delta t L(u^n) \\ u^{n+1} = \frac{1}{2} u^n + \frac{1}{2}\left(u^{(1)} + \Delta t L(u^{(1)})\right) \end{array}$$

- Third order in time and $C = 1$.

$$\begin{array}{l} u^{(1)} = u^n + \Delta t L(u^n) \\ u^{(2)} = \frac{3}{4} u^n + \frac{1}{4}\left(u^{(1)} + \Delta t L(u^{(1)})\right) \\ u^{n+1} = \frac{1}{3} u^n + \frac{2}{3}\left(u^{(2)} + \Delta t L(u^{(2)})\right) \end{array}$$

In these examples, the number of stages is equal to the order of the scheme. It can be shown [44] that there exists no 4th order SSP RK scheme with four stages. Spiteri and Ruuth [45]developed fourth order methods with $m = 5, 6, 7$ and 8 stages: for example

$$\begin{array}{rcl} u^{(1)} & = & u_n + 0.391752226571890\ \Delta t L(u^n), \\ u^{(2)} & = & 0.444370493651235\ u^n + 0.555629506348765\ u^{(1)} + 0.368410593050371\ \Delta t L(^{(1)}), \\ u^{(3)} & = & 0.620101851488403\ u^n + 0.379898148511597\ u^{(2)} + 0.251891774271694\ \Delta t L(^{(2)}), \\ u^{(4)} & = & 0.178079954393132\ u^n + 0.821920045606868\ u^{(3)} + 0.544974750228521\ \Delta t L(^{(3)}), \\ u^{n+1} & = & 0.517231671970585\ u^{(2)} + 0.096059710526147\ u^{(3)} + 0.063692468666290\ \Delta t L(^{(3)}) \\ & & +0.386708617503269\ u^{(4)} + 0.226007483236906\ \Delta t L(^{(4)}) \end{array}$$

for which $C = 1.508$.

A rather complete discussion can be found in [43, 46, 45]. Error estimates for explicit Runge Kutta time stepping can be found in [47, 48, 49, 50].

3 A different setting : residual distribution

We now consider the framework known today as *residual distribution*. Its roots can be found the seminal work of P.L. Roe [51, 52] on *fluctuation splitting*, and in all the contributions of the 90s on wave decomposition, hyperbolic elliptic splitting, and multidimensional upwind methods [53, 54, 55, 56, 57, 58, 59, 60], and see also [61].

We present it here as a general framework to study and unify the "more classical" approaches recalled in section 2, while giving some additional flexibility to construct new "non classical" discretizations.

3.1 Steady hyperbolic problems

Consider the scalar steady state advection equation

$$\vec{a} \cdot \nabla u = 0 \quad \text{on} \quad \Omega \subset \mathbb{R}^2 \tag{18}$$

where $\nabla \cdot \vec{a} = 0$, and with boundary conditions

$$\int_{\partial\Omega} (\vec{a} \cdot \hat{n})^- (g - u) = \int_{\partial\Omega^-} \vec{a} \cdot \hat{n} (g - u) = 0 \tag{19}$$

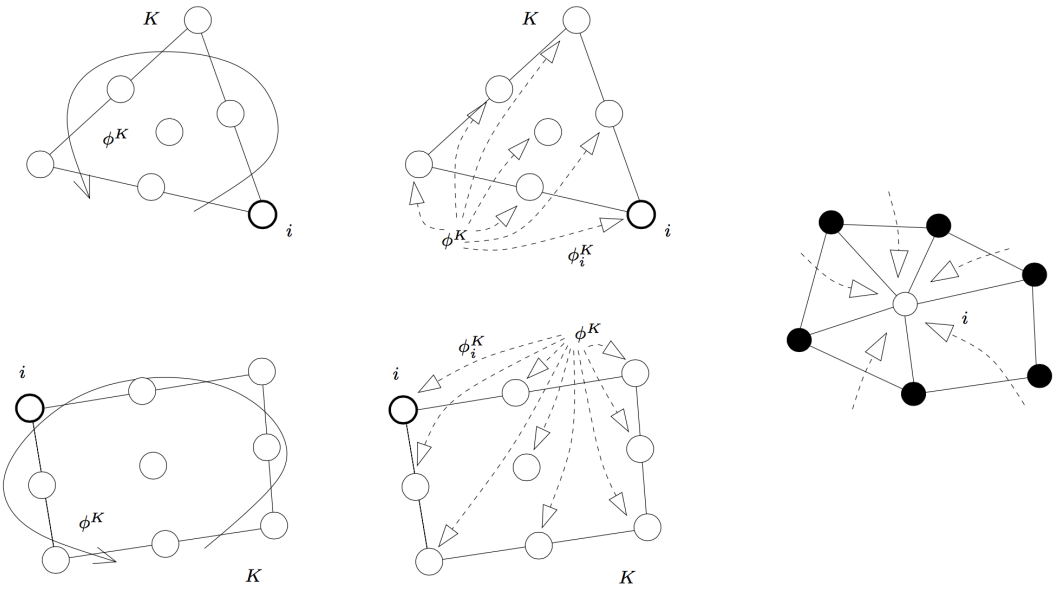


Figure 2: Residual distribution

To find a numerical approximation of the solution of (18)-(19), on a tessellation of the spatial domain Ω_h, we use a generalization of the fluctuation splitting strategy put forward by P.L. Roe. In particular, we start by considering u_h, a continuous nodal finite element approximation of the solution

$$u_h = \sum_{i \in \Omega_h} \varphi_i u_i = \sum_{K \in \Omega_h} u_i \varphi_i \Big|_K \tag{20}$$

For a given degree of freedom i of the continuous collocated finite element expansion, let K_i be the set of elements sharing i as a node, and similarly, let F_i be the set of mesh faces sharing i. Given an initial guess for the degrees of freedom, we proceed as follows

1. For all elements K, compute the *fluctuation/residual*

$$\phi^K = \int_K \vec{a} \cdot \nabla u_h|_K \, d\mathbf{x} \left(\approx - \int_K \partial_t u_h \, d\mathbf{x} \right) \tag{21}$$

2. For all elements K, distribute the fluctuation to the three nodes of K. Let ϕ_j^K denote the *amount* of fluctuation sent to node $j \in K$, then the *conservation/consistency* requirement is

$$\sum_{j=1}^{j=j_K} \phi_j^K = \phi^K \tag{22}$$

3. For all nodes $i \in \Omega_h$, assemble *signals* from surrounding elements and evolve toward steady state by some iterative procedure such as e.g.

$$u_i^{n+1} = u_i^n - \omega_i \sum_{K \in K_i} \phi_i^K \tag{23}$$

The method described by (21)-(22)-(23) aims at providing a solution to the discrete algebraic system

$$\sum_{K \in K_i} \phi_i^K = 0 \,, \quad \forall \, i \in \Omega_h \tag{24}$$

As formulated, it does not include boundary conditions (BCs). The most general way to introduce them is to consider, for any face $f \in \partial\Omega_h$ the *face fluctuations*.

$$\phi^f = \int_f \vec{a} \cdot \mathbf{n}(g_h^* - u_h)\, df \tag{25}$$

where, to embed the compatibility condition implicit in (19), we have introduced the numerical flux :

$$(g^* \vec{a}) \cdot \mathbf{n} = \vec{a} \cdot \mathbf{n} \left[\frac{1 + \text{sign}(\vec{a} \cdot \mathbf{n})}{2} u + \frac{1 - \text{sign}(\vec{a} \cdot \mathbf{n})}{2} g \right]$$

Face fluctuations can be split to the degrees of freedom $j \in f$ by means of distributed residuals ϕ_j^f such that

$$\sum_{j=1}^{j_f} \phi_j^f = \phi^f \tag{26}$$

Finally, the complete discrete fluctuation splitting/residual distribution steady equations read

$$\sum_{K \in K_i} \phi_i^K + \sum_{f \in F_i \bigcap \partial\Omega_h} \phi_i^f = 0 \tag{27}$$

3.1.1 Accuracy conditions

The first formulation of these schemes, on linear triangular elements, relied for the construction of second order discretizations on the so-called *linearity preservation* property [53, 62] defined as follows.

Definition 3.1 (Linearity preservation)**.** *Let $\{\beta_j^K\}_{j \in K}$ be a set of distribution coefficients* uniformly bounded *with respect to h, u_h, ϕ^K, and with respect to the data of the problem ($\vec{a}$, boundary data, etc), and verifying the consistency property*

$$\sum_{j=1}^{j_K} \beta_j^K = 1 \tag{28}$$

A Linearity Preserving *scheme is one for which*

$$\phi_i^K = \beta_i^K \phi^K \tag{29}$$

Proposition 3.2. *Linearity preserving schemes are second order accurate.*

The simple property stated in definition 3.1 and in proposition 3.2 has been known since the late 80s, but it has taken more than a decade to be formally understood. A more general characterisation of the accuracy of these schemes, due to [63] and generalized in [64, 61, 65, 66, 67], is the following.

Definition 3.3 (Truncation error and accuracy)**.** *Let ψ be smooth function, $\psi \in C^{r+1}(\Omega)$. Let Ω_h be an unstructured grid composed of non-overlapping elements. On the generic element $K \in \Omega_h$ consider the $r-$th degree continuous polynomial approximation* (20)*. Let in particular $\psi_h = \sum_{j \in K} \psi_j \varphi_j$ be the $r-$th degree polynomial approximation of type* (20) *of ψ, the values ψ_j being obtained by Galerkin projection. Consider now an exact smooth function $u \in H^{r+1}$ verifying* (18)-(19) *in a classical sense : $\vec{a} \cdot \nabla u = 0$ in Ω, and $u = g$ on $\partial\Omega^-$. Let u_h be its polynomial approximation of degree r of type* (20)*, obtained by Galerkin projection. Let now $\phi_j^K(u_h)$ and $\phi_j^f(u_h)$ the value of the split residuals obtained when replacing the nodal values of the solution obtained with the scheme by the values u_j of the Galerkin projection of u. We define the integral truncation error $\epsilon(u_h, \psi)$*

$$\epsilon(u_h, \psi) = \sum_{j \in \Omega_h} \psi_j \left\{ \sum_{K \in K_j} \phi_j^K(u_h) + \sum_{f \in F_j} \phi_j^f(u_h) \right\} = \sum_{K \in \Omega_h} \sum_{j=1}^{j_K} \psi_j \phi_j^K(u_h) + \sum_{f \in \partial\Omega_h} \sum_{j=1}^{j_f} \psi_j \phi_j^f(u_h) \tag{30}$$

We shall say that a scheme is $r+1$ order accurate if it verifies the truncation error estimate

$$|\epsilon(u_h, \psi)| \leq C(\Omega_h)\, h^{r+1}$$

The following general characterization is possible.

Proposition 3.4. *In d spatial dimensions, a sufficient condition for scheme* (27) *to be* $r+1$ *order accurate in the sense of definition 3.3 is to simultaneously have*

$$\begin{aligned} \left|\phi_i^K(u_h)\right| &\leq C_{\Omega_h} h^{r+d} \quad \forall K \in \Omega_h \,,\ \forall\, i \in K \\ \left|\phi_i^f(u_h)\right| &\leq C_{\partial\Omega_h} h^{r+d-1} \quad \forall f \in \partial\Omega_h \,,\ \forall\, i \in f \end{aligned} \tag{31}$$

The proof of this property is omitted for brevity. The interested reader can refer to [64, 61, 65, 66, 67] for details. The importance of this characterization is that it allows to provide some design conditions. To see this, one must first recall that for the solution u of definition 3.3, and for its Galerkin projection u_h on the r-degree finite element polynomial space, we can use classical approximation results [68, 69] to show that

$$\|\vec{a}\cdot\nabla u_h - \vec{a}\cdot\nabla u\| \leq C_u h^r \quad \text{in} \quad \Omega_h$$

and, provided that the boundary $\partial\Omega$ is also smooth enough, provided that $\partial\Omega_h$ is a high order polynomial rendering of the exact boundary [66, 67], we also have[1]

$$\|(g^*\vec{a}\cdot\mathbf{n})_h - (g^*\vec{a}\cdot\mathbf{n})\| \leq C_{u,\mathbf{n}} h^{r+1}$$

where the norms used are standard L norms, such as the L^2 or the max norm, with no derivatives involved. With the regularity hypotheses made on the mesh, we also have that $|K| = \mathcal{O}(h^d)$ and $|f| = \mathcal{O}(h^{d-1})$, for and K and any f. This, and the fact that $\vec{a}\cdot\nabla u = 0$ and $g^* - u = 0$ for the exact solution, leads to the conclusion that: a sufficient condition for a scheme to be $r+1$ order accurate in the sense of definition 3.3 is that we can find for any $K \in \Omega_h$ and for any $f \in \partial\Omega_h$ sets of uniformly bounded *test functions* ω_i^K and ω_i^f, such that

$$\sum_{j=1}^{j_K} \omega_i^K = 1 \quad \text{and} \quad \sum_{j=1}^{j_f} \omega_i^f = 1$$

and that the distribution can be obtained as

$$\phi_i^K(u_h) = \int_K \omega_i^K\, \vec{a}\cdot\nabla u_h \, d\mathbf{x} \quad \text{and} \quad \phi_i^f(u_h) = \int_f \omega_i^f\, \vec{a}\cdot\mathbf{n}\,(g_h^* - u_h)\, df \tag{32}$$

Clearly, the linearity preserving schemes of definition 3.1 are obtained as the particular case in which the test functions are constant within each element !

As we will see immediately, this consistency analysis, applies trivially to classical continuous Galerkin discretizations, as well as to their stabilized counterparts, as well as to discontinuous Galerkin methods.

3.1.2 Stability and convergence

The above consistency analysis gives conditions under which that *if* convergence with respect to the mesh parameter h is obtained, $r+1$ convergence rates are expected w.r.t. h for a $r-$th degree polynomial approximation, and in correspondence of sufficiently smooth solutions. The missing piece of information is : how do we make sure that convergence is indeed achieved ? A classical finite element convergence analysis would need two main ingredients [69] : a consistency estimate, which we have provided, and a stability condition, which we have not. If we could provide a stability statement which ensures *e.g.* that $\forall u_h$ in our approximation space

$$|\sum_K \sum_{j=1}^{j_K} u_j \phi_j^K(u_h) + \sum_f \sum_{j=1}^{j_f} u_j \phi_j^f(u_h)| \geq C' \|u_h\|^2 \,, \quad \text{with} \ \ 0 < C' < \infty \,, \tag{33}$$

then using more or less classical arguments [69], we could infer the existence of the discrete solution, and derive more rigorous estimates for the error associated to this solution.

Unfortunately, *to this day residual distribution schemes lack a framework for stability analysis*. Some weaker results showing the decay of the solution energy (L^2 norm) during iterations (23) have been shown in several works [70, 71, 61]. These conditions are however not sufficient to say more on the discrete solution.

[1] this aspect can be explained in more rigorous terms and made systematic if taken into account from the start

On the other hand, we are able to rule out some schemes as the following property shows in two space dimensions[2].

Proposition 3.5 (Fall of the $\beta\phi$ paradigm, 2d advection)**.** *Consider the solution of*

$$\vec{a} \cdot \nabla u = 0$$

in two space dimensions, with $\vec{a}$ constant, and with $\partial\Omega$ a collection of straight sides. Any scheme of the form

$$0 = \sum_{K \in K_i} \beta_i^K \phi^K + \sum_{f \in F_i \bigcap \partial\Omega_h} \beta_i^f \phi^f$$

cannot be freed of high frequency spurious modes whatever the form of β_i^K, if K is a P^k Lagrange triangle with $k > 2$ and if K is a Q^k Lagrange quadrilateral $\forall k \geq 1$.

Proof. For all elements considered, we explicitly show one spurious mode exact solution of the discrete problem with homogeneous boundary conditions. This mode can be added to any grid function without the scheme detecting its presence, thus preserving this unphysical perturbation.

First recall that for homogeneous boundary conditions and using the hypothesis on $\partial\Omega$

$$\phi^K = \oint_{\partial K} \vec{a} \cdot \mathbf{n} u_h = \sum_{f \in \partial K} \vec{a} \cdot \mathbf{n} \int_f u_h \, df$$

and

$$\phi^f = -\int_f \frac{1 - \text{sign}(\vec{a} \cdot \mathbf{n})}{2} \vec{a} \cdot \mathbf{n} \, u_h \, df = -\frac{1 - \text{sign}(\vec{a} \cdot \mathbf{n})}{2} \vec{a} \cdot \mathbf{n} \int_f u_h \, df$$

so we focus on the approximation of the integrals of u_h over element faces. Denote by let the number of freedom on each face $f \in \partial K$ be $\mathrm{C}_f + 2$. We consider the mode defined by $u_j = 1$ if j is a vertex, otherwise on each $f \in \partial K$ we set $\forall j \neq v$

$$u_j = -\frac{2}{\mathrm{C}_f} \frac{\int_f \varphi_v \, df}{\int_f \varphi_j \, df}$$

having denoted with v one of the two vertices forming face f. The mode is compatible with the continuity of the representation, and with the adoption of hybrid meshes. For P^k triangles with $k \geq 3$ and Q^k elements with $k \geq 2$, the value of the solution at nodes within the elements remains arbitrary. For this mode, one easily checks that $\phi^K = 0\,, \forall K$, and that $\phi^f = 0\,, \forall\, f \in \partial\Omega_h$.

The only remaining element is the Q^1 quadrilateral which is easily checked to suffer from the checkerboard spurious mode in which u oscillates between -1 and 1 on every face. □

The important consequence of proposition 3.5 is that we have to start looking for schemes exploiting the sub-elemental variation of the discrete solution. A well known example of such a scheme, perfectly fitting the framework presented, is the SUPG scheme of T.J.R. Hughes and collaborators [72, 73, 74] obtained by setting

$$\phi_i^K = \int_K \varphi_i \vec{a} \cdot \nabla u_h \, d\mathbf{x} + \overbrace{\int_K \vec{a} \cdot \nabla \varphi_i \, \tau_K \, \vec{a} \cdot \nabla u_h \, d\mathbf{x}}^{\text{Streamline Dissipation}} \qquad \text{and} \qquad \phi_i^f = \int_f \varphi_i \vec{a} \cdot \mathbf{n} \, (g_h^* - u_h) \, df \tag{34}$$

Stability results for the SUPG scheme can be obtained in the classical sense discussed in the beginning of this section (see for example [75, 76, 77, 78, 42, 74] and references therein), and are based on the positive-semidefinite nature of the bi-linear form associated to the Streamline-dissipation term.

Other examples of schemes allowing to overcome the flaw of proposition 3.5 will be given in the following. In general, guidelines to construct such methods can be obtained by considering the convergence

[2] A similar explicit construction can be done in the three space dimensions including high prder tets, prisms, and hexas. Details are left out of this manuscript

of iteration (23). In the simplest setting of scalar advection, if we recast this iteration as the following update for the array of degrees of freedom U

$$U^{n+1} = U^n - \omega(A_h U^n - F)$$

convergence requires that for some $r < 1$ and for all V

$$\|(\mathrm{Id} - \omega A_h)V\|^2 \leq r\|V\|^2$$

which can be developed into

$$V^t A_h V \geq \frac{1-r}{2\omega}\|V\|^2 + \frac{\omega}{2}\|A_h V\|^2 \geq C_h\|V\|^2 \geq 0$$

leading back to a condition of type (33), and to the necessary (albeit not sufficient) condition

$$V^t A_h V \geq 0 \tag{35}$$

which we will use in the following.

3.1.3 Embedding a discrete maximum principle

Monotonicity of the numerical solution is retained by the so-called *local positivity* contraints for the distribution. This property is related to positive coefficient theory which has replaced the TVD theory to construct high order schemes [79, 80, 81].

Definition 3.6 (Positive scheme)**.** *A* (locally) positive *scheme is one for which*

$$\phi_i^K = \sum_{\substack{j\in K\\ j\neq i}} c_{ij}(u_i - u_j)\,, \quad c_{ij} \geq 0\ \forall j \in K \tag{36}$$

Positivity is the key to the construction of non-oscillatory schemes [53, 62] :

Proposition 3.7 (Local Positivity and *discrete maximum principle*)**.** *Locally positive schemes, combined with the evolution step* (23) *verify the discrete maximum principle*

$$\min_{j\in K_i} u_j^n \leq u_i^{n+1} \leq \max_{j\in K_i} u_j^n \quad \forall\, i \in \Omega_h$$

under the following condition

$$\min_{i\in\Omega_h}\left(\omega_i \sum_{K\in K_i}\sum_{\substack{j\in K\\ j\neq i}} c_{ij}\right) \leq 1$$

Proof. The proof follows from the positivity of the c_{ij}s and time step restriction, and from

$$u_i^{n+1} = \left(1 - \omega_i \sum_{K\in K_i}\sum_{\substack{j\in K\\ j\neq i}} c_{ij}\right)u_i^n + \omega_i \sum_{\substack{j\in K_i\\ j\neq i}}\sum_{K\in K_i \bigcap K_j} c_{ij}u_j^n$$

□

This characterization can be generalized to fully consistent time dependent discretizations as we will show later.

3.1.4 A general framework : relation with classical discretization approaches

The formalism introduced in the previous sections for the scalar advection equation can be easily generalized to (systems of) steady nonlinear conservation laws of the form

$$\text{div } \mathbf{f}(\mathbf{w}) = 0 \quad \text{on} \quad \Omega \subset \mathbb{R}^2 \tag{37}$$

with the appropriate boundary conditions on $\partial\Omega$. We now look for a solution satisfying

$$\sum_{K \in K_i} \boldsymbol{\phi}_i^K + \sum_{f \in F_i \bigcap \partial\Omega_h} \boldsymbol{\phi}_i^f = 0 \tag{38}$$

where in every element K

$$\sum_{i \in K} \boldsymbol{\phi}_i^K = \oint_{\partial K} \mathbf{f}(\mathbf{w}_h) \cdot \mathbf{n} \, d\partial K = \int_K \text{ div } \mathbf{f}(\mathbf{w}_h) d\mathbf{x} \tag{39}$$

while on a boundary face f we have

$$\sum_{i \in f} \boldsymbol{\phi}_i^f = \int_f \big(\hat{\mathbf{f}}(\mathbf{w}_h, \mathbf{g}, \mathbf{n}) - \mathbf{f}(\mathbf{w}_h) \cdot \mathbf{n}\big) df \tag{40}$$

where the numerical flux $\hat{\mathbf{f}}(\mathbf{w}_h, \mathbf{g}, \mathbf{n})$ accounts for the boundary conditions.

This framework defines a sort of *super class* of methods, which allows to embed, and has relations, with all the discretization approaches introduced in section 2.

Continuous FEM as residual distribution

The simplest example is perhaps that of the stabilized finite elements discussed in section 2.1.3. In particular, given a continuous collocated finite element expansion for which we can write $V_h = \text{ span}\{\varphi_i\}_{i \in \Omega_h}$, then the (un-stabilized) continuous Galerkin method is obtained simply by setting

$$\boldsymbol{\phi}_i^K = \int_K \varphi_i \text{ div } \mathbf{f}(\mathbf{w}_h) \, d\mathbf{x} \,, \quad \boldsymbol{\phi}_i^f = \int_f \varphi_i \big(\hat{\mathbf{f}}(\mathbf{w}_h, \mathbf{g}, \mathbf{n}) - \mathbf{f}(\mathbf{w}_h) \cdot \mathbf{n}\big) df$$

For nodal finite elements, the relation $\sum_{i \in K} \varphi_i = 1$ ensures that consistency is satisfied.

There is, however, a slight catch which is worth underlying. The relation between the last definitions, and the consistency condition (39) requires that exact integration is performed, w.r.t. the assumed polynomial variation of $\mathbf{w}_h$ and the definition of the nonlinear flux $\mathbf{f}$. This is required to go from the integral of the flux divergence, to the boundary integral (39), so that the variational formulation (15b) is recovered. In practice, exact quadrature is never used. The practical way to handle this issue, is to introduce a high order polynomial representation of the flux $\mathbf{f}_h$. Based on the accuracy of a given quadrature formula, we can uniquely identify the polynomial degree of such an expansion, built starting from the reconstructed values of $\mathbf{w}_h$ at a sufficient number of flux evaluation points, exactly as done in the so called quadrature free approaches used in Discontinuous Galerkin [82, 83], and in the most recent flux reconstruction methods (cf [84] and references therein). Based on the exact evaluation of the integrals of this polynomial flux, we can re-formulate our consistency conditions as being

$$\sum_{i \in K} \boldsymbol{\phi}_i^K = \oint_{\partial K} \mathbf{f}_h(\mathbf{w}_h) \cdot \mathbf{n} \, d\partial K = \int_K \text{ div } \mathbf{f}_h(\mathbf{w}_h) d\mathbf{x} \tag{41}$$

and

$$\sum_{i \in f} \boldsymbol{\phi}_i^f = \int_f \big(\hat{\mathbf{f}}^h(\mathbf{w}_h, \mathbf{g}, \mathbf{n}) - \mathbf{f}_h(\mathbf{w}_h) \cdot \mathbf{n}\big) df \tag{42}$$

The choice of the polynomial degree has to respect at least some accuracy constraints which are easily deduced from the analysis of section 3.1.1. In particular, for this analysis to apply in the nonlinear case, one must ensure that for a given smooth flux $\mathbf{f}$, and for a r-th degree finite element approximation

$$\| \text{ div } \mathbf{f}_h(\mathbf{w}_h) - \text{ div } \mathbf{f} \|_K \leq C_u h^r \,, \quad \| \mathbf{f}_h \cdot \mathbf{n} - \mathbf{f} \cdot \mathbf{n} \|_f \leq C_{u,\mathbf{n}} h^{r+1}$$

This requires the flux polynomial to be of degree $r_{\mathbf{f}} \geq r$.

With this modification, the (un-stabilised) continuous Galerkin approximation will read

$$\boldsymbol{\phi}_i^K = \int_K \varphi_i \text{ div } \mathbf{f}_h(\mathbf{w}_h)\, d\mathbf{x}\,, \quad \boldsymbol{\phi}_i^f = \int_f \varphi_i \big(\hat{\mathbf{f}}^h(\mathbf{w}_h, \mathbf{g}, \mathbf{n}) - \mathbf{f}_h(\mathbf{w}_h)\cdot\mathbf{n}\big) df$$

while a stabilised variant is readily obtained by including in $\boldsymbol{\phi}_i^K$ the streamline dissipation term (cf. section 2.1.3), leading to a SUPG distribution :

$$\boldsymbol{\phi}_i^K = \int_K \varphi_i \text{ div } \mathbf{f}_h(\mathbf{w}_h)\, d\mathbf{x} + \int_K \big(\nabla_{\mathbf{w}}\mathbf{f}(\mathbf{w}_h)\cdot\nabla\varphi_i\big)\mathcal{T}_K\big(\nabla_{\mathbf{w}}\mathbf{f}(\mathbf{w}_h)\cdot\nabla\mathbf{w}_h\big) d\mathbf{x} \tag{43}$$

where the relation $\sum_{i\in K}\varphi_i = 1$ allows to show that (42)-(41) are met.

Although different in spirit, the edge-stabilized schemes discussed e.g. in [85, 86] can be also recast in the formalism above by setting

$$\boldsymbol{\phi}_i^K = \int_K \varphi_i \text{ div } \mathbf{f}_h(\mathbf{w}_h)\, d\mathbf{x} + \oint_{\partial K} \gamma^{\partial K}(\mathbf{w}_h)[\nabla\mathbf{w}_h]\cdot[\nabla\varphi_i] \tag{44}$$

where again consistency is a consequence of the partition of unity property, while one ca easily demonstrate that the accuracy conditions are met provided that $\gamma^{\partial K}(\mathbf{w}_h) = \mathcal{O}(h^2)$, as in (16c).

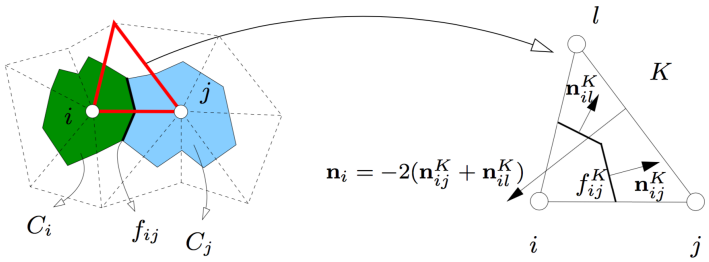


Figure 3: Node centered finite volume

Finite volume vs residual distribution : local conservation and continuous FEM

We recall here the analogy between residual distribution node centered finite volume schemes on median dual cells. With the notation of section 2.1.1, and with reference to figure 3, we can write the semi-discrete evolution equation for $\mathbf{w}_i$, the average of $\mathbf{w}$ on cell C_i, as

$$|C_i|\frac{d\mathbf{w}_i}{dt} = -\sum_j \int_{f_{ij}} \hat{f}\big(\mathcal{R}(\mathbf{w}_h)_i, \mathcal{R}(\mathbf{w}_h)_j, \mathbf{n}_{ij}\big) = -\sum_{K\in K_i}\sum_{\substack{j\in K\\ j\neq i}} \int_{f_{ij}^K} \hat{f}\big(\mathcal{R}(\mathbf{w}_h)_i, \mathcal{R}(\mathbf{w}_h)_j, \mathbf{n}_{ij}\big) = -\sum_{K\in K_i}\sum_{\substack{j\in K\\ j\neq i}} \hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij} \tag{45}$$

Local conservation is equivalent now to the condition

$$\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij} + \hat{\mathbf{f}}_{ji}^K\cdot\mathbf{n}_{ji} = 0 \tag{46}$$

Since C_i is a closed polygon, we also have $\sum_{K\in K_i}\sum_{\substack{j\in K\\ j\neq i}} \mathbf{n}_{ij} = 0$, which allows to recast (45) as

$$|C_i|\frac{d\mathbf{w}_i}{dt} = -\sum_{K\in K_i}\sum_{\substack{j\in K\\ j\neq i}} (\hat{\mathbf{f}}_{ij}^K - \mathbf{f}_i)\cdot\mathbf{n}_{ij} \tag{47}$$

If we now set

$$\boldsymbol{\phi}_i^K = \sum_{\substack{j\in K\\ j\neq i}} (\hat{\mathbf{f}}_{ij}^K - \mathbf{f}_i)\cdot\mathbf{n}_{ij} \tag{48}$$

we find that the local conservation property (46) implies

$$\sum_{j\in K}\boldsymbol{\phi}_i^K=\frac{1}{2}\sum_{j\in K}\mathbf{f}_j\cdot\mathbf{n}_j=\oint_{\partial K}\mathbf{f}_h\cdot\mathbf{n}\,df=\boldsymbol{\phi}^K$$

with $\mathbf{f}_h=\sum_j\varphi_j\mathbf{f}_j$.

This shows that for any given higher order finite volume discretization we may define a residual distribution method consistent with a second order polynomial flux approximation. While this was known for a certain time, the reverse is not. In particular, given a definition of the split residuals $\{\boldsymbol{\phi}_j^K\}_{j\in K}$ we may ask if there exist a definition of consistent fluxes expressing local conservation over the median dual cell for the residual distribution method. If we require these fluxes to satisfy (48), the we may write the system

$$\begin{aligned}
\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij}+\hat{\mathbf{f}}_{il}^K\cdot\mathbf{n}_{il}&=\boldsymbol{\phi}_i^K-\frac{1}{2}\mathbf{f}_i\cdot\mathbf{n}_i\\
\hat{\mathbf{f}}_{ji}^K\cdot\mathbf{n}_{ji}+\hat{\mathbf{f}}_{jl}^K\cdot\mathbf{n}_{jl}&=\boldsymbol{\phi}_j^K-\frac{1}{2}\mathbf{f}_j\cdot\mathbf{n}_j\\
\hat{\mathbf{f}}_{li}^K\cdot\mathbf{n}_{li}+\hat{\mathbf{f}}_{lj}^K\cdot\mathbf{n}_{lj}&=\boldsymbol{\phi}_l^K-\frac{1}{2}\mathbf{f}_l\cdot\mathbf{n}_l
\end{aligned}\tag{49}$$

using local conservation, and setting $\boldsymbol{\Psi}_i^K=\boldsymbol{\phi}_i^K-\mathbf{f}_i\cdot\mathbf{n}_i/2$, we obtain a linear system for $(\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij},\hat{\mathbf{f}}_{jl}^K\cdot\mathbf{n}_{jl},\hat{\mathbf{f}}_{li}^K\cdot\mathbf{n}_{li})$

$$\begin{pmatrix}1&0&-1\\-1&1&0\\0&-1&1\end{pmatrix}\begin{pmatrix}\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij}\\\hat{\mathbf{f}}_{jl}^K\cdot\mathbf{n}_{jl}\\\hat{\mathbf{f}}_{li}^K\cdot\mathbf{n}_{li}\end{pmatrix}=\begin{pmatrix}\boldsymbol{\Psi}_i^K\\\boldsymbol{\Psi}_j^K\\\boldsymbol{\Psi}_l^K\end{pmatrix}\tag{50}$$

The associated matrix has rank 2. We can easily find particular solutions setting to zero one of the unknowns and solving the resulting sub system. Averaging out the three particular solutions obtained (for symmetry) er en up with the following multidimensional numerical fluxes w.r.t. which the residual distribution scheme is locally conservative on the median dual cell :

$$\begin{aligned}
\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij}=\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij}(\mathbf{w}_i,\mathbf{w}_j,\mathbf{w}_l)&=\frac{1}{3}(\boldsymbol{\Psi}_i^K-\boldsymbol{\Psi}_j^K)=\frac{1}{3}(\boldsymbol{\phi}_i^K-\boldsymbol{\phi}_j^K)-\frac{1}{6}(\mathbf{f}_i\cdot\mathbf{n}_i-\mathbf{f}_j\cdot\mathbf{n}_j)\\
\hat{\mathbf{f}}_{jl}^K\cdot\mathbf{n}_{jl}=\hat{\mathbf{f}}_{jl}^K\cdot\mathbf{n}_{jl}(\mathbf{w}_i,\mathbf{w}_j,\mathbf{w}_l)&=\frac{1}{3}(\boldsymbol{\Psi}_j^K-\boldsymbol{\Psi}_l^K)=\frac{1}{3}(\boldsymbol{\phi}_j^K-\boldsymbol{\phi}_l^K)-\frac{1}{6}(\mathbf{f}_j\cdot\mathbf{n}_j-\mathbf{f}_l\cdot\mathbf{n}_l)\\
\hat{\mathbf{f}}_{li}^K\cdot\mathbf{n}_{li}=\hat{\mathbf{f}}_{li}^K\cdot\mathbf{n}_{li}(\mathbf{w}_i,\mathbf{w}_j,\mathbf{w}_l)&=\frac{1}{3}(\boldsymbol{\Psi}_l^K-\boldsymbol{\Psi}_i^K)=\frac{1}{3}(\boldsymbol{\phi}_l^K-\boldsymbol{\phi}_i^K)-\frac{1}{6}(\mathbf{f}_l\cdot\mathbf{n}_l-\mathbf{f}_i\cdot\mathbf{n}_i)
\end{aligned}\tag{51}$$

These are *three states multidimensional numerical fluxes*. Consistency can be formulated as

$$\hat{\mathbf{f}}_{ij}^K\cdot\mathbf{n}_{ij}(\mathbf{w},\mathbf{w},\mathbf{w})=\mathbf{f}(\mathbf{w})\cdot\frac{\mathbf{n}_j-\mathbf{n}_i}{6}$$

as for a constant state over the element we always have $\boldsymbol{\phi}_i^K=0\,\forall\,i$. Other standard properties of numerical fluxes, e.g. Lipschitz continuity, are inherited from the properties of the physical flux $\mathbf{f}$, and of the split residuals $\boldsymbol{\phi}_i^K$.

A similar construction can be repeated for schemes written on high order finite elements. To understand how this works, we start from the $\mathbb{P}^2$ case. We consider the set-up defined of figure 4 : the element is split first into 4 sub-triangles K_1, K_2, K_3, and K_4. From this sub-triangulation, we can construct a dual mesh as in the P^1 case. The dual mesh is a collection of cells C_j whose intersection with an element K defines 6 sub-zones, represented by dashed lines in the figure. The notation used in this case is similar to the one used before : in the sub-triangle K_i, we denote by $\mathbf{n}_{ij}^{K_i}$ the normal to the portion of the face separating the median dual cells C_i and C_j, and by $\mathbf{f}_{ij}^{K_i}\cdot\mathbf{n}_{ij}^{K_i}$ the corresponding local numerical flux.

We can now write down the finite volume equations for each control cell C_j, and then proceed as in the P^1 case to determine contributions associated to each sub-element K_i. To relate these sub-elemental residuals to the $\mathbb{P}$ distributed residuals, we sum for each node the contribution from the sub-elements to

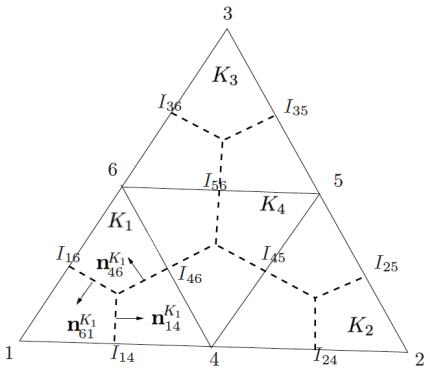


Figure 4: $\mathbb{P}^2$ residual distribution and finite volumes

which the node belongs. As before, this leads to a system of equations, which reads :

$$\begin{aligned}
\hat{\mathbf{f}}_{14}^{K_1}\cdot\mathbf{n}_{14}^{K_1}-\hat{\mathbf{f}}_{61}^{K_1}\cdot\mathbf{n}_{61}^{K_1}&=\boldsymbol{\phi}_1^K-\mathbf{F}_1^K\\
-\hat{\mathbf{f}}_{42}^{K_2}\cdot\mathbf{n}_{42}^{K_2}+\hat{\mathbf{f}}_{25}^{K_2}\cdot\mathbf{n}_{25}^{K_2}&=\boldsymbol{\phi}_2^K-\mathbf{F}_2^K\\
-\hat{\mathbf{f}}_{53}^{K_3}\cdot\mathbf{n}_{53}^{K_3}+\hat{\mathbf{f}}_{36}^{K_3}\cdot\mathbf{n}_{36}^{K_3}&=\boldsymbol{\phi}_3^K-\mathbf{F}_3^K\\
-\hat{\mathbf{f}}_{14}^{K_1}\cdot\mathbf{n}_{14}^{K_1}+(\hat{\mathbf{f}}_{41}^{K_1}\cdot\mathbf{n}_{46}^{K_1}-\hat{\mathbf{f}}_{64}^{K_4}\cdot\mathbf{n}_{64}^{K_4})+(\hat{\mathbf{f}}_{45}^{K_4}\cdot\mathbf{n}_{45}^{K_4}-\hat{\mathbf{f}}_{54}^{K_2}\cdot\mathbf{n}_{54}^{K_2})+\hat{\mathbf{f}}_{42}^{K_2}\cdot\mathbf{n}_{42}^{K_2}&=\boldsymbol{\phi}_4^K-\mathbf{F}_4^K\\
-\hat{\mathbf{f}}_{25}^{K_2}\cdot\mathbf{n}_{25}^{K_2}+(\hat{\mathbf{f}}_{54}^{K_2}\cdot\mathbf{n}_{54}^{K_2}-\hat{\mathbf{f}}_{45}^{K_4}\cdot\mathbf{n}_{45}^{K_4})+(\hat{\mathbf{f}}_{56}^{K_4}\cdot\mathbf{n}_{56}^{K_4}-\hat{\mathbf{f}}_{65}^{K_3}\cdot\mathbf{n}_{65}^{K_3})+\hat{\mathbf{f}}_{53}^{K_3}\cdot\mathbf{n}_{53}^{K_3}&=\boldsymbol{\phi}_5^K-\mathbf{F}_5^K\\
-\hat{\mathbf{f}}_{36}^{K_3}\cdot\mathbf{n}_{36}^{K_3}+(\hat{\mathbf{f}}_{65}^{K_3}\cdot\mathbf{n}_{65}^{K_3}-\hat{\mathbf{f}}_{56}^{K_4}\cdot\mathbf{n}_{56}^{K_4})+(\hat{\mathbf{f}}_{64}^{K_4}\cdot\mathbf{n}_{64}^{K_4}-\hat{\mathbf{f}}_{46}^{K_1}\cdot\mathbf{n}_{46}^{K_1})+\hat{\mathbf{f}}_{61}^{K_1}\cdot\mathbf{n}_{61}^{K_1}&=\boldsymbol{\phi}_6^K-\mathbf{F}_6^K
\end{aligned}\tag{52}$$

having set

$$\mathbf{F}_i^K=\int\limits_{C_i\bigcap\partial K}\mathbf{f}(\mathbf{w}_h)\cdot\mathbf{n}\,d\Gamma$$

with $\mathbf{w}_h$ the $\mathbb{P}$ finite element solution, and with the obvious relation

$$\sum_{K\in K_i}\mathbf{F}_i^K=0$$

due the the continuity of the flux. If now we set, to simplify the notation

$$\begin{aligned}
\hat{\mathbf{f}}_{14}&:=\hat{\mathbf{f}}_{14}^{K_1}\cdot\mathbf{n}_{14}^{K_1}\ ,\quad \hat{\mathbf{f}}_{61}:=\hat{\mathbf{f}}_{61}^{K_1}\cdot\mathbf{n}_{61}^{K_1}\ ,\quad \hat{\mathbf{f}}_{64}:=\hat{\mathbf{f}}_{64}^{K_4}\cdot\mathbf{n}_{64}^{K_4}-\hat{\mathbf{f}}_{46}^{K_1}\cdot\mathbf{n}_{46}^{K_1}\\
\hat{\mathbf{f}}_{42}&:=\hat{\mathbf{f}}_{42}^{K_2}\cdot\mathbf{n}_{42}^{K_2}\ ,\quad \hat{\mathbf{f}}_{25}:=\hat{\mathbf{f}}_{25}^{K_2}\cdot\mathbf{n}_{25}^{K_2}\ ,\quad \hat{\mathbf{f}}_{45}:=\hat{\mathbf{f}}_{45}^{K_4}\cdot\mathbf{n}_{45}^{K_4}-\hat{\mathbf{f}}_{54}^{K_2}\cdot\mathbf{n}_{54}^{K_2}\\
\hat{\mathbf{f}}_{53}&:=\hat{\mathbf{f}}_{53}^{K_3}\cdot\mathbf{n}_{53}^{K_3}\ ,\quad \hat{\mathbf{f}}_{36}:=\hat{\mathbf{f}}_{36}^{K_3}\cdot\mathbf{n}_{36}^{K_3}\ ,\quad \hat{\mathbf{f}}_{56}:=\hat{\mathbf{f}}_{56}^{K_4}\cdot\mathbf{n}_{56}^{K_4}-\hat{\mathbf{f}}_{65}^{K_3}\cdot\mathbf{n}_{65}^{K_3}
\end{aligned}$$

and $\boldsymbol{\Psi}_i^K=\boldsymbol{\phi}_i^K-\mathbf{F}_i^K$, we obtain

$$\begin{aligned}
\hat{\mathbf{f}}_{14}-\hat{\mathbf{f}}_{61}&=\boldsymbol{\Psi}_1^K\\
-\hat{\mathbf{f}}_{42}+\hat{\mathbf{f}}_{25}&=\boldsymbol{\Psi}_2^K\\
-\hat{\mathbf{f}}_{53}+\hat{\mathbf{f}}_{36}&=\boldsymbol{\Psi}_3^K\\
-\hat{\mathbf{f}}_{14}-\hat{\mathbf{f}}_{64}+\hat{\mathbf{f}}_{45}+\hat{\mathbf{f}}_{42}&=\boldsymbol{\Psi}_4^K\\
-\hat{\mathbf{f}}_{25}-\hat{\mathbf{f}}_{45}+\hat{\mathbf{f}}_{56}+\hat{\mathbf{f}}_{53}&=\boldsymbol{\Psi}_5^K\\
-\hat{\mathbf{f}}_{36}-\hat{\mathbf{f}}_{56}+\hat{\mathbf{f}}_{64}+\hat{\mathbf{f}}_{61}&=\boldsymbol{\Psi}_6^K
\end{aligned}\tag{53}$$

System (53) has a very neat interpretation : the sub-triangulation of figure 4 defines a triangulation of element K associated to its degrees of freedom. For any edge between two degrees of freedom, say $[i,j]$, we look for fluxes $\hat{\mathbf{f}}_{ij}$ satisfying (53), with the constraint $\hat{\mathbf{f}}_{ij}+\hat{\mathbf{f}}_{ji}=0$.

In the $\mathbb{P}$ case, one easily shows that the matrix associated to (53) has rank 5, which can be used to obtain a definition of the equivalent finite volume fluxes as done for linear elements. We only sketch the generalization to P^k elements which relies on the following main elements

- construct a triangulation of K which vertices are the degrees of freedom of the interpolation ;
- associate to his sub-triangulation a dual tessellation to be used to define local conservation equations ;
- set
$$\mathbf{\Psi}_i^K = \boldsymbol{\phi}_i^K - \int\limits_{C_i \bigcap \partial K} \mathbf{f}(\mathbf{w}_h) \cdot \mathbf{n}\, d\Gamma$$
- write equations for conservative edge fluxes $\hat{\mathbf{f}}_{ij}$. Assemble a liner system for a subset $\mathcal{F}$ of the ordered couples (i,j) associated to the edges of the sub-triangulation, with $\mathcal{F}$ containing either (i,j), or (j,i) for any two fixed nodes. We have that
  1. the matrix coefficients of the linear system are
$$\theta_{ij} = \begin{cases} 0 & (i,j) \text{ is no an edge} \\ 1 & (i,j) \text{ is an edge and } (i,j) \in \mathcal{F} \\ -1 & (i,j) \text{ is an edge and } (i,j) \notin \mathcal{F} \end{cases}$$
  2. the i-th right hand side of the system is equal to $\mathbf{\Psi}_i^K$
  3. the rank of the system matrix is equal to $n_{dof} - 1$

Our analysis can be also generalized to three space dimensions, and to other type of elements, since it only relies on the possibility of constructing a set of connected dual cells, which is possible for any mesh. This analysis shows that *whatever the type of element, the approximation of the residual distribution spatial discretisation can be reformulated by means of a finite volume approximation defined by multidimensional fluxes function of n_{dof} states. Hence all continuous finite element schemes admitting a residual distribution re-formulation are locally conservative.*

3.1.5 WENO-RD and bridge with DG

One can slightly extend the RD formalism. In what is written above, the main assumption is that the approximation is globally continuous. This assumption can be relaxed. Assume that, as for DG, the trial function space is made of functions that are polynomials on each elements, but we relax the continuity assumption. This problem has been studied in [87, 88, 89]. In each element K, we assume that we have N_K degrees of freedom, say $\{i_{i_j}, j = 1, N_K\}$, and assume that we have constructed residuals $\Phi^K_{i_{i_j}}(u^h)$. The conservation relation must be relaxed into

$$\sum_{j=1}^{N_K} \Phi^K_{i_{i_j}}(\mathbf{w}^h) = \int_{\partial K} \hat{f}(\mathbf{w}^{h,+}, \mathbf{w}^{h,-}, \mathbf{n}) d\partial K \tag{54}$$

where $\hat{f}$ is a consistent numerical flux, $\mathbf{w}^{h,+}, \mathbf{w}^{h,-}$ are the states on the two sides of the faces that makes ∂K. In [88], it is shown how to reformulate a DG method with $\mathbb{P}^1$ elements in this framework. Though limited to $\mathbb{P}^1$ element, this approach can easily be extended to higher order of approximation, see also [88] for a more systematic (than [87]) technique. In the discontinuous case, a simple variant can be found, see [88]. This remark make it possible to use the technique that we describe in section 3.1.7.

A completely different approach has been pursued in [90]. Starting from a finite difference-like grid and using WENO reconstruction, these authors have been successful in developing RD-like approximation for hyperbolic problems.

3.1.6 A general Lax-Wendroff result

One of the key constraint a RD scheme must fullfil is that for any element or face, the sum of the sub-residual must equal the total residuals, these are the conditions (22) and (26).[3] These conditions guaranty a Lax-Wendroff like result, see [64]. More precisely, if we assume that:

[3]The results above can easily be generalised to the conditions 54 provided $\hat{f}$ is Lipschitz continuous.

Assumption 3.1. *The mesh $\mathcal{T}_h$ is conformal and regular. By regular we mean that all elements are roughly the same size, more precisely that there exist constants C_1 and C_2 such that for any element*

$$K,\ \ C_1 \leq \sup_{K\in\mathcal{T}_h} \frac{h^2}{|K|} \leq C_2$$

We introduce the spaces:

$$V_h^k = \{\mathbf{v}_h \in C^0(\mathbb{R}^d)^p; \mathbf{v}_{h|K} \text{ polynomial of degree } k, \forall\ K \in \mathcal{T}_h\}$$

$$X^h = \{\mathbf{v}_h; v_{h|C} \text{ constant } \ \in \mathbb{R}^p, \forall\ C \in \mathcal{C}_h\}.$$

Here, $\mathbf{f}_{|K}$ denotes the restriction of $\mathbf{f}$ to K. The second assumption is on the residuals.

Assumption 3.2. *Let $\mathcal{T}_h$ be a triangulation satisfying the assumption 3.1. For any $C \in \mathbb{R}^+$, there exists $C'(C, \mathcal{T}_h) \in \mathbb{R}^+$ which depends only on C and $\mathcal{T}_h$ such that for any $\mathbf{w} \in X^h$, with $||\mathbf{w}||_{L^\infty(\mathbb{R}^2)} \leq C$ we have*

$$\forall K, \forall i, ||\Phi_i^K|| \leq C'(C, \mathcal{T}_h)\ h \sum_{j\in K} ||\mathbf{w}(j) - \mathbf{w}(i)||. \tag{55}$$

We assume that the residual $\Phi_i^{K'_K}$ and the numerical solution satisfy the following conditions.

Assumption 3.3. *There exists an approximation $\mathbf{f}^h$ of the flux $\mathbf{f}$ such that*

(i) $\forall \mathbf{w}^h \in X^h$, $\Phi^K := \displaystyle\int_K div\ \mathbf{f}^h(\mathbf{w}^h)dx = \sum_{i\in K} \Phi_i^K(\mathbf{w}^h)$,

(ii) $\forall \mathbf{w}^h \in X^h$, $\forall K_1, K_2$ *neigbors,*

$$\mathbf{f}^h(\mathbf{w}^h)_{|K_1}\ .\ \vec{n} = \mathbf{f}^h(\mathbf{w}^h)_{|K_2}\ .\ \vec{n}\ \ a.e.\ on\ K_1 \cap K_2$$

where $\vec{n}$ is a normal of $K_1 \cap K_2$.

(iii) For any $C > 0$, there exists $C'(C)$ such that for any $\mathbf{w}^h \in X^h$ with $||\mathbf{w}^h||_{L^\infty(\mathbb{R}^2)} \leq C$, one has for $K \in \mathcal{T}_h$ and $\mathbf{f}_K^h = \mathbf{f}_{|K}^h$, $||\ div\ \mathbf{f}_K^h(u^h)|| \leq \dfrac{C'}{h} \displaystyle\sum_{i,j} ||\mathbf{w}_i^h - \mathbf{w}_j^h||$ a.e. on K.

(iv) For any sequence $(\mathbf{w}^h)_h$ bounded in $L^\infty(\mathbb{R}^2 \times \mathbb{R}^+)^p$ independently of h and convergent in $L^2_{loc}(\mathbb{R}^2 \times \mathbb{R}^+)^p$ to $\mathbf{w}$, we have

$$\lim_h ||\mathbf{f}^h(\mathbf{w}^h) - \mathbf{f}(\mathbf{w})||_{L^1_{loc}(\mathbb{R}^d\times\mathbb{R}^+)^p} = 0.$$

We have the following result,

Theorem 3.4. *Let be $\mathbf{w}_0 \in L^\infty(\mathbb{R}^d)^p$ and $\mathbf{w}^h$ the approximation given by*

$$\mathbf{w}_i^{n+1} = \mathbf{w}_i^n - \frac{\Delta}{|C_i|}\left(\sum_{K\ni i} \boldsymbol{\phi}_i^K(\mathbf{w}^{n+1}) + \sum_{F\cap\partial\Omega\ni i} \boldsymbol{\phi}_i^F\right)$$
$$\mathbf{w}_i^0 = \mathbf{w}_0(i).$$

We assume that the scheme satisfies the assumptions 3.2 and 3.3. We also assume there exists a constant C that depends only on C_1, C_2 and u_0 and a function $\mathbf{w} \in (L^2(\mathbb{R}^d \times \mathbb{R}^+))^p$ such that

$$\sup_h \sup_{x,y,t} |\mathbf{w}^h(x, y, t)| \leq C$$

$$\lim_h ||\mathbf{w} - \mathbf{w}_h||_{L^2_{loc}(\mathbb{R}^d\times\mathbb{R}^+)^p} = 0$$

Then $\mathbf{w}$ is a weak solution of

$$\frac{\partial \mathbf{w}}{\partial t} + \ div\ \mathbf{f}(\mathbf{w}) = 0$$
$$\mathbf{w}(\mathbf{x}, 0) = \mathbf{w}_0(\mathbf{x})$$

The proof is given in [64].

3.1.7 Construction of non-classical high order schemes

Well posed linear schemes

The simplest method is obtained by splitting the element residuals in a symmetric manner, as e.g.

$$\boldsymbol{\phi}_i^K = \frac{1}{n_{dof}} \boldsymbol{\phi}^K$$

This definition verifies trivially all the accuracy criteria, and the conditions for the Lax-Wendroff theorem. Nevertheless, it is flawed by the existence of spurious mode discussed in section 3.1.2 which is a clear symptom of a lack of stability. An example of a stabilized method is the Lax-Friedrich's type distribution

$$\boldsymbol{\phi}_i^{\mathrm{LF}} = \frac{1}{n_{dof}} \boldsymbol{\phi}^K + \alpha_K (\mathbf{w}_i - \overline{\mathbf{w}}_K) \tag{56}$$

with $\overline{\mathbf{w}}_K$ the arithmetic average of the solution values in K. In the scalar case, we can easily prove the stability of this method in both the L^2 and L^∞ norms (cf. section 3.1.3). This method does not verify the accuracy conditions of section 3.2.1.

To obtain a stable high order method we can follow the ideas of [91, 92], and add to an unstable high order method a streamline dissipation term

$$\boldsymbol{\phi}_i^K = \beta_i^K \boldsymbol{\phi}^k + \theta_K \int_K (\nabla_{\mathbf{w}} \mathbf{f} \cdot \nabla \varphi_i) \mathcal{T}_K (\nabla_{\mathbf{w}} \mathbf{f} \cdot \nabla \mathbf{w}_h) \, d\mathbf{x} \tag{57}$$

with θ_K a scalar coefficient, and where, for generality, we have replaced $1/n_{dof}$ by a generic bounded distribution coefficient. Following [91, 92], we seek for rules to define the term $\theta_K \mathcal{T}_K$ for scalar advection. We start by recasting the method obtained with (57) as

$$-\int_{\Omega_h} u_h \vec{a} \cdot \varphi_i \, d\mathbf{x} + \sum_{K \in K_i} \int_K (\beta_i^K - \varphi_i)(\vec{a} \cdot \nabla u_h) \, d\mathbf{x} + \sum_{K \in K_i} \theta_K \int_K (\vec{a} \cdot \nabla \varphi_i) \mathcal{T}_K (\vec{a} \cdot \nabla u_h) \, d\mathbf{x} = \text{boundary cond.s}$$

We can associate to this method the bi-linear form

$$a(v^h, u_h) + \sum_K b_K(v^h, u_h) = l_{\mathrm{bc.s}}(v^h)$$

where

$$a(v^h, u_h) = a^{\mathrm{Galerkin}}(v^h, u_h) + \sum_K a_K(v^h, u_h)$$

with

$$a_K(v^h, u_h) = \int_K (v_K^\beta - v^h)(\vec{a} \cdot \nabla u_h) \, d\mathbf{x}$$

and where $b_K(v^h, u_h)$ is the streamline dissipation term

$$b_K(v^h, u_h) = \int_K (\vec{a} \cdot \nabla v^h) \theta_K \mathcal{T}_K (\vec{a} \cdot \nabla u_h) \, d\mathbf{x}$$

For simplicity, and following the ideas of [91], we now express the increment $v_K^\beta - v^h$ as a function of ∇v^h, and of a (scheme dependent) element length h_K^β and direction ξ_K^β

$$a_K(v^h, u_h) + b_K(v^h, u_h) = \int_K (\xi_K^\beta \cdot \nabla v^h) \, h_K^\beta \, (\vec{a} \cdot \nabla u_h) \, d\mathbf{x} + \int_K (\vec{a} \cdot \nabla v^h) \, \theta_K \mathcal{T}_K \, (\vec{a} \cdot \nabla u_h) \, d\mathbf{x}$$

Finally, we want to define the coefficient $\theta_K \mathcal{T}_K$ such that

$$a_K(v^h, u_h) + b_K(u_h, u_h) = \int_K (\xi_K^\beta \cdot \nabla u_h) \, h_K^\beta \, (\vec{a} \cdot \nabla u_h) \, d\mathbf{x} + \int_K \theta_K \mathcal{T}_K \, (\vec{a} \cdot \nabla u_h)^2 \, d\mathbf{x} \geq 0 \tag{58}$$

In particular, to have (58) satisfied in practice, we consider the fully discrete evaluation of the streamline dissipation term

$$\int_K (\vec{a}\cdot\nabla\varphi_i)\,\theta_K\mathcal{T}_K\,(\vec{a}\cdot\nabla u_h)\,d\mathbf{x} \approx \theta_K\mathcal{T}_K|K|\sum_{x_{quad}}\omega_{quad}(\vec{a}\cdot\nabla\varphi_i(x_{quad}))\,(\vec{a}\cdot\nabla u_h(x_{quad})) \tag{59}$$

having assumed for simplicity that a constant value of $\mathcal{T}_K$ is used over each element. We seek now guidelines to choose a quadrature formula.

A necessary condition to have (58) is that the quadratic form

$$q_K(u_h) := |K|\sum_{x_{quad}}\omega_{quad}(\vec{a}\cdot\nabla u_h(x_{quad}))^2$$

must be positive whenever $\vec{a}\cdot\nabla u_h \neq 0$. A sufficient condition for this to be true is that

$$\text{if}\quad \vec{a}\cdot\nabla u_h(x_{quad}) = 0\;\forall\, x_{quad}\;\text{then}\quad \vec{a}\cdot\nabla u_h = 0 \tag{60}$$

In this case, we can find positive bounded constants such that

$$C_{1,q}q_K(u_h) \leq \int_K (\vec{a}\cdot\nabla u_h)^2\,d\mathbf{x} \leq C_{2,q}q_K(u_h)$$

Since $V_h = \text{span}\{\varphi_i\}$ is a finite-dimensional spaces, the discrete quantity

$$Q(u_h) = \sum_K q_K(u_h)$$

defines on V_h a norm equivalent to $u_h \mapsto \int_\Omega (\vec{a}\cdot\nabla u_h)^2\,d\mathbf{x}$ This allows to prove the following result.

Proposition 3.8 (Quadrature of the streamline dissipation)**.** *Independently on the values of the weights ω_{quad}, provided that the number of points used ti evaluate* (59) *is large enough to guarantee* (60) *then we can find $(\theta_K\mathcal{T}_K)_0$ such that the scheme obtained with* (57) *is well posed whenever $\theta_K\mathcal{T}_K \geq (\theta_K\mathcal{T}_K)_0$.*

Proof. See [91]. □

In light of this analysis, the set of points used to evaluate (59) does not need necessarily be a set of quadrature points, as the relevant condition is not to evaluate the streamline dissipation term exactly, but to ensure (60). In this light, term (59) can be seen as a sort of *filtering term* allowing to remove spurious modes and guarantee the well-posedness of the method. In particular, even for constant scalar advection, the number of points sufficient to have (60) is smaller than that of most quadrature/cubature formulas providing an exact evaluation of the streamline dissipation term, and in any case simpler point values can be used, as e.g. the x_{dof} (cf. [91, 92]).

Another path to avoid the flaw associated to proposition 3.5, is to "bring the distribution coefficient β under the integral". The classical definition associated to SUPG (43) is one example of such a method. However, we can also provide similar generalizations of the so-called multidimensional upwind methods, which have constituted one of the elements of originality of the residual distribution approach (cf. [61] and references therein). In particular, we consider the method defined in the scalar case by

$$\phi_i^K = \int_K (\vec{a}\cdot\nabla\varphi_i)^+\,\gamma_K\,(\vec{a}\cdot\nabla u_h)\,d\mathbf{x} \tag{61}$$

with $\gamma_K > 0$. If, without loss of generality, we consider the linear advection problem admitting a solution $u > 0$[4], we may assume that we seek a discrete solution $u_h \in V_h^+ = \{u_h \in \text{span}\{\varphi_i\} | u_h > 0\}$, we can in this case write

$$\sum_{i\in\Omega_h} u_i \sum_{K\in K_i}\phi_i^K = \sum_{i\in\Omega_h}\int_K (\vec{a}\cdot\nabla u_h)^+\gamma_K\vec{a}\cdot\nabla u_h\,d\mathbf{x} = \sum_{i\in\Omega_h}\int_K (\vec{a}\cdot\nabla u_h)^+\gamma_K\vec{(\vec{a}}\cdot\nabla u_h)^+\,d\mathbf{x} \geq 0$$

[4]which we may always do in the linear case by rescaling the boundary data g by $M' = |\min_{\mathbf{x}\in\Omega} g| + M$, $M > 0$

as $\vec{a}\cdot\nabla u_h = (\vec{a}\cdot\nabla u_h)^+ + (\vec{a}\cdot\nabla u_h)^-$, and $(\vec{a}\cdot\nabla u_h)^+(\vec{a}\cdot\nabla u_h)^- = 0$. This shows that condition (35) is met, giving an indication of the well-posedness of the method, confirmed by all numerical evidence. Note that for the method to satisfy (39), we need to set

$$\gamma_K = \left(\sum_{j\in K}(\vec{a}\cdot\nabla\varphi_j)^+\right)^{-1}$$

which can be show to reduce in the P^1 case leads to

$$\phi_i^K = \beta_i^K\phi^K\,, \qquad \beta_i^K = (\vec{a}\cdot\mathbf{n}_i)^+/\big(\sum_{j\in K}(\vec{a}\cdot\mathbf{n}_j)^+\big)$$

which is nothing else that the well known multidimensional upwind LDA scheme introduced by Roe, Deconinck and co-workers in the 90's [61]. The generalization (61) is obtained by formally replacing the so-called upwind parameters $k_i = \vec{a}\cdot\mathbf{n}_i/2$ by the i-th streamline component of the solution gradient $k_i = \vec{a}\cdot\nabla\varphi_i$, and by performing the distribution of the local residual, instead of distributing the integrated cell residual ϕ^K. The extension to nonlinear problems is obtained by replacing in (61) the residual $\vec{a}\cdot\nabla u_h$, with $\nabla\cdot\mathbf{f}_h$, with $\mathbf{f}_h$ a higher order polynomial, of at degree at least $k+1$ (one higher than the solution), built starting from the values of u_h. Refer to [93, 94] for more details.

Non-oscillatory methods

The analysis of section 3.1.3 constitutes the basic artillery used in the past years to construct methods allowing a non-oscillatory approximation of discontinuous solutions. The key element of these constructions is some low (first) order linear scheme, satisfying (36). A typical example of such a scheme is given by (56). For this scheme, and in the scalar case, we can readily prove that (36) holds in d space dimensions, as soon as $\alpha_K > h_K^{d-1}\|\nabla_{bbw}\mathbf{f}\|_{L^\infty(K)}$, with h_K a characteristic length scale of the element.

A neat way of producing a formally high order method starting from (56), is to fabricate uniformly bounded distribution coefficients by applying a nonlinear mapping to the quantities $\phi_i^{\rm LF}/\phi^K$. An example of such a mapping is the well known "PSI limiter" [61]

$$\beta_i^{\rm LLF} = \frac{(\phi_i^{\rm LF}/\phi^K)^+}{\sum\limits_{j\in K}(\phi_j^{\rm LF}/\phi^K)^+} = \frac{(\phi_i^{\rm LF}\phi^K)^+}{\sum\limits_{j\in K}(\phi_j^{\rm LF}\phi^K)^+} \tag{62}$$

For the corresponding scheme, one easily shows that

$$\phi_i^{\rm LLF} = \beta_i^{\rm LLF}\phi^K = \gamma_i\phi_i^{\rm LF}\,, \qquad \gamma_i = \frac{(\phi_i^{\rm LF}/\phi^K)^+}{\sum\limits_{j\in K}(\phi_j^{\rm LF}/\phi^K)^+}\phi^K/\phi_i^{\rm LF} \;\in[0,\,1] \tag{63}$$

Thus, this limited Lax-Friedrich's distribution is by construction stable in the L^∞ norm. Unfortunately, we also know from proposition 3.5 that this scheme will not in general converge, and anyways may not converge to the right solution. A cure to this problem has been suggested in [95, 91, 92], and consists in adding the filtering term (59) in smooth regions. The resulting method reads

$$\phi_i^{\rm LLFs} = \beta_i^{\rm LLF}\phi^K + \theta(u_h)|K|\sum_{x_{quad}}\omega_{quad}(\vec{a}\cdot\nabla\varphi_i(x_{quad}))\,\mathcal{T}_K\,(\vec{a}\cdot\nabla u_h(x_{quad})) \tag{64}$$

where $\theta(u_h)$ is defined such that the conditions of proposition 3.8 are met in smooth regions, while $\theta < \mathcal{O}(h_K)$ in vicinity of discontinuities. Practical definitions of this term can be found in [95, 92, 96, 97]. The extension of this construction to systems is performed by computing the limiter (62) either equation by equation, or by a prior projection of the residuals on characteristic directions, and by replacing the advection vector in (64) by the flux Jacobian matrices. A common definition of the scaling matrix $\mathcal{T}_K$ is

$$\mathcal{T}_K = |K|\big(\sum_v(\nabla_{\mathbf{w}}\mathbf{f}(\mathbf{w}_v)\cdot\nabla\varphi_v(\mathbf{x}_v))^+\big)^{-1}$$

with v the vertices of element K.

An alternative construction consists in adding to a linear high order and stable scheme a local amount of shock capturing dissipation. This approach dates back a long way [98]. In the framework of residual distribution schemes it has been reformulated by means of a technique reminiscent of flux limiting in the finite volume context : the nonlinear blending of a linear high order method with a linear low (first) order positive coefficient one. For example, blending (56) with a high order stabilized method would lead to

$$\boldsymbol{\phi}_i^K = \frac{1}{n_{dof}}\boldsymbol{\phi}^K + \delta(\mathbf{w}_h)\alpha_K(\mathbf{w}_i - \overline{\mathbf{w}}_K) + (\mathrm{Id} - \delta(\mathbf{w}_h))\theta_K|K| \sum_{x_{quad}} \omega_{quad}(\vec{a}\cdot\nabla\varphi_i(x_{quad}))\,\mathcal{T}_K\left(\vec{a}\cdot\nabla u_h(x_{quad})\right) \tag{65}$$

where different forms of the stabilization as selected depending on whether $\mathbf{w}_h$ is smooth, in which case $\delta(\mathbf{w}_h) \leq \mathcal{O}(h_K)$, or discontinuous, in which case $\mathrm{Id} - \delta(\mathbf{w}_h) \leq \mathcal{O}(h_K)$. More involved constructions considering replacing (56) in the blending by (63) have also been proposed e.g. in [99].

3.1.8 Handling source terms

The extension of the above framework to the approximation of solutions of

$$\mathrm{div}\ \mathbf{f}(\mathbf{w}) + \mathbf{s}(\mathbf{w}, \mathbf{x}) = 0 \tag{66}$$

is based on the inclusion of the source in the re-definition of the local element residual, leading to the requirement

$$\sum_{i\in K} \boldsymbol{\phi}_i^K = \oint_{\partial K} \mathbf{f}_h(\mathbf{w}_h)\cdot\mathbf{n}\, d\partial K + \int_K \mathbf{s}_h(\mathbf{w}_h, \mathbf{x})\, d\mathbf{x} \tag{67}$$

All of the methods described earlier can be extended to this more general setting.

Interesting results can be shown when $\mathbf{s}$ depends on some given data, say a given field $f(\mathbf{x})$:

$$\mathbf{s}(\mathbf{w}, \mathbf{x}) = \mathbf{s}(\mathbf{w}, f(\mathbf{x}))$$

This is the case in some environnemental applications (e.g. shallow water equations), or when considering the solution of the differential problem on a manifold (see e.g. [100] and references therein). Such problems often embed some particular solutions which are characterized by the existence of a set of invariants $\mathbf{v} = \mathbf{v}(\mathbf{w}, f)$ constant throughout the spatial domain. Assuming a sufficient smoothness of f, of the solution, and of the mapping $(\mathbf{v}, f) \mapsto \mathbf{w}(\mathbf{v}, f)$, in this case we can write that

$$\mathrm{div}\ \mathbf{f}(\mathbf{w}) = (\nabla_{\mathbf{w}}\mathbf{f}\nabla_{\mathbf{v}}\mathbf{w})\,\nabla\mathbf{v} + \underbrace{(\nabla_{\mathbf{w}}\mathbf{f}\nabla_f\mathbf{w})}_{\Lambda(\mathbf{v},f)}\nabla f$$

Solutions characterized by the invariance relation $\mathbf{v} = \mathbf{v}_0 = c^t\ \forall\mathbf{x}$, satisfy (cf. (66))

$$\Lambda(\mathbf{v}_0, f)\nabla f + \mathbf{s}(\mathbf{v}_0, f) = 0 \tag{68}$$

An interesting result concerning this class of solutions for schemes defined by

$$\boldsymbol{\phi}_i^K = \int_K \omega_i^K(\nabla\cdot\mathbf{f}_h(\mathbf{v}_h, f) + \mathbf{s}_h(\mathbf{v}_h, f)) \tag{69}$$

thus based on a direct approximation of the invariant states, instead of the conserved variables $\mathbf{w}$. The following is shown in [101, 99, 102].

Proposition 3.9 (Steady invariants and superconsistency)**.** *Under standard regularity assumptions on the mesh, provided that* (69) *is true for some test function* ω_i^K *which is uniformly bounded w.r.t. h,* $\mathbf{v}_h$*, element residuals, and w.r.t. to the data of the problem, then for exact integration the scheme defined by* (69) *preserves exactly the equilibrium* (68)*. For approximate integration, assuming that a flux quadrature exact for approximate polynomial fluxes of degree* p_f *is used, and a source quadrature exact for approximate polynomial sources of degree* p_v*, and assuming that* $f \in H^{p+1}$ *with* $\nabla f \in H^p$*, and* $p > \min(p_f, p_v)$*, then the scheme defined by* (69) *is super-cosnsistent w.r.t. solutions characterized by* (68)*, and in particular, its consistency is of order* $r = \min(p_f + 2, p_v + 3)$*.*

This result shows the potential of residual framework considered here in guaranteeing the balance of the flux divergence with complex source terms. Similar and stronger results, especially on simple particular solutions encountered in shallow water flows, have already been recalled in the previous chapter [61]. Some applications of this property will be shown in the results section.

3.1.9 Handling viscous terms

When considering the advection diffusion equation (with $\nabla \cdot \vec{a} = 0$):

$$\vec{a} \cdot \nabla u - \nabla \cdot (\mathcal{K}(u) \cdot \nabla u) = 0 \tag{70}$$

with $\mathcal{K}(u)$ a positive semi-definite diffusion matrix coefficient, the first idea is to look at it as a standard conservation relation with an enhanced flux, now

$$\mathbf{f}(u, \nabla u) = \vec{a}u - \mathcal{K}(u) \cdot \nabla u,$$

to which, one could apply the same construction as before. However, there is a fundamental difference: if the approximation of the solution is sought to be piecewise polynomial and globally continuous, its gradient will still be piecewise polynomial, but will not be globally continuous anymore. One of the fundamental requirement of the previous developments is that the flux on the boundary of the element is single valued. This cannot be anymore the case here unless something is done.

There are two ways of solving this issue. Both are similar what is done in LDG methods. The first step is in each case to rewrite the partial differential equation into a, possibly hyperbolic, first order system of PDEs. For the two-dimensional advection-diffusion equation, setting $\mathcal{K} = \nu\,\mathrm{Id}$, we consider the hyperbolic first order system

$$\begin{aligned} \frac{\partial \mathbf{u}}{\partial t} + \vec{a} \cdot \nabla u &= \nu \left(\partial_x p + \partial_y q\right) \\ \frac{\partial p}{\partial t} &= \frac{1}{T_r} \left(\partial_x u - p\right) \\ \frac{\partial q}{\partial t} &= \frac{1}{T_r} \left(\partial_y u - q\right) \end{aligned} \tag{71}$$

where p and q are the gradient variables and T_r is a relaxation time. At the steady state the system (71) is equivalent to the original equation (70), independently of the parameter T_r, and p,q become equivalent to the derivatives of the unknown. The idea of reformulating a parabolic problem with second order derivatives as a hyperbolic system such as e.g. (71), is not new, as it dates back to the work of Vertnotte and Cattaneo [103, 104] to study the heat equation. This idea has been efficiently exploited by H. Nishikawa and A. Mazaheri to construct schemes for the steady and time dependent diffusion, advection-diffusion, advetcion-diffuction and Navier-Stokes equations, see e.g. [105, 106, 107, 108] and [109, 110] for recent formulations of residual distribution and Discontinuous Galerkin based on this approach.

There are two ways to approach system (71). The first is to make use of its hyperbolicity, and reuse all the artillery already available. In this case, the overhead of having to introduce the gradient variables, can be compensated, by a careful design of the scheme which may guarantee the same accuracy for both the solution *and* its derivatives. This may have an impact on the computation of e.g. forces and heat fluxes and allow the use of coarser meshes to provide accurate values of these quantities. This is a path being followed in [109, 110], but the demonstration of its feasibility for practical application is still a work in progress.

Another way to exploit system (71) has been suggested in [107, 108], and developed from scalar advection diffusion up to laminar Navier-Stokes and RANS equations in [111, 112, 113]. In this alternative approach only the first discrete equation or u_h is kept. This is of course a discretization of (70), which however depends on values of p and q. These values are now replaced by an appropriate high order *reconstruction* of the solution derivatives starting from u_h. Simple solutions are possible, as e.g. the use of simple arithmetic averages for the viscous fluxes on element boundaries, see for example [114]. However, these simple choices lead to suboptimal accuracy, mostly because one order of accuracy is lost in the evaluation of the gradient. In [111], a systematic study of possible recovery methods (arithmetic average, least square, ect) has been conducted, and the best solution is to take into advantage, via a local least square minimization algorithm, of the existence of super convergence points in element. At these points, as put forward by Zienkiewiscz and Zhu [115], the gradient is approximated at full order. In the following, this reconstruction will be nicknamed as SPR-ZZ.

Example of a non-classical scheme. We want to show how to use these ideas to generalise the schemes of section §3.1.7 for the solution of (70). The schemes obtained are those used in the numerical results we will discuss later.

So we start from scheme (57), assuming for simplicity $\beta_i^K = 1/n_{dof}$. If we assume to be in the purely diffusive case, apply this scheme to system (71), and we only look at the local distributed residuals only for the first equation we have

$$\phi_i^{\text{fos}} = \frac{1}{n_{dof}}\phi^{K,\nu} + \int_K \tau_\nu \nabla\varphi_i \cdot (\nabla u_h - (p_h, q_h))$$

having assumed for the stabilization matrix $\mathcal{T}_K = \delta_\nu \,\mathrm{Id}$, and where

$$\phi^{K,\nu} = -\oint_{\partial K} \nu(p_h, q_h)\cdot \boldsymbol{n}$$

Note also that the effect of the relaxation time T_r has been embedded in the δ_ν coefficient.

The trick is now to replace the nodal values of the gradients p and q by accurately reconstructed ones, which we obtain with the SPR-ZZ procedure recalled above. The important part is the definition of the total residual. From the Lax-Wendroff theorem recalled in section §3.1.6 however we know that the numerical approximations of bot these fluxes must be edge-continuous. The simplest way to achieve that is to use for the viscous flux the finite element approximation based on the reconstructed nodal gradients. We denote this quantity by $\widetilde{\nabla u_h}$. So, for pure diffusion, the scheme is finally defined by

$$\phi_i^{K,\nu} = \frac{1}{n_{dof}}\phi^{K,\nu} + \int_K \delta_\nu \nabla\varphi_i \cdot \left(\nabla u_h - \widetilde{\nabla u_h}\right)$$

where

$$\phi^{K,\nu} = -\oint_{\partial K} \nu\widetilde{\nabla u_h}\cdot \boldsymbol{n}$$

and with δ_ν having the dimensions of a diffusion coefficient.

For advection diffusion, we can apply the same procedure. Starting with the total residual

$$\phi^K = \oint_{\partial K} \left(\vec{a}\, u_h - \nu\widetilde{\nabla u_h}\right)\cdot \boldsymbol{n}\,\mathrm{d}\partial\Omega,$$

we can deduce from the first order system formulation two types of regularization terms leading to local nodal residual[5]

$$\phi_i^K = \frac{1}{n_{dof}}\phi^K + \int_K \tau_a \vec{a}\cdot\nabla\varphi_i\,(\vec{a}\cdot\nabla u_h - \nu\nabla\cdot\nabla u_h) + \int_K \delta_\nu\nabla\varphi_i\cdot\left(\nabla u_h - \widetilde{\nabla u_h}\right)$$

The optimal choice of the scaling parameters τ_a and δ_ν has been shown to require some dependence on the elemental Re number $Re = \frac{\|\boldsymbol{\lambda}\|h}{\nu}$, see for example [111, 106, 116] and references therein. This is taken into account by setting

$$\begin{aligned}\phi_i^K = \frac{1}{n_{dof}}\phi^K &+ \xi(Re)\int_K \left(\vec{a}\cdot\nabla\varphi_i\right)\tau\left(\vec{a}\cdot\nabla u_h - \nabla\cdot(\nu\nabla u_h)\right)\mathrm{d}\Omega \\ &+ \left(1-\xi(Re)\right)\int_K \frac{\nu\delta}{2}\left(\nabla u_h - \widetilde{\nabla u_h}\right)\cdot\nabla\varphi_i\,\mathrm{d}\Omega,\end{aligned} \tag{72}$$

where the function $\xi(Re)$ is such that $\xi(Re)\to 0$ in the diffusion limit ($Re\to 0$) and $\xi(Re)\to 1$ in the advection limit ($Re\to\infty$).

To account for non-smooth solutions, one can use the same technique discussed in section §3.1.7, replace the centered contribution by a nonlinear limited residual, and pre-multiply the stabilisation

[5] having set here $\mathcal{T}_K = \mathrm{diag}(\tau_a, \delta_\nu, \delta_\nu)$

terms by some smoothness sensor, so that the scheme can be generally written in the finale general form

$$\begin{aligned}\phi_i^K =&\beta_i^K \phi^K \\ &+\theta(u_h)\xi(Re)\int_K \Big(\vec{a}\cdot\nabla\varphi_i\Big)\,\tau\,\Big(\vec{a}\cdot\nabla u_h - \nabla\cdot(\nu\nabla u_h)\Big)\,\mathrm{d}\Omega \\ &+\theta(u_h)\Big(1-\xi(Re)\Big)\int_K \frac{\nu\delta}{2}\,\Big(\nabla u_h - \widetilde{\nabla u_h}\Big)\cdot\nabla\varphi_i\,\mathrm{d}\Omega,\end{aligned} \tag{73}$$

where β_i^K is computed following the limiting procedure discussed in section 3.1.7 in the non-smooth case.

The numerical scheme obtained for the advection-diffusion scalar equation is then extended to the compressible Navier-Stokes equations. The governing equations read

$$\frac{\partial \mathbf{w}}{\partial t} + \boldsymbol{\nabla}\cdot\mathbf{f}^a(\mathbf{w}) - \boldsymbol{\nabla}\cdot\mathbf{f}^v(\mathbf{w},\boldsymbol{\nabla}\mathbf{w}) = 0,$$

where $\mathbf{w}$ and $\mathbf{f}^a(\mathbf{w})$ are the vector of the conservative variables and the advective flux function, respectively, as defined for the Euler equations, while $\mathbf{f}^v(\mathbf{w},\boldsymbol{\nabla}\mathbf{w}) = (\,\mathsf{f}_x^v,\,\mathsf{f}_y^v\,)^{\mathrm{T}}$ is the viscous flux function

$$\mathsf{f}_x^v(\mathbf{w},\boldsymbol{\nabla}\mathbf{w}) = \begin{pmatrix} 0 \\ \tau_{xx} \\ \tau_{xy} \\ u\tau_{xx} + v\tau_{xy} - q_x \end{pmatrix}, \quad \mathsf{f}_y^v(\mathbf{w},\boldsymbol{\nabla}\mathbf{w}) = \begin{pmatrix} 0 \\ \tau_{xy} \\ \tau_{yy} \\ u\tau_{xy} + v\tau_{yy} - q_y \end{pmatrix},$$

where

$$\tau_{xx} = \mu\left(\frac{4}{3}\frac{\partial v_x}{\partial x} - \frac{2}{3}\frac{\partial v_y}{\partial y}\right), \quad \tau_{yy} = \mu\left(\frac{4}{3}\frac{\partial v_y}{\partial y} - \frac{2}{3}\frac{\partial v_x}{\partial x}\right), \quad \tau_{xy} = \tau_{yx} = \mu\left(\frac{\partial v_x}{\partial y} + \frac{\partial v_y}{\partial x}\right)$$

are the components of the stress tensor, with μ the dynamic viscosity of the fluid, and q_x, q_y are the components of the heat flux $\boldsymbol{q}$ which is defined as

$$\boldsymbol{q} = k\boldsymbol{\nabla}T,$$

where T is the temperature and k is the thermal conductivity coefficient. It is well know that the viscous flux function $\mathbf{f}^v$ is homogeneous with respect to the gradient of the conservative variable $\boldsymbol{\nabla}\mathbf{w}$

$$\mathbf{f}^v(\mathbf{w},\boldsymbol{\nabla}\mathbf{w}) = \mathbb{K}(\mathbf{w})\boldsymbol{\nabla}\mathbf{w},$$

with the homogeneity tensor $\mathbb{K}(\mathbf{w}) = \dfrac{\partial \mathbf{f}^d}{\partial \mathbf{w}}$.

The discretization of the Navier-Stokes equations is straightforward. The total residual on a generic element K is given by

$$\boldsymbol{\Phi}^K = \oint_{\partial K}\Big(\mathbf{f}^a(\mathbf{w}) - \mathbb{K}(\mathbf{w})\widetilde{\boldsymbol{\nabla}\mathbf{w}}\Big)\cdot\boldsymbol{n},$$

with $\widetilde{\boldsymbol{\nabla}\mathbf{w}}$ the reconstructed gradient of the conservative variables and the boundary integral computed by the means of a quadrature rule. The total residual is first distributed to all the DOF of the element using the low order Rusanov scheme, subsequently the limitation procedure is applied to obtain an high order residual as described in the section 4.2.1. In the last step the filtering term is added together with the dumping term acting for the viscous part. The complete scheme reads

$$\begin{aligned}\boldsymbol{\Phi}_i^K = \tilde{\boldsymbol{\Phi}}_i^K &+ \xi(Re)\int_K \Big(\boldsymbol{A}\cdot\boldsymbol{\nabla}\varphi_i\Big)\,\boldsymbol{\tau}\,\Big(\boldsymbol{A}\cdot\boldsymbol{\nabla}\mathbf{w}_h - \boldsymbol{\nabla}\cdot\big(\mathbb{K}\boldsymbol{\nabla}\mathbf{w}_h\big)\Big)\,\mathrm{d}\Omega \\ &+ \Big(1-\xi(Re)\Big)\int_E \tfrac{1}{2}\mathbb{K}\,\Big(\boldsymbol{\nabla}\mathbf{w}_h - \widetilde{\boldsymbol{\nabla}\mathbf{w}}_h\Big)\cdot\boldsymbol{\nabla}\varphi_i\,\mathrm{d}\Omega.\end{aligned} \tag{74}$$

with $\tilde{\boldsymbol{\Phi}}_i^K$ denoting the (unfiltered) centered or nonlinear distribution.

3.2 Time dependent problems

In this chapter, we consider the approximation of time dependent solutions to a system of conservation laws reading

$$\partial_t \mathbf{w} + \nabla \cdot \mathbf{f}(\mathbf{w}) = 0 \quad \text{on} \quad \Omega \times [0\,, \mathrm{T_{fin}}] \subset \mathbb{R}^d \times \mathbb{R}^+ \tag{75}$$

As shown in [117], and then in [118, 119], and in [120, 121, 122, 123] (cf. also the chapter [61]), to obtain high order schemes for this case, one must carefully design a coupling between the stencil used to approximate the integral of the time derivative and the flux divergence. Some approaches to obtain this coupling are recalled, and a more general prototype is analyzed. The links with other methods are briefly recalled. The first part of the section is devoted to fully implicit methods. We then discuss a path allowing the contruction of explicit apporaches which do-not require the inversion of a mass matrix, or for which this matrix reduces to the symmetric positive-definite Galerkin one.

Note that, beside classical stabilized finite element schemes (SUPG, GLS, etc), here the status of residual distribution type methods is less advanced. Here we discuss some of the most interesting ideas toward generalizing the methods presented for steady state. Some research directions to push the limits of the existing construtions will be discussed later.

3.2.1 Implicit prototype for time dependent solutions

We introduce the time discretized version of (75) by means of an $r+1$th order time integration scheme

$$\Gamma^{n+1}(\mathbf{w}) = \sum_{i=0}^{p} \alpha_i \frac{\delta \mathbf{w}^{n+1-i}}{\Delta t} + \sum_{j=0}^{q} \theta_j \nabla \cdot \mathbf{f}^{n+1-j} \tag{76}$$

where $\Delta t = \min_n (t^{n+1} - t^n)$, with $\Delta t^{n+1} = t^{n+1} - t^n$, with $\delta \mathbf{w}^{n+1} = \mathbf{w}^{n+1} - \mathbf{w}^n$, $\mathbf{f}^{n+1-j} = \mathbf{f}^{n+1-j}(\mathbf{w}^{n+1-j})$, and with the α_i and θ_j coefficients given by a time integration scheme of choice. This may be a generic stage of a multi-stage method, or a multi-step scheme. Space time schemes can be embedded in the analysis that follows by appropriate definitions of the α_is, and of the $\delta \mathbf{w}^{n+1-i}$ to embed eventually jumps in the time direction, when using discontinuous in time space-time elements. An important assumption, is that the time stepping verifies the conservation identity

$$\sum_{n=0}^{N} \sum_{i=0}^{p} \alpha_i \delta \mathbf{w}^{n+1-i} = \mathbf{w}^N - \mathbf{w}^0 = \mathbf{w}(\mathrm{T_{fin}}) - \mathbf{w}_0 \tag{77}$$

We set on every $K \in \Omega_h$

$$\Phi^K = \int_K \Gamma^{n+1}(\mathbf{w}_h) = \int_K \left(\sum_{i=0}^{p} \alpha_i \frac{\delta \mathbf{w}_h^{n+1-i}}{\Delta t} + \sum_{j=0}^{q} \theta_j \nabla \cdot \mathbf{f}_h^{n+1-j} \right) \tag{78}$$

with $\mathbf{w}_h$ and $\mathbf{f}_h$ continuous finite element polynomial approximations of degree k and (at least) k respectively. Similarly, on each boundary face f we set

$$\boldsymbol{\phi}^f = \int_f \sum_{j=0}^{q} \theta_j (\hat{\mathbf{g}} - \mathbf{f}_h)^{n+1-j} \cdot \vec{n} \tag{79}$$

with $\hat{\mathbf{g}}$ a numerical flux consistent with the BCs.

Similarly to the previous sections, we consider the scheme that computes $\mathbf{w}_h$ as the solution of

$$\sum_{K \in K_i} \Phi_i^K + \sum_{f \in F_i} \boldsymbol{\phi}_i^f = 0 \tag{80}$$

where $\forall\, K$ and $\forall\, f$

$$\sum_{j \in K} \Phi_j^K = \Phi^K \quad \text{and} \quad \sum_{j \in f} \boldsymbol{\phi}_j^f = \boldsymbol{\phi}^f \tag{81}$$

Consistency analysis

To begin with, we generalize the consistency conditions. To simplify the notation we consider the scalar case, and we neglect the boundary conditions, which can be easily embedded in the spatial operator as shown.We will assume some classical regularity properties for the mesh and the time stepping strategy, namely

$$C_0 \leq \sup_{K\in\Omega_h} \frac{h^2}{|K|} \leq C_1\,, \quad C_0' \leq \frac{\Delta t}{h} \leq C_1' \tag{82}$$

Let now $\mathbf{w} \in C^{l+1}$ be an exact classical solution of (75), with $l \geq \max(r,k)$, and such that

$$\sum_{i=0}^{p} \alpha_i \frac{\delta \mathbf{w}^{n+1-i}}{\Delta t} + \sum_{j=0}^{q} \theta_j \nabla\cdot \mathbf{f}^{n+1-j} = \partial_t \mathbf{w} + \nabla\cdot\mathbf{f} + \mathcal{O}(\Delta t^{r+1}) \tag{83}$$

We denote by $\mathbf{w}_h^m$, the kth degree continuos finite element projection/interpolation of $\mathbf{w}^m$.

Consider now $\psi \in C_0^1(\Omega\times[0,\mathrm{T_{fin}}])$, a smooth test function with $\psi|_{\partial\Omega}=0$. Let ψ_h be its k degree polynomial finite element projection/interpolation, with ψ_i^n the corresponding values at the choses degrees of freedom. It is also assumed that [68, 69] there exist constants C_0'', C_1'', C_2 such that

$$\begin{aligned} \|\partial_t\psi_h\|_{L^\infty(\Omega_h)} \leq C_0''\,, \quad \|\psi_h(t+\Delta t)-\psi_h(t)\|_{L^\infty(\Omega_h)} \leq C_0''\Delta t \\ \|\psi_h\|_{L^\infty(\Omega_h)} \leq C_1''\,, \quad |\psi_i-\psi_j| \leq \|\nabla\psi_h\|_{L^\infty(\Omega_h)}\, h \leq C_2 h \end{aligned} \tag{84}$$

We define the following truncation error for scheme (80)

$$\epsilon(\mathbf{w}_h,\psi) := \sum_{n=0}^{N}\sum_{i\in\Omega_h} \Delta t^{n+1}\psi_i^{n+1} \sum_{K\in K_i} \Phi_i^K(\mathbf{w}_h) = \sum_{n=0}^{N}\sum_{K\in\Omega_h}\sum_{i\in K}\int_{t^n}^{t^{n+1}} \psi_i^{n+1}\Phi_i^K(\mathbf{w}_h) \tag{85}$$

We introduce the Galerkin splitting in space

$$\Phi_i^{\mathrm{G}} = \int_K \varphi_i \Gamma^{n+1}$$

and note that

$$\sum_{j\in K}(\Phi_j^K - \Phi_j^{\mathrm{G}}) = 0$$

This allows to recast the error as

$$\epsilon(\mathbf{w}_h,\psi) = \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}} \left\{ \int_{\Omega_h} \psi_h^{n+1}\Gamma^{n+1}(\mathbf{w}_h) + \frac{1}{\mathrm{C}_K}\sum_{K\in\Omega_h}\sum_{i,j\in K}(\psi_i-\psi_j)(\Phi_i^K-\Phi_i^{\mathrm{G}}) \right\} \tag{86}$$

Multiplying (83) by ψ_h and integrating over space and time we can get

$$\sum_{n=0}^{N}\int_{t^n}^{t^{n+1}} \int_{\Omega_h} \psi_h^{n+1}\Gamma^{n+1}(\mathbf{w}) = \sum_{n=0}^{N} \Delta t\mathcal{O}(\Delta t^{r+1}) = \mathcal{O}(\Delta t^{r+1})$$

So the error can be estimated as follows

$$\begin{aligned} \epsilon(\mathbf{w}_h,\psi) &= \mathrm{I} + \mathrm{II} + \mathrm{III} + \mathcal{O}(\Delta t^{r+1}) \\ \mathrm{I} &= \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}} \int_{\Omega_h} \psi_h^{n+1} \sum_{i=0}^{p}\alpha_i \frac{\delta(\mathbf{w}_h-\mathbf{w})^{n+1-i}}{\Delta t} \\ \mathrm{II} &= \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}} \sum_{j=0}^{q}\int_{\Omega_h} \psi_h^{n+1}\nabla\cdot(\mathbf{f}_h-\mathbf{f})^{n+1-j} \\ \mathrm{III} &= \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}} \frac{1}{\mathrm{C}_K}\sum_{K\in\Omega_h}\sum_{i,j\in K}(\psi_i-\psi_j)(\Phi_i^K-\Phi_i^{\mathrm{G}}) \end{aligned}$$

Estimating each of the terms we get to the conditions of the cell and boundary splittings allowing to preserve the $\mathcal{O}(\Delta t^{r+1})$ appearing on the right hand side. This is readily done by using the hypotheses on the regularity of u and standard interpolation results [68, 69]. In particular, for term I we can use hypothesis (77) to write

$$\begin{aligned}
\mathrm{I} = \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}}\int_{\Omega_h}\sum_{i=0}^{p}\alpha_i\frac{\delta(\psi_h\mathbf{w}_h-\psi_h\mathbf{w})^{n+1-i}}{\Delta t} + \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}}\int_{\Omega_h}\sum_{i=0}^{p}\alpha_i(\psi_h^{n+1}-\psi_h^{n-i+1/2})\frac{\delta(\mathbf{w}_h-u)^{n+1-i}}{\Delta t} \\
-\sum_{n=0}^{N}\int_{t^n}^{t^{n+1}}\int_{\Omega_h}(\mathbf{w}_h-u)^{n-i+1/2}\sum_{i=0}^{p}\alpha_i\frac{\delta(\psi_h)^{n+1-i}}{\Delta t} = \\
\int_{\Omega_h}(\psi_h(\mathbf{w}_h-u))(\mathrm{T_{fin}}) - \int_{\Omega_h}(\psi_h(\mathbf{w}_h-u))_0 + \sum_{n=0}^{N}\int_{t^n}^{t^{n+1}}\int_{\Omega_h}\sum_{i=0}^{p}\alpha_i(\psi_h^{n+1}-\psi_h^{n-i+1/2})\frac{\delta(\mathbf{w}_h-u)^{n+1-i}}{\Delta t} \\
-\sum_{n=0}^{N}\int_{t^n}^{t^{n+1}}\int_{\Omega_h}(\mathbf{w}_h-u)^{n-i+1/2}\sum_{i=0}^{p}\alpha_i\frac{\delta(\psi_h)^{n+1-i}}{\Delta t}
\end{aligned}$$

Using (84), and the regularity of u, we can now bound this term as

$$|\mathrm{I}| = \mathcal{O}(h^{k+1}) + C\frac{\mathrm{T_{fin}}}{\Delta t}\,\Delta t\,\mathcal{O}(h^{k+1})C_0''\sup_{i=1,p}|\alpha_i| = \mathcal{O}(h^{k+1})$$

The analysis of the remaining terms is practically identical to the one of section §3.1.1, and omitted for brevity (the interested reader can refer to [102] for details). The final result is the following.

Proposition 3.10 (Accuracy of RD, unsteady case)**.** *Under assuption* (82) *on the time stepping, given a $k+1$th order continuous polynomial approximation of the unknown ad of the fluxes, and a $r+1$th order accurate time integration scheme, scheme* (80) *verifies the truncation error estimate*

$$|\epsilon(\mathbf{w}_h,\psi)| \leq \mathcal{O}(h^{p+1}), \quad p=\min(k,\,r)$$

provided that

$$\sup_{K\in\Omega_h}\sup_{i\in K}|\Phi_i^K(\mathbf{w}_h)| = \mathcal{O}(h^{p+d}) \tag{87}$$

whenever $\mathbf{w}_h$ is the finite element projection/interpolation of a smooth exact solution. In this case we say that the scheme is $p+1$th order accurate.

Moreover we have the following estimate.

Lemma 3.11 (Consistency estimate, time dependent case)**.** *Under the hypotheses of proposition 3.10 the following consistency estimates hold.*

$$\Gamma^{n+1}(\mathbf{w}_h) = \mathcal{O}(h^k) + \mathcal{O}(\Delta t^{r+1}), \quad \Phi^K(\mathbf{w}_h) = \mathcal{O}(h^{p+d}) \tag{88}$$

Proof. The proof is easily obtained by considering that due to (83)

$$\Gamma^{n+1}(\mathbf{w}_h) = \mathcal{O}(\Delta t^{r+1}) + \Gamma^{n+1}(\mathbf{w}_h) - \Gamma^{n+1}(\mathbf{w})$$

By its definition, and under the hypotheses made, one easily checks that $\Gamma^{n+1}(\mathbf{w}_h) - \Gamma^{n+1}(\mathbf{w}) = \mathcal{O}(h^k)$. The estimate on $\Phi^K(\mathbf{w}_h)$ is trivially obtained upon integration of Γ^{n+1}.

□

As a consequence we have the following corollary.

Corollary 3.12 (High order residual schemes)**.** *Under the hypotheses of proposition 3.10, a sufficient condition for a scheme of the form* (80) *to be $p+1$th order accurate if there exist a test function ω_i uniformly bounded w.r.t. h, $\mathbf{w}_h$, $\Gamma^{n+1}(\mathbf{w}_h)$, and w.r.t the data of the problem, such that*

$$\Phi_i^K(\mathbf{w}_h) = \int_K \omega_i\Gamma^{n+1}(\mathbf{w}_h) \tag{89}$$

Examples of implicit high order schemes

Typical examples of high order methods are obtained with the natural extension to the time dependent case case of SUPG-type methods (see e.g. [124, 125, 74, 126] and references therein). Some notable examples of less classical high order schemes exploit (89) with $\omega_i = \beta_i^K$ constant per element. The fist of such *accuracy preserving* schemes are found in the work of D. Caraeni [118, 119], up to third order of accuracy for the Navier-Stokes equations, and more recently in [127] where the third order scheme of Caraeni has been blended with a monotone one via a FCT procedure to provide oscillation free high order solutions of the compressible Euler equations.

Other non-classical constructions have tried to exploit the similarities between stabilized finite elements, and RD methods with constant distribution coefficients. The objective of these works is to find clever definitions of mass matrices/test functions guaranteeing the satisfaction of t (89). This was the initial idea behind the work of Maerz and Ferrante at the von Karman Institute [120, 121] later pursued first in [122, 128], and [129, 130, 123, 131, 96] (see also [61]). This has provided interesting results, but so far only for second order methods.

Finally, examples of space time RD schemes up to third order are discussed in [132, 133] with monotonicity preserving extensions discussed in [122, 134, 123, 135].

All these works, use almost exactly the same techniques developed for steady problems, either treating the time derivative as a source term, or as an additional space direction. The potential of these methods is that they may allow preservation of monotonicity unconditionally w.r.t the time step size, which is very interesting when considering local mesh refinement (see e.g. [135, 136]), or stiff problems (viscous terms, chemical reactions etc.). The drawback of this formulation, is that the nonlinear stabilization involved depends on the unknown solution at the new time level, thus ruling out *a-priori* simpler genuinely explicit time marching methods, often preferred in the hyperbolic case. Some exceptions to this rule exist, as for example Taylor-Galerkin, and Lax-Wendroff type methods which can also be recast in a residual distribution formalism (see e.g. [137, 138, 139], and [61]).

A technique to side step this issue, and construct some non-classical genuinely explicit monotone and high order residual methods is discussed in the next section.

3.2.2 Genuinely explicit time advancement for residual methods

The main idea here is to start from a prototype high order scheme, which we will write in general as (boundary conditions are neglected for simplicity)

$$\int_{\Omega_h} \omega_i(\mathbf{w}_h)\left(\partial_t \mathbf{w}_h + \nabla\cdot\mathbf{f}_h(\mathbf{w}_h)\right) + \sum_{K\in\Omega_h}\oint_{\partial K}\gamma^{\partial K}(\mathbf{w}_h)[\nabla\mathbf{w}_h]\cdot[\nabla\varphi_i] = 0$$

with $[\cdot]$ a jump of a quantity, as in (16c). The weight ω_i in the first term is better expressed as a composition of local restrictions $\omega_i = \sum_K \omega_i^K$, and depends on the specific method chosen. For SUPG-type schemes we can write $\omega_i^K = \varphi_i^K + \gamma_i^K(\mathbf{w}_h)$ with the first term only depending on the mesh. For other methods such as RD schemes, similar decompositions may be invoked, however these are not unique [97]. Other definitions can be obtained by considering Variational Multi-Scale stabilization techniques, or bubble functions (see [74] for a review). The last term in the method is one of the possible forms of an edge stabilization [85, 86]. Due to the presence of the $\partial_t\mathbf{w}_h$ term in the residual $r(\mathbf{w}_h)$, and to the continuity of the approximation, the first term will led to a global mass matrix in the resulting system of ODEs. This matrix in general depends on the discrete solution $\mathbf{w}_h$, and, in the case of RD schemes, is not uniquely defined, nor guaranteed to be invertible [97].

The first idea to simplify things, comes originally from [97], and requires the introduction of some discrete approximation of the ODE system. As done before, we consider a semi-discretization in time, and the semi-discrete residual which we write here as

$$\Gamma^{n+1} = \Gamma^{n+1}(\mathbf{w}_h^{n+1};\{\mathbf{w}_h^{(s)}\}) = \alpha_{-1}\mathbf{w}_h^{n+1} + \sum_{s=0}^{S}\alpha_s\mathbf{w}_h^{(s)} + \Delta t\sum_{s=0}^{S}\theta_s\nabla\cdot\mathbf{f}_h(\mathbf{w}_h^{(s)}) \tag{90}$$

with the $\mathbf{w}^{(s)}$ values being either those computed from previous time steps (multi-step scheme) or from some previous predictor stages (multi-stage). Note that the two summations on the right hand side

are independent on the unknown $\mathbf{w}^{n+1}$. As before, for a rth order accurate method in time, the local truncation error relation will be of the type $\Gamma^{n+1} = \mathcal{O}(\Delta t^{r+1}) = \mathcal{O}(h^{r+1})$, if (82) hold, as is the case always for explicit time schemes. If we proceeded as in the last section, we would plug Γ^{n+1} in the spatial discretization, and the term $\alpha_{-1}\mathbf{w}_h^{n+1}$ would lead to the inversion of a (nonlinear) mass matrix. However, in [97] it is proved that *given a kthe order accurate approximation in space, and a rth order accurate approximation in time, provided that the ratio $\Delta t/h$ is uniformly bounded, the space-time discretization*

$$\int_{\Omega_h} \varphi_i(\Gamma^{n+1}(\mathbf{w}_h^{n+1};\{\mathbf{w}_h^{(s)}\}) - \tilde{\Gamma}^{n+1}(\{\mathbf{w}_h^{(s)}\})) = -\int_{\Omega_h} \omega_i(\mathbf{w}_h)\tilde{\Gamma}^{n+1}(\{\mathbf{w}_h^{(s)}\}) \\ -\Delta t\sum_{s=0}^{S}\beta_s \sum_{K\in\Omega_h}\oint_{\partial K}\gamma^{\partial K}(\mathbf{w}_h^{(s)})[\nabla\mathbf{w}_h]\cdot[\nabla\varphi_i]$$

verifies a truncation error/consistency estimate of the type $\epsilon = \mathcal{O}(h^p)$, with $p = \min(k+1, r+1)$, provided that for a smooth exact solution, the modified semi-discrete residual $\tilde{\Gamma}^{n+1}$ verifies the consistency estimate

$$\tilde{\Gamma}^{n+1} = \mathcal{O}(h^{p-1})$$

The first practical use of this reduced consistency requirement for $\tilde{\Gamma}^{n+1}$ is to modify a given time discretization to obtain residual expressions one order lower. For example, for the classical third order RK3 method one has [97]

$$\text{first step} \quad \begin{cases} \Gamma_{RK3}^{(1)} = \mathbf{w}^{(1)} - \mathbf{w}^n + \nabla\cdot\mathbf{f}(\mathbf{w}^n) \\ \tilde{\Gamma}_{RK3}^{(1)} = \nabla\cdot\mathbf{f}(\mathbf{w}^n) \end{cases}$$

$$\text{second step} \quad \begin{cases} \Gamma_{RK3}^{(2)} = \mathbf{w}^{(2)} - \mathbf{w}^n + \dfrac{\Delta t}{4}(\nabla\cdot\mathbf{f}(\mathbf{w}^n) + \nabla\cdot\mathbf{f}(\mathbf{w}^{(1)})) \\ \tilde{\Gamma}_{RK3}^{(2)} = \dfrac{\mathbf{w}^{(1)} - \mathbf{w}^n}{2} + \dfrac{\Delta t}{4}(\nabla\cdot\mathbf{f}(\mathbf{w}^n) + \nabla\cdot\mathbf{f}(\mathbf{w}^{(1)})) \end{cases}$$

$$\text{final step} \quad \begin{cases} \Gamma_{RK3}^{n+1} = \mathbf{w}^{n+1} - \mathbf{w}^n + \dfrac{\Delta t}{6}(\nabla\cdot\mathbf{f}(\mathbf{w}^n) + 4\nabla\cdot\mathbf{f}(\mathbf{w}^{(2)}) + \nabla\cdot\mathbf{f}(\mathbf{w}^{(1)})) \\ \tilde{\Gamma}_{RK3}^{n+1} = 2(\mathbf{w}^{(2)} - \mathbf{w}^n) + \dfrac{\Delta t}{6}(\nabla\cdot\mathbf{f}(u^n) + 4\nabla\cdot\mathbf{f}(\mathbf{w}^{(2)}) + \nabla\cdot\mathbf{f}(\mathbf{w}^{(1)})) \end{cases}$$

For the extrapolated backward differencing method (eBDf3) one finds [140, 141]

$$\Gamma_{eBDf3}^{n+1} = \frac{11}{6}\mathbf{w}^{n+1} - 3\mathbf{w}^n + \frac{3}{2}\mathbf{w}^{n-1} - \frac{1}{3}\mathbf{w}^{n-2} + \Delta t(3\nabla\cdot\mathbf{f}(\mathbf{w}^n) - 3\nabla\cdot\mathbf{f}(\mathbf{w}^{n-1}) + \nabla\cdot\mathbf{f}(\mathbf{w}^{n-2}))$$
$$\tilde{\Gamma}_{eBDf3}^{n+1} = \frac{5}{2}\mathbf{w}^{n} - 4\mathbf{w}^{n-1} + \frac{3}{2}\mathbf{w}^{n-2} + \Delta t(3\nabla\cdot\mathbf{f}(u^n) - 3\nabla\cdot\mathbf{f}(\mathbf{w}^{n-1}) + \nabla\cdot\mathbf{f}(\mathbf{w}^{n-2}))$$

Equation (90), can be also seen as a defect correction method, in which a lower order residual is used as a means of approximating solutions of a high order one.

Note however, that relation (90) still requires the inversion of the Galerkin mass matrix which, even though symmetric positive-definite, is not an inverse monotone matrix. This may destroy all the efforts made in the construction of a shock capturing mechanism in the method. The solution is to constrain the choice of finite element spaces to those allowing to lump this matrix. Several choices exist, either based on standard Lagrange elements on a cubature grids with strictly positive cubature weights [142, 143, 144, 145], or using non-Lagrange elements having a similar property, as the Bezier basis proposed in [66] (see also [146] chapter 5). Whatever the choice, this approach leads to a fully explicit space-time discretization reading

$$|\mathcal{V}_i|\left(\alpha_{-1}\mathbf{w}_i^{n+1} + \sum_{s=0}^{S}\tilde{\alpha}_s\mathbf{w}_i^{(s)}\right) = -\int_{\Omega_h}\omega_i(\mathbf{w}_h)\tilde{\Gamma}^{n+1}(\{\mathbf{w}_h^{(s)}\}) \\ -\Delta t\sum_{s=0}^{S}\theta_s\sum_{K\in\Omega_h}\oint_{\partial K}\gamma^{\partial K}(\mathbf{w}_h^{(s)})[\nabla\mathbf{w}_h]\cdot[\nabla\varphi_i]$$

with $|\mathcal{V}_i|$ a nodal volume depending on the areas of the surrounding elements and on the quadrature weights induced by the finite element basis, and with the $\tilde{\alpha}_s$ obtained from the "defect-correction" in time $\Gamma - \tilde{\Gamma}$.

This construction provides genuinely explicit variants of all well known stabilized continuous finite elements (SUPG, GLS, VMS, etc), as well as of nonlinear residual distribution schemes discussed in this chapter. Thorough numerical validations have been reported in [97, 99, 140, 141]. Some examples will be provided in the comping sections.

4 Applications

4.1 Scalar examples

We start with a few scalar convergence tests, to check some of the theoretical aspects discussed in the chapter. consider the approximation of solutions of the steady scalar advection equation (18) on the domain $\Omega = [0,\, 1]^2$, with $\vec{a} = (0,\, 1)$, and with inlet condition $u(x, 0) = \sin^2(\kappa\pi x)$.

We start by a result taken from [91]. The test aims at verifying the analysis of section §3.1.7. The grid convergence has been run for $\kappa = 1$ with the nonlinear LLFs scheme (64), and with different evaluation strategies for the streamline dissipation, or *filtering* term. In particular, the discrete term in (59) is taken as the arithmetic average of its value in a certain set of points. Note that, with the exceptions of linear polynomials, this evaluation does not give in general any k_{extact} quadrature formula. Table 1 shows the impact of under evaluating this term. For a $\mathbb{P}^2$ finite element approximation, first order of accuracy us obtained, unless a three points stencil is used. Similarly, for $\mathbb{P}^3$ finite element approximation, a stencil of at least 6 points is required. Provided that the number of points is large enough, we see that indeed we recover the expected second, third and fourth order rates, even though the expressions used to evaluate the streamline dissipation are not obtained form a high order quadrature formula.

	$k=1$ filter : $\mathbb{P}^0$ dof	$k=2$ filter : $\mathbb{P}^0$ dof	$k=2$ filter : $\mathbb{P}^1$ dof	$k=3$ filter : $\mathbb{P}^1$ dof	$k=3$ filter: $\mathbb{P}^2$ dof
h	L^2	L^2	L^2	L^2	L^2
1/25	0.50493E-02	0.25122E-01	0.32612E-04	2.17274E-02	0.12071E-05
1/50	0.14684E-02	0.12935E-01	0.48741E-05	1.13486E-02	0.90642E-07
1/100	0.41019E-03	0.83978E-02	0.66019E-06	5.83347 E-03	0.53860E-08
average rate	1.790	0.7904	2.812	0.9292	3.914

Table 1: Scalar advection : grid convergence for the LLFs scheme (64). Verification of the analysis of section §3.1.7: impact of the number of evaluation points for the "filtering term" (from [91]).

The next example, taken from [147] (see also [94, 93]) aims at verifying the convergence rates obtained with the "variable β" LDA (61). Polynomial approximations up to degree $k = 7$ are tested using meshes with roughly the same number of degrees of freedom in all cases (from ≈ 2000 for the coarsest mesh to $\approx 32,000$ for the finest). The simulations are run with $\kappa = 5$. The results, summarized on table 2, show that indeed the method converges with a rate in between $k + 1/2$ and $k + 1$. For $k = 6$ and $k = 7$, converging results have only been obtained by using the optimized collocation of the degrees of freedom based on the warp and blend procedure discussed in [148]. Computations on standard Lagrange elements with equally spaced degrees of freedom did not converge for $k > 5$.

4.2 External aerodynamics

In this section, we report a couple of results from [92] for compressible fluids without viscous effect (Euler equations) and [67, 149] for the Navier Stokes case. The interested reader may consult [150] for informations and results for the turbulent case (Spalart and Allmaras model).

k	N_{dof}	h	ϵ_{L^2}	rate
1	2094	0.02185	3.49E-02	–
	8124	0.01109	7.44E-03	2.24
	32546	0.00554	1.36E-03	2.46
2	2189	0.02137	1.37E-02	–
	8217	0.01103	2.19E-03	2.65
(equi-spaced)	32181	0.00557	3.04E-04	2.88
3	2113	0.02175	5.20E-03	–
	8347	0.01095	2.94E-04	4.16
(equi-spaced)	33520	0.00546	2.11E-05	3.89
4	2017	0.02227	2.57E-03	–
	8593	0.01079	9.94E-05	4.71
(equi-spaced)	32553	0.00554	4.28E-06	4.55
5	2381	0.02049	1.15E-03	–
	8611	0.01078	2.83E-05	5.36
(equi-spaced)	33546	0.00546	5.32E-07	5.75
6	2317	0.02077	6.68E-04	–
	8293	0.01098	7.30E-06	7.01
(warp-blend)	33073	0.00550	7.29E-08	6.67
7	2633	0.01949	3.82E-04	–
	9430	0.01030	2.44E-06	7.92
(warp-blend)	34427	0.00539	9.86E-09	8.51

Table 2: Scalar advection : convergence for the variable β LDA (61) (courtesy of M. Vymazal [147], see also [94, 93]).

4.2.1 Euler equations

Method: from scalar to systems So far, we have only dealt with scalar problems: the computation of the residual distribution parameters is done via arithmetic and logical operations on scalar. This cannot be as simple for systems, vecause dividing vectors has no meaning.

The method that is followed has been introduced in [95]. The idea is as follows: given an element K, we first consider an average state $\overline{\mathbf{w}}$. The choice of thsi average states does not seem to be essential, and we take the arithmetic mean. From this, one can evaluate the Jacobians of the flux at this state, say $A(\overline{bbu})$. The next step is to choose a direction $\mathbf{d}$. Again the choice does not seem to be essential and for fluid dynamic problems, we consider the normalised velocity except when the velocity vanishes. In that case we take an arbitrary direction. Once this is done, we compute the eigenvectors $\{r_j\}_{j=1,\cdots,m}$ of the matrix $A(\overline{\mathbf{w}}) \cdot \mathbf{d}$. Any vector X can be decomposed on this basis as:

$$X = \sum_{i=1}^{m} \ell_j(X) r_j.$$

The eigenvectors r_j are often called the right eigenvectors, while the linear forms ℓ_j are often called the left eigenvectors of $A(\overline{\mathbf{w}}) \cdot \mathbf{d}$.

Starting from the LLF residual, $\{\Phi_j\}_{j=1,K}$ where K is the number of degrees of freedom in K. For any eigenvector r_i, we consider the quantities

$$\{\ell_i(\Phi_j)\}_{j=1,\ldots,K}$$

that clearly satisfy:

$$\sum_{j=1}^{N} \ell_i(\Phi_j) = \ell(\Phi).$$

Because of this we interpret these quantities as residual and we can apply the technique of section 3.1.7 to evaluate, for any $j = 1, \dots K$,

$$(\ell_i(\Phi_j))^\star = \beta_j^i \ell(\Phi)$$

where, for example β_j^i is evaluated via PSI "limiter" (62). Once this is done, we define

$$\Phi_j^\star = \sum_{i=1}^{m} (\ell_i(\Phi_j))^\star,$$

this satisfies the accuracy requirements. If needed (and this generally the case), one can add a least square filtering term,

$$\Phi_j^{\star\star} = \Phi_j^\star + \theta(\mathbf{w}_h)|K| \sum_{x_{quad}} \omega_{quad} (A(\overline{\mathbf{w}}) \cdot \nabla \varphi_i(x_{quad})) \, \mathcal{T}_K \, (A(\overline{\mathbf{w}}) \cdot \nabla \mathbf{w}_h(x_{quad})$$

where

$$\mathcal{T}_K^{-1} = \sum_{j=1}^{N} \left| A(\mathbf{w}) \cdot \nabla \varphi_i(x_{quad}) \right|.$$

In [151] that the matrix $\sum_{j=1}^{N} \left| A(\overline{\mathbf{w}}) \cdot \nabla \varphi_i(x_{quad}) \right|$ is always invertible except when the velocity defined by $\overline{\mathbf{w}}$ is zero. However, it that case, the matrices

$$A(\overline{\mathbf{w}} \cdot \nabla \varphi_i(x_{quad}) \, \mathcal{T}_K$$

can always be defined, see [151] for details.

Applications. These results are taken from [92]. The meshes use triangles only unless specified.

In our first example, the domain is a square $\Omega = [0, 1]^2$. The boundary conditions are :

- If $y > 0.5$ and $x = 0$, the Mach number is set to $M_\infty = 4$, the density is $\rho_\infty = 0.5$ and the velocity is $(u_\infty = M_\infty c_\infty, 0)$ with $c_\infty = \sqrt{\gamma p_\infty / \rho_\infty}$.
- If $y \leq 0.5$ and $x = 0$, the Mach number is set to 2.4, the velocity is $(u_\infty, 0)$ and the density set to 1
- The other boundary are assumed to be supersonic.

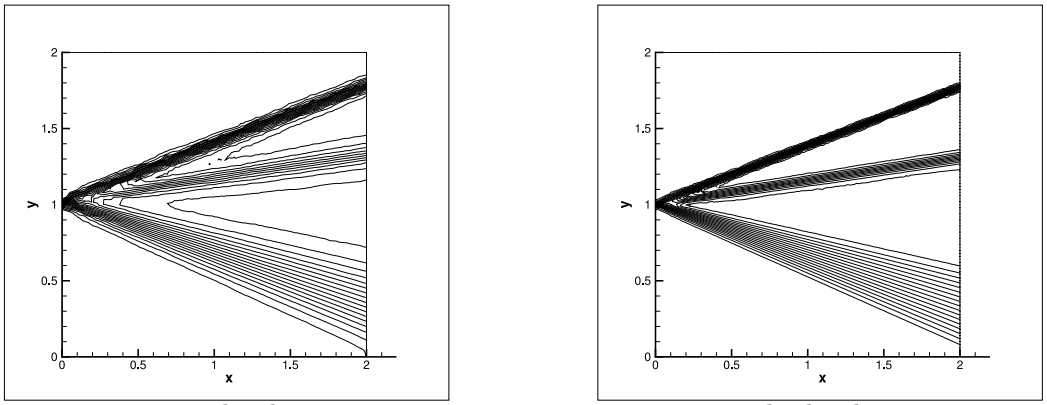
In such a configuration, the flow is steady and supersonic. We have a shock wave on the bottom, followed by a slip line and then a fan, see figure 5. Since the flow is supersonic, the $x-$ coordinate plays the role of time : if one makes a cross–section $x = \text{const}$, we have a self-similar solution of the same type as what one gets for a one dimensional shock tube. It is clear that there is no oscillation at all on the density. The same conclusion holds for the other variables (not displayed).

The next example is the classical flow at $M_\infty = 0.35$ over a sphere. In that case, the flow is symmetric with respect to the x–axis of the domain, but also with respect to the y axis. We have run this case with a second order scheme, a third order scheme, and again the second order scheme on the mesh that has the same degrees of freedom as those of the $\mathbb{P}^2$ scheme. In other words, we subdivide each triangle into 4 smaller triangles which vertices are those of the large triangle and the mid–edges points. The initial mesh has 2719 nodes, 5308 elements and 100 nodes on cylinder. It is displayed on figure 6.

We see on Figure 7 which displays the pressure coefficient isolines the improvement of the solution quality when the scheme is upgraded from second order to third order. More important, the same Figure indicates clearly that the second order scheme on the refined mesh gives less accurate results than the third order one. Note that we have the same degrees of freedom in both cases.

We have re-run this test case on an hybrid mesh using the second order and the third order schemes. In both cases, the same degrees of freedom are used (i.e. we use the DOFs of the sub-triangulation for the second order scheme). The results are shown on figure 8. The mesh use 81 points on the sphere. We get the same conclusions as before.

Our next examples is a flow over a NACA012 airfoil. It is transonic, and has the following conditions at infinity: $M = 0.8$, angle of attack of 1.25° The mesh has 10959 points and 21591. This corresponds to 43509 degrees of freedom.



second order third order

Figure 5: Jet problem : isolines of the density, second and third order LLxFf scheme. All the degrees of freedom are plotted. and the same isolines are plotted

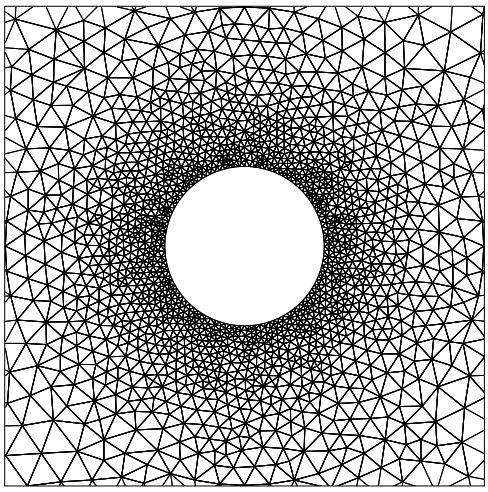
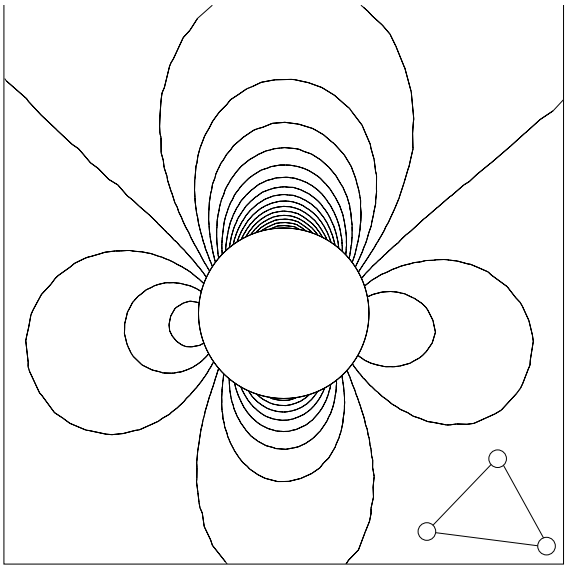
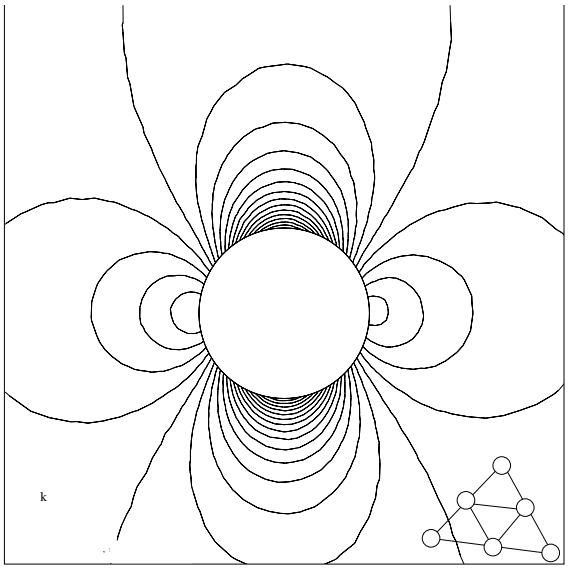


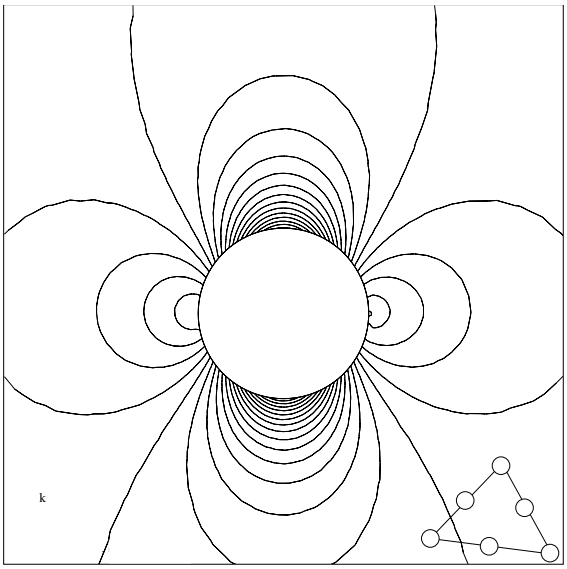
Figure 6: Subsonic sphere problem : Zoom of the mesh for the sphere problem. The mesh has no symetry.



Second order

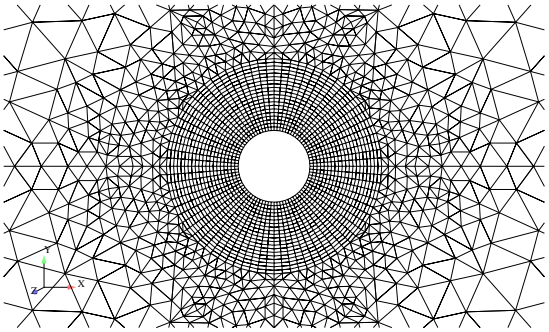


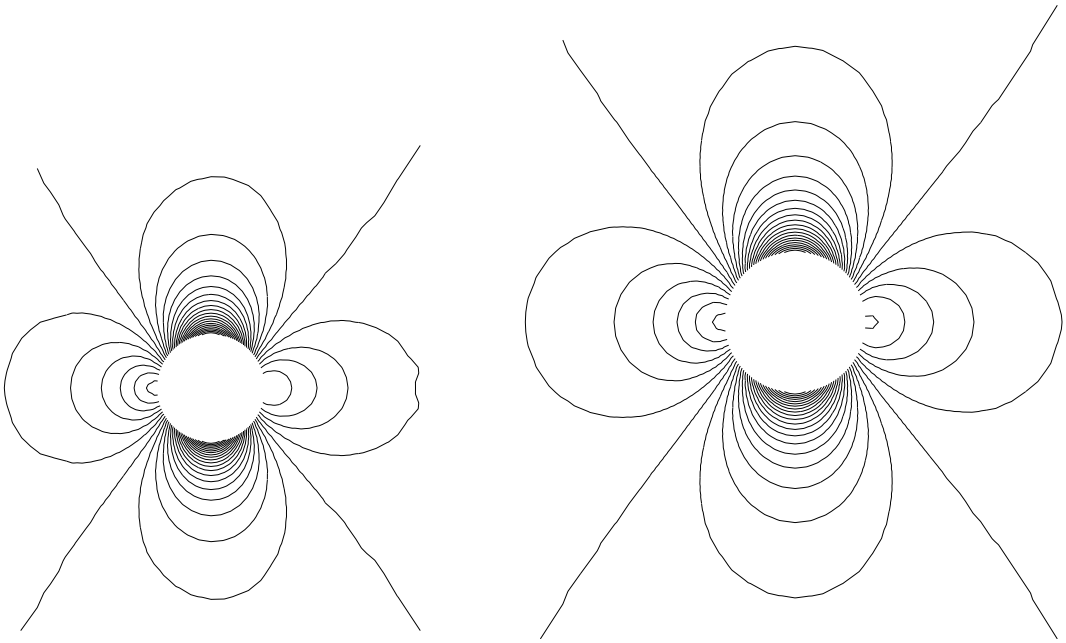
Second order using the P^2 dofs



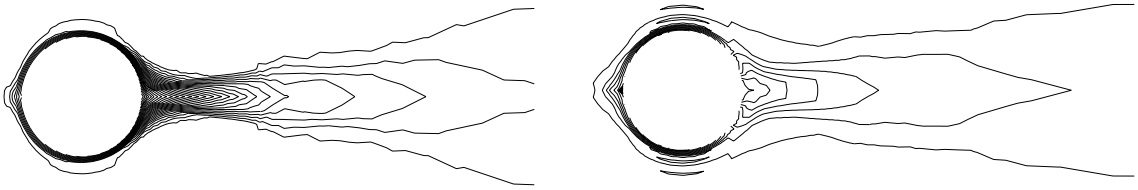
third order scheme

Figure 7: Subsonic sphere problem : Isolines of the pressure coefficient. We have the same isolines on each figure.





Pressure coefficient



second order third order

Entropy variation

Figure 8: Subsonic sphere problem, hybrid mesh : Pressure coefficient and entropy variation on an hybrid mesh, $M_\infty = 0.35$.

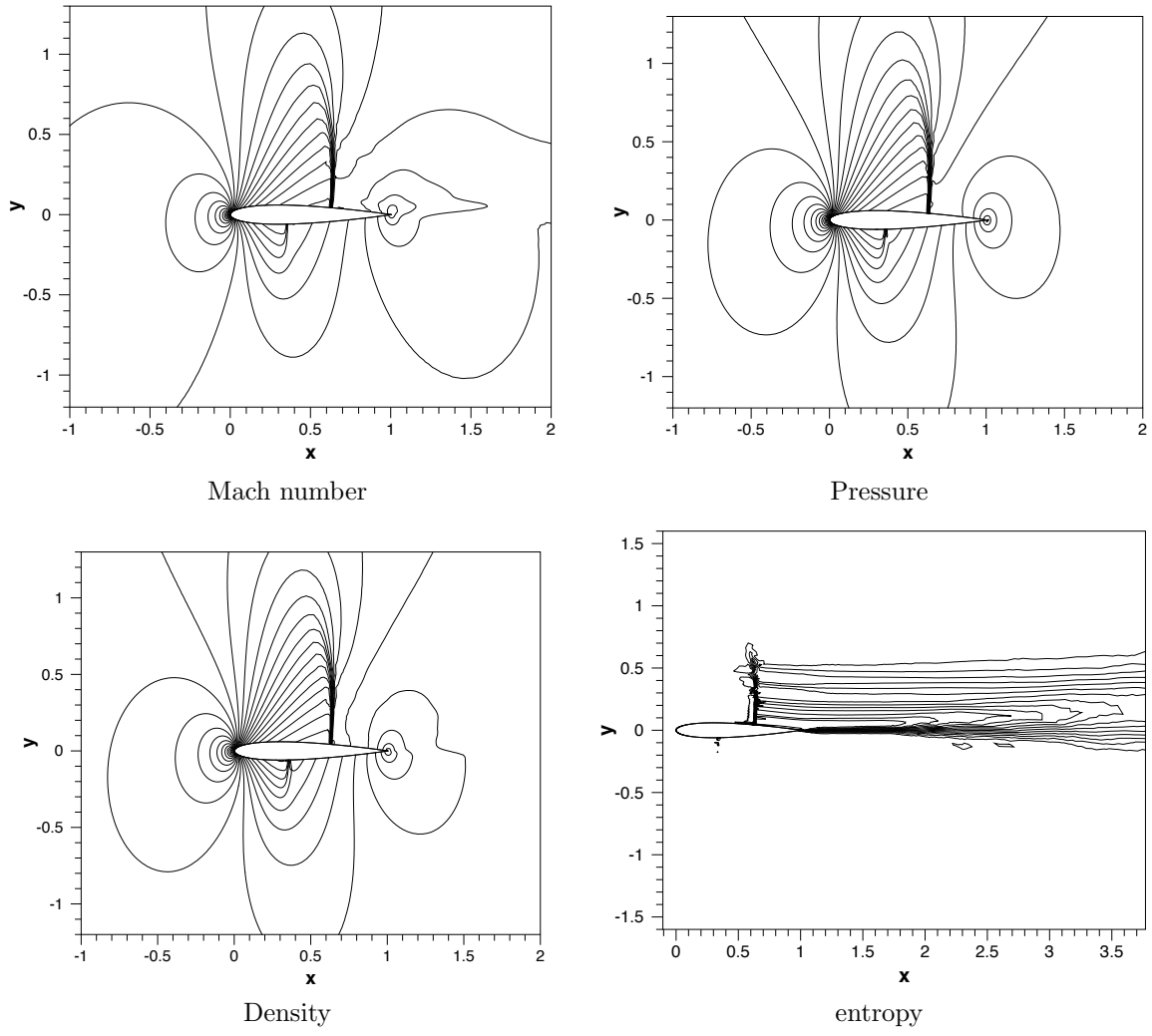


Figure 9: Transonic NACA012 problem. Isolines of the Mach number, pressure, density and entropy for the NACA012 case.

On figure 9, we have displayed the Mach number, the pressure coefficients en relative entropy deviation for the third order version of the scheme.The solutions are fine. Note however a non physical overshoot in the entropy across the upper shock.

We have run many other tests as the following (results not shown). If we compare the second order solution run with a mesh constructed from the mesh we have used where the element is sub-triangulated so that we have the same number of degrees of freedom, we can see an excellent agreement between the solutions with a main difference however. In both cases, the shock with is one element, but one element for the third order solution is roughly twice as large as an element for the second order one. Hence, the shock look more diffused in the third order case. However, the entropy levels are much lower, as we have already seen in the two sphere subsonic case.

Another case is the Ringleb flow. It has been devised by F. Ringleb [152] in 1940, see [153] for a derivation of more general solutions. This is an isentropic, irrotational two dimensional flow. It is defined from the streamline function (θ is the velocity angle with respect to a given direction and v is the norm of the velocity) $\psi = \frac{\sin\theta}{v}$. From this, it is possible to get the explicit form of the streamlines

$$x = \frac{1}{2}\frac{1}{\rho}\left(\frac{1}{v^2} - \frac{2}{k^2}\right) + \frac{J}{2}$$
$$y = \pm\frac{1}{k\rho v}\sqrt{1 - \left(\frac{q}{k}\right)^2}$$

with

$$k = \frac{1}{\phi} \text{ a constant on any stream line, } J = \frac{1}{c} + \frac{1}{3c^2} + \frac{1}{5c^2} - \frac{1}{2}\log\left(\frac{1+c}{1-c}\right)$$
$$c = \sqrt{1 - \frac{\gamma-1}{2}q^2}, \rho = c^{2/(\gamma-1)}$$

The pressure is determined by the equal entropy assumption. We see that the isotach lines are the circles

$$\left(x - \frac{J}{2}\right)^2 + y^2 = \frac{1}{4\rho^2 q^4}$$

From this it is possible to determine the exact solution: given a point (x, y), we determine the speed of sound c such that (x, y) belongs to the circle of center $(J(c)/2, 0)$ and radius $1/2(\rho q^2)$. Once this is done, we can get all the other values.

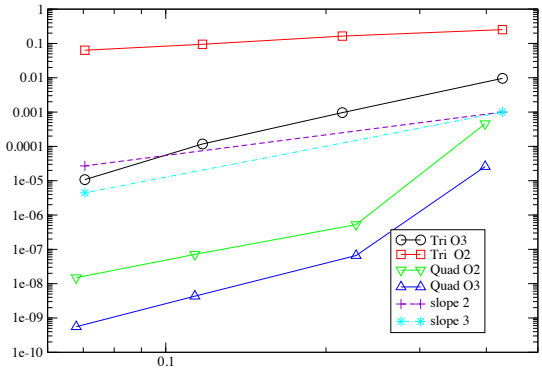


Figure 10: Ringleb flow problem. L^2 error on the density for the Ringleb flow. Tri stands for triangle, Quad for quadrangle. O2 stands for second order, O3 for third order.

We have run this case in the (symmetric) domain defined by

- the circle $q = 0.3$ on the top and the bottom,
- the extreme stream lines $k = 0.4$ and $k = 0.8$.

The simulation has been conducted with two series of meshes. The first one is made of quads cut into two triangles, always in the same direction. The mesh is then made symetric. In the second one, we only consider the quads. In both cases, we have $2 \times P$ points on the streamlines $k = 0.3$ and 0.8 and P points on the circles $q = 0.3$. Here we have taken $P = 15$, 30, 60 and 100. The error (in the L^2 norm for the density are shown on figure 10. We see a slope of -3 for the third order scheme and -1.5 for the second order scheme. We also note that though the formal accuracy in both case is as expected, the effective accuracy on the quad meshes is much superior to what is obtained for triangle meshes.

We have run the same scheme on a scramjet–like configuration using an hybrid mesh as shown on Figure 11. This example has already been run in [95]. The inflow mach number is set to 3.5. The

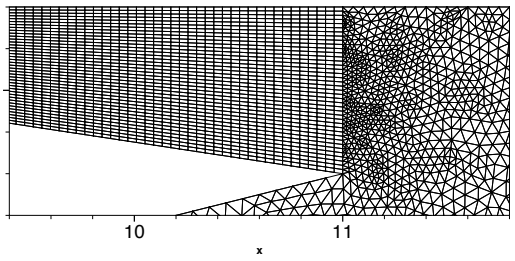
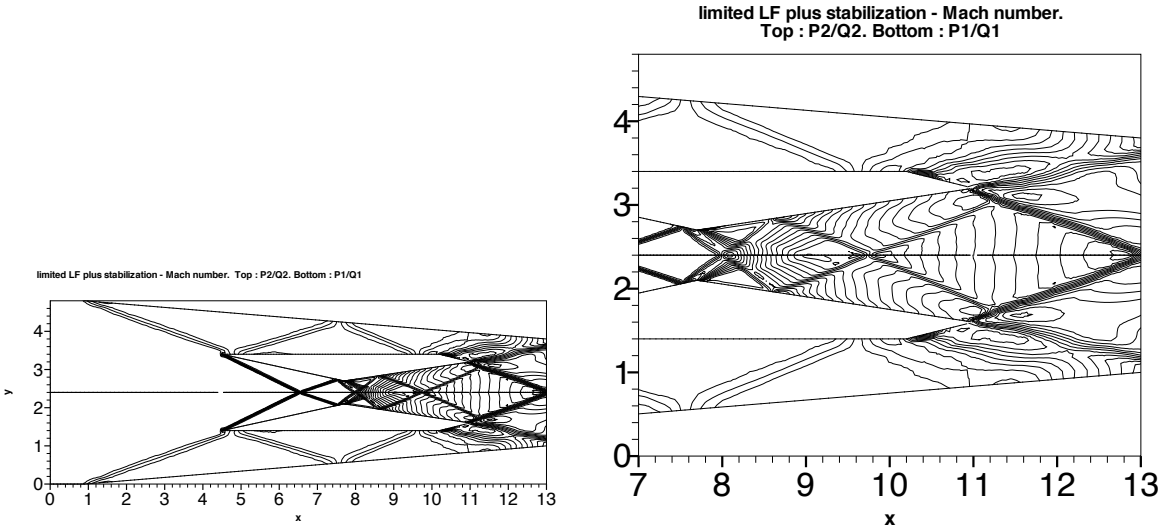


Figure 11: Zoom of the mesh for the scramjet problem.

geometry is such that many waves coexist and interact in very complex flow patterns. This situation is particularly clear on the upper part of the internal body where shocks, fans and their reflection due to wall interact. Again, in both cases, the same number of degrees of freedom have been used. Once again, the scheme has been run starting from a uniform flow configuration. Figure 12 shows the Mach number isolines. As expected, there is no real difference between the solutions since the flow is basically made



zoom

Figure 12: Scramjet problem. Mach number distribution. Top : the third order solution, bottom the second order solution. The same isolines are plotted.

of shock, fans, slip lines and constant states : this is not an accuracy case, but a case that shows that, despite the flow complexity, the third order scheme is robust.

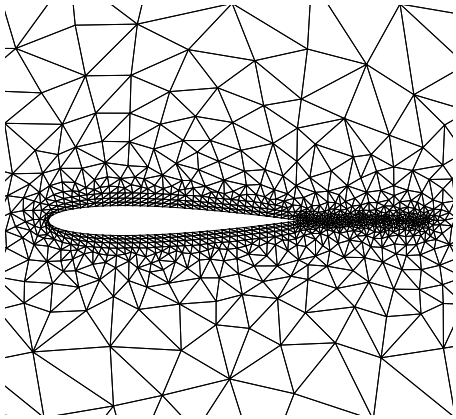


Figure 13: An example of computational grid used for the NACA-0012 test case.

However, one can see a small difference between the solutions : the slip line created by the interaction of two shocks after the blade is a little bit more twisted for the third order scheme than the second order one. We also see that the resolution of the discontinuities is in both case approximately one cell width.

4.2.2 Navier Stokes equations

We rereport here results taken from [149]. The scheme and problems have already been discussed. For more details, the reader may also consult [67]. The filtering term has to be more elaborate in order to take into account the viscous terms.

The first example is the classical test case consisting of a subsonic viscous flow over a NACA-0012 airfoil at zero angle of attack. The free stream mach number is 0.5 and Reynolds number is 5 000. This is a widely used test case for two dimensional laminar flows; a distinctive feature of this test case is a steady separation bubble near the trailing edge of the airfoil. An example of computational grid is displayed in figure 13. The grid extends about 50 chords away from the airfoil. The airfoil boundary is considered adiabatic, no-slip and is represented by piece-wise quadratic elements, the far-field boundary condition is applied on the outer boundary of the domain, see [112] for a precise description of the boundary conditions approximation, as well as details on the steady state solver. The steady state is considered to be reached when the L^2 norm of the density residual is drop by ten orders of magnitudes respect to the initial value.

In figure 14 are depicted the solutions computed with the linear scheme and the SPR-ZZ gradient reconstruction, for $\mathbb{P}^1$ and $\mathbb{P}^2$ elements. The solution with the $\mathbf{\Phi}^1$ elements has been computed on a grid obtained from that with $\mathbb{P}^2$ elements (4 216 elements) and splitting each $\mathbb{P}^2$ triangle with four $\mathbb{P}^1$ triangles, in such a way the number of DOFs for the second and third order simulation is exactly the same. Note, in figure 14, that although there is not much difference in the Mach number contours between the second and the third order simulations, the streamlines near the trailing edge are very different, and only the third order scheme is able to reproduce the symmetric recirculation bubble. For the same simulations, in figure 15 and 16 are reported the pressure and skin friction coefficients profiles, respectively. Note the more regularity of the solution of the third order simulation respect to the second order one, for the same number of DOFs.

The second example is a steady, laminar flow at high angle of attack, around a delta wing with sharp edges. As the flow passes the leading edge, it rolls up and creates a big vortex structure which is convected far behind the wing, at the same time, near the leading edge a smaller secondary vortex appears. A free stream Mach number $\mathrm{M} = 0.5$ is considered, the Reynolds number, based on the root chord of the wing is $\mathrm{Re} = 4\,000$, the angle of attack is $\alpha = 12.5°$.

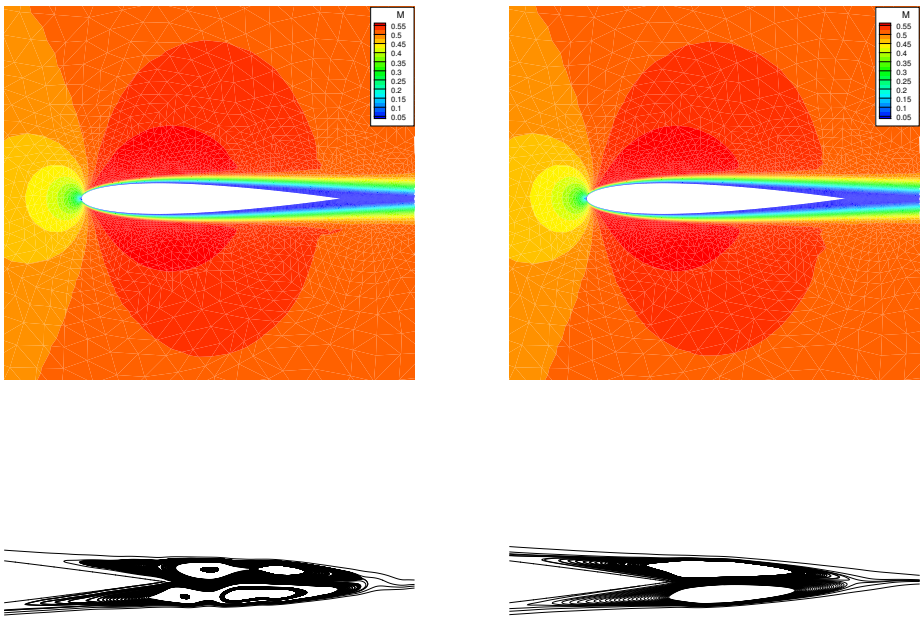


Figure 14: Mach Number contours (top) and streamlines near the trailing edge (bottom) for the second (left) and third (right) order linear scheme.

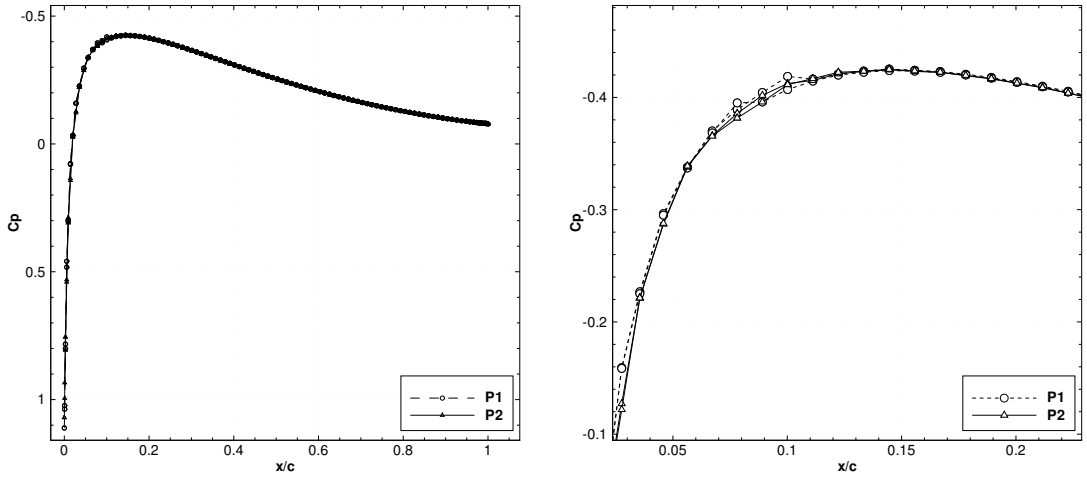


Figure 15: Pressure coefficient along the whole NACA-0012 airfoil for the second and third order simulations, with the same number of DOFs.

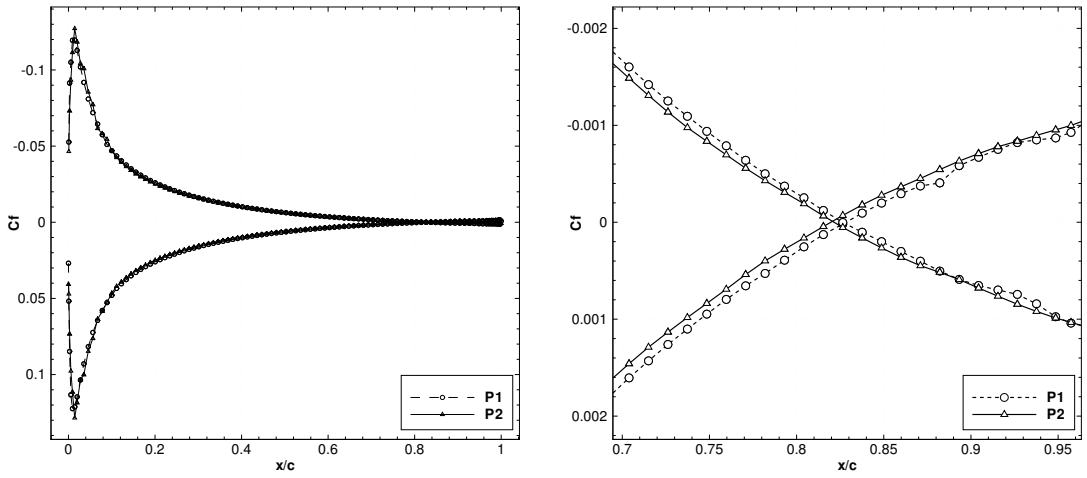


Figure 16: Skin friction coefficient along the whole NACA-0012 airfoil for the second and third order simulations, with the same number of DOFs.

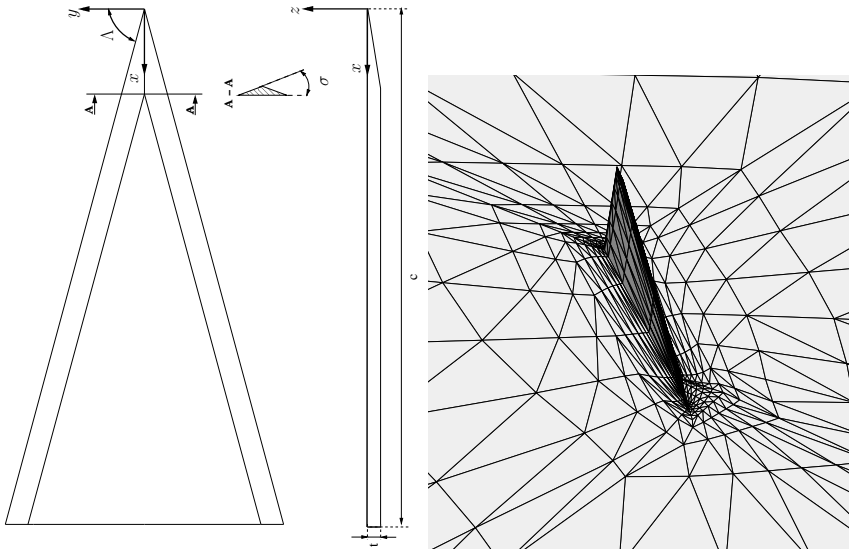


Figure 17: Left: Bottom and side views of the model of the delta wing: $\Lambda = 75°$, $\sigma = 60°$ and $t/c = 0.024$. Right: a coarse mesh of tetrahedra use for the simulations.

The geometry of the delta wing is depicted in figure 17, together with an example of a coarse grid used for the simulations. The grid consists of tetrahedra; finer levels of grids are obtained by uniformly splitting each tetrahedron of the coarser level with eight tetrahedra. Note the presence of very stretched elements on the wing. The wing surface is treated as no-slip adiabatic wall, the vertical plane intersecting the root of wing is treated as a symmetry plane, while far field boundary conditions are applied on the outer boundary of the domain.

The solution is initialized with an uniform flow, the lower order solution is used as initial solution for the third order computation. For this test case the linear scheme is used with the SPR-ZZ gradient reconstruction method; in figure 18 are reported the streamlines and Mach number contours, at different stations, of the third order solution on the finest grid.

In figure 19 are reported the drag and lift coefficients computed with linear and quadratic elements, on three uniformly refined grids. For comparison, are reported also the reference values computed in [154] by extrapolating the results obtained with a higher order DG method. Observing the convergence of the drag coefficient in term of DOFs, it can be noted that there is no significant gain in using a higher order approximation, with respect to the second order. This behavior can be caused by the singularity at the leading edge of the wing, which might mask the benefits of a higher order approximation with an uniform mesh refinement. Regarding the convergence of the lift coefficient, it could be observed a clear benefit of using a higher order approximation, because the big vortex structure over the wing is better captured with higher order elements.

As last test example, the interaction of an oblique shock wave with a laminar boundary layer is considered. The aim of this test is to show the non-oscillatory properties of the non-linear scheme in presence of discontinuities of the solution and at the same time, the capability to maintain the accuracy required for the discretization of the boundary layer.

The test consists in a laminar boundary layer developing over a flat plate and an incident shock impinging the boundary layer. Since the flow is supersonic, a shock appears at the leading edge of the

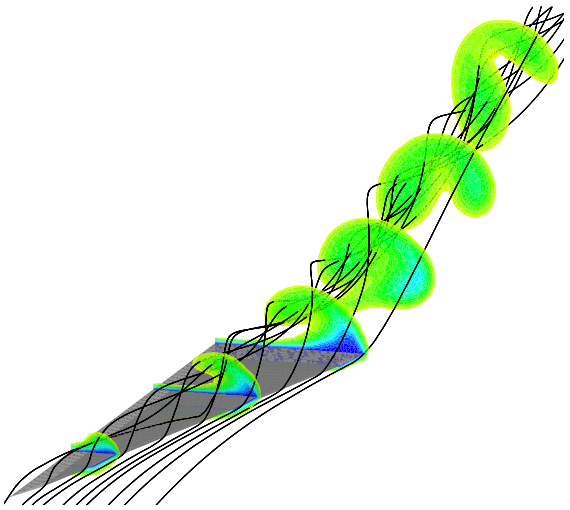


Figure 18: Streamlines and slices of Mach number contours along and behind the delta wing, for a third order simulation on a fine grid.

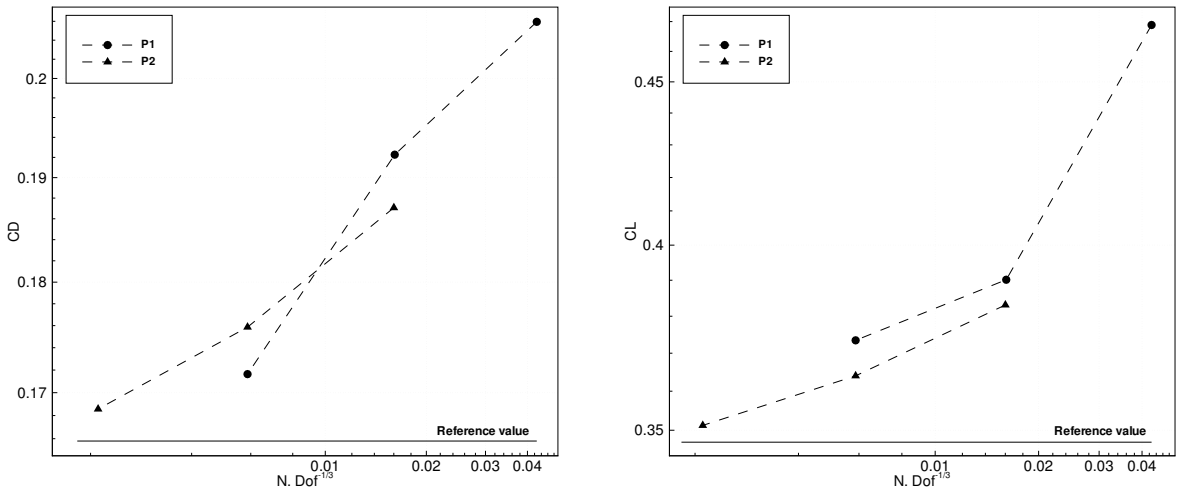


Figure 19: Drag (left) and lift (right) coefficients as function of DOFs for the delta wing simulation, with linear and quadratic elements.

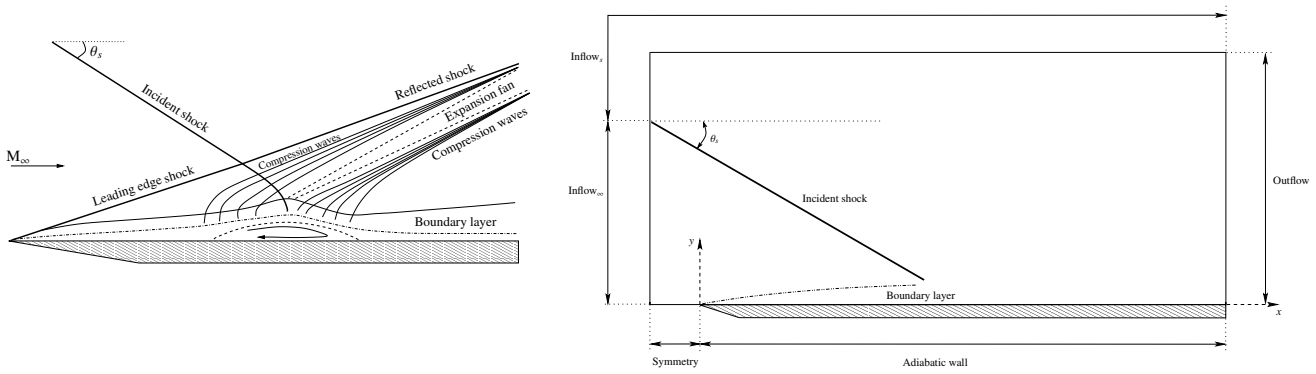


Figure 20: Schematic representation of the waves pattern (left) and computational domain with boundary conditions (right) for the shock-wave/boundary layer interaction problem

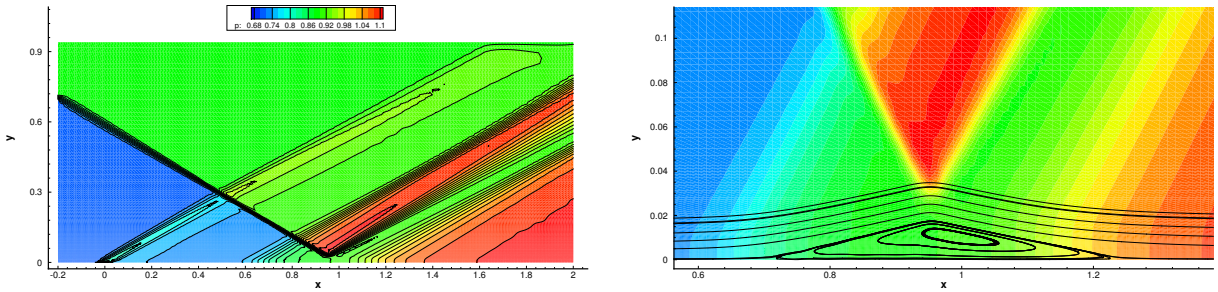
flat plate, that interacts with the oblique shock. Furthermore, at the impinging point, the incident shock produces a separation of the boundary layer, the shock is then reflected and an expansion fan appears, turning the flow toward the wall and causing a reattachment of the boundary layer, as it is depicted in figure 20.

In the numerical simulations, the oblique shock is generated by imposing the incoming supersonic flow state on the lower part of left boundary, while another supersonic state is imposed on the upper part of the left boundary and on the top boundary; this state is computed using the relations of the oblique shocks, such that the incident shock has a certain angle of incidence θ_s. The height of the computational domain is 0.94, while the range of the computational domain in the x direction is $[-0.2, 2]$, the flat plate has length $L = 2$ with the leading edge of flat plate at $x = 0$. Along the plate, the no-slip adiabatic wall boundary condition is applied, while the symmetry boundary condition is applied on the remaining part of the bottom boundary. On the right boundary, the outflow boundary condition is applied, see figure 20. The inflow states are chosen such that the free-stream Mach number is $M = 2.15$ and the angle of the incident shock is $\theta_s = 30.8^\circ$, in this case the impingement point would be at center of the plate for an inviscid fluid. The Reynolds number based on the free-stream values and the distance between the plate leading edge and the inviscid shock impingement point is 1×10^5.

The non-linear scheme with the SPR-ZZ gradient recovery strategy is used to perform the numerical simulations at second and third order of accuracy. The computational domain is generated from the triangulation of a 90×85 structured grid; the first number refers to the number of elements on the horizontal boundaries, with 80 elements along the plates, the second number refers to the number of elements on the vertical boundaries. The element distribution is uniform on the x direction, while along the y direction a non-uniform distribution of the elements is used, with a mesh spacing $\Delta y = 0.5 \times 10^{-3}$ near the bottom boundary. For comparison, a second order simulation is also performed on a finer grid with the same number of DOFs of the third order simulation on the coarse grid. The simulation is initialized with an uniform solution, and the second order solution is used as initial solution for the third order approximation. Except the case of the second order simulation on the coarse grid, for which the initial residual is reduces by ten orders of magnitude, the residual for the third order simulation and the second order one on the finer grid could not be reduced by more than eight orders of magnitude.

In figure 21-(a) are shown the contours of the pressure for the third order simulation; all the features of this problem are well represented. In figure 21-(b) is reported a zoom of the solution where the incident shock impinges the boundary layer. Two features are evident: the reflection of the incident shock and the recirculation bubble as a consequence of the separation and subsequent reattachment of the boundary layer produced by the incident shock and the expansion fan, respectively.

The profiles of density, pressure and Mach number along the lines at $y = 0.29$ and $y = 0.15$ are reported in figure 22. Note that the third order scheme gives a very sharp and monotone representation of the discontinuities and also smooth portions of the solution are better represented compared to the second order solution. It is important to remember that smooth and discontinuous solutions are treated within



(a) (b)

Figure 21: Left: contours of the pressure obtained with the third order scheme for the shock/boundary layer interaction. Right: zoom of the solution near the impinging point of the shock with the boundary layer, streamlines are also reported to show the separation bubble.

the same non-linear scheme without any special treatment or tuning parameter. For a fair comparison, it is also reported the solution obtained with the second order scheme on a finer mesh; it is worth noticing that, although a mesh refinement produces an improvement of the numerical solution, the level of accuracy obtained with the second order scheme is still lower than that obtained with the third order scheme, for the same number of DOFs.

Finally, in figure 23 are reported the values of the pressure and of the friction coefficient along the plate. The oscillations near the point $x = 0$ are due to the singularity of the solution at the leading edge of the flat plate, but they are limited only in small region around the leading edge. The third order scheme seems less sensitive to this singularity compared to the second order simulations. The separation bubble can easily detected by the negative values of friction coefficients, note also the pressure plateau in the detached zone.

4.3 Free surface flows

The framework presented in this chapter has proved quite interesting to construct discrete approximations of systems of PDEs modelling free surface flows, namely the shallow water equations and dispersive enhancements (Boussinesq and/or Green-Naghdi equations). Early work on steady hydrostatic flows had been done in the PhD of M. Hubbard (see e.g. [155, 156], and also the paper by Brufau and Garcia-Navarro [157]). More recent work, combining high order of accuracy in space and time, the preservation of moving steady states, robust handing of dry areas, and dispersive extensions is discussed in [65, 96, 101, 136, 158, 99, 159].

4.3.1 Inundation of a complex three-dimensional beach

The first example we consider involves the solution of the shallow water equations, and is taken from [99]. In this paper, the nonlinear stabilized Lax-Friedrich's method has been combined with the fully explicit time marching strategy discussed in section §3.2, and modified to allow a (provable) preservation of the non-negativity of the water depth. To illustrate the capabilities of the method obtained we consider a standard benchmark in the oceanography community involving the Tsunami runup onto a complex three-dimensional beach. The so-called Monai valley benchmark aims a simulating a scaled down laboratory experiment reproducing the impact of the tsunami wave that hit the Okushiri island in Japan in 1993. The bathymetry, and inlet data are available on the web page of the third international workshop on long wave runup models [160] (see also [161, 162]), with the data relative to the time series of the wavet level in three gauges close to the shore. The shape of the bathymetry and of the inlet wave, as well as the position of the three wave gauges are shown in the left, middle and right pictures in figure 24. In the observations [160, 161, 162] the highest runup is of 32 $[m]$, and it occurs in the region of the Monay valley where the bathymetry is steepest. For clarity, this region is encircled in the results presented in the following.

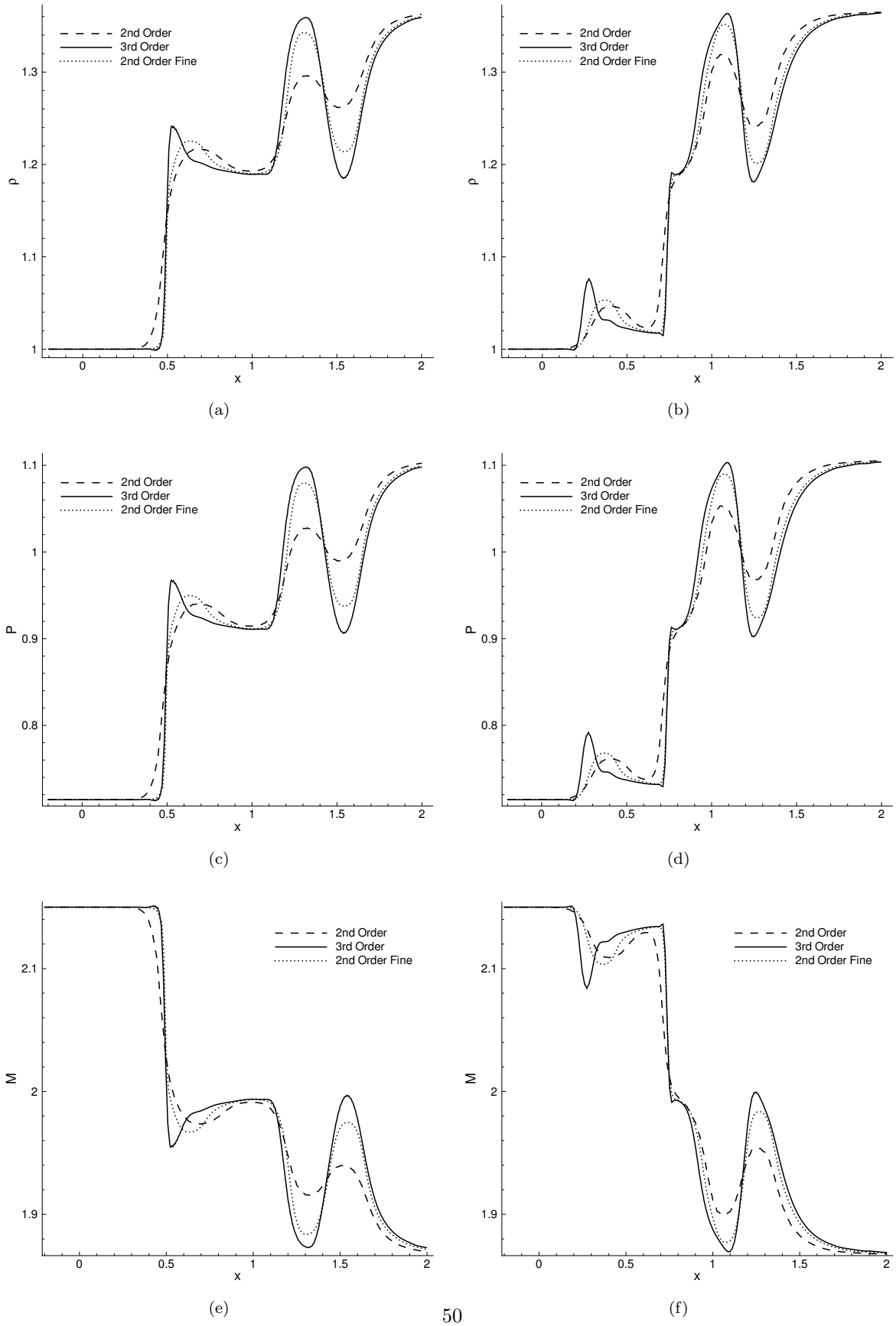
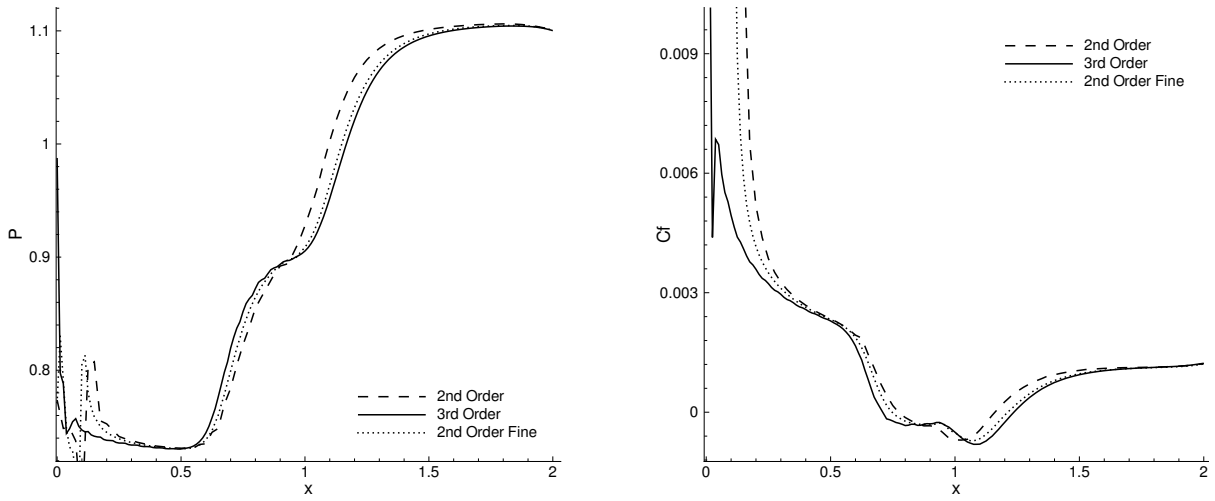


Figure 22: Density, pressure and Mach number profiles along the line $y = 0.29$ (a, c, e) and the line $y = 0.15$ (b, d, f) for the shock/boundary layer interaction problem.



(a) (b)

Figure 23: Pressure (a) and skin friction (b) profiles along the flat plate for the shock/boundary layer interaction problem.

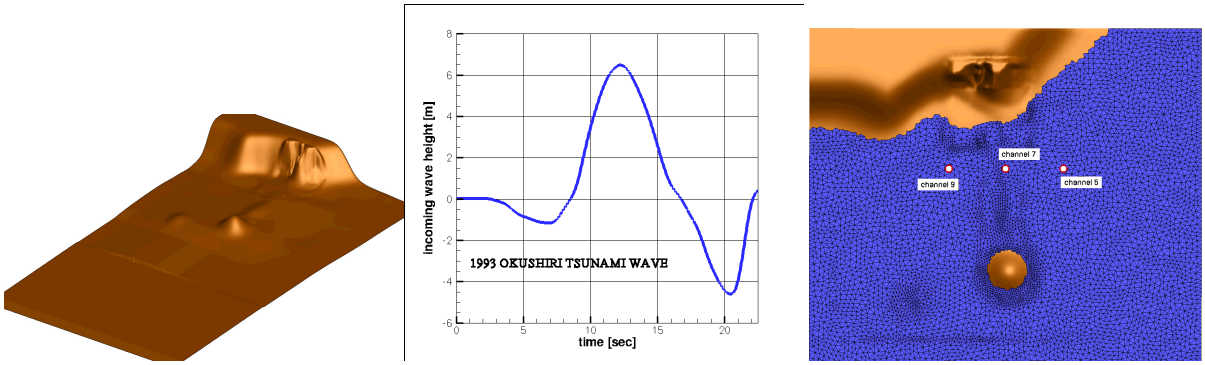


Figure 24: Monai valley benchmark. Left: bathymetry. Middle: incoming wave. Right: position of the wave gauges.

The results obtained are summarized on figures 25 and 26. The top row in the first figure shows the initial withdrawing of the water followed by the arrival of the main wave The bottom row shows how, after hitting the beach, the wave reflects, and a large wave travels toward the right to hit the steepest slopes in the region of the Monai village. As already said, the highest runup observed is about $32\,[m]$ and it has been observed in the region of the Monai valley, highlighted by a yellow circle in the figure. This is well reproduced by the simulations.

Lastly we report on figure 26 the time history of the water level in gauge 5, comparing simulated and measured values, and the runup plot, showing clearly that the deepest inundation point is the region of the Monai village.

4.3.2 Approximation of moving steady states

The super consistency property discussed in section §3.1.8 also has applications in shallow water flows. In this case, the state vector $\mathbf{w}$ is defined by the quantities H, the water depth, and $\vec{q}$, the volume flux $\vec{q} = H\vec{u}$, with $\vec{u}$ the depth-averaged flow velocity. A known steady state involving moving water,

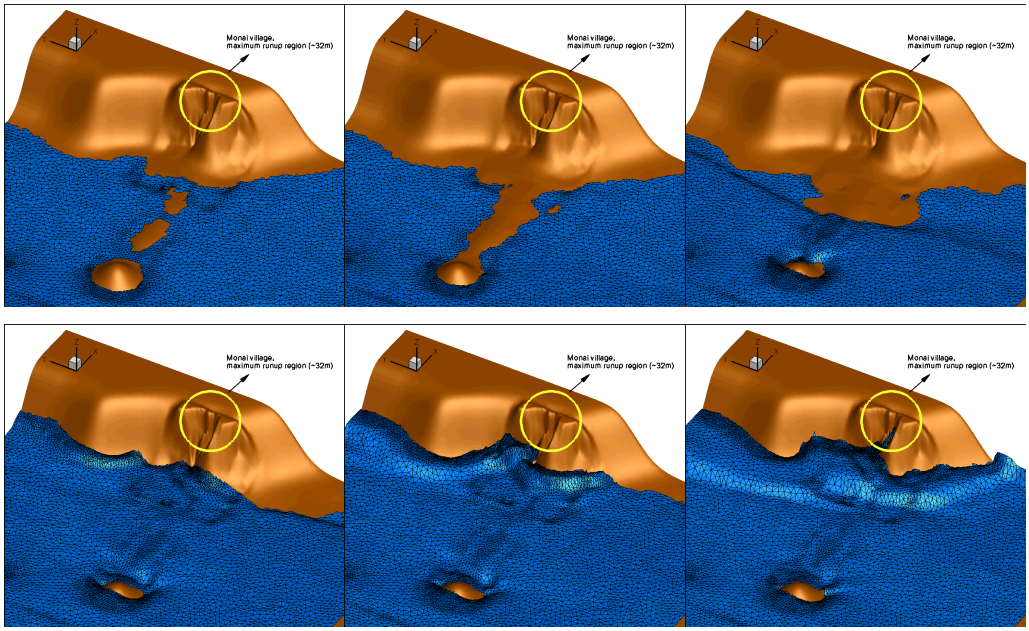


Figure 25: Monai valley benchmark. 3D view of the inundation process.

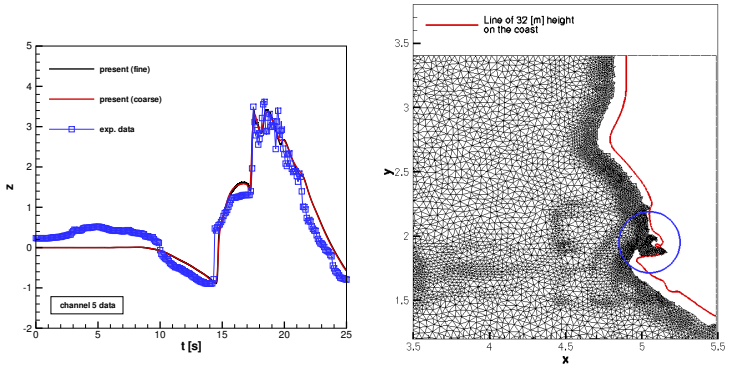


Figure 26: Monai valley benchmark. Time series in gauge 5 (left), and runup plot (right).

is the pseudo-one dimensional flow characterized by $\vec{q} = \vec{q}_0 = c^t$, and $\mathcal{E} = \mathcal{E}_0$, with $\mathcal{E}$ the total energy $g(H+b) + \vec{u}\cdot\vec{u}/2$, and $b = b(x,y)$ the bathymetry. This solution allows to check numerically proposition 3.9. To do this, we consider the tests discussed in [101, 99].

The first, involves a small perturbation of the initial steady state over a bathymetry with C^1 regularity obtained as a series of ribs defined by truncated $\sin^2$ functions. The evolution of the perturbation on an irregular triangulation is studied. The typical result is shown on 27 showing a 3D view of the free surface level. The left picture is obtained with a standard scheme based on a $\mathbb{P}^1$ approximation of the state vector $\mathbf{w}$ and of the bathymetry. The right result is instead obtained with the scheme based on a direct approximation of the total energy $\mathcal{E}$ and of the flux $\vec{q}$, and with a higher order approximation and quadrature of the bathymetric gradient. The improvement is quite remarkable.

The second test consists in verifying the property of proposition 3.9 by computing, on irregular triangulations, the solution error at a finite time when starting from the exact nodal steady state. This is done with bathymetries of increasing smoothness, and with volume and edge quadrature strategies of increasing accuracy. The results are sumamrized on figure 28 in which figures (a), (b), (c), and (d) show

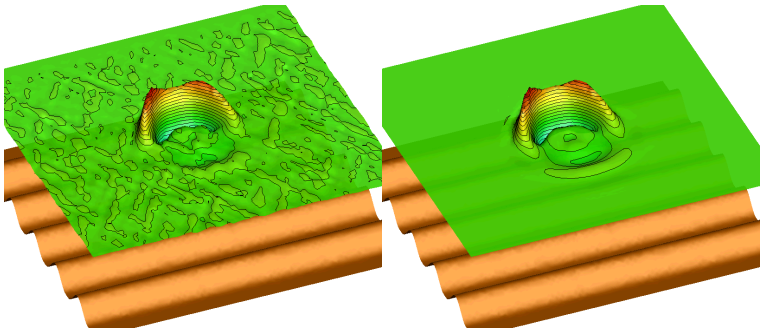


Figure 27: Moving steady states: evolution of a small perturbation in a homo-energetic steady state. 3D plot of the free surface. Left: approximation in physical variables. Right: approximation in steady invariants.

the grid convergence obtained on bathymetries with different regularity when using quadrature strategies with errors of orders h^2, h^4, h^6, and h^8 respectively. The last column show the error convergence on a fixed mesh when increasing the accuracy of the quadrature. In articular, picture (e) is obtained on the coarsest mesh used in the convergence study, while picture (f) on the finest. The underlying approximation is P^1. Not only this result confirms the super consistency analysis, but it also shows that for exact quadrature, the residual approach would yield exact preservation of the steady state.

For additional examples involving other steady state solutions, the interested reader is referred to [101, 99, 102].

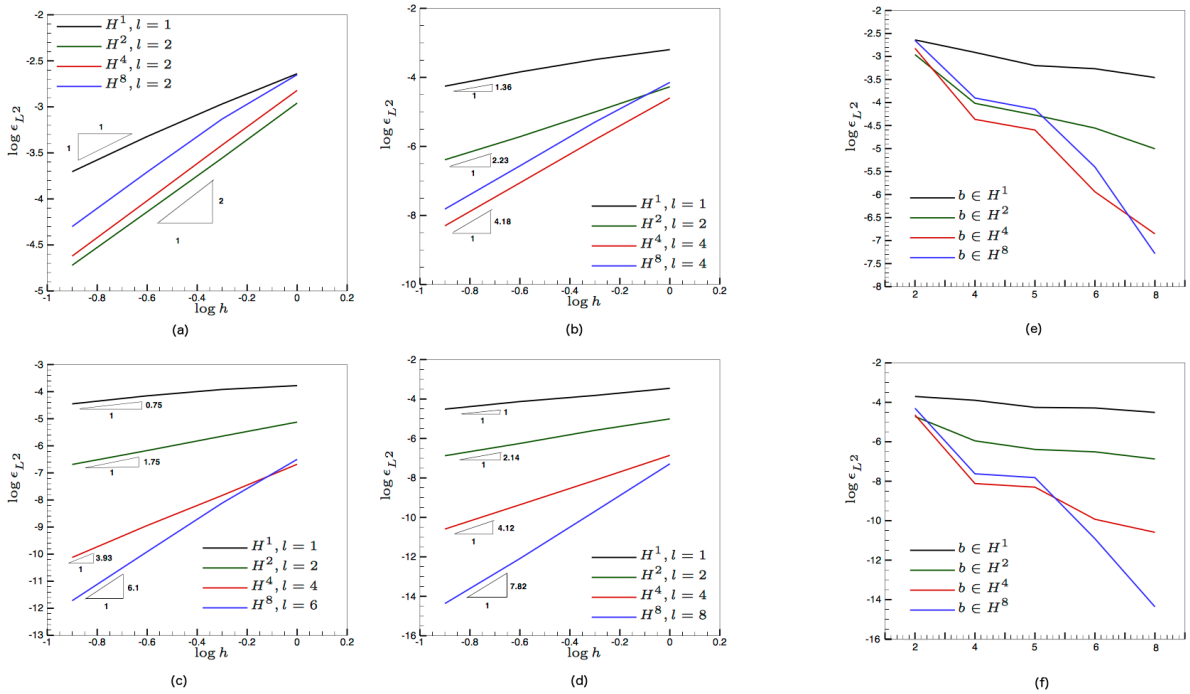


Figure 28: Moving steady states: super consistency of the scheme. Left and middle column: grid convergence for different quadrature strategies. Right column: quadrature convergence on the coarsest (top) and finest (bottom) grid.

4.3.3 Residual based stabilized methods for dispersive waves

Another challenging application in free surface flows in the inclusion of non-hydrostatic effects in depth-averaged models. The interested reader may consult the review papers [163, 164] and the book [165] for an overview of the modelling issues. Concerning numerics, the typical form of a depth averaged Boussinesq-type model is

$$\partial_t K + \nabla \cdot \mathbf{f}(\mathbf{w}) + \mathbf{s}(\mathbf{w}, \mathbf{x}) = 0 \tag{91}$$

where the quantity $K(\mathbf{w})$ is related to the state vector by

$$\mathbf{w} - \mathcal{T}(\mathbf{w}) = K \tag{92}$$

where $\mathcal{T}(\cdot)$ is a non-linear elliptic operator. Here physical dispersion is present in the PDE. The challenge is thus to design a numerical method with low dissipation and very low dispersion errors to allow long time integration of propagating waves, while however guaranteeing a sufficient degree of dissipation to avoid spurious modes. The use of some stability mechanism is also required as the term $\mathcal{T}(\mathbf{w})$ is often neglected locally to recover the hyperbolic shallow water equations, and model breaking regions as moving bores [166, 167, 168, 159]. The requirement is then to have a low dissipation/dispersion method, capable of handling both the parabolic Boussinesq equations, and the hyperbolic shallow water ones, with eventually capabilities for capturing of shocks and dry areas.

This has led to the work presented in [158, 169, 159] which has tried to extend upwind and multidimensional upwind residual based stabilization techniques to these systems. The main idea is to decouple the approximation of the two sub problems above. The elliptic step (92) is solved with a standard C^0 Galerkin method, while an upwind scheme is used in the evolution step (91). The work discussed in the references shows evidence that this approach is a sound one, and that provided that the hyperbolic step is solved with al least third order of accuracy, the elliptic phase can be solved with a second order method without affecting the dispersion accuracy. This generalizes on unstructured grids, and to residual based stabilized method, an idea proposed in the finite difference context by Wei and Kirby in [170]. The schemes obtained, all reduce in 1D to a streamline upwind method stabilizing the Galerkin approximation of the first order PDE (91) with cell integrals depeding on the residual of (91), and on the sign of the shallow water Jacobians. In two space dimensions, both a standard stramline upwind formulation, and a multidimensional upwind variant based on the LDA method (61) have been proposed in [158].

We present here three results. The first is the characterization of the accuracy of the methods obtained. Figure 29 shows results relative to a second order Galerkin approximation of (92), and a third order stramline upwind (SU) or or fourth order Galerkin (cG) approximation of (91). In particular, the left picture provides a numerical convergence study on a propagating solitary wave. Despite the second order treatment of the elliptic term, the overall accuracy measured for a propagating solution is three. More importantly, the middle and right pictures study the dispersion errors of the schemes and compare them to second and fourth order finite differencing. The result shows two important features: both the cG and SU are as good or better than the fourth order finite difference method; for propagating solutions, the upwind SU stabilization actually improves the dispersion properties of the scheme providing lower dispersion errors, especially for shorter waves.

The second result tests the ability of the proposed method to correctly reproduce the energy exchange between different harmonics when monochromatic waves shoal on a two dimensional circular shelf. This is a standard benchmark for multidimensional Boussinesq-type codes (see [171, 172, 173, 174, 175, 176, 177] and references therein). In [178] experiments have been conducted in several configurations involving values of period and amplitude for the incoming monochromatic wave. Here we discuss the results for: case (a) with T=2 s, A = 0.0075 m, h_0/λ = 0.117; case (b) with T=1 s, A = 0.0195 m, h_0/λ = 0.306. The first case has a relatively weak dispersive character, but presents an important energy transfer to higher harmonics. The second case is quite demanding as it involves a higher dispersion degree, outside the validity of the most simple Boussinesq models. Figure 30 summarizes the results obtained solving the enhanced Boussinesq equations of [179] on unstrucutred triangulations. Both a 'classical' streamline upwind stabilization and a multidimensional one based on the LDA distribution (61) have been tested. The pictures clearly show that these multidimensional stabilized methods have a high potential in resolving the energy transfer between harmonics, also in the more demanding cases.

Finally, we show the results obtained on an experiment carried on [180] and involving the refraction and diffraction of monochromatic waves over a complex bathymetry. A sketch of the experiment is

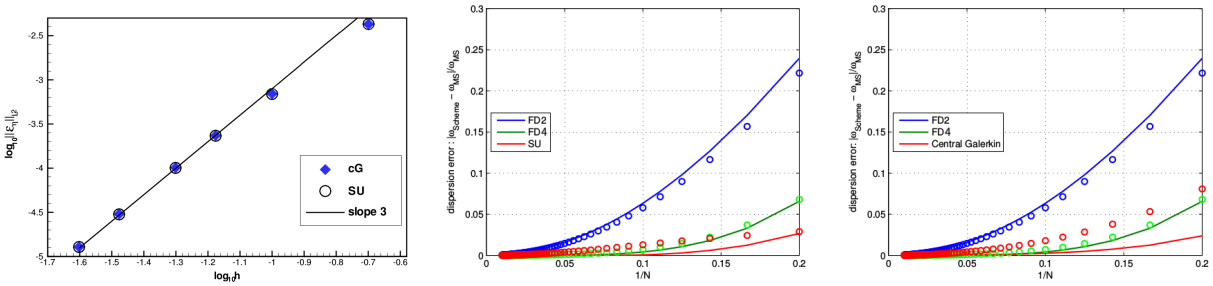


Figure 29: Accuracy of residual approximation of dispersive equations. Left: grid convergence study on an exact solitary wave solution. Middle and left: dispersion error in function of the number of nodes per wavelength for the upwind stabilized and unstabilized scheme, and comparison with second and fourth order finite differencing. Solid line: $kh_0 = 0.5$ (long wave). Circles: $kh_0 = 2.6$ ("short" wave).

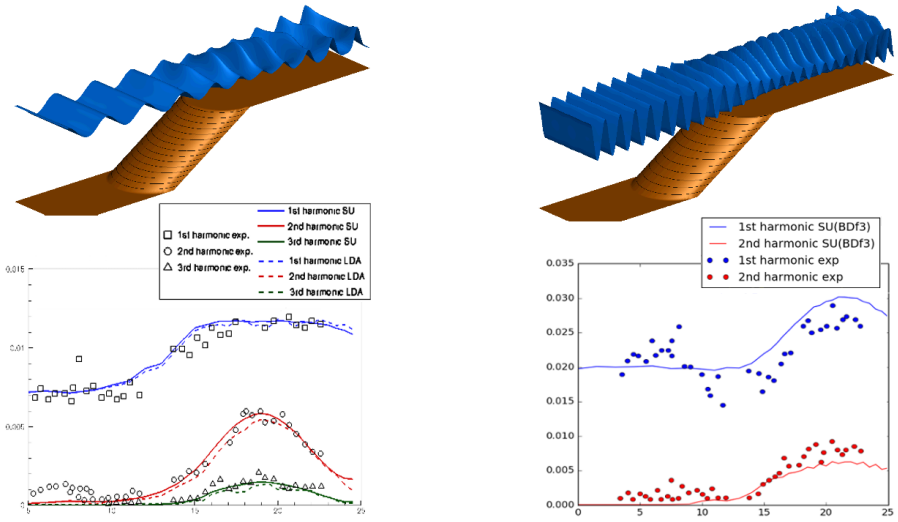


Figure 30: Wave diffraction on a circular shoal, case (a) (left) and case (b) (right).

reported on the top-left picture on figure 31. The bathymetry involves a shoal presenting a constant angle with the main incoming wave direction, with an elliptic bump which leads to a complex multidimensional wave pattern which involves dispersive effects both in the main wave direction, and along the orthogonal. As shown in the sketch on figure 31, the experiments provide the the normalized time-average of the water height in 8 different sections. Profiting from the general formulation used here, the problem is solved on an unstructured triangulation refined in correspondence of the sampling region, as shown in the picture on the top-right on figure 31.

On the same figure, the typical instantaneous wave pattern obtained is shown. One can clearly see the effect of the submerged feature in diffracting the incoming waves. To provide a more quantitative appreciation of the result, the comparison with the experiments is shown for thee of the eight sections on figure 32. The results, again obtained with two different upwind (and multidimensional upwind) stabilization approaches, confirm the potential of residual based methods in capturing complex dispersive wave phenomena.

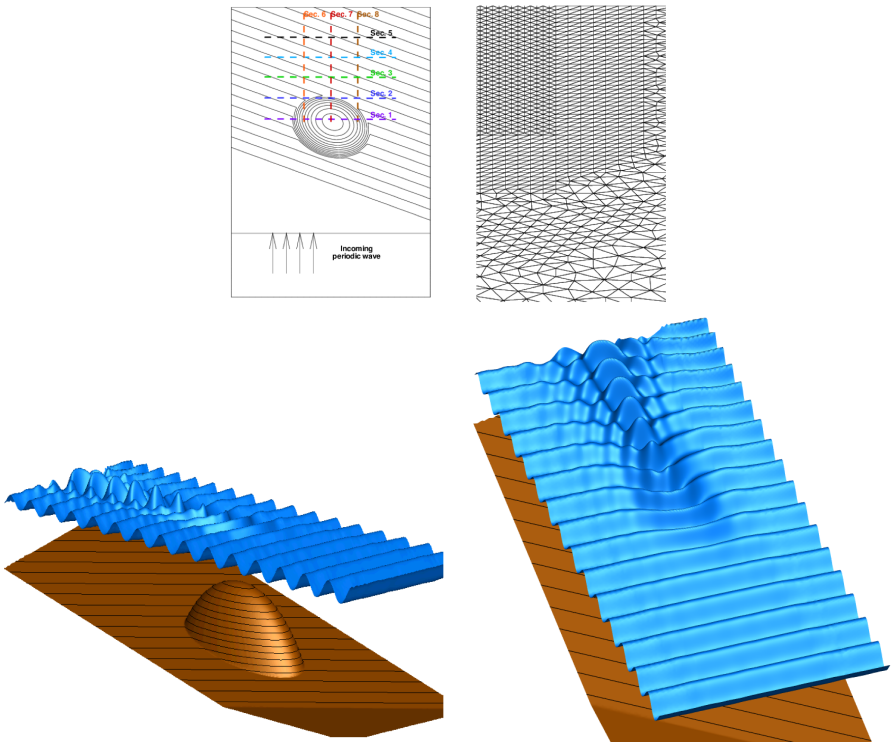


Figure 31: Wave scattering on an elliptic shoal. Problem sketch (top-left), close up of the grid (top-right), and instantaneous wave patterns (bottom).

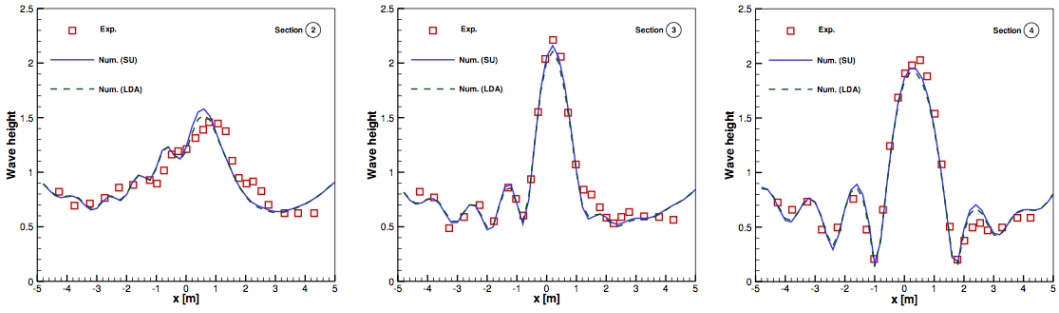


Figure 32: Wave scattering on an elliptic shoal. Time-average of the water height: simulations *vs* experiments.

5 Conclusion, Open challenges

Over the years, the Residual distribution technology has proved that continuous finite elements allow the same flexibility as discontinuous finite elements. The stencils are comparable, in particular for viscous calculations, and less degrees of freedom are always needed, eventhough the difference between completely discontinuous approxiamtion and continuous ones tend to become smaller and smaller as the polynomial degree increases. We have also shown that all these methods are *locally conservative* contrarily to a common belief. The techniques developed here shows that the schemes is very robust. We could not show all possible results, but simulation for hypersonic flows are possible without major difficulties. We have also shown (see reference), that iterative convergence to machine zero is possible even for turbulent flows, see [150].

However, all the problems have not been solved so far:

- High order and unsteady problems: this chapter has displayed a couple of solutions for geophysical flows. Other examples, related to compressible flow problems, can be found in [97] where a fully explicit method is described, or [136, 181] for implicit technique. Considering now higher than second order in time, research is still needed, however see [182] for a fully explicit (i.e. mass matrix free) technique for linear problems. The same technique can be applied for non linear problems.
- Error estimation and adaptation. Some work on adjoint problems in the RD framework has been done by [93, 183].
- p-adaptation : continuous and discontinuous approximation. Some work in that direction has been done in [184, 185]

Acknowledgements

This work would not have been possible without the contributions, suggestions and help of our collaborators, students, and friends (in alphabetical order): P. Congedo (INRIA), H. Deconinck (VKI), S. D'Angelo (VKI, now Credo Consulting), D. De Santis (INRIA, now Stanford University), A. Filippini (Inria) M. Hubbard (Nottingham University), A. Larat (INRIA, now Ecole Centrale Paris), M. Vymazal (VKI, nom Imperial College) and many others. RA has been conducting this work with the partial support of SNFS grant # 200021_153604. MR has been partly funded by the TANDEM project (reference ANR-11-RSNR-0023-01 French *Programme Investissements d'Avenir*.

References

[1] R. Abgrall and C.W. Shu, editors. *Chapters for the Handbook of Numerical Methods for Hyperbolic Problems*, volume 17 of *Handbook of Numerical Analysis*. Elsevier, 2016.

[2] D. A. di Pietro and A. Ern. *Mathematical aspects of Discontinuous Galerkin methods*, volume 69 of *Mathématiques et Applications*. Springer, 2012.

[3] E. Godlewski and P.A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118 of *Applied Mathematical Sciences*. Springer, 1996.

[4] F. Dubois and P. Le Floch. Boundary conditions for nonlinear hyperbolic systems of conservation laws. *Journal of Differential Equations*, 71:93–122, 1988.

[5] A. Harten. On the symmetric form of conservation laws with entropy. *J. Comp. Phys.*, 49:151–164, 1983.

[6] Th.J.R. Hughes, L.P. Franca, and M. Mallet. Finite element formulation for Computational Fluid Dynamics : I symmetric forms of the compressible Euler and Navier Stokes equations and the secoond law of thermodynamics. *Computer Methods in Applied Mechanics and Engineering*, 54:223–234, 1986.

[7] R. LeVeque, editor. *Finite volume methods for hyperbolic problems*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2002.

[8] E.F. Toro. *Riemann solvers and numerical methods for fluid dynamics*. Springer, 1997.

[9] B. van Leer. Toward the ultimate conservative difference scheme V. a second order sequel to Godunov's method. *Journal of Computational Physics*, 32(101), 1979.

[10] P.K. Sweby. High resolution schemes using limiter for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, 995-1011, 1984.

[11] S.R. Chakravarthy and S.J. Osher. A new class of high accuracy tvd schemes for hyperbolic conservation laws. *AIAA paper*, 85:0363, 1985.

[12] J. Goodman and R. LeVeque. On the accuracy of stable schemes for 2D scalar conservation laws. *Mathematics of Computation*, 45(171):15–21, 1985.

[13] A. Harten and S. Osher. Uniformly high-order accurate nonoscillatory schemes I. *SIAM J. Numer. Anal.*, 24(2):279–309, 1987.

[14] A. Harten, B. Engquist, S. Osher, and S.R. Chakravarthy. Uniformly high order accurate essentially non-oscillatory schemes, III. *Journal of computational physics*, 71(2):231–303, 1987.

[15] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77(2):439–471, 1988.

[16] Chi-Wang Shu and Stanley Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes, ii. *Journal of Computational Physics*, 83(1):32–78, 1989.

[17] R. Abgrall. On essentially non-oscillatory schemes on unstructured meshes: Analysis and implementation. *Journal of Computational Physics*, 114(1):45–58, 1994.

[18] Xu-Dong Liu, Stanley Osher, and Tony Chan. Weighted essentially non-oscillatory schemes. *Journal of computational physics*, 115(1):200–212, 1994.

[19] Chi-Wang Shu. High order weighted essentially nonoscillatory schemes for convection dominated problems. *SIAM review*, 51(1):82–126, 2009.

[20] O. Friedrich. Weighted essentially non-oscillatory schemes for the interpolation of mean values on unstructured grids. *Journal of Computational Physics*, 144(1):194–212, July 1998.

[21] C. Hu and C.-W. Shu. Weighted essentially non-oscillatory schemes on triangular meshes. *Journal of Computational Physics*, 150:97–127, 1999.

[22] J. Zhu, J. Qiu, C. Shu, and M. Dumbser. Runge-Kutta discontinuous Galerkin method using WENO limiters II: unstructured meshes. *J. Comput. Phys.*, 227(9):4330–435, 2008.

[23] W.H. Reed and T.R. Hill. Triangular mesh methods for the neutron transport equation. Technical Report LA-UR-73-0479, Los Alamos Scientific Laboratory, 1973. `http://lib-www.lanl.gov/cgi-bin/getfile?00354107.pdf`.

[24] P. Lesaint and P.A. Raviart. On a finite element method for solving the neutron transport equation. In *Mathematical Aspects of Finite Elements in partial Differential Equations*, number 33 in Math. Res. Center, pages 89–123, New York, 1974. University of Wisconsin-Madison, Academic Press.

[25] C. Johnson and J. Pitkäranta. An analysis of the discontinous galerkin method for a scalar hyperbolic equation. *Mathematics of Computation*, 46:1–26, 1986.

[26] G. Chavent and B. Cockburn. The local projection p^0p^1-discontinuous galerkin finite element method for scalar conservation laws. *RAIRO Modél. Math. Anal. Numér.*, 23(4):565–592, 1989.

[27] B. Cockburn and C.W. Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element methid for conservation laws. III. *Journal of Computational Physics*, 84(1):90–113, 1989.

[28] B. Cockburn and C.W. Shu. The local Runge-Kutta local projection discontinuous finite element method for scalar conservation laws. *RAIRO Modél. Math. Anal. Numér.*, 25(3):337–361, 1991.

[29] B. Cockburn, G.E. Karniadakis, and C.W. Shu, editors. *Discontinuous Galerkin Methods- Theory, computation and applications*, volume 11 of *Lecture Notes in Computer Sciences and Engineering*. Springer, 2000.

[30] F. Vilar, P.H. Maire, and R. Abgrall. A discontinuous galerkin discretization for solving the two-dimensional gas dynamics equations written under total lagrangian formulation on general unstructured grids. *Journal of Computational Physics*, 276(1):188–234, 2014.

[31] G. Jiang and C.W. Shu. On a cell entropy for Discontinuous Galerkin Methods. *Math. of Comp.*, 62(206):531–538, 1994.

[32] S. Osher. Rieman solvers, the entropy condition and difference approximations. *SIAM J. Numer. Anal.*, 21:217–235, 1984.

[33] B. Cockburn and C.W. Shu. TVD Runge Kutta local projection discontinuous Galerkin Finite element method for conservation laws II : general framework. *Math. of Comp.*, 52(186):411–435, 1989.

[34] R. Biswas, K. D. Devine, and J.E. Flaherty. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14:255–283, 1994.

[35] A. Burbeau, P. Sagaut, and Ch.-H. Bruneau. A problem-independent limiter for high-order Runge-Kutta discontinuous Galerkin methods. *Journal of Computational Physics*, 169(1):111–150, 2001.

[36] L. Krivodonova, J. Xin, J.-F. Remacle, N. Chevaugeonand, and J.E. Flaherty. Shock detection and limiting with discontinuous galerkin methods for hyperbolic conservation laws. *Applied Numerical Mathematic*, 48:323–338, 2004.

[37] J. Qiu and C.-W. Shu. A comparison of trouble cell indicators for runge-kutta discontinuous galerkin method using weno limiters. *SIAM J. Sci. Comput.*, 27:995–1013, 2005.

[38] G. Li and J. Qiu. Hybrid weighted essentially non-oscillatory schemes with different indicators. *J. Comput. Phys.*, 229:8105–8129, 2010.

[39] J. Qiu and C.-W. Shu:. Hermite weno schemes and their application as limiters for runge-kutta discontinuous galerkin method ii: Two dimensional case. *Computers & Fluids*, 34:642–663, 2005.

[40] D. Balsara, C. Altmann, C.D. Munz, and M. Dumbser. A sub-cell based indicator for troubled zones in RKDG schemes and a novel class of hybrid RKDG+HWENO schemes. *J. Comput. Phys.*, 226(1):586–620, 2007.

[41] M. Dumbser. Arbitrary high order PNPM schemes on unstructured meshes for the compressible navier-stokes equations. *Computers & Fluids*, 39(1):60–76, 2010.

[42] E. Burman. Consistent supg-method for transient transport problems: Stability and convergence. *Computer Methods in Applied Mechanics and Engineering*, 199(17-20):1114 – 1123, 2010.

[43] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong Stability-Preserving High-Order Time Discretisation Methods. *SIAM Review*, 43(1):89–112, 2001.

[44] S. Gottlieb and C.-W. Shu. Total variation diminishing runge-kutta schemes,. *Mathematics of Computation*, 67:73–85, 1998.

[45] Steven J Ruuth and Raymond J Spiteri. Two barriers on strong-stability-preserving time discretization methods. *Journal of Scientific Computing*, 17(1-4):211–220, 2002.

[46] Sigal Gottlieb. On high order strong stability preserving runge–kutta and multi step time discretizations. *Journal of Scientific Computing*, 25(1):105–128, 2005.

[47] E. Burman, A. Ern, and M.A. Fernandez. Explicit Runge-Kutta schemes and finite elements with symetric stabilization for first-order linear PDE systems. *SIAM J. Numer. Anal.*, 48(6):2019–2042, 2010.

[48] Q. Zhang and C.W. Shu. Error estimates to smooth solutions of Runge-Kutta discontinuous Galerkin methods for scalar conservation laws. *SIAM J. Numer. Anal.*, 42(2):641–666, 2004.

[49] Q. Zhang and C.W. Shu. Stability analysis and a priori error estimates for the third order explicit runge-kutta discontinuous galerkin method for scalar conservation laws. *SIAM J. Numer. Anal.*, 48(3):1038–1063, 2010.

[50] X. Meng, C.W. Shu, and B. Wu. Optimal error estimates for dicontinuous galerkin methods based on upwind-biased fluxes for linear hyperbolic equations. *Mathematics of Computation*, 2015.

[51] P. L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43:357–372, 1981.

[52] P.L. Roe. Fluctuations and signals - a framework for numerical evolution problems. In K.W. Morton and M.J. Baines, editors, *Numerical Methods for Fluids Dynamics*, pages 219–257. Academic Press, 1982.

[53] P. L. Roe. Linear advection schemes on triangular meshes. Technical Report CoA 8720, Cranfield Institute of Technology, 1987.

[54] P. L. Roe. "optimum" upwind advection on a triangular mesh. Technical Report ICASE 90-75, ICASE, 1990.

[55] P.L. Roe. Multidimensional upwinding : motivation and concepts. VKI-LS 1994-05, 1994. VKI Lecture series : Computational Fluid Dynamics.

[56] P.L. Roe and D. Sidilkover. Optimum positive linear schemes for advection in two and three dimensions. *SIAM J. Numer. Anal.*, 29(6):1542–1568, 1992.

[57] D. Sidilkover and P.L. Roe. Unification of some advection schemes in two dimensions. *Technical Report 95-10, ICASE*, 1995.

[58] R. Struijs, H. Deconinck, P. De Palma, P.L. Roe, and K.G. Powell. Progress on multidimensional upwind Euler solvers for unstructured grids. AIAA paper 91–1550, 1991.

[59] H. Deconinck, P.L. Roe, and R. Struijs. A multidimensional generalization of Roe's difference splitter for the Euler equations. *Computers and Fluids*, 22(2/3):215–222, 1993.

[60] H. Nishikawa, M. Rad, and P.L. Roe. A third-order fluctuation splitting scheme that preserves potential flow. 15th AIAA Computational Fluid Dynamics Conference, Anaheim, CA, USA, June 2001.

[61] H. Deconinck and M. Ricchiuto. Residual distribution schemes: foundation and analysis. In E. Stein, R. de Borst, and T.J.R. Hughes, editors, *Encyclopedia of Computational Mechanics*. John Wiley & Sons, Ltd., 2007. DOI: 10.1002/0470091355.ecm054.

[62] H. Paillere and H. Deconinck. Compact cell vertex convection schemes on unstructured meshes. In H Deconinck and B Koren, editors, *Notes on Numerical Fluid Mechanics*, pages 1–50. Vieweg-Verlag, Braunschweig, Germany, 1997.

[63] R. Abgrall. Toward the ultimate conservative scheme : Following the quest. *J. Comput. Phys*, 167(2):277–315, 2001.

[64] R. Abgrall and P.L. Roe. High-order fluctuation schemes on triangular meshes. *J. Sci. Comput.*, 19(3):3–36, 2003.

[65] M. Ricchiuto, R. Abgrall, and H. Deconinck. Application of conservative residual distribution schemes to the solution of the shallow water equations on unstructured meshes,. *J. Comput. Phys.*, 222:287–331, 2007.

[66] R. Abgrall and J. Trefilik. An example of high order residual distribution scheme using non-lagrange elements. *Journal of Scientific Computing*, 45:3–25, 2010.

[67] R. Abgrall, D. de Santis, and M. Ricchiuto. High-order preserving residual distribution schemes for advection-diffusion scalar problems on arbitrary grids. *SISC - SIAM Journal of Scientific Computing*, 36(3):A955–A983, 2014.

[68] P.G. Ciarlet and P.A. Raviart. General lagrange and hermite interpolation in $\mathbb{R}^n$ with applications to finite element methods. *Arch.Ration.Mech.Anal.*, 46:177–199, 1972.

[69] A. Ern and J.-C. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied Mathematical Sciences*. Springer, 2004.

[70] T.J. Barth. An energy look at the N scheme. Working notes, NASA Ames research center, CA, USA, 1996.

[71] R. Abgrall and T.J. Barth. Residual distribution schemes for conservation laws via adaptive quadrature. *SIAM J. Sci. Comput.*, 24(3):732–769, 2002.

[72] T.J.R. Hughes and A. Brook. Streamline upwind Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. *Comp. Meth. Appl. Mech. Engrg.*, 32:199–259, 1982.

[73] L.P. Franca, S.L. Frey, and T.J.R. Hughes. Stabilized finite element methods : I. application to the advective-diffusive model. Technical Report RR-1300, INRIA, 1990.

[74] T.J.R. Hughes, G. Scovazzi, and L.P. Franca. *Multiscale and Stabilized Methods*. John Wiley & Sons, Ltd, 2004.

[75] A. Szepessy. Convergence of a shock-capturing streamline diffusion finite element method for a scalar conservation law in two space dimensions. *Math. Comp.*, 53:527–545, 1989.

[76] C. Johnson, A. Szepessy, and P. Hansbo. On the convergence of shock capturing streamline diffusion finite element methods for hyperbolic conservation laws. *Math. Comp.*, 54:107–129, 1990.

[77] C. Johnson and A. Szepessy. On the convergence of a finite element method for a nonlinear hyperbolic conservation law. *Math. Comp.*, 49:427–444, 1990.

[78] P.B. Bochev, M.D. Gunzburger, and J.N. Shadid. Stability of the supg finite element method for transient advection-diffusion problems. *Computer Methods in Applied Mechanics and Engineering*, 193(23-26):2301 – 2323, 2004.

[79] J.B. Goodman and R.J. LeVeque. On the accuracy of stable schemes for 2d scalar conservation laws. *Mathematics of Computation*, 45(171):15–2, 1985.

[80] S.P. Spekreijse. Multigrid solution of monotone second-order discretizations of hyperbolic conservation laws. *Math. Comp.*, 49:135–155, 1987.

[81] T.J. Barth. Numerical methods for conservation laws on structured and unstructured meshes. *VKI LS 2003-05, 33rd Computational Fluid dynamics Course, von Karman Institute for Fluid Dynamics*, 2003.

[82] H.L. Atkins and C.-W. Shu. Quadrature-Free Implementation of Discontinuous Galerkin Method for Hyperbolic Equations. *AIAA Journal*, 36:775–782, 1998.

[83] D.P. Lockard and H.L. Atkins. Efficient implementations of the quadrature free Discontinuous Galerkin method. In *14th AIAA CFD Conference, Norfolk (VA)*, 1999.

[84] H.T. Huynh, Z.J. Wang, and P.E. Vincent. High-order methods for computational fluid dynamics: A brief review of compact differential formulations on unstructured grids. *Computers & Fluids*, 98:209 – 220, 2014.

[85] E. Burman, A. Ern, and M.A. Fernandez. Explicit runge-kutta schemes and finite elements with symmetric stabilization for first-order linear pde systems. *SIAM J. Numer. Anal.*, 48(6):2019–2042, 2010.

[86] E. Burman, A. Quarteroni, and B. Stamm. Interior penalty continuous and discontinuous finite element approximations of hyperbolic equations. *Journal of Scientific Computing*, 43(3):293–312, 2008.

[87] R. Abgrall and C.-W. Shu. Development of residual distribution schemes for the discontinuous galerkin method: the scalar case with linear elements. *Communications in Computational Physics*, 5(2-4):376–390, 2009.

[88] R. Abgrall. A residual method using discontinuous elements for the computation of possibly non smooth flows. *Adv. Appl. Math. Mech*, 2(1):32–44, 2010.

[89] M. Hubbard. Discontinuous fluctuation distribution. *Journal of Computational Physics*, 227(24):10125 – 10147, 2008.

[90] C.-S. Chou and C.-W. Shu. High order residual distribution conservative finite difference weno schemes for steady state problems on non-smooth meshes. *J. Comp. Phys.*, 214(3):698–724, 2006.

[91] R. Abgrall, A. Larat, M. Ricchiuto, and C. Tavé. A simple construction of very high order non-oscillatory compact schemes on unstructured meshes. *Computers & Fluids*, 38(7):1314 – 1323, 2009.

[92] R. Abgrall, A. Larat, and M. Ricchiuto. Construction of very high order residual distribution schemes for steady inviscid flow problems on hybrid unstructured meshes. *Journal of Computational Physics*, 230(11):4103 – 4136, 2011.

[93] S. D'Angelo, M. Ricchiuto, and H. Deconinck. Adjoint-based error estimation for adaptive petrov-galerkin finite element methods. *38th Lecture Series on Advanced Computational Fluid Dynamics - Adjoint methods and their application in CFD*, 2015. von Karman Institute for Fluid Dynamics.

[94] M. Vymazal, L. Koloszar, S. D'Angelo, N. Villedieu, M. Ricchiuto, and H. Deconinck. High-order residual distribution and error estimation for steady and unsteady compressible flow. In Norbert Kroll, Charles Hirsch, Francesco Bassi, Craig Johnston, and Koen Hillewaert, editors, *IDIHOM: Industrialization of High-Order Methods - A Top-Down Approach*, volume 128 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, pages 381–395. Springer International Publishing, 2015.

[95] R. Abgrall. Essentially non oscillatory residual distribution schemes for hyperbolic problems. *J. Comput. Phys*, 214(2):773–808, 2006.

[96] M. Ricchiuto and A. Bollermann. Stabilized residual distribution for shallow water simulations. *J. Comput. Phys*, 228(4):1071–1115, 2009.

[97] M. Ricchiuto and R. Abgrall. Explicit runge-kutta residual distribution schemes for time dependent problems: Second order case. *Journal of Computational Physics*, 229(16):5653 – 5691, 2010.

[98] T.J.R. Hughes and M. Mallet. A new finite element formulation for CFD IV: a discontinuity-capturing operator for multidimensional advective-diffusive systems. *Comp. Meth. Appl. Mech. Engrg.*, 58:329–336, 1986.

[99] M. Ricchiuto. An explicit residual based approach for shallow water flows. *Journal of Computational Physics*, 280:306–344, 2015.

[100] J.A. Rossmanith, , D.S. Bale, and R.J. LeVeque. A wave propagation algorithm for hyperbolic systems on curved manifolds. *J.Comput.Phys.*, 199:61–662, 2004.

[101] M. Ricchiuto. On the c-property and generalized c-property of residual distribution for the shallow water equations. *Journal of Scientific Computing*, 48:304–318, 2011.

[102] M. Ricchiuto. *Contributions to the development of residual discretizations for hyperbolic conservation laws with application to shallow water flows.* Habilitation à diriger des recherches, Université Sciences et Technologies - Bordeaux I, December 2011.

[103] P. Vernotte. Les paradoxes de la theorie continue de l?equation de la chaleur. *C.R. Acad.Sci.*, 246, 1958.

[104] C. Cattaneo. A form of heat-conduction equations which eliminates the paradox of instantaneous propagation. *C.R. Acad.Sci.*, 247, 1958.

[105] H. Nishikawa. A first-order system approach for diffusion equation. i: Second-order residual-distribution schemes. *Journal of Computational Physics*, 227(1):315 – 352, 2007.

[106] H. Nishikawa. A first-order system approach for diffusion equation. ii: Unification of advection and diffusion. *Journal of Computational Physics*, 229(11):3989 – 4016, 2010.

[107] H. Nishikawa. Beyond interface gradient: A general principle for constructing diffusion scheme. In *40th Fluid Dynamics Conference and Exhibit*. AIAA Paper 2010-5093, 2010.

[108] H. Nishikawa. Robust and accurate viscous discretization via upwind scheme I: Basic principle. *Computers & Fluids*, 49:62–86, 2011.

[109] A. Mazaheri and H. Nishikawa. Improved second-order hyperbolic residual-distribution scheme and its extension to third-order on arbitrary triangular grids. *Journal of Computational Physics*, 300:455 – 491, 2015.

[110] A. Mazaheri and H. Nishikawa. Efficient high-order discontinuous galerkin schemes with first-order hyperbolic advection-diffusion system approach. *Journal of Computational Physics*, 2016. available online.

[111] R. Abgrall, M. Ricchiuto, and D de Santis. High-order preserving residual distribution schemes for advection-diffusion scalar problems on arbitrary grids. *SIAM J. Scientific Computing*, 36(3):A955–A983, 2014.

[112] R. Abgrall and D. de Santis. Linear and non-linear high order accurate residual distribution schemes for the discretization of the steady compressible navier-stokes equations. *Journal of Computational Physics*, 283:329–359, 2015.

[113] D. De Santis. High-order linear and non-linear residual distribution schemes for turbulent compressible flows. *Computer Methods in Applied Mechanics and Engineering*, 285:1 – 31, 2015.

[114] R. Abgrall, A. Krust, M. Ricchiuto, and D. de Santis. Numerical approximation of parabolic problems by means of residual distribution schemes. *International Journal on numerical Methods in Fluids*, 71(9):1191–1206, 2013.

[115] O. C. Zienkiewicz and J. Z. Zhu. A simple error estimator and adaptive procedure for practical engineerng analysis. *International Journal for Numerical Methods in Engineering*, 24(2):337–357, 1987.

[116] M. Ricchiuto, N. Villedieu, R. Abgrall, and H. Deconinck. On uniformly high-order accurate residual distribution schemes for advection-diffusion. *Journal of Computational and Applied Mathematics*, 215(2):547 – 556, 2008.

[117] R. Struijs. *A Multi-Dimensional Upwind Discretization Method for the Euler Equations on Unstructured Grids.* PhD thesis, University of Delft, Netherlands, 1994.

[118] D.A. Caraeni. *Development of a Multidimensional Upwind Residual Distribution Solver for Large Eddy Simulation of Industrial Turbulent Flows*. PhD thesis, Lund Institute of Technology, 2000.

[119] D. Caraeni and L. Fuchs. Compact third–order multidimensional upwind scheme for Navier-Stokes simulations. *Theoretical and Computational Fluid Dynamics*, 15:373–401, 2002.

[120] J. Maerz and G. Degrez. Improving time accuracy of residual distribution schemes. Technical Report VKI-PR 96-17, von Karman Institute for Fluid Dynamics, 1996.

[121] A. Ferrante and H. Deconinck. Solution of the unsteady Euler equations using residual distribution and flux corrected transport. Technical Report VKI-PR 97-08, von Karman Institute for Fluid Dynamics, 1997.

[122] R. Abgrall and M. Mezine. Construction of second-order accurate monotone and stable residual distribution schemes for unsteady flow problems. *J. Comput. Phys.*, 188:16–55, 2003.

[123] M. Ricchiuto, Á. Csík, and H. Deconinck. Residual distribution for general time dependent conservation laws. *J. Comput. Phys*, 209(1):249–289, 2005.

[124] T.J.R. Hughes and T.E. Tezduyar. Development of time-accurate finite element techniques for first order hyperbolic systems with emphasys on the compressible euler equations. *Comp. Meth. Appl. Mech. Engrg.*, 45(1-3):217–284, 1984.

[125] F. Shakib and T.J.R. Hughes. A new finite element formulation for computational fluid dynamics: Ix. fourier analysis of space-time galerkin/least-squares algorithms. *Comp. Meth. Appl. Mech. Engrg.*, 87(1):35–58, 1991.

[126] F. Chalot and P.-E. Normand. Higher-order stabilized finite elements in an industrial navier-stokes code. In Norbert Kroll, Heribert Bieler, Herman Deconinck, Vincent Couaillier, Harmen van der Ven, and Kaare Srensen, editors, *ADIGMA - A European Initiative on the Development of Adaptive Higher-Order Variational Methods for Aerospace Applications*, volume 113 of *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, pages 145–165. Springer Berlin / Heidelberg, 2010.

[127] G. Rossiello, P. De Palma, G. Pascazio, and M. Napolitano. Third-order-accurate fluctuation splitting schemes for unsteady hyperbolic problems. *J. Comput. Phys*, 222(1):332–352, 2007.

[128] M. Mezine. *Conception de Schémas Distributifs pour l'aérodynamique stationnaire et instationnaire*. PhD thesis, École doctorale de mathématiques et informatique, Université de Bordeaux I, 2002.

[129] M. Mezine, M. Ricchiuto, R. Abgrall, and H. Deconinck. Monotone and stable residual distribution schemes on prismatic space-time elements for unsteady conservation laws. *VKI LS 2003-05, 33rd Computational Fluid dynamics Course, von Karman Institute for Fluid Dynamics* -, 2003.

[130] M. Ricchiuto, Á. Csík, and H. Deconinck. Conservative residual distribution schemes for general unsteady systems of conservation laws. In *ICCFD3 International Conference on Computational Fluid Dynamics 3*, Toronto, Canada, July 2004.

[131] M. Ricchiuto and R. Abgrall. Stable and convergent residual distribution for time-dependent conservation laws. In *ICCFD4 International Conference on Computational Fluid Dynamics 4*, Ghent, Belgium, July 2006.

[132] M. Ricchiuto, R. Abgrall, and H. Deconinck. Construction of very high order residual distribution schemes for unsteady advection: preliminary results. *VKI LS 2003-05, 33rd Computational Fluid dynamics Course, von Karman Institute for Fluid Dynamics*, 2003.

[133] L. Koloszár, N. Villedieu, T. Quintino, P. Rambaud, H. Deconinck, and J. Anthoine. Residual distribution method for aeroacoustics. *AIAA Journal*, 49(5):1021–1037, 2011.

[134] R. Abgrall, N. Andrianov, and M. Mezine. Towards very high-order accurate schemes for unsteady convection problems on unstructured meshes. *International Journal for Numerical Methods in Fluids*, 47(8-9):679–691, 2005.

[135] M. Hubbard and M. Ricchiuto. Discontinuous upwind residual distribution: A route to unconditional positivity and high order accuracy. *Computers and Fluids*, 46(1):263 – 269, 2011.

[136] D. Sarmany, M. Hubbard, and M. Ricchiuto. Unconditionally stable space-time discontinuous residual distribution for shallow water flows. 2013.

[137] M. Hubbard and P.L. Roe. Compact high resolution algorithms for time dependent advection problems on unstructured grids. *Int. J. Numer. Methods Fluids*, 33(5):711–736, 2000.

[138] M. Ricchiuto and H. Deconinck. Time accurate solution of hyperbolic partial differential equations using fct and residual distribution. VKI report VKI SR1999-33, September 1999.

[139] G. Rossiello, P. De Palma, G. Pascazio, and M. Napolitano. Second-order-accurate explicit fluctuation splitting schemes for unsteady problems. *Computers & Fluids*, 38(7):1384 – 1393, 2009.

[140] J. Klosa. Extrapolated BDF residual distribution schemes for the shallow water equations. Master thesis, 2012.

[141] J. Klosa, M. Ricchiuto, and R. Abgrall. Multi-step and multi-stage explicit time stepping for residual distribution. application to shallow water flows. in preparation, 2016.

[142] G. Cohen, P. Joly, J.E. Roberts, and N. Tordjman. High order triangular finite elements with mass lumping for the wave equation. *SIAM J. Numer. Anal.*, 38(6):2047–2078, 2001.

[143] F.X. Giraldo and M.A. Taylor. A diagoanl mass-matrix triangular spectral element method based on cubature points. *J. Eng. Math.*, 56:307–322, 2006.

[144] Y. Xu. Gauss-Lobatto integration on the triangle. *SIAM J.Numer.Anal.*, 49:541–548, 2011.

[145] W.A. Mulder. New triangular mass-lumped finite elements of degree six for wave propagation. *Progress in Electromagnetics Research*, 141:671–692, 2013.

[146] D.F. Rogers. *An Introduction to NURBS: with Historical Perspectives*. Morgan Kaufman, San Mateo, 2001.

[147] M. Vymazal. *Very high order residual distribution schemes via variable distribution coefficients*. PhD thesis, Université Libre de Bruxelles and von Karman Institute for Fluid Dynamics, 2016. under review.

[148] T. Warburton. An explicit construction of interpolation nodes on the simplex. *Journal of Engineering Mathematics*, 56(3):247–262, 2006.

[149] R. Abgrall and D. de Santis. Linear and non-linear high order accurate residual distribution schemes for the discretization of the steady compressible navier-stokes equations. *Journal of Computational Physics*, 283:329–359, 2015.

[150] D. De Santis. High-order linear and non-linear residual distribution schemes for turbulent compressible flows.

[151] R. Abgrall. Toward the ultimate conservative scheme: Following the quest. *J. Comput. Phys*, 167(2):277–315, 2001.

[152] F. Ringleb. Exakte Lösungen der Differentialgleichungen einer abadiatischen Gassströmung. *ZAMM*, 20(4):185–198, 1940.

[153] R. von Mises. Dover, 1958. Unabridged republication of the work first published by Academic Press Inc.

[154] T. Leicht and R. Hartmann. Error estimation and anisotropic mesh refinement for 3d laminar aerodynamic flow simulations. *Journal of Computational Physics*, 229(19):7344 – 7360, 2010.

[155] P. Garcia-Navarro, M.E. Hubbard, and A. Priestley. Genuinely multidimensional upwinding for the 2d shallow-water equations. *J.Comput.Phys*, 121(1):79–93, 1995.

[156] M.E. Hubbard and M.J. Baines. Conservative multidimensional upwinding for the steady two-dimensional shallow water equations. *J.Comput.Phys*, 138(2):419–448, 1997.

[157] P. Brufau and P. Garcia-Navarro. Unsteady free surface flow simulation over complex topography with a multidimensional upwind technique. *Journal of Computational Physics*, 186(2):503–526, 2003.

[158] M. Ricchiuto and A.G. Filippini. Upwind residual discretization of enhanced Boussinesq equations for wave propagation over complex bathymetries. *Journal of Computational Physics*, 271:306–341, 2014.

[159] A.G. Filippini, M. Kazolea, and M. Ricchiuto. A flexible genuinely nonlinear approach for wave propagation, breaking and runup. *J.Comput.Phys*, 310:381–417, 2016.

[160] Benchmark problem #2, Tsunami runup onto a complex three-dimensional beach. The third international workshop on long-wave runup models, http://isec.nacse.org/workshop/2004_cornell/bmark2.html.

[161] Tsunami runup onto a complex three-dimensional beach; Monai valley. Benchmarks of the NOAA Center for tsunami research, http://nctr.pmel.noaa.gov/benchmark/Laboratory/Laboratory_MonaiValley/.

[162] P. L.-F. Liu, H. Yeh, and C. Synolakis, editors. *Advanced Numerical Models for Simulating Tsunami Waves and Runup*, volume 10 of *Advances in Coastal and Ocean Engineering*. World Scientific, 2008.

[163] J.T. Kirby. Chapter 1 boussinesq models and applications to nearshore wave propagation, surf zone processes and wave-induced currents. In V.C. Lakhan, editor, *Advances in Coastal Modeling*, volume 67 of *Elsevier Oceanography Series*, pages 1–41. Elsevier, 2003.

[164] M. Brocchini. A reasoned overview on boussinesq-type models: the interplay between physics, mathematics and numerics. *Proc. Royal Soc. A*, 469, 2014.

[165] D. Lannes. *The Water Waves Problem: Mathematical Analysis and Asymptotics*. American Mathematical Society, Providence, Rhode Island, 2013.

[166] M. Tonelli and M. Petti. Simulation of wave breaking over complex bathymerties by a Boussinesq model. *J. Hydraulic Res*, 49:473–486, 2011.

[167] P. Bonneton, F. Chazel, D. Lannes, F. Marche, and M. Tissier. A spliting approach for the fully nonlinear and weakly dispersive green-naghdi model. *J.Comput.Phys*, 230:1479–1498, 2011.

[168] M. Kazolea, A. I. Delis, and C. E. Synolakis. Numerical treatment of wave breaking on unstructured finite volume approximations for extended Boussinesq-type equations. *J.Comput.Phys.*, 271:281–305, 2014.

[169] P. Bacigaluppi, M. Ricchiuto, and P. Bonneton. A 1d stabilized finite element model for non-hydrostatic wave breaking and run-up. In J. Fuhrmann, M. Ohlberger, and C. Rohde, editors, *Finite Volumes for Complex Applications VII*, volume 77 of *Springer Proceedings in Mathematics and Statistics*. Springer, 2014.

[170] G. Wei and J. T. Kirby. A time-dependent numerical code for extended Boussinesq equations. *Journal of Waterway, Port, Coastal, and Ocsean Engineering*, 120:251–261, 1995.

[171] O.R. Sørensen P.A. Madsen. A new form of the boussinesq equations with improved linear dispersion characteristics. part 2: a slowing varying bathymetry. *Coastal Engineering*, 18:183–204, 1992.

[172] S. Beji and K. Nadaoka. A formal derivation and numerical modeling of the improved boussinesq equations for varying depth. *Ocean Engineering*, 23, 1996.

[173] M.A. Walkley and M. Berzins. A finite element method for the two-dimensional extended boussinesq equations. *Int. J. Numer. Meth Fluids*, 39, 2002.

[174] L. Sørensen O.R. Sørensen, H. Schaffer. Boussinesq-type modelling using an unstructured finite element technique. *Coastal Engineering*, 50, 2004.

[175] C. Eskilsson, S.J. Sherwin, and L. Bergdhal. An unstructured spectral/*hp* element model for enhanced boussinesq-type euations. *Coastal Engineering*, 53, 2006.

[176] M. Tonelli and M. Petti. Hybrid finite volume - finite difference scheme for 2dh improved boussinesq equations. *Coastal Engineering*, 56, 2009.

[177] M. Kazolea, A.I. Delis, I.K. Nikolos, and C.E. Synolakis. An unstructured finite volume numerical scheme for extended 2d boussinesq-type equations. *Coastal Engineering*, 69:42–66, 2012.

[178] R.W. Whalin. The limit of applicability of linear wave refraction theory in a convergence zone. Res.Rep.H-71-3, USACE,Waterways Expt. Station, Vicksburg, MS, 1971.

[179] H.A. Schaffer and P.A. Madsen. Further enhancements of boussinesq-type equations. *Coastal Engineering*, 26:1–14, 1995.

[180] J.C.W. Berkhoff, N. Booy, and A.C. Radder. Verification of numerical wave propagation models for simple harmonic linear water waves. *Coastal Engineering*, 6:255–279, 1982.

[181] R. Abgrall, N. Andrianov, and M. Mezine. Towards very high-order accurate schemes for unsteady convection problems on unstructured meshes. *Int. J. Numer. Methods Fluids*, 47(8-9):679–691, 2005.

[182] R. Abgrall, P. Pacigaluppi, and S. Tokareva. How to avoid mass matrix for linear hyperbolic problems. In B. Karasözen, M. Manguoglu, M.and Tezer-Sezgin, S. Göktepe, and Ö. Ugur, editors, *Numerical Mathematics and Advanced Applications- ENUMATH 2015*, Lecture notes in Computational Science and Engineering. Springer, 2016.

[183] S. D'Angelo, M. Ricchiuto, R. Abgrall, and H. Deconinck. Generalized framework for adjoint error estimation of PG method in linear problems. Research Report RR-7613, INRIA, 2011.

[184] R. Abgrall, H. Beaugendre, C. Dobzinsky, and Q. Viville. $p-$ adaptation is possible with continuous finite elements: the Euler equations case. *Journal of Scientific Computing*, 2016. in revision.

[185] R. Abgrall, H. Beaugendre, C. Dobzinsky, and Q. Viville. $p-$ adaptation is possible with continuous finite elements: the Navier Stokes equations case. 2016. in preparation.