A simple construction of very high order non-oscillatory compact schemes on unstructured meshes

R. Abgrall *, A. Larat, M. Ricchiuto, C. Tavé

Project ScAlAppx, INRIA Futurs and Institut de Mathématiques Université Bordeaux 1, 33405 Talence Cedex, France

A R T I C L E   I N F O
Article history:
Received 1 September 2007
Accepted 23 January 2008
Available online 20 April 2008

A B S T R A C T
In [Abgrall R, Roe PL. High order fluctuation schemes on triangular meshes. J Sci Comput 2003;19(1–3):3–36] have been constructed very high order residual distribution schemes for scalar problems. They were using triangle unstructured meshes. However, the construction was quite involved and was not very flexible. Here, following [Abgrall R. Essentially non-oscillatory residual distribution schemes for hyperbolic problems. J Comput Phys 2006;214(2):773–808], we develop a systematic way of constructing very high order non-oscillatory schemes for such meshes. Applications to scalar and systems problems are given.

1. Introduction
We are interested in the approximation of the following model problem
\[ \begin{align*}
\vec{J} \cdot \nabla u &= f \quad x \in \Omega \\
u &= 0 \quad x \in \Gamma^- 
\end{align*} \] (1)
where \( \Omega \subset \mathbb{R}^d \) is a polygonal set, \( \Gamma^- \) is the inflow boundary
and \( \vec{h}_i \) is the local normal at the point \( x \in \partial \Omega \).

We consider a conformal triangulation \( \mathcal{F}_h \) which elements \( K \) are triangles, quads in 2D or tetrahedrons/hex in 3D. More general elements could, in principle, be considered. The parameter \( h \) denotes the maximum value of all the diameters \( h_K \) of the circumscribed circle/sphere to the elements \( K \) of \( \mathcal{F}_h \). We also assume that the meshes are regular. The mesh \( \mathcal{F}_h \) are assumed to be adapted to (1), i.e. \( \Gamma^- \) is a collection of edges/faces of \( \mathcal{F}_h \). The space \( \mathbb{V}_h^p \) is the set of continuous functions that, on each element \( K \), are polynomials of degree \( p \) that vanishes on \( \Gamma^- \).

The Eq. (1) is discretized by a variational formulation of the type:
Let us give \( p, q \in \mathbb{N}^\circ \), find \( v^h \in \mathbb{V}_h^{p} \) such that for all \( \phi^h \in \mathbb{V}_h^{q} \),
\[ a(v^h, \phi^h) = \ell(v^h). \] (2)

The examples we are interested in are the SUPG schemes [5,6] and the stabilized residual distribution schemes [1] which are in general non-linear schemes, even for linear problems.

In the SUPG schemes, we take \( p = q \in \mathbb{N}^\circ \) and the relation (2) writes
\[ \int_{\Omega} \phi^h (\vec{J} \cdot \nabla u^h - f(x)) dx + h \int_{\Omega} (\vec{J} \cdot \nabla \phi^h) (\vec{J} \cdot \nabla u^h - f(x)) dx = 0 \] (3)
i.e.
\[ a(v^h, u^h) = \int_{\Omega} \phi^h (\vec{J} \cdot \nabla u^h) dx + h \int_{\Omega} (\vec{J} \cdot \nabla \phi^h) (\vec{J} \cdot \nabla u^h) dx \] (4)
\[ \ell(v^h) = \int_{\Omega} \phi^h f(x) dx + h \int_{\Omega} (\vec{J} \cdot \nabla \phi^h) f(x) dx \]

In the second example, the Residual Distribution schemes (RD schemes for short), we also take \( q = p \in \mathbb{N}^\circ \) but the formulation is completely different in order to account, for example, of a maximum principle. These schemes are described in Section 2.

The solutions of (2) are obtained by an iterative scheme. The convergence of the iterative procedure is important for two reasons
1. The uniqueness of the solutions of (2) is essential to have a well posed problem, and one wishes to obtain a good approximation of the solution of (2).
2. One can show, and we recall this later in the text, that if the problem (2) is not solved with enough precision, the formal accuracy of the scheme (2) is lost.

In this respect, the SUPG scheme is dissipative, and coercive in a proper norm, so that existence and uniqueness is guaranteed. However, in the case of the RD schemes, this may be no longer true, at least for their unstabilized version. Indeed, a RD scheme can be constructed so that it is positivity preserving but in general, the solution of (2) may not be unique, see [1].

Coming back to the SUPG scheme, we see that the forms \( a \) and \( b \) are the sum of two terms, the forms

**Corresponding author. Tel.: +33 540 00 60 68; fax: +33 540 00 26 26. E-mail address: remi.abgrall@math.u-bordeaux1.fr (R. Abgrall).**
\( a'(u^h, v^h) = \int_\Omega v^h (\tilde{\nabla} \cdot \nabla u^h) \, dx \)

and

\( c'(v^h) = \int_\Omega v^h f \, dx \)

date that define the Galerkin formulation of (1). The problem (2) with \( \alpha \) and \( \beta \) is known to be very unstable. It is stabilized by adding a dissipative term \( q \) and a linear form \( b \) to keep the consistency of the scheme,

\[
q(u^h, v^h) = h \sum_{K} \mathcal{D}_K(v^h, u^h), \quad \mathcal{D}_K(v^h, u^h) = \int_K (\tilde{\nabla} v^h) (\tilde{\nabla} u^h) \, dx \\
b(v^h) = h \int_K (\tilde{\nabla} v^h) f(x) \, dx
\]

The exact evaluation of \( \mathcal{D}_K \) may be quite costly in practice. If in the case of \( p = 1 \), the terms \( (\tilde{\nabla} \cdot \nabla u^h)(\tilde{\nabla} \cdot \nabla v^h) \) can be evaluated with second order accuracy with only one point (the centroid). For \( p = 2 \), the components of \( \nabla v_3 \) are of degree one, and an exact quadrature formula (for a constant velocity) \( \tilde{\nabla} \) is obtained with 3 quadrature points (the mid-points of the edges of the triangle) of a 5 point formula as indicated in Table 1. When \( p = 3 \), \( (\tilde{\nabla} \cdot \nabla u^h)(\tilde{\nabla} \cdot \nabla v^h) \) is of degree 4 and 7 quadrature points are needed. This can be seen from Table 1 where weights and quadrature points are displayed for triangular elements. For \( Q^k \) elements, the situation is worse.

The question we are interested in this paper is the following. Given a scheme of the type (2),

\( a(u^h, v^h) = \beta(v^h) \)

what are the requirements about the forms \( q \) and \( b \) such that the scheme

\( a(u^h, v^h) + q(u^h, v^h) + \epsilon(v^h) + b(v^h) \)

is well posed, has provable error estimates in a well behaved norm? How can \( q \) and \( b \) be chosen such that the evaluation of these terms is as simple as possible with the minimal number of operations?

The schemes we are interested in, like (3) or the RD scheme, share several formal properties in common. Namely,

1. If \( u \) is a smooth solution of (1), then for any \( v^h \in V^h_h \), we have

\( a(u, v^h) = \beta(v^h) \)  

Moreover, if \( u^h \) denotes the solution of scheme, we have

\( a(u - u^h, v^h) = 0. \)

Note that this property, which is well known for the SUPG scheme, is also true for the RDS scheme, even if the RDS scheme is non-linear.

2. From this, if \( u^h \) denotes now the interpolant of the exact solution \( u \) of (1), then the equivalent equation of the scheme is

\( a(u^h, v^h) + \epsilon(v^h) = O(h^{p-d}) \)

from which we deduce the formal order of accuracy. Of course, in the case of the SUPG scheme, things can be made more rigorous.

These properties must remain intact.

In the first section, we explain in detail what is a RD scheme. The SUPG schemes are particular cases. The second section is devoted to the describe and discuss natural necessary conditions. The third section is devoted to examples and numerical results.

2. Examples of “unstabilised schemes”

The example of the Galerkin formulation of (1) is well known so we skip it. We give some details on the RD schemes that are less known.

We consider a conformal mesh, the generic element is denoted by \( K \). The degrees of freedom are denoted by \( x_{r} \). In the case of a \( P^k \) interpolant, they are just the vertices of the mesh. For a \( P^{k+1} \) interpolant, we have to add the mid-edge points, etc. Obvious generalization can be described for other continuous elements such as the \( P^k \) or \( Q^k \) elements.

In order to construct a RD scheme for (1), on has first to construct “residuals” \( \Phi^\sigma_\sigma \) such that the two conditions are met :

1. Compact stencil condition: \( \Phi^\sigma_\sigma(u^h) = \Phi^\sigma_\sigma(u^h) \) only depends on the values of \( u \) at the degrees of freedom in \( K \).
2. Conservation condition : \( \Phi^\sigma_\sigma(u^h) = 0. \)

This is a conservation constraint.

The function \( u^h \) has to be solution of

\[
\sum_{r \in K} \Phi^\sigma_\sigma \right|_{x=r} \, u^h \, dx = \int_K f(x) \, dx := \Phi^K
\]

for any \( \sigma \), \( \sum_{r \in K} \Phi^\sigma_\sigma = 0. \)

As said previously, the SUPG schemes are examples of RD schemes, since they are exactly (10) with

\[
\Phi^\sigma_\sigma = \int_K \delta_{\sigma}(\tilde{\nabla} u^h - f) \, dx + h \int_K (\tilde{\nabla} \phi_{\sigma})(\tilde{\nabla} u^h - f) \, dx
\]
Of course these conditions (conservation and a compact stencil) are not enough to provide a working scheme in term of stability and accuracy. Here we focus on the $L^\infty$ stability and the residual property 8a, 8b which ensure formal accuracy. These two additional constraints are achieved by the following procedure.

One starts from a monotone scheme, say the Lax–Friedrichs one,
\[ \Phi_{1}^{x} = \frac{1}{N_{K}} \left( \Phi_{x}^{K} + \alpha_{x} \sum_{\sigma \in K} (u_{\sigma} - u_{x}) \right) \]
which is only first order, $N_{K}$ represents the number of degree of freedom in $K$. Then we define
\[ X_{\sigma} := \frac{\Phi_{x}^{\sigma}}{\Phi_{x}^{K}}. \]
they sum up to unity thanks to the conservation relation and
\[ \beta_{x}^{K} := \sum_{\sigma \in K} X_{\sigma} \beta_{x}^{\sigma}. \] (11)
There is no problem in the definition of $\beta_{x}^{\sigma}$ since
\[ \sum_{\sigma \in K} X_{\sigma} \geq \sum_{\sigma \in K} X_{\sigma} = 1. \]
The RD scheme is then defined by (10) with
\[ \Phi_{x}^{K} = \beta_{x}^{K} \Phi_{x}^{K}. \] (12)

The solution of (18)–(12) is sought for by an iterative method. The simplest one is
\[ u_{x}^{n+1} = u_{x}^{n} - \phi_{x}^{K} \sum_{k \in K} \Phi_{x}^{k} \text{ for all } \sigma \] (13)
with $u_{x}^{0} = 0$ for example, and one hopes that $u_{\sigma} = \lim_{n \to \infty} u_{x}^{n}$. Thanks to the definition of $\beta_{x}^{\sigma}$, one can see that the sequence \[ \{u_{x}^{n}\}_{n=0}^{\infty} \]
maximal principle provided a CFL-like condition
\[ 0 \leq \alpha_{x} \max_{k \in K} \max_{\sigma \in K} \max_{x \in K} (\nabla \phi_{x}(x)) \leq 1. \]
Note that sharper estimates can be given, but this is not the point here.

The variational formulation of (10)–(12) is easily obtained. If one multiply (10) by $\psi_{x}(x)$ and sum over all the degrees of freedom, one obtains
\[ 0 = \sum_{K} \Psi_{K} \]
with
\[ \Psi_{K} = \sum_{\sigma \in K} \phi_{x}^{\sigma} \int_{K} (\nabla u_{x}^{h} - f) dx \]
\[ = \int_{K} v(\nabla u_{x}^{h} - f) dx + \sum_{\sigma \in K} \int_{K} (\beta_{x}^{\sigma} - \phi_{x}^{K}) (\nabla u_{x}^{h} - f) dx \]
Since by definition
\[ \sum_{\sigma \in K} \beta_{x}^{\sigma} = 1 = \sum_{\sigma \in K} \phi_{x}^{K}, \]
we have
\[ \sum_{\sigma \in K} \phi_{x}^{K} \int_{K} (\beta_{x}^{\sigma} - \phi_{x}^{K}) (\nabla u_{x}^{h} - f) dx \]
\[ = \frac{1}{(N_{K} - 1)!} \sum_{\sigma \in K} \phi_{x}^{K} (\psi_{x}^{\sigma} - \phi_{x}^{K}) \int_{K} (\nabla u_{x}^{h} - f) dx \]
\[ = \frac{h}{(N_{K} - 1)!} \sum_{\sigma \in K} \phi_{x}^{K} (\psi_{x}^{\sigma} - \phi_{x}^{K}) \int_{K} (\nabla u_{x}^{h} - f) dx \]
with $\phi_{x}^{K} = ||x_{0} x_{\sigma}|| / h$ which is bounded since the mesh is regular, $\psi_{x}^{\sigma} = \phi_{x}^{\sigma} - \phi_{x}^{K}$ and $\psi_{x}^{\sigma} = \phi_{x}^{K} - \phi_{x}^{\sigma}$.

The form $\alpha$ is
\[ a(u^{h}, v^{h}) = \int_{K} v^{h}(\nabla u^{h}) dx + h \sum_{\sigma \in K} \sum_{\sigma \in K} \phi_{x}^{K} \int_{K} (\nabla u^{h} - f) dx \]
\[ \times \int_{K} (\psi_{x}^{\sigma} - \psi_{x}^{K}) (\nabla u^{h}) dx \]
and $\ell$ is
\[ \ell(v^{h}) = \int_{K} v^{h} f dx + \frac{h}{(N_{K} - 1)!} \sum_{\sigma \in K} \phi_{x}^{K} \int_{K} (\nabla u^{h} - f) dx \]
which have the same structure as (4).

The problem of this scheme is that even though the iteration (13) is $L^\infty$ stable, it does not converge in general. The same conclusion holds for more involved iterative scheme and the reason is that (10)–(12) is not well posed except for very special situations.

An example is given for a second order (hence $P^2$ interpolation) using the local Lax–Friedrichs scheme (12) on
\[ -y \frac{\partial u}{\partial x} + x \frac{\partial u}{\partial y} = 0 \quad (x, y) \in [0, 1]^2 \] (16)
The convergence history and a solution is given on Fig. 1. The solution of Fig. 1 is obviously not a second order accurate approximation of (16). The next section is devoted to describe a simple modification of the scheme. This problem has already been solved in [1] for second order schemes, we show how to extend the method in a simple and efficient way.

3. Construction and discussion

Let us consider the problem
\[ a(u^{h}, v^{h}) = \ell(v^{h}) \] (17a)
with $\alpha$ and $\ell$ given by
\[ a(u^{h}, v^{h}) = \sum_{k \in K} \alpha_{k} (u^{h}, v^{h}) \]
\[ a_{k}(u^{h}, v^{h}) = \int_{K} v_{k} (\nabla u^{h}) dx + h b_{k}(u_{h}, v_{h}) \]
\[ \ell(v^{h}) = \sum_{k \in K} \ell_{k}(v^{h}) \]
\[ \ell_{k}(v^{h}) = \int_{K} v_{k} f dx + h b_{k}(v_{h}) \]
We assume that $\alpha$, $\ell$, $a_{k}$ and $\ell_{k}$ satisfy the following assumptions :

Assumption 3.1.

1. $a_{k}$ and $\ell_{k}$ are linear in $v_{h}$.
2. if $u$ is the solution of (1),
\[ a(u, v^{h}) = \ell(v^{h}) \]
for any $v^{h} \in V_{h}$ and
\[ a(u - u^{h}, v^{h}) = 0 \]
for any $v^{h} \in V_{h}$. More precisely, because of the structure of the forms $a$ and $b$, we assume that for any $K$, and any $v_{h} \in V_{h}^{0}$,
\[ a_{k}(u, v^{h}) = \ell_{k}(v_{h}) \]
and
\[ h b_{k}(u - u^{h}, v_{h}) = O(h^{3-d}). \]

Note that other RD scheme exist, they do not satisfy a $L^\infty$ stability property. An example is the SUPG scheme, another one is the LDA scheme, see [4,10].
These assumptions are true for the SUPG and RD schemes. Moreover, for these two schemes, we have the conservation constraint
\[ a_e(u^h, 1) = I_k(1) = 0 \]
for any \( K \).

**Remark 3.1 (About the linearity assumption).** The problem (1) is linear. All what is said here can be extended to the non-linear case, and the linearity assumption still holds.

The scheme writes in the RD form (14). To see this, we consider the list of degrees of freedom \( \{ x_e \} \). For piecewise linear interpolant and triangular elements or \( Q^1 \) interpolant, they are just the vertices of the mesh. For quadratic interpolant and triangular meshes, they are the vertices of the mesh and the mid-edges points, etc. The Lagrange interpolant of degree \( p \) associated to a given degree of freedom \( x_e \) is denoted as \( \varphi^p_e \). We have

1. \( \varphi^p_e(x_e) = \delta^p_e \),
2. \( \varphi^p_e \) is continuous,
3. for any element \( T \), the restriction of \( \varphi^p_e \) to \( T \) is a polynomial of degree \( d \).

By definition, any \( u^h \in V^0_h \) can be written as
\[ u^h = \sum_{\sigma} u^h(x_e) \varphi^p_e, \]
and \( v^h \in V^0_h \) can be written as
\[ v^h = \sum_{\sigma} v^h(x_e) \varphi^p_e, \]
so that the scheme can be reformulated as finding \( u^h \in V^0_h \) such that for any \( \sigma \), we have
\[ \sum_{\sigma \subset T} \left\{ \int_{\sigma} \varphi^p_e \langle \hat{\mathbf{i}} \cdot \nabla u^h \rangle \, dx + h_e a_e(\varphi^p_e, u^h) \right. \
+ \left. \int_{\sigma} f(x) \varphi^p_e(x) \, dx + h_e k_e(\varphi^p_e) \right\} = 0 \tag{18} \]

Such a scheme is rewritten as a RD scheme with the residual
\[ \Phi^T = \sum_{\sigma \subset T} \varphi^p_e \left( \hat{\mathbf{i}} \cdot \nabla u^h - f \right) \, dx + h_e k_e(\varphi^p_e, u^h) - h_e k_e(\varphi^p_e) \tag{19} \]
the conservation constraint is automatically satisfied because \( \sum_{\sigma \subset T} \varphi^p_e = 1 \) and using the linearity with respect to \( v^h \),
\[ \sum_{\sigma \subset T} \Phi^p_e = \int_T \left( \hat{\mathbf{i}} \cdot \nabla u^h - f \right) \, dx + h_e \left( a_e(u^h, \sum_{\sigma \subset T} \varphi^p_e) - k_e \left( \sum_{\sigma \subset T} \varphi^p_e \right) \right) \]
\[ = \int_T \left( \hat{\mathbf{i}} \cdot \nabla u^h - f \right) \, dx = \Phi_k \]

Formally, the scheme is accurate because if that if \( w^h \in V^0_h \) is the interpolant of the exact solution of (1) assumed to be smooth, then \( \Phi^T(w^h) = 0 \). In fact, for RD scheme, this property is true because \( \Phi^T(w^h) = 0 \) since
\[ \Phi^T(w^h) = \int_T \hat{\mathbf{i}} \cdot \nabla w^h \, dx - \int_T f(x) \, dx = \int_T \hat{\mathbf{i}} \cdot \nabla w^h(x) - u(x) \, dx \]
\[ = O(h^{d-1}) \times O(h^{p+1}) = O(h^{d-p-1}). \]

Using this remark, the relation (9) follows for regular meshes. A first scheme of the type (7) can be obtained by perturbing the residual (19) as
\[ (\Phi^T)^\sigma = (\Phi^p)^\sigma + \psi^\sigma \]
with the constraint \( \sum_{\sigma \subset T} \psi^\sigma = 0 \) to ensure conservation. The formal accuracy property is also conserved if \( \psi^\sigma(w^h) = 0 \) whenever \( \hat{\mathbf{i}} \cdot \nabla w^h = 0 \).

A first example is obviously given by
\[ \psi^\sigma = \vartheta \int_T (\hat{\mathbf{i}} \cdot \nabla \varphi^p_e)(\hat{\mathbf{i}} \cdot \nabla u^h - f) \, dx \]
\[ \text{where} \ \vartheta \ \text{is chosen such that} \]
\[ \int_T (\hat{\mathbf{i}} \cdot \nabla u)(\hat{\mathbf{i}} \cdot \nabla u) \, dx + \theta \int_T (\hat{\mathbf{i}} \cdot \nabla u)^2 \, dx \geq 0. \tag{23} \]
Under this condition, the iterative scheme (13) is convergent when \( n \to +\infty \). In (22), \( \psi^\sigma \) is evaluated by a quadrature formula of exact order,\n\[ \int_T (\hat{\mathbf{i}} \cdot \nabla \varphi^p_e)(\hat{\mathbf{i}} \cdot \nabla u^h) \, dx = \sum_{\sigma \subset T} \omega_{\sigma \subset T} (\hat{\mathbf{i}} \cdot \nabla \varphi^p_e)(\hat{\mathbf{i}} \cdot \nabla u^h) \]
\[ \times (\hat{\mathbf{i}} \cdot \nabla u^h(x_{\sigma \subset T} \quad \text{and} \\( f(x_{\sigma \subset T}) \)) \tag{24} \]

The question is now : given a formula of the type
\[ \psi^\sigma = \vartheta \int_T \left( \sum_{\sigma \subset T} \omega_{\sigma \subset T} (\hat{\mathbf{i}} \cdot \nabla \varphi^p_e)(\hat{\mathbf{i}} \cdot \nabla u^h(x_{\sigma \subset T} \quad \text{and} \\( f(x_{\sigma \subset T}) \)) \right) \tag{25} \]
what are the requirements on the points $x_{\text{quad}}$ and the weights $\omega_{\text{quad}}$, so that we still have the inequality (23) and the accuracy condition (21)?

**Accuracy constraint.** Assuming that the solution of (1) is smooth enough, we have on $T$

$$w^h - u = O(h^{p-1})$$

and for a regular mesh

$$\nabla \varphi_a = O(h^{-1})$$

so that for any $x_{\text{quad}}$,

$$\left( \tilde{h} \cdot \nabla \varphi_a^s(x_{\text{quad}}) \right) \left( \tilde{h} \cdot \nabla \varphi_h(x_{\text{quad}}) - f(x_{\text{quad}}) \right) = O(h^{p-1})$$

so that

$$\psi_T^h = h \times O(h^4) = O(h^{p+4})$$

The formal accuracy is automatically guaranteed.

**Constraints on the weights and the points $x_{\text{quad}}$:** In order to have the inequality (23), a necessary condition is that the quadratic form

$$q_k(v_h) := \sum_{x_{\text{quad}}} \omega_{x_{\text{quad}}} \left( \tilde{h} \cdot \nabla \varphi_h(x_{\text{quad}}) \right)^2$$

must be positive definite whenever the polynomial $\tilde{h} \cdot \nabla v_h \neq 0$.

A sufficient condition is

for all $x_{\text{quad}}, \omega_{\text{quad}} > 0$

$$\tilde{h} \cdot \nabla \varphi_h(x_{\text{quad}}) = 0 \text{ for each } x_{\text{quad}}, \text{ then } \tilde{h} \cdot \nabla \varphi_h = 0.$$

(26)

Under these conditions, there exist constants $C_{1,q}$ and $C_{2,q}$ such that

$$C_{1,q} q_k(v_h) \leq h_k \int (\tilde{h} \cdot \nabla \varphi_h)^2 dx \leq C_{2,q} q_k(v_h)$$

(27)

because $P^d$ is a finite dimensional space, hence

$$Q(v_h) = \sum_K q_k(v_h)$$

defines a norm on $V_h$ which is equivalent to the norm $V_h \mapsto \int (\tilde{h} \cdot \nabla v_h)^2 dx$.

We have shown the following result

**Proposition 3.2.** If $a$ and $\ell$ are defined by (17b), under the Assumption 3.1 and provided that the conditions (26) hold, for each element $K$, there exists $\theta_{q,k} > 0$ such that the scheme (17a) for $\theta_K > \theta_{q,k}$ is well posed.

**Proof.** The scheme writes in variational formulation: find $u_h \in V_h$ such that for all $v_h \in V_h$,

$$a'(t_h, v_h) = \ell(v_h)$$

with

$$a' = a + \beta, \quad \ell' = \ell$$

where

$$b(u_h, v_h) = \sum_{K \subset \mathcal{T}_h} \sum_{\sigma \in K} v_{\sigma} \psi_{\sigma}^b$$

with $\psi_{\sigma}^b$ defined by (25) and using the scalar product

$$\langle u_h, v_h \rangle = \sum_{K \subset \mathcal{T}_h} |K| \left( \sum_{\sigma \in K} u_{\sigma} v_{\sigma} \right).$$

The iterative scheme (13) writes,

$$u_h^{n+1} = u^n - \omega(Au^n - F)$$

with

$$\langle Au^h, v_h \rangle = a(u^h, v_h), \quad \langle F, v_h \rangle = \ell(v_h).$$

The scheme is convergent if

$$\|Id - \omega A\| < 1.$$ 

A necessary condition is that for any $v_h$,

$$-2\|Av^h, v_h\| + \omega\|Av^h\| \leq 0$$

for some $\omega > 0$. This condition needs

$$a(v_h, v_h) = \langle Av^h, v_h \rangle > 0$$

(28)

for any $v_h$. Coming back to the problem,

$$a(v_h, v_h) = \int_{\Omega} \tilde{h} \cdot \nabla v_h^2 dx + \sum_K h_k(a_k(v_h, v_h) + \theta q_k(v_h))$$

hence a necessary condition for having (28) is that on any $K$ we have

$$a_k(v_h, v_h) + \theta q_k(v_h) > 0.$$

From Assumption 3.1, we see that

$$\text{Ker} a_k = \{ v_h \in P_K, a_k(v_h) = 0 \} \subseteq \text{Ker} \theta K = \{ v_h \in P_K, a_k(v_h) = 0 \},$$

so that which means that, since $q_k$ is positive definite, the scheme is contractant provided that

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Accuracy results for (29)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 1/N$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>Second order accurate results</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>$0.50493 \times 10^{-2}$</td>
</tr>
<tr>
<td>50</td>
<td>$0.14684 \times 10^{-2}$</td>
</tr>
<tr>
<td>75</td>
<td>$0.74684 \times 10^{-3}$</td>
</tr>
<tr>
<td>100</td>
<td>$0.41019 \times 10^{-3}$</td>
</tr>
<tr>
<td>Third order accurate results</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>$0.32612 \times 10^{-4}$</td>
</tr>
<tr>
<td>50</td>
<td>$0.48741 \times 10^{-5}$</td>
</tr>
<tr>
<td>75</td>
<td>$0.13334 \times 10^{-5}$</td>
</tr>
<tr>
<td>100</td>
<td>$0.66019 \times 10^{-6}$</td>
</tr>
<tr>
<td>Fourth order accurate results</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>$0.20860 \times 10^{-5}$</td>
</tr>
<tr>
<td>50</td>
<td>$0.17001 \times 10^{-5}$</td>
</tr>
<tr>
<td>75</td>
<td>$0.27027 \times 10^{-6}$</td>
</tr>
<tr>
<td>100</td>
<td>$0.91462 \times 10^{-7}$</td>
</tr>
</tbody>
</table>

The third order accurate scheme uses the vertices in (25). The fourth order scheme uses the vertices and the mid-points (25). In each case, $\theta = 1$.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Accuracy results for (29)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 1/N$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>25</td>
<td>$0.25122 \times 10^{-3}$</td>
</tr>
<tr>
<td>50</td>
<td>$0.12935 \times 10^{-3}$</td>
</tr>
<tr>
<td>100</td>
<td>$0.83978 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

The “third” order accurate scheme uses the gravity center in (25). In each case, $\theta = 1$.

<table>
<thead>
<tr>
<th>Table 4</th>
<th>Accuracy results for (29)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$h = 1/N$</td>
<td>$L^2$</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
</tr>
<tr>
<td>25</td>
<td>$2.17274 \times 10^{-2}$</td>
</tr>
<tr>
<td>50</td>
<td>$1.13486 \times 10^{-2} $</td>
</tr>
<tr>
<td>100</td>
<td>$5.83347 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

The “fourth” order accurate scheme uses the vertices in (25). In each case, $\theta = 1$. 
\[ \theta_H > \theta_{K,0} = \min \left( 0, -\frac{\sup_{v \in I^K} a(v^h, \psi^h)}{\inf_{v \in I^K, \psi \in \ker q_K} q_K \left( \psi^h \right)} \right) \in \mathbb{R} \]

for each \( K \).

4. Examples and numerical illustrations

4.1. Accuracy study

We apply the method on a simple linear problem, namely

\[ u, \min = -0.5, \max = 1.5 \]

\[ u, \min = -0.5255, \max = 1.526 \]

\[ u, \min = -0.5538, \max = 1.549 \]

Fig. 2. Isolines of the solution of (29) when 1 point or 3 points are used in (25). The baseline scheme is formally third order. All the degrees of freedom are represented.

\[ \theta = 0 \]
\[ \theta = 1, 4 \text{ quadrature points} \]

\[ \theta \text{ defined by (31), 4 quad points} \]
\[ \theta \text{ defined by (31), 3 quad points} \]

Fig. 3. Results obtained for problem (30) with various choices of \( h \) and quadrature points. In the case of 4 quadrature points we have chosen the centroid (weight \( -27/48 \)), and the points of coordinates \((0.6,0.2,0.2),(0.2,0.6,0.2),(0.2,0.2,0.6)\) with weights \( 25/48 \). In the case of 3 quadrature points, they are simply the vertices of the triangle with the weights \( 1/3 \).
\[ \frac{\partial u}{\partial y} = 0 \quad (x, y) \in [0, 1]^2 \]
\[ u(x, y) = \sin(\pi y)^2 \quad x = 0 \]

for which the solution is simply \( u(x, y) = \sin(\pi \sqrt{x^2 + y^2})^2 \).

We have run the formally second order scheme, third order and fourth order schemes with respectively 1, 3, 6 “quadrature points” in (25).

For the third order scheme, the “quadrature” points are simply the vertices of the triangle. For the third order case, we have chosen the vertices and the mid-point edges. The weights are 1, 1/3 and 1/6 respectively. The results, in term of accuracy, are independent of choices of the “quadrature” points, provided that (24) defines a positive definite quadrature form. The constant \( \theta \) in (25) is set to unity. The results are displayed in Table 2. We see that the expected order of accuracy is met in each case. If now we repeat the same experiment with a smaller number of “quadrature” points, the accuracy is degraded and the results are only first order accurate or the scheme is only consistent, see Tables 3 and 4. This can also be seen visually on Fig. 2.

![Figure 4](image-url)  
**Fig. 4.** Cross-section of the solution at \( y = 0.25 \) and \( y = 0.75 \). For \( y = 0.25 \), the solution corresponding to 3 and 4 quadrature points and \( \theta \neq 0 \) are undistinguishable. The curve labelled \( \Theta(Res) \) corresponds to the choice (31). Some difference appears for the cross-section at \( y = 0.75 \). The choice (31) appears to be a good compromise.

![Figure 5](image-url)  
**Fig. 5.** Supersonic jet, third order solution.
4.2. The non-linear case

The second example is the solution of the problem

\[
\frac{1}{2} \frac{\partial u^2}{\partial x} + u \frac{\partial u}{\partial x} = 0 \text{ if } x \in [0, 1]^2
\]

\[u(x, y) = 1.5x - 0.5 \text{ when } y = 0 \text{ or } x \in \{0, 1\}\]

The solution consists in a compression merging into a shock which foot is located at (0.5, 0.75). Several schemes are tested. We only represent the solutions obtained by the formally third order scheme since the behavior for the fourth order one is the same. The "quadrature" points are again the vertices of the elements with and without the centroid depending on if we take 3 or 4 points.

On Fig. 3, we represent the isolines of the scheme when \(\theta\) is set to 0, 1 or

\[
\theta = \frac{1}{\sum_{\text{vertices}} |k_i|} \min \left( 1, \frac{\sqrt{\sum_{\text{vertices}} |k_i|}}{\int_{\Gamma_T} (n_y \frac{\partial u}{\partial x} + n_x u) dl + \epsilon} \right).
\]

In (31), if \(n_x^i\) and \(n_y^i\) are the components of the inward normal opposite the vertex \(i\) in the triangle, \(k_i = n_x^i \frac{\partial u}{\partial x} + n_y^i u\) and \(\epsilon = 10^{-10}\). Once again, the same conclusions hold: 3 points are necessary to get accurate results. We compare the solutions depending on which option is chose (3/4 quad points, the choice of \(\theta\)). To do this, we make

**Fig. 6.** With and without dissipation, density isolines.

**Fig. 7.** Comparison between 1st, second and third order, same degrees of freedom.
cross-section at \( y = 0.25 \), i.e. in the fan, and \( y = 0.75 \), i.e. in the discontinuity.

The results of Fig. 4 show that if \( h = 0 \), the oscillations visible in Fig. 3 are not a manifestation of an instability, the scheme is over-compressive. When \( h = 1 \) or is chosen as (31), there is no difference in the solution, whatever the number of quadrature points. In Fig. 4, we plot the result at \( y = 0.75 \). If we add the addition of the term (25), the scheme is no longer formally monotonicity preserving, but the Fig. 4 indicate that no undershoot nor overshoot are created. The same figure also indicate that the choice (31) is the best compromise between accuracy and stability. The effect of this term is that when the solution is smooth, \( h \approx 1 \) while \( h \approx 0 \) in the discontinuity.

4.3. Euler equations

The last examples that we show are for the Euler equations. Details about the scheme can be found in [2] in particular about the way Eq. (12) is implemented. The method has been implemented only for \( P_2 \) element so far again with 3 “quadrature” points. The first example is a supersonic jet with \( M = 2.4 \) on the bottom and \( M = 4.4 \) on the top. The solution, see Fig. 5, is made of a shock wave followed by a contact and a fan. On Fig. 6, we show the effect of adding and removing the term (25). We can also see the increase of accuracy. On Fig. 7, we have run the first order, second order and third order RD schemes with the same number of degrees of freedom, namely the vertices and the mid-points of the mesh. A last example is a 4 state shock tube problem (configuration 12 of [7]). This case is time dependant, but we can compute the solution at time \( t \) since the solution is self-similar, \( U(x, y, t) = V(\tfrac{\zeta}{t}) \). The function \( V(\zeta, \nu) \) satisfies

\[
-\zeta V_\zeta - \nu V_\nu + \text{div}(\zeta, \nu) F(V) = 0.
\]

The case has been chosen that the boundary condition can easily be computed analytically. The scheme is the same as before, but we modify the definition of the total residual by

\[
\Phi^T := \int_{\Gamma} (-\zeta V_\zeta - \nu V_\nu + \text{div}(\zeta, \nu) F(V)) d\zeta d\nu.
\]

This integral is evaluated by

\[
\Phi^T = \int_{\Gamma} (F(V) \cdot \tilde{n} - (\zeta, \nu) \cdot \tilde{n}) dl + \int_{\Gamma} V(\zeta, \nu) d\zeta d\nu.
\]

Again we see the improvement obtained by adding the term (25). The scheme is very robust and non-oscillatory, despite the interaction between many waves.

5. Conclusion

In this paper, we have discussed a simple way to construct simple and accurate very high order residual distribution schemes. A theoretical discussion is provided which is confirmed by numerical experiments on scalar problems and the Euler equations. We have focussed on schemes like Residual Distribution schemes, but we believe however that the method we present in this paper can be adapted to other type of schemes. Note also that it shares common features with the work of Corre and Lerat, see [8,9,3] for example (see Figs. 8 and 9).
Acknowledgements

The work of A. Larat is supported by the European Project ADIG-MA on the development of innovative solution algorithms for aerodynamic simulations. The work of C. Tavé was supported by CNES.

The first author would like to take this opportunity to thank Professor Alain Lerat for the major impact he had on his scientific life. I was lucky to have him as a professor during my master studies, and he made me discover that computing could be fascinating. After this, there was no possible return to pure mathematics, my original background.

References


