Analysis of high-order residual-based dissipation for unsteady compressible flows

Karim Grimich^{*}, Paola Cinnella^{*} and Alain Lerat^{*} Corresponding author: karim.grimich@ensam.eu

* Arts et Metiers-ParisTech, DynFluid Laboratory, 151 bd. de l'Hopital, 75013 Paris, France.

Abstract: A comprehensive study of the numerical properties of high-order residual-based dissipation terms for unsteady compressible flows leads to the design of well-behaved, low dissipative schemes of third-, fifth- and seventh-order accuracy. The dissipation and dispersion properties of the schemes are then evaluated theoreticaly, through Fourier space analysis, and numerically, through selected test cases including the inviscid Taylor-Green Vortex flow.

Keywords: Residual-based scheme, High-order, Unsteady, Dissipation.

1 Introduction

Classical methods for calculating compressible flows on a structured mesh rely on a directional approach in which space derivatives are approximated independently direction by direction. In the present work, we study compact approximations that provide a high accuracy not for each space derivatives treated apart but for the complete residual r, *i.e.* the sum of all the terms in the governing equations. For unsteady problems, r also includes the time derivative. Schemes of this type are said to be Residual-Based Compact (RBC) and have been developed in the last ten years ([1, 2] for instance) with application to realistic flow configurations in aerodynamics and aeroacoustics. A special feature of RBC schemes is the structure of their numerical dissipation also constructed from the complete residual r. This dissipation exhibits interesting properties that have not been fully analysed for unsteady problems so far. Here, we present a detailed study of the residual-based dissipation involved in high-order RBC schemes for the unsteady Euler equations. This study allows a better understanding of the dissipation mechanism and leads to some improvements of the existing RBC schemes. Given the importance of numerical dissipation in Computational Fluid Dynamics, it is also hoped that the present work could help the development of other classes of high-order schemes.

The paper is organised as follows. In Section 2 we briefly recall the general design principle of RBC spatial discretization for convective problems, then we focus on selected RBC schemes of third-, fifth- and seventh-order of accuracy. In section 3 we identify the effective dissipation operator, which is not obvious for a RBC scheme since the dissipation operator and derive a criterion ensuring dissipation of RBC schemes. Then in section 4 we derive the spectral counterparts of the RBC schemes under investigation and discuss their dissipation and dispersion properties. Finally, Section 5 presents some numerical experiments supporting the preceding theoretical analysis.

2 High-order RBC schemes

In this Section, we recall the design principles of RBC approximations of the space derivatives for a hyperbolic system of conservation laws. For brevity and clarity, we will focus on two-dimensional problems even if there is no restriction for extending the analyses below to multi-dimensional hyperbolic problems. At this stage, we treat time derivatives exactly, i.e. we focus on semi-discrete approximations in space.

2.1 Concept of residual-based scheme

Let us consider an initial-value problem for the hyperbolic system of conservation laws:

$$w_t + f_x + g_y = 0 \text{ on } \mathbb{R}^2 \times \mathbb{R}^+ \tag{1}$$

with initial conditions

$$w(x, y, 0) = w_0(x, y)$$

where t is the time, x and y are Cartesian space coordinates, w is the state vector and f = f(w), g = g(w)are flux components depending smoothly on w. The Jacobian matrices of the flux are denoted A = df/dwand B = dg/dw. System (1) is approximated in space on a uniform mesh $(x_j = j\delta x, y_k = k\delta y)$ with steps δx and δy of the same order of magnitude, say $\mathcal{O}(h)$, using the basic difference and average operators:

$$(\delta_1 v)_{j+\frac{1}{2},k} = v_{j+1,k} - v_{j,k} \qquad (\delta_2 v)_{j,k+\frac{1}{2}} = v_{j,k+1} - v_{j,k}$$
$$(\mu_1 v)_{j+\frac{1}{2},k} = \frac{1}{2}(v_{j+1,k} + v_{j,k}) \qquad (\mu_2 v)_{j,k+\frac{1}{2}} = \frac{1}{2}(v_{j,k+1} + v_{j,k})$$

where j and k are integers or half integers.

A residual-based scheme can be expressed in terms of approximations of the exact residual:

$$r := w_t + f_x + g_y \tag{2}$$

More precisely, such a scheme is of the following form:

$$(\tilde{r}_0)_{j,k} = \tilde{d}_{j,k} \tag{3}$$

where \tilde{r}_0 is a space-centered approximation of r called the *main residual* and \tilde{d} is a residual-based dissipation term defined in terms of first-order differences of the residual as:

$$\tilde{d}_{j,k} = \frac{1}{2} [\delta_1(\Phi_1 \tilde{r_1}) + \delta_2(\Phi_2 \tilde{r_2})]_{j,k}$$
(4)

where $\tilde{r_1}$ and $\tilde{r_2}$, respectively defined at $j + \frac{1}{2}$, k and $j, k + \frac{1}{2}$, are also space-centered approximations of r called the *mid-point residuals*, and Φ_1 , Φ_2 are numerical viscosity matrices (defined at the same location as the corresponding mid-point residuals). These matrices depend only on the eigensystems of the Jacobian matrices A and B and on the steps δx and δy . They are designed once for all [3] and use no tuning parameters nor limiters. Since the matrices Φ_1 and Φ_2 remain $\mathcal{O}(1)$ as δx and δy tend to zero, the dissipation \tilde{d} represents, to the leading order, a numerical approximation of the second-order partial differential term:

$$d = \frac{\delta x}{2} (\Phi_1 r)_x + \frac{\delta y}{2} (\Phi_2 r)_y \tag{5}$$

This leading term of the expansion, that is only first order accurate, vanishes for an exact solution (r = 0), so that \tilde{d} is actually consistent with a high-order dissipation term that will be discussed later. In the following, the residuals are discretized in a compact way so that:

 $\begin{aligned} & (\tilde{r_0})_{j,k} &= r_{j,k} + \mathcal{O}(h^{2p}) \\ & (\tilde{r_1})_{j+\frac{1}{2},k} = r_{j+\frac{1}{2},k} + \mathcal{O}(h^{2p-2}) \\ & (\tilde{r_2})_{j,k+\frac{1}{2}} = r_{j,k+\frac{1}{2}} + \mathcal{O}(h^{2p-2}). \end{aligned}$

with $p \ge 2$. Thus, the dissipation term (4) satisfies:

$$\tilde{d}_{j,k} = d_{j,k} + \mathcal{O}(h^{2p-1})$$

and the truncation error of the semi-discrete scheme (3) is

$$\varepsilon_{j,k} = r_{j,k} + \mathcal{O}(h^{2p}) - d_{j,k} + \mathcal{O}(h^{2p-1}).$$

Since the exact residual r and the leading term d of the residual-based dissipation (5) are null for an exact unsteady solution, we finally obtain:

$$\varepsilon_{j,k} = \mathcal{O}(h^{2p-1}). \tag{6}$$

The preceding truncation error analysis shows that approximating the main residual at order 2p and the mid-point residuals at order 2p - 2 leads to a Residual-Based Compact scheme of order q = 2p - 1. Such a scheme is called RBCq. In the following, we focus on schemes using 5×5 -point stencils at most, which corresponds to p = 2, 3 and 4. More precisely, RBC3 schemes can be constructed with 3×3 points only and RBC5 and RBC7 schemes with 5×5 points.

2.2 Approximation of the main residual

The main residual \tilde{r}_0 is approximated through a difference operator of the form:

$$(\tilde{r}_0)_{j,k} = \left(\overline{D_1}\ \overline{D_2}w_t + \overline{D_2}\ \overline{N_1}\frac{\delta_1\mu_1f}{\delta x} + \overline{D_1}\ \overline{N_2}\frac{\delta_2\mu_2g}{\delta y}\right)_{j,k} \tag{7}$$

where $\overline{D_m}$ and $\overline{N_m}$ (m = 1, 2) are formal polynomials of the second difference operator in the m^{th} direction:

$$\overline{N_m} = I + \overline{a}\delta_m^2, \quad \overline{D_m} = I + \overline{b}\delta_m^2 + \overline{c}\delta_m^4, \quad \overline{a}, \overline{b}, \overline{c} \in \mathbb{R},$$
(8)

where I is the identity operator and

$$(\delta_m^p f)_{j,k} = \underbrace{\delta_m(\delta_m(\dots(\delta_m f)))}_{p \text{ times}} \tag{9}$$

The degrees of the polynomials are chosen in such a way that the scheme stencil is limited to 5×5 space points at most. Operator (7) is obtained by replacing space derivatives in each direction by Pade operators:

$$f_x = (\overline{D_1})^{-1} \overline{N_1} \frac{\delta_1 \mu_1 f}{\delta x} + \mathcal{O}(\delta x^{2p})$$

$$g_y = (\overline{D_2})^{-1} \overline{N_2} \frac{\delta_2 \mu_2 g}{\delta y} + \mathcal{O}(\delta y^{2p})$$
(10)

and subsequently applying the operator $\overline{D_1} \ \overline{D_2}$ to the whole left-hand side of the equation. The truncation error of (7) is

$$\varepsilon_{j,k} = \left[I + \mathcal{O}(h^2)\right] \left[r_{j,k} + \mathcal{O}(h^{2p})\right] \tag{11}$$

where the first term at the right-hand side vanishes for an exact unsteady solution. This represents a substantial difference of RBC schemes with respect to standard Pade approximations and avoid the inversion of linear systems per each space direction if a suitable time-integration technique is selected [1, 4, 5].

An approximation of the main residual of order 2p = 4 on a 3×3 stencil, order 2p = 6 or 2p = 8 on a 5×5 -point stencil can be obtained using the set of coefficients given in [3, 2]. Note that, because we use purely centered operators, no damping effects are introduced at this stage. Thus, the dissipation properties of RBC schemes are actually governed by the right-hand side operator \tilde{d} of Eq.(3).

2.3 Residual-based dissipation operator

As anticipated at the beginning of this Section, the dissipation operator $d_{j,k}$ is given by Eq.(4), and involves mid-point residual approximations. Precisely, the following difference operators are used:

$$(\tilde{r_{1}})_{j+\frac{1}{2},k} = \left[N_{1}^{\mu} \mu_{1} \left(D_{2} w_{t} + N_{2} \frac{\delta_{2} \mu_{2} g}{\delta y} \right) + N_{1}^{\delta} D_{2} \frac{\delta_{1} f}{\delta x} \right]_{j+\frac{1}{2},k}$$

$$(\tilde{r_{2}})_{j,k+\frac{1}{2}} = \left[N_{2}^{\mu} \mu_{2} \left(D_{1} w_{t} + N_{1} \frac{\delta_{1} \mu_{1} f}{\delta x} \right) + N_{2}^{\delta} D_{1} \frac{\delta_{2} g}{\delta y} \right]_{j,k+\frac{1}{2}}$$

$$(12)$$

based again on the use of formal polynomials of the difference operators:

$$N_m^{\delta} = I + a^{\delta} \delta_m^2 , \quad N_m^{\mu} = I + a^{\mu} \delta_m^2 , \quad N_m = I + a \delta_m^2 , \quad D_m = I + b \delta_m^2 + c \delta_m^4$$

with $m = 1, 2$, and $a^{\delta}, a^{\mu}, a, b, c \in \mathbb{R}$ (13)

The dissipation matrices Φ_1 , Φ_2 have been designed in [1, 3] in order to introduce some form of upwinding of the numerical scheme with respect to the local advection direction. For a linear scalar problem with advection velocities A and B in the x and y direction respectively, they satisfy the conditions

$$\Phi_1 A \ge 0, \quad \Phi_2 B \ge 0
\delta x \Phi_1 B = \delta y \Phi_2 A$$
(14)

with

$$\Phi_1 = \operatorname{sgn}(A)\Phi, \quad \Phi_2 = \operatorname{sgn}(B)\Psi,$$

$$\Phi = \min\left(1, \frac{1}{\alpha}\right), \quad \Psi = \alpha\Phi \quad \text{with} \quad \alpha = \frac{\delta x|B|}{\delta y|A|}.$$
 (15)

The interested reader is referred to [1, 3, 6] for more details about the dissipation matrices Φ_1 , Φ_2 and their extension to systems of conservation laws.

3 The χ -criterion for dissipation

During the time evolution, d is null and \tilde{d} is consistent with some high-order term. Thus the numerical dissipation \tilde{d} should not be viewed as consistent with the operator d, but with another operator induced by the discretization of d. In the following, we determine the true differential operator consistent with \tilde{d} and find the conditions under which this operator is dissipative or not, since this feature affects the stability of the scheme.

For the sake of simplicity, we continue to name \tilde{d} the dissipation term before establishing the conditions justifying this designation.

For a better understanding of the role of the different contributions in the general dissipation (4), we proceed in two stages. First, we restrict our attention to the role of the x-discretization in $\tilde{r_1}$ and of the y-discretization in $\tilde{r_2}$.

Thus, we define:

$$(\tilde{r_1}^x)_{j+\frac{1}{2},k} = \left[N_1^{\mu} \mu_1(w_t + g_y) + N_1^{\delta} \frac{\delta_1 f}{\delta x} \right]_{j+\frac{1}{2},k}$$
(16)

This residual depends only on two constant coefficients a^{μ} and a^{δ} and involves four points at most (i.e. j-1, j, j+1 and j+2). Similarly, we set:

$$(\tilde{r_2}^y)_{j,k+\frac{1}{2}} = \left[N_2^{\mu} \mu_2(w_t + f_x) + N_2^{\delta} \frac{\delta_2 g}{\delta y} \right]_{j,k+\frac{1}{2}}$$
(17)

We now carry out a Taylor expansion of $\tilde{r_1}^x$ around $(j + \frac{1}{2})\delta x$. Provided the exact residual is sufficiently smooth and null everywhere for an exact unsteady solution, $\tilde{r_1}^x$ at $(j + \frac{1}{2}, k)$ reduces to:

$$\tilde{r_1}^x = \delta x^2 \left(a^\delta - a^\mu - \frac{1}{12} \right) f_{xxx} + \frac{\delta x^4}{24} \left(3a^\delta - 5a^\mu - \frac{1}{20} \right) f_{5x} + \frac{\delta x^6}{5760} \left(39a^\delta - 91a^\mu - \frac{3}{28} \right) f_{7x} + \mathcal{O}(\delta x^8)$$
(18)

A similar result is obtained for $\tilde{r_2}^y$. Inserting the semi-discrete residuals (16) and (17) in the dissipation (4) gives the partial dissipation term:

$$\tilde{d}_{j,k}^{x,y} = \frac{1}{2} [\delta_1(\Phi_1 \tilde{r_1}^x) + \delta_2(\Phi_2 \tilde{r_2}^y)]_{j,k}$$
(19)

We now consider the complete space discretization. For the mid-point residuals, this means that f_x in \tilde{r}_2 and g_y in \tilde{r}_1 are discretized through Pade approximations. By applying the operator D_2 (resp. D_1) to all the terms of \tilde{r}_1 (resp. \tilde{r}_2) we complete the approximations of the mid-point residuals as in (12). A Taylor expansion of the complete mid-point residual \tilde{r}_1 can be easily obtained by noting that

$$N_2 \frac{\delta_2 \mu_2 g}{\delta y} = D_2 (g_y + \varepsilon_2)$$

so that $\tilde{r_1}$ can be related to $\tilde{r_1}^x$ as

$$\tilde{r_1} = D_2(\tilde{r_1}^x + N_1^\mu \mu_1 \varepsilon_2)$$

Since N_1^{μ} and D_2 are consistent with the identity plus second order terms, whereas $\tilde{r_1}^x$ and ε_2 are $\mathcal{O}(h^{2p-2})$ for $p \ge 2$, we get

$$\tilde{r_1} = \tilde{r_1}^x + \mu_1 \varepsilon_2 + \mathcal{O}(h^{2p}).$$
⁽²⁰⁾

Similarly, we get

$$\tilde{r_2} = \tilde{r_2}^y + \mu_2 \varepsilon_1 + \mathcal{O}(h^{2p}).$$
⁽²¹⁾

These simple relations are applied below to obtain the full dissipation term \tilde{d} for RBC3, RBC5 and RBC7. The effective dissipation term \tilde{d} induced by the discretization of the simple operator

$$d = \frac{\delta x}{2} [\Phi_1(w_t + f_x + g_y)]_x + \frac{\delta y}{2} [\Phi_2(w_t + f_x + g_y)]_y$$

in RBC3, RBC5 and RBC7 schemes has been identified (see [7] for details) and all these expressions can be cast in a general form. For any integer $p \ge 2$, let us denote q = 2p - 1 and

$$f_{qx} = \frac{\partial^q f}{\partial x^q}, \quad g_{qy} = \frac{\partial^q g}{\partial y^q}$$

The dissipation of a RBCq scheme can be expressed as:

$$\tilde{d} = d_q + \mathcal{O}(h^{q+2}) \tag{22}$$

with

$$d_q = (-1)^{p-1} \kappa \left\{ \delta x [\Phi_1(\delta x^{q-1} f_{qx} + \chi \delta y^{q-1} g_{qy})]_x + \delta y [\Phi_2(\delta y^{q-1} g_{qy} + \chi \delta x^{q-1} f_{qx})]_y \right\}$$
(23)

where $\kappa > 0$ and χ are two constant coefficients.

We need to determine whether the above multidimentional operator, involving cross derivatives, is dissipative or not. Let us first introduce a definition of dissipation. Consider linear fluxes

$$f = Aw, \quad g = Bw$$

where A and B are scalar constants. The expression (23) reduces to

$$d_q = \mathscr{D}_q w$$

with the linear partial differential operator:

$$\begin{aligned} \mathscr{D}_{q} &= (-1)^{p-1} \kappa \left(\, \delta x^{2p-1} \Phi_{1} A \frac{\partial^{2p}}{\partial x^{2p}} + \chi \delta x \delta y^{2p-2} \Phi_{1} B \frac{\partial^{2p}}{\partial x \partial y^{2p-1}} \right. \\ &+ \left. \delta y^{2p-1} \Phi_{2} B \frac{\partial^{2p}}{\partial y^{2p}} + \chi \delta y \delta x^{2p-2} \Phi_{2} A \frac{\partial^{2p}}{\partial y \partial x^{2p-1}} \right) \end{aligned}$$

Note that this linear version contains the highest derivatives in (23). All the derivatives in \mathscr{D}_q being even, its Fourier symbol is real. It is denoted by $\hat{\mathscr{D}}_q(\xi, \eta)$, where ξ and η are the wave numbers (Fourier variables).

Definition 3.1. The operator (23) is said to be dissipative (in the broad sense) if:

$$\forall \xi \in \mathbb{R}, \quad \forall \eta \in \mathbb{R}, \qquad \hat{\mathscr{D}}_q(\xi, \eta) \leqslant 0.$$

Theorem 3.1. (χ -criterion) The operator (23) is dissipative for any order q = 2p - 1 ($p \ge 2$), any advection direction (A, B) and any functions Φ_1 , Φ_2 satisfying the conditions (14) if and only if $\chi = 0$. (see [7] for proof)

Taking into account the accuracy order and the conditions for the stencil to have a minimal extent (3×3 points for RBC3 and 5×5 points for RBC5 and RBC7), the χ -criterion leads to the following conclusions. Dissipation is ensured in any situation by a unique set of coefficients for RBC3 and RBC7, and a twoparameter family of coefficients for RBC5 (see [7] for details). Note that coefficients for RBC3 and RBC7 are different from those given in former works [3, 2], where their calculation was based on accuracy and minimal complexity considerations only. Those schemes were advanced in time via a linear backward multistep method coupled with an implicit dual time stepping technique and proved to be inconditionally stable in pseudo time. However, violation of the dissipation criterion of Theorem 3.1 leads to a slight instability with respect to physical time for some wave numbers. This is discussed in the following.

4 Spectral properties of RBC schemes

Given mesh functions w and w_t such that $w(x_j, y_k, t) = w_{j,k}(t)$ and $w_t(x_j, y_k, t) = \frac{\partial w}{\partial t}\Big|_{j,k}$, the RBCq spatial discretization scheme of order q = 3, 5, 7 can be expressed as:

$$\tilde{r_0}(w, w_t) = \tilde{d}(w, w_t) \tag{24}$$

where \tilde{r}_0 is the centered residual operator and \tilde{d} is the numerical dissipation operator. They can both be split in a part dependent on the state vector w only and another one that is a function of the time derivative w_t only. Precisely, calling:

$$\mathfrak{I}_{\tilde{r}_{0}} = \overline{D_{1}} \ \overline{D_{2}}, \qquad \mathfrak{I}_{\tilde{d}} = \frac{1}{2} \left(\delta_{1} \Phi_{1} N_{1}^{\mu} \mu_{1} D_{2} + \delta_{2} \Phi_{2} N_{2}^{\mu} \mu_{2} D_{1} \right)$$
(25)

the linear difference operators that apply to w_t and:

$$\Re_{\tilde{r}_{0}}(w) = \overline{D_{2}} \ \overline{N_{1}} \frac{\delta_{1}\mu_{1}f(w)}{\delta x} + \overline{D_{1}} \ \overline{N_{2}} \frac{\delta_{2}\mu_{2}g(w)}{\delta y},$$
$$\Re_{\tilde{d}}(w) = \frac{1}{2} \left[\delta_{1}\Phi_{1} \left(N_{1}^{\delta}D_{2} \frac{\delta_{1}f(w)}{\delta x} + N_{1}^{\mu}\mu_{1}N_{2} \frac{\delta_{2}\mu_{2}g(w)}{\delta y} \right) + \delta_{2}\Phi_{2} \left(N_{2}^{\delta}D_{1} \frac{\delta_{2}g(w)}{\delta y} + N_{2}^{\mu}\mu_{2}N_{1} \frac{\delta_{1}\mu_{1}f(w)}{\delta x} \right) \right]$$
(26)

the operators that apply to functions of w only, Eq. (24) re-writes:

$$(\mathfrak{I}_{\tilde{r}_0} - \mathfrak{I}_{\tilde{d}})w_t = -\mathfrak{R}_{\tilde{r}_0}(w) + \mathfrak{R}_{\tilde{d}}(w)$$

or, in a more compact way:

$$v_t = -\mathfrak{I}^{-1}\mathfrak{R}(w), \text{ with } \mathfrak{I} = \mathfrak{I}_{\tilde{r}_0} - \mathfrak{I}_{\tilde{d}} \text{ and } \mathfrak{R} = \mathfrak{R}_{\tilde{r}_0} - \mathfrak{R}_{\tilde{d}}$$
 (27)

or

$$w_t = \mathcal{S}(w), \text{ with } \mathcal{S} = -\mathfrak{I}^{-1}\mathfrak{R}$$
 (28)

Eq. (28) represents now a system of ordinary differential equations and, completed by initial conditions $w(.,0) = w_0(.)$, it defines a Cauchy problem. Thus, the stability properties of the semi-discrete scheme depend only on properties of the operator S.

To study these properties, we consider the linear problem

1

$$w_t + Aw_x + Bw_y = 0 \tag{29}$$

where A and B are scalar constants. The Fourier transform of Eq. (29) is:

$$\widehat{w}_t = -i(Ak_x + Bk_y)\widehat{w} = -i\mathbf{A} \cdot \mathbf{k}\widehat{w} \tag{30}$$

where $\mathbf{A} = (A, B)$ is the advection velocity vector and $\mathbf{k} = (k_x, k_y)$ a 2D wave vector. Denoting $\dot{\mathbf{A}} = (\dot{A}, \dot{B})$ a vector whose components are the CFL numbers in the x and y directions, respectively:

$$\dot{A} = A\Delta t/\delta x, \qquad \dot{B} = B\Delta t/\delta y$$

and introducing the reduced wave number aligned with $\mathbf{A}/|\mathbf{A}|$:

$$\xi_{ heta} = rac{\dot{\mathbf{A}}}{|\dot{\mathbf{A}}|} \cdot \boldsymbol{\xi}$$

with $\boldsymbol{\xi} = (\xi, \eta) = (k_x \delta x, k_y \delta y)$ the reduced wave vector, Eq. (30) can be rewritten as:

$$\widehat{w}_t = -i|\dot{\mathbf{A}}|\xi_\theta \frac{\widehat{w}}{\Delta t} \tag{31}$$

The notation ξ_{θ} refers to the local advection direction with respect to the mesh, where $\cos(\theta) = \dot{A}/|\dot{\mathbf{A}}|$ and $\sin(\theta) = \dot{B}/|\dot{\mathbf{A}}|$. The right-hand side of Eq. (31) represents the exact transport operator, a pure imaginary number, and ξ_{θ} is a pure real number. We now consider the semi-discrete counterpart to Eq. (31). Taking the Fourier transform of the semi-discrete system (28) applied to the linear problem (29), we obtain:

$$\widehat{w}_t = \widehat{\mathcal{S}}(\boldsymbol{\xi}, \dot{\mathbf{A}}) \frac{\widehat{w}}{\Delta t}$$
(32)

with $\widehat{\mathcal{S}}$ the Fourier symbol of the spatial operator. Introducing the modified wave number:

$$\xi_{\theta}^* = i \frac{\widehat{S}(\boldsymbol{\xi}, \dot{\mathbf{A}})}{|\dot{\mathbf{A}}|} \tag{33}$$

Eq. (32) can be rewritten as:

$$\widehat{w}_t = -i|\dot{\mathbf{A}}|\xi_\theta^* \frac{\widehat{w}}{\Delta t} \tag{34}$$

which represents the numerical counterpart of the exact operator (31). Eq. (33) represents the approximate dispersion relation of the semi-discrete scheme and relates a given reduced wave number ξ_{θ} to its numerical representation ξ_{θ}^* . Differently from the exact wave number, the modified wave number ξ_{θ}^* is a complex number, with a non-zero imaginary part in general. If the imaginary part is positive, then any Fourier mode set as an initial condition to the ordinary differential equation (28) will grow without bound.

As a consequence, a necessary condition for the Cauchy-stability of the semi-discrete system of equations is

 $\sup_{\xi_{\theta} \in [-\pi,\pi]} \Im(\xi_{\theta}^{*}) \leqslant 0, \, \forall \dot{\mathbf{A}} \in \mathbb{R}^{2}$ (35)

This comes to require that the Fourier symbol of the spatial discretization \hat{S} always lies in the left-hand side of the complex plane. Any spatial discretization satisfying the stability condition (35) leads to an unconditionally stable fully discrete scheme whenever it is combined to an A-stable time integration method. In addition to stability analysis, the modified wave number may be used as an indicator of how accurately a given wave number is represented by the difference operator. Specifically, we define the error with respect to the exact wave number:

$$\mathcal{E} = \xi_{\theta}^* - \xi_{\theta} = (\Re(\xi_{\theta}^*) - \xi_{\theta}) + i\Im(\xi_{\theta}^*)$$
(36)

More precisely, following previous work [8, 9, 10] for 1-D problems, we consider the multidimensional counterparts of the normalised phase error $\mathcal{P}_{\xi_{\theta}}$ and of the scheme damping function $\mathcal{D}_{\xi_{\theta}}$:

$$\mathcal{P}_{\xi_{\theta}} = \frac{|\Re e(\xi_{\theta}^*) - \xi_{\theta}|}{\pi}, \ \mathcal{D}_{\xi_{\theta}} = 1 - \exp[\Im m(\xi_{\theta}^*)]$$
(37)

For an infinitely accurate scheme, $\mathcal{P}_{\xi_{\theta}} = 0$ and $\mathcal{D}_{\xi_{\theta}} = 0$. In the following, we apply the preceding measures to the spectral operators of RBC schemes presented in Section 2. To this purpose, we first establish the approximate dispersion relation for an RBC scheme. We start by taking the Fourier transform of Eq. (27), which gives:

$$\widehat{w}_t = -\widehat{\Im}^{-1}\widehat{\Re}\widehat{w}, \text{ with } \widehat{\Im} = \widehat{\Im}(\boldsymbol{\xi}, \dot{\mathbf{A}}) \text{ and } \widehat{\Re} = \widehat{\Re}(\boldsymbol{\xi}, \dot{\mathbf{A}})$$

As a consequence:

$$\xi_{\theta}^{*} = -i\Delta t \frac{\widehat{\mathfrak{I}}^{-1}\widehat{\mathfrak{R}}}{|\dot{\mathbf{A}}|} \tag{38}$$

For the sake of brevity, we omit the mathematical expressions for $\hat{\mathfrak{I}}$ and $\hat{\mathfrak{R}}$. Given the complexity of the analytical expression of the modified wave number for an RBCq scheme, a computer code has been written to study its mathematical properties. First of all, we numerically check that RBCq schemes using coefficients deduced from Thm. 3.1 are Cauchy stable. To this purpose, the imaginary part of ξ^*_{θ} is computed for CFL numbers \dot{A} and \dot{B} ranging in [-2, 2] with a step $\Delta \dot{A} = \Delta \dot{B} = 1/100$ and the wave vector $\boldsymbol{\xi}$ ranging in $[-\pi, \pi] \times [0, \pi]$ with steps $\Delta \boldsymbol{\xi} = \Delta \eta = \pi/50$. Fig. 1 displays criterion (35) for several RBC schemes. Precisely, Fig. 1(a) to 1(c) show that RBC schemes with coefficients satisfying the dissipation criterion of Thm. 3.1 verify condition (35) for any CFL and any wave number, whereas the third- and seventh-order RBC schemes with coefficients given in [2] violate the stability criterion over a wide range of wave numbers. Then, we proceed to the analysis of the dispersion error and damping function for dissipative RBC schemes of different orders. Fig. 2 shows the dispersion error and damping functions in the case of advection aligned with one grid direction. Specifically, we choose $\dot{\mathbf{A}} = (\dot{A}, 0)$.

In this case, the error does not depend on η . Then, we consider a 1D cut of figure 2 for all schemes under investigation, and represent the errors on the same graph for comparison purposes. Precisely, Fig. 3(a) shows the phase errors (in log scale), for RBC schemes of different orders. Tab. 1 provides the wave number corresponding to a normalized phase error equal to 10^{-3} . RBCq schemes of fifth- and seventh-order accuracy exhibit a cut-off wave number, ξ_c , very close to $\pi/2$, the smallest resolvable wave number being close to $2\pi/5$. Similarly, Fig. 3(b) displays 1D cuts of the damping function $\mathcal{D}_{\xi_0^*}$. RBC5 and RBC7 exhibit a damping function of less than 10^{-3} up to cut-off wave numbers of 1.03 and 1.24, respectively. The conclusion of the preceding analysis is that RBC schemes of order 5 and 7 can accurately resolve a given wavelength by means of less than 5 mesh cells, $\lambda_c/\delta x$, whilst RBC3 requires approximately 16 mesh cells to meet the prescribed accuracy requirements on dissiplation errors and about 9 mesh points for dispersion errors. Nevertheless, this requirements are much lower than those e.g. of the third-order upwind scheme, for which 20 to 10 mesh cells per wavelength are required to meet the accuracy criterion on dissipation and dispersion, respectively. Figure 3(b) shows that the intrinsic numerical dissipation of high-order RBCq schemes acts as a selective filter with a shap cut-off at high frequency. It efficiently damps out grid-to-grid oscillations that can lead to numerical instabilities without affecting the resolved wave numbers.

Then, we investigate the spectrum of RBCq schemes in the case of multi-dimensional advection. Specifi-

that:



Figure 1: Representation in the complex plane of the Cauchy-stability criterion (33) for RBC schemes with different choices of Pade coefficients for the numerical dissipation term.

	ξ_c	$\lambda_c/\delta x$		ξ_c	λ_c/δ
RBC3	0.74	8.47	RBC3	0.40	15.5
RBC5	1.39	4.53	RBC5	1.03	6.08
RBC7	1.54	4.07	RBC7	1.24	5.06

Table 1: Dispersion accuracy limit of RBC schemes.

Table 2: Dissipation accuracy limit of RBC schemes.

cally, we consider the case $\dot{\mathbf{A}} = (\dot{A}, \dot{A})$, *i.e.* $\theta = \pi/4$, which corresponds to to advection velocity aligned with a mesh diagonal. Fig. 4 shows the phase error and damping function contours for this case that illustrate well the multidimensional nature of RBCq schemes. The bold black contour corresponds to the error criterion 10^{-3} . RBCq schemes display very law errors up to small reduced wave numbers in the diagonal direction. There is almost no dispersion in the direction normal to the diagonal. Inspection of contours of the damping function, displayed in Fig. 4 show that the multidimensional nature of RBCq schemes allows minimising numerical dissipation along the advection direction while dispersion is minimized in the direction orthogonal to the advection direction. This effect is more evident for higher order RBC5 and RBC7 schemes compared to RBC3.

5 Numerical experiments

For the present computations, the time derivative in the main residual and in the mid-point residuals is discretized by a Linear Multistep Method of order two, which is A-stable. Such a method is unconditionally



Figure 2: Contours of the phase error $\mathcal{P}_{\xi_0^*}$ (top) and damping function $\mathcal{D}_{\xi_0^*}$ (bottom) for pure advection in the *x* direction, $\dot{\mathbf{A}} = (\dot{A}, 0)$ *i.e.* $\theta = 0$.



Figure 3: 1D cut of \mathcal{P}_{ξ_0} and $\mathcal{D}_{\xi_0^*}$ for RBC schemes.

stable when the spatial approximation is dissipative. The time-discretization being fully-implicit, it is solved by using a dual-time stepping approach.



Figure 4: Contours of the phase error $\mathcal{P}_{\xi^*_{\pi/4}}$ (top) and damping function $\mathcal{D}_{\xi^*_{\pi/4}}$ (bottom) for pure advection in the diagonal direction, $\dot{\mathbf{A}} = (\dot{A}, \dot{A})$ *i.e.* $\theta = \pi/4$.

5.1 Advection of a sine wave

5.1.1 Validation of the the χ -criterion

We consider the initial-value problem:

$$\begin{cases} w_t + w_x + w_y = 0\\ w(x, y, 0) = \sin(2\pi(x+y)), \ -1 \le x \le 1, \ -1 \le y \le 1, \end{cases}$$

with periodic boundary conditions. The initial condition is shown on Fig. 5. In the diagonal direction, the wavelength is $\sqrt{2}/2$ and the advection speed is $\sqrt{2}$. The computational domain $[-1,1]^2$ is discretized by 25×25 square cells ($\delta x = \delta y = 0.08$), which corresponds to 12.5 points per wavelength. The time step is chosen such that: $\Delta t/\delta x = 0.05$. As expected, wave amplitude grows significantly when the χ -criterion is violated. The growth is faster with RBC3 than with RBC7. On the contrary, the sine wave is damped out when the χ -criterion is satisfied. The damping is very small for RBC7: 1.6% after 5000 time-iterations (t = 20), corresponding to a diagonal advection over a distance of 40 wavelengths.

5.1.2 Resolvability of RBC schemes

The sine wave advection can also be used to study numerically dissipation and dispersion errors for comparison with theoretical results of Section 4. Precisely, we want to investigate the minimum number of mesh points per wavelength required by RBC schemes. The computational domain $[-1, 1]^2$ is discretised by 48×48



Figure 5: Initial wave on the 25×25 mesh.

Figure 6: Wave amplitude versus time-iterations.

square cells ($\delta x = \delta y = 1/24$) with periodic boundary conditions and the number of points per wave length are chosen to be 6, 8 and 16, which corresponds to the theoretical accuracy limits of RBC schemes as stated in Section 4. A very small time step is used in order to rule out errors due to the time integration scheme $\Delta t = 0.002 \ (\Delta t/\delta x = 8.3 \times 10^{-5}).$

We consider the initial-value problem:

$$\begin{cases} w_t + Aw_x + Bw_y = 0\\ w(x, y, 0) = \sin\left[2\pi\left(\frac{x}{\delta x \ n_x} + \frac{\beta y}{\delta y \ n_y}\right)\right], \ -1 \leqslant x \leqslant 1, \ -1 \leqslant y \leqslant 1, \end{cases}$$

where w is a scalar quantity, $\mathbf{A} = (A, B)$ is the advection velocitie vector chosen equal to (1,0) and (1,1) corresponding to the advection directions θ equal to 0 and $\pi/4$ respectively, n_x and n_y are the number of points per wavelength in the x- and y-direction and β is equal to 1 for a 2D-sine wave and 0 for a sine wave in the x-direction.



Figure 7: Discretization of the sine wave with 6, 8 and 16 points per wavelength.

Fig. 7 shows representations of the initial 1D-sine wave using 6, 8 and 16 points per wavelength. The choice 6 or 8 points per wavelength gives a rathely rough representation of the wave. The amplitude of the sine wave versus the number of wavelength traveled by the sine wave in the x-direction is displayed for the three values of n_x and several RBC schemes in Fig. 8 for the case of the advection along a single mesh direction, so that the damping of the sine wave only depends on the number of mesh points along this direction. Fig. 9 shows two cases with rather different discretizations in the direction orthogonal to



Figure 8: Wave amplitude versus number of wavelengths traveled in the case of $\mathbf{A} = (1, 0)$ *i.e.* $\theta = 0$.

the advection direction giving the same damping error in time even with RBC3 which totally damps a wave discretized by 6 points per wavelength. This confirms the results of Fig. 2 since errors only come from the advection direction in the case of an advection along a mesh direction. It has to be stressed that this behaviour comes from the use the dissipation matrices Φ_i (i = 1, 2, 3) which adapt the dissipation to the characteristics of the flow [11]. On the contrary of the behaviour of a filter is less costly but do not adapt to flow and filter in all directions whatever the flow. Fig. 10 shows that even after 48 wavelengths traveled their is almost no dispersion errors in the computed solutions. This confirms that dissipation error predominates compared to dispersion for RBCq schemes.



Figure 9: Initial conditions inducing the same dissipation and dispersion error in the case of $\theta = 0$.

5.2 Converging cylindrical shock

When the χ -criterion is violated in an RBC scheme, the lack of dissipation occurs in some oblique flow directions. So, we consider a test case involving all the flow directions and a large range of wave numbers, that is a 2-D simulation of a converging cylindrical shock on a uniform Cartesian mesh. Of course, this axisymmetric problem could be solved more easily as a 1-D problem in polar coordinates. Here, the 2-D Euler equations, for a perfect gas with a specific heat ratio $\gamma = 1.4$, are solved in a square domain $[-0.5, 0.5]^2$.

At time t = 0, a cylindrical shock (satisfying the Rankine-Hugoniot relations) lies on a circle of center (x, y) = (0, 0) and radius $r_0 = 0.25$. Inside the cylindrical shock (state 0), the fluid is at rest and at pressure



Figure 10: Sine wave after an advection over 48 wavelengths in the case of $\theta = 0$ with n_x equal 6 and 16.

 p_0 . The pressure just behind the shock is $p_1 = 2.4 p_0$ at t = 0. Outside the cylindrical shock, the initial state corresponds to a steady converging flow, *i.e.* the flow at a radius $r > r_0$ is related to the state 1 just behind the shock by the conservation of mass ($\rho V r = \rho_1 V_1 r_0$ where ρ is the density and V the radial velocity), the conservation of total enthalpy and of the entropy. For improving the initial representation of the shock on the Cartesian mesh, the vector w of conservative variables is defined as follows in the mesh cells intersecting the shock:

$$w_* = (1 - \theta)w_0 + \theta w_1, \quad 0 \le \theta \le 1$$

where $\theta \delta x \delta y$ is the cell area fraction in state 1.

During the evolution, the cylindrical shock increases in strength as it converges towards the axis. When the shock reaches the axis, it is reflected as a divergent shock. At the very instant of reflection, the pressure at the axis becomes infinite in the Euler model. To avoid a numerical difficulty, the Cartesian mesh is set so that the axis corresponds to a cell vertex and not to a cell center. This prevents the computation of any unphysical quantity on the axis. Note also that the outside boundary is not affected by the perturbations coming from the shock motion in the duration of the present simulation.

Chisnell [12] gave in 1957 an analytical estimation of the pressure behind a moving cylindrical shock, the theoretical arguments of which were improved by Whitham [13]. According to this theory, the Mach number M of the shock wave (relative to the fluid at rest) at radius r is solution of the differential equation:

$$\frac{dM}{dr} = -\frac{1}{r} \frac{(M^2 - 1).K(M)}{2M}$$
(39)

where

$$\begin{split} K(M) &= 2 \left[\left(1 + \frac{2}{\gamma + 1} \frac{1 - \mu^2}{\mu} \right) \left(2\mu + 1 + \frac{1}{M^2} \right) \right]^{-1} \\ \mu &= \left[\frac{(\gamma - 1)M^2 + 2}{2\gamma M^2 - (\gamma - 1)} \right]^{\frac{1}{2}}. \end{split}$$

For $\gamma = 1.4$, the function K(M) decreases slowly from 0.5 for M = 1 to $14/(17 + 7\sqrt{7}) \approx 0.394$ for $M \to \infty$. Starting from the initial condition $M_0 = M(r_0)$, the equation (39) can easily be solved numerically with a high accuracy. An exact solution is also available [12], but its expression is very complicated and defined in the form r = r(M).

The pressure behind the shock in motion is deduced from M = M(r) using the Rankine-Hugoniot relations:

$$p_1 = \frac{2\gamma M^2 - (\gamma - 1)}{\gamma + 1} p_0 \tag{40}$$

The converging cylindrical shock problem is solved by the RBC schemes on a 800×800 Cartesian mesh with $\Delta t/\delta x = 0.21$. When the χ -criterion is violated, the computation fails after a few time iterations (one iteration for RBC3 with b = 0 and 26 for RBC7 with c = 0). When the χ -criterion is satisfied, the computation succeeds, even after the shock reflection on the axis. In this case, the pressure profiles along the x-axis are shown on Fig. 11 for the RBC3, RBC5 and RBC7 schemes at different times, together with the analytical pressure behind the shock deduced from (39)-(40). The agreement between the numerical solution and the Chisnell theory is very good. The shape of the converging shock computed by the RBC7 scheme at different times is shown on Fig. 12. This shape has been defined as the isobar lines of level $\frac{1}{2}(p_1+p_0)$ at each time. The converging shock appears to be perfectly circular on the Cartesian mesh. Clearly, Fig. 11 reveals the oscillatory nature of the shock profiles computed by the present high order schemes, specially by RBC5 and RBC7. It should be noted that the computations have been achieved by a strict use of the method described in the present paper: there is no limiter, no entropy correction, no filtering or other additive. In these conditions, it appears that a good design of the dissipative operator and low dispersion allow the calculation of a difficult test case, even if the discrete shock is not perfectly represented.



Figure 11: Pressure along the x-axis at different times for RBC schemes satisfying the χ -criterion.



Figure 12: Shape of the converging shock at different times computed by RBC7.

5.3 Inviscid Taylor-Green vortex

The use of high accuracy schemes is of utmost importance for numerical simulation of turbulent flows since they enable capturing flow structures from large to small scales at an acceptable computational cost. The Taylor-Green vortex is a classical example of nonlinear fluid flow with kinetic energy transfert from large to small scales and is a good milestone to assess the applicability of numerical schemes to Large Eddy Simulation (LES).

A three-dimesional vortex is set as an initial condition for 3D-computation. Because of vortex-stretching and vortex tilting mechanisms, the vortex breaks up, giving origin to smaller and smaller structures. At finite Reynolds number, the kinetic energy is transfered from larger to smaller scales and dissipated by the smallest one; the test case gives thereby a simple model of the energy cascade. The case of the inviscid Taylor-Green vortex (Fig. 13) does not completely model the energy cascade since the energy is not dissipated through physical viscosity.

The initial conditions of the computation are:

$$\begin{cases} u(x, y, z, 0) &= \sin(x)\cos(y)\cos(z) \\ v(x, y, z, 0) &= -\cos(x)\sin(y)\cos(z) \\ w(x, y, z, 0) &= 0 \\ \rho(x, y, z, 0) &= 1 \\ p(x, y, z, 0) &= p_0 + \frac{\rho}{16} [(\cos(2z) + 2)(\cos(2x) + \cos(2y)) - 2] \end{cases}$$

Computation parameters are exactly the same as those taken by Shu *et al.* in [14]. This way our results



Figure 13: Iso surface Q=0 colored by k (computed with RBC5 on the 128^3 mesh). The figure show initial phases of vortex break-up.

can be compared to those of Shu *et al.* [14] using the WENO-5 scheme and the Fourier collocation method with a sharp-cuttof filter (F-SF-23N) which can be considered as a benchmark.

Comparisons are done on integral quantities such as enstrophy (Fig. 14) and kinetic energy (Fig. 15). This quanties are normalised by their initial values and defined in [14]. Integral quantities are more difficult to match since all the errors committed in the computational domain will be taken into account. However, RBC schemes have a good resolvability, thanks to their low dissipation and dispersion errors. Even the third-order accurate scheme gives better results than the WENO-5 scheme.

Morover, the flow physic of the flow is well captured, and vortex stretching mechanisms are clearly visible (see Figure 16). These results are encouraging showing a good resolvability of high-order RBC schemes and their aptitude to compute a flow with a large range of scales.

6 Conclusions

A comprehensive study of the dissipation properties of a family of Residual-Based Compact schemes has been presented for 2-D and 3-D hyperbolic systems of conservation laws. The residual-based numerical dissipation operator has been shown to be the counterpart of a high-order differential operator based on pure and mixed derivatives of even order. A general criterion (Thm. 3.1) has been established for this operator to be dissipative. The resolvability of the RBC schemes has been quantified through its spectral



Figure 14: Comparison of the time evolution of the normalized total enstrophy on a 64^3 mesh (a) and on a 128^3 mesh (b) with different numerical schemes.



Figure 15: Comparison of the time evolution of the normalized total kinetic energy on a 64^3 mesh (a) and on a 128^3 mesh (b) with different numerical schemes.



Figure 16: Visualization of the vortex stretching through iso-contours of vorticity (computed with RBC5 on the 128^3 mesh).

properties.

Numerical tests confirm the theoretical results and demonstrate the importance of a well-designed dissipation operator for numerical simulations in gas dynamics. The low dissipative and dispersive errors introduced of RBC schemes make them excellent candidates for compressible turbulent flow simulations, since they combine good shock capturing capabilities with high resolvability of fine flow structures. Specifically, RBC5 and RBC7 schemes provide very encouraging results, and their applicability to complex unsteady flow will be further investigated in future work.

Acknowledgements

This research has been done within the framework of the European project IDIHOM (Industrialization of High Order Methods) which aims to promote the use of high-order numerical methods by the European aerospace industry.

References

- [1] Alain Lerat and Christophe Corre. A residual-based compact scheme for the compressible Navier-Stokes equations. *Journal of Computational Physics*, 170:642–675, 2001.
- [2] Christophe Corre, Fabrice Falissard, and Alain Lerat. High-order residual-based compact schemes for compressible inviscid flows. Computers & Fluids, 36:1567–1582, 2007.
- [3] Alain Lerat and Christophe Corre. Higher order residual-based compact schemes on structured grids, chapter CFD-higher order discretization methods, pages 1–111. 34th Comput. Fluid Dyn. Course, von Karman Institute for Fluid Dynamics. VKI LS 2006-1, 2006.
- [4] Alain Lerat and Christophe Corre. Residual-based compact schemes for multidimensional hyperbolic systems of conservation laws. Computers & Fluids, 31:639–661, 2002.
- [5] Christophe Corre, Grégoire Hanss, and Alain Lerat. A residual-based compact scheme for the unsteady compressible Navier-Stokes equations. Computers & Fluids, 34:561–580, 2005.
- [6] Bertrand Michel. Contribution à la simulation numérique efficace des écoulements dans les prises d'air supersoniques. PhD thesis, Arts et Métiers ParisTech, 2004.
- [7] Alain Lerat, Karim Grimich, and Paola Cinnella. On the design of high order residual-based dissipation for unsteady compressible flows. *Journal of Computational Physics*, 2012. submitted.
- [8] Sanjiva K. Lele. Compact finite difference schemes with spectral-like resolution. Journal of Computational Physics, 103:16–42, 1992.
- [9] Christopher K. W. Tam and Jay C. Webb. Dispersion-relation-preserving finite difference schemes for computational acoustics. *Journal of Computational Physics*, 107:262–281, 1993.
- [10] C. Bailly and C. Bogey. A family of low dispersive and low dissipatibe explicit schemes for flow and noise computations. *Journal of Computational Physics*, 194:194–214, 2004.
- [11] Ying Huang and Alain Lerat. Second-order upwinding through a characteristic time-step matrix for compressible flow calculations. *Journal of Computational Physics*, 142:445–472, 1998.
- [12] R. F. Chisnell. The motion of a shock wave in a channel, with applications to cylindrical and spherical shock waves. *Journal of Fluid Mechanics*, 2:286–298, 1957.
- [13] G. B. Whitham. On the propagation of shock waves through regions of non uniform area or flow. Journal of Fluid Mechanics, 4:337–360, 1958.
- [14] Chi-Wang Shu, Wai-Sun Don, David Gottlieb, Oleg Schilling, and Leland Jameson. Numerical convergence study of nearly incompressible, inviscid taylor-green vortex flow. *Journal of Scientific Computing*, 24(1), July 2005.