Journal of Computational Physics **170,** 642–675 (2001)
doi:10.1006/jcph.2001.6755, available online at http://www.idealibrary.com on IDEAL®

A Residual-Based Compact Scheme for the Compressible Navier–Stokes Equations

Alain Lerat and Christophe Corre

SINUMEF Laboratory, ENSAM, 151, Boulevard de l'Hopital, 75013 Paris, France
E-mail: lerat@paris.ensam.fr, corre@paris.ensam.fr

Received May 10, 2000; revised February 14, 2001

A simple and efficient time-dependent method is presented for solving the steady compressible Euler and Navier–Stokes equations with third-order accuracy. Owing to its residual-based structure, the numerical scheme is compact without requiring any linear algebra, and it uses a simple numerical dissipation built on the residual. The method contains no tuning parameter. Accuracy and efficiency are demonstrated for 2-D inviscid and viscous model problems. Navier–Stokes calculations are presented for a shock/boundary layer interaction, a separated laminar flow, and a transonic turbulent flow over an airfoil. © 2001 Academic Press

Key Words: high accuracy; compact scheme; multidimensional; Navier–Stokes equations; compressible flow; aerodynamics.

---

1. INTRODUCTION

The calculation of some compressible flow problems requires the use of a high-accuracy numerical scheme. Classical examples are the direct or large eddy simulation of turbulent flows, instability phenomena such as aerodynamic buffeting, aeroacoustic problems, etc.

To construct a high-accuracy scheme, two methods are classically considered: one simply increases the grid stencil of the scheme; the other uses compact or Pade approximations. Compact schemes are attractive because of their narrow numerical dependence domain. In computational fluid dynamics, they have been mainly developed with centered schemes (see for instance [1, 3, 13, 17, 25, 29, 30]) using artificial viscosity, limiters, or numerical filters for shock capturing. Upwind compact schemes have also been proposed for calculating compressible flows (see for instance [11, 20, 28]).

Here, we present an efficient compact time-dependent method for solving the steady compressible Euler and Navier–Stokes equations with high accuracy. In this method, the process for getting high accuracy as well as the construction of the numerical dissipation relies strongly on the residual vanishing at steady state. Starting from a basic centered scheme, the idea of increasing the accuracy order is not to cancel the leading term of

0021-9991/01 $35.00

the truncation error, but to add a new error term so that the resulting truncation error is expressed in terms of derivatives of the residual only. This does not lead to a correction for each derivative in the governing equations but to a compact approximation that cannot be derived by Pade formulae and requires no linear algebra. Similarly, the numerical dissipation is residual-based and is consistent with an operator vanishing at steady state.

For simplicity and robustness, we use a first-order dissipation operator, which is compatible with any accuracy order. More precisely, by applying this operator, a centered scheme accurate at order $2p$ becomes a dissipative scheme accurate at order $2p - 1$ (dissipation at steady state comes from the truncation error of the dissipation operator). This paper studies the case $p = 2$. The stencil of the dissipative scheme is then composed of 3×3 points for the 2-D Euler equations and $(3 \times 3) + 4 = 13$ points for the 2-D Navier–Stokes equations.

For constructing the dissipation operator, we follow along the lines of some of our previous studies on second-order-accurate schemes of the Lax–Wendroff type [18, 19, 21] and recent attempts to improve them by introducing a so-called characteristic time-step matrix [7, 14, 15]. In this work, we present a rational construction of numerical dissipation not derived from a modification of the Lax–Wendroff approach. This viewpoint brings a better understanding of the dissipation operator and allows its application to high-accuracy schemes. It should be noted that the dissipation operator is parameter free and that we add no limiters or any other correction. In addition, the dissipation operator (and thus the numerical solution) is independent of the time step used to reach the steady state.

The paper is organized as follows. Section 2 presents the design principles of the scheme for a two-dimensional hyperbolic system of conservation laws. Section 3 describes the complete determination of the dissipation operator for the two-dimensional case. Section 4 presents the fully discrete form of the scheme and its stability domain. Section 5 extends the method to the compressible Navier–Stokes equations with a real third-order accuracy. Section 6 briefly presents the main features of an implicit version of the scheme. Section 7 describes precise validations of the accuracy order and the efficiency for four model test cases: a 2-D circular advection problem, a 1-D boundary layer model, a 2-D Poiseuille-type problem and a 1-D inviscid Bürgers model. Section 8 presents first applications to the 2-D compressible Navier–Stokes equations: a laminar flow over a NACA0012 airfoil; a shock/laminar boundary layer interaction on a flat plate; and a turbulent transonic flow over a RAE2822 airfoil.These applications show the high accuracy and shock-capturing property of the third-order scheme. Finally, some conclusions are drawn and possible extensions of the method are discussed.

2. DESIGN PRINCIPLES

2.1. *Residual-Based Compactness*

Consider an initial-value problem for the hyperbolic system of conservation laws,

$$w_t + f_x + g_y = 0, \tag{1}$$

where t is the time, x and y are space coordinates, $w = w(x, y, t)$ is the state vector, and $f = f(w)$ and $g = g(w)$ are flux components depending smoothly on w. The Jacobian matrices of the fluxes are

$$A = \frac{df}{dw} \quad \text{and} \quad B = \frac{dg}{dw}.$$

Suppose that system (1) is discretized by a conservative scheme of the simple form

$$\frac{1}{\Delta t}(w^{n+1} - w^n) + \tilde{r}^n = 0, \tag{2}$$

where w^n is the numerical solution at time level $n\Delta t$, and $\tilde{r}^n$ denotes some centered difference approximation to the steady state residual

$$r = f_x + g_y. \tag{3}$$

Let $v_{j,k}$ be a mesh function defined on a uniform Cartesian mesh ($x_j = j\,\delta x$, $y_k = k\,\delta y$), with steps δx and δy of the same order, say $O(h)$. Consider the basic difference and average operators:

$$\begin{aligned} (\delta_1 v)_{j+\frac{1}{2},k} &= v_{j+1,k} - v_{j,k} & (\delta_2 v)_{j,k+\frac{1}{2}} &= v_{j,k+1} - v_{j,k} \\ (\mu_1 v)_{j+\frac{1}{2},k} &= \frac{1}{2}(v_{j+1,k} + v_{j,k}) & (\mu_2 v)_{j,k+\frac{1}{2}} &= \frac{1}{2}(v_{j,k+1} + v_{j,k}). \end{aligned}$$

For smooth solutions, the simplest centered discretization of r is

$$\tilde{r}_0 = \frac{1}{\delta x}\delta_1\mu_1 f + \frac{1}{\delta y}\delta_2\mu_2 g, \tag{4}$$

where the subscripts j, k are omitted for concision. For an exact steady solution ($r = 0$), the truncation error of (4) reads

$$\epsilon_0 = \frac{\delta x^2}{6} f_{xxx} + \frac{\delta y^2}{6} g_{yyy} + O(h^4).$$

The second-order approximation (4) can be improved by correcting the leading error term. Classically, this can be achieved either in compact or noncompact formulation.

Noncompact fourth-order approximation (NC). Here, the second-order error term is corrected explicitly as

$$\tilde{r}_{NC} = \frac{1}{\delta x}\delta_1\mu_1\left(I - \frac{1}{6}\delta_1^2\right) f + \frac{1}{\delta y}\delta_2\mu_2\left(I - \frac{1}{6}\delta_2^2\right) g, \tag{5}$$

where I is the identity operator. This formula involves five points per direction. For $r = 0$, the truncation error of (5) is

$$\epsilon_{NC} = -\frac{1}{30}(\delta x^4 f_{xxxxx} + \delta y^4 g_{yyyyy}) + O(h^6).$$

Pade–compact fourth-order approximation (PC). Here, the second-order term is corrected implicitly using Pade fractions,

$$\nabla_1^{\text{Pade}} f = \frac{\delta_1\mu_1}{I + \frac{1}{6}\delta_1^2} f, \quad \nabla_2^{\text{Pade}} g = \frac{\delta_2\mu_2}{I + \frac{1}{6}\delta_2^2} g$$

that are computed via the numerical solution of linear algebraic systems. The residual is approximated by

$$\tilde{r}_{PC} = \frac{1}{\delta x}\nabla_1^{\text{Pade}} f + \frac{1}{\delta y}\nabla_2^{\text{Pade}} g. \tag{6}$$

It involves three points per direction. For $r = 0$, the truncation error of (6) is

$$\epsilon_{PC} = -\frac{1}{180}(\delta x^4 f_{xxxxx} + \delta y^4 g_{yyyyy}) + O(h^6).$$

As is well known, the numerical constant in the leading term of the truncation error is smaller with this Pade-type scheme than with the noncompact fourth-order approximation.

Residual-based compact approximation (RBC). We now introduce a compact discretization of r that does not require the solution of a linear system. The idea is no longer to correct the truncation error of (4), but to add a new error term in each difference formula so that the leading error term on r is expressed in terms of derivatives of r only. Namely, r is discretized by summing two centered differences, each of second-order accuracy:

$$\tilde{r}_{RBC} = \frac{1}{\delta x}\delta_1\mu_1\left(I + \frac{1}{6}\delta_2^2\right) f + \frac{1}{\delta y}\delta_2\mu_2\left(I + \frac{1}{6}\delta_1^2\right) g. \tag{7}$$

For $r = 0$, the truncation error of (7) is

$$\begin{aligned}\epsilon_{RBC} &= \frac{\delta x^2}{6} f_{xxx} + \frac{\delta y^2}{6} f_{xyy} + \frac{\delta y^2}{6} g_{yyy} + \frac{\delta x^2}{6} g_{xxy} + O(h^4)\\ &= \frac{\delta x^2}{6} r_{xx} + \frac{\delta y^2}{6} r_{yy} + O(h^4) = O(h^4).\end{aligned}$$

Expanding the truncation error up to order 6 gives

$$\begin{aligned}\epsilon_{RBC} &= \frac{\delta x^4}{120} r_{xxxx} + \frac{\delta x^2\delta y^2}{36} r_{xxyy} + \frac{\delta y^4}{120} r_{yyyy} + \frac{\delta x^4}{180} g_{xxxxy} + \frac{\delta y^4}{180} f_{xyyyy} + O(h^6)\\ &= \frac{1}{180}(\delta x^4 g_{xxxxy} + \delta y^4 f_{xyyyy}) + O(h^6).\end{aligned}$$

Since $r = 0$, this truncation error can also be written as

$$\epsilon_{RBC} = -\frac{1}{180}(\delta x^4 f_{xxxxx} + \delta y^4 g_{yyyyy}) + O(h^6),$$

and we note that the numerical constant is the same as in ϵ_{PC}.

Fourth-order accuracy is obtained with RBC using 3×3 points only and *no linear algebraic system*. Compared to NC, the RBC formula (7) uses the near points $j \pm 1, k \pm 1$ (corner points) instead of the more distant points $j \pm 2, k$ and $j, k \pm 2$. This stencil narrowing has several advantages:

(1) unconditionally stable implicit versions can be easily derived—see Section 6;
(2) the numerical treatment of boundary conditions is straightforward;
(3) the truncation error constant is divided by 6;
(4) accurate extension to irregular meshes is easier on a compact stencil.

Note that an approximation similar to (7) can be found in a paper by Abarbanel and Kumar [1].

Higher-order. RBC approximation can be defined at any order of accuracy. For instance, the formula

$$\tilde{r}^{(6)}_{RBC} = \frac{1}{\delta x}\delta_1\mu_1\left(I + \frac{1}{6}\delta_2^2 - \frac{1}{180}\delta_2^4\right) f + \frac{1}{\delta y}\delta_2\mu_2\left(I + \frac{1}{6}\delta_1^2 - \frac{1}{180}\delta_1^4\right) g$$

is sixth-order accurate at steady state (see the truncation error of the fourth-order RBC approximation). It uses 5 × 5 points only and no linear algebraic system.

2.2. *Residual-Based Dissipation*

Of course the present high-accuracy schemes cannot be used without some numerical dissipation. With the present Euler explicit time discretization, they would even be unstable. There are two classical (and equivalent) ways to introduce dissipation:

- find an upwind expression of high accuracy for $\tilde{r}$
- add a dissipative term of high order to the scheme.

Here, we follow a different approach and add a dissipative term of order 1 that vanishes at steady state.

More precisely, we now consider

$$\frac{1}{\Delta t}(w^{n+1} - w^n) + \tilde{r}^n = \tilde{P}^n, \tag{8}$$

where $\tilde{P}^n$ is a centered difference approximation to a partial differential operator P such that:

- P brings “parabolicity” or “dissipation” (see Section 3)
- $P = O(h)$, for robustness in the transient phase
- P vanishes at steady state ($r = 0 \Rightarrow P = 0$)
- P is independent of Δt.

A simple choice of P satisfying these criteria is

$$P = \frac{\delta x}{2}(\Phi_1 r)_x + \frac{\delta y}{2}(\Phi_2 r)_y = \frac{\delta x}{2}[\Phi_1(f_x + g_y)]_x + \frac{\delta y}{2}[\Phi_2(f_x + g_y)]_y, \tag{9}$$

where Φ_1 and Φ_2 are matricial coefficients depending only on the space increments and on the flux Jacobian matrices

$$\Phi_1 = \Phi_1(A, B; \delta x, \delta y) = O(1)$$
$$\Phi_2 = \Phi_2(A, B; \delta x, \delta y) = O(1).$$

Functions Φ_1 and Φ_2 will be precisely defined in the next section.

Choosing a 3 × 3-point stencil for scheme (8), we define the discrete operator $\tilde{P}$ from classical centered formulae of second-order accuracy and, for $\tilde{r}$, we set

$$\tilde{r} = \frac{1}{\delta x}\delta_1\mu_1\left(I + \frac{\theta}{6}\delta_2^2\right) f + \frac{1}{\delta y}\delta_2\mu_2\left(I + \frac{\theta}{6}\delta_1^2\right) g \tag{10}$$

to include the simply centered approximation (4)—for $\theta = 0$—and the RBC approximation (7)—for $\theta = 1$.

For an exact steady solution ($w_t = r = 0$), the truncation error of scheme (8) is

$$\epsilon = \frac{\delta x^2}{6} r_{xx} + \frac{\delta y^2}{6} r_{yy} + \frac{\theta - 1}{6}(\delta x^2 g_x + \delta y^2 f_y)_{xy} - \frac{\delta x}{2}(\Phi_1 r)_x - \frac{\delta y}{2}(\Phi_2 r)_y + O(h^3).$$

Thus ϵ is $O(h^2)$ for $\theta \neq 1$ and $O(h^3)$ for $\theta = 1$. Consequently, the dissipative scheme can be third-order accurate at steady state on a 3 × 3-point stencil.

Remark.

(a) The Lax–Wendroff scheme includes an operator similar to P but dependent on Δt and not so efficient as the one constructed in Section 3.

(b) The exact operator P is not dissipative at steady state ($r = 0 \Rightarrow P = 0$). However numerical dissipation exists at steady state and comes from the truncation error of $\tilde{P}$ which is $O(h^3)$ and involves fourth derivatives (see Section 4).

(c) Clearly, by choosing a 5 × 5-point stencil, using for $\tilde{r}$ a sixth-order approximation and improving similarly the approximation for $\tilde{P}$, we can reach fifth-order accuracy at steady state. In fact, the use of the present kind of dissipation does not bound the order of accuracy at steady state.

3. CHOOSING Φ_1 AND Φ_2

3.1. *Method*

The question now is to determine functions Φ_1 and Φ_2 ensuring the stability of the time-dependent scheme (8). During the unsteady evolution, this scheme is first-order accurate. More precisely, it approximates

$$w_t + \frac{\Delta t}{2} w_{tt} + O(\Delta t^2) + f_x + g_y + O(h^2) = P + O(h^3), \tag{11}$$

where $P = O(h)$. Assuming that $\Delta t = O(h)$, this gives

$$w_t + f_x + g_y = O(h).$$

By derivating with respect to time,

$$\begin{aligned} w_{tt} &= -(f_{tx} + g_{ty}) + O(h) = -(Aw_t)_x - (Bw_t)_y + O(h) \\ &= [A(f_x + g_y)]_x + [B(f_x + g_y)]_y + O(h). \end{aligned} \tag{12}$$

By substituting (12) into (11) and using the expression (9) for P, we find that scheme (8) is a second-order approximation to

$$w_t + f_x + g_y = P' \tag{13}$$

with

$$\begin{aligned} P' &= \frac{1}{2}[C_1(f_x + g_y)]_x + \frac{1}{2}[C_2(f_x + g_y)]_y \\ &= \frac{1}{2}(C_1 A w_x + C_1 B w_y)_x + \frac{1}{2}(C_2 A w_x + C_2 B w_y)_y, \end{aligned}$$

where

$$C_1 = \delta x \Phi_1 - \Delta t A, \quad C_2 = \delta y \Phi_2 - \Delta t B.$$

A necessary condition for the stability of scheme (8) is that the differential operator P' be dissipative.

3.2. *Single Conservation Law*

Suppose that Eq. (1) is a single conservation law, i.e., w, f, g, A, B, Φ_1, and Φ_2, are scalar-valued functions, and consider the quadratic form Q' associated with the operator P',

$$Q' = \frac{1}{2}[C_1 A \xi^2 + (C_1 B + C_2 A)\xi\eta + C_2 B \eta^2] = \frac{1}{2} X^T M' X,$$

with

$$X = \begin{bmatrix} \xi \\ \eta \end{bmatrix} \qquad M' = \begin{bmatrix} C_1 A & \frac{1}{2}(C_1 B + C_2 A) \\ \frac{1}{2}(C_1 B + C_2 A) & C_2 B \end{bmatrix}.$$

Dissipativity (in a broad sense) of P' means that the quadratic form Q' is nonnegative definite, i.e., the eigenvalues of the symmetric matrix M' are positive or null, which is equivalent to

$$\begin{cases} C_1 B = C_2 A \\ C_1 A \geq 0, \quad C_2 B \geq 0. \end{cases}$$

This can also be written as

$$\frac{\delta x \Phi_1}{A} = \frac{\delta y \Phi_2}{B} \tag{14}$$

$$\delta x \Phi_1 A \geq \Delta t A^2, \quad \delta y \Phi_2 B \geq \Delta t B^2. \tag{15}$$

Since (15) shows that Φ_1 is of the sign of A and that Φ_2 is of the sign of B, we introduce the notations

$$\Phi_1 = \operatorname{sgn}(A)\Phi, \quad \Phi \geq 0$$
$$\Phi_2 = \operatorname{sgn}(B)\Psi, \quad \Psi \geq 0$$

and

$$\alpha = \frac{\delta x |B|}{\delta y |A|}. \tag{16}$$

With these notations, the conditions (14)–(15) become:

$$\Psi = \alpha \Phi, \tag{17}$$

$$\frac{\Delta t}{\delta x}|A| \leq \Phi, \quad \frac{\Delta t}{\delta y}|B| \leq \Psi. \tag{18}$$

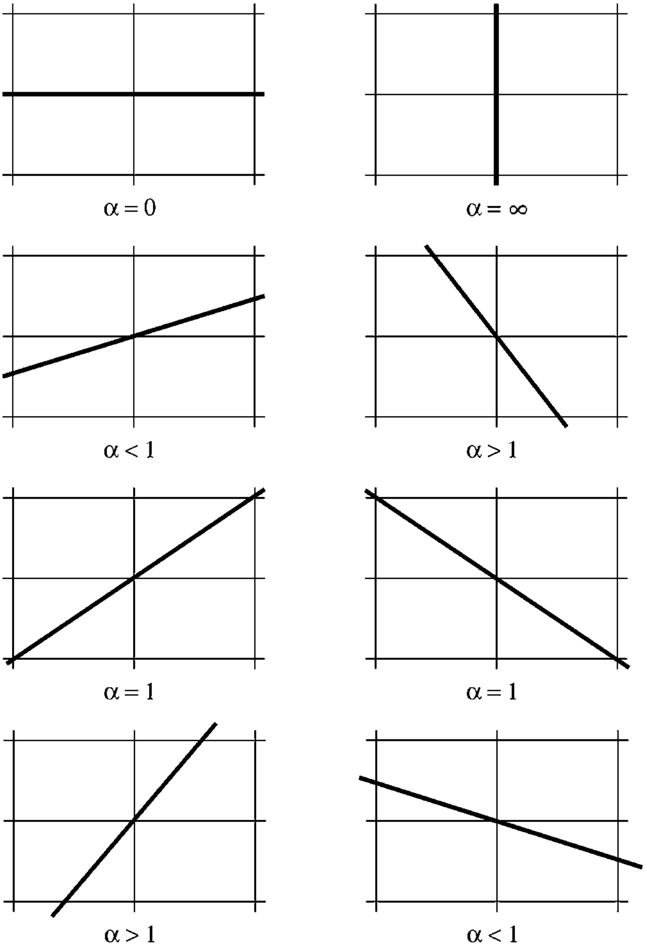

FIG. 1. Parameter $\alpha = \frac{\delta x}{\delta y}\frac{|B|}{|A|}$ representing the local advection orientation.

Note that the parameter α characterizes the local advection direction with respect to the mesh. For $\alpha < 1$, that is $\frac{|A|}{\delta x} > \frac{|B|}{\delta y}$, the advection direction is between the x-direction and a mesh diagonal (see Fig. 1). For $\alpha = 1$, the advection is along a mesh diagonal. For $\alpha > 1$, the advection direction is between the y-direction and a mesh diagonal. Note also that, due to (17) and $\frac{|B|}{\delta y} = \alpha\frac{|A|}{\delta x}$, the two inequalities (18) are equivalent.

Expressed with the functions Φ and Ψ, the dissipation operator reads as follows:

$$P = \frac{\delta x}{2}[\Phi \operatorname{sgn}(A)(f_x + g_y)]_x + \frac{\delta y}{2}[\Psi \operatorname{sgn}(B)(f_x + g_y)]_y. \qquad (19)$$

Consider now two choices of the pair (Φ, Ψ) that are compatible with (17)–(18) and lead to the simplest expression of P. They are as follows:

(a) $\Phi = 1, \Psi = \alpha$.

Due to (18), the time step is bounded by $\frac{\Delta t}{\delta x}|A| \leq 1$. Unfortunately, for large values of α (quasi-vertical advection), expression (19) shows that the dissipation in the y-direction becomes very high, i.e., the scheme becomes highly dissipative in the advection direction, which is unacceptable.

(b) $\Phi = 1/\alpha$, $\Psi = 1$.

The time step is bounded by $\frac{\Delta t}{\delta y}|B| \leq 1$. Now, for small values of α (quasi-horizontal advection), the dissipation in the x-direction becomes very high, which is also unacceptable.

Since the dissipation condition (17) prevents us from choosing $\Phi = \Psi = 1$, except in the very special case where $\alpha = 1$ (diagonal advection), we keep Φ and Ψ as close as possible to 1. More precisely, for each $\alpha > 0$, we look for $\Phi \geq 0$ minimizing the distance

$$|\Phi - 1| + |\Psi - 1| = |\Phi - 1| + |\alpha\Phi - 1|.$$

The optimal solution is found to be

$$\Phi = \begin{cases} 1 & \text{if } \alpha \leq 1 \\ \frac{1}{\alpha} & \text{if } \alpha > 1 \end{cases}$$

and

$$\Psi = \alpha\Phi = \begin{cases} \alpha & \text{if } \alpha < 1 \\ 1 & \text{if } \alpha \geq 1. \end{cases}$$

These functions are plotted in Fig. 2. They can also be written as

$$\begin{aligned} \Phi &= \min\left(1, \frac{1}{\alpha}\right) = \min\left(1, \frac{\delta y|A|}{\delta x|B|}\right) \\ \Psi &= \min(1, \alpha) = \min\left(1, \frac{\delta x|B|}{\delta y|A|}\right). \end{aligned} \tag{20}$$

For the above Φ and Ψ functions, condition (18) is equivalent to

$$\frac{\Delta t}{\delta x}|A| \leq 1 \quad \text{and} \quad \frac{\Delta t}{\delta y}|B| \leq 1,$$

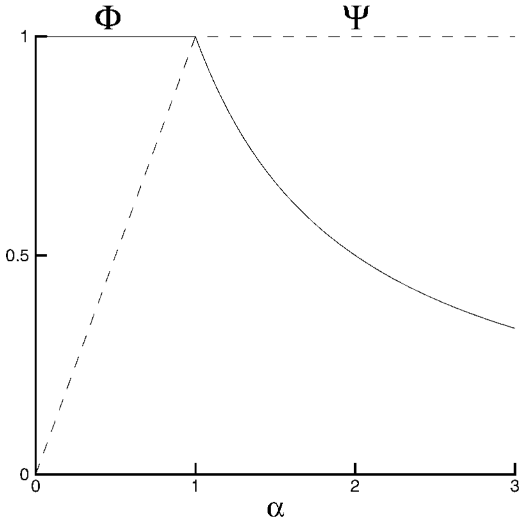

FIG. 2. Coefficients Φ and Ψ versus advection orientation.

that is,

$$\Delta t \leq \min\left(\frac{|A|}{\delta x}, \frac{|B|}{\delta y}\right).$$

Remember that this is only a necessary condition for stability. The true stability domain of the scheme will be given in Section 4.

3.3. *System of Conservation Laws*

For an easy implementation of scheme (8) applied to the general system (1), the matrices Φ_1 and Φ_2 are constructed through a direct extension of Section 3.2: the eigenvectors of Φ_1 are those of A; the eigenvectors of Φ_2 are those of B; and their eigenvalues are deduced from the scalar definitions of Φ_1 and Φ_2. More precisely, let T_A (respectively T_B) be a matrix whose column vectors are the right eigenvectors of the Jacobian matrix A (respectively B) and let $a^{(i)}$ (respectively $b^{(i)}$) be the eigenvalues of A (respectively B). Owing to the hyperbolicity of (1), $a^{(i)}$ and $b^{(i)}$ are real, and T_A and T_B are regular, so that

$$A = T_A Diag\left[a^{(i)}\right] T_A^{-1}, \quad B = T_B Diag\left[b^{(i)}\right] T_B^{-1},$$

where $Diag[d^{(i)}]$ denotes a diagonal matrix with diagonal entries $d^{(i)}$.

The matrices Φ_1 and Φ_2 are defined by

$$\Phi_1 = T_A Diag\left[\phi_1^{(i)}\right] T_A^{-1}, \quad \Phi_2 = T_B Diag\left[\phi_2^{(i)}\right] T_B^{-1}, \tag{21}$$

with

$$\phi_1^{(i)} = \operatorname{sgn}(a^{(i)})\phi^{(i)}, \quad \phi_2^{(i)} = \operatorname{sgn}(b^{(i)})\psi^{(i)} \tag{22}$$

$$\phi^{(i)} = \min\left(1, \frac{\delta y\, |a^{(i)}|}{\delta x\, m(B)}\right), \quad \psi^{(i)} = \min\left(1, \frac{\delta x\, |b^{(i)}|}{\delta y\, m(A)}\right), \tag{23}$$

where $m(A) = \min_i(|a^{(i)}|)$ and $m(B) = \min_i(|b^{(i)}|)$.

4. FULLY DISCRETE SCHEME FOR THE EULER EQUATIONS

On a uniform Cartesian mesh, the proposed scheme is of the form (8) where $\tilde{r}$ is given by (7) and $\tilde{P}$ is a centered difference approximation to the dissipative term P defined by (9) with Φ_1 and Φ_2 given by (21)–(23).

To ensure linear stability and dissipation for the fully discrete scheme, it is important to discretize P properly. Replacing all the first derivatives in (9) by classical centered differences would lead to instability (and also to an overly large scheme stencil). As in the Lax–Wendroff method, it is necessary to use different centered expressions for the "external" and "internal" first derivatives in (9). Moreover, multidimensional dissipation in the sense of Kreiss cannot be achieved by using a rotated Richtmyer (cf. [2]) or a Ni [23] type discretization of P—which comes to define a predictor step at location $(j + \frac{1}{2})\delta x$, $(k + \frac{1}{2})\delta y$. This is why we use a special approximation to P involving a predictor step for each space direction, the stability and dissipation of which were proved in [19] for the Lax–Wendroff method in any number of space dimensions.

To be more specific, the dissipative term is discretized as

$$\tilde{P} = \frac{1}{2}\delta_1\left[\Phi_1\left(\frac{\delta_1 f}{\delta x} + \mu_1\frac{\delta_2\mu_2 g}{\delta y}\right)\right] + \frac{1}{2}\delta_2\left[\Phi_2\left(\mu_2\frac{\delta_1\mu_1 f}{\delta x} + \frac{\delta_2 g}{\delta y}\right)\right], \tag{24}$$

where Φ_1 at cell interface $j+\frac{1}{2}, k$ is computed from (21)–(23) by replacing A and B by their Roe averages $(A_R)_{j+1/2,\,k}$ and $(B_R)_{j,\,k+1/2}$ [27] and similarly computing Φ_2 at cell interface $j, k+\frac{1}{2}$.

It is convenient to express the full scheme in several stages. First a predictor Π_1 at the cell interface $j+\frac{1}{2}, k$ and a predictor Π_2 at the cell interface $j, k+\frac{1}{2}$ are defined. They represent the terms in the brackets of (24). Then, the numerical fluxes are obtained from the expression (7) of $\tilde{r}$ and the predictors. Finally, the new cell values at the cell centers can be deduced as

(a) predictors

$$\begin{aligned}(\Pi_1)^n_{j+\frac{1}{2},\,k} &= \left(\delta_1 f + \frac{\delta x}{\delta y}\delta_2\mu_1\mu_2 g\right)^n_{j+\frac{1}{2},\,k}\\ (\Pi_2)^n_{j,\,k+\frac{1}{2}} &= \left(\frac{\delta y}{\delta x}\delta_1\mu_1\mu_2 f + \delta_2 g\right)^n_{j,\,k+\frac{1}{2}};\end{aligned} \tag{25}$$

(b) numerical fluxes

$$\begin{aligned}(h_1)^n_{j+\frac{1}{2},\,k} &= \left[\left(I + \frac{\theta}{6}\delta_2^2\right)\mu_1 f - \frac{1}{2}\Phi_1\Pi_1\right]^n_{j+\frac{1}{2},\,k}\\ (h_2)^n_{j,\,k+\frac{1}{2}} &= \left[\left(I + \frac{\theta}{6}\delta_1^2\right)\mu_2 g - \frac{1}{2}\Phi_2\Pi_2\right]^n_{j,\,k+\frac{1}{2}};\end{aligned} \tag{26}$$

and

(c) new cell-values

$$w^{n+1}_{j,\,k} = w^n_{j,\,k} - \Delta t\left(\frac{\delta_1 h_1}{\delta x} + \frac{\delta_2 h_2}{\delta y}\right)^n_{j,\,k}, \tag{27}$$

where a parameter θ has been introduced for future reference.

The above 3×3-point scheme is first-order accurate for an unsteady problem. At steady state, it is third-order accurate for $\theta = 1$ and second-order accurate otherwise. For $\theta = 0$, $\tilde{r}$ is approximated by the classical centered difference (4) and one recovers the scheme proposed in [15] from a modification of the Lax–Wendroff scheme.

Note also that for a 1-D problem in the x-direction, we get $\Phi_1 = \text{sgn}(A_R)$ and $\Phi_2 = 0$, so that the numerical flux reduces to

$$h_1 = \mu_1 f - \frac{1}{2}\text{sgn}(A_R)\delta_1 f,$$

which corresponds to the classical Roe scheme [27]. At steady state, the formal order of accuracy of this 1-D scheme is infinite.

The L_2-stability of the scheme (25)–(27) can be studied for the linear fluxes $f = Aw$, $g = Bw$, where A and B are scalar constants. In this case, the amplification factor of the scheme reads

$$G = 1 - (R + iS),$$

where $i^2 = -1$ and

$$\begin{aligned} R &= \Phi_1 \dot{A}(1 - \cos\xi) + \Phi_2 \dot{B}(1 - \cos\eta) + \frac{1}{2}(\Phi_1 \dot{B} + \Phi_2 \dot{A}) \sin\xi \sin\eta \\ S &= \dot{A}\left[1 - \frac{\theta}{3}(1 - \cos\eta)\right] \sin\xi + \dot{B}\left[1 - \frac{\theta}{3}(1 - \cos\xi)\right] \sin\eta, \end{aligned}$$

with

$$\dot{A} = \frac{\Delta t}{\delta x} A, \quad \dot{B} = \frac{\Delta t}{\delta y} B,$$

and ξ, η are the reduced wave numbers in the x and y directions.

The stability domain of the scheme, that is the domain in which $|G| \leq 1$ in terms of $\dot{A}$ and $\dot{B}$, is shown in Figs. 3a and 3b for $\theta = 0$ (second-order) and $\theta = 1$ (third-order), respectively. Note that the stability domain is larger for the third-order version than for the second-order one.

5. FULLY DISCRETE SCHEME FOR THE NAVIER–STOKES EQUATIONS

The Navier–Stokes equations can be written in the form

$$w_t + (f^E - f^V)_x + (g^E - g^V)_y = 0, \tag{28}$$

where $f^E = f^E(w)$ and $g^E = g^E(w)$ stand for the Euler fluxes, and $f^V = f^V(w, w_x, w_y)$ and $g^V = g^V(w, w_x, w_y)$ stand for the viscous fluxes.

System (28) is discretized in the same way as system (1), that is,

$$\frac{1}{\Delta t}(w^{n+1} - w^n) + \tilde{r}^n = \tilde{P}^n, \tag{29}$$

where $\tilde{r}^n$ denotes a centered approximation to the steady-state residual

$$r = (f^E - f^V)_x + (g^E - g^V)_y, \tag{30}$$

and $\tilde{P}^n$ is a centered approximation to

$$P = \frac{\delta x}{2}(\Phi_1 r)_x + \frac{\delta y}{2}(\Phi_2 r)_y. \tag{31}$$

Only the definition of r has changed. Note that since the operator P now contains third-order derivatives, it is no longer purely dissipative.

To simplify the presentation of the fully discrete scheme, we now assume that the viscous fluxes read

$$f^V = \nu^{(1)} w_x + \nu^{(2)} w_y, \quad g^V = \nu^{(3)} w_x + \nu^{(4)} w_y, \tag{32}$$

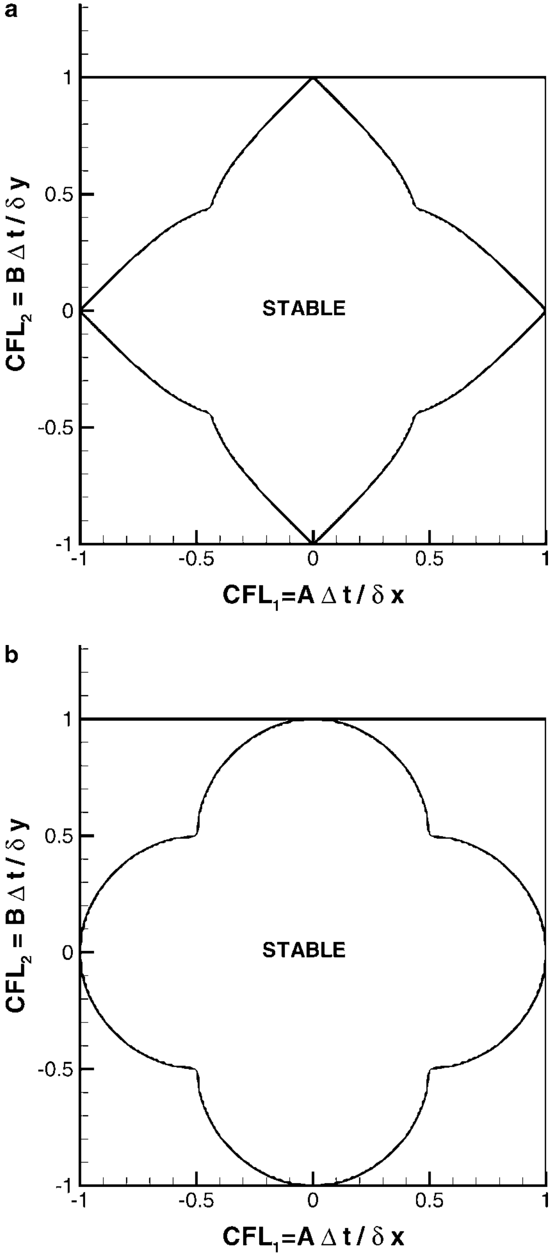

FIG. 3. Advection case. (a) Stability domain of the residual-based second-order scheme ($\theta = 0$). (b) Stability domain of the residual-based third-order scheme ($\theta = 1$).

where $\nu^{(1)}$, $\nu^{(2)}$, $\nu^{(3)}$, and $\nu^{(4)}$ are constant viscosity coefficients ensuring physical dissipation, i.e.,

$$\nu^{(1)} \geq 0, \quad \nu^{(4)} \geq 0, \quad \left(\nu^{(2)} + \nu^{(3)}\right)^2 \leq 2\,\nu^{(1)}\nu^{(4)}.$$

The only work left to do is to revise our compact approximation on the left-hand side of

(29) to include the second derivative. Here, the discrete residual $\tilde{r}$ should be

$$\begin{aligned}\tilde{r} = {}& \frac{1}{\delta x}\delta_1\left[\left(I + \frac{1}{6}\delta_2^2\right)\left(\mu_1 f^E - \nu^{(1)}\frac{\delta_1 w}{\delta x}\right) - \nu^{(2)}\mu_1\mu_2\frac{\delta_2 w}{\delta y} - \frac{1}{12}\nu^{(1)}\frac{\delta_1^3 w}{\delta x}\right] \\ & + \frac{1}{\delta y}\delta_2\left[\left(I + \frac{1}{6}\delta_1^2\right)\left(\mu_2 g^E - \nu^{(4)}\frac{\delta_2 w}{\delta y}\right) - \nu^{(3)}\mu_1\mu_2\frac{\delta_1 w}{\delta x} - \frac{1}{12}\nu^{(4)}\frac{\delta_2^3 w}{\delta y}\right]. \qquad (33)\end{aligned}$$

A Taylor expansion of (33) gives

$$\tilde{r} = r + \tau_e + \tau_a + O(h^4),$$

where r is the exact residual,

$$r = f_x + g_y = \left(f^E - \nu^{(1)}w_x - \nu^{(2)}w_y\right)_x + \left(g^E - \nu^{(3)}w_x - \nu^{(4)}w_y\right)_y;$$

τ_e is the leading term of the truncation error for the classical centered approximation to r,

$$\begin{aligned}\tau_e = {}& \frac{\delta x^2}{6} f^E_{xxx} - \frac{\delta x^2}{12}\left(\nu^{(1)}w_x\right)_{xxx} - \frac{\delta x^2}{6}\left(\nu^{(2)}w_y\right)_{xxx} - \frac{\delta y^2}{6}\left(\nu^{(2)}w_y\right)_{xyy} \\ & + \frac{\delta y^2}{6} g^E_{yyy} - \frac{\delta y^2}{12}\left(\nu^{(4)}w_y\right)_{yyy} - \frac{\delta y^2}{6}\left(\nu^{(3)}w_x\right)_{yyy} - \frac{\delta x^2}{6}\left(\nu^{(3)}w_x\right)_{xxy};\end{aligned}$$

and τ_a is the contribution of the second-order terms *added* to the classical centered formulae in (33) in order to produce derivatives of r in the total discretization error,

$$\tau_a = \frac{\delta y^2}{6}\left(f^E - \nu^{(1)}w_x\right)_{xyy} - \frac{\delta x^2}{12}\left(\nu^{(1)}w_x\right)_{xxx} + \frac{\delta x^2}{6}\left(g^E - \nu^{(4)}w_y\right)_{xxy} - \frac{\delta y^2}{12}\left(\nu^{(4)}w_y\right)_{yyy}.$$

Clearly,

$$\begin{aligned}\tau_e + \tau_a = {}& \frac{\delta x^2}{6}\left(f^E - \nu^{(1)}w_x - \nu^{(2)}w_y\right)_{xxx} + \frac{\delta x^2}{6}\left(g^E - \nu^{(3)}w_x - \nu^{(4)}w_y\right)_{xxy} \\ & + \frac{\delta y^2}{6}\left(f^E - \nu^{(1)}w_x - \nu^{(2)}w_y\right)_{xyy} + \frac{\delta y^2}{6}\left(g^E - \nu^{(3)}w_x - \nu^{(4)}w_y\right)_{yyy} \\ = {}& \frac{\delta x^2}{6} r_{xx} + \frac{\delta y^2}{6} r_{yy},\end{aligned}$$

so that one finally gets

$$\tilde{r} = r + \frac{\delta x^2}{6} r_{xx} + \frac{\delta y^2}{6} r_{yy} + O(h^4), \qquad (34)$$

exactly as in Section 2.

Thus, the fully discrete scheme reads as follows:

(a) predictors:

$$\begin{aligned}(\Pi_1)^n_{j+\frac{1}{2},k} &= \left[\delta_1\left(f^E - \nu^{(1)}\mu_1\frac{\delta_1 w}{\delta x} - \nu^{(2)}\mu_2\frac{\delta_2 w}{\delta y}\right)\right.\\ &\quad \left.+\frac{\delta x}{\delta y}\delta_2\left(\mu_1\mu_2 g^E - \nu^{(3)}\mu_2\frac{\delta_1 w}{\delta x} - \nu^{(4)}\mu_1\frac{\delta_2 w}{\delta y}\right)\right]^n_{j+\frac{1}{2},k}\\ (\Pi_2)^n_{j,k+\frac{1}{2}} &= \left[\frac{\delta y}{\delta x}\delta_1\left(\mu_1\mu_2 f^E - \nu^{(1)}\mu_2\frac{\delta_1 w}{\delta x} - \nu^{(2)}\mu_1\frac{\delta_2 w}{\delta y}\right)\right.\\ &\quad \left.+\delta_2\left(g^E - \nu^{(3)}\mu_1\frac{\delta_1 w}{\delta x} - \nu^{(4)}\mu_2\frac{\delta_2 w}{\delta y}\right)\right]^n_{j,k+\frac{1}{2}};\end{aligned} \tag{35}$$

(b) numerical fluxes:

$$\begin{aligned}(h_1)^n_{j+\frac{1}{2},k} &= \left[\left(I+\frac{\theta}{6}\delta_2^2\right)\left(\mu_1 f^E - \nu^{(1)}\frac{\delta_1 w}{\delta x}\right)\right.\\ &\quad \left.-\nu^{(2)}\mu_1\mu_2\frac{\delta_2 w}{\delta y} - \frac{\theta}{12}\nu^{(1)}\frac{\delta_1^3 w}{\delta x} - \frac{1}{2}\Phi_1\Pi_1\right]^n_{j+\frac{1}{2},k}\\ (h_2)^n_{j,k+\frac{1}{2}} &= \left[\left(I+\frac{\theta}{6}\delta_1^2\right)\left(\mu_2 g^E - \nu^{(4)}\frac{\delta_2 w}{\delta y}\right)\right.\\ &\quad \left.-\nu^{(3)}\mu_1\mu_2\frac{\delta_1 w}{\delta x} - \frac{\theta}{12}\nu^{(4)}\frac{\delta_2^3 w}{\delta y} - \frac{1}{2}\Phi_2\Pi_2\right]^n_{j,k+\frac{1}{2}};\end{aligned} \tag{36}$$

(c) new cell values:

$$w^{n+1}_{j,k} = w^n_{j,k} - \Delta t\left(\frac{\delta_1 h_1}{\delta x} + \frac{\delta_2 h_2}{\delta y}\right)^n_{j,k}. \tag{37}$$

At steady state, the above scheme is third-order accurate for $\theta = 1$ and second-order accurate otherwise. For $\theta = 0$, we recover the second-order scheme proposed in [7].

Remark.

(a) The second- and third-order schemes share the same $(3 \times 3) + 4 = 13$ point stencil. This stencil is sufficient to approximate the corrective terms specific to the third-order scheme for the Navier–Stokes equations.

(b) For viscous flow problems, the functions Φ_1 and Φ_2 could take into account the physical viscosity in order to reduce the numerical dissipation in viscous layers. In the present work however, we retain for viscous applications the functions Φ_1, Φ_2 defined for the Euler equations: this choice has no effect on the order of accuracy.

The stability domain of scheme (35)–(37) is displayed in Figs. 4a and 4b, for a choice of fluxes corresponding to a model boundary layer problem: $f^E = Aw, g^E = 0$, $f^V = 0, g^V = \nu w_y$. Note that the present explicit scheme of third-order accuracy has a large stability domain.

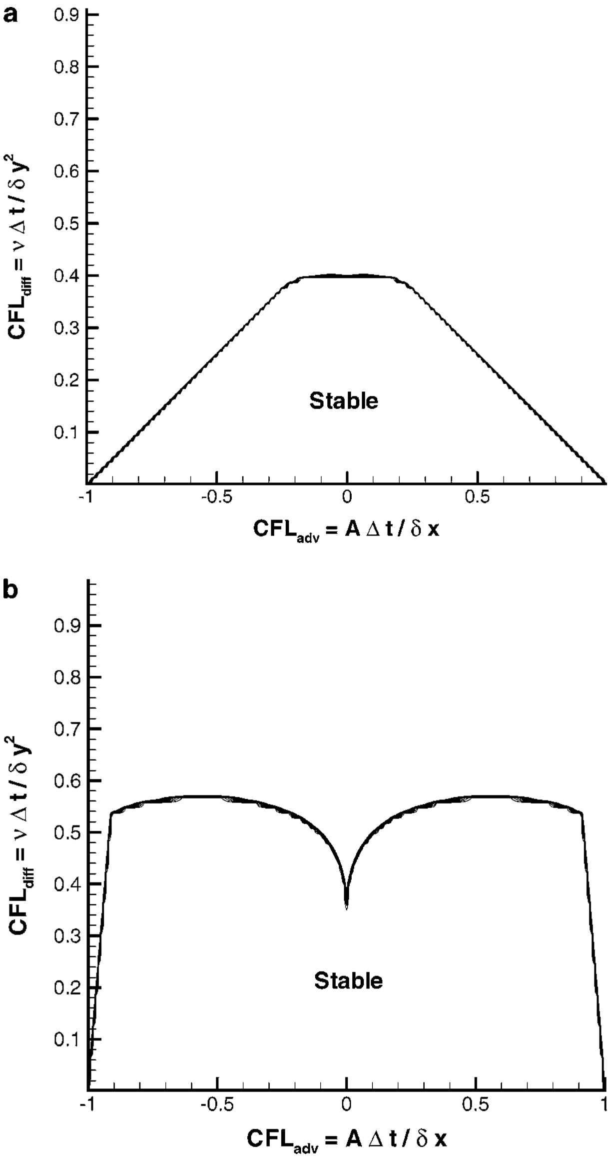

FIG. 4. Advection-diffusion case. (a) Stability domain of the residual-based second-order scheme ($\theta = 0$). (b) Stability domain of the residual-based third-order scheme ($\theta = 1$).

6. IMPLICIT VERSION OF THE SCHEME

The efficiency of steady-state calculations may be increased by adding a suitable implicit stage to the previously presented scheme. In this section, we shortly present the implicit version of the scheme. Let us first consider the case of the Euler equations. Applying the Euler backward time-discretization to the explicit scheme (25)–(27), linearizing the numerical fluxes taken at time $(n + 1)\Delta t$, and simplifying the resulting linear implicit operator to retain only derivatives in the x and y directions that can be discretized on a three-point stencil (implicit mixed second-order derivatives coming from the dissipative term are dropped for both the second- and third-order schemes and so are the implicit mixed third-order derivatives specific to the third-order scheme) yields an implicit stage of the Harten type [12] (with $\Phi|A|$, $\Psi|B|$ in the second-order terms instead of $|A|$, $|B|$ for classical Roe-type schemes),

$$\Delta w_{j,k} + \frac{\Delta t}{\delta x}\left[\mu_1(A\delta_1\Delta w^n)_{j,k} - \frac{1}{2}\delta_1(\Phi|A|\delta_1\Delta w^n)_{j,k}\right] + \frac{\Delta t}{\delta y}\left[\mu_2(B\delta_2\Delta w^n)_{j,k} - \frac{1}{2}\delta_2(\Psi|B|\delta_2\Delta w^n)_{j,k}\right] = \Delta w^{expl}_{j,k}, \tag{38}$$

where $\Delta w^n_{j,k} = w^{n+1}_{j,k} - w^n_{j,k}$, and $\Delta w^{expl}_{j,k}$ is the similar time increment given by the explicit stage. From now on, we will refer to the second- (respectively third) order scheme as scheme (38) with $\theta = 0$ (respectively $\theta = 1$) in the explicit stage (25)–(27); strictly speaking, second- or third-order accuracy is reached only at steady state.

The implicit scheme (38) is linearly stable for $\theta = 0$ and $\theta = 1$ and offers good damping properties when used with large CFL numbers. The amplification factor of the scheme is displayed in Figs. 5a and 5b in the case of a model advection problem, $w_t + aw_x + bw_y = 0$, with a and b fixed at constant values. The second-order scheme leads to better damping

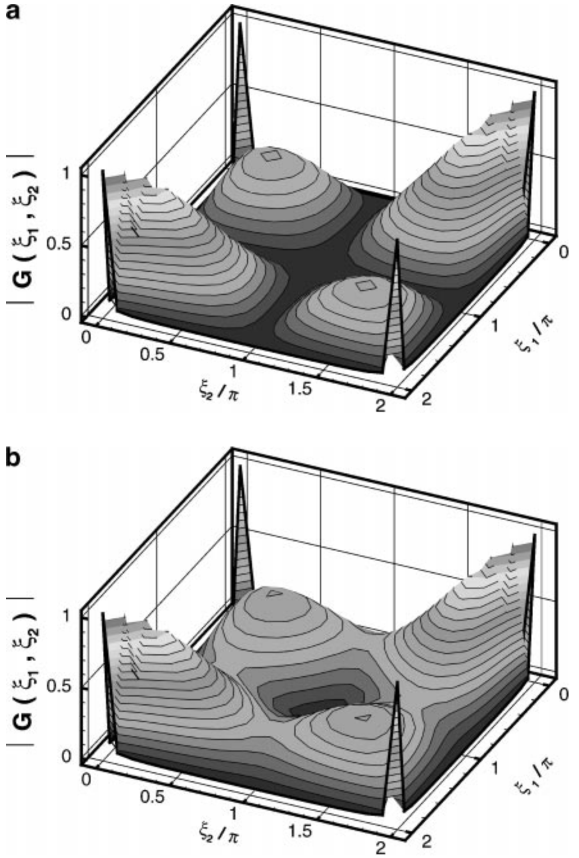

FIG. 5. Isovalues of the amplification factor of scheme (38) in the reduced wave-numbers plane. 2-D advection problem: $a\Delta t/\delta x = b\Delta t/\delta y = 100$. (a) Second-order scheme ($\theta = 0$). (b) Third-order scheme ($\theta = 1$).

than the third-order scheme, which was expected since the third-order explicit stage contains third derivatives which were not made implicit in order to keep the scheme simple. However, thanks to the method's compactness, the simplifications performed on the implicit stage do not significantly degrade the efficiency of the scheme, which would not be the case if a third-order MUSCL-type explicit stage were coupled with the implicit stage retained in (38) (see [26] for instance). For practical applications, it is usually not possible to directly solve the implicit stage (38) due to CPU cost and memory storage issues, and approximate solution techniques such as alternate-line relaxation of the Jacobi or Gauss-Seidel type are typically used instead. Tools to analyze the stability and efficiency of relaxed implicit schemes were presented in [6]. For a 2-D model advection problem, it can be shown, using these tools, that scheme (38) is linearly unconditionally stable when its implicit stage is solved using an alternate-line Gauss–Seidel relaxation; moreover, only a few inner iterations of the relaxation process (typically one or two) are required to ensure a damping very close to that of the direct solver.

In the case of the Navier–Stokes equations, an implicit stage is built from the explicit scheme (35)–(37) following the same lines as in the Euler case: only derivatives in the x and y directions that can be approached using three-point formulae are retained in the implicit stage in order to keep it simple. This choice leads to the following implicit scheme:

$$\Delta w_{j,k} + \frac{\Delta t}{\delta x}\left[\mu_1(A\delta_1\Delta w^n)_{j,k} - \delta_1\left(\left(\frac{1}{2}\Phi|A| + \frac{A^V}{\delta x}\right)\delta_1\Delta w^n\right)_{j,k}\right] + \frac{\Delta t}{\delta y}\left[\mu_2(B\delta_2\Delta w^n)_{j,k} - \delta_2\left(\left(\frac{1}{2}\Psi|B| + \frac{B^V}{\delta y}\right)\delta_2\Delta w^n\right)_{j,k}\right] = \Delta w^{expl}_{j,k}, \qquad (39)$$

with $A^V = \partial f^V(w, w_x, w_y)/\partial w_x$, $B^V = \partial g^V(w, w_x, w_y)/\partial w_y$.

The implicit scheme (39) is linearly stable for $\theta = 0$ and $\theta = 1$. The efficiency of the scheme was studied for a 2-D linear advection–diffusion problem, $w_t + aw_x = \nu w_{yy}$, with the following choice of numerical parameters: $\Delta x = \Delta y = h$, $\frac{a\Delta t}{h} = CFL \min(1, \frac{\mathrm{Re}_h}{2})$, $\frac{\nu\Delta t}{h^2} = CFL \min(\frac{1}{2}, \frac{1}{\mathrm{Re}_h})$, where $\mathrm{Re}_h = \frac{ah}{\nu}$ denotes the cell Reynolds number. It was observed that the second-order scheme offers a better damping than the third-order scheme, which was expected since a greater number of simplifications was performed on this latter when building the implicit stage. However, in spite of these numerous simplifications, the third-order scheme offers a reasonably good damping at large time steps as soon as the cell Reynolds number is not too small, which is indeed the case for large regions of flows in practical problems (see Figs. 6a–6d). The analysis of a relaxed version of the implicit stage (39) led to the same conclusion as in the Euler case: a single iteration of an alternate Gauss–Seidel relaxation technique is sufficient to provide an efficiency very close to that of a direct solution of (39). This relaxation technique will be applied for all the computations presented in the following sections.

7. VALIDATIONS FOR MODEL PROBLEMS

The theoretical properties of the residual-based scheme are studied in this section through the calculation of several model problems, the exact steady solutions of which are known, allowing the computation of the actual accuracy orders and a detailed assessment of the extra precision provided by the third-order method.

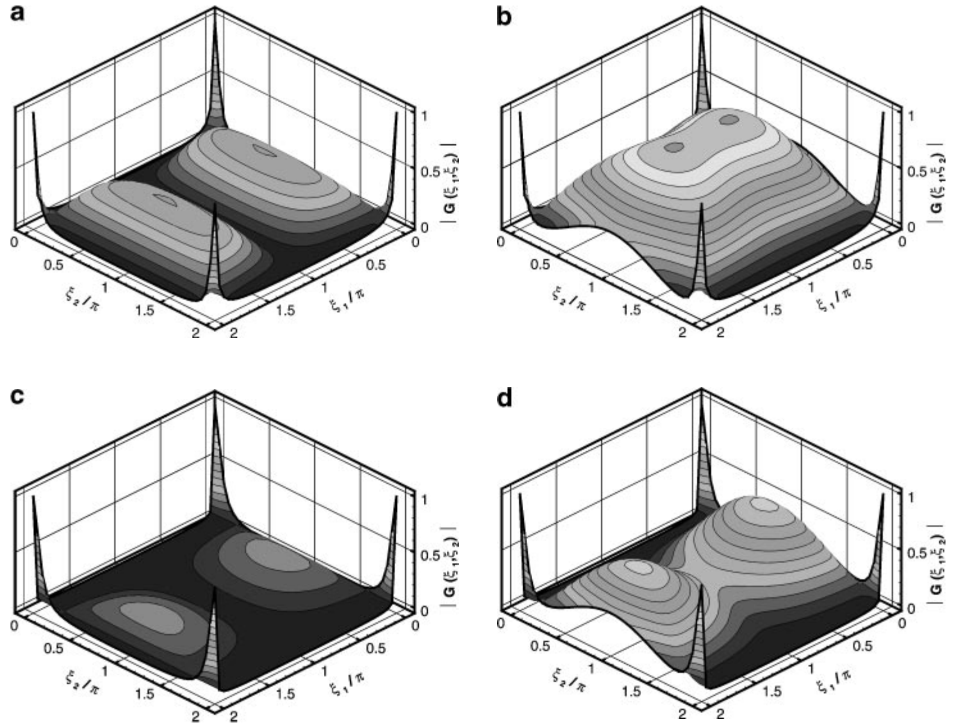

FIG. 6. Isovalues of the amplification factor of scheme (39) in the reduced wave-numbers plane. 2-D advection–diffusion problem. (a) $CFL = 100$, $\mathrm{Re}_h = 0.5$. Second-order scheme ($\theta = 0$). (b) $CFL = 100$, $\mathrm{Re}_h = 0.5$. Third-order scheme ($\theta = 1$). (c) $CFL = 100$, $\mathrm{Re}_h = 5$. Second-order scheme ($\theta = 0$). (d) $CFL = 100$, $\mathrm{Re}_h = 5$. Third-order scheme ($\theta = 1$).

7.1. *Rotational Advection*

We consider the linear advection with rotation around the point ($x = 1$, $y = 0$) of a smooth Gaussian profile over the square domain $[0, 1] \times [0, 1]$, similar to [24]. More precisely, we look for the steady solution of the following initial boundary-value problem:

$$\begin{cases} \frac{\partial w}{\partial t} + y\frac{\partial w}{\partial x} + (1 - x)\frac{\partial w}{\partial y} = 0, & 0 < x < 1, \quad 0 < y < 1; \\ w(x, y, 0) = 0, & 0 < x < 1, \quad 0 < y < 1; \\ w(x, 0, t) = \exp\left(-50\left(x - \frac{1}{2}\right)^2\right), & 0 \le x \le 1, \quad t \ge 0; \\ w(x, 1, t) = 0, & 0 \le x \le 1, \quad t \ge 0; \\ w(0, y, t) = 0, & 0 \le y \le 1, \quad t \ge 0. \end{cases} \tag{40}$$

On the boundary part ($x = 1$, $0 \le y \le 1$), no condition is required (outflow boundary). Numerically, the boundary values are obtained from extrapolation of the interior values. The exact steady solution is $w = const$ on any circle of center $(1, 0)$. The numerical solutions obtained with the second-order scheme (defined by formulae (25)–(27), (38) with $\theta = 0$) and the third-order scheme (defined by the same formulae with $\theta = 1$) using a $(J = 81) \times (K = 81)$ uniform Cartesian mesh are displayed in Fig. 7a, along with the exact steady solution. The iso-lines of the numerical solution are nearly perfect circles and cannot be distinguished from those of the exact solution. The numerical solution along the mesh diagonal $y = 1 - x$

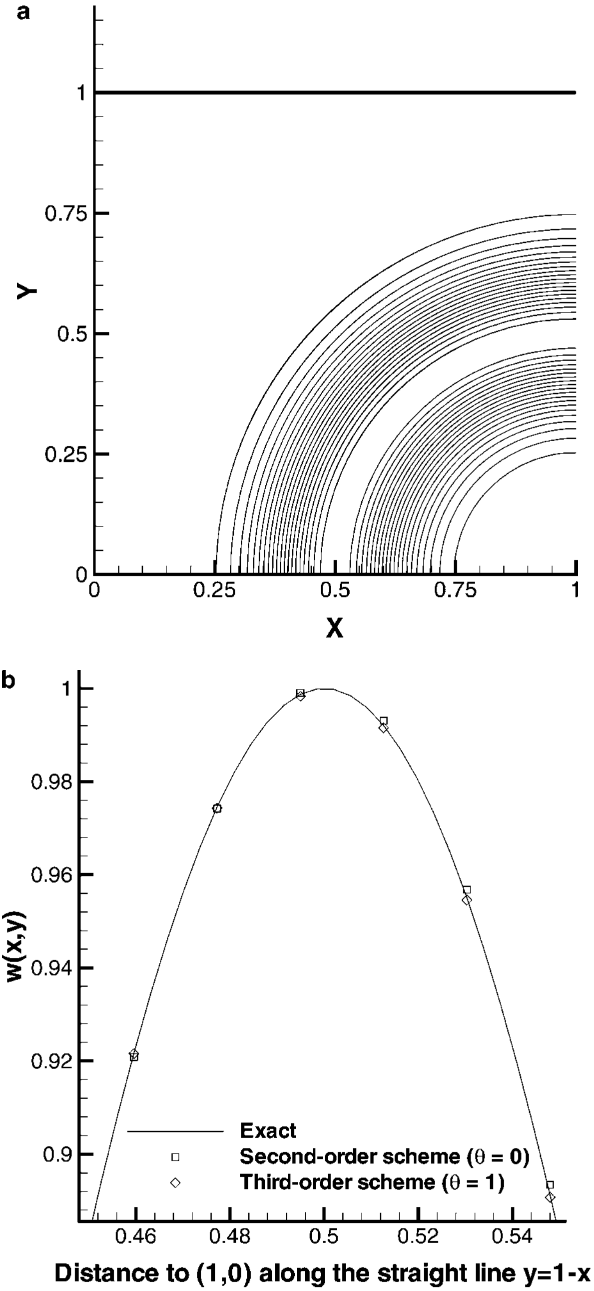

FIG. 7. Rotational advection. (a) Isovalues of $w(x, y)$ on an 80×80 grid. (b) Close-up of the distribution of $w(x, y)$ along the diagonal $y = 1 - x$.

is also presented in Fig. 7b and allows us to observe the effective higher accuracy of the third-order method with respect to the second-order one. More precisely, the L_2-norm errors between the exact solution and the numerical solutions corresponding to various mesh sizes ($\Delta x = \frac{1}{(J-1)} = \Delta y = \frac{1}{(K-1)}$) are plotted in Fig. 8: the second-order accuracy of the $\theta = 0$ scheme and the third-order accuracy of the $\theta = 1$ scheme are clearly demonstrated. The

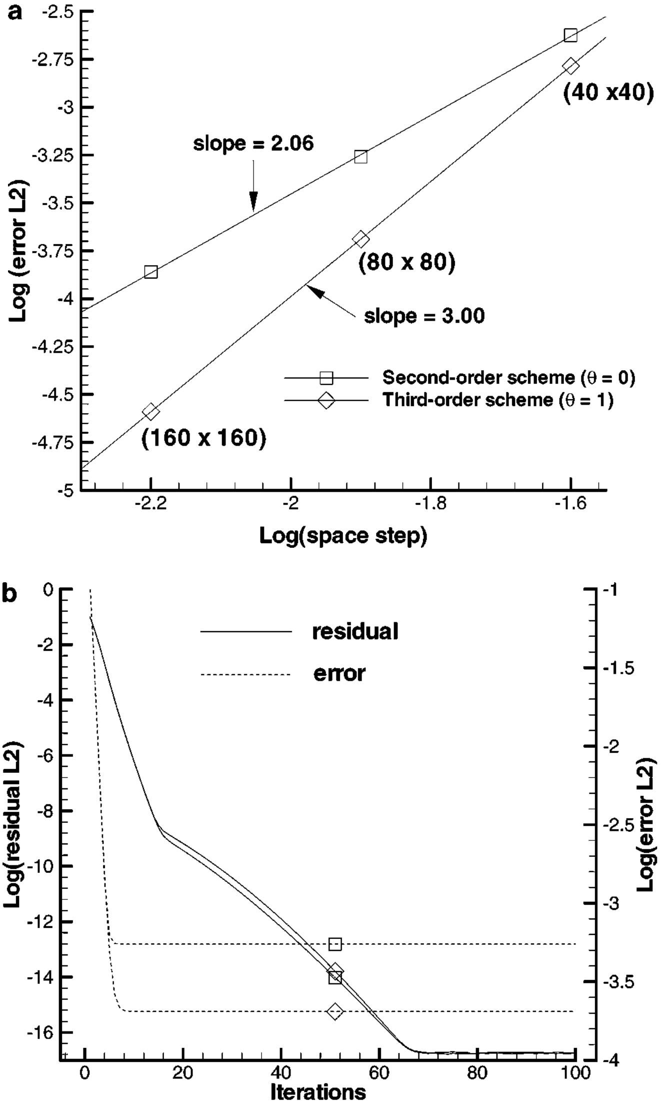

FIG. 8. Rotational advection. (a) Computed orders of accuracy. (b) Convergence and error history on an 80 × 80 grid.

convergence history of the second- and third-order computations is monitored through the L_2-norm of $\frac{\Delta w}{\Delta t}$ on the computational domain and displayed in Figs. 8a and 8b. For both schemes, the implicit stage is solved using a single iteration of an alternate-line symmetric Gauss–Seidel relaxation and a CFL number equal to 100 (larger CFL numbers can be used but an asymptotic behavior is reached), yielding fast convergence to steady state. Now, an important concern is the following: since the present schemes are higher order only at steady state, *what is the effect of stopping at some level of the residual on the actual accuracy of the solution*? This can be checked in Fig. 8. On the 80 × 80 grid, one observes that after 10 time iterations only, the residual is equal to 10^{-4}, while the actual error has already converged to its steady state value, i.e., third-order accuracy is really achieved. Due to this error-convergence efficiency, the residual-based approach appears to be working concept.

7.2. *1-D Boundary-Layer Model*

We now consider the following 1-D viscous flow model problem:

$$\begin{cases} w_t + a w_x = \nu w_{xx}, & 0 < x < L; \\ w(x,0) = 0, & 0 < x < L; \\ w(0,t) = 0, & t \geq 0; \\ w(L,t) = 1, & t \geq 0. \end{cases} \tag{41}$$

The exact steady solution of (41) on a grid $(j\Delta x)_{0\leq j\leq J}$ is given by $w_j = (1-\exp(j\mathrm{Re}_\delta))/(1-\exp(\mathrm{Re}))$, where Re denotes the Reynolds number, $\mathrm{Re} = aL/\nu$, and $\mathrm{Re}_\delta = a\delta x/\nu = \mathrm{Re}/J$ denotes the cell Reynolds number. In the present computations, $a = 1$, $\nu = 0.1$, and $L = 1$, so that $\mathrm{Re} = 10$ and $\mathrm{Re}_\delta = 10/J$. The numerical solutions obtained with the 1-D second-order scheme (deduced from formulae (35)–(37), (39) with $\theta = 0$) and the third-order scheme (defined by the same formulae with $\theta = 1$) using a uniform grid of 20 cells ($\mathrm{Re}_\delta = 1/2$) are presented in Fig. 9a, along with the above exact steady solution they approach closely. A closeup of these solutions (Fig. 9b) shows the higher precision of the third-order scheme. The actual orders of accuracy of the schemes, based on several computations using different grid refinements, are plotted in Fig. 10a: they are close to the theoretical second and third orders. The convergence and error history of the schemes is displayed in Fig. 10b: third-order accuracy for this viscous problem is obtained after very few iterations, prior to full convergence of the residual to machine accuracy steady state.

7.3. *2-D Poiseuille Flow Model*

We are now looking for an approximate steady solution of the following problem:

$$\begin{cases} w_t + a w_x = \nu w_{yy}, & 0 < x < L, \quad 0 < y < L; \\ w(x,y,0) = 1, & 0 < x < L, \quad 0 < y < L; \\ w(x,0,t) = w(x,L,t) = 0, & 0 < x < L, \quad t \geq 0; \\ w(0,y,t) = w_0(y), & 0 < y < L, \quad t \geq 0. \end{cases} \tag{42}$$

At steady state, problem (42) may be viewed as modelling a Poiseuille flow: a profile, prescribed at inflow $y = 0$, is advected and diffused between two solid walls on which its value is fixed to zero (see Fig. 11a). The inflow is chosen equal to $w_0(y) = \sin(\pi\frac{y}{L})$, so that the exact steady solution of (42) is given by $w(x,y) = \exp(-\frac{\pi^2}{\mathrm{Re}}\frac{x}{L})\sin(\pi\frac{y}{L})$, where $\mathrm{Re} = aL/\nu$ denotes the Reynolds number, equal to 10 in our case ($a = 1$, $\nu = 0.1$, $L = 1$). The w-profiles along the outflow boundary $x = 1$ given by scheme (35)–(37), (39) on a uniform Cartesian mesh of $(J = 20) \times (K = 20)$ points are displayed in Fig. 11b along with the above exact solution. Once again the excellent accuracy of the third-order scheme is observed. The actual orders of accuracy of the schemes are plotted in Fig. 12a: they are in very good agreement with the theoretical orders. The convergence and error history of the schemes is displayed in Fig. 12b: the slower convergence of the third-order scheme is in accordance with the amplification factor analysis in Section 6. Note that the slower decay of the residual only slightly affects the convergence on the error, which reaches a steady state anyway after few iterations. Computations performed on the same grid with a smaller diffusion coefficient ν, i.e., a larger cell Reynolds number, led to a much smaller difference

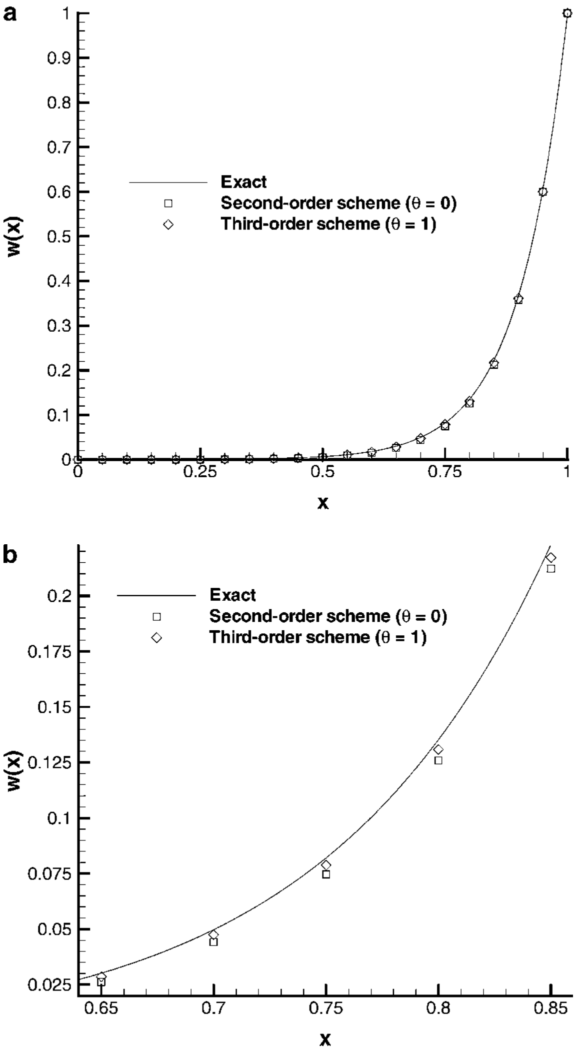

FIG. 9. 1-D boundary-layer model. (a) Exact and numerical solutions on a grid of 20 cells. (b) Closeup of the solutions.

of convergence rate between the second- and third-order schemes, as predicted in Section 6 from the amplification factor study.

7.4. *Inviscid Bürgers Model*

Here, we simply check the shock-capturing property of the scheme on the simple nonlinear problem:

$$\begin{cases} w_t + (w^2/2)_x = 0, & 0 < x < 1; \\ w(x,0) = 1 - 2x, & 0 \leq x \leq 1; \\ w(0,t) = 1, & t > 0; \\ w(1,t) = -1, & t > 0. \end{cases} \tag{43}$$

The exact steady solution of (43) is a shock between $w = 1$ and $w = -1$. The numerical

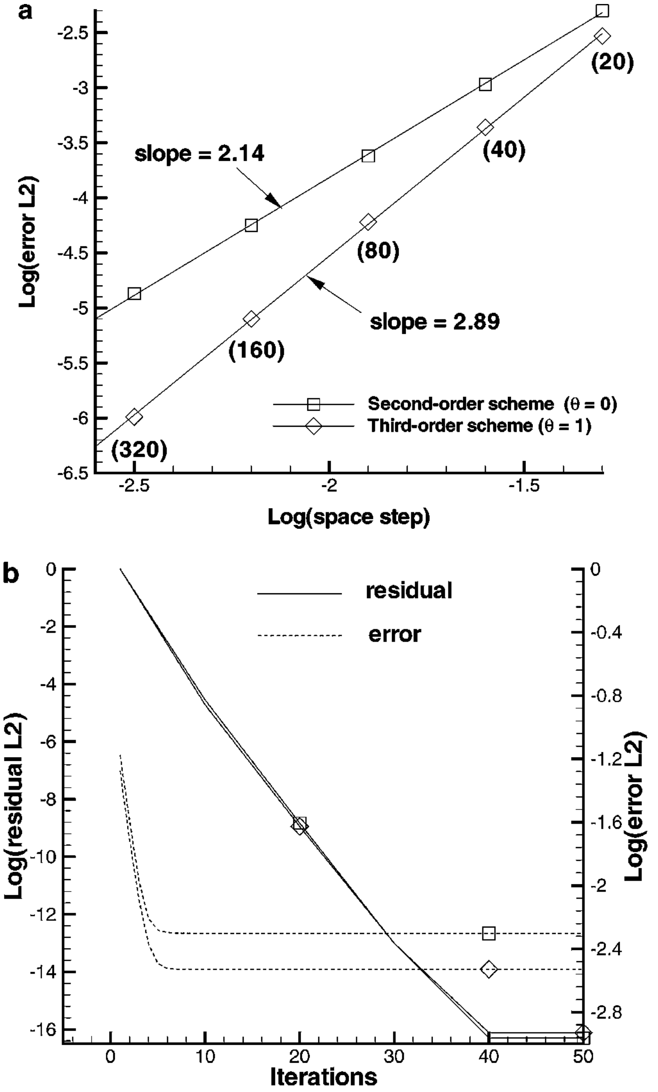

FIG. 10. 1-D boundary-layer model. (a) Computed orders of accuracy. (b) Convergence and error history on a 20-point grid ($CFL = 100$).

solution obtained with the third-order RBC scheme using $\delta x = 1/20$ is shown in Fig. 13: the shock profile is spread on two mesh cells without oscillation. At CFL $= 100$, the residual is 10^{-6} after 20 time iterations and it reaches machine-zero in about 40 time iterations.

The inviscid and viscous results obtained in this section for model problems validate the new design concept of the schemes presented in this paper: residual-based compactness and dissipation lead to efficient and high-accuracy methods for multidimensional inviscid and viscous problems.

8. APPLICATIONS TO VISCOUS COMPRESSIBLE FLOW PROBLEMS

In this section, some viscous flow computations performed with the proposed residual-based scheme are presented. For these first applications, the extension of the scheme to

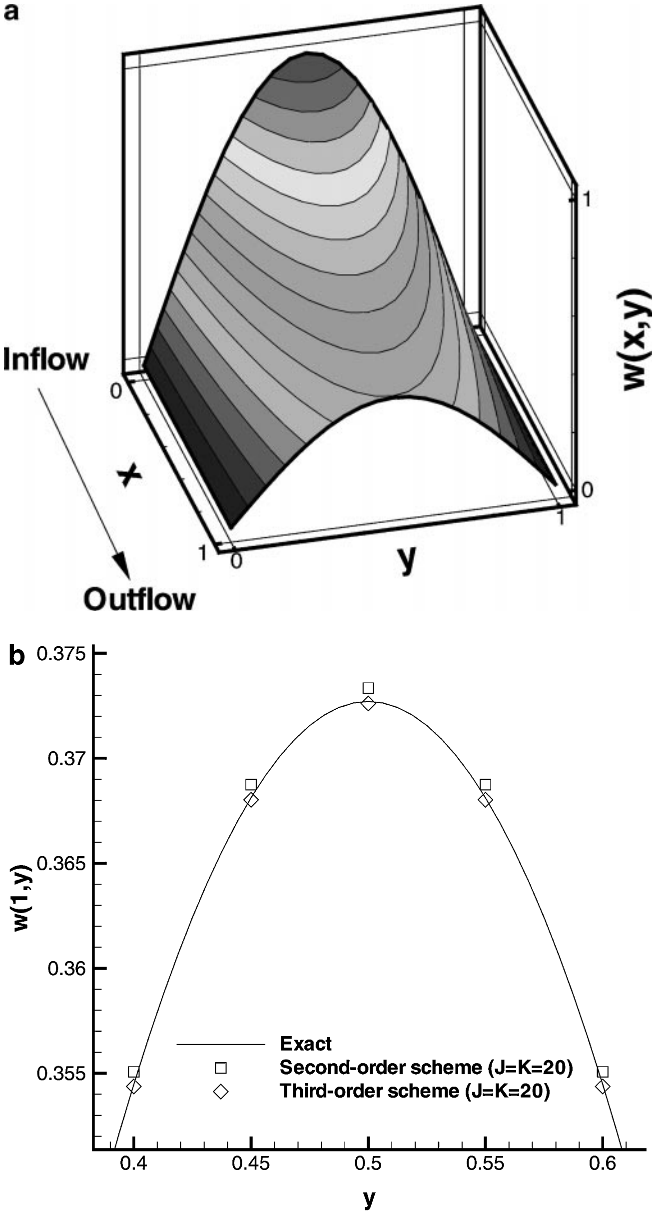

FIG. 11. 2-D Poiseuille flow model. (a) Isovalues of $w(x, y)$. (b) Closeup on $w(x, y)$ along the outflow boundary.

curvilinear meshes was realized in a straightforward finite-volume formulation that does not necessarily guarantee a true second- or third-order accuracy on nonregular meshes. We are not able to compute the real order of accuracy anyway-since the exact solutions are here unknown. The present analysis aims rather at assessing the ability of the third-order scheme ($\theta = 1$) to provide accurate solutions for realistic applications and whether to use it instead of the second-order scheme ($\theta = 0$) which has already been successfully applied to the computation of viscous flows over airfoils and wings in previous studies [5, 7].

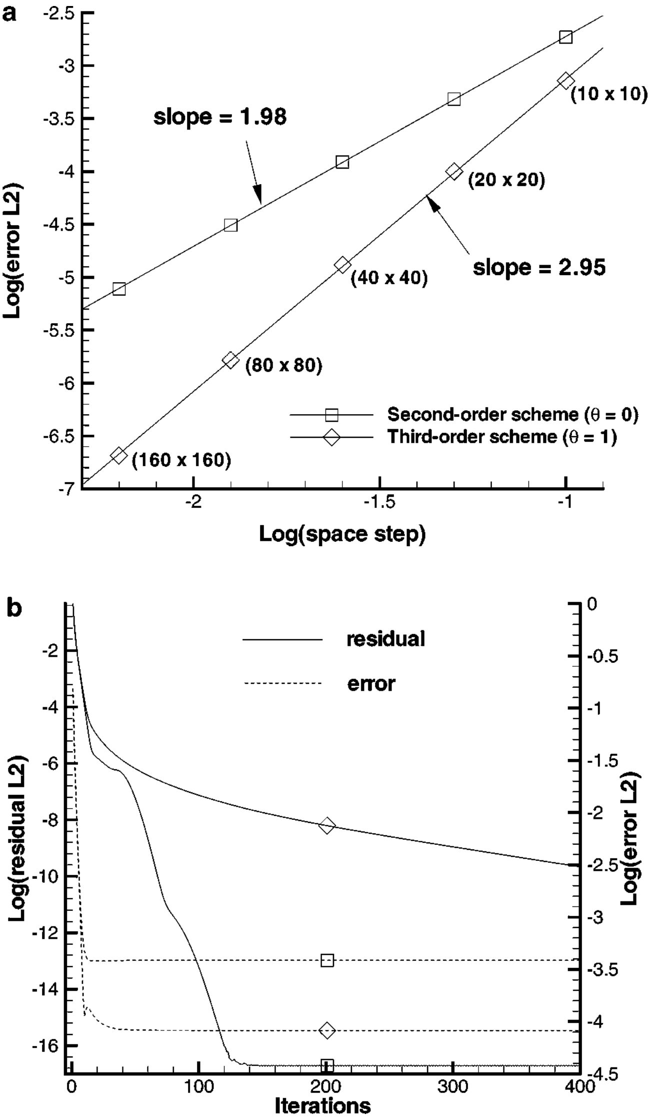

FIG. 12. 2-D Poiseuille flow model. (a) Computed orders of accuracy. (b) Convergence and error history on a 20 × 20 grid ($CFL = 100$).

8.1. *Separated Laminar Flow over a NACA0012 Airfoil*

The laminar flow at $M_\infty = 0.5$ and $\mathrm{Re} = 5000$ over a thermally insulated NACA0012 airfoil at zero incidence is computed using a series of C meshes containing the same number of points (150 × 64)—only the upper part of the airfoil is considered owing to the symmetry of the problem—but with different mesh spacing at the wall, varying from 10^{-4} chord for the finest mesh up to 10^{-2} chord for the coarsest mesh. For this problem, the main flow feature is a separation region occurring near the airfoil trailing edge (see the Mach contours displayed in Fig. 14, where the recirculation bubble appears as a region of low Mach number flow). We first checked that the results provided by the residual-based schemes compared well with other published results. The pressure and viscous drag coefficients as well as the

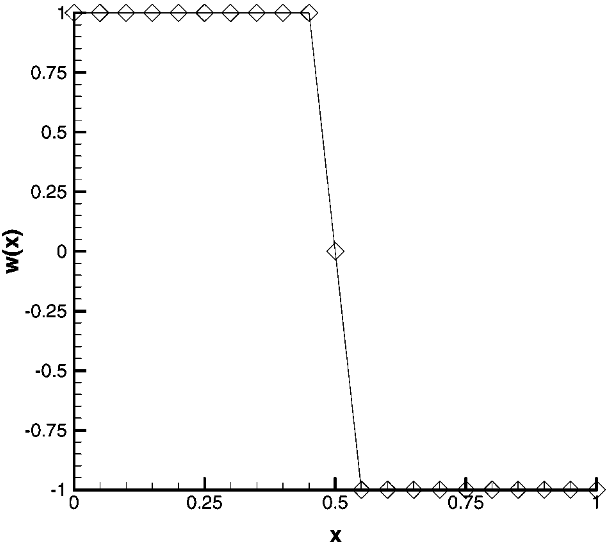

FIG. 13. 1-D Bürgers problem. Typical numerical shock structure.

location of the separation point computed using the third-order scheme on the finest grid are indeed close to the values obtained in [8, 22] on grids with similar refinement at the wall (see Table I). Note also that on this finest mesh the second- and third-order schemes yield almost the same results so that using the third-order scheme does not seem especially attractive.

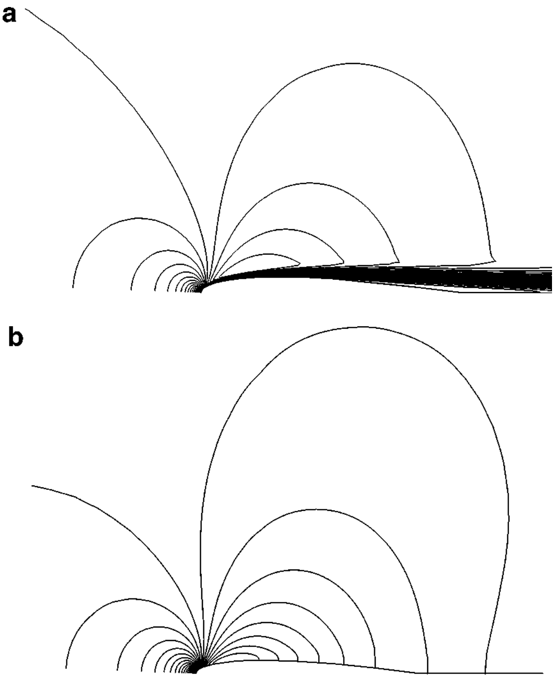

FIG. 14. Laminar flow over a NACA0012 airfoil. Third-order computation on finest mesh. (a) Mach contours. (b) Pressure contours.

TABLE I
Viscous and Inviscid Drag Coefficients and Separation Location (Chord %) for NACA0012 Airfoil[a]

Method	CD_f	CD_p	x_s
Present 2nd-order scheme	0.03235	0.02271	80.3
Present 3rd-order scheme	0.03230	0.02277	80.7
[8]	0.0325	0.0226	81.9
[22]	0.0332	0.0229	81.4

[a] ($M = 0.5$, Re $= 5000$, $\alpha = 0$ deg).

However, let us now investigate the effect of grid coarsening on the second- and third-order numerical solutions. All the computations were performed using a relaxed version of the implicit schemes and $CFL = 100$, which provided a fast convergence to steady state in all cases, except for the second-order scheme on the coarsest grid, which did not converge properly because of a lack of spatial resolution in the separation region; decreasing the CFL number did not allow the second-order scheme to converge on this grid, for which the mesh size at the wall leads to less than 10 points in the boundary layer. On the same grid, the third-order scheme reached a machine accuracy steady state in 150 iterations. It is clear from Figs. 15a and 15b that the third-order scheme is much less sensitive to grid coarsening than the second-order scheme: the shift on the viscous drag prediction when multiplying the mesh size at the wall by 75 (both schemes reach steady state for a mesh size at the wall between $1 \times 10^{-4}c$ (or below) and $7.5 \times 10^{-3}c$) is about 12.5% for the third-order scheme and 25% for the second-order scheme; the shift on the inviscid drag does not exceed 2% for the third-order scheme while it is about 30% for the second-order scheme. The low sensitivity of the third-order scheme to grid coarsening is also depicted in Figs. 16a and 16b where the computed skin-friction is plotted for different grid sizes. The coarsening tends to lower the skin-friction peak located near the leading edge for both schemes but with a slower decrease for the third-order scheme. More importantly, the coarsening has only a weak influence on the third-order prediction outside the leading-edge region: the separation point, identified as the point where skin-friction becomes zero, shifts only slightly, even on very coarse grids. Meanwhile, the second-order prediction is much more affected by the coarsening on the whole airfoil length and eventually fails to converge.

8.2. *2-D Shock/Boundary Layer Interaction*

In order to assess the robustness of the residual-based scheme, the interaction of an oblique shock wave with a laminar boundary layer is computed. The test-case considered has been experimentally and numerically studied by Degrez *et al.* [9], and numerical simulations were performed by other authors [16]. The freestream Mach number is 2.2; the Reynolds number based on the distance X_{sh} between the plate leading edge and the shock impingement point is 9.8645×10^4; the shock angle with respect to the horizontal is -30.027^o (see [9] for more details). The computational grid is the one used in [16]: it contains 74 points uniformly distributed in the streamwise direction, between $x = -0.18$ and $x = 2.01$ (the plate leading edge is located at $x = 0$), and 62 points in the direction normal to the plate, with a mesh size at the wall about 1.5×10^{-4}. The CFL number is equal to 10.

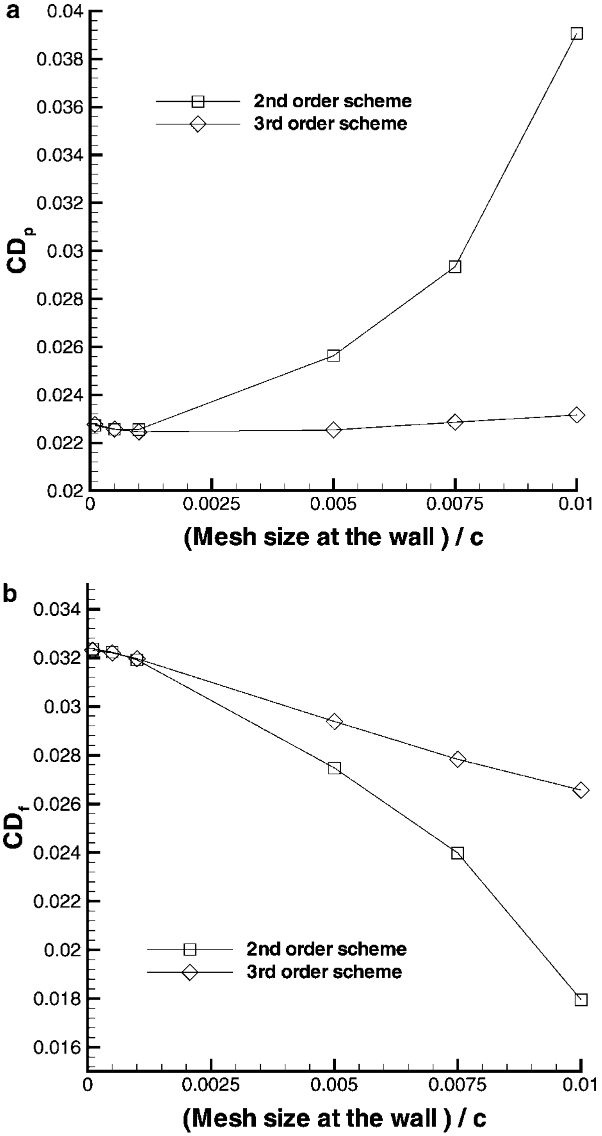

FIG. 15. Laminar flow over a NACA0012 airfoil. (a) Effect of mesh coarsening on the computed pressure drag. (b) Effect of mesh coarsening on the computed skin-friction drag.

The separation and reattachment location of the boundary layer may be identified on the computed skin-friction distribution (see Fig. 17a) as those points where C_f changes sign. Their computed values agree reasonably well with the experiment and compare favourably with the values obtained in [16] using a third-order-biased upwind total variation diminishing (TVD) scheme on the same Cartesian mesh (see Table II). The pressure distributions along the plate are displayed in Fig. 17b: to allow a comparison with [9] the pressures have been normalized by the minimum pressure p_0 just upstream of the interaction. The computations compare well with the experiment but the predicted pressure gradient at reattachment is too steep with respect to this experiment. However, the third-order distribution more closely approaches the distribution measured in [9].

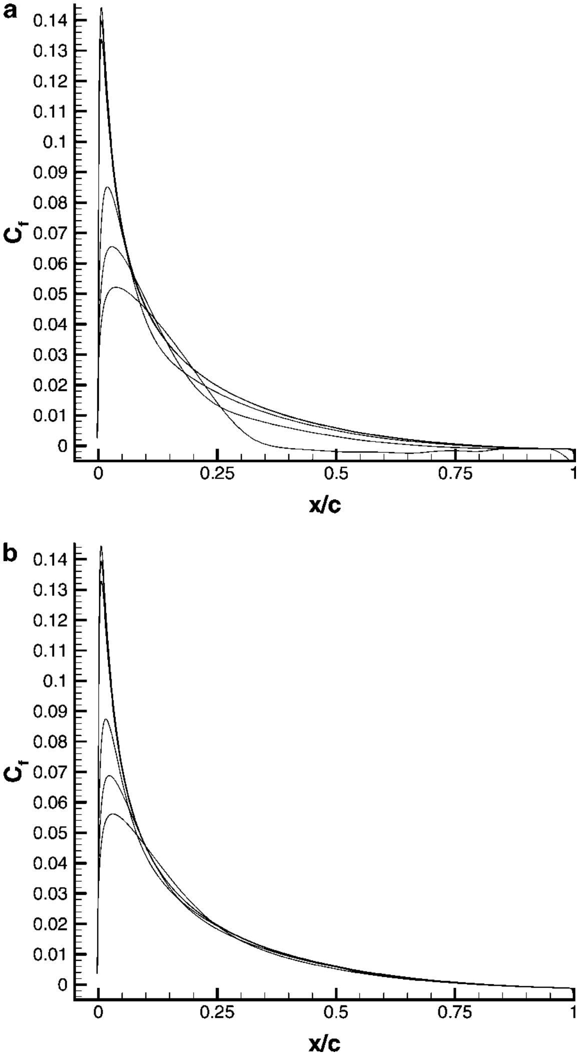

FIG. 16. Laminar flow over a NACA0012 airfoil. Effect of mesh coarsening on the computed skin-friction distribution. (a) Second-order scheme. (b) Third-order scheme.

8.3. *Turbulent Transonic Flow over a RAE 2822 Airfoil*

The last problem considered is the turbulent flow over a RAE2822 airfoil, Case 9 of Cook *et al.* [4], assumed to correspond to $\mathrm{Re} = 6.5 \times 10^6$, $M_\infty = 0.734$, and $\alpha = 2.54$ deg, following the wind tunnel corrections suggested in [10]. The two-layer algebraic eddy

TABLE II

Separation and Reattachment in Shock/Boundary Layer Interaction

	Separation (X/X_{sh})	Reattachment (X/X_{sh})
Present 2nd-order scheme	0.78	1.22
Present 3rd-order scheme	0.76	1.22
[16]	0.745	1.39
Computation [9]	0.79	1.24
Experiment [9]	0.78	1.28

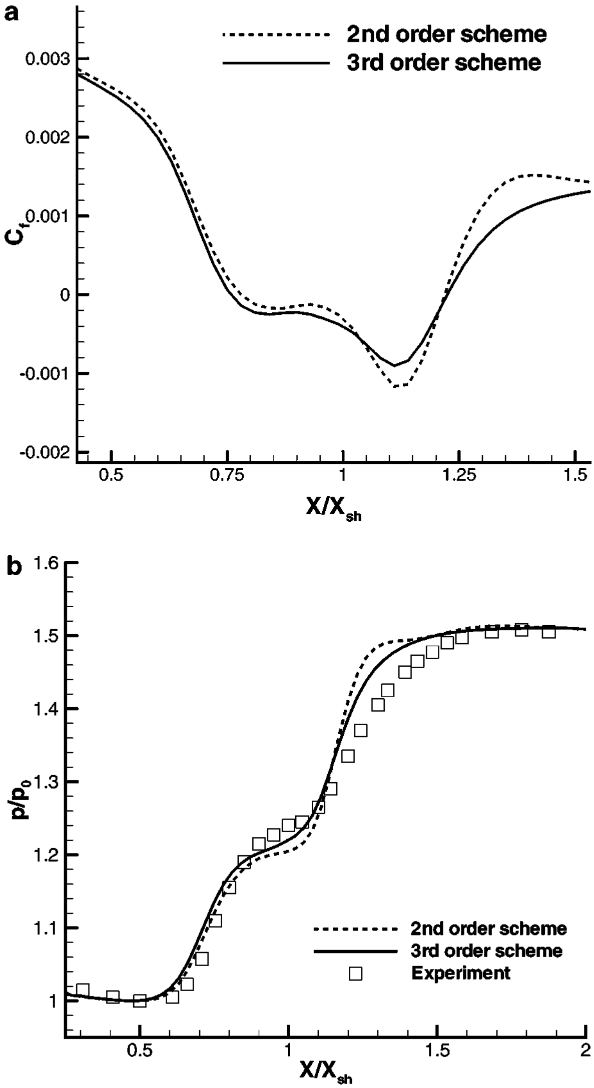

FIG. 17. Shock/boundary layer interaction. (a) Skin-friction distribution along the plate. (b) Pressure distribution along the plate.

viscosity model of Baldwin–Lomax is used. The computational C-mesh contains (320 $\times$ 64) points, with 260 points on the airfoil; the far-field is located at 10 chord lengths from the airfoil and the first mesh interval normal to the body is about 6×10^{-6} chord length. The CFL number is equal to 20. A comparison of the results with the experiment is shown in Fig. 18: there is close agreement with the experiment on the pressure coefficient, and the computed pressure distribution compares well with the results presented by contributors to the EUROVAL project [10]; in particular, the residual-based schemes provide a good shock resolution without requiring any tuning parameter. Though not shown here for space reasons, the skin-friction distribution and velocity profiles at different locations along the airfoil are also in good agreement with both the experimental values given in [4] and the numerical results obtained in [10]. Note that, for this test case, we did not see any noticeable

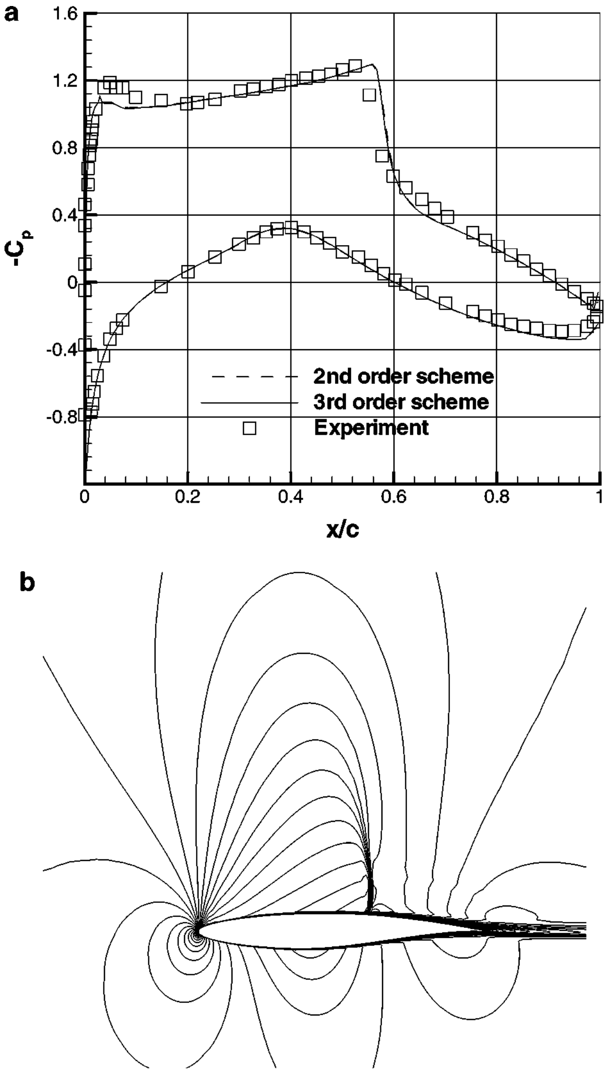

FIG. 18. Turbulent flow over a RAE2822 airfoil. (a) Wall-pressure distribution. (b) Third-order scheme Mach contours.

difference between the second- and third-order results. Possibly, in such a calculation, a rigorous implementation of the residual-based correction on a curvilinear mesh would be of interest.

9. CONCLUSION

A robust and efficient compact method for solving the steady compressible Euler and Navier–Stokes equations with high accuracy has been presented. Compactness and high accuracy are obtained by expressing the truncation error in terms of derivatives of the residual, vanishing at steady state. Robustness and good shock-capturing properties are achieved thanks to a first-order dissipation, also based on the residual, that becomes a high-order term at steady state only. Compactness and first-order dissipation allow the construction of an efficient implicit version of the scheme. These properties were assessed on model advection and advection–diffusion problems using uniform Cartesian meshes, which demonstrated true third-order-accurate solutions, stability, and quick convergence to the steady state. Then, calculations of viscous compressible flow problems, on nonuniform Cartesian meshes and curvilinear grids, showed the ability of the third-order scheme to handle flows with shocks and to provide reasonably good solutions on coarse grids. The low sensitivity of the third-order scheme to grid coarsening, clearly demonstrated on a separated laminar flow problem, could be exploited in, for instance, a multigrid strategy.

In the present work, true high accuracy was not ensured on nonregular grids because, in a first step, the truncation error corrections developed on a regular grid were retained. However, greater accuracy could be achieved by expressing the scheme truncation error on such nonregular grids and correcting this expression in order to make derivatives of the residual appear. Work is in progress on this subject.

The residual-based scheme could also be applied to other applications which do not necessarily require distorted grids, such as the direct simulation of turbulent flows or aeroacoustic problems. Concerning this type of application, it should be noted that the ideas developed in this study to build a third-order scheme are not limited to this order of accuracy, and higher-order schemes could be devised using the strategy suggested in Section 2.1.

Note also that the extension of the scheme to three-dimensional problems is very straightforward. For instance, the residual-based compact approximation (7) will become

$$\begin{aligned}\tilde{r}_{RBC} = {}& \frac{1}{\delta x}\delta_1\mu_1\left(I + \frac{1}{6}\delta_2^2 + \frac{1}{6}\delta_3^2\right) f + \frac{1}{\delta y}\delta_2\mu_2\left(I + \frac{1}{6}\delta_1^2 + \frac{1}{6}\delta_3^2\right) g\\ & + \frac{1}{\delta z}\delta_3\mu_3\left(I + \frac{1}{6}\delta_1^2 + \frac{1}{6}\delta_2^2\right) h,\end{aligned}$$

using obvious notations.

Finally, the generalization of the concept to unsteady problems is easy to devise: the physical time-derivative must simply be included in the residual when building the high-order correction and the dissipative term in order to ensure high-order space accuracy. A dual time step approach will conveniently allow a high-order accuracy in time.

REFERENCES

1. S. Abarbanel and A. Kumar, Compact high-order schemes for the Euler equations, *J. Sci. Comput.* **3**, 275 (1988).
2. S. Z. Burstein, High order accurate difference methods in hydrodynamics, in *Nonlinear Partial Differential Equations*, edited by W. F. Ames (Academic Press, San Diego, 1967), p. 279.
3. B. Cockburn and C. W. Shu, Nonlinearly stable compact schemes for shock calculations, *SIAM J. Numer. Anal.* **31**, 607 (1994).
4. T. J. Cook, M. A. McDonald, and M. C. P. Firmin, *Airfoil RAE2822 Pressure Distributions and Boundary Layer and Wake Measurements,* AGARD Advisory Report 138 (1979).
5. C. Corre, Y. Huang, and A. Lerat, A dramatic improvement of an implicit Lax–Wendroff scheme for steady compressible viscous flow calculations, Proc. of the 15th ICNMFD, *Lecture Notes in Physics* **490**, 219 (1997).
6. C. Corre, K. Khalfallah, and A. Lerat, Line-relaxation methods for a class of centred schemes, *Comput. Fluid Dyn. J.* **5**, 213 (1996).
7. C. Corre and A. Lerat, Efficient calculation of 3-D turbulent transonic flows, Proc. of the 16th ICNMFD, *Lecture Notes in Physics* **515**, 19 (1998).

8. P. I. Crumpton, J. A. Mackenzie, and K. W. Morton, Cell vertex algorithms for the compressible Navier–Stokes equations, *J. Comput. Phys.* **109**, 1 (1993).
9. G. Degrez, C. H. Boccadoro, and J. F. Wendt, The interaction of an oblique shock wave with a laminar boundary layer revisited. An experimental and numerical study, *J. Fluid Mech.* **177**, 247 (1987).
10. EUROVAL, *An European Initiative on Validation of CFD Codes*. *Notes on Numerical Fluid Mechanics* (Vieweg, Wiesbaden, 1993).
11. D. Fu and Y. Ma, High resolution schemes, in *Computational Fluid Dynamics Review 1995,* edited by J. Wiley (Wiley, New York, 1995), p. 234.
12. A. Harten, On a class of high resolution total-variation stable finite-difference schemes, *SIAM J. Numer. Anal.* **21**, 4 (1984).
13. R. S. Hirsh, High order accurate difference solutions of fluid mechanics problems by a compact differencing technique, *J. Comput. Phys.* **19**, 90 (1975).
14. Y. Huang, K. Khalfallah, and A. Lerat, A second-order upwind scheme without limiter for the steady Euler equations, in *Computational Fluid Dynamics Review 1994,* edited by S. Wagner *et al.* (Wiley, New York, 1994), p. 587.
15. Y. Huang and A. Lerat, Second-order upwinding through a characteristic time-step matrix for compressible flow calculations, *J. Comput. Phys.* **142**, 445 (1998).
16. H. D. Kim and N. S. Liu, A time-accurate high-resolution TVD scheme for solving the Navier–Stokes equations, *Comp. Fluids* **22**, 4/5, 517 (1993).
17. S. K. Lele, Compact finite difference schemes with spectral-like resolution, *J. Comput. Phys.* **103**, 16 (1992).
18. A. Lerat, Implicit methods of second order accuracy for Euler equations, *AIAA Paper* **83-1925** (1983); *AIAA J.* **23**, 33 (1985).
19. A. Lerat, Multidimensional centred schemes of the Lax–Wendroff type, in *Computational Fluid Dynamics Review 1995,* edited by M. Hafez and K. Oshima (Wiley, New York, 1995), p. 124.
20. A. Lerat and A. Rezgui, Schémas dissipatifs précis à l'ordre trois pour les systèmes hyperboliques, *C.R. Acad. SC. Paris* **323**, **II b**, 397 (1996).
21. A. Lerat and J. Sidès, Efficient solution of the steady Euler equations with a centered implicit method, in *Numerical Methods for Fluid Dynamics 3,* edited by K. W. Morton and M. J. Baines (Clarendon, Oxford, 1988), p. 65.
22. D. J. Mavriplis and A. Jameson, Multigrid solution of the Navier–Stokes equations on triangular meshes, *AIAA J*. **28**(8), 1415 (1990).
23. R. H. Ni, A Multiple-grid scheme for solving the Euler equations, *AIAA Paper* **81-1025** (1981); *AIAA J.* **20**, 1565 (1982).
24. H. Paillère, E. Van der Weide, and H. Deconinck, Multidimensional upwind methods for inviscid and viscous compressible flows, in *VKI Lecture Series 1995-02* (Von Karman Institute, Brussels, 1995).
25. R. Peyret, Une méthode aux différences finies de haute précision pour la résolution des équations de Navier–Stokes stationnaires, *C.R. Acad. SC. Paris* **286 A**, 59 (1978).
26. T. H. Pulliam, R. W. MacCormack, and S. Venkateswaran, Convergence Characteristics of Approximate Factorization Methods, Proc. of the 16th ICNMFD, *Lecture Notes in Physics* **515,** 409 (1998).
27. P. L. Roe, Approximate Riemann solvers, Parameter vectors, and difference schemes, *J. Comput. Phys.* **43**, 357 (1981).
28. A. I. Tolstykh, *High Accuracy Noncentered Compact Difference Schemes for Fluid Dynamics Applications* (World Scientific, Singapore, 1994).
29. M. R. Visbal and D. V. Gaitonde, High-order accurate methods for complex unsteady subsonic flows, *AIAA J.* **37**(10), 1231 (1999).
30. H. C. Yee, Explicit and implicit multidimensional compact high-accurate resolution shock-capturing methods: formulation, *J. Comput. Phys.* **131**, 216 (1997).