Implicit Upwind Schemes for Low Mach Number Compressible Flows

Cécile Viozat

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Implicit Upwind Schemes for Low Mach Number Compressible Flows

Cécile VIOZAT

N° 3084

Janvier 1997

THÈME 4

ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

Implicit Upwind Schemes for Low Mach Number Compressible Flows

Cécile VIOZAT*

Thème 4 — Simulation et optimisation
de systèmes complexes
Projet SINUS

Rapport de recherche n° 3084 — Janvier 1997 — 67 pages

Abstract: At low Mach number, the Roe scheme presents an excess of artificial viscosity. A correction of this scheme, which uses the preconditioning of Turkel, leads to an improvement of the solution. We refer to this new scheme as the Roe-Turkel scheme. We show that for the Roe-Turkel scheme the convergence of the numerical solution towards the exact solution depends only on the mesh size parameter whereas that of the Roe scheme depends on the ratio between mesh size parameter and Mach number. We also show that for the computation of steady state solutions, when the Roe-Turkel scheme is introduced in the physical phase and in the mathematical phase, the iterative convergence of the implicit Defect-Correction method is only moderately affected at low Mach number as compared with the transonic regime. The proposed formulation is consistent in time (applicable to unsteady flows). It enables a solution at Mach 10^{-6} of the same quality on a same mesh as that obtained at Mach 10^{-1} for flow around a NACA0012 airfoil. The results obtained at Mach 10^{-3} are close to those given by an incompressible Navier-Stokes code on comparable mesh for shear-driven cavity flow.

Key-words: low Mach number flow, Roe scheme, upwind scheme, compressible flow, iterative implicit method, Defect-Correction method

(Résumé : tsvp)

* E-mail:Cecile.Viozat@sophia.inria.fr

Schémas décentrés implicites pour les écoulements à petit nombre de Mach

Résumé : A petit nombre de Mach, le schéma de Roe présente un excès de viscosité numérique. Une correction de ce schéma utilisant le préconditionnement de Turkel permet de supprimer cet excès et entraîne une amélioration de la précision de la solution. Nous nommons ce nouveau schéma le schéma de Roe-Turkel. Nous montrons que pour le schéma de Roe-Turkel la convergence de la solution numérique vers la solution exacte ne dépend que de la finesse du maillage alors que celle du schéma de Roe dépend du rapport entre finesse du maillage et nombre de Mach. Nous montrons également que pour le calcul de solutions stationnaires la convergence itérative de la méthode implicite Defect-Correction devient pratiquement aussi bonne à Mach petit qu'en transsonique lorsque le schéma de Roe-Turkel est introduit dans la phase physique et dans la phase mathématique. La formulation proposée est consistante en temps (applicable à de l'instationnaire). Elle permet d'obtenir une solution à Mach 10^{-6} de même qualité sur un même maillage que celle obtenue à Mach 10^{-1} pour le cas test d'un profil d'aile NACA0012. Les résultats pour Mach 10^{-3} sont proches de ceux donnés par un code pour les écoulements incompressibles à maillage comparable pour le cas test de la cavité carrée.

Mots-clé : écoulement à petit nombre de Mach, schéma de Roe, schéma décentré, écoulement compressible, méthode itérative implicite, méthode Defect-Correction

Contents

Nomenclature

$a = \sqrt{\frac{p\,\gamma}{\rho}}$	Speed of sound
$A = \frac{\partial F}{\partial W}$, $B = \frac{\partial G}{\partial W}$, $C = \frac{\partial H}{\partial W}$	Inviscid flux Jacobians
$E = \rho\,\frac{q^2}{2} + \frac{p}{\gamma - 1}$	Total energy per unit volume for a perfect gas
F, G, H	Inviscid flux vector
$H = \frac{q^2}{2} + \frac{a^2}{\gamma - 1} = \frac{E + p}{\rho}$	Total enthalpy per unit volume for a perfect gas
h	Space step
$M = \frac{q}{a}$	Mach number
p	Pressure
P_e	Preconditioning matrix for entropic variables
P_c	Preconditioning matrix for conservative variables
$q = \sqrt{u^2 + v^2 + w^2}$	Magnitude of velocity
$R = \frac{\partial U}{\partial W}$	Jacobian matrix
Re	Reynolds number
$S = ln\frac{p}{\rho^\gamma}$	Entropic variable
T	Temperature
u, v, w	Velocity components
$V = (u,\ v,\ w)$	Velocity vector
W	Conservative variables (ρ, $\rho\,u$, $\rho\,v$, $\rho\,w$, E)
β	Preconditioning parameter
$\bar{\beta}$	Upwinding parameter for spatial second order accuracy
γ	Ratio of specific heats
λ_i	i^{th} eigenvalue of $A\,\nu_x + B\,\nu_y + C\,\nu_z$
ν_x, ν_y, ν_z	Propagation directions
ρ	Density

1 Introduction

The performance of many existing compressible codes, originally built for transonic or supersonic problems, degrades as the Mach number of the computed flow vanishes; the lower the Mach number is, the more the accuracy of the solution and the more the iterative convergence degrades. However, the need of computing low Mach number flows or locally compressible flows is more and more frequently encountered in engineering. This is the case for cryogenic rocket engines including supersonic jets and slower recirculations, for powder rocket engines and also for piston engines especially at the beginning of the exhaust stroke or at the ignition phase.

The obvious physical difficulty of low Mach number compressible flow computations is the presence of two very different time scales associated with acoustics and gas motion. Prior a low Mach number gas flow is established, the acoustic waves travel many times through the computational domain. The representation of these transitory states being of no interest for steady state computations, the first works on this topic aimed at changing the ratio of wave speeds: preconditioning methods have been developed to accelerate the iterative time convergence. These methods enable computation of steady flows at low Mach number in conditions almost as good (accuracy, efficiency) as for transonic flows.

We study a method which is consistent in time, having in mind unsteady flow applications. We show how a simple modification of the upwind scheme enables us to obtain an accurate solution, while preserving a good behaviour of the implicit algorithm for arbitrary Mach number provided it is positive.

We consider a flow governed by the compressible Navier-Stokes equations for a perfect gas. An upwind discretisation relying on the flux splitting of Roe is employed for the inviscid terms and a centred approximation is employed for the viscous terms. We use a mixed finite volume/finite element method for the spatial approximation of the convective and diffusive fluxes.

After recalling various models for the low Mach number flows, we present a short review of the iterative preconditioning methods. Then, we present the preconditioning of the explicit Euler equations and we propose a formulation which is consistent in time and involves only a modification of the upwind scheme. We study next the uniform consistence with regard to the Mach number of the Roe scheme with and without preconditioning of Turkel. We also study the iterative convergence of the implicit Defect-Correction method at low Mach number. Finally, we examine the accuracy of the results obtained with the new scheme at low Mach number for a shear-driven cavity flow.

2 Review of various models for the low Mach number regime

We first recall some definitions.

A **compressible** flow is a flow whose density changes with pressure and temperature.

A **dilatable** flow is a flow whose density changes with temperature.

In the literature two definitions are found for an **incompressible** flow.
The first one is that it is a flow whose density does not change with pressure [1] [27]. It is thus a fluid which may be dilatable.
The second definition is that it is a flow whose density is constant [17] [3].
In this report, we use the latter definition because as one refers to the incompressible flow equations one usually thinks of the condition $\nabla \cdot V = 0$ which holds when the density is constant.

To compute low Mach number flows several possibilities arise. One can solve the compressible flow equations. One can solve the dilatable flow equations, also called low Mach number equations. In some cases, another alternative consists of solving the incompressible flow equations. In this section, we briefly present these various models.

2.1 The compressible Navier-Stokes equations

In the general case where one wishes to compute the flow of a compressible fluid, like for example the flow around an aircraft, one uses the compressible Navier-Stokes equations. These equations in their nondimensional form are given by,

$$\left\{\begin{array}{rcl} \rho_t + \nabla \cdot (\rho V) & = & 0 \\ \rho(V_t + (V\nabla)V) + \dfrac{\nabla p}{\gamma M^2} & = & \dfrac{1}{Re}\nabla \cdot \tau \\ \rho\left(T_t + V \cdot \nabla T\right) - \dfrac{\gamma - 1}{\gamma}(p_t + V \cdot \nabla p) & = & \dfrac{1}{Re\,Pr}\nabla \cdot \lambda \nabla T + \dfrac{M^2\,(\gamma-1)}{Re}\,(\tau \cdot \nabla V) \\ p & = & \rho T, \end{array}\right. \tag{1}$$

where Re, Pr, λ and μ are respectively the Reynolds number, the Prandtl number, the thermic conductivity and the dynamic viscosity of the fluid.

The stress tensor is (I denoting the identity matrix)

$$\tau = \mu(\nabla V + (\nabla V)^t - 2/3(\nabla \cdot V)I).$$

These equations become stiff and difficult to solve at low Mach number, because of the term $\frac{1}{M^2}$ present in the momentum equation.

2.2 The low Mach Navier-Stokes equations

Majda [19] (among others) showed that as the Mach number goes to zero, the compressible Navier-Stokes equations tend to the incompressible flow equations. When the variations of the temperature cannot be neglected the incompressible flow equations can no longer be used even if the Mach number is low. However, a model derived from the compressible equations by assuming that the Mach number is low can be built. This model called the low Mach number model is easier to solve than the full Navier-Stokes equations. It is often used in combustion problems. Asymptotic expansions of the unknowns of the Navier-Stokes equations in terms of the Mach number [20] [16] are performed

$$\begin{aligned} \rho &= \rho^{(0)} + \gamma M^2 \rho^{(1)} + (\gamma M^2)^2 \rho^{(2)} + O\left((\gamma M^2)^3\right) \\ p &= p^{(0)} + \gamma M^2 p^{(1)} + (\gamma M^2)^2 p^{(2)} + O\left((\gamma M^2)^3\right), \end{aligned} \tag{2}$$

where the unknowns of each order are assumed to be of order unity and $\gamma M^2 \ll 1$.

These asymptotic expansions are substituted into the compressible Navier-Stokes equations and the powers of M are equated. The terms of order M^{-2} in the momentum equation imply $\nabla p^{(0)} = 0$. The quantity $p^{(0)}$ is thus constant in space.

At the zero order, the low Mach number Navier-Stokes equations are obtained

$$\left\{ \begin{aligned} \rho_t^{(0)} + \nabla \cdot (\rho^{(0)} V^{(0)}) &= 0 \\ \rho^{(0)}(V_t^{(0)} + (V^{(0)} \cdot \nabla) V^{(0)}) &= -\nabla p^{(1)} + \frac{1}{Re} \nabla \cdot \tau^{(0)} \\ \rho^{(0)} \left(T_t^{(0)} + V^{(0)} \cdot \nabla T^{(0)}\right) - \frac{\gamma - 1}{\gamma} \frac{dp^{(0)}}{dt} &= \frac{1}{Re\, Pr} \nabla \cdot \lambda \nabla T^{(0)} \\ p^{(0)} &= \rho^{(0)} T^{(0)} \end{aligned} \right. \tag{3}$$

where it is assumed that $\gamma M^2 \ll Re$ and $\gamma M^2 \ll Re\, Pr$, and where $p^{(0)}$ is computed using the boundary conditions [12].

The quantity $p^{(1)}$ is called dynamic pressure because it is directly linked to the speed of the fluid and the quantity $p^{(0)}$ is called thermodynamic pressure because it appears in the temperature equation and in the state law. We note then that in the low Mach number Navier-Stokes equations the density $\rho^{(0)}$ does not change with the dynamic pressure $p^{(1)}$, but it depends on the thermodynamic pressure $p^{(0)}$.

2.3 The incompressible Navier-Stokes equations

In the absence of temperature and density variations, the flow can be solved with the incompressible equations which are given by,

$$\begin{cases} \nabla \cdot V = 0 \\ V_t + V \cdot \nabla V = -\nabla p + \dfrac{1}{Re} \nabla \cdot \tau \end{cases} \tag{4}$$

where the stress tensor is given by

$$\tau = \mu \nabla^2 V. \tag{5}$$

These equations are usually employed to simulate low Mach number flows at constant temperature, like for example flows around a car or a building. They are also used for liquids.

2.4 Motivation for this study

Because of the difficulty of solving directly the compressible Navier-Stokes equations for low Mach number flows, one usually simplifies these equations with the assumption that the Mach number is low or that no variation of the temperature occurs. Depending on the envisaged application, different equations are solved with different codes. It would be interesting to be able to handle with only one code all these applications. Moreover, this code would also enable computation of flows involving supersonic parts and low Mach number parts. This is why certain authors have recently evaluated the performance of compressible Navier-Stokes codes at low Mach number [36] [15]. This performance is poorer than with incompressible flows. The problems arise mainly from the fact that at low Mach number some variables or their fluctuations are very small and others very large. In this report, we study a problem which arises in certain upwind schemes. These schemes blend the wave speeds in the artificial viscosity term. When the Mach number is of order unity this is without serious consequence. On the other hand, when the Mach number is low, the wave speeds are of different magnitudes; this has dramatic consequences as we shall see.

3 Review of the pseudo-unsteady methods with iterative preconditioning

The scheme that we analyse in this report originates from preconditioning methods. In this section, we briefly present the evolution of these methods. A detailed review has been recently given by Turkel in [29] and by Radespiel-Turkel in [24].

3.1 The method of artificial compressibility to solve the incompressible equations

In the inviscid case, the incompressible Euler equations can be written as

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} p \\ u \\ v \end{bmatrix}_t + \begin{bmatrix} 0 & 1 & 0 \\ 1 & u & 0 \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} p \\ u \\ v \end{bmatrix}_x + \begin{bmatrix} 0 & 0 & 1 \\ 0 & v & 0 \\ 1 & 0 & v \end{bmatrix} \begin{bmatrix} p \\ u \\ v \end{bmatrix}_y = 0 \tag{6}$$

whereas the compressible Euler equations can be written as

$$\begin{bmatrix} p \\ u \\ v \\ S \end{bmatrix}_t + \begin{bmatrix} u & \rho a^2 & 0 & 0 \\ \rho^{-1} & u & 0 & 0 \\ 0 & 0 & u & 0 \\ 0 & 0 & 0 & u \end{bmatrix} \begin{bmatrix} p \\ u \\ v \\ S \end{bmatrix}_x + \begin{bmatrix} v & 0 & \rho a^2 & 0 \\ 0 & v & 0 & 0 \\ \rho^{-1} & 0 & v & 0 \\ 0 & 0 & 0 & v \end{bmatrix} \begin{bmatrix} p \\ u \\ v \\ S \end{bmatrix}_y = 0 \tag{7}$$

and the state equation $\rho = \rho(p, S)$ closes the system, S being an entropic variable.
In these equations p, u, v, t, x, y, ρ and a refer respectively to pressure, velocity components, time, spatial coordinates, density and speed of sound.

Chorin in 1967 [6] proposed the method of artificial compressibility to solve the incompressible Navier-Stokes equations when a steady state solution is sought. This method consists of adding a density time derivative to the continuity equation in order to restore the hyperbolic type in this equation and thus to convert a system of mixed elliptic/hyperbolic type to a totally hyperbolic system. Thus, the continuity equation

$$u_x + v_y = 0, \tag{8}$$

is replaced by

$$\begin{cases} \rho_t + u_x + v_y = 0, \\ \rho = \delta\, p \end{cases} \tag{9}$$

where the relation $\rho = \delta\, p$ plays the role of the state equation. The parameter δ is called the *artificial compressibility.*

Equation (9) can be written as

$$\delta\, p_t + u_x + v_y = 0.$$

Thus, the time derivative is multiplied by the following *preconditioning matrix*

$$P^{-1} = \begin{bmatrix} \delta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{10}$$

This method is not consistent in time, but this has no consequences when one is interested in the steady state solution only. The equation is advanced in time until a steady limit solution of the incompressible equations is reached. The steady state solution of the preconditioned system is the same as the solution of the original system. The parameter δ and the time steps are chosen so that the scheme is stable and that the convergence towards the steady state is as rapid as possible. Chorin solved this system with an explicit scheme using a centred finite difference discretisation for time and space derivatives. He applied this method to a thermal convection problem.

3.2 Preconditioning of the equations for low Mach number

The idea of multiplying the time derivative by an artificial term in order to accelerate the convergence to the steady state has been studied by several authors not only for the incompressible equations, but also for the compressible equations. Indeed, the system of the incompressible Euler equations in which terms have been introduced as factors of the time derivative form a totally hyperbolic system, i.e. it is of the same type as the system of the compressible equations. Thus, the problem of solving the incompressible equations and the one of solving the compressible equations at low Mach number meet and can be solved with the same method when a steady state solution is sought. In the mid '80s, these methods were developed among others by Viviand [35], for the compressible Euler equations in the isoenergetic case. For more details about the origin of the pseudo-unsteady methods for the Euler and Navier-stokes equations the reader should consult the book by Peyret and Taylor [22].

There is not just one preconditioner (preconditioning matrix), but several families of preconditioners improving the condition of the fluid mechanic equations because the preconditioner can be derived using different variables and in different ways. The preconditioners are usually derived in primitive variables, p, u, v, plus another variable, because this choice simplifies the equations.

For example, Choi and Merkle [4] proposed the following diagonal preconditioner in a formulation with primitive variables $[\rho,\, u,\, v,\, p]$

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & \beta^{-2} \end{bmatrix}, \tag{11}$$

where β is taken equal to the Mach number.

With an implicit Euler algorithm this preconditioner allows one of the error terms in the momentum equation to remain bounded. This term was originally proportional to the square of the speed of sound and thus became excessively large as the Mach goes to zero. Choi and Merkle have numerically compared the rate of the convergence towards the steady state of the preconditioned compressible equations to the one of the preconditioned incompressible equations. They showed that the convergence of the preconditioned compressible equations is as rapid as the convergence of the incompressible equations solved with the method of the artificial compressibility. For the Navier-Stokes equations, they derived [5] a preconditioner in the variables $[\rho,\, u,\, v,\, T]$ and they showed that it significantly accelerates the convergence rate. This preconditioner is used to compute flows around a cylinder, in a nozzle and in a square wall driven cavity.

The wave speeds of the preconditioned Euler equations solved by Choi and Merkle are given by the eigenvalues of $P\,(A\,\nu_x + B\,\nu_y)$ where P denotes the inverse of P^{-1} given in (11), A and B are the inviscid flux Jacobians with respect to the primitive variables, and where ν_x and ν_y are two real values corresponding to the propagation directions of the waves. These wave speeds are

$$\left\{ \begin{array}{rcl} \lambda_{1,2} & = & u\,\nu_x + v\,\nu_y \\ \lambda_{3,4} & = & \dfrac{1}{2}\left[(1\,+\,\beta^2)\,\lambda_{1,2} \pm \sqrt{[(1-\beta^2)\,\lambda_{1,2}]^2 + [2\,\beta\,a\ \|\ \nu\ \|]^2}\right]. \end{array} \right. \tag{12}$$

The speed of the material waves is unchanged by the preconditioning; however, at low Mach number the acoustic waves are slowed down and stay almost at the same speed as the material waves. Indeed, if we assume $\beta \sim M$, $u \sim 1$, $\|\ \nu\ \| \sim 1$ and $M \to 0$, we obtain

$$\lambda_{3,4} \sim \frac{1 \pm \sqrt{5}}{2}\,\lambda_{1,2}.$$

The condition number of the inviscid flux matrices is equal to the ratio of the largest to the smallest eigenvalue. It is about 2.6 for the preconditioned system while for the original one it is inversely proportional to the Mach number; the condition number becomes larger as the Mach number tends to zero.

Turkel in 1987 [28] proposed a preconditioner which reduces the condition number to one in the low Mach number case. This preconditioner is derived in the entropic variables [p, u, v, S].

Turkel generalised the method developed by Chorin by adding pressure time derivatives to the momentum equations and not just to the continuity equation. He used two free parameters, β in front of the time derivative in the continuity equation and α in front of the time derivative in the momentum equations. Thus, the preconditioner proposed by Turkel is given by

$$P^{-1} = \begin{bmatrix} \dfrac{1}{\rho\,\beta^2} & 0 & 0 & 0 \\ \dfrac{\alpha\,u}{\rho\,\beta^2} & 1 & 0 & 0 \\ \dfrac{\alpha\,v}{\rho\,\beta^2} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{13}$$

where β^2 is of the order of $u^2 + v^2$.

The wave speeds are then

$$\begin{cases} \lambda_{1,2} &= u\,\nu_x + v\,\nu_y \\ \lambda_{3,4} &= \dfrac{1}{2}\left[\left(1 - \alpha + \dfrac{\beta^2}{a^2}\right)\lambda_{1,2} \pm \sqrt{\left[\left(1 - \alpha + \dfrac{\beta^2}{a^2}\right)^2 \lambda_{1,2}\right]^2 + 4\left[1 - \dfrac{\lambda_{1,2}^2}{a^2}\right]\beta^2}\ \right]. \end{cases} \tag{14}$$

When $\alpha = 0$, one obtains the same eigenvalues as with the preconditioner of Choi and Merkle, and the condition number is about 2.6 at low Mach number.

When $\alpha = 1$, the condition number is 1, that is all the waves propagate at the same speed. This is the optimum condition number.

3.3 Preconditioning of the equations for a large range of Mach numbers

Without preconditioning, the wave speeds are

$$\begin{cases} \lambda_{1,2} &= u\,\nu_x + v\,\nu_y \\ \lambda_{3,4} &= \lambda_{1,2} \pm a \parallel \nu \parallel, \end{cases} \tag{15}$$

and the condition number of the system is large not only at low Mach number, but also, as noticed in Lee's thesis [18], in the transonic case for example. Indeed, when the flow

speed is very close to the sound speed, $\lambda_{1,2} \sim a \parallel \nu \parallel$ implies that the ratio λ_3 over λ_4, $\frac{\lambda_{1,2} + a \parallel \nu \parallel}{\lambda_{1,2} - a \parallel \nu \parallel}$, is large and thus the convergence rate is slow. Knowing that a condition number close to 1 leads to a better convergence than a large condition number, certain authors investigated preconditioners applicable to a large range of Mach numbers.

The preconditioner proposed by van Leer-Lee-Roe [33] [14] equalises the eigenvalues whatever the Mach number. This preconditioner is derived in the stream-aligned symmetrisation variables defined by their differentials: $[\frac{dp}{\rho a}, du, dv, dp - a^2 d\rho]$.

3.4 Preconditioning and upwind schemes

When the equations are preconditioned and a Roe upwind scheme [25] is used, van Leer-Lee-Roe [33] observed that the numerical fluxes must be modified.

The numerical fluxes for the original 1D equations are

$$F_{i+1/2} = \frac{1}{2}\left[F(W_i) + F(W_{i+1})\right] - \frac{1}{2} \mid (\tilde{A}_c)_{i+1/2} \mid (W_{i+1} - W_i), \tag{16}$$

where $\tilde{A}_c$ is the Roe matrix in conservative variables W.

For the preconditioned equations these fluxes become

$$F_{i+1/2} = \frac{1}{2}\left[F(W_i) + F(W_{i+1})\right] - \frac{1}{2} P_c^{-1} \mid P_c\,(\tilde{A}_c)_{i+1/2} \mid (W_{i+1} - W_i), \tag{17}$$

where P_c is the preconditioner in conservative variables.

Van Leer-Lee-Roe [33] noted also that this modification is not necessary with the flux vector splitting of van Leer [32]. This modification improves the numerical solution, and then the preconditioned compressible equations tend to the preconditioned incompressible equations [30].

3.5 Robustness

A problem encountered with certains preconditioners, particularly the one of Turkel and the one of Leer-Lee-Roe, is that they are singular at stagnation points and along sonic lines. In order to avoid these singularities, certains terms must be smoothed [34] or bounded so that the non-normality of the eigenvectors appearing at very low Mach number is eliminated [7].

4 Explicit scheme: preconditioning for the Euler equations

We present the preconditioning of the equations in the case where an explicit method with an upwind scheme is used. We then study four different ways of introducing a given preconditioner in these equations.

4.1 The Euler equations in conservative variables

For flow simulations of practical interest, possibly with shocks, we consider the conservative variables.

The 2D Euler equations in their conservative differential form can be written as

$$\frac{\partial W}{\partial t} + \frac{\partial F(W)}{\partial x} + \frac{\partial G(W)}{\partial y} = 0,$$

$$W = \begin{bmatrix} \rho \\ \rho\, u \\ \rho\, v \\ E \end{bmatrix},\ F(W) = \begin{bmatrix} \rho\, u \\ \rho\, u^2 + p \\ \rho\, u\, v \\ (E + p)\, u \end{bmatrix},\ G(W) = \begin{bmatrix} \rho\, v \\ \rho\, u\, v \\ \rho\, v^2 + p \\ (E + p)\, v \end{bmatrix}, \tag{18}$$

where ρ is density, u and v are components of the velocity vector, p is pressure and E is the total energy per unit volume. The energy E is obtained by the state equation for a perfect gas

$$E = \rho\, \frac{u^2 + v^2}{2} + \frac{p}{\gamma - 1}, \tag{19}$$

where γ is the ratio of specific heats ($\gamma = 1.4$ for air).

System (18) can also be written in non conservative form:

$$W_t + A_c\, W_x + B_c\, W_y = 0, \tag{20}$$

where A_c and B_c are the inviscid flux Jacobians given by

$$A_c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ \dfrac{\gamma - 3}{2}\, u^2 + \dfrac{\gamma - 1}{2}\, v^2 & (3 - \gamma)\, u & (1 - \gamma)\, v & \gamma - 1 \\ -u\, v & v & u & 0 \\ (\gamma - 1)\, q^2\, u - \dfrac{\gamma\, E\, u}{\rho} & \dfrac{\gamma\, E}{\rho} - \dfrac{\gamma - 1}{2}(3\, u^2 + v^2) & (1 - \gamma)\, u\, v & \gamma\, u \end{bmatrix},$$

and

$$B_c = \begin{bmatrix} 0 & 0 & 1 & 0 \\ -u\,v & v & u & 0 \\ \dfrac{\gamma-1}{2}\,u^2 + \dfrac{\gamma-3}{2}\,v^2 & (1-\gamma)\,u & (3-\gamma)\,v & \gamma-1 \\ (\gamma-1)\,q^2\,v - \dfrac{\gamma\,E\,v}{\rho} & (1-\gamma)\,u\,v & \dfrac{\gamma\,E}{\rho} - \dfrac{\gamma-1}{2}\,(u^2+3\,v^2) & \gamma\,v \end{bmatrix},$$

where q is the magnitude of velocity.

4.2 Iterative preconditioning and upwind schemes

4.2.1 The one-dimensional case

Recall that in the scalar and linear case, a first-order upwind scheme is obtained when approximating the first derivative $a\,\dfrac{\partial w}{\partial x}$ with

$$\begin{array}{lll} a\dfrac{w_{i+1}-w_i}{h} & \text{if} & a<0 \\ a\dfrac{w_i-w_{i-1}}{h} & \text{if} & a\geq 0, \end{array} \tag{21}$$

or equivalently by

$$a\left(\frac{w_{i+1}-w_{i-1}}{2\,h}\right) - \frac{h}{2}\mid a\mid \left(\frac{w_{i-1}-2\,w_i+w_{i+1}}{h^2}\right). \tag{22}$$

The second part of this sum is the stabilisation term (or artificial viscosity); it enables the scheme to be stable.

Approximation (22) can also be written as

$$\frac{f_{i+1/2}-f_{i-1/2}}{h} \tag{23}$$

where

$$\begin{aligned} f_{i+1/2} &= a\,\frac{w_{i+1}+w_i}{2} - \frac{1}{2}\mid a\mid (w_{i+1}-w_i)\,, \\ f_{i-1/2} &= a\,\frac{w_i+w_{i-1}}{2} - \frac{1}{2}\mid a\mid (w_i-w_{i-1})\,. \end{aligned} \tag{24}$$

Expanding (24) in Taylor's series yields:

$$\frac{f_{i+1/2}-f_{i-1/2}}{h} = a\,\frac{\partial w}{\partial x} - \frac{h}{2}\mid a\mid \frac{\partial^2 w}{\partial x^2} + O(h^2). \tag{25}$$

Thus, the time-continuous "equivalent" [22] or "modified" [37] equation of the advection equation

$$\frac{\partial w}{\partial t} + a\frac{\partial w}{\partial x} = 0, \tag{26}$$

is given by

$$\frac{\partial w}{\partial t} + a\frac{\partial w}{\partial x} = \frac{h}{2} \mid a \mid \frac{\partial^2 w}{\partial x^2} + O(h^2). \tag{27}$$

4.2.2 The two-dimensional case

We use the finite volume method to solve the Euler equation. This method consists of dividing the domain into cells and of expressing conservation of the fluxes going in and out in the course of time.

Applying the first-order accurate upwind scheme of the previous section to the Euler equation (20) is equivalent to solving exactly the following "equivalent" equation (a *cartesian mesh* with space steps equal in both directions, $\Delta x = \Delta y = h$ has been assumed)

$$W_t + A_c W_x + B_c W_y - \frac{h}{2}\left[(\mid A_c \mid W_x)_x + (\mid B_c \mid W_y)_y\right] = 0, \tag{28}$$

in which second-order terms have been neglected.

In this case, for computing $\mid A_c \mid$ and $\mid B_c \mid$, the matrices A_c and B_c are diagonalised, and the absolute value of each eigenvalue of the diagonal matrix is taken.

The preconditioned Euler equations can be written as

$$P_c^{-1} W_t + A_c W_x + B_c W_y = 0, \tag{29}$$

or in the hyperbolic form

$$W_t + P_c A_c W_x + P_c B_c W_y = 0. \tag{30}$$

Applying the first-order upwind scheme to (30) is equivalent to solving exactly the following "equivalent" equation

$$W_t + P_c A_c W_x + P_c B_c W_y - \frac{h}{2}\left[(\mid P_c A_c \mid W_x)_x + (\mid P_c B_c \mid W_y)_y\right] = 0, \tag{31}$$

in which second-order terms have been neglected.

After multiplying this equation by P_c^{-1} we obtain

$$P_c^{-1} W_t + A_c W_x + B_c W_y - \frac{h}{2}\left[(P_c^{-1} \mid P_c A \mid W_x)_x + (P_c^{-1} \mid P_c B_c \mid W_y)_y\right] = 0. \tag{32}$$

Let us consider the following Partial Differential Equation

$$\int_{cell} W_t + \sum_{flux} (F(W)\,\nu_x + G(W)\,\nu_y) - \sum_{flux} \frac{1}{2} \mid A_c\,\nu_x + B_c\,\nu_y \mid \delta W = 0 \tag{33}$$

from which we seek the steady state solution

$$\sum_{flux} (F(W)\,\nu_x + G(W)\,\nu_y) - \sum_{flux} \frac{1}{2} \mid A_c\,\nu_x + B_c\,\nu_y \mid \delta W = 0. \tag{34}$$

The time advancing of (33) towards a solution of (34) may be excessively slow because the wave speeds of (33) are of very different magnitudes. Therefore, equation (33) is preconditioned as equation (32). We obtain

$$P_c^{-1} \int_{cell} W_t + \sum_{flux} (F(W)\,\nu_x + G(W)\,\nu_y) - \sum_{flux} \frac{1}{2} P_c^{-1} \mid P_c\,A_c\,\nu_x + P_c\,B_c\,\nu_y \mid \delta W = 0. \tag{35}$$

This equation can be written as

$$\int_{cell} W_t + P_c \left[\sum_{flux} (F(W)\,\nu_x + G(W)\,\nu_y) - \sum_{flux} \frac{1}{2} P_c^{-1} \mid P_c\,A_c\,\nu_x + P_c\,B_c\,\nu_y \mid \delta W \right] = 0. \tag{36}$$

In conclusion, two modifications can be performed in order to precondition the Euler equations. The first is to multiply at each time step the numerical fluxes by the preconditioner written in the conservative variables P_c. The second modification is on the stabilisation term.

4.3 Preconditioner

4.3.1 Formulation with entropic variables

In the sequel, we consider a preconditioner close to the one of Turkel [28] for the case $\alpha = 0$ in (13). The entropic variables $U = [p,\ u,\ v,\ S]$ where $S = ln\frac{p}{\rho^\gamma}$ are used. With these variables the Euler equation can be written as

$$U_t + A_e\,U_x + B_e\,U_y = 0, \tag{37}$$

where

$$A_e = \begin{bmatrix} u & \rho\,a^2 & 0 & 0 \\ \rho^{-1} & u & 0 & 0 \\ 0 & 0 & u & 0 \\ 0 & 0 & 0 & u \end{bmatrix}, B_e = \begin{bmatrix} v & 0 & \rho\,a^2 & 0 \\ 0 & v & 0 & 0 \\ \rho^{-1} & 0 & v & 0 \\ 0 & 0 & 0 & v \end{bmatrix}. \tag{38}$$

We consider

$$P_e = \begin{bmatrix} \beta^2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{39}$$

where β is of the order of the Mach number.

4.3.2 Formulation with conservative variables

To obtain the preconditioner with conservative variables, P_c, one determines first the Jacobian matrix $R = \dfrac{\partial U}{\partial W}$ and its inverse, then one computes P_c by

$$P_c = R^{-1}\, P_e\, R.$$

The Jacobian matrices are given by

$$R^{-1} = \begin{bmatrix} \dfrac{1}{a^2} & 0 & 0 & -\dfrac{\rho}{\gamma} \\ \dfrac{u}{a^2} & \rho & 0 & -\dfrac{\rho\, u}{\gamma} \\ \dfrac{v}{a^2} & 0 & \rho & -\dfrac{\rho\, v}{\gamma} \\ \dfrac{H}{a^2} & \rho\, u & \rho\, v & -\dfrac{\rho\, q^2}{2\,\gamma} \end{bmatrix}, \; R = \begin{bmatrix} \dfrac{(\gamma-1)\, q^2}{2} & -u\,(\gamma-1) & -v\,(\gamma-1) & \gamma-1 \\ -\dfrac{u}{\rho} & \dfrac{1}{\rho} & 0 & 0 \\ -\dfrac{v}{\rho} & 0 & \dfrac{1}{\rho} & 0 \\ \dfrac{(\gamma-1)\, q^2}{2\, p} - \dfrac{\gamma}{\rho} & -\dfrac{u\,(\gamma-1)}{p} & -\dfrac{v\,(\gamma-1)}{p} & \dfrac{\gamma-1}{p} \end{bmatrix}.$$

where H is the total enthalpy per unit volume

$$H = \frac{q^2}{2} + \frac{a^2}{\gamma - 1}.$$

After computations one obtains the preconditioner with the conservative variables

$$P_c = Id + (\beta^2 - 1)\, \frac{\gamma - 1}{a^2} \begin{bmatrix} \dfrac{q^2}{2} & -u & -v & 1 \\ \dfrac{q^2}{2}\, u & -u^2 & -v\, u & u \\ \dfrac{q^2}{2}\, v & -u\, v & -v^2 & v \\ \dfrac{q^2}{2}\, H & -u\, H & -v\, H & H \end{bmatrix}.$$

4.4 Modification of the stabilisation term

The modification of the stabilisation term consists in puting $P_c^{-1} \mid P_c\, A_c\, \nu_x + P_c\, B_c\, \nu_y \mid$ in place of $\mid A_c\, \nu_x + B_c\, \nu_y \mid$.

To simplify the diagonalisation to perform, we write

$$P_c^{-1} \mid P_c\, A_c\, \nu_x + P_c\, B_c\, \nu_y \mid = R^{-1}\, P_e^{-1} \mid P_e\, A_e\, \nu_x + P_e\, B_e\, \nu_y \mid R. \tag{40}$$

This leads to the diagonalisation of the matrix $T = P_e\, A_e\, \nu_x + P_e\, B_e\, \nu_y$ which is simpler than the matrix $P_c\, A_c\, \nu_x + P_c\, B_c\, \nu_y$.

The matrix T is given by

$$T = \begin{bmatrix} \beta^2 (u\, \nu_x + v\, \nu_y) & \rho\, \beta^2\, a^2\, \nu_x & \rho\, \beta^2\, a^2\, \nu_y & 0 \\ \dfrac{\nu_x}{\rho} & u\, \nu_x + v\, \nu_y & 0 & 0 \\ \dfrac{\nu_y}{\rho} & 0 & u\, \nu_x + v\, \nu_y & 0 \\ 0 & 0 & 0 & u\, \nu_x + v\, \nu_y \end{bmatrix}.$$

The diagonalisation of T is performed in the following way

$$T = M\, \Lambda\, M^{-1}. \tag{41}$$

By definition

$$\mid T \mid = M \mid \Lambda \mid M^{-1}. \tag{42}$$

The diagonal matrix is

$$\Lambda = diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \tag{43}$$

where the eigenvalues $\lambda_1, \lambda_2, \lambda_3$ and λ_4 are given in (12).

The matrices of the right eigenvectors and of the left eigenvectors are respectively given by

$$M = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & -\widehat{\nu}_y & \dfrac{\widehat{r}\, \widehat{\nu}_x}{\rho\, \beta^2\, a^2} & \dfrac{\widehat{s}\, \widehat{\nu}_x}{\rho\, \beta^2\, a^2} \\ 0 & \widehat{\nu}_x & \dfrac{\widehat{r}\, \widehat{\nu}_y}{\rho\, \beta^2\, a^2} & \dfrac{\widehat{s}\, \widehat{\nu}_y}{\rho\, \beta^2\, a^2} \\ 1 & 0 & 0 & 0 \end{bmatrix}, \; M^{-1} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -\widehat{\nu}_y & \widehat{\nu}_x & 0 \\ \dfrac{\widehat{s}}{2\, \widehat{t}} & -\dfrac{\rho\, \beta^2\, a^2}{2\, \widehat{t}}\, \widehat{\nu}_x & -\dfrac{\rho\, \beta^2\, a^2}{2\, \widehat{t}}\, \widehat{\nu}_y & 0 \\ -\dfrac{\widehat{r}}{2\, \widehat{t}} & \dfrac{\rho\, \beta^2\, a^2}{2\, \widehat{t}}\, \widehat{\nu}_x & \dfrac{\rho\, \beta^2\, a^2}{2\, \widehat{t}}\, \widehat{\nu}_y & 0 \end{bmatrix}. \tag{44}$$

where

$$r = \lambda_3 - \lambda_1\, \beta^2, \quad s = \lambda_4 - \lambda_1\, \beta^2, \quad t = \frac{\lambda_4 - \lambda_3}{2},$$

and where the stressed variables $\widehat{x}$ are defined by $\widehat{x} = x/\sqrt{\nu_x^2 + \nu_y^2}$.

Let us note that when $\beta = 1$ we have

$$\widehat{r} = a, \quad \widehat{s} = -a, \quad \widehat{t} = -a,$$

while when $M \to 0$ and $\beta \sim M$ the variables $\widehat{r}$, $\widehat{s}$ and $\widehat{t}$ are of the order of the flow velocity.

From (40) and (41), we have

$$P_c^{-1} \mid P_c A_c \nu_x + P_c B_c \nu_y \mid = R^{-1} P_e^{-1} M \mid \Lambda \mid M^{-1} R, \tag{45}$$

or equally

$$P_c^{-1} \mid P_c A_c \nu_x + P_c B_c \nu_y \mid = T_{rt}^g \mid \Lambda \mid T_{rt}^d \tag{46}$$

with

$$T_{rt}^g = \begin{bmatrix} 1 & 0 & \dfrac{1}{2\,\beta^2\,a^2} & \dfrac{1}{2\,\beta^2\,a^2} \\ u & \widehat{\nu}_y & \dfrac{u + \widehat{r}\,\widehat{\nu}_x}{2\,\beta^2\,a^2} & \dfrac{u + \widehat{s}\,\widehat{\nu}_x}{2\,\beta^2\,a^2} \\ v & -\widehat{\nu}_x & \dfrac{v + \widehat{r}\,\widehat{\nu}_y}{2\,\beta^2\,a^2} & \dfrac{v + \widehat{s}\,\widehat{\nu}_y}{2\,\beta^2\,a^2} \\ \dfrac{q^2}{2} & u\,\widehat{\nu}_y - v\,\widehat{\nu}_x & \dfrac{H + \widehat{r}\,\widehat{\lambda}_1}{2\,\beta^2\,a^2} & \dfrac{H + \widehat{s}\,\widehat{\lambda}_1}{2\,\beta^2\,a^2} \end{bmatrix} \tag{47}$$

and

$$T_{rt}^d = \begin{bmatrix} 1 - \dfrac{\gamma-1}{a^2}\dfrac{q^2}{2} & \dfrac{\gamma-1}{a^2}\,u & \dfrac{\gamma-1}{a^2}\,v & -\dfrac{\gamma-1}{a^2} \\ v\,\widehat{\nu}_x - u\,\widehat{\nu}_y & \widehat{\nu}_y & -\widehat{\nu}_x & 0 \\ \dfrac{\widehat{s}\,\frac{q^2}{2}\,(\gamma-1) + \beta^2\,a^2\,\widehat{\lambda}_1}{\widehat{t}} & -\dfrac{\widehat{s}\,u\,(\gamma-1) + \beta^2\,a^2\,\widehat{\nu}_x}{\widehat{t}} & -\dfrac{\widehat{s}\,v\,(\gamma-1) + \beta^2\,a^2\,\widehat{\nu}_y}{\widehat{t}} & \dfrac{\widehat{s}\,(\gamma-1)}{\widehat{t}} \\ -\dfrac{\widehat{r}\,\frac{q^2}{2}\,(\gamma-1) + \beta^2\,a^2\,\widehat{\lambda}_1}{\widehat{t}} & \dfrac{\widehat{r}\,u\,(\gamma-1) + \beta^2\,a^2\,\widehat{\nu}_x}{\widehat{t}} & \dfrac{\widehat{r}\,v\,(\gamma-1) + \beta^2\,a^2\,\widehat{\nu}_y}{\widehat{t}} & -\dfrac{\widehat{r}\,(\gamma-1)}{\widehat{t}} \end{bmatrix}. \tag{48}$$

When $\beta = 1$, we recognize the well-known eigenvector matrices. We note that only two components of the fluxes are modified by the preconditioning because the first two columns of T_{rt}^g and the first two lines of T_{rt}^d are unchanged by the preconditioning. Only the acoustic wave propagation is modified.

The matrices R^{-1}, R, T, M, M^{-1}, T_{rt}^g, T_{rt}^d and the eigenvalues are given in Appendix A for the 3D case. The stabilisation term $P_c^{-1} \mid P_c A_c \mid \delta W$ is also computed in this appendix.

4.5 Numerical tests

In this section, the numerical method employed is first presented, then four formulations are proposed according to the preconditioning. Finally, the third formulation is tested on a shock tube problem and all four formulations on a NACA0012 airfoil flow problem.

4.5.1 Numerical method

We introduce the MUSCL-FEM second-order accurate upwind approximation [8] used on a non structured triangulation. This method combines the "upwind-TVD Finite Volume " approach for the inviscid terms and the "Finite Element" approach for the viscous terms. We briefly remind ourselves here of the spatial discretisation of hyperbolic terms.

A dual mesh is defined by building around each node i of the mesh a control cell C_i (see in Figure 1) from the triangles having i for vertex. The boundary of C_i is denoted by ∂C_i and is obtained by joining the middle points of the edges ij to the centroids G of each triangle having i and j as common vertices.

The integration of the numerical fluxes is performed by splitting ∂C_i into interfaces $\partial C_{ij} = [G_{1,ij}, I_{ij}] \cup [I_{ij}, G_{2,ij}]$ (see Figure 2). The outward normal to ∂C_{ij} is denoted η_{ij}:

$$\eta_{ij} = \int_{\partial C_{ij}} \vec{\nu} d\sigma.$$

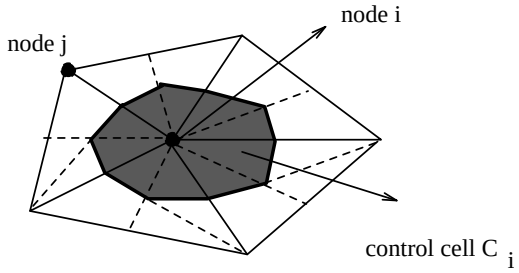

Figure 1: Control cell C_i

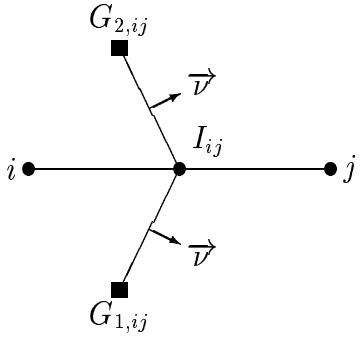

Figure 2: Interface ∂C_{ij} separating nodes i and j

The first-order numerical fluxes can be written as

$$\Phi_{ij} = \Phi(W_i, W_j, \eta_{ij}). \tag{49}$$

A second-order spatial accurate scheme for hyperbolic terms can be obtained using the MUSCL interpolation technique introduced by van Leer [31]. To reach the second-order accuracy the numerical fluxes are evaluated with extrapolated values W_{ij} and W_{ji} at the interface ∂C_{ij}. Thus, the function Φ remains the same, only its arguments are modified:

$$\Phi_{ij} = \Phi(W_{ij}, W_{ji}, \eta_{ij}). \tag{50}$$

The quantities W_{ij} and W_{ji} are computed by

$$\begin{aligned} W_{ij} &= W_i + \frac{1}{2}\nabla W_{ij} \cdot \vec{ij} \\ W_{ji} &= W_j + \frac{1}{2}\nabla W_{ji} \cdot \vec{ji}. \end{aligned} \tag{51}$$

A nodal gradient centred at the middle of the edge ij (see in figure 3) is defined by

$$(\nabla W)_{ij}^{Cent} \cdot \vec{ij} = W_j - W_i \tag{52}$$

and an upwind nodal gradient is defined by

$$\begin{aligned} (\nabla W)_{ij}^{Decent} \cdot \vec{ij} &= \nabla W \mid_{T_{ij}} \\ (\nabla W)_{ji}^{Decent} \cdot \vec{ji} &= \nabla W \mid_{T_{ji}}, \end{aligned} \tag{53}$$

where T_{ij} and T_{ji} are defined in Figure 3.

Figure 3: Downstream and upstream triangles T_{ij} and T_{ji}

We use a "β-scheme" which combines the centred and fully upwind gradient to obtain

$$\nabla W_{ij} \cdot \vec{ij} = (1 - \bar{\beta})(\nabla W)_{ij}^{Cent} \cdot \vec{ij} + \bar{\beta}\,(\nabla W)_{ij}^{Decent} \cdot \vec{ij} \tag{54}$$

where $\bar{\beta}$ is the parameter of upwinding included in interval $[0, 1]$. In our test case, we took $\bar{\beta} = \frac{1}{2}$.

The scheme described above is not monotone. It can introduce extrema which would not exist, particularly in the case of transonic and supersonic flows. To reduce the oscillations in the solution a slope-limiting procedure can be used. In our test case, we did not use limiters.

4.5.2 The four formulations of the equations

We consider the following four formulations in which the stabilisation terms are represented by their differential equivalents

1- Without preconditioning

$$W_t + A_c W_x + B_c W_y - \frac{h}{2}(\mid A_c \mid W_x)_x - \frac{h}{2}(\mid B_c \mid W_y)_y = 0. \tag{55}$$

2- With iterative preconditioning but without preconditioning of the stabilisation term

$$P_c^{-1} W_t + A_c W_x + B_c W_y - \frac{h}{2}\left[(\mid A_c \mid W_x)_x + (\mid B_c \mid W_y)_y\right] = 0, \tag{56}$$

or in an equal way

$$W_t + P_c A_c W_x + P_c B_c W_y - \frac{h}{2}\left[P_c(\mid A_c \mid W_x)_x + P_c(\mid B_c \mid W_y)_y\right] = 0. \tag{57}$$

3- With iterative preconditioning and with preconditioning of the stabilisation term [33]

$$W_t + P_c A_c W_x + P_c B_c W_y - \frac{h}{2}\left[(\mid P_c A_c \mid W_x)_x + (\mid P_c B_c \mid W_y)_y\right] = 0, \tag{58}$$

or in an equal way

$$P_c^{-1} W_t + A_c W_x + B_c W_y - \frac{h}{2}\left[(P_c^{-1} \mid P_c A_c \mid W_x)_x + (P_c^{-1} \mid P_c B_c \mid W_y)_y\right] = 0. \tag{59}$$

4- With preconditioning of the stabilisation term only

$$W_t + A_c W_x + B_c W_y - \frac{h}{2}\left[P_c^{-1}(\mid P_c A_c \mid W_x)_x + P_c^{-1}(\mid P_c B_c \mid W_y)_y\right] = 0. \tag{60}$$

This is the preconditioning that we propose to study in the sequel.

Remark:
Only Formulations 1 and 4 are consistent in time and enable the prediction of unsteady flows. Formulations 2 and 3 enable to reach steady states more quickly.
Formulation 1 represents the original Roe scheme with conservative variables. Formulation 2 results from a modification of the time derivative and converges faster towards a steady state solution which is identical to the solution obtained with Formulation 1. The third formulation represents a Roe-type scheme applied to the *preconditioned hyperbolic problem* since the total fluxes ($P_c\,A_c$,$P_c\,B_c$) are considered in the stabilisation term. In the fourth formulation the stabilisation of the third formulation is kept but the consistency in time is recovered. ■

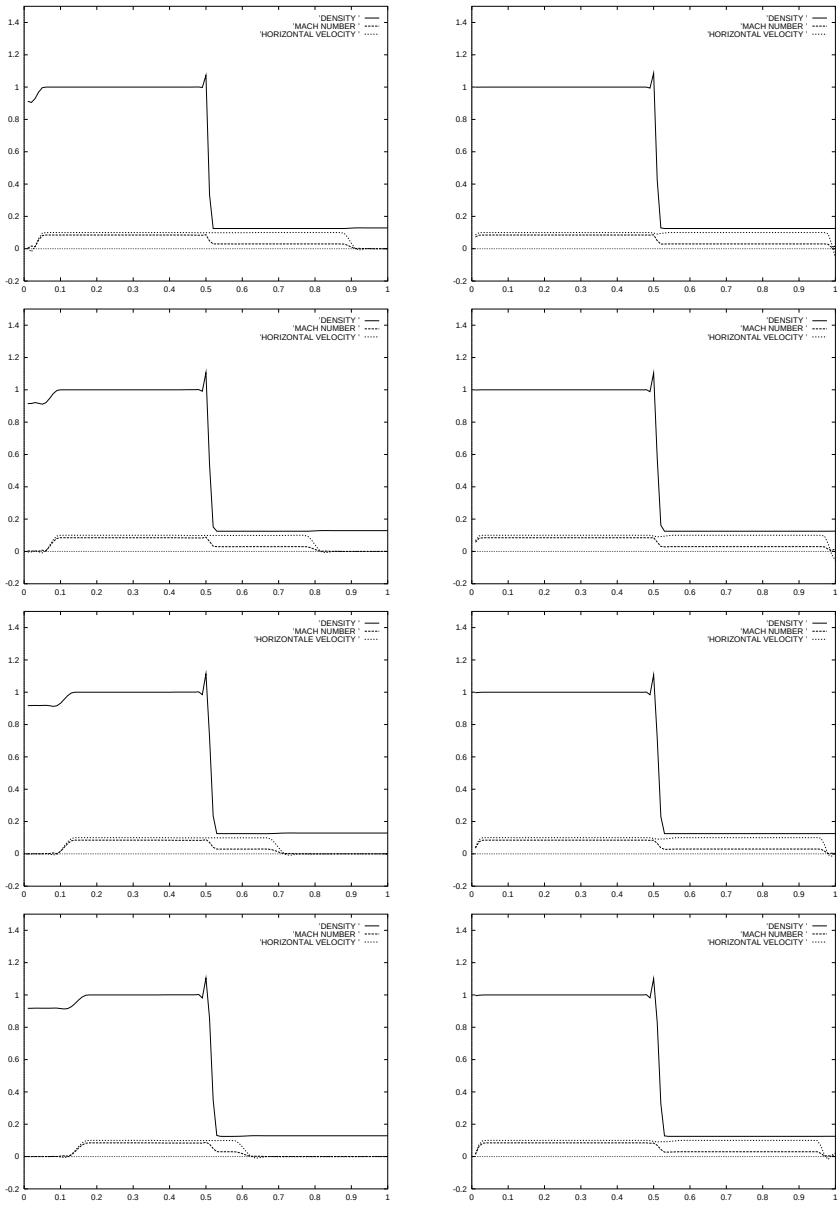

Figure 4: Evolution of the acoustic waves (Mach, horizontal velocity) and of the material waves (density) in a shock tube at low Mach number without preconditioning (left; Formulation 1) and with preconditioning (right; Formulation 3), at time intervals of 0.03 starting at t = 0. Second-order spatial accuracy, without limiters.

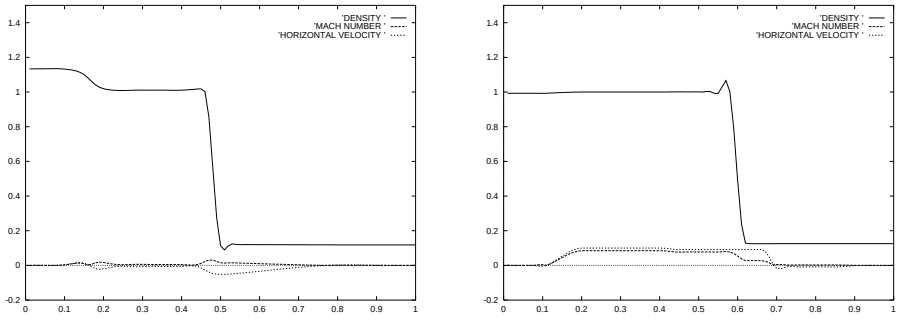

Figure 5: Evolution of the acoustic waves (Mach, horizontal velocity) and of the material waves (density) in a shock tube at low Mach number without preconditioning (left; Formulation 1) and with preconditioning (right; Formulation 3), at time t = 1. Second-order spatial accuracy, without limiters.

4.5.3 Shock tube

The shock tube problem we consider is one-dimensional and the computations are made with a 2D code in a square domain $\Omega = [0,1]^2$. The initial conditions are the following:

$$\begin{array}{cccc} \rho_1 = 1 & u_1 = 0.1 & v_1 = 0 & E_1 = 2.5 \\ \rho_2 = 0.125 & u_2 = 0.0125 & v_2 = 0 & E_2 = 2.5. \end{array} \tag{61}$$

The variation of the pressure will instigate an horizontal movement and the Mach number will at the most be of the order of 0.1. We thus take $\beta = 0.1$ for the tests performed on this shock tube.

Although this problem is unsteady, we apply the third formulation in order to see the evolution of the material and acoustic waves, and thus to understand better the effect of the iterative preconditioning.

We use an explicit scheme. The time step must be chosen inversely proportional to the largest eigenvalue of the system. At low Mach number when the equations are not preconditioned this eigenvalue is approximately equal to the speed of sound. Therefore, very small time steps must be used in order to see the acoustic wave moves. Moreover, the material waves, which are advancing much slower, do not change a lot during one time step. This is why a very large number of time steps can be necessary to see the evolution of the material waves when the speed of the waves is of very different orders of magnitude. This is illustrated in Figure 4, left column, where the acoustic wave moves quickly while the material wave almost does not move. In Figure 5 we show that after a certain time, the acoustic wave did several round trips while the material wave barely moved.

The effect of the iterative preconditioning is illustrated in Figure 4 in the right column. We note that the iterative preconditioning slows down the acoustic wave. After a certain time, we see that the material wave moved from $x = 0.5$ to $x = 0.6$ (see Figure 5).

4.5.4 NACA0012 airfoil

We are now interested in the 2D steady problem of a low Mach number flow around an airfoil with no incidence.

For this test case we chose a coarse mesh of 800 nodes (= vertices) composed of non-structured triangles.

With this problem, we test on the one hand the approximation, and on the other hand the efficiency of the explicit algorithm for the four formulations presented above.

Figure 6 shows the isovalues of the velocity at Mach 0.1, with and without preconditioning of the stabilisation term. The scheme is second-order accurate. From the approximation point of view, only the modification of the upwind fluxes enables us to obtain a solution of acceptable accuracy (taking into account the fact that the mesh is very coarse). The plots of the density convergence are presented in Figure 7 for each one of the four formulations given earlier. The CFL number is 2. The time step is local to accelerate the convergence to the steady state. From the algorithmic point of view, the standard explicit algorithm (Formulation 1) does not converge, and on the contrary presents a cycle limit. The totally preconditioned equations (Formulation 3) enable us to reach a residual decreased by seven orders of magnitude although in 7000 iterations which is still prohibitive.

4.6 Conclusion

It has been observed by van Leer and we showed it again in this section that when a Roe scheme is used the total preconditioning of the Euler equations (Formulation 3) achieves not only its first aim, improving the convergence towards the steady state solution, but moreover it improves the accuracy of the solution thanks to the modification of $\mid A_c\nu_x + B_c\nu_y \mid$ into $P_c^{-1} \mid P_c \left(A_c\nu_x + B_c\nu_y\right) \mid$.

In the sequel, we propose to use Formulation 4 because this method is consistent in time and thus applicable to unsteady cases. We will refer to the Roe scheme with the modification of Turkel as the Roe-Turkel scheme. It has been observed in the numerical experiments that this scheme is more accurate than the Roe scheme at low Mach number. A study of the consistency of the Roe scheme and the Roe-Turkel scheme will enable us to understand why.

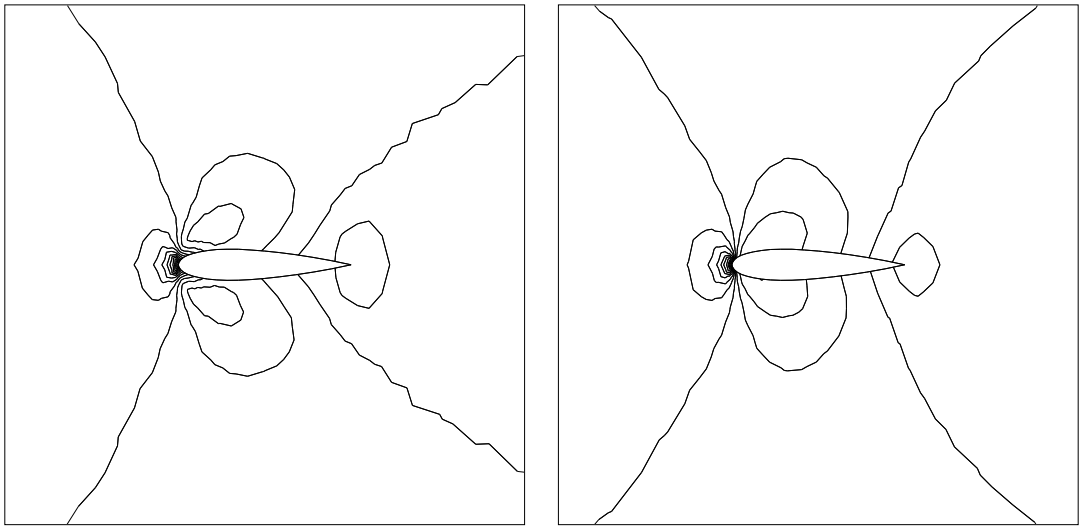

Figure 6: Isovalues of the velocity around a NACA0012 at $M_\infty = 0.1$. Without preconditioning of the stabilisation term (left; Formulations 1 and 2) and with (right; Formulations 3 and 4). Second-order accurate, without limiters. Interval between isovalues: 0.05. Min/Max: 0, 2.

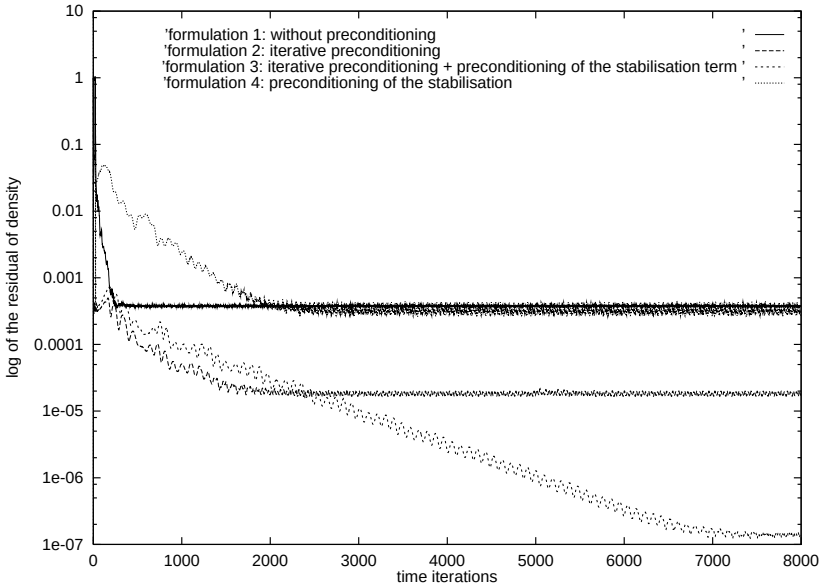

Figure 7: Convergence at second-order accuracy of the density according to the preconditioning of the equations, $M_\infty = 0.1$, explicit case.

5 Consistency study

Ideally, one would wish to show that the discretized solution converges towards the exact solution, $u_h \to u$. An intermediate step consists of showing the consistency of the scheme, that is $A_h\, u \to f$, in the case of a uniform mesh. It is with this kind of approach, very standard, that we propose to elucidate the behaviour of the Roe upwind scheme approximations, with and without modification of Turkel.

Two classes of upwind schemes can be distinguished; one is based on a Riemann solver approach (or FDS, flux difference splitting) and the other is based on a flux decomposition (FVS, flux vector splitting). We are interested here in the Roe scheme which relies on an approximate solution of the Riemann problem.

5.1 The Roe scheme

The first upwind scheme developed to solve a hyperbolic linear equation was the scheme of Courant-Isaacson-Rees (CIR) proposed in 1952 [23].

The Roe scheme [25] is a generalisation of the CIR scheme to a non-linear coupled system of hyperbolic equations. It consists of solving the following equation

$$\frac{\partial W}{\partial t} + \tilde{A}_c \frac{\partial W}{\partial x} = 0. \tag{62}$$

The discretised solution is piecewise constant on each control volume. At the cell interfaces, a Riemann problem must be solved. At the cell interface $i+1/2$, the matrix $\tilde{A}_c$ must satisfy the following properties

- $\delta F = \tilde{A}_c(W_L, W_R)\, \delta W$ où $\delta = ()_R - ()_L$
- $\tilde{A}_c(W, W) = A_c(W)$
- $\tilde{A}_c$ has a set of real eigenvalues and linearly independent eigenvectors.

The numerical fluxes can be written as

$$F_{i+1/2} = \frac{1}{2}\left[F(W_i) + F(W_{i+1})\right] - \frac{1}{2} \mid (\tilde{A}_c)_{i+1/2} \mid (W_{i+1} - W_i) \tag{63}$$

where $(\tilde{A}_c)_{i+1/2}$ is the Jacobian matrix A_c evaluated at a state defined by the Roe average

$$(\cdot)_{i+1/2} = \frac{\sqrt{\rho_i}(\cdot) + \sqrt{\rho_{i+1}}(\cdot)}{\sqrt{\rho_i} + \sqrt{\rho_{i+1}}}, \tag{64}$$

and $\rho_{i+1/2} = \sqrt{\rho_i\, \rho_{i+1}}$.

5.2 The Roe-Turkel scheme

The only difference of the "Roe-Turkel" scheme as we define it in comparison with the Roe scheme is that the stabilisation term $\mid (\tilde{A}_c)_{i+1/2} \mid$ is replaced by $P_c^{-1} \mid P_c\,(\tilde{A}_c)_{i+1/2} \mid$. This modification was detailed in Section 4.4. For the Roe-Turkel scheme, β is taken as the order of the Mach number (when $\beta = 1$ we obtain the Roe scheme).

Lemma 1: When the mesh size h and the Mach number M tend to zero, the truncation error of the first-order accurate Roe scheme, applicable to the momentum equations, is only $O(\frac{h}{M})$ whilst that of the Roe-Turkel scheme is O(h) independently of the Mach number.

Proof:

For simplicity, we present here the 1D case. In 2D and 3D, we obtain similar results.

We show first that the accuracy of the Roe-Turkel scheme depends on the order of magnitude of the matrix $P_c^{-1} \mid P_c\,A_c \mid$, then we compute this matrix and its magnitude at low Mach number.

- **The accuracy of the Roe-Turkel scheme depends on the order of magnitude of the entries of the matrix $P_c^{-1} \mid P_c\,A_c \mid$**

  With the Roe scheme, the term $\tilde{A}_c \dfrac{\partial W}{\partial x}$ is discretised by

  $$\frac{F_{i+1/2} - F_{i-1/2}}{h} \tag{65}$$

  where

  $$\begin{aligned} F_{i+1/2} &= \frac{F(W_{i+1}) + F(W_i)}{2} - \frac{1}{2} \mid (\tilde{A}_c)_{i+1/2} \mid (W_{i+1} - W_i), \\ F_{i-1/2} &= \frac{F(W_i) + F(W_{i-1})}{2} - \frac{1}{2} \mid (\tilde{A}_c)_{i-1/2} \mid (W_i - W_{i-1}), \end{aligned} \tag{66}$$

  where $(\tilde{A}_c)_{i+1/2} = \tilde{A}_c(W_i, W_{i+1})$ and $(\tilde{A}_c)_{i-1/2} = \tilde{A}_c(W_{i-1}, W_i)$ are the Roe matrices.

  The approximation of $\tilde{A}_c \dfrac{\partial W}{\partial x}$ by the Roe scheme can be equally written as

  $$\begin{aligned} \frac{F_{i+1/2} - F_{i-1/2}}{h} = & \frac{F(W_{i+1}) - F(W_{i-1})}{2\,h} \\ & - \frac{\mid (\tilde{A}_c)_{i+1/2} \mid (W_{i+1} - W_i) - \mid (\tilde{A}_c)_{i-1/2} \mid (W_i - W_{i-1})}{2\,h}. \end{aligned} \tag{67}$$

Let us perform a Taylor development of the second order on the Roe matrices by supposing that their eigenvalues do not change sign from $i-1$ to i and from i to $i+1$ (we assume that this change is seldom enough to be negligible); we obtain

$$\begin{aligned} \mid (A_c)_{i-1/2} \mid &= \mid (A_c)_i \mid - \frac{h}{2} \frac{\partial \mid (A_c)_i \mid}{\partial x} + O(h^2) \\ \mid (A_c)_{i+1/2} \mid &= \mid (A_c)_i \mid + \frac{h}{2} \frac{\partial \mid (A_c)_i \mid}{\partial x} + O(h^2); \end{aligned} \tag{68}$$

thus,

$$\frac{F_{i+1/2} - F_{i-1/2}}{h} = \frac{F(W_{i+1}) - F(W_{i-1})}{2\,h} - \frac{h}{2} \mid (A_c)_i \mid \frac{W_{i+1} - 2W_i + W_{i-1}}{h^2} + O(h^2). \tag{69}$$

We conclude that the truncation error of the Roe scheme at point i is given by

$$\epsilon_i = \frac{h}{2} \mid (A_c)_i \mid \frac{\partial^2 W}{\partial x^2} + O(h^2). \tag{70}$$

In a similar way, the truncation error of the Roe-Turkel scheme at point i is given by

$$\epsilon_i = \frac{h}{2} P_c^{-1} \mid P_c\,(A_c)_i \mid \frac{\partial^2 W}{\partial x^2} + O(h^2). \tag{71}$$

Therefore, the truncation error of the Roe-Turkel scheme is in $O(P_c^{-1} \mid P_c\,(A_c)_i \mid h)$,

$$\epsilon_i < K\,P_c^{-1} \mid P_c\,(A_c)_i \mid h \tag{72}$$

where K is a constant.

- **Computation of the matrix $P_c^{-1} \mid P_c\,A_c \mid$**

  In a similar way to the 2D case presented in Section 4.4; we can write

$$P_c^{-1} \mid P_c\,A_c \mid = T_{rt}^g \mid \Lambda \mid T_{rt}^d \tag{73}$$

where

$$\Lambda = Diag(\lambda_1, \lambda_2, \lambda_3) \tag{74}$$

with

$$\left\{ \begin{array}{rcl} \lambda_1 & = & u \\ \lambda_{2,3} & = & \frac{1}{2}\left[(\beta^2+1)u \pm \sqrt{(\beta^2+1)^2\,u^2 - 4\,\beta^2\,(u^2-a^2)} \;\right]. \end{array} \right. \tag{75}$$

The eigenvector matrices T^g_{rt} and T^d_{rt} are given by

$$T^g_{rt} = \begin{bmatrix} 1 & 1 & 1 \\ u & u+r & u+s \\ \dfrac{u^2}{2} & H+ru & H+su \end{bmatrix}, \tag{76}$$

$$T^d_{rt} = \begin{bmatrix} 1-\dfrac{(\gamma-1)}{a^2}\dfrac{u^2}{2} & \dfrac{\gamma-1}{a^2}u & -\dfrac{\gamma-1}{a^2} \\ \dfrac{s\dfrac{u^2}{2}(\gamma-1)+\beta^2 a^2 u}{2\beta^2 a^2 t} & -\dfrac{su(\gamma-1)+\beta^2 a^2}{2\beta^2 a^2 t} & \dfrac{s(\gamma-1)}{2\beta^2 a^2 t} \\ -\dfrac{r\dfrac{u^2}{2}(\gamma-1)+\beta^2 a^2 u}{2\beta^2 a^2 t} & \dfrac{ru(\gamma-1)+\beta^2 a^2}{2\beta^2 a^2 t} & -\dfrac{r(\gamma-1)}{2\beta^2 a^2 t} \end{bmatrix}, \tag{77}$$

where

$$r = \lambda_2 - \lambda_1\,\beta^2, \quad s = \lambda_3 - \lambda_1\,\beta^2, \quad t = \frac{\lambda_3-\lambda_2}{2}. \tag{78}$$

Remark:
For the Roe scheme we have $\lambda_{2,3} = u \pm a$, $r = a$, $s = -a$ and $t = -a$, these variables are of the order of the speed of sound whilst for the Roe-Turkel scheme they are of the order of the flow velocity u. ■

Thus, the entries of the matrix $P_c^{-1} \mid P_c A_c \mid$ are

$$\begin{aligned}
c_{11} &= \mid \lambda_1 \mid \left(1 - \frac{(\gamma-1)\,M^2}{2}\right) + (\gamma-1)\,\frac{M^2}{2}\,d_2 + u\,d_1 \\
c_{21} &= \mid \lambda_1 \mid \left(1 - \frac{(\gamma-1)\,M^2}{2}\right) u + (\gamma-1)\,\frac{M^2}{2}\,d_4 + u\,d_3 \\
c_{31} &= \mid \lambda_1 \mid \left(1 - \frac{(\gamma-1)\,M^2}{2}\right) \frac{u^2}{2} + (\gamma-1)\,\frac{M^2}{2}\,d_6 + u\,d_5
\end{aligned}$$

$$\begin{aligned}
c_{12} &= \mid \lambda_1 \mid \frac{(\gamma-1)}{u}\,M^2 - (\gamma-1)\,\frac{M^2}{u}\,d_2 - d_1 \\
c_{22} &= \mid \lambda_1 \mid (\gamma-1)\,M^2 - (\gamma-1)\,\frac{M^2}{u}\,d_4 - d_3 \\
c_{32} &= \mid \lambda_1 \mid \frac{(\gamma-1)}{2}\,M^2\,u - (\gamma-1)\,\frac{M^2}{u}\,d_6 - d_5
\end{aligned} \tag{79}$$

$$\begin{aligned}
c_{13} &= -\mid \lambda_1 \mid \frac{(\gamma-1)}{u^2}\,M^2 + \frac{(\gamma-1)}{u^2}\,M^2\,d_2 \\
c_{23} &= -\mid \lambda_1 \mid \frac{(\gamma-1)}{u}\,M^2 + \frac{(\gamma-1)}{u^2}\,M^2\,d_4 \\
c_{33} &= -\mid \lambda_1 \mid \frac{(\gamma-1)}{2}\,M^2 + \frac{(\gamma-1)}{u^2}\,M^2\,d_6
\end{aligned}$$

where

$$\begin{aligned}
d_1 &= \frac{\mid \lambda_2 \mid - \mid \lambda_3 \mid}{2\,t} \\
d_2 &= \frac{s \mid \lambda_2 \mid - r \mid \lambda_3 \mid}{2\,\beta^2\,t} \\
d_3 &= \frac{(u+r) \mid \lambda_2 \mid - (u+s) \mid \lambda_3 \mid}{2\,t} \\
d_4 &= \frac{s\,(u+r) \mid \lambda_2 \mid - r\,(u+s) \mid \lambda_3 \mid}{2\,\beta^2\,t} \\
d_5 &= \frac{(H + r\,u) \mid \lambda_2 \mid - (H + s\,u) \mid \lambda_3 \mid}{2\,t} \\
d_6 &= \frac{s\,(H + r\,u) \mid \lambda_2 \mid - r\,(H + s\,u) \mid \lambda_3 \mid}{2\,\beta^2\,t}.
\end{aligned} \tag{80}$$

The absolute values are computed by assuming $\mid \lambda_1 \mid = \lambda_1$, $\mid \lambda_2 \mid = \lambda_2$ and $\mid \lambda_3 \mid = -\lambda_3$.

The quantities r, s and t, and the eigenvalues are written in terms of u, M and β (see equations (75) and (78)). They are then substituted into (80). Thus

$$\begin{aligned}
d_1 &= -\frac{\beta^2+1}{\sqrt{X}} \\
d_2 &= \frac{u\,(\beta^2-1+2/M^2)}{\sqrt{X}} \\
d_3 &= -\frac{2\,u}{\sqrt{X}}\,(1+\frac{\beta^2}{M^2}) \\
d_4 &= \frac{u^2}{\sqrt{X}}\,(\beta^2-1+3/M^2+\beta^2/M^2) \\
d_5 &= \frac{1}{\sqrt{X}}\left[-H\,(\beta^2+1)+u^2\,(\beta^2-1-\frac{2\,\beta^2}{M^2})\right] \\
d_6 &= \frac{u}{\sqrt{X}}\left[H\,(\beta^2-1+2/M^2)+\frac{u^2}{M^2}(\beta^2+1)\right],
\end{aligned}$$

where $X = \beta^4 - 2\,\beta^2 + 1 + 4\frac{\beta^2}{M^2}$.

For the particular case where $\beta = M$ and $\beta = 1$ the above terms are given in Table 1. In the third column of this table the terms of (80) are given when the absolute value of the eigenvalues is suppressed; we thus obtain the corresponding terms of matrix A_c.

Remark:
We observe that the terms of Table 1 occurring in $P_c^{-1} \mid P_c\,A_c \mid$, $\mid A_c \mid$ and A_c are not of the same order of magnitude when $M \to 0$, which is why these three matrices have entries of different magnitudes at low Mach number. ■

- **Order of magnitude of the entries of the matrix $P_c^{-1} \mid P_c\,A_c \mid$ when the Mach number tends to zero**

We assume that the Mach number tends to zero. In reality, when the Mach number goes to zero, the flow velocity is low and the speed of sound remains bounded. For the analysis of the order of magnitude of the truncation error performed in this section, we consider the equations adimensioned in the following manner (coherent with the options of the numerical experiments presented further): the flow velocity is of order 1 at infinity, and the speed of sound tends to infinity when the Mach goes to zero. The density is also bounded to 1 at infinity and changes little at low Mach number. Table 2 presents the order of magnitude of variables occurring the matrix $P_c^{-1} \mid P_c\,A_c \mid$.

	$\beta = M$	$\beta = 1$	*corresponding terms in* A_c
d_1	$-\dfrac{M^2+1}{\sqrt{X}}$	$-M$	-1
d_2	$\dfrac{u\,(M^2-1+2/M^2)}{\sqrt{X}}$	$\dfrac{u}{M}$	u
d_3	$-\dfrac{4\,u}{\sqrt{X}}$	$-u\,(M+1/M)$	$-2\,u$
d_4	$\dfrac{u^2}{\sqrt{X}}\,(M^2+3/M^2)$	$2\,\dfrac{u^2}{M}$	$u^2\,(1+1/M^2)$
d_5	$\dfrac{1}{\sqrt{X}}\,\left[-H\,(M^2+1)+u^2\,(M^2-3)\right]$	$-\left(M\,H+\dfrac{u^2}{M}\right)$	$-(H+u^2)$
d_6	$\dfrac{u}{\sqrt{X}}\left[H\,(M^2-1+2/M^2)+\dfrac{u^2}{M^2}(M^2+1)\right]$	$\dfrac{u}{M}\,(H+u^2)$	$u\left(H+\dfrac{u^2}{M^2}\right)$
X	$M^4-2\,M^2+5$	$\dfrac{4}{M^2}$	

Table 1: Terms occurring in the matrix $P_c^{-1} \mid P_c\,A_c \mid$ (case where $\beta = M$ and $\beta = 1$) and in the matrix A_c. We observe that when the Mach number is low the stabilisation terms of the Roe scheme (2^{nd} column) are not of the same order as the fluxes (3^{rd} column). On the other hand the stabilisation terms of the Roe-Turkel scheme (1^{st} column) are of the same order as the fluxes (3^{rd} column).

a	u, ρ, λ_1	$\lambda_2, \lambda_3, r, s, t, \beta\, a$	H, p, E
$O\left(\frac{1}{M}\right)$	$O(1)$	$O\left(\frac{1}{M}\right)$ $si\ \beta = 1$ $O(1)$ $si\ \beta = M$	$O\left(\frac{1}{M^2}\right)$

Table 2: Order of magnitude of variables occurring in the stabilisation term

Now, we substitute the terms (80) into (79) and we use the order of magnitude of the variables given in Table 2 to simplify by neglectable terms.

For the Roe scheme at low Mach number ($\beta = 1$ and $M \sim 0$) we obtain

$$| A_c | \sim \begin{pmatrix} u & M\,(2-\gamma) & (\gamma-1)\,\dfrac{M}{u} \\ -\dfrac{u^2}{M} & \dfrac{u}{M} & 2\,(\gamma-1)\,M \\ \dfrac{(\gamma-3)\,u\,M\,H}{2} - \dfrac{u^3}{M} & (2-\gamma)\,M\,H + \dfrac{u^2}{M} & \dfrac{\gamma-1}{u}\,M\,H \end{pmatrix}. \tag{81}$$

For the Roe-Turkel scheme at low Mach number ($\beta = M$ and $M \sim 0$) we obtain

$$P_c^{-1}\ |\ P_c\, A_c\ | \sim \begin{pmatrix} u\,(1+\dfrac{\gamma-2}{\sqrt{5}}) & \dfrac{3-2\,\gamma}{\sqrt{5}} & \dfrac{2\,(\gamma-1)}{u\,\sqrt{5}} \\ u^2\,\left(1+\dfrac{3\,\gamma-11}{2\,\sqrt{5}}\right) & \dfrac{(-3\,\gamma+7)\,u}{\sqrt{5}} & \dfrac{3\,(\gamma-1)}{\sqrt{5}} \\ \dfrac{(\gamma-2)\,u\,H}{\sqrt{5}} & \dfrac{(3-2\,\gamma)\,H}{\sqrt{5}} & \dfrac{2\,(\gamma-1)\,H}{u\,\sqrt{5}} \end{pmatrix}. \tag{82}$$

Thus, for the Roe scheme we have

$$| A_c | \sim \begin{bmatrix} O(1) & O(M) & O(M) \\ O\left(\frac{1}{M}\right) & O\left(\frac{1}{M}\right) & O(M) \\ O\left(\frac{1}{M}\right) & O\left(\frac{1}{M}\right) & O\left(\frac{1}{M}\right) \end{bmatrix}. \tag{83}$$

and for the Roe-Turkel scheme we have

$$P_c^{-1} \mid P_c A_c \mid \sim \begin{bmatrix} O(1) & O(1) & O(1) \\ O(1) & O(1) & O(1) \\ O\left(\frac{1}{M^2}\right) & O\left(\frac{1}{M^2}\right) & O\left(\frac{1}{M^2}\right) \end{bmatrix} \tag{84}$$

while we have

$$A_c \sim \begin{bmatrix} 0 & O(1) & 0 \\ O(1) & O(1) & O(1) \\ O(\frac{1}{M^2}) & O(\frac{1}{M^2}) & O(1) \end{bmatrix}. \tag{85}$$

- **Conclusion**

  For the momentum equation the truncation error is thus $O(\frac{h}{M})$ for the Roe scheme and $O(h)$ for the Roe-Turkel scheme. This completes the proof. ■

With the Roe scheme the wave speeds are mixed up in the stabilisation terms. Therefore, at low Mach number, the stabilisation terms are not of the same order of magnitude as the flux terms because the wave speeds are then of different magnitudes. The Roe-Turkel scheme cures this fault by reducing the distance between the wave speeds.

This study shows that for a same mesh the Roe scheme will present numerical errors larger and larger when the Mach tends to zero, whilst the correction of Turkel enables an accurate and consistent approximation independently of the Mach number.

5.3 Remark concerning other upwind schemes

It seems that, as in the Roe scheme, the other flux difference splittings (Steger-Warming, Osher..) suffer from the fault of mixing up the wave speeds in the stabilisation terms. In the case where the wave speeds are of different magnitudes, as is the case at low Mach number, this can lead to inaccuracy. This inaccuracy should be corrected by a preconditioning of the diffusive terms of Roe-Turkel type.

5.4 Numerical illustration

At low Mach number, this is when the speed of sound is large compared to the flow velocity and at constant temperature the density is almost constant. The solutions of the Euler compressible equations approximate the ones of the incompressible equations [19]. A correctly adimensioned solution is in particular almost independent of the Mach number

$$If \quad M << 1 \quad then \quad u^{compressible} \sim u^{incompressible}. \tag{86}$$

Observing the isovalues of the velocity we show in the context of an eulerian flow around a NACA0012 airfoil that this property is verified for the Roe-Turkel scheme whilst it is not verified for the Roe scheme.

We carried out two sets of numerical experiments illustrating Lemma 1.

The first set of experiments illustrates the fact that the standard Roe and the Roe-Turkel schemes are both consistent in the usual meaning of the term. For a fixed Mach number ($M_\infty = 0.1$), we made the mesh size h tend to zero. The converged solution (normalised residual of ρu to 10^{-8}) is computed (Figure 8) on meshes of 800 nodes, 3114 nodes and 12284 nodes. We observe that the solutions obtained with both schemes converge towards a plausible solution when the space step h tends to zero. Both schemes seem to be consistent in space.

The second set of experiments illustrates the fact that the standard Roe scheme is not *uniformly consistent* with regard to the Mach number whilst the Roe-Turkel scheme is in the following meaning

$$\exists\eta \quad \forall\varepsilon \quad \exists h \quad : \quad \forall M < \eta \quad | A_h^M u^M - f_h | < \varepsilon, \tag{87}$$

where u^M is the continuous solution and A_h^M the corresponding discrete operator.

For a fixed mesh (800 nodes), we tend the Mach number to zero. The converged solution (normalised residual of ρu to 10^{-8}) is computed (Figure 9) for Mach 10^{-1}, Mach 10^{-2} and Mach 10^{-3}. For the Roe scheme we observe that on a same mesh the lower the Mach number, the more inaccurate is the solution. On the other hand for the Roe-Turkel scheme the solution remains almost the same (velocity modulus contours are then indistinguishable) independently for all Mach numbers between 0.1 and 10^{-6}. The Roe scheme is undoubtedly a convergent scheme, but the effort necessary in mesh fineness for a given accuracy grows inversely proportionally with the Mach number. In contrast, the uniform consistence with regard to the Mach number of the Roe-Turkel is observed.

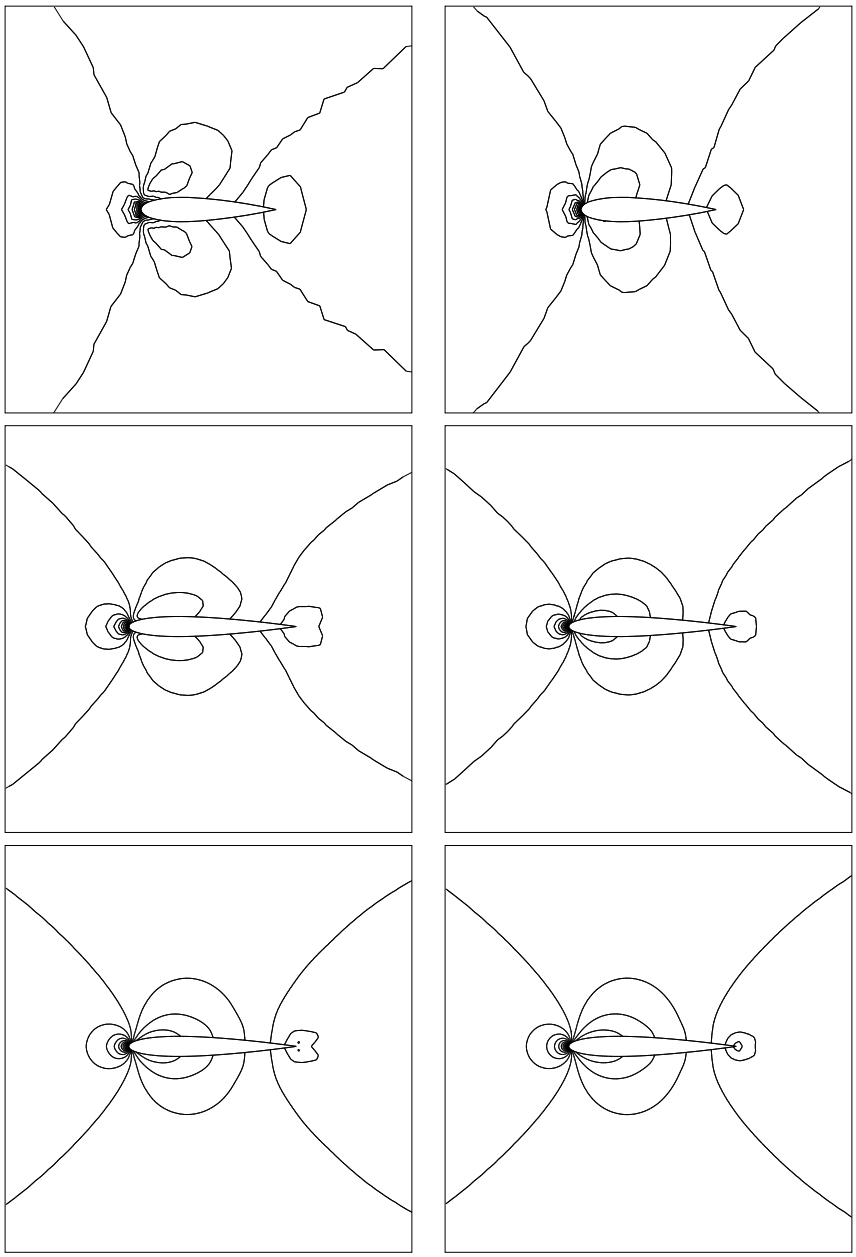

Figure 8: Isovalues of the velocity at second-order accuracy, without limiters, for $M_\infty = 0.1$ with the Roe scheme (left) and the Roe-Turkel scheme (right) on a mesh of 800 nodes (top), 3114 nodes (middle) and 12284 nodes (bottom). Interval between isovalues: 0.05. Min/Max: 0, 2.

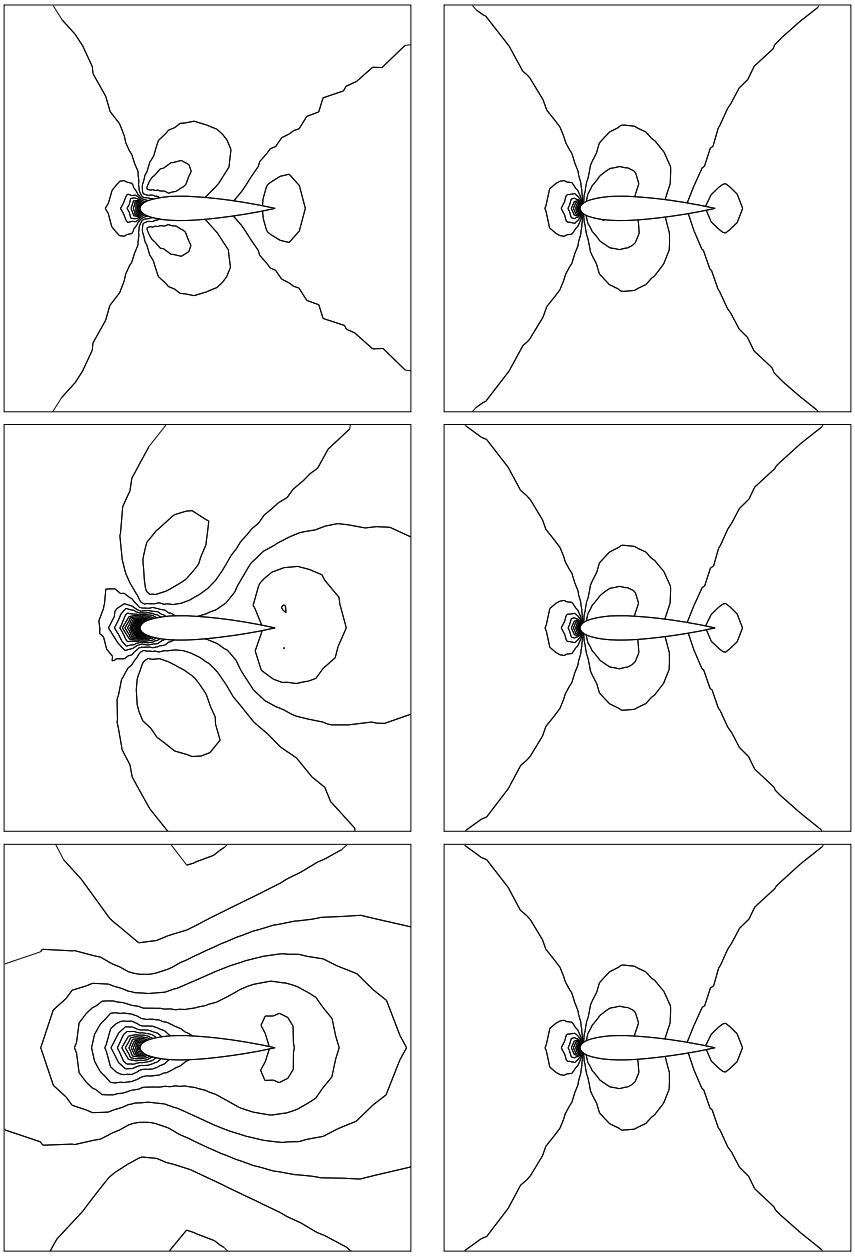

Figure 9: Isovalues of the velocity at second-order accuracy, without limiters, on a coarse mesh (800 nodes) for $M_\infty = 0.1$ (top), $M_\infty = 0.01$ (middle), $M_\infty = 0.001$ (bottom) with the Roe scheme (left) and the Roe-Turkel scheme (right). For $10^{-6} < M_\infty < 0.1$ it has been observed with the Roe-Turkel scheme that the velocity contours change only a little. Interval between isovalues: 0.05. Min/Max: 0, 2.

6 Implicit scheme: iterative convergence study

In numerous unsteady applications, and obviously for the search of steady state solutions, it is not necessary to accurately compute the acoustic details. Therefore, we now propose to build an implicit scheme capable of being used with large time steps. First, we present a spatial first-order accurate implicit scheme, then, we present a spatial second-order accurate implicit scheme to compute steady state solutions.

The implicit backward Euler scheme can be written as

$$\mathcal{M}\,\frac{W^{n+1}-W^n}{\Delta t}+\Psi(W^{n+1})=0, \tag{88}$$

where $\mathcal{M}=diag(area(C_i))$ is the mass matrix, $area(C_i)$ being the area of the cell C_i and where $\Psi(W^{n+1})$ is the function of the numerical fluxes at time $n+1$.

In linearising the function of the numerical fluxes $\Psi(W^{n+1})$, when it is differentiable, thanks to a Taylor development of the first order in time, equation (88) becomes

$$\mathcal{M}\,\frac{W^{n+1}-W^n}{\Delta t}+\Psi(W^n)+\Psi'(W^n)\,(W^{n+1}-W^n)=0, \tag{89}$$

where $\Psi'(W^n)$ denotes the Jacobian of $\Psi(W^n)$.

The linearised implicit scheme obtained then can be written in the following δ **form**

$$\left(\frac{\mathcal{M}}{\Delta t}+\Psi'(W^n)\right)\,\delta W_I^{n+1}=-\Psi(W^n), \tag{90}$$

where $\delta W_I^{n+1}=W^{n+1}-W^n$.

In order to distinguish the two phases of the algorithm (physical phase and mathematical phase) we write equation (90) under the following form

$$\begin{cases} \mathcal{M}\,\delta W_E^{n+1}=-\Delta t\,\Psi(W^n) \\ \mathcal{K}(W^n)\,\delta W_I^{n+1}=\mathcal{M}\,\dfrac{\delta W_E^{n+1}}{\Delta t}, \end{cases} \tag{91}$$

where $\delta W_I^{n+1}=\delta W_E^{n+1}=W^{n+1}-W^n$ and $\mathcal{K}(W^n)=\dfrac{\mathcal{M}}{\Delta t}+\Psi'(W^n)$.

The **physical phase** (or explicit phase) consists of determining the term $\mathcal{M}\,\delta W_E$ by solving the same equation as for an explicit scheme.

The **mathematical phase** (or implicit phase) consists of building the implicit matrix, $\mathcal{K}(W^n)$, then solving with an iterative method the following linear system

$$\mathcal{K}\,\delta W_I = \mathcal{M}\,\frac{\delta W_E}{\Delta t}, \tag{92}$$

where the matrix $\mathcal{K}$ plays the role of a preconditioner when only the steady state is of interest.

At **the next time step**, the solution is determined by

$$W^{n+1} = W^n + \delta W_I. \tag{93}$$

6.1 Spatial first-order accurate implicit scheme

For a first order upwind scheme the flux $\Psi^{(1)}(W^n)$ is computed by summation of the elementary fluxes for each edge of the vertex i and j by

$$\Phi^{(1)}_{seg}(W_i^n, W_j^n, \vec{\nu}) = \frac{\mathcal{F}(W_i^n, \vec{\nu}) + \mathcal{F}(W_j^n, \vec{\nu})}{2} - d(W_i^n, W_j^n, \vec{\nu}) \tag{94}$$

$$d(W_i^n, W_j^n, \vec{\nu}) = \frac{1}{2} \mid A(W_i^n, W_j^n, \vec{\nu}) \mid (W_j^n - W_i^n), \tag{95}$$

where $A = \dfrac{\partial F(W)}{\partial W}\,\nu_x + \dfrac{\partial G(W)}{\partial W}\,\nu_y$ and $\mathcal{F} = F(W)\,\nu_x + G(W)\,\nu_y$, ν_x and ν_y are the components of the outward normal to the interface ∂C_{ij} separating nodes i and j (see Figure 2).

In order to avoid the computation of a complicated Jacobian, which is moreover not unique since the Roe flux is non-differentiable, the Jacobian matrix $(\Psi^{(1)})'(W^n)$ is approximated by a linear operator denoted $\mathcal{A}^{(1)}(W^n)$. This operator is obtained by freezing the Roe matrix $A(W^n)$ in equation (94) when the Jacobian is computed

$$(\Psi^{(1)})'(W^n) \approx \mathcal{A}^{(1)}(W^n) = \sum_{seg} \frac{1}{2}\mathcal{F}'(W_i^n, \vec{\nu}) - \frac{1}{2} \mid A(W_i^n, W_j^n, \vec{\nu}) \mid, \tag{96}$$

where the sum is taken over all the edges ij joining two vertices.

After the approximation of the Jacobian matrix and at first order, equation (90) becomes

$$\left(\frac{\mathcal{M}}{\Delta t} + \mathcal{A}^{(1)}(W^n)\right)\,\delta W_I^{n+1} = -\Psi^{(1)}(W^n). \tag{97}$$

Remark:
When $\Delta t \to \infty$ equation (97) becomes

$$\mathcal{A}^{(1)}(W^n)\,(W^{n+1} - W^n) = -\Psi^{(1)}(W^n). \tag{98}$$

We recognise a modified Newton method. Indeed, if $\mathcal{A}^{(1)}$ was the exact Jacobian of $\Psi^{(1)}$, that is $\mathcal{A}^{(1)} = (\Psi^{(1)})'$, then the algorithm (98) would reduce to the Newton method which enables us to compute the zeros of $\Psi(W^n) = 0$ and whose convergence we know to be quadratic. Thus, we expect that the algorithm (98) converges quickly. ■

6.2 Convergence study of the Defect-Correction method

We previously distinguished the physical phase which controls the accuracy of the steady state solution and the mathematical phase which controls the convergence of the evolution method. In order to obtain a spatial second-order accurate scheme, the right hand side of the equation for the physical phase is discretised by a first derivative accurate at second order. In order to conserve the spatial accuracy for unsteady computations, a second-order accurate discretisation should also be used in the mathematical phase [21]. However, this modification being complex and expensive, here we just compute steady state solutions with a second-order accurate explicit flux and implicitly precondition with a first order operator. This method is a particular case of the DeC (Defect-Correction) method which consists of preconditioning a complex non-linear operator with a simpler non-linear preconditioner

$$\left\{ \begin{array}{l} \mathcal{M}\,\delta W_E^{n+1} = -\Delta t\,\Psi^{(2)}(W^n) \\ \mathcal{K}^{(1)}(W^n)\,\delta W_I^{n+1} = \mathcal{M}\,\dfrac{\delta W_E^{n+1}}{\Delta t}, \end{array} \right. \tag{99}$$

where $\delta W_I^{n+1} = \delta W_E^{n+1} = W^{n+1} - W^n$, $\mathcal{K}^{(1)}(W^n) = \dfrac{\mathcal{M}}{\Delta t} + \mathcal{A}^{(1)}(W^n)$. For computing the second-order accurate numerical flux, $\Psi^{(2)}(W^n)$, the MUSCL technique described in Section 4.5.1 is employed.

The rate of the iterative convergence of the DeC method towards the steady state solution was studied by Désidéri and Hemker [9] [10]. They showed that a scheme built on a central or fully-upwind discretisation of the second-order accurate fluxes does not converge, but for a combination of these two discretisations, the best rate of convergence is $\dfrac{1}{2}$.

The modification of Turkel in the physical phase enables us to improve the accuracy of the solution at low Mach number as we showed it previously. The modification of Turkel in the mathematical phase is necessary at low Mach number for the preconditioning to be stable and efficient.

From Lemma 1, the equivalent equation of the first-order upwind scheme, can be written for the quantities w advected at speed u (ex: the entropy)

$$u\,w_x = \eta\,\frac{h}{2}\;|\,u\,|\;w_{xx} + \cdots \tag{100}$$

where h is the time step, $\eta = O(1)$ for the Roe-Turkel scheme (η is independent of the Mach number), and $\eta = O(\frac{1}{M})$ for the Roe scheme when the Mach number tends to zero. The algebraic analysis of Désidéri and Hemker was for the case where $\eta = 1$ and thus is no longer valid for the Roe scheme at low Mach number. We generalise here the Fourier analysis of the convergence in the general case where η is arbitrary.

Lemma 2: The Implicit Defect-Correction schemes built on (100) have a convergence predicted by Fourier analysis which is mesh independent and which depends only on η. It is more than 0.5 for values of η greater than 1; the convergence ratio degrades to 1 as η increases to infinity.

Proof:

We perform a 1D scalar analysis since it predicts already well enough the convergence observed in the numerical experiments with the 2D Euler equations.

The scalar 1D model of the periodic advection equation can be written as

$$\left\{ \begin{array}{rcl} u_t + a\, u_x & = & 0 \\ u(x+2\pi,t) & = & u(x,t), \quad \text{(periodicity in space)} \end{array} \right. \tag{101}$$

We consider the discretisation of a line.

- **Mathematical phase: order 1**

  The first derivative in space, u_x, is discretised by a first upwind derivative accurate at first order, $\delta_{x,1} = Trid(-1,1,0)$. This derivative can be written as the sum of a first centred derivative accurate at the second order and a second centred derivative accurate at the first order

  $$\delta_{x,1} = \frac{1}{2} Trid(-1,0,1) + \frac{1}{2} Trid(-1,2,-1). \tag{102}$$

  We consider here the general case where η varies. The second derivative of equation (102) is thus multiplied by the parameter η_1 in the following way

  $$\delta_{x,1} = \frac{1}{2} Trid(-1,0,1) + \frac{1}{2} \eta_1 Trid(-1,2,-1).$$

- **Physical phase: order 2**

  The spatial first derivative, u_x, can be discretised by a centred derivative accurate at the second order or by a fully upwind derivative accurate at the second order. We consider here a linear combination of these two discetisations whose upwinding rate is controled by parameter $\bar{\beta}$

  $$\delta_{x,2} = (1-\bar{\beta})\, Trid(-\frac{1}{2},0,\frac{1}{2}) + \bar{\beta}\, Penta(\frac{1}{2},-2,\frac{3}{2},0,0). \tag{103}$$

Note that in this section $\bar{\beta}$ denotes the upwinding parameter for spatial second order accuracy (β - scheme) and no longer the parameter of Turkel.

The fully upwind derivative of equation (103) can be written as the sum of a third centred derivative and a fourth centred derivative. The parameter η_2 is inserted in front of the fourth derivative

$$\begin{aligned}\delta_{x,2} \;=\; & \frac{1}{2}\,Penta(0,-1,0,1,0) \\ & + \frac{1}{2}\,\bar{\beta}\,\left[\left\{\frac{1}{2}\,Penta(1,-3,3,-1,0) \;+\; \frac{1}{2}\,Penta(0,1,-3,3,-1)\right\}\right. \\ & \left. +\; \eta_2\,\left\{\frac{1}{2}\,Penta(1,-3,3,-1,0) - \frac{1}{2}\,Penta(0,1,-3,3,-1)\right\}\right].\end{aligned}$$

The stability study of the DeC method is performed with **Fourier analysis** which consists of developing the numerical solution into discrete Fourier series. We note u_i^n the discrete Fourier modes at time n and at discretisation point i. We have

$$u_j^n = e^{-j\,i\,\omega}$$

where $\omega = 2\,k\,\pi\,h$ with $k = 0, ..., N$ and $j^2 = -1$.

Details of the following computations are given in Appendix B .

Fourier transform for the first-order operator is given by

$$F(\delta_{x,1}) = 2\,j\,\sin\frac{\omega}{2}\,\left(\cos\frac{\omega}{2} - j\,\eta_1\,\sin\frac{\omega}{2}\right),$$

and for the second-order operator by

$$F(\delta_{x,2}) = 2\,j\,\sin\frac{\omega}{2}\,\left(\cos\frac{\omega}{2} + 2\,\bar{\beta}\,\sin^2\frac{\omega}{2}\,\left[\cos\frac{\omega}{2} - \eta_2\,j\,\sin\frac{\omega}{2}\right]\right).$$

We denote G_k the amplification operator which links state n to state $n+1$.

The scheme is stable if the von Neumann criterion is verified

$$\max_{k\in[0,N]} \mid G_k(\Delta t) \mid \leq 1.$$

The amplification operator when $t \to \infty$ is given by

$$G_\infty = I - \left(\delta_{x,1}^u\right)^{-1}\,\delta_{x,2}^{\bar{\beta}}.$$

Thus

$$F(G_\infty) = I - (F(\delta_{x,1})(\omega))^{-1}\ F(\delta^{\bar\beta}_{x,2})(\omega).$$

The modulus of $F(G_\infty)$ can be written as

$$\mid F(G_\infty) \mid = \frac{\sqrt{\sin^4\frac{\omega}{2}\left(\eta_1^2 - 2\bar\beta\,(\cos^2\frac{\omega}{2} + \eta_1\,\eta_2\sin^2\frac{\omega}{2})\right)^2 + \eta_2^2\,\cos^2\frac{\omega}{2}\,\sin^2\frac{\omega}{2}}}{\cos^2\frac{\omega}{2} + \eta_1^2\sin^2\frac{\omega}{2}}.$$

Thus

$$S = \sup_{\omega\in\{-\pi/h,\pi/h\}} \mid F(G_\infty) \mid = \sup_{s\in(0,1)} \sqrt{\frac{\eta_1^2 + 4\,\bar\beta\,(\bar\beta - \eta_1\,\eta_2)\,s + 4\,\bar\beta^2(\eta_2^2-1)\,s^2}{1 + (\eta_1^2-1)\,s}\,s}$$

where $s = \sin^2\frac{\omega}{2}$.

The function S is plotted in Figure 10 when the same scheme is employed in the two phases of the algorithm (η_1=η_2=η). For the Roe scheme ($\eta = 1$) we find $\frac{1}{2}$ for the rate of convergence as in the analysis of Hemker and Desideri. For the Roe-Turkel scheme ($\eta = \frac{1}{M}$), the larger the value of η (the lower the Mach number), the more the convergence rate degrades. For $\eta \to \infty$ we have $S \to 1$. This completes the proof. ■

6.2.1 Theoretical prediction of the convergence rate

In this section, we predict more accurately the convergence rate of the DeC method.

- **Roe scheme in physical and mathematical phases**

  In this case, $\eta_1 = \eta_2 = \eta$ with $\eta \gg 1$ at low Mach number.

  The convergence can then be analytically estimated. Indeed,

  $$\lim_{\eta\to\infty} \mid F(G_\infty) \mid = \mid 1 - 2\,\bar\beta\,s \mid .$$

  Hence

  $$S = \sup_{t\in\{0,1\}} \mid F(G_\infty) \mid \sim 1.$$

  The implicit Defect-Correction algorithm does not converge well at very low Mach number with the Roe scheme.

  More accurately, it is shown in Figure 10 that the larger the value of η with regard to 1, the more the convergence degrades. This degradation is very quick; as soon as $\eta = 100$ the method converges badly independently of $\bar\beta$.

Mach number M	0.1	0.01	0.001
Parameter $\eta = M^{-1}$	10	100	1000
Amplification operator S	0.9	0.99	about 1

Table 3: Amplification operator versus the Mach number with the second-order accurate Roe scheme ($\bar{\beta} = 0.5$). Prediction of Fourier analysis.

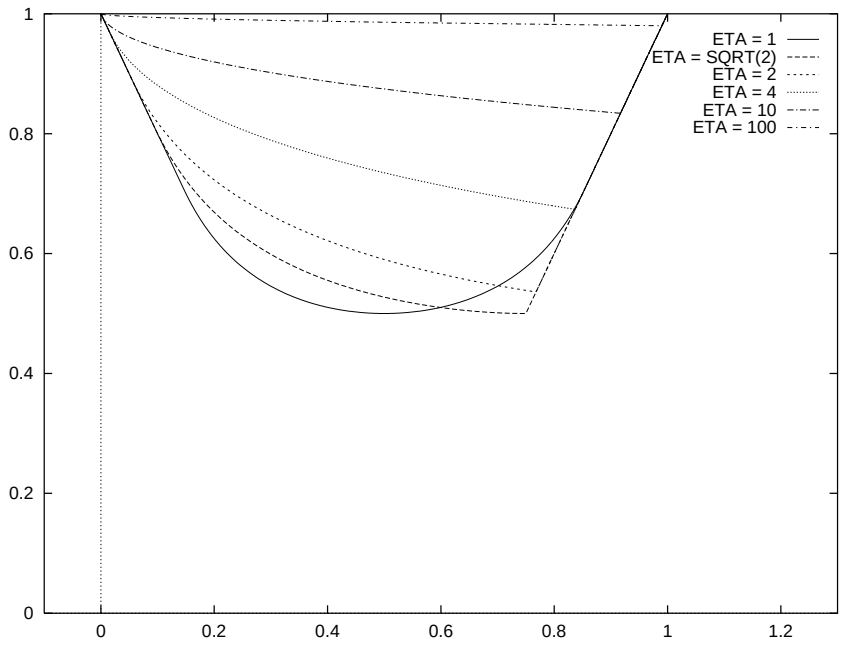

Figure 10: Amplification operator predicted by Fourier analysis S versus the upwinding parameter $\bar{\beta}$ for various values of η. The lower the Mach number, the closer to 1 is the amplification operator η.

The parameter η is of the order of $\frac{1}{M}$. The prediction rates are summarised in table 3 from the plots in Figure 10. This table predicts the convergence rate of the DeC method at second order accuracy and for an upwinding parameter $\bar{\beta} = 0.5$.

- **Roe-Turkel scheme in physical and mathematical phases**

  In this case, $\eta_1 = \eta_2 = \eta = 1$. Thus

  $$\mid F(G_\infty) \mid = \sqrt{s\,\left(1 + 4\,\bar{\beta}\,(\bar{\beta} - 1)\,s\right)}.$$

  This is also the case of the Roe scheme for non-low Mach number.

Mach number M	0.1	0.01	0.001
convergence rate	0.95	0.97	0.99

Table 4: Convergence rate versus Mach number with the Roe scheme. Given by results of the numerical experiments (see Figure 11) .

By taking $\gamma = 1.4$ in equation (82) we obtain

$$P_c^{-1} \mid P_c A_c \mid \sim \begin{pmatrix} 0.73\,u & 0.09 & \dfrac{0.36}{u} \\ 0.52\,u^2 & 1.25\,u & 0.54 \\ -0.27\,u\,H & 0.09\,H & \dfrac{0.36\,H}{u} \end{pmatrix}. \tag{104}$$

The parameter η is equal to 1.25 independently of the Mach number. From Figure 10 for $\bar{\beta} = 1/2$ we estimate the amplification operator S to be 0.53.

6.2.2 Numerical experiments

- **Roe scheme**

  As predicted by analysis, the lower the Mach number, the more the convergence rate degrades. For a mesh of 800 nodes (Figure 11), we obtain the convergence rates given in table 4.

- **Roe-Turkel scheme**

  For a mesh of 800 nodes (Figure 12) the convergence rate is 0.93 independently of the Mach number. For finer meshes of 3114 and 12284 nodes (Figure 12), the rate of convergence seems to be about 0.76 independently of the fineness of the mesh. This experimental value should be compared with the value 0.53 predicted on the 1D advective model.

The independence of the convergence rate with regard to the Mach number is verified for all sizes of mesh. The independence of the convergence rate with regard to the mesh is verified for cases where the mesh is fine enough (more than 3000 vertices).

Remark:
For fine meshes and a low Mach number, the convergence of the residual is limited by the zero machine at a level dependent on the Mach number if the average values of certain variables are not introduced [26]. ■

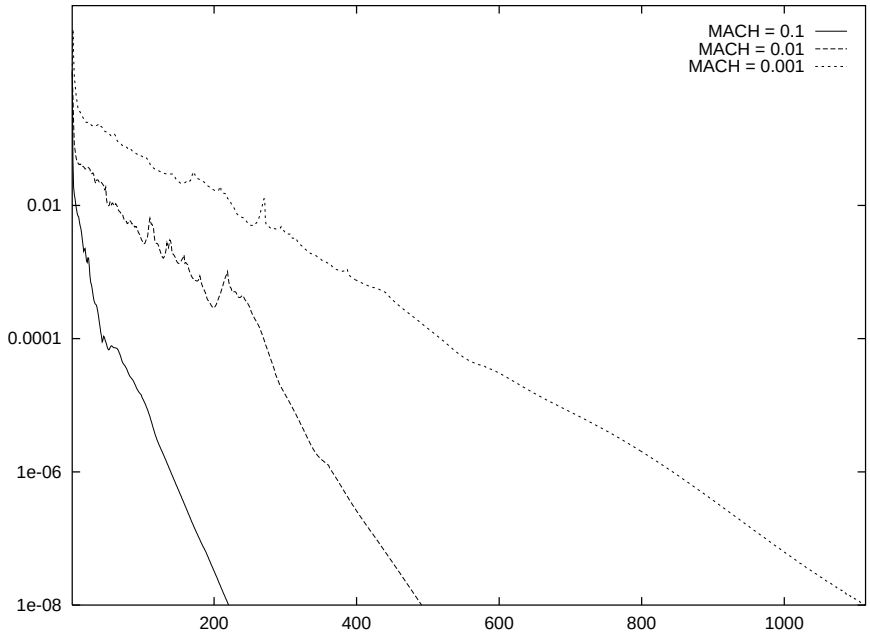

Figure 11: Convergence of the DeC method with the second-order accurate Roe scheme on a mesh of 800 nodes for flow arounds a NACA0012 without incidence.

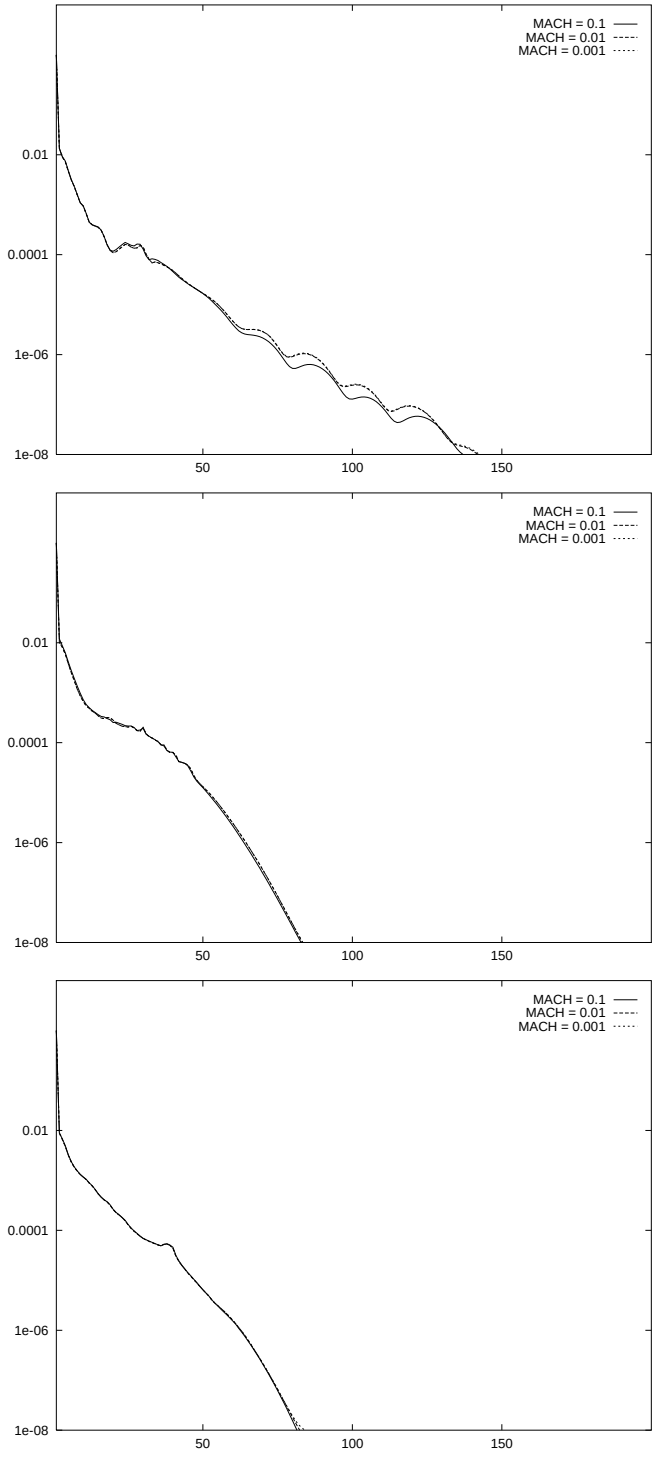

Figure 12: Convergence of the implicit DeC method with the second-order accurate Roe-Turkel scheme on a mesh of 800 nodes (top), 3114 nodes (middle) and 12284 nodes (top).

7 Numerical tests: shear-driven cavity flows

We apply the Roe-Turkel scheme to a shear-driven cavity flow. The Navier-Stokes incompressible equations are commonly solved for this model problem. Thus, this test case enables us to compare our results to those given in the literature and particularly with those of Ghia *et al.* [13] and with those of Botella and Peyret [2]. Ghia *et al.* used a centred finite difference discretisation and their computations were performed on a very thin mesh of 129×129 nodes. More recently, Botella and Peyret obtained more accurate results by a Chebyshev collocation method on a mesh of 49×49 for $Re = 100$ and of 97×97 for $Re = 1000$.

The computational domain is a square defined on $\Omega = [0, 1]^2$. The upper wall is driven at a constant horizontal velocity while the other walls are motionless. The initial velocity field is uniform in all the domain. The Mach number is equal to 10^{-3} in all the domain. The walls being adiabatic, the system certainly receives (kinetic) energy without losses, thus it should not have a steady state solution. In practice, however, we will say that a steady state solution is a solution for which the normalised residual of the momentum decreased of a factor 10^{10}.

We performed two tests corresponding to different Reynolds numbers; one at $Re = 100$ and the other at $Re = 1000$. We used cartesian meshes where each square was split into two triangles. The spatial approximation is second-order accurate.

For a Reynolds number of 100, a coarse mesh of $441 = 21 \times 21$ nodes is enough to obtain a solution almost identical to the one given by Ghia *et al.* and by Botella and Peyret for the horizontal velocity profile in the median vertical axis of the cavity (Figure 13), and for the vertical velocity profile in the median horizontal axis of the cavity (Figure 14). The residual of the horizontal momentum of the implicit DeC method decreases by 10 orders of magnitude in 60 time steps (Figure 15). The CPU time to compute these 60 time steps is 58 seconds on a Dec Alpha 600 / 266 Mhz workstation. At each time step, the linear system was converged to a normalised residual of 10^{-3} which is adequate since no amelioration of the time iterative convergence is obtained if the linear system is totally converged.

For a Reynolds number of 1000, the profile of the velocity in the median vertical axis of the cavity was computed on several meshes (Figure 16). Whatever the mesh, the steady state is reached in only 2500 time iterations because large time steps could not be used without obtaining negative densities. The 61×61 mesh is 1.5 times finer than the 41×41 mesh. Therefore, the distance between the solution obtained with a 61×61 mesh and the reference solution should be 1.5^α (where α is the order of accuracy) times smaller than the distance between the solution obtained with a 41×41 mesh and the reference solution. In the same way, the 81×81 mesh is 2 times finer than the 41×41 mesh. Therefore, the distance between the solution obtained with a 81×81 mesh and the reference solution should be 2^α times smaller than the distance between the solution obtained with a 41×41 mesh and

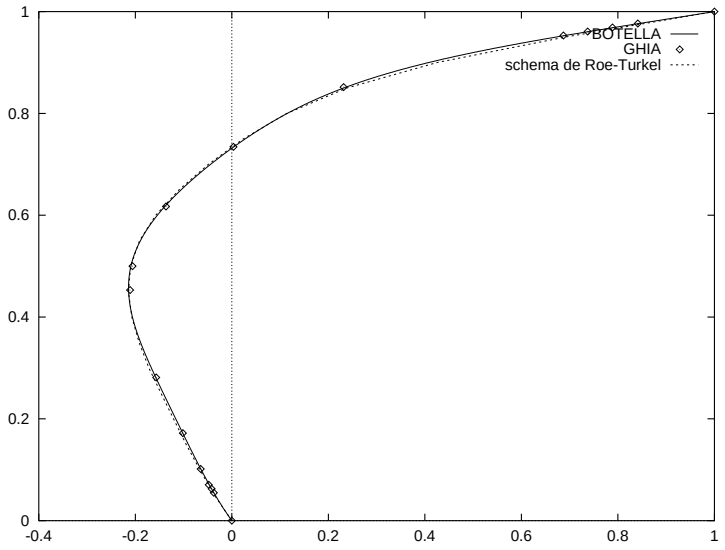

Figure 13: Profile of the horizontal velocity at $x = 0.5$ along the y-axis, second-order accurate Roe-Turkel scheme, $M_\infty = 10^{-3}$, $Re = 100$, uniform mesh, 21×21, no limiters.

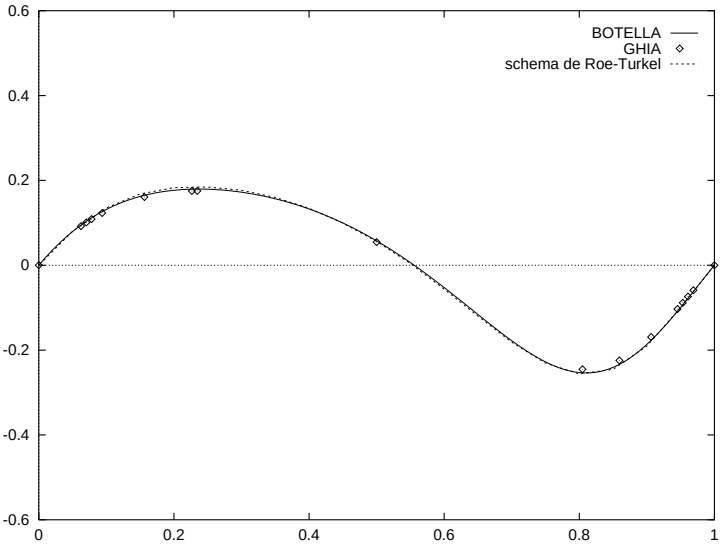

Figure 14: Vertical velocity at $y = 0.5$ along the x-axis, second-order accurate Roe-Turkel scheme, $M_\infty = 10^{-3}$, $Re = 100$, uniform mesh, 21×21, no limiters.

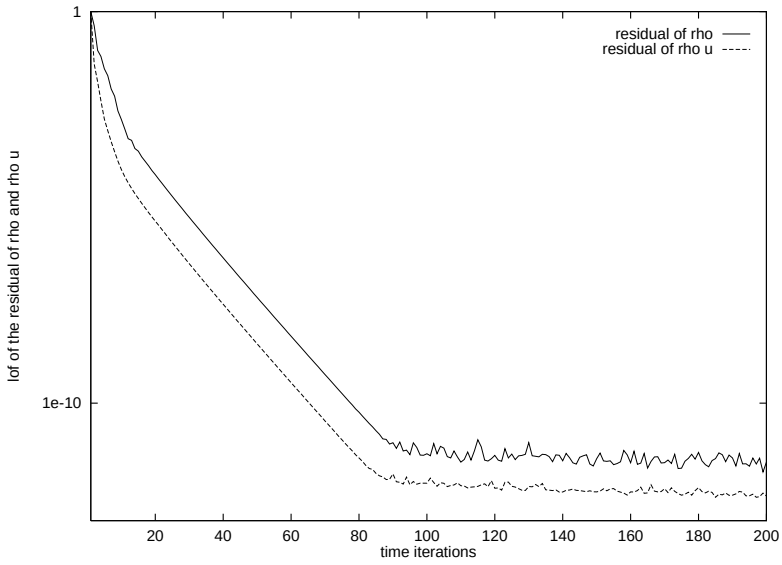

Figure 15: Convergence of the DeC implicit method with the second-order accurate Roe-Turkel scheme, $M_\infty = 10^{-3}$, $Re = 100$, uniform mesh, 21×21.

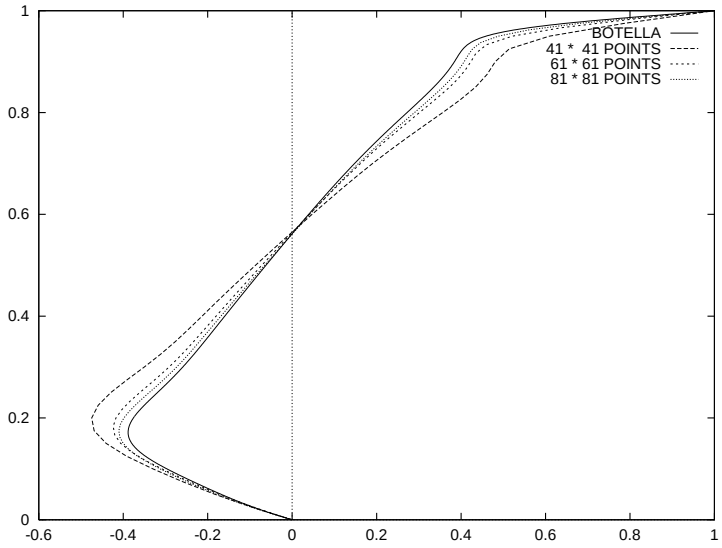

Figure 16: Profile of the horizontal velocity at $x = 0.5$ along the y-axis, second-order accurate Roe-Turkel scheme, $M_\infty = 10^{-3}$, $Re = 1000$, uniform meshes finer and finer, no limiters.

	41×41	61×61	81×81	*spectral* [2]
$u(0.5, 0.2)$	-0.4747	-0.4172	-0.3985	-0.3592
numerical order		1.7	1.6	
$u(0.5, 0.9)$	0.4811	0.4232	0.4074	0.3843
numerical order		2.2	2.0	

Table 5: Numerical order for a given mesh and a given location is computed by comparison of the derivation with a spectral value as compared to a analogous figure obtains on the coarser mesh. Results obtained from the experiments presented in Figure 16

the reference solution. The experiment results presented in Figure 16 seem to confirm a behaviour close to second order (see Table 5). The location of the center of the main vortex is correctly determined with all meshes, but the scheme involves a certain dissipation which does not enable us to capture the upper boundary layer even with the 81×81 mesh. In next figures we present the results of a numerical experiment performed on a mesh originally of size 81×81 whose two node lines near the upper wall have been refined. On this mesh, the profiles of the velocity in the median vertical axis and in the median horizontal axis of the cavity are presented respectively in Figures 17 and 18. The solution is close to the reference solutions. The approximation converges towards the solution of the incompressible Navier-Stokes equations.

The Mach number is less than 10^{-3} throughout the flow (see Figure 19). The fluctuations of the density and of the pressure represented in Figure 20 are respectively computed by $\rho_{\text{represented}} = (\rho - \rho_{min})/(\rho_{max} - \rho_{min})$ and $p_{\text{represented}} = (p - p_{min})/(p_{max} - p_{min})$. The density is of the order of 1 and it has a variation of the order of M^2 as expected. We present in Figure 21 the trajectories of the particles that we obtain, and in Figure 22 the streamfunction computed by Botella and Peyret. The secondary vortices are almost of the same size.

Remark:
With the original Roe scheme, on a mesh of size 81×81 the profile of the velocity obtained at $Re = 1000$ is inaccurate and similar to the one obtained with the Roe-Turkel scheme for $Re = 100$. In order to obtain with the original Roe scheme results as accurate as those obtained with the Roe-Turkel scheme, an extremely fine mesh should be necessary.

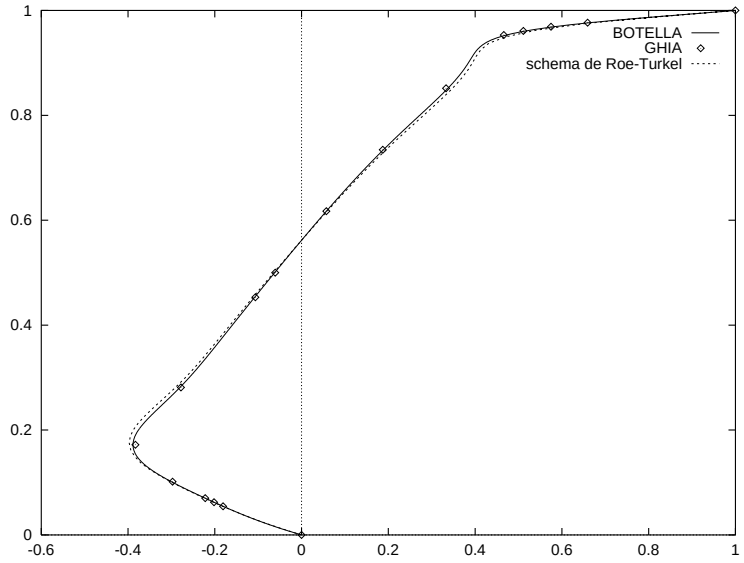

Figure 17: Profile of the horizontal velocity at $x = 0.5$ along the y-axis, second-order accurate Roe-Turkel scheme, $M_\infty = 10^{-3}$, $Re = 1000$, no limiters, mesh 81×81 refined near the upper wall (7586 nodes).

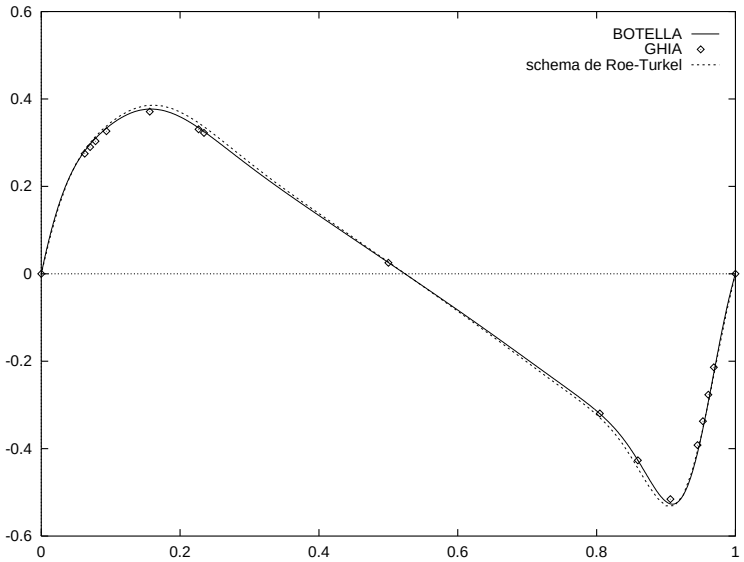

Figure 18: Vertical velocity at $y = 0.5$ along the x-axis, second-order accurate Roe-Turkel scheme, $M_\infty = 10^{-3}$, $Re = 1000$, no limiters, mesh 81×81 refined near the upper wall (7586 nodes).

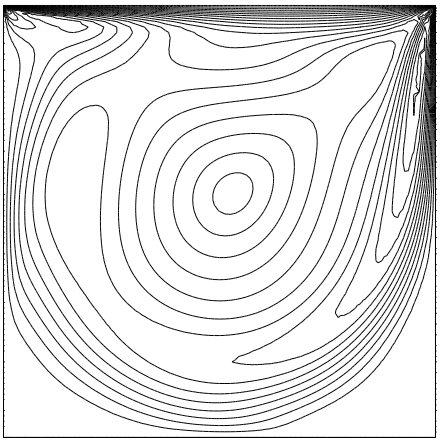

Figure 19: Isovalues of Mach number. Interval between isovalues: 5.10^{-5}. Min/Max: 10^{-5}, 10^{-3}. Second-order accurate Roe-Turkel scheme, no limiters, $Re = 1000$. Mesh 81 $\times$ 81 refined near the upper wall (7586 nodes).

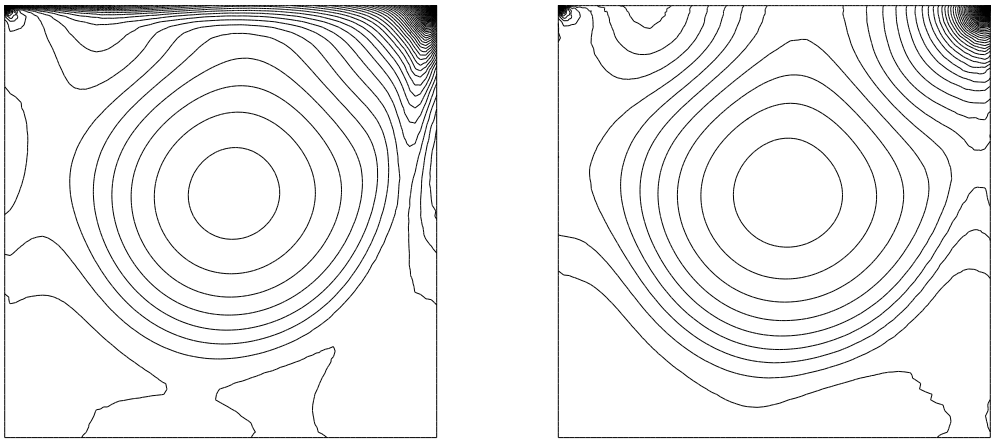

Figure 20: Isovalues of the fluctuations ($x_{\text{represented}} = \frac{x - x_{min}}{x_{max} - x_{min}}$) of density (left) and of pressure (right). $\rho_{min} = 0.99930016$, $\rho_{max} = 0.99930925$. $p_{min} = 714283$, $p_{max} = 714290$. Second-order accurate Roe-Turkel scheme, no limiters, $Re = 1000$, M=10^{-3}. Interval between isovalues: 0.002. Min/Max: 0.11, 0.14. Mesh 81 $\times$ 81 refined near the upper wall (7586 nodes).

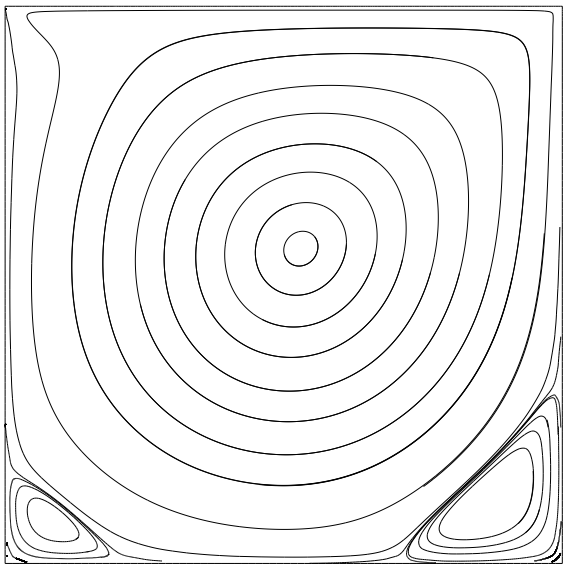

Figure 21: Trajectories of the particles. Second-order accurate Roe-Turkel scheme, without limiters, $Re = 1000$, M=10^{-3}. Mesh 81 $\times$ 81 refined near the upper wall (7586 nodes).

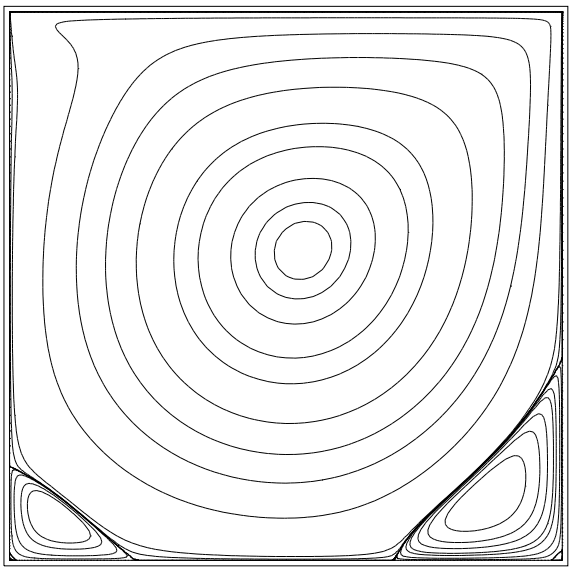

Figure 22: Isovalues of the streamfunction obtained with a spectral approximation (Chebyshev, 97 $\times$ 97) of the incompressible solution in reference [2] for $Re = 1000$.

8 Conclusion

In this study, we have elected a base preconditioner due to Turkel, for its simplicity and its generality. Other preconditioners, for example those of van Leer-Lee-Roe and of Choi-Merkle, should lead to similar results.
The preconditioner has been applied to the flux splitting of Roe but it could likely be adapted to most upwinding schemes of characteristic type (Steger-Warming, Osher, ...).

This work enabled us to elucidate two problems linked to the approximation of the Roe scheme and of the MUSCL technique to low Mach number flows.

On the one hand we showed that the Roe scheme is not uniformly consistent with regard to the Mach number and that at low Mach number it is advisable to use what we call the Roe-Turkel scheme which is uniformly consistent with regard to the Mach number.

On the other hand we showed that when the Roe scheme is used the rate of the iterative convergence towards the steady state solution of the implicit Defect-Correction scheme degrades as the Mach number tends to zero. In contrast, with the Roe-Turkel scheme in both phases of the algorithm the convergence is independent of the Mach number when this parameter is small.

The improvements presented in this report have been introduced in a 2D and 3D industrial code (cf Appendix A) and 3D computations are planned. The proposed method already enables a certain number of computations with an accuracy comparable to the one of incompressible models, at a somewhat greater although not prohibitive cost.

In order to use very intensively such a scheme for low Mach number flow it would still be advisable to look into the following problems:

- **Rounding errors**
  The plots of the convergence of the Defect-Correction method end with a level consisting of small random oscillations around a value of the residual. This level is due to rounding errors of the computer. The lower the Mach number, the higher this level is with respect to residuals. In particular, at low Mach number (in practice the limit level is 10^{-6} in double precision), the pressure is of order of $\frac{1}{M^2}$ whilst the fluctuations of the pressure are of order of 1; thus, rounding errors appear when two pressures are subtracted for example. This problem is known and to cure it one should work with a reference pressure [26].

- **Robustness of mixed flows. Low Mach/high Mach**
  The computations that we presented are globally at low Mach numbers. The preconditioning involves a parameter β to which a value must be given. For the tests presented here, we took either β equal to the maximum Mach number in all the flow, or for a larger accuracy β equal to the maximum between the local Mach number and a small value.
  For mixed flow low Mach/transonic only the option β variable is possible. However, when β varies we observe in numerical experiments that the Roe-Turkel scheme is less robust than the Roe scheme. This point is under investigation.

- **Efficiency of the algorithm of the linear resolution**
  At each iteration of the Defect-Correction method a linear system must be solved and we use the Jacobi method. The rate of the iterative convergence of the Jacobi method degrades with the Mach number. This phenomenon accelerates with the Roe-Turkel scheme. The analysis of this difficulty is also under investigation.

9 Acknowledgements

I thank Alain Dervieux for guiding my work, and J. Antoine Désidéri, Stéphane Lanteri, Hervé Guillard and Roger Peyret for their interest in this study and various comments. I also thank Jérome Francescatto for reviewing an early draft of the report. One part of the computations was performed with the code N3S-MUSCL of the firm SIMULOG thanks to the help of Didier Chargy and Agnès Merlo.

This work was part of a thesis supported by CNES and RENAULT.

References

[1] G. K. Batchelor, *An Introduction to Fluid Dynamics*, cambridge University Press, 1967.

[2] O. Botella and R. Peyret, Computations of Singular Solutions of the Navier-Stokes Equations using the Chebyshev Collocation Method, technical report INRIA, in preparation, 1996.

[3] S. Candel, *Mécanique des fluides – cours*, DUNOD, 1995 (in French).

[4] D. Choi and C. L. Merkle, Application of Time-iterative Schemes to Incompressible Flow, *A.I.A.A Journal*, vol. 23, pp. 1518 – 1524, 1985.

[5] Y.-H. Choi and C. L. Merkle, The Application of Preconditioning in Viscous Flows, *J. Comp. Phys.*, vol. 105, pp. 207 – 223, 1993.

[6] A. J. Chorin, A Numerical Method for Solving Incompressible Viscous Flow Problems, *J. Comp. Phys.*, vol. 2, pp. 12 – 26, 1967.

[7] D.L. Darmofal and P.J. Schmid, The Importance of Eigenvectors for Local Preconditioners of the Euler Equations, *J. Comp. Phys.*, pp. 346 – 362, 1996.

[8] A. Dervieux and J.-A. Désidéri, Compressible Flow Solvers Using Unstructured Grids, Technical Report INRIA, RR1732, 1992.

[9] J.-A. Désidéri and P. W. Hemker, Analysis of the Convergence of Iterative Implicit and Defect-Correction Algorithms for Hyperbolic Problems, Technical Report INRIA, RR1200, 1990.

[10] J.A. Désidéri and P.W. Hemker, Convergence Analysis of the Defect-Correction Iteration for Hyperbolic Problems, *SIAM J. Sci. Comput.*, pp. 88 – 118, 1995.

[11] J.A. Désidéri, C. Hirsch, P. Le Tallec, M. Pandolfi, and J. Periaux, editors, *Proceedings of the Third ECCOMAS CFD Conference*, Wiley, 1996.

[12] J. Fröhlich, *Résolution numérique des équations de Navier-Stokes à faible nombre de Mach par méthode spectrale*, PhD thesis, University of Nice-Sophia-Antipolis, 1990 (in French).

[13] U. Ghia, K. N. Ghia, and C. T. Shin, High-Re Solutions for Incompressible Flow Using the Navier-Stokes Equations and a Multigrid Method, *J. Comp. Phys.*, vol. 48, pp. 387 – 411, 1982.

[14] A. G. Godfrey, R. W. Walters, and B. van Leer, Preconditioning for the Navier-Stokes Equations with Finite-rate Chemistry, *AIAA Paper 93-0535*, 1993.

[15] W. E. Milholen II, N. Chokani, and J. Al-Saadi, Performance of Three-dimensional Compressible Navier-Stokes Codes at Low Mach Numbers, *A.I.A.A Journal*, vol. 34, no 7, pp. 1356 – 1362, 1996.

[16] R. Klein, Semi-implicit Extension of a Godunov-type Scheme Based on Low Mach Number Asymptotics i: One-dimensional Flow, *J. Comp. Phys.*, vol. 121, pp. 213 – 237, 1995.

[17] L. Landau and E. Lifchitz, *Physique Théorique Tome VI: Mécanique des fluides*, Editions MIR, 1971 (in French).

[18] W.-T. Lee, *Local Preconditioning of the Euler Equations*, PhD thesis, University of Michigan, 1992.

[19] A. Majda, *Compressible Fluid Flow and Systems of Conservation Laws in Several Space Variables*, Springer-Verlag, 1984.

[20] A. Majda and J. Sethian, The Derivation and Numerical Solution of the Equations for Zero Mach Number Combustion, *Combust. Sci. and Tech.*, vol. 42, pp. 185 – 205, 1985.

[21] R. Martin, *Développement de méthodes de calculs efficaces pour des écoulements instationnaires en géométrie déformable*, PhD thesis, University of Nice-Sophia-Antipolis, 1996 (in French).

[22] R. Peyret and T. Taylor, *Computational Methods for Fluid Flow*, Springer, New York, 1983.

[23] Courant R., Isaacson E., and Rees M, On the Solution of Nonlinear Hyperbolic Differential Equations by Finite Differences, *Communications Pure and Appl. Math.*, 1952.

[24] R. Radespiel and E. Turkel, A Comparison of Preconditioning Methods, In CRM en collaboration avec CERCA, editor, *Conference on Numerical Methods for the Euler and Navier-Stokes Equations*, 1995.

[25] P. L. Roe, Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes, *J. Comp. Phys.*, vol. 43, pp. 357 – 372, 1981.

[26] J. Sesterhenn, B. Müller, and H. Thomann, On the Cancellation Problem in Calculating Compressible Low Mach Number Flows, In *Proceedings of the Third ECCOMAS CFD Conference*, 1996.

[27] D. J. Tritton, *Physical Fluid Dynamics*, Oxford Science Publication, second edition edition, 1988.

[28] E. Turkel, Preconditioned Methods for Solving the Incompressible and Low Speed Compressible Equations, *J. Comp. Phys.*, vol. 72, pp. 277 – 298, 1987.

[29] E. Turkel, Review of Preconditioning Methods for fluid dynamics, *Applied Numerical Mathematics*, vol. 12, pp. 257 – 284, 1993.

[30] E. Turkel, A. Fiterman, and B. van Leer, Preconditioning and the Limit to the Incompressible Flow Equations for Finite Difference Schemes, In M. Hafez and D. Caughey, editors, *Computing the Future: Advances and Prospects for Computational Aerodynamics*, pp. 215 – 234, John Wiley and Sons, 1994.

[31] B. van Leer. Towards the Ultimate Conservative Difference Scheme i. the Quest of Monotonicity, *Lectures notes in Physics*, vol. 18, pp. 163, 1972.

[32] B. van Leer, Flux-vector Splitting for the Euler Equations, *Lectures notes in Physics*, 170, 1982.

[33] B. van Leer, W.-T. Lee, and P. Roe, Characteristic Time-stepping or Local Preconditioning of the Euler Equations, *AIAA paper 91-1552*, 1991.

[34] B. van Leer, L. Mesaros, C-H Tai, and Turkel, Local Preconditioning in a Stagnation Point, *AIAA paper 95-1654*, 1995.

[35] H. Viviand, Pseudo-unsteady Systems for Steady Inviscid Flow Calculation, In F. Angrand, A. Dervieux, J.A. Désidéri, and R. Glowinski, editors, *Numerical Methods for the Euler Equations of Fluid Dynamics*, pages 334 – 368. SIAM Philadelphia, 1985.

[36] G. Volpe, Performance of Compressible Flow Codes at Low Mach Numbers, *A.I.A.A Journal*, vol 31, no 1, pp. 49 – 56, Jan. 1993.

[37] R. F. Warming and B. J. Hyett, The Modified Equation Approach to the Stability and Accuracy Analysis of Finite-Difference Methods, *J. Comp. Phys.*, vol. 14, pp. 159 – 179, 1974.

A Computation of the stabilisation term of the Roe-Turkel scheme in 3D

A.1 Computation of the matrix $P_c^{-1} \mid P_c\, D_c \mid$

The stabilisation term of the Roe scheme written using conservative variables, $W = [\rho, \rho\, u, \rho\, v, \rho\, w, E]$, is given by

$$\mid D_c \mid = \mid A_c\, \nu_x + B_c\, \nu_y + C_c\, \nu_z \mid . \tag{105}$$

To obtain the Roe-Turkel scheme this term is replaced by

$$P_c{}^{-1} \mid P_c\, D_c \mid . \tag{106}$$

Using the entropic variables, $U = [p,\ u,\ v,\ w,\ S]$, we have $P_e = Diag(\beta^2, 1, 1, 1, 1)$ where the parameter β is taken of the order of the local Mach number.

The computation of the absolute value of a matrix requires the diagonalisation of this matrix. Since the diagonalisation of (106) is simpler performed using the entropic variables than using the conservative variables, we write

$$P_c{}^{-1} \mid P_c\, D_c \mid = R^{-1}\, P_e{}^{-1} \mid P_e\, D \mid R \tag{107}$$

where the Jacobian matrices $R^{-1} = \dfrac{\partial W}{\partial U}$ and $R = \dfrac{\partial U}{\partial W}$ are given by

$$R = \begin{bmatrix} (\gamma-1)\dfrac{q^2}{2} & -u\,(\gamma-1) & -v\,(\gamma-1) & -w\,(\gamma-1) & \gamma-1 \\ -\dfrac{u}{\rho} & \dfrac{1}{\rho} & 0 & 0 & 0 \\ -\dfrac{v}{\rho} & 0 & \dfrac{1}{\rho} & 0 & 0 \\ -\dfrac{w}{\rho} & 0 & 0 & \dfrac{1}{\rho} & 0 \\ \dfrac{\gamma-1}{p}\dfrac{q^2}{2} - \dfrac{\gamma}{\rho} & -\dfrac{(\gamma-1)\,u}{p} & -\dfrac{(\gamma-1)\,v}{p} & -\dfrac{(\gamma-1)\,w}{p} & \dfrac{\gamma-1}{p} \end{bmatrix}, \tag{108}$$

where $q = \sqrt{u^2 + v^2 + w^2}$ is the magnitude of velocity, and

$$R^{-1} = \begin{bmatrix} \frac{1}{a^2} & 0 & 0 & 0 & -\frac{\rho}{\gamma} \\ \frac{u}{a^2} & \rho & 0 & 0 & -\frac{\rho\, u}{\gamma} \\ \frac{v}{a^2} & 0 & \rho & 0 & -\frac{\rho\, v}{\gamma} \\ \frac{w}{a^2} & 0 & 0 & \rho & -\frac{\rho\, w}{\gamma} \\ \frac{H}{a^2} & \rho\, u & \rho\, v & \rho\, w & -\frac{\rho\, q^2}{2\,\gamma} \end{bmatrix}. \tag{109}$$

where a is the speed of sound and

$$H = \frac{q^2}{2} + a^2\ (\gamma - 1)^{-1} \tag{110}$$

is enthalpy,
and where the matrix $P_e\, D$ is given by

$$P_e\, D = \begin{bmatrix} \lambda_1\, \beta^2 & \rho\, \beta^2\, a^2\, \nu_x & \rho\, \beta^2\, a^2\, \nu_y & \rho\, \beta^2\, a^2\, \nu_z & 0 \\ \frac{\nu_x}{\rho} & \lambda_1 & 0 & 0 & 0 \\ \frac{\nu_y}{\rho} & 0 & \lambda_1 & 0 & 0 \\ \frac{\nu_z}{\rho} & 0 & 0 & \lambda_1 & 0 \\ 0 & 0 & 0 & 0 & \lambda_1 \end{bmatrix} \tag{111}$$

where $\lambda_1 = u\, \nu_x + v\, \nu_y + w\, \nu_z$.

The diagonalisation of $\mid P_e\, D \mid$ can be written as

$$\mid P_e\, D \mid = M \mid \Lambda \mid M^{-1} \tag{112}$$

where the columns of the matrix M are the right eigenvectors of the matrix $P_e\, D$, and M^{-1} is the inverse of the matrix M.

The matrices M and M^{-1} are given by

$$M = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & -\widehat{\nu}_y & 0 & \dfrac{\widehat{r}\,\widehat{\nu}_x}{\rho\,\beta^2\,a^2} & \dfrac{\widehat{s}\,\widehat{\nu}_x}{\rho\,\beta^2\,a^2} \\ 0 & \widehat{\nu}_x & -\widehat{\nu}_z & \dfrac{\widehat{r}\,\widehat{\nu}_y}{\rho\,\beta^2\,a^2} & \dfrac{\widehat{s}\,\widehat{\nu}_y}{\rho\,\beta^2\,a^2} \\ 0 & 0 & \widehat{\nu}_y & \dfrac{\widehat{r}\,\widehat{\nu}_z}{\rho\,\beta^2\,a^2} & \dfrac{\widehat{s}\,\widehat{\nu}_z}{\rho\,\beta^2\,a^2} \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}, \tag{113}$$

$$M^{-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & -\dfrac{\widehat{\nu}_z^2+\widehat{\nu}_y^2}{\widehat{\nu}_y} & \widehat{\nu}_x & \dfrac{\widehat{\nu}_x\,\widehat{\nu}_z}{\widehat{\nu}_y} & 0 \\ 0 & -\dfrac{\widehat{\nu}_x\,\widehat{\nu}_z}{\widehat{\nu}_y} & -\widehat{\nu}_z & \dfrac{\widehat{\nu}_y^2+\widehat{\nu}_x^2}{\widehat{\nu}_y} & 0 \\ \dfrac{\widehat{s}}{2\,\widehat{t}} & -\dfrac{\rho\,\beta^2\,a^2\,\widehat{\nu}_x}{2\,\widehat{t}} & -\dfrac{\rho\,\beta^2\,a^2\,\widehat{\nu}_y}{2\,\widehat{t}} & -\dfrac{\rho\,\beta^2\,a^2\,\widehat{\nu}_z}{2\,\widehat{t}} & 0 \\ -\dfrac{\widehat{r}}{2\,\widehat{t}} & \dfrac{\rho\,\beta^2\,a^2\,\widehat{\nu}_x}{2\,\widehat{t}} & \dfrac{\rho\,\beta^2\,a^2\,\widehat{\nu}_y}{2\,\widehat{t}} & \dfrac{\rho\,\beta^2\,a^2\,\widehat{\nu}_z}{2\,\widehat{t}} & 0 \end{bmatrix}, \tag{114}$$

with

$$r = \lambda_3 - \lambda_1\,\beta^2, \quad s = \lambda_4 - \lambda_1\,\beta^2, \quad t = \frac{\lambda_5 - \lambda_4}{2},$$

and where the stressed variables x are defined by $\widehat{x} = x/\sqrt{\nu_x^2+\nu_y^2+\nu_z^2}$.

The diagonal matrix, composed of the eigenvalues of $P_e\,D$, is given by

$$\Lambda = Diag(\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \tag{115}$$

with

$$\begin{cases} \lambda_1 = \lambda_2 = \lambda_3 & = & u\,\nu_x + v\,\nu_y + w\,\nu_z & = & \vec{U}\cdot\vec{\nu} \\ \lambda_{4,5} & = & \dfrac{1}{2}\left[(1\,+\,\beta^2)\,\lambda_1 \pm \sqrt{[(1-\beta^2)\,\lambda_1]^2 + [2\,\beta\,a\;\|\;\nu\;\|]^2}\right] \end{cases} \tag{116}$$

where $\|\;\nu\;\| = \sqrt{{\nu_x}^2 + {\nu_y}^2 + {\nu_z}^2}$.

Remark:
In the case where $\beta = 1$ (Roe scheme) we have λ_1, λ_2 and λ_3 unchanged and we have

$$\lambda_{4,5} = \lambda_1 \pm a\;\|\;\nu\;\|, \tag{117}$$

thus the variables $\widehat{r}$, $\widehat{s}$ and $\widehat{t}$ are of the order of the speed of sound

$$\widehat{r} = a, \quad \widehat{s} = -a, \quad \widehat{t} = -a,$$

whilst when $M \to 0$ and $\beta \sim M$ (Roe-Turkel scheme) we have the variables $\widehat{r}$, $\widehat{s}$ and $\widehat{t}$ of the order of the flow velocity. ■

In conclusion, for low Mach number flow the term $\mid D_c \mid$ (Roe scheme) should be replaced by (Roe-Rurkel scheme):

$$\begin{array}{rcl} {P_c}^{-1} \mid P_c \, D_c \mid & = & R^{-1} \, P_e^{-1} \, M \mid \Lambda \mid M^{-1} \, R \\ & = & T_{rt}^g \mid \Lambda \mid T_{rt}^d \end{array} \tag{118}$$

where the matrices T_{rt}^g and T_{rt}^d are given by

$$T_{rt}^g = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 \\ u & 1 & 0 & u + \widehat{r}\,\widehat{\nu}_x & u + \widehat{s}\,\widehat{\nu}_x \\ v & -\dfrac{\widehat{\nu}_x}{\widehat{\nu}_y} & -\dfrac{\widehat{\nu}_z}{\widehat{\nu}_y} & v + \widehat{r}\,\widehat{\nu}_y & v + \widehat{s}\,\widehat{\nu}_y \\ w & 0 & 1 & w + \widehat{r}\,\widehat{\nu}_z & w + \widehat{s}\,\widehat{\nu}_z \\ \dfrac{q^2}{2} & u - v\,\dfrac{\widehat{\nu}_x}{\widehat{\nu}_y} & w - v\,\dfrac{\widehat{\nu}_z}{\widehat{\nu}_y} & H + \widehat{r}\widehat{\lambda}_1 & H + \widehat{s}\widehat{\lambda}_1 \end{bmatrix} \tag{119}$$

$$T_{rt}^d = \begin{bmatrix} 1 - \dfrac{(\gamma-1)}{a^2}\dfrac{q^2}{2} & \dfrac{\gamma-1}{a^2}\,u & \dfrac{\gamma-1}{a^2}\,v & \dfrac{\gamma-1}{a^2}\,w & -\dfrac{\gamma-1}{a^2} \\ \widehat{\lambda}_1\,\widehat{\nu}_x - u & \widehat{\nu}_y^2 + \widehat{\nu}_z^2 & -\widehat{\nu}_x\,\widehat{\nu}_y & -\widehat{\nu}_x\,\widehat{\nu}_z & 0 \\ \widehat{\lambda}_1\,\widehat{\nu}_z - w & -\widehat{\nu}_x\,\widehat{\nu}_z & -\widehat{\nu}_y\,\widehat{\nu}_z & \widehat{\nu}_x^2 + \widehat{\nu}_y^2 & 0 \\ \dfrac{\widehat{s}\frac{q^2}{2}(\gamma-1)+\beta^2 a^2 \widehat{\lambda}_1}{2\beta^2 a^2 \widehat{t}} & -\dfrac{\widehat{s}u(\gamma-1)+\beta^2 a^2 \widehat{\nu}_x}{2\beta^2 a^2 \widehat{t}} & -\dfrac{\widehat{s}v(\gamma-1)+\beta^2 a^2 \widehat{\nu}_y}{2\beta^2 a^2 \widehat{t}} & -\dfrac{\widehat{s}w(\gamma-1)+\beta^2 a^2 \widehat{\nu}_z}{2\beta^2 a^2 \widehat{t}} & \dfrac{\widehat{s}(\gamma-1)}{2\beta^2 a^2 \widehat{t}} \\ -\dfrac{\widehat{r}\frac{q^2}{2}(\gamma-1)+\beta^2 a^2 \widehat{\lambda}_1}{2\beta^2 a^2 \widehat{t}} & \dfrac{\widehat{r}u(\gamma-1)+\beta^2 a^2 \widehat{\nu}_x}{2\beta^2 a^2 \widehat{t}} & \dfrac{\widehat{r}v(\gamma-1)+\beta^2 a^2 \widehat{\nu}_y}{2\beta^2 a^2 \widehat{t}} & \dfrac{\widehat{r}w(\gamma-1)+\beta^2 a^2 \widehat{\nu}_z}{2\beta^2 a^2 \widehat{t}} & -\dfrac{\widehat{r}(\gamma-1)}{2\beta^2 a^2 \widehat{t}} \end{bmatrix}. \tag{120}$$

A.2 Computation of $P_c^{-1} \mid P_c \, D_c \mid \delta W$

We perform the computation of ${P_c}^{-1} \mid P_c \, D_c \mid$ by adding the artificial viscosity brought by each eigenvalue

$$ {P_c}^{-1} \mid P_c \, D_c \mid = T_{rt}^g \left(\Sigma_{n=1}^5 \mid \Lambda_n \mid \right) T_{rt}^d \tag{121}$$

where the matrix Λ_n has only one non-zero element equal to λ_n and located at the n^{th} entry of the diagonal.

We can also write $P_c^{-1} \mid P_c D_c \mid$ in the following form

$$P_c^{-1} \mid P_c D_c \mid \;=\; \Sigma_{n=1}^{5} \left(P_c^{-1} \mid P_c D_c \mid\right)^{(n)} \tag{122}$$

where

$$\begin{aligned} \left(P_c^{-1} \mid P_c D_c \mid\right)^{(n)} &= T_{rt}^{g} \mid \Lambda_n \mid T_{rt}^{d} \\ &= \mid \lambda_n \mid \left(i^{\text{th}}\, column\, of\, T_{rt}^{g}\right) \left(i^{\text{th}}\, line\, of\, T_{rt}^{d}\right). \end{aligned}$$

The stabilisation term of the numerical fluxes is given by

$$F \;= \Sigma_{i=1}^{5} F^{(n)} \tag{123}$$

where

$$F^{(n)} = \left(P_c^{-1} \mid P_c D_c \mid\right)^{(n)} \delta W \tag{124}$$

and

$$\delta W = \begin{bmatrix} \delta W_1 \\ \delta W_2 \\ \delta W_3 \\ \delta W_4 \\ \delta W_5 \end{bmatrix} = \begin{bmatrix} \rho_2 - \rho_1 \\ \rho_2\, u_2 - \rho_1\, u_1 \\ \rho_2\, v_2 - \rho_1\, v_1 \\ \rho_2\, w_2 - \rho_1\, w_1 \\ E_2 - E_1 \end{bmatrix}. \tag{125}$$

The artificial viscosity introduced by each eigenvalue is simply computed by

$$F^{(n)} = \mid \lambda_n \mid \alpha_{\lambda_n} \left(i^{\text{th}}\, column\, of\, T_{rt}^{g}\right) \tag{126}$$

where the eigenforms α_{λ_n} are given by

$$\alpha_{\lambda_n} = \left(i^{\text{th}}\, line\, of\, T_{rt}^{d}\right) \delta W. \tag{127}$$

We have

$$F^{(1)} \;=\; \mid \lambda_1 \mid \alpha_{\lambda_1} \begin{bmatrix} 1 \\ u \\ v \\ w \\ \dfrac{q^2}{2} \end{bmatrix}, \qquad F^{(2)} \;=\; \mid \lambda_2 \mid \alpha_{\lambda_2} \begin{bmatrix} 0 \\ 1 \\ -\dfrac{\widehat{\nu}_x}{\widehat{\nu}_y} \\ 0 \\ u - v\, \dfrac{\widehat{\nu}_x}{\widehat{\nu}_y} \end{bmatrix}, \tag{128}$$

$$F^{(3)} \;=\; \mid \lambda_1 \mid B \begin{bmatrix} 0 \\ 0 \\ -\dfrac{\widehat{\nu}_z}{\widehat{\nu}_y} \\ 1 \\ w - v \dfrac{\widehat{\nu}_z}{\widehat{\nu}_y} \end{bmatrix}, \tag{129}$$

$$F^{(4)} = \mid \lambda_4 \mid \alpha_{\lambda_4} \begin{bmatrix} 1 \\ u + \widehat{r}\,\widehat{\nu}_x \\ v + \widehat{r}\,\widehat{\nu}_y \\ w + \widehat{r}\,\widehat{\nu}_z \\ H + \widehat{r}\,\widehat{\lambda}_1 \end{bmatrix}, \quad F^{(5)} = \mid \lambda_5 \mid \alpha_{\lambda_5} \begin{bmatrix} 1 \\ u + \widehat{s}\,\widehat{\nu}_x \\ v + \widehat{s}\,\widehat{\nu}_y \\ w + \widehat{s}\,\widehat{\nu}_z \\ H + \widehat{s}\,\widehat{\lambda}_1 \end{bmatrix}. \tag{130}$$

with

$$\begin{aligned} \alpha_{\lambda_1} &= \delta W_1 - \frac{A}{a^2} \\ \alpha_{\lambda_2} &= \left(\widehat{\lambda}_1\,\widehat{\nu}_x - u\right)\delta W_1 + \left(\widehat{\nu}_y^2 + \widehat{\nu}_z^2\right)\delta W_2 - \widehat{\nu}_x\,\widehat{\nu}_y\,\delta W_3 - \widehat{\nu}_x\,\widehat{\nu}_z\,\delta W_4 \\ \alpha_{\lambda_3} &= \left(\widehat{\lambda}_1\,\widehat{\nu}_z - w\right)\delta W_1 - \widehat{\nu}_x\,\widehat{\nu}_z\,\delta W_2 - \widehat{\nu}_z\,\widehat{\nu}_y\,\delta W_3 + \left(\widehat{\nu}_y^2 + \widehat{\nu}_x^2\right)\delta W_4 \\ \alpha_{\lambda_4} &= \frac{\widehat{s}\,A + \beta^2\,a^2\,B}{2\,\beta^2\,a^2\,\widehat{t}}, \\ \alpha_{\lambda_5} &= -\frac{\widehat{r}\,A + \beta^2\,a^2\,B}{2\,\beta^2\,a^2\,\widehat{t}}. \end{aligned} \tag{131}$$

where

$$\begin{aligned} A &= \left(\frac{q^2}{2}\,\delta W_1 - u\,\delta\,W_2 - v\,\delta W_3 - w\,\delta W_4 + \delta W_5\right)(\gamma - 1) \\ B &= \widehat{\lambda}_1\,\delta W_1 - \widehat{\nu}_x\,\delta\,W_2 - \widehat{\nu}_y\,\delta W_3 - \widehat{\nu}_z\,\delta W_4. \end{aligned}$$

After simplification, we obtain

$$\begin{aligned} A &= \delta\,p \\ B &= -\sqrt{\rho_1\,\rho_2}(\widehat{\nu}_x\,\delta\,u + \widehat{\nu}_y\,\delta\,v + \widehat{\nu}_z\,\delta\,w) = -\sqrt{\rho_1\,\rho_2}\,\delta\,\widehat{\lambda}_1 \end{aligned} \tag{132}$$

Remark:
The quantities occurring in $F^{(i)}$ (λ_1, λ_4, λ_5, p, a, H ...) are obtained from the average (Roe average) quantities of ρ, u, v, w. ■

Remark:
The modification of Turkel affects only the fluxes due to the acoustic wave speeds ($F^{(4)}$ et $F^{(5)}$). ■

Thus, $F = P_c^{-1} \mid P_c D_c \mid \delta W$ is given by

$$F = \left[\begin{array}{lcllcl}
\mid \lambda_1 \mid \alpha_{\lambda_1} & + & \mid \lambda_4 \mid \alpha_{\lambda_4} & + & \mid \lambda_5 \mid \alpha_{\lambda_5} \\
\mid \lambda_1 \mid (\alpha_{\lambda_1} u + \alpha_{\lambda_2}) & + & \mid \lambda_4 \mid \alpha_{\lambda_4} (u + \widehat{r}\,\widehat{\nu}_x) & + & \mid \lambda_5 \mid \alpha_{\lambda_5} (u + \widehat{s}\,\widehat{\nu}_x) \\
\mid \lambda_1 \mid (\alpha_{\lambda_1} v + \alpha) & + & \mid \lambda_4 \mid \alpha_{\lambda_4} (v + \widehat{r}\,\widehat{\nu}_y) & + & \mid \lambda_5 \mid \alpha_{\lambda_5} (v + \widehat{s}\,\widehat{\nu}_y) \\
\mid \lambda_1 \mid (\alpha_{\lambda_1} w + \alpha_{\lambda_3}) & + & \mid \lambda_4 \mid \alpha_{\lambda_4} (w + \widehat{r}\,\widehat{\nu}_z) & + & \mid \lambda_5 \mid \alpha_{\lambda_5} (w + \widehat{s}\,\widehat{\nu}_z) \\
\mid \lambda_1 \mid \left(\alpha_{\lambda_1} \dfrac{q^2}{2} + \alpha_{\lambda_2} u + \alpha v + \alpha_{\lambda_3} w \right) & + & \mid \lambda_4 \mid \alpha_{\lambda_4} \left(H + \widehat{r}\,\widehat{\lambda}_1 \right) & + & \mid \lambda_5 \mid \alpha_{\lambda_5} \left(H + \widehat{s}\,\widehat{\lambda}_1 \right)
\end{array} \right],$$

where

$$\alpha = -\frac{\alpha_{\lambda_2}\,\widehat{\nu}_x + \alpha_{\lambda_3}\,\widehat{\nu}_z}{\widehat{\nu}_y}.$$

Remark:
In the above formula, the parameter β occurs only in the following variables: λ_4, λ_5, α_{λ_4}, α_{λ_5}, $\widehat{r}$ and $\widehat{s}$. ■

B Fourier analysis

B.1 Fourier transform for the first order operator

$$\delta_{x,1} = \frac{1}{2} Trid(-1,0,1) + \frac{1}{2}\eta\, Trid(-1,2,-1)$$

$$\begin{aligned} F(\delta_{x,1}) &= \frac{1}{2}(-e^{-j\omega} + e^{j\omega}) + \frac{1}{2}\eta\,(-e^{-j\omega} + 2 - e^{j\omega}) \\ &= j\,\sin\omega + 2\,\eta\,\sin^2\tfrac{\omega}{2} \\ &= 2\,j\,\sin\tfrac{\omega}{2}\left(\cos\tfrac{\omega}{2} - j\,\eta\,\sin\tfrac{\omega}{2}\right) \end{aligned}$$

B.2 Fourier transform for the second order operator

$$\begin{aligned} \delta_{x,2} = \frac{1}{2} Penta(0,-1,0,1,0) + \frac{1}{2}\bar{\beta}\Big[&\Big\{\frac{1}{2} Penta(1,-3,3,-1,0) + \frac{1}{2} Penta(0,1,-3,3,-1)\Big\} \\ &+ \eta\,\Big\{\frac{1}{2} Penta(1,-3,3,-1,0) - \frac{1}{2} Penta(0,1,-3,3,-1)\Big\}\Big] \end{aligned}$$

$$\begin{aligned} F(\delta_{x,2}) &= \frac{1}{2}(-e^{-j\omega} + e^{j\omega}) + \frac{1}{4}\bar{\beta}\,[\{(e^{-2j\omega} - 3e^{-j\omega} + 3 - e^{j\omega}) + (e^{-j\omega} - 3 + 3e^{j\omega} - e^{2\,j\omega})\} \\ &\qquad + \eta\,\{(e^{-2j\omega} - 3e^{-j\omega} + 3 - e^{j\omega}) - (e^{-j\omega} - 3 + 3e^{j\omega} - e^{2\,j\omega})\}] \\ &= \frac{1}{2}(-e^{-j\omega} + e^{j\omega}) + \frac{\bar{\beta}}{4}\left(e^{-3j\frac{\omega}{2}} - 3e^{-j\frac{\omega}{2}} + 3e^{j\frac{\omega}{2}} - e^{3j\frac{\omega}{2}}\right)\left[e^{-j\frac{\omega}{2}} + e^{j\frac{\omega}{2}} + \eta\,\left(e^{-j\frac{\omega}{2}} - e^{j\frac{\omega}{2}}\right)\right] \\ &= 2\,j\,\sin\tfrac{\omega}{2}\cos\tfrac{\omega}{2} + \frac{\bar{\beta}}{4}\left(8\,j\,\sin^3\tfrac{\omega}{2}\right)\left[2\cos\tfrac{\omega}{2} - 2\,\eta\,j\,\sin\tfrac{\omega}{2}\right] \\ &= 2\,j\,\sin\tfrac{\omega}{2}\left(\cos\tfrac{\omega}{2} + 2\,\bar{\beta}\,\sin^2\tfrac{\omega}{2}\left[\cos\tfrac{\omega}{2} - \eta\,j\,\sin\tfrac{\omega}{2}\right]\right) \end{aligned}$$

B.3 Amplification operator

$$\begin{aligned} F(G_\infty) &= I - (F(\delta_{x,1})(\omega))^{-1}\ F(\delta^{\bar\beta}_{x,2})(\omega) \\ &= 1 - \frac{\cos\frac{\omega}{2}\,(\cos\frac{\omega}{2} + j\,\eta\,\sin\frac{\omega}{2}) + 2\,\bar\beta\,\sin^2\frac{\omega}{2}\,(\cos^2\frac{\omega}{2} + \eta^2\sin^2\frac{\omega}{2})}{\cos^2\frac{\omega}{2} + \eta^2\sin^2\frac{\omega}{2}} \\ &= \frac{\eta^2\,\sin^2\frac{\omega}{2} - j\,\eta\,\cos\frac{\omega}{2}\sin\frac{\omega}{2} - 2\,\bar\beta\sin^2\frac{\omega}{2}\left(\cos^2\frac{\omega}{2} + \eta^2\sin^2\frac{\omega}{2}\right)}{\cos^2\frac{\omega}{2} + \eta^2\sin^2\frac{\omega}{2}} \end{aligned}$$

The modulus of $F(G_\infty)$ can be written as

$$\mid F(G_\infty)\mid = \frac{\sqrt{\sin^4\frac{\omega}{2}\,\left(\eta^2 - 2\bar\beta\,(\cos^2\frac{\omega}{2} + \eta^2\sin^2\frac{\omega}{2})\right)^2 + \eta^2\,\cos^2\frac{\omega}{2}\,\sin^2\frac{\omega}{2}}}{\cos^2\frac{\omega}{2} + \eta^2\sin^2\frac{\omega}{2}}$$

or similarly as

$$\begin{aligned} \mid F(G_\infty)\mid &= \frac{\sqrt{s^2\,\left(\eta^2 - 2\,\bar\beta(1 + (\eta^2-1)\,s)\right)^2 + \eta^2 s\,(1-s)}}{1 + (\eta^2-1)\,s} \\ &= \sqrt{\frac{\eta^2 + 4\,\bar\beta\,(\bar\beta - \eta^2)\,s + 4\,\bar\beta^2(\eta^2-1)\,s^2}{1 + (\eta^2-1)\,s}}\,s \end{aligned}$$

where $s = \sin^2\frac{\omega}{2}$.

Thus

$$S = \sup_{\omega\in\{-\pi/h,\pi/h\}} \mid F(G_\infty)\mid = \sup_{s\in(0,1)} \sqrt{\frac{\eta^2 + 4\,\bar\beta\,(\bar\beta - \eta^2)\,s + 4\,\bar\beta^2(\eta^2-1)\,s^2}{1 + (\eta^2-1)\,s}}\,s.$$

Unité de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unité de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unité de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unité de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unité de recherche INRIA Sophia Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
ISSN 0249-6399