



**HAL**  
open science

## STAGGERED SCHEMES FOR ALL SPEED FLOWS

Raphaelé Herbin, Walid Kheriji, Jean-Claude Latché

► **To cite this version:**

Raphaelé Herbin, Walid Kheriji, Jean-Claude Latché. STAGGERED SCHEMES FOR ALL SPEED FLOWS. ESAIM: Proceedings, 2011, Congrès National de Mathématiques Appliquées et Industrielles, 35, pp.122-150. <10.1051/proc/201235008>. <hal-00636173>

**HAL Id: hal-00636173**

**<https://hal.science/hal-00636173v1>**

Submitted on 26 Oct 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

## STAGGERED SCHEMES FOR ALL SPEED FLOWS

R. HERBIN<sup>1</sup>, W. KHERIJI<sup>2</sup> AND J.-C. LATCHE<sup>3</sup>

**Abstract.** We present in this paper a class of schemes for the numerical simulation of compressible flows. In order to ensure the stability of the discretizations in a wide range of Mach numbers and introduce sufficient decoupling for the numerical resolution, we choose to implement and study pressure correction schemes on staggered meshes. The implicit version of the schemes is also considered for the theoretical study. We give both algorithms for the barotropic Navier Stokes equations, for the full Navier-Stokes equations and for the Euler equations. In each case, we summarize the theoretical results that were recently obtained concerning the stability and consistency of the schemes and present some numerical results which confirm their good performance.

## INTRODUCTION

The aim of this paper is to review some recent contributions to the development of a class of schemes for the simulation of industrial compressible flows. The considered systems of governing equations are coupled and strongly nonlinear, and the applications in view involve complex geometry and flows, possibly combining quasi-steady states with quick transient phases, with strong physical properties contrasts (in particular, for the viscosity, the compressibility and the density). Accordingly, the algorithms are developed so as to realize a compromise between two main requirements: preserve the stability in a wide range of Mach numbers and introduce sufficient decoupling to facilitate the resolution of discrete algebraic systems. Pressure correction methods seem to be a good choice to address these requirements.

This class of schemes was first introduced in the framework of incompressible flows a long time ago [6, 54], and such algorithms are now quite widespread and well understood in this context (see, for example, [44] for an introduction and [20] for a review of most of the variants). Pressure correction schemes are less popular in the context of compressible flows, even though their application to compressible Navier-Stokes equations may also be traced back to the late sixties, with the seminal work of Harlow and Amsden [25, 26], who developed an iterative algorithm (the so-called ICE method) including an elliptic corrector step for the pressure. Later on, pressure correction equations appeared in numerical schemes proposed by several researchers, essentially in the finite-volume framework, using either a collocated [11, 35, 38, 45, 50, 52] or a staggered arrangement [3, 5, 8, 33, 34, 36, 56, 57, 59–61] of unknowns; in the first case, some corrective actions are to be foreseen to avoid the usual odd-even decoupling of the pressure in the low Mach number regime. Some of these algorithms are essentially implicit, the end-of-step solution being then obtained by SIMPLE-like iterative processes [11, 35, 36, 38, 45, 52, 58]. The other schemes [3, 8, 33, 34, 50, 56, 57, 59, 61, 62] are predictor-corrector methods, where basically two steps are performed sequentially: first a semi-explicit decoupled prediction of the momentum or velocity (and possibly

---

<sup>1</sup> Université d'Aix-Marseille, herbin@cmi.univ-mrs.fr

<sup>2</sup> Institut de Radioprotection et de Sûreté Nucléaire, kheriji.walid@gmail.com

<sup>3</sup> Institut de Radioprotection et de Sûreté Nucléaire, jean-claude.latche@irsn.fr

energy, for non-barotropic flows) and, second, a correction step where the end-of step pressure is evaluated and the momentum and velocity are corrected, as in projection methods for incompressible flows. The Characteristic-Based Split (CBS) scheme (see [48] for a recent review or [63] for the seminal paper) was developed in the finite-element context and belongs to this latter class of methods.

In this work, implicit-in-time discretizations are addressed for their (relative) simplicity in view of the theoretical studies; however, non-iterative pressure correction schemes are our main concern for practical computations. We consider here staggered-in-space discretizations, with the aim to build schemes which are stable and accurate at all Mach numbers and, in particular, which boil down when the Mach number tends to zero to a usual algorithm for incompressible flows (or, more generally, for the asymptotic model of vanishing Mach number flows [43]). This last requirement also implies that, if we implement upwinding techniques (and we will have to for stability reasons), upwinding may have to be performed for each equation separately and with respect to the material velocity only. This is in contradiction with the most common strategy adopted for hyperbolic systems, where upwinding is built from the wave structure of the system (see *eg.* [4, 55] for surveys and [10, 23, 24] for analysis of these schemes at low Mach number), and yields algorithms which are used in practice (see, *eg.*, the so-called AUSM family of schemes [41, 42]), but scarcely studied from a theoretical point of view. One of our main concerns here will thus be to bring, as far as possible, theoretical arguments supporting our numerical developments. Let us first recall a (possible) common skeleton of convergence studies in the finite volume context [12]. The proof may usually be decomposed into three steps:

- (i) The first step is to get the existence and some *a priori* estimates on the approximate solution, or, in other words, to obtain stability results for the scheme.
- (ii) Next, up to the extraction of a subsequence, compactness arguments yield the existence of a (possibly weak) limit to a sequence of discrete solutions obtained with a sequence of discretizations the space step and, for unsteady problems, the time step of which tend to zero. At this point, *a priori* estimates may imply some regularity of the limit.
- (iii) Finally, the fact that the limit is a solution to (a weak form) of the continuous problem is proven by passing to the limit in (a weak formulation of) the scheme.

For the problems studied here, namely the compressible Navier-Stokes or Euler equations, the realization of the complete program seems out of reach, due to the lack of control (Step (i)) of space translates of the unknown; hence we obtain a convergence of sequence of discrete solutions (Step (ii)) in a sense too weak to allow the passage to the limit in the scheme (Step (iii)). There is thus no hope at the present stage to prove the convergence of the schemes in the general cases (*i.e.* except for the barotropic viscous Navier-Stokes equations, see [15, 40, 49] for theoretical analysis of the continuous problem and [13, 14, 17] for scheme convergence analysis in the simplified case of the steady Stokes problem), and our theoretical analyses are then necessarily somewhat incomplete. However, in both the barotropic and the non-barotropic cases, and at least for most variants of the schemes, we do get the following results:

- (i) We show that the discrete solution satisfies discrete analogues of the estimates known in the continuous case: positivity of the density and, in the non-barotropic case, of the internal energy, decrease of the total energy, and, for the viscous barotropic flows, control of the velocity in the  $L^2(H^1)$  norm. These estimates allow to prove the existence of at least one solution to the scheme, by topological degree arguments.
- (ii) *Supposing the convergence of the scheme in strong enough norms*, we then show that the limits of sequences of solutions are weak solutions to the continuous problem, which may be seen (and is referred to hereafter) as a consistency property of the schemes.

Finally, we confort these theoretical experiments by numerical tests, performed with the open-source software ISIS [32], developed at IRSN on the basis of the software component library and programming environment PELICANS [51].

This paper is organized as follows. We first introduce the considered space discretizations (Section 1). Then we turn to the barotropic Navier-Stokes equations (Section 2), to the "complete" Navier-Stokes equations

(Section 3), and, finally, to the Euler equations (Section 4); for each case, we present the schemes, summarize the theoretical results and report some numerical tests.

In several theoretical developments, we are lead to use a derived form of a discrete finite volume convection operator (for instance, typically, a convection operator for the kinetic energy, possibly with residual terms, obtained from the finite volume discretization of the convection of the velocity components); an abstract presentation of such computations is given in the Appendix.

## 1. MESHES AND UNKNOWNNS

Let the computational domain  $\Omega$  be an open polygonal subset of  $\mathbb{R}^d$ ,  $d \leq 3$ , and  $\mathcal{M}$  be a partition of  $\Omega$ , supposed to be regular in the usual sense of the finite element literature (eg. [7]). The cells may be:

- for a general domain  $\Omega$ , either convex quadrilaterals ( $d = 2$ ) or hexahedra ( $d = 3$ ) or simplices, both types of meshes being possibly combined in a same mesh,
- for a domain the boundaries of which are hyperplanes normal to a coordinate axis, rectangles ( $d = 2$ ) or rectangular parallelepipeds ( $d = 3$ ) (the faces of which, of course, are then also necessarily normal to a coordinate axis).

By  $\mathcal{E}$  and  $\mathcal{E}(K)$  we denote the set of all  $(d - 1)$ -faces  $\sigma$  of the mesh and of the element  $K \in \mathcal{M}$  respectively. The set of edges included in the boundary of  $\Omega$  is denoted by  $\mathcal{E}_{\text{ext}}$  and the set of internal ones (*i.e.*  $\mathcal{E} \setminus \mathcal{E}_{\text{ext}}$ ) is denoted by  $\mathcal{E}_{\text{int}}$ ; a face  $\sigma \in \mathcal{E}_{\text{int}}$  separating the cells  $K$  and  $L$  is denoted by  $\sigma = K|L$ . The outward normal vector to a face  $\sigma$  of  $K$  is denoted by  $\mathbf{n}_{K,\sigma}$ . For  $K \in \mathcal{M}$  and  $\sigma \in \mathcal{E}$ , we denote by  $|K|$  the measure of  $K$  and by  $|\sigma|$  the  $(d - 1)$ -measure of the face  $\sigma$ . For  $1 \leq i \leq d$ , we denote by  $\mathcal{E}^{(i)} \subset \mathcal{E}$  the subset of the faces of  $\mathcal{E}$  which are perpendicular to the  $i^{\text{th}}$  unit vector of the canonical basis of  $\mathbb{R}^d$ .

The space discretization is staggered, using either the Marker-And Cell (MAC) scheme [26, 27], or non-conforming low-order finite element approximations, namely the Rannacher and Turek element (RT) [53] for quadrilateral or hexahedric meshes, or the Crouzeix-Raviart (CR) element [9] for simplicial meshes.

For all these space discretizations, the degrees of freedom for the pressure, the density and the internal energy are associated to the cells of the mesh  $\mathcal{M}$ , and are denoted by:

$$\{p_K, \rho_K, e_K, K \in \mathcal{M}\}.$$

Let us then turn to the degrees of freedom for the velocity.

- **Rannacher-Turek** or **Crouzeix-Raviart** discretizations – The degrees of freedom for the velocities are located at the center of the faces of the mesh, and we choose the version of the element where they represent the average of the velocity through a face. The set of degrees of freedom reads:

$$\{\mathbf{u}_{\sigma,i}, \sigma \in \mathcal{E}, 1 \leq i \leq d\}.$$

- **MAC** discretization – The degrees of freedom for the  $i^{\text{th}}$  component of the velocity, defined at the centres of the face  $\sigma \in \mathcal{E}^{(i)}$ , are denoted by:

$$\{\mathbf{u}_{\sigma,i}, \sigma \in \mathcal{E}^{(i)}, 1 \leq i \leq d\}.$$

For the definition of the schemes, we need a dual mesh which is defined as follows.

- **Rannacher-Turek** or **Crouzeix-Raviart** discretizations – For the RT or CR discretization, the dual mesh is the same for all the velocity components. When  $K \in \mathcal{M}$  is a simplex, a rectangles or a cuboid, for  $\sigma \in \mathcal{E}(K)$ , we define  $D_{K,\sigma}$  as the cone with basis  $\sigma$  and with vertex the mass center of  $K$ . We thus obtain a partition of  $K$  in  $m$  sub-volumes, where  $m$  is the numbers of faces of the mesh, each sub-volume having the same measure  $|D_{K,\sigma}| = |K|/m$ . We extend this definition to general quadrangles and hexahedra, by supposing that we have built a partition still of equal-volume sub-cells, and with the same connectivities;

note that this is of course always possible, but that such a volume  $D_{K,\sigma}$  may be no longer a cone, since, if  $K$  is far from a parallelogram, it may not be possible to build a cone having  $\sigma$  as basis, the opposite vertex lying in  $K$  and a volume equal to  $|K|/m$ . The volume  $D_{K,\sigma}$  is referred to as the half-diamond cell associated to  $K$  and  $\sigma$ .

For  $\sigma \in \mathcal{E}_{\text{int}}$ ,  $\sigma = K|L$ , we now define the diamond cell  $D_\sigma$  associated to  $\sigma$  by  $D_\sigma = D_{K,\sigma} \cup D_{L,\sigma}$ .

- **MAC** discretization – For the MAC scheme, the dual mesh depends on the component of the velocity. For each of them, its definition differs from the RT or CR one only by the choice of the half-diamond cell, which, for  $K \in \mathcal{M}$  and  $\sigma \in \mathcal{E}(K)$ , is now the rectangle of basis  $\sigma$  and of measure  $|D_{K,\sigma}|$  equal to half the measure of  $K$ .

We denote by  $|D_\sigma|$  the measure of the dual cell  $|D_\sigma|$ , and by  $\varepsilon = D_\sigma|D_{\sigma'}$  the face separating two diamond cells  $D_\sigma$  and  $D_{\sigma'}$  (see Figure 1).

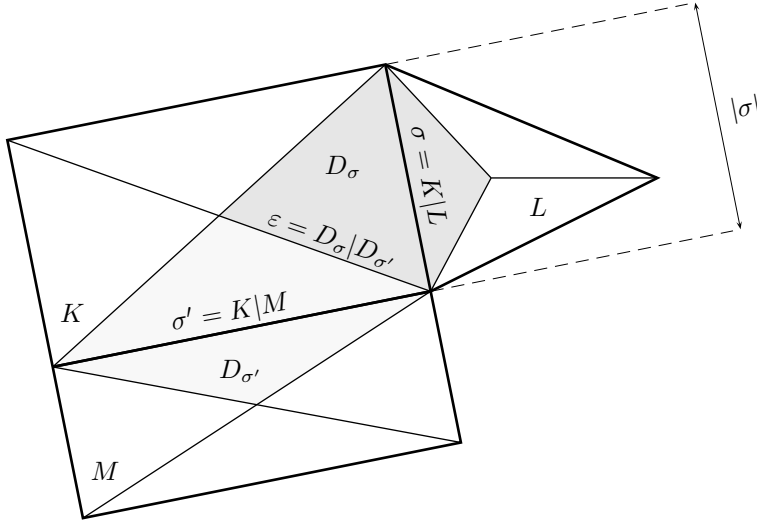


FIGURE 1. Primal and dual meshes for the Rannacher-Turek and Crouzeix-Raviart elements.

## 2. COMPRESSIBLE BAROTROPIC NAVIER-STOKES EQUATIONS

The addressed problem in this section reads:

$$\partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0, \quad (1a)$$

$$\partial_t(\rho \mathbf{u}) + \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p - \operatorname{div}(\boldsymbol{\tau}(\mathbf{u})) = 0, \quad (1b)$$

$$\rho = \varrho(p), \quad (1c)$$

where  $t$  stands for the time,  $\rho$ ,  $\mathbf{u}$  and  $p$  are the density, velocity, and pressure in the flow,  $\boldsymbol{\tau}(\mathbf{u})$  stands for the shear stress tensor, and the function  $\varrho$  is the equation of state. The problem is supposed to be posed over  $\Omega \times (0, T)$ , where  $(0, T)$  is a finite time interval. This system must be supplemented by suitable boundary conditions, and initial conditions for  $\rho$  and  $\mathbf{u}$ , the initial condition for  $\rho$  being supposed positive. The closure relation for  $\boldsymbol{\tau}(\mathbf{u})$  is assumed to be:

$$\boldsymbol{\tau}(\mathbf{u}) = \mu(\nabla \mathbf{u} + \nabla^t \mathbf{u}) - \frac{2\mu}{3} \operatorname{div} \mathbf{u} I,$$

where  $\mu$  stands for a non-negative parameter, possibly depending on  $\mathbf{x}$ . When the viscous term  $\boldsymbol{\tau}(\mathbf{u})$  vanishes, the system (1) becomes hyperbolic.

Let us denote by  $E_k$  the kinetic energy  $E_k = \frac{1}{2}\rho|\mathbf{u}|^2$ . Taking the inner product of (1b) by  $\mathbf{u}$  yields, after formal compositions of partial derivatives and using (1a):

$$\partial_t E_k + \operatorname{div}(E_k \mathbf{u}) + \nabla p \cdot \mathbf{u} = \operatorname{div}(\boldsymbol{\tau}(\mathbf{u})) \cdot \mathbf{u}. \quad (2)$$

This relation is referred to as the kinetic energy balance.

Let us now define the function  $\mathcal{P}$ , from  $(0, +\infty)$  to  $\mathbb{R}$ , as a primitive of  $s \mapsto \wp(s)/s^2$ , where  $\wp = \varrho^{-1}$ ; this quantity is often called the elastic potential. Let  $\mathcal{H}$  be the function defined by  $\mathcal{H}(s) = s\mathcal{P}(s)$ ,  $\forall s \in (0, +\infty)$ ; it may easily be checked that  $\rho\mathcal{H}'(\rho) - \mathcal{H}(\rho) = \wp(\rho)$ ; therefore, by a formal computation detailed in the appendix (see Equation (34)), multiplying (1a) by  $\mathcal{H}'(\rho)$  yields:

$$\partial_t(\mathcal{H}(\rho)) + \operatorname{div}(\mathcal{H}(\rho) \mathbf{u}) + p \operatorname{div}(\mathbf{u}) = 0. \quad (3)$$

Let us denote by  $\mathcal{S}$  the quantity  $\mathcal{S} = E_k + \mathcal{H}(\rho)$ . Summing (2) and (3), we get:

$$\partial_t \mathcal{S} + \operatorname{div}((\mathcal{S} + p) \mathbf{u}) - \operatorname{div}(\boldsymbol{\tau}(\mathbf{u}) \mathbf{u}) = -\boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u}. \quad (4)$$

This shows that, in the hyperbolic case,  $\mathcal{S}$  is an entropy of the system, and an entropy solution to (1) is thus required to satisfy:

$$\int_0^T \int_{\Omega} [-\mathcal{S} \partial_t \varphi - (\mathcal{S} + p) \mathbf{u} \cdot \nabla \varphi] \, d\mathbf{x} \, dt - \int_{\Omega} \mathcal{S}(\mathbf{x}, 0) \varphi(\mathbf{x}, 0) \, d\mathbf{x} \leq 0, \quad \forall \varphi \in C_c^\infty(\Omega \times [0, T]), \varphi \geq 0. \quad (5)$$

Then, formally, if we suppose that the velocity is prescribed to zero at the boundary (the normal velocity, in the hyperbolic case), integrating (4) yields, since the viscous dissipation term  $\boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u}$  is non-negative:

$$\frac{d}{dt} \int_{\Omega} \left[ \frac{1}{2} \rho |\mathbf{u}|^2 + \mathcal{H}(\rho) \right] \, d\mathbf{x} \leq 0. \quad (6)$$

Since the function  $\mathcal{P}$  is increasing, Inequality (6) provides an estimate of the solution.

We study two schemes for the numerical solution of System (1) which differ by the time discretization: the first one is implicit, and the second one is a non-iterative pressure-correction scheme introduced in [16]. This latter algorithm (and, by an easy extension, also the first one) was shown in [16] to have at least one solution, to provide solutions satisfying  $\rho > 0$  (and so  $p > 0$ ) and to be unconditionally stable, in the sense that its (their) solution(s) satisfies a discrete analogue of Inequality (6). The results presented in this section complement this work in several directions. For the implicit scheme:

- We first derive discrete analogues of (2) and (3), the first (local) balance equation, *i.e.* the discrete kinetic energy balance, being obtained on dual cells, and the second one, *i.e.* the elastic potential balance, on primal cells.

These equations are used a first time to obtain the stability of the scheme by a simple integration in space (*i.e.* summation over the primal and dual control volumes).

- Second, in one space dimension and for the hyperbolic case, we prove that the limit of any convergent sequence of solutions to the scheme is a weak solution to the problem (in fact, satisfies the Rankine-Hugoniot conditions, and thus exhibits "correct" shocks).
- Finally, passing to the limit on the discrete kinetic energy and elastic potential balances, we show that such a limit also satisfies the entropy inequality (5).

For the pressure correction scheme, the results are essentially the same: the scheme is unconditionally stable, and the passage to the limit in the scheme shows that, in case of convergence, the predicted and end-of-step velocities necessarily tend to the same function, and that the limit is a weak solution to the problem, satisfying the entropy inequality.

Numerical tests, performed with the pressure correction scheme, confort these theoretical results.

We first summarize in this section the obtained theoretical results (Sections 2.1 and 2.2.3) which are detailed in [28]. Then we show results of a numerical test (Section 2.2.4), extracted from a more comprehensive study also addressing an extension of the scheme to two-phase flows, presented in [37].

## 2.1. An implicit scheme scheme

### 2.1.1. The scheme

Let us consider a uniform partition  $0 = t_0 < t_1 < \dots < t_N = T$  of the time interval  $(0, T)$ , and let  $\delta t = t_{n+1} - t_n$  for  $n = 0, 1, \dots, N - 1$  be the constant time step.

We begin with the discretization of the mass balance equation (1a). For both the MAC and RT or CR discretizations, let us denote by  $\mathbf{u}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma}$  the outward normal velocity to the face  $\sigma$  of  $K$ , which is computed, for the RT and CR elements, by taking the inner product of the velocity at the face with the outward normal vector (as implied by the notation) and which is given, for the MAC scheme, by the value of the component of the velocity at the center of the face (up to a change of sign). The discrete equations are obtained by an upwind finite volume discretization and read:

$$\forall K \in \mathcal{M}, \quad \frac{|K|}{\delta t} (\rho_K^{n+1} - \rho_K^n) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma} = 0, \quad \text{with } F_{K,\sigma} = |\sigma| \mathbf{u}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma} \tilde{\rho}_\sigma^{n+1}, \quad (7)$$

and where  $\tilde{\rho}_\sigma^{n+1}$  is the upwind approximation of  $\rho^{n+1}$  at the face  $\sigma$  with respect to  $\mathbf{u}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma}$ . This approximation ensures that  $\rho^{n+1} > 0$  if  $\rho^n > 0$  and if the density is prescribed to a positive value at inflow boundaries.

For both MAC and RT or CR discretizations, we denote by  $(\text{div}\tau(\mathbf{u}^{n+1}))_{\sigma,i}$  an approximation of the  $i$ -th component of the viscous term associated to  $\sigma$ , and we denote by  $(\nabla p^n)_{\sigma,i}$  the  $i$ -th component of the discrete pressure gradient at the face  $\sigma$ . With these notations, we are able to write the following general form of the approximation of the momentum balance equation:

$$\begin{aligned} \frac{|D_\sigma|}{\delta t} (\rho_\sigma^{n+1} \mathbf{u}_{\sigma,i}^{n+1} - \rho_\sigma^n \mathbf{u}_{\sigma,i}^n) + \sum_{\varepsilon \in \mathcal{E}(D_\sigma)} F_{\sigma,\varepsilon}^{n+1} \mathbf{u}_{\varepsilon,i}^{n+1} \\ + |D_\sigma| (\nabla p^{n+1})_{\sigma,i} - |D_\sigma| (\text{div}\tau(\mathbf{u}^{n+1}))_{\sigma,i} = 0, \end{aligned} \quad (8)$$

for  $1 \leq i \leq d$ , and for  $\sigma \in \mathcal{E}$  in the case of the RT or CR discretizations, and  $\sigma \in \mathcal{E}^{(i)}$  in the case of the MAC scheme. In this relation,  $\rho_\sigma^{n+1}$  and  $\rho_\sigma^n$  stand for an approximation of the density on the face  $\sigma$  at time  $t^{n+1}$  and  $t^n$  respectively (which must not be confused with the upstream density  $\tilde{\rho}_\sigma$  used in the mass balance),  $F_{\sigma,\varepsilon}^{n+1}$  is the discrete mass flux through the dual face  $\varepsilon$  outward  $D_\sigma$ , and  $\mathbf{u}_{\varepsilon,i}^{n+1}$  stands for an approximation of  $\mathbf{u}_i^{n+1}$  on  $\varepsilon$  which may be chosen either centred or upwind.

The finite element discretization of the  $i$ -th component of the pressure gradient term reads:

$$|D_\sigma| (\nabla p^{n+1})_{\sigma,i} = - \sum_{M \in \mathcal{M}} \int_M p^{n+1} \text{div}\varphi_\sigma^{(i)} \, d\mathbf{x},$$

with  $\varphi_\sigma^{(i)}$  reads  $\varphi_\sigma^{(i)} = \varphi_\sigma \mathbf{e}^{(i)}$ , where  $\varphi_\sigma$  is the finite element shape function associated to  $\sigma$  and  $\mathbf{e}^{(i)}$  stands for the  $i^{\text{th}}$  vector of the canonical basis of  $\mathbb{R}^d$ . Since the pressure is piecewise constant, using the definition of the RT or CR shape functions, an easy computation yields for an internal face  $\sigma = K|L$  :

$$|D_\sigma|(\nabla p^{n+1})_{\sigma,i} = |\sigma| (p_L^{n+1} - p_K^{n+1}) \mathbf{n}_{K,\sigma} \cdot \mathbf{e}^{(i)},$$

This expression coincides with the discrete gradient in the MAC discretization.

The finite element discretization of the viscous term  $(\text{div}\tau(\mathbf{u}^{n+1}))_{\sigma,i}$ , associated to  $\sigma$  and to the component  $i$ , reads:

$$|D_\sigma|(\text{div}\tau(\mathbf{u}^{n+1}))_{\sigma,i} = -\mu \sum_{K \in \mathcal{M}} \int_K \nabla \mathbf{u}^{n+1} \cdot \nabla \varphi_\sigma^{(i)} - \frac{\mu}{3} \sum_{K \in \mathcal{M}} \int_K \text{div} \mathbf{u}^{n+1} \text{div} \varphi_\sigma^{(i)}.$$

The MAC discretization of this same viscous term is detailed in [2].

The main motivation to implement a finite volume approximation for the first two terms in (8) is to obtain a discrete equivalent of the kinetic energy balance (see next section). For this result to be valid, the necessary condition is that the convection operator vanishes for a constant velocity, *i.e.* that the following discrete mass balance over the diamond cells is satisfied [1, 16]:

$$\forall \sigma \in \mathcal{E}_{\text{int}}, \quad \frac{|D_\sigma|}{\delta t} (\rho_\sigma^{n+1} - \rho_\sigma^n) + \sum_{\varepsilon \in \mathcal{E}(D_\sigma)} F_{\sigma,\varepsilon}^{n+1} = 0. \quad (9)$$

This governs the choice for the definition of the density approximation  $\rho_\sigma$  and the mass fluxes  $F_{\sigma,\varepsilon}$ . The density  $\rho_\sigma$  is defined by a weighted average:  $\forall \sigma \in \mathcal{E}_{\text{int}}, \sigma = K|L, |D_\sigma| \rho_\sigma = |D_{K,\sigma}| \rho_K + |D_{L,\sigma}| \rho_L$ . For a dual edge  $\varepsilon$  included in the primal cell  $K$ , the flux  $F_{\sigma,\varepsilon}$  is computed as a linear combination (with constant coefficients, *i.e.* independent of the edge and the cell) of the mass fluxes through the faces of  $K$ , *i.e.* the quantities  $(F_{K,\sigma}^{n+1})_{\sigma \in \mathcal{E}(K)}$  appearing in the discrete mass balance (7). We do not give here this set of coefficients, and refer to [1, 18, 31] for a detailed construction of this approximation.

### 2.1.2. Kinetic energy balance, elastic potential identity and stability

We begin by deriving a discrete kinetic energy balance equation. Let  $\delta^{\text{up}}$  be a coefficient defined by  $\delta^{\text{up}} = 1$  if an upwind discretization is used for the convection term in the momentum balance equation (8) and  $\delta^{\text{up}} = 0$  in the centered case. With this notation, the momentum balance equation reads:

$$\begin{aligned} \frac{|D_\sigma|}{\delta t} (\rho_\sigma^{n+1} \mathbf{u}_{\sigma,i}^{n+1} - \rho_\sigma^n \mathbf{u}_{\sigma,i}^n) + \sum_{\varepsilon = D_\sigma | D_{\sigma'}} \frac{1}{2} F_{\sigma,\varepsilon}^{n+1} (\mathbf{u}_{\sigma,i}^{n+1} + \mathbf{u}_{\sigma',i}^{n+1}) \\ + \delta^{\text{up}} \sum_{\varepsilon = D_\sigma | D_{\sigma'}} \frac{1}{2} |F_{\sigma,\varepsilon}^{n+1}| (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma',i}^{n+1}) + |D_\sigma| (\nabla p^{n+1})_{\sigma,i} - |D_\sigma| (\text{div}\tau(\mathbf{u}^{n+1}))_{\sigma,i} = 0. \end{aligned}$$

Taking the inner product of this equation with the corresponding velocity unknown, *i.e.*  $\mathbf{u}_{\sigma,i}^{n+1}$ , yields  $T_{\sigma,i}^{\text{conv}} + T_{\sigma,i}^{\text{up}} + T_{\sigma,i}^{p,\tau} = 0$ , with:

$$\begin{aligned} T_{\sigma,i}^{\text{conv}} &= \left[ \frac{|D_\sigma|}{\delta t} (\rho_\sigma^{n+1} \mathbf{u}_{\sigma,i}^{n+1} - \rho_\sigma^n \mathbf{u}_{\sigma,i}^n) + \sum_{\varepsilon = D_\sigma | D_{\sigma'}} \frac{1}{2} F_{\sigma,\varepsilon}^{n+1} (\mathbf{u}_{\sigma,i}^{n+1} + \mathbf{u}_{\sigma',i}^{n+1}) \right] \mathbf{u}_{\sigma,i}^{n+1}, \\ T_{\sigma,i}^{\text{up}} &= \delta^{\text{up}} \left[ \sum_{\varepsilon = D_\sigma | D_{\sigma'}} \frac{1}{2} |F_{\sigma,\varepsilon}^{n+1}| (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma',i}^{n+1}) \right] \mathbf{u}_{\sigma,i}^{n+1}, \\ T_{\sigma,i}^{p,\tau} &= |D_\sigma| (\nabla p^{n+1})_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1} - |D_\sigma| (\text{div}\tau(\mathbf{u}^{n+1}))_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1}. \end{aligned}$$

Lemma A.2, applied on the dual mesh, yields:

$$T_{\sigma,i}^{\text{conv}} = \frac{1}{2} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^{n+1} (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^n (\mathbf{u}_{\sigma,i}^n)^2 \right] + \frac{1}{2} \sum_{\varepsilon=D_\sigma|D_{\sigma'}} F_{\sigma,\varepsilon}^{n+1} \mathbf{u}_{\sigma,i}^{n+1} \mathbf{u}_{\sigma',i}^{n+1} + \frac{|D_\sigma|}{2\delta t} \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma,i}^n)^2.$$

Let us define  $R_{\sigma,i}^{n+1}$  by the sum of  $T_{\sigma,i}^{\text{up}}$  and the last term of  $T_{\sigma,i}^{\text{conv}}$ :

$$R_{\sigma,i}^{n+1} = \frac{1}{2} \frac{|D_\sigma|}{\delta t} \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma,i}^n)^2 + \delta^{\text{up}} \left[ \sum_{\varepsilon=D_\sigma|D_{\sigma'}} \frac{1}{2} |F_{\sigma,\varepsilon}^{n+1}| (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma',i}^{n+1}) \right] \mathbf{u}_{\sigma,i}^{n+1}. \quad (10)$$

With this notation, we thus obtain the following relation:

$$\begin{aligned} \frac{1}{2} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^{n+1} (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^n (\mathbf{u}_{\sigma,i}^n)^2 \right] + \frac{1}{2} \sum_{\varepsilon=D_\sigma|D_{\sigma'}} F_{\sigma,\varepsilon}^{n+1} \mathbf{u}_{\sigma,i}^{n+1} \mathbf{u}_{\sigma',i}^{n+1} \\ + |D_\sigma| (\nabla p^{n+1})_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1} - |D_\sigma| (\text{div} \tau(\mathbf{u}^{n+1}))_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1} = -R_{\sigma,i}^{n+1}. \end{aligned} \quad (11)$$

We recognize at the left-hand side a discrete kinetic energy balance, *i.e.* a reasonable discretization of Equation (2), with a conservative finite volume discretization of the kinetic energy convection terms. The right-hand side consists in a numerical residual, the sign of which will be studied later.

We now turn to the elastic potential balance. Multiplying the discrete mass balance equation (7) by  $\mathcal{H}'(\rho_K)$  and invoking Lemma A.1 yields,  $\forall K \in \mathcal{M}$ :

$$\frac{|K|}{\delta t} (\mathcal{H}(\rho_K^{n+1}) - \mathcal{H}(\rho_K^n)) + \sum_{\sigma \in \mathcal{E}(K)} |\sigma| [\mathcal{H}(\bar{\rho}_\sigma^{n+1}) + p_K] \mathbf{u}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma} = -R_K^{n+1}, \quad (12)$$

with:

$$R_K^{n+1} = \frac{1}{2} \frac{|K|}{\delta t} \mathcal{H}''(\bar{\rho}_K^{n,n+1}) (\rho_K^{n+1} - \rho_K^n)^2 - \frac{1}{2} \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{u}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma} \mathcal{H}''(\bar{\rho}_\sigma^{n+1}) (\rho_\sigma^{n+1} - \rho_K^{n+1})^2,$$

where  $\bar{\rho}_K^{n,n+1} \in [\min(\rho_K^{n+1}, \rho_K^n), \max(\rho_K^{n+1}, \rho_K^n)]$  and, for  $\sigma \in \mathcal{E}(K)$ ,  $\bar{\rho}_\sigma^{n+1} \in [\min(\rho_\sigma^{n+1}, \rho_K^{n+1}), \max(\rho_\sigma^{n+1}, \rho_K^{n+1})]$ .

Equation (12) is a finite volume discretization of the (non conservative) elastic potential balance (3), with a non positive residual term, thanks to the fact that the function  $\mathcal{H}$  is convex and that an upwind approximation of the density is used in the mass balance.

The stability of the scheme is then obtained by summing:

- (i) Equation (11) over the components  $i$  and the faces  $\sigma \in \mathcal{E}$  for the RT or CR discretizations, and over  $i$  and  $\sigma \in \mathcal{E}^{(i)}$  for the MAC scheme,
- (ii) Equation (12) over  $K \in \mathcal{M}$ ,
- (iii) and, finally, the two obtained relations.

Let us suppose that the velocity vanishes at the boundary, and let us then invoke three arguments. First, the discrete gradient and divergence operators are dual with respect to the  $L^2$  inner product, in the sense that:

$$\sum_{i,\mathcal{E}} |D_\sigma| (\nabla p^{n+1})_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1} + \sum_{K \in \mathcal{M}} p_K \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \mathbf{u}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma} = 0,$$

where the notation  $\sum_{i,\mathcal{E}}$  means that we sum over the component index  $i$  and on  $\sigma \in \mathcal{E}$  for the RT and CR discretizations, and on  $i$  and  $\sigma \in \mathcal{E}^{(i)}$  for the MAC scheme. Second, we suppose that (see Section 3):

$$\sum_{i,\mathcal{E}} |D_\sigma| (\operatorname{div} \tau(\mathbf{u}^{n+1}))_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1} \leq 0.$$

Third, reordering the summations yields, for the part of the remainder of the momentum balance equation associated to the upwinding:

$$\sum_{i,\mathcal{E}} T_{\sigma,i}^{\text{up}} = \delta^{\text{up}} \sum_{i,\bar{\mathcal{E}}} \sum_{(\varepsilon=D_\sigma|D_{\sigma'})} \frac{1}{2} |F_{\sigma,\varepsilon}^{n+1}| (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma',i}^{n+1})^2 \geq 0,$$

where the notation  $\sum_{i,\mathcal{E}} \sum_{(\varepsilon=D_\sigma|D_{\sigma'})}$  means that we perform the sum over  $i$  and the faces of the dual mesh associated to the component  $i$  of the velocity, and that, for a face  $\varepsilon$  in the sum, the two adjacent dual cells are denoted by  $D_\sigma$  and  $D'_{\sigma'}$ . Finally, since the conservative fluxes vanish in the summation, we thus get:

$$\frac{1}{2} \sum_{i,\mathcal{E}} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^{n+1} (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^n (\mathbf{u}_{\sigma,i}^n)^2 \right] + \sum_{K \in \mathcal{M}} \frac{|K|}{\delta t} (\mathcal{H}(\rho_K^{n+1}) - \mathcal{H}(\rho_K^n)) \leq 0, \quad (13)$$

which is a discrete analogue to (6).

### 2.1.3. Passing to the limit in the scheme (1D case)

We focus in this section on the inviscid 1D form of Problem (1), and show that, if a sequence of solutions is controlled in suitable norms and converges to a limit, this latter necessarily satisfies a (part of the) weak formulation of the continuous problem.

Let  $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m \in \mathbb{N}}$  be a sequence of meshes and time steps, such that the time step  $\delta t^{(m)}$  and the size  $h^{(m)}$  of the mesh  $\mathcal{M}^{(m)}$ , defined by:

$$h^{(m)} = \sup_{K \in \mathcal{M}^{(m)}} \operatorname{diam}(K),$$

tend to zero as  $m \rightarrow \infty$ .

Let  $\rho^{(m)}$ ,  $p^{(m)}$  and  $u^{(m)}$  be the solution given by the scheme with the mesh  $\mathcal{M}^{(m)}$  and the time step  $\delta t^{(m)}$ , or, more precisely speaking, a 1D version of the scheme which may be obtained by taking the MAC variant, only one horizontal stripe of meshes, supposing that the vertical component of the velocity (the degree of freedom of which are located on the top and bottom boundaries) vanishes, and that the measure of the faces is equal to 1. To the discrete unknowns, we associate piecewise constant functions on time intervals and on primal or dual meshes, so the density  $\rho^{(m)}$ , the pressure  $p^{(m)}$  and the velocity  $u^{(m)}$  are defined almost everywhere on  $\Omega \times (0, T)$  by:

$$\begin{aligned} \rho^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{K \in \mathcal{M}} (\rho^{(m)})_K^n \mathcal{X}_K \mathcal{X}_{(n,n+1)}, & p^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{K \in \mathcal{M}} (p^{(m)})_K^n \mathcal{X}_K \mathcal{X}_{(n,n+1)}, \\ u^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{\sigma \in \mathcal{E}} (u^{(m)})_\sigma^n \mathcal{X}_{D_\sigma} \mathcal{X}_{(n,n+1)}, \end{aligned}$$

where  $\mathcal{X}_K$ ,  $\mathcal{X}_{D_\sigma}$  and  $\mathcal{X}_{(n,n+1)}$  stand for the characteristic function of  $K$ ,  $D_\sigma$  and the interval  $(t^n, t^{n+1})$  respectively.

We suppose a uniform control on the translates in space and time of the sequence of solutions, which we now state. For discrete function  $q$  and  $v$  defined on the primal and dual mesh, respectively, we define a discrete

$L^1((0, T); BV(\Omega))$  norm by:

$$\|q\|_{\mathcal{T},x,BV} = \sum_{n=0}^N \delta t \sum_{\sigma \in \mathcal{E}, \sigma=K|L} |q_L^n - q_K^n|, \quad \|v\|_{\mathcal{T},x,BV} = \sum_{n=0}^N \delta t \sum_{\varepsilon \in \bar{\mathcal{E}}, \sigma=D_\sigma|D'_\sigma} |v_\sigma^n - v_{\sigma'}^n|,$$

and a discrete  $L^1(\Omega; BV((0, T)))$  norm by:

$$\|q\|_{\mathcal{T},t,BV} = \sum_{K \in \mathcal{M}} h_K \sum_{n=0}^{N-1} |q_K^{n+1} - q_K^n|, \quad \|v\|_{\mathcal{T},t,BV} = \sum_{\sigma \in \mathcal{E}} h_\sigma \sum_{n=0}^{N-1} |v_\sigma^{n+1} - v_\sigma^n|,$$

where, for  $\sigma = K|L$ ,  $h_\sigma = (h_K + h_L)/2$ . We suppose the following uniform bounds of the sequence of solutions with respect to these two norms:

$$\|\rho^{(m)}\|_{\mathcal{T},x,BV} + \|p^{(m)}\|_{\mathcal{T},x,BV} + \|u^{(m)}\|_{\mathcal{T},x,BV} \leq C, \quad \forall m \in \mathbb{N}, \quad (14)$$

and:

$$\|u^{(m)}\|_{\mathcal{T},t,BV} \leq C, \quad \forall m \in \mathbb{N}. \quad (15)$$

A weak solution to the continuous problem satisfies, for any  $\varphi \in C_c^\infty([0, T] \times \Omega)$ :

$$- \int_{\Omega \times (0, T)} \left[ \rho \partial_t \varphi + \rho u \partial_x \varphi \right] dx dt - \int_{\Omega} \rho(x, 0) \varphi(x, 0) dx = 0, \quad (16a)$$

$$- \int_{\Omega \times (0, T)} \left[ \rho u \partial_t \varphi + (\rho u^2 + p) \partial_x \varphi \right] dx dt - \int_{\Omega} \rho(x, 0) u(x, 0) \varphi(x, 0) dx = 0, \quad (16b)$$

$$\rho = \varrho(p). \quad (16c)$$

Note that these relations are not sufficient to define a weak solution to the problem, since they do not imply anything about the boundary conditions. However, they allow to derive the Rankine-Hugoniot conditions ; so, if we show that they are satisfied by the limit of a sequence of solutions to the discrete problem, this implies, loosely speaking, that *the scheme computes the right shocks*, which is the result we are seeking. It is stated in the following theorem.

**Theorem 2.1.** *Let  $\Omega$  be an open bounded interval of  $\mathbb{R}$ . Let  $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m \in \mathbb{N}}$  be a sequence of meshes and time steps, such that  $h^{(m)}$  and  $\delta t^{(m)}$  tend to zero as  $m$  tends to infinity. Let  $(\rho^{(m)}, p^{(m)}, u^{(m)})_{m \in \mathbb{N}}$  be the corresponding sequence of solutions. We suppose that this sequence satisfies (15) and (14) and converges in  $L^r((0, T) \times \Omega)^3$ , for  $1 \leq r < \infty$ , to  $(\bar{\rho}, \bar{p}, \bar{u}) \in L^\infty((0, T) \times \Omega)^3$ .*

*Then the limit  $(\bar{\rho}, \bar{p}, \bar{u})$  satisfies the system (16) and the entropy condition (5).*

*Proof.* The passage to the limit in the equations of the scheme is technical, but invokes rather standard arguments.

Obtaining the entropy condition is more intricate. We need to pass to the limit in the kinetic energy balance (11) and in the elastic potential balance (12) simultaneously. To this purpose, for  $\varphi \in C_c^\infty([0, T] \times \Omega)$ , we define two interpolates: one is defined over the dual cells and is used as a test function for (11) and the second one is defined over the primal cells, and is used as a test function for (12). We then pass to the limit in the "differential terms" of these discrete equations, and disregard the non-negative residuals (at the left-hand side). A problem is posed by the residual associated to the upwinding, which reads:

$$R_{\sigma,i}^{n+1} = \left[ \sum_{\varepsilon=D_\sigma|D_{\sigma'}} \frac{1}{2} |F_{\sigma,\varepsilon}^{n+1}| (\mathbf{u}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma',i}^{n+1}) \right] \mathbf{u}_{\sigma,i}^{n+1},$$

and the sign of which is unknown. To get an intuition of how to deal with this term, let us remark that it may be seen as a discrete analogue to a diffusion term  $-\mu u \Delta u$  with a numerical viscosity  $\mu$  tending to zero as the space step. Let us now compare this term to  $\mu |\nabla u|^2$ , in the sense of distributions. For  $\psi$  a regular function with a compact support, remarking that  $-\mu u \Delta u - \mu |\nabla u|^2 = -\operatorname{div}(\mu u \nabla u)$ , we get:

$$\int_0^T \int_{\Omega} [-\mu u \Delta u - \mu |\nabla u|^2] \psi \, d\mathbf{x} \, dt = \int_0^T \int_{\Omega} \mu u \nabla u \cdot \nabla \psi \, d\mathbf{x} \, dt \leq C_{\psi} \|u\|_{L^{\infty}} \|u\|_{W^{1,1}},$$

and therefore, if  $\|u\|_{L^{\infty}}$  and  $\|u\|_{W^{1,1}}$  are bounded, the difference between  $-\mu u \Delta u$  and  $\mu |\nabla u|^2$  behaves like  $\mu$ . Returning at the discrete level, this computation suggests that  $R_{\sigma}$  behaves at the limit as a dissipation term (*i.e.* a discrete equivalent of  $\mu |\nabla u|^2$ ), the sign of which is guaranteed. The same argument is used in a different way in the non-barotropic case: the "viscous term"  $R_{\sigma}$  is compensated in the internal energy balance by a "dissipation term" (see Section 4.1).  $\square$

**Remark 2.2** (Control of the translates). In the assumptions of Theorem 2.1, we can sharpen (14) and (15). Indeed, to prove that the limit is a weak solution, it is sufficient to have:

$$\lim_{m \rightarrow +\infty} h^{(m)} \left[ \|\rho^{(m)}\|_{\mathcal{T},x,BV} + \|p^{(m)}\|_{\mathcal{T},x,BV} + \|u^{(m)}\|_{\mathcal{T},x,BV} \right] = 0.$$

In addition, this estimate may be proven (and not supposed) by adding to the scheme a numerical diffusion scaled by  $(h^{(m)})^{\beta}$ , with  $0 < \beta < 2$ . To obtain that the limit is the entropy weak solution, the following assumption is sufficient:

$$\lim_{m \rightarrow +\infty} \delta t \left[ \|u^{(m)}\|_{\mathcal{T},t,BV} \right] = 0.$$

## 2.2. A pressure correction scheme

### 2.2.1. The scheme

In this section, we derive the pressure correction scheme from the implicit scheme. The first step, as usual, is to compute a tentative velocity by solving the momentum balance equation with the beginning-of-step pressure. Then, the velocity is corrected and the other variables are advanced in time, in the so-called correction step. For stability reasons, or, in other words, to be able to derive a kinetic energy balance, we need that the mass balance over the dual cells (9) holds; since the mass balance is not yet solved when performing the prediction step, this leads us to do a time shift of the density at this step.

In the time semi-discrete setting, the proposed algorithm reads:

1 - **Pressure renormalization step** – Solve the following elliptic problem for  $\tilde{p}^{n+1}$ :

$$\operatorname{div} \left[ \frac{1}{\rho^n} \nabla \tilde{p}^{n+1} \right] = \operatorname{div} \left[ \frac{1}{(\rho^n \rho^n)^{1/2}} \nabla p^n \right] \quad (17)$$

2 - **Prediction step** – Solve the following semi-discrete linearized momentum balance equation for  $\tilde{\mathbf{u}}^{n+1}$ :

$$\frac{\rho^n \tilde{\mathbf{u}}^{n+1} - \rho^{n-1} \mathbf{u}^n}{\delta t} + \operatorname{div}(\rho^n \tilde{\mathbf{u}}^{n+1} \otimes \mathbf{u}^n) + \nabla \tilde{p}^{n+1} - \operatorname{div}(\boldsymbol{\tau}(\tilde{\mathbf{u}}^{n+1})) = 0. \quad (18)$$

3 - **Correction step** – Solve (simultaneously) the following non linear equations for  $p^{n+1}$ ,  $\mathbf{u}^{n+1}$  and  $\rho^{n+1}$ :

$$\rho^n \frac{\mathbf{u}^{n+1} - \tilde{\mathbf{u}}^{n+1}}{\delta t} + \nabla(p^{n+1} - \tilde{p}^{n+1}) = 0, \quad (19a)$$

$$\frac{\rho^{n+1} - \rho^n}{\delta t} + \operatorname{div}(\rho^{n+1} \mathbf{u}^{n+1}) = 0, \quad (19b)$$

$$\rho^{n+1} = \varrho(p^{n+1}). \quad (19c)$$

The solution of Step 3 is performed by combining equations (19a) and (19b), therefore obtaining a non-linear elliptic problem for the pressure, which reads in the time semi-discrete setting:

$$\frac{\varrho(p^{n+1}) - \rho^n}{\delta t^2} - \operatorname{div} \left[ \frac{\rho^{n+1}}{\rho^n} \nabla(p^{n+1} - \tilde{p}^{n+1}) \right] = -\frac{1}{\delta t} \operatorname{div}(\rho^{n+1} \tilde{\mathbf{u}}^{n+1}).$$

The fully discrete equations are obtained from the implicit scheme by a mere change in time levels, except for Equations (17) and (19a), which are new. The first one is obtained by using the discrete gradient and divergence operators already introduced, and reads:

$$\forall K \in \mathcal{M}, \quad \sum_{\sigma=K|L} \frac{1}{\rho_\sigma^n} \frac{|\sigma|^2}{|D_\sigma|} (\tilde{p}_K^{n+1} - \tilde{p}_L^{n+1}) = \sum_{\sigma=K|L} \frac{1}{\sqrt{\rho_\sigma^n \rho_\sigma^{n-1}}} \frac{|\sigma|^2}{|D_\sigma|} (p_K^n - p_L^n).$$

Relation (19a) is discretized similarly to the momentum balance (8), *i.e.* a finite volume technique is used for the unsteady term in both the MAC, RT and CR discretizations:

$$\frac{|D_\sigma|}{\delta t} \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1} - \tilde{\mathbf{u}}_{\sigma,i}^{n+1}) + |D_\sigma| \left[ (\nabla p^{n+1})_{\sigma,i} - (\nabla \tilde{p}^{n+1})_{\sigma,i} \right] = 0,$$

for  $1 \leq i \leq d$ , and for  $\sigma \in \mathcal{E}$  in the case of the RT or CR discretizations, and  $\sigma \in \mathcal{E}^{(i)}$  in the case of the MAC scheme.

### 2.2.2. Stability and kinetic energy balance equation

We repeat the process that we followed for the implicit scheme, to prove the stability of the scheme and derive a discrete kinetic energy balance equation. To this purpose, we multiply the velocity prediction equation by the corresponding degree of freedom of the predicted velocity  $\tilde{\mathbf{u}}_{\sigma,i}^{n+1}$ , to obtain:

$$\begin{aligned} \frac{|D_\sigma|}{\delta t} (\rho_\sigma^n \tilde{\mathbf{u}}_{\sigma,i}^{n+1} - \rho_\sigma^{n-1} \mathbf{u}_{\sigma,i}^n) \tilde{\mathbf{u}}_{\sigma,i}^{n+1} + \sum_{\varepsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon}^n \tilde{\mathbf{u}}_{\varepsilon,i}^{n+1} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} \\ + |D_\sigma| (\nabla \tilde{p}^{n+1})_{\sigma,i} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} - |D_\sigma| (\operatorname{div} \tau(\mathbf{u}^{n+1}))_{\sigma,i} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} = 0. \end{aligned} \quad (20)$$

We then write the velocity correction equation as:

$$\left[ \frac{|D_\sigma|}{\delta t} \rho_\sigma^n \right]^{1/2} \mathbf{u}_{\sigma,i}^{n+1} + \frac{[|D_\sigma| \delta t]^{1/2}}{(\rho_\sigma^n)^{1/2}} (\nabla p^{n+1})_{\sigma,i} = \left[ \frac{|D_\sigma|}{\delta t} (\rho_\sigma^n) \right]^{1/2} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} + \frac{[|D_\sigma| \delta t]^{1/2}}{(\rho_\sigma^n)^{1/2}} (\nabla \tilde{p}^{n+1})_{\sigma,i},$$

and square this relation, sum with (20) and get, applying Lemma A.2 (again on the dual mesh) to the first two terms of (20):

$$\begin{aligned} & \frac{1}{2} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^{n-1} (\mathbf{u}_{\sigma,i}^n)^2 \right] + \frac{1}{2} \sum_{\varepsilon=D_\sigma|D_{\sigma'}} F_{\sigma,\varepsilon}^n \tilde{\mathbf{u}}_{\sigma,i}^{n+1} \tilde{\mathbf{u}}_{\sigma',i}^{n+1} + |D_\sigma| (\nabla p^{n+1})_{\sigma,i} \mathbf{u}_{\sigma,i}^{n+1} \\ & - |D_\sigma| (\operatorname{div} \tau(\mathbf{u}^{n+1}))_{\sigma,i} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} + \frac{|D_\sigma| \delta t}{\rho_\sigma^n} \left[ |(\nabla p^{n+1})_{\sigma,i}|^2 - |(\nabla \tilde{p}^{n+1})_{\sigma,i}|^2 \right] = R_{\sigma,i}^{n+1}, \quad (21) \end{aligned}$$

where  $R_{\sigma,i}^{n+1}$  takes the same expression as in the implicit case (*i.e.* is given by Equation (10)), replacing  $\mathbf{u}^{n+1}$  by  $\tilde{\mathbf{u}}^{n+1}$ . Summing Relation (21) over the components and edges, Relation (12) over the cells and finally the two resulting equations together yields:

$$\begin{aligned} & \frac{1}{2} \sum_{i,\varepsilon} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^{n+1} (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^n (\mathbf{u}_{\sigma,i}^n)^2 \right] + \sum_{K \in \mathcal{M}} \frac{|K|}{\delta t} (\mathcal{H}(\rho_K^{n+1}) - \mathcal{H}(\rho_K^n)) \\ & + \sum_{i,\varepsilon} \frac{|D_\sigma| \delta t}{\rho_\sigma^n} \left[ |(\nabla p^{n+1})_{\sigma,i}|^2 - |(\nabla \tilde{p}^{n+1})_{\sigma,i}|^2 \right] \leq 0, \end{aligned}$$

which would be a discrete analogue to (6), up to a detail: to obtain a difference of the same quantity taken at two consecutive time steps, we need to change  $\rho_\sigma^n |(\nabla \tilde{p}^{n+1})_{\sigma,i}|^2$  to  $\rho_\sigma^{n-1} |(\nabla p^n)_{\sigma,i}|^2$ . This is the purpose of the pressure renormalization step, which was already introduced in [19]; we finally get:

$$\begin{aligned} & \frac{1}{2} \sum_{i,\varepsilon} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^{n+1} (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^n (\mathbf{u}_{\sigma,i}^n)^2 \right] + \sum_{K \in \mathcal{M}} \frac{|K|}{\delta t} (\mathcal{H}(\rho_K^{n+1}) - \mathcal{H}(\rho_K^n)) \\ & + \sum_{i,\varepsilon} |D_\sigma| \delta t \left[ \frac{1}{\rho_\sigma^n} |(\nabla p^{n+1})_{\sigma,i}|^2 - \frac{1}{\rho_\sigma^{n-1}} |(\nabla p^n)_{\sigma,i}|^2 \right] \leq 0. \end{aligned}$$

Note that this inequality yields a control on ( $\delta t$  times) a  $H^1$  discrete semi-norm of the pressure, confornting the robustness of the scheme, but also increasing its dissipation. In our numerical experiments, the pressure renormalization step did not appear to have a significant influence on the results, and was then systematically omitted.

### 2.2.3. Passing to the limit in the scheme (1D case)

We obtain for the pressure correction scheme results which are similar to the implicit scheme ones. They are stated in the following theorem.

**Theorem 2.3.** *Let  $\Omega$  be an open bounded interval of  $\mathbb{R}$ . Let  $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m \in \mathbb{N}}$  be a sequence of meshes and time steps, such that  $h^{(m)}$  and  $\delta t^{(m)}$  tend to zero as  $m$  tends to infinity. Let  $(\rho^{(m)}, p^{(m)}, u^{(m)}, \tilde{u}^{(m)})_{m \in \mathbb{N}}$  be the corresponding sequence of solutions. We suppose that this sequence satisfies the control over the time and space estimates given by (14), (15) and:*

$$\|\tilde{u}^{(m)}\|_{T,x,BV} \leq C, \quad \forall m \in \mathbb{N}.$$

*We assume in addition that it converges in  $L^r((0, T) \times \Omega)^4$ , for  $1 \leq r < \infty$ , to  $(\bar{\rho}, \bar{p}, \bar{u}, \bar{\tilde{u}}) \in L^\infty((0, T) \times \Omega)^4$ . Then we have  $\bar{\tilde{u}} = \bar{u}$ , and the triplet  $(\bar{\rho}, \bar{p}, \bar{u})$  satisfies the system (16) and the entropy condition (5).*

### 2.2.4. Numerical experiments

We now describe the behaviour of the pressure correction scheme for a Riemann problem, *i.e.* an inviscid monodimensional problem, the initial condition of which consists in two uniform left (L) and right (R) states,

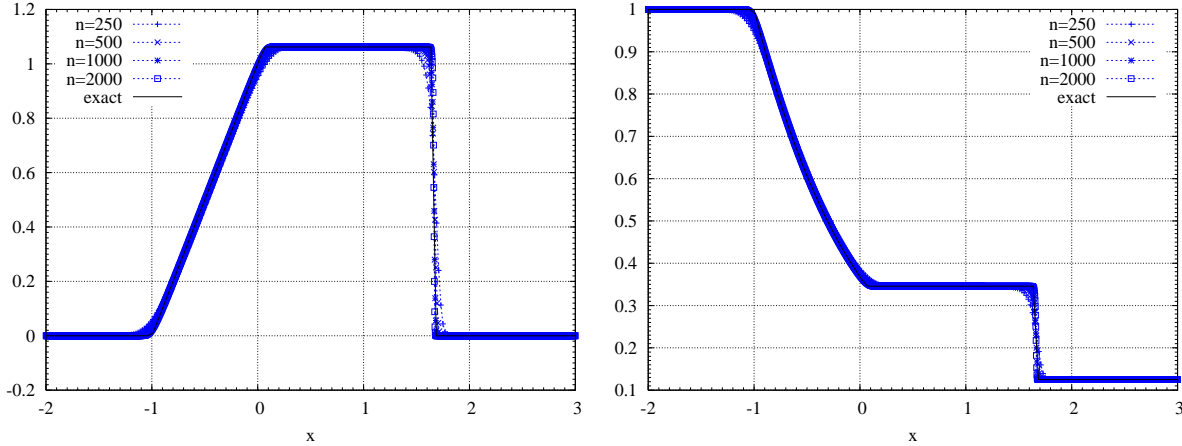


FIGURE 2. Sod shock tube problem – Centred scheme – Exact solution and numerical solution of the problem at  $t = 1$  with  $CFL=1$ . Velocity (left) and pressure (right).

separated by a discontinuity, located by convention at the origin  $x = 0$ . The two initial constant states are given by:

$$\begin{pmatrix} \rho \\ \mathbf{u} \end{pmatrix}_L = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} \rho \\ \mathbf{u} \end{pmatrix}_R = \begin{pmatrix} 0.125 \\ 0 \end{pmatrix},$$

and the equation of state is given by  $p = \rho$ . The problem is posed over the interval  $(-2, 3)$ . The solution of this problem consists in a rarefaction wave travelling to the left and a shock travelling to the right.

The problem is solved with a one dimensional scheme, which may be obtained from the previous exposition by taking one horizontal stripe of cells (of constant size) with the MAC discretization, and applying perfect slip boundary conditions at the top and bottom boundary.

On Figure 2, we show the solution at  $t = 1$  obtained with various meshes and time steps. These latter parameters are adjusted to have  $CFL = 1$ , taking as reference velocity the sum of the maximum velocity  $v = 1$  and the speed of sound  $a = 1$ . In these computations, we use a centred discretization of the convection term in the momentum balance equation, surprisingly without observing any spurious oscillations. However, note that results obtained with the CR and RT discretizations (not shown here) differ in this respect: the introduction of a residual viscosity (either physical or by upwinding) is necessary to avoid the odd-even decoupling phenomenon, as usually observed with centred approximations of the convection operator.

We then report, on Figure 3, the obtained numerical error as a function of the time and space step. The observed order of convergence is close to 0.9, for both the velocity and the pressure.

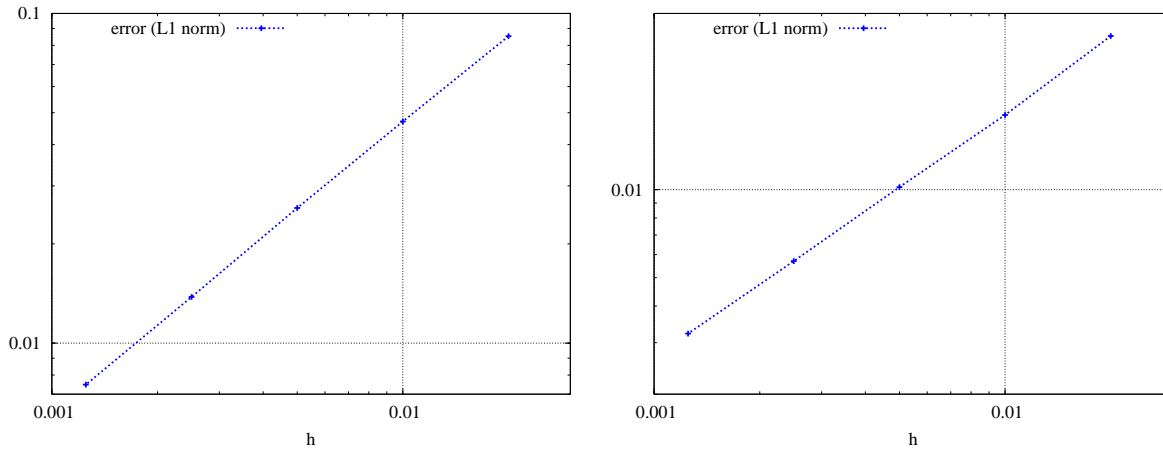


FIGURE 3. Sod shock tube problem – Centred scheme –  $L^1$  norm of the error between the numerical solution and the exact solution at  $t = 1$ , as a function of the mesh (or time) step, for CFL=1. Velocity (left) and pressure (right).

### 3. COMPRESSIBLE NAVIER-STOKES EQUATIONS

We now address the compressible Navier-Stokes equations (22).

$$\partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0, \quad (22a)$$

$$\partial_t(\rho \mathbf{u}) + \operatorname{div}(\rho \mathbf{u} \otimes \mathbf{u}) + \nabla p - \operatorname{div}(\boldsymbol{\tau}(\mathbf{u})) = 0, \quad (22b)$$

$$\partial_t(\rho E) + \operatorname{div}(\rho E \mathbf{u}) + \operatorname{div}(p \mathbf{u}) + \operatorname{div}(\mathbf{q}) = \operatorname{div}(\boldsymbol{\tau}(\mathbf{u}) \cdot \mathbf{u}), \quad (22c)$$

$$\rho = \varrho(p, e), \quad E = \frac{1}{2} |\mathbf{u}|^2 + e, \quad (22d)$$

where  $E$  and  $e$  are the total energy and internal energy in the flow, and  $\mathbf{q}$  stands for the heat conduction flux, assumed to be given by:

$$\mathbf{q} = -\lambda \nabla e,$$

with  $\lambda \geq 0$ . We suppose that the equation of state may be set under the form  $p = \wp(\rho, e)$  with  $\wp(\cdot, 0) = 0$  and  $\wp(0, \cdot) = 0$ . This system must be complemented by suitable boundary conditions and initial conditions for  $u$ ,  $\rho$  and  $e$ , which we suppose positive for the two latter unknowns.

Let us suppose that the solution is regular. Subtracting the kinetic energy balance equation from the total energy balance, we obtain the internal energy balance equation:

$$\partial_t(\rho e) + \operatorname{div}(\rho e \mathbf{u}) + p \operatorname{div}(\mathbf{u}) - \operatorname{div}(\lambda \nabla e) = \boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u}. \quad (23)$$

Since,

- (i) the viscous dissipation term  $\boldsymbol{\tau}(\mathbf{u}) : \nabla \mathbf{u}$  is non-negative,
- (ii) thanks to the mass balance equation, the first two terms may be recast as a transport operator:  $\partial_t(\rho e) + \operatorname{div}(\rho e \mathbf{u}) = \rho [\partial_t e + \mathbf{u} \cdot \nabla e]$ ,
- (iii) and, finally, because, from the assumption on the equation of state, the pressure vanishes when  $e = 0$ , this equation implies that  $e$  remains non-negative at all times.

In the framework of incompressible or low Mach number flows, the natural energy balance equation is the internal energy equation (23), so discretizing (23) instead of the total energy balance (22c) is a reasonable choice in order to obtain an algorithm which is valid for all flow regimes. In addition, it presents two advantages:

- first, it avoids the space discretization of the total energy, which is rather unnatural for staggered schemes since the velocity and the scalar variables are not collocated,
- second, a suitable discretization of (23) may yield, "by construction" of the scheme, the positivity of the internal energy.

However, integrating (22c) over  $\Omega$  yields a stability estimate for the solution, which reads, if we suppose for short that  $\mathbf{u}$  is prescribed to zero on the whole boundary  $\partial\Omega$ , and that the system is adiabatic, *i.e.*  $\nabla \mathbf{q} \cdot \mathbf{n} = 0$  on  $\partial\Omega$ :

$$\frac{d}{dt} \int_{\Omega} \left[ \frac{1}{2} \rho |\mathbf{u}|^2 + \rho e \right] d\mathbf{x} \leq 0, \quad (24)$$

and we would like (an analogue of) this stability estimate to hold at the discrete level.

In fact, the bridge between the discretization of (23) and this latter inequality is once again the kinetic energy balance equation, and the tools which were developed in the previous sections will readily yield the desired stability result if, at the discrete level, we are able:

- (i) to identify the integral of the dissipation term at the right-hand side of the discrete counterpart of (23) with what is obtained from the (discrete)  $L^2$  inner product between the velocity and the diffusion term in the discrete momentum balance equation (22b).

(ii) to prove that the right-hand side of (23) is non-negative in order to preserve the positivity of the internal energy.

Both properties are quite natural for finite element discretizations, but may be not so easy to obtain for the MAC scheme; for this latter case, a way to build an approximation of the viscous and dissipation terms to get this property is proposed in [30] ( see also [2]).

Two unconditionally stable schemes for the compressible Navier-Stokes equations are built, on the basis of these arguments [30]: the first one is implicit, and the second one, used in practice, is a pressure correction scheme. We only describe here this latter, which reads:

**Pressure renormalization step** – Solve for  $\tilde{p}^{n+1}$ :

$$\forall K \in \mathcal{M}, \quad \sum_{\sigma=K|L} \frac{1}{\rho_\sigma^n |D_\sigma|} (\tilde{p}_K^{n+1} - \tilde{p}_L^{n+1}) = \sum_{\sigma=K|L} \frac{1}{\sqrt{\rho_\sigma^n \rho_\sigma^{n-1}} |D_\sigma|} (p_K^n - p_L^n), \quad (25a)$$

**Prediction step** – Solve for  $\tilde{\mathbf{u}}^{n+1}$ :

$$\text{For } 1 \leq i \leq d, \quad \left| \begin{array}{l} \forall \sigma \in \mathcal{E}^{(i)} \text{ in the MAC case,} \\ \forall \sigma \in \mathcal{E} \text{ otherwise,} \end{array} \right.$$

$$\begin{aligned} \frac{|D_\sigma|}{\delta t} (\rho_\sigma^n \tilde{\mathbf{u}}_{\sigma,i}^{n+1} - \rho_\sigma^{n-1} \mathbf{u}_{\sigma,i}^n) + \sum_{\varepsilon \in \mathcal{E}(D_\sigma)} F_{\sigma,\varepsilon}^n \tilde{\mathbf{u}}_{\varepsilon,i}^{n+1} + |D_\sigma| (\nabla \tilde{p}^{n+1})_{\sigma,i} \\ - |D_\sigma| (\text{div} \tau(\tilde{\mathbf{u}}^{n+1}))_{\sigma,i} = 0, \end{aligned} \quad (25b)$$

**Correction step** – Solve for  $\rho^{n+1}$ ,  $p^{n+1}$ ,  $e^{n+1}$  and  $\mathbf{u}^{n+1}$ :

$$\text{For } 1 \leq i \leq d, \quad \left| \begin{array}{l} \forall \sigma \in \mathcal{E}^{(i)} \text{ in the MAC case,} \\ \forall \sigma \in \mathcal{E} \text{ otherwise,} \end{array} \right.$$

$$\frac{|D_\sigma|}{\delta t} \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1} - \tilde{\mathbf{u}}_{\sigma,i}^{n+1}) + |D_\sigma| [(\nabla p^{n+1})_{\sigma,i} - (\nabla \tilde{p}^{n+1})_{\sigma,i}] = 0, \quad (25c)$$

$$\forall K \in \mathcal{M}, \quad \frac{|K|}{\delta t} (\rho_K^{n+1} - \rho_K^n) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^{n+1} = 0, \quad (25d)$$

$$\forall K \in \mathcal{M},$$

$$\begin{aligned} \frac{|K|}{\delta t} (\rho_K^{n+1} e_K^{n+1} - \rho_K^n e_K^n) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^{n+1} e_\sigma^{n+1} + |K| (p^{n+1} \text{div} \tilde{\mathbf{u}}^{n+1})_K \\ - \text{div}(\lambda \nabla e)_K = |K| (\boldsymbol{\tau}(\tilde{\mathbf{u}}^{n+1}) : \nabla \tilde{\mathbf{u}}^{n+1})_K, \end{aligned} \quad (25e)$$

$$\forall K \in \mathcal{M}, \quad p_K^{n+1} = (\gamma - 1) \rho_K^{n+1} e_K^{n+1}. \quad (25f)$$

The construction of this scheme relies on the same ingredients as in the barotropic case, in particular the time shift of the densities.

The equation (25e) is a approximation of the internal balance over the primal mesh  $K$ , which ensures the positivity of the internal energy, thanks to two essential arguments:

- first, the approximation of the convection operator  $e \mapsto \partial_t(\rho e) + \text{div}(\rho e \mathbf{u})$  is upwind (*i.e.*  $e_\sigma^{n+1} = e_K^{n+1}$  if  $F_{K,\sigma}^{n+1} \geq 0$  and  $e_L^{n+1}$  otherwise) and this operator satisfies a consistency property with the mass balance  $\partial_t \rho + \text{div}(\rho \mathbf{u}) = 0$  which may be stated as the fact that it vanishes if  $e$  is constant.

This property is, of course, necessary for an operator to satisfy a discrete maximum principle (constants are necessarily solutions to an equation obeying a maximum principle. . . ); it is also classically shown [39] to be sufficient.

- second, the internal energy balance is coupled to the algorithm in such a way that the pressure in the discretization of the term  $p \text{div} \mathbf{u}$  obeys the equation of state, and thus, in particular vanishes when  $e < 0$  (see [46] for another pressure-correction algorithm using the same coupling).

The technique used to ensure the positivity of  $e$  is to define:

$$|K| (p^{n+1} \text{div} \tilde{\mathbf{u}}^{n+1})_K = \wp(\rho_K^{n+1}, (e_K^{n+1})^+) \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \tilde{\mathbf{u}}_\sigma^{n+1} \cdot \mathbf{n}_{K,\sigma}, \quad (26)$$

where  $(e_K^{n+1})^+$  stands for the positive part of  $e_K^{n+1}$ , *i.e.*  $(e_K^{n+1})^+ = \max(e_K^{n+1}, 0)$ . Testing then the internal energy balance by the negative part of  $e_K^{n+1}$ , designed by  $(e_K^{n+1})^- = -\min(e_K^{n+1}, 0)$ , and summing over  $K \in \mathcal{M}$ . Supposing, for short, that the normal velocity vanishes on the boundaries, Lemma A.2 yields:

$$\sum_{K \in \mathcal{M}} \left[ \frac{|K|}{\delta t} (\rho_K^{n+1} e_K^{n+1} - \rho_K^n e_K^n) + \sum_{\sigma \in \mathcal{E}(K)} F_{K,\sigma}^{n+1} e_\sigma^{n+1} \right] (e_K^{n+1})^- \geq \frac{-1}{\delta t} \sum_{K \in \mathcal{M}} \varrho_K^{n+1} [(e_K^{n+1})^-]^2 - \varrho_K^n [(e_K^n)^-]^2,$$

while,  $\forall K \in \mathcal{M}$ ,  $(p^{n+1} \text{div} \tilde{\mathbf{u}}^{n+1})_K (e_K^{n+1})^- = 0$  and the right-hand side is non-negative, which yields the result. A topological degree argument, applied to the algebraic system corresponding to the whole correction step, yields the existence of at least one solution and, since, for this solution,  $e \geq 0$ ,  $(e_K^{n+1})^+ = e_K^{n+1}$  and the discretization (26) is consistent.

The obtained stability result is stated in the following theorem.

**Theorem 3.1.** *There exists a solution to the scheme which satisfies  $\rho > 0$ ,  $e > 0$  and for all  $n \leq N$ , the following inequality holds:*

$$\begin{aligned} \sum_{K \in \mathcal{M}} |K| \rho_K^n e_K^n + \frac{1}{2} \sum_{\sigma \in \mathcal{E}_{\text{int}}} |D_\sigma| \rho_\sigma^{n-1} |u_\sigma^n|^2 + \frac{\delta t^2}{2} |p^n|_{\rho^{n-1}, \mathcal{M}}^2 \\ \leq \sum_{K \in \mathcal{M}} |K| \rho_K^0 e_K^0 + \frac{1}{2} \sum_{\sigma \in \mathcal{E}_{\text{int}}} |D_\sigma| \rho_\sigma^{-1} |u_\sigma^0|^2 + \frac{\delta t^2}{2} |p^0|_{\rho^{-1}, \mathcal{M}}^2, \end{aligned}$$

where, for any discrete pressure  $q$ :

$$|q|_{\rho, \mathcal{M}}^2 = \sum_{\sigma=K|L} \frac{1}{\rho_\sigma} \frac{|\sigma|^2}{|D_\sigma|} (p_L - p_K)^2.$$

#### 4. EULER EQUATIONS

For solutions with shocks, Equation (23) is not equivalent to (22c); more precisely speaking, one can show that, at a shock location, a positive measure should replace  $\boldsymbol{\tau}(\mathbf{u}) : \boldsymbol{\nabla} \mathbf{u}$  (which formally vanishes since  $\mu = 0$ ) at the right-hand side of Equation (23). Discretizing (23) instead of (22c) may thus yield a scheme which does not compute the correct weak discontinuous solutions, the manifestation of this non-consistency being that the numerical solutions present shocks which do not satisfy the Rankine-Hugoniot conditions associated to (22c). The essential result of this section is to provide solutions to circumvent this problem.

This study is closely related to the analysis performed in the barotropic case. Indeed, it may be checked that the entropy of the barotropic problem takes an expression similar to the total energy  $E$  (in fact, if the equation of state in the barotropic case is derived by supposing that the flow is isentropic, we have the exact equality  $\mathcal{H} = \rho e$ ); the elastic potential balance (in the barotropic case) plays the same role as the internal energy balance (in the non-barotropic case). The only difference is that the entropy condition is an inequality while the total energy is an equality: in other words, while, for the barotropic case, we just checked that residual terms were non-positive, we now have to ensure that they vanish with the discretization steps. To this purpose, we thus follow a strategy quite similar to Section 2:

- Starting from the discrete momentum balance equation, with an *ad hoc* discretization of the convection operator, we derive a discrete kinetic energy balance; residual terms are present in this relation, which do not tend to zero with space and time steps (they are the discrete manifestations of the above mentioned measures).
- These residual terms are then compensated by source terms added to the internal energy balance.

We provide a theoretical justification of this process by showing that, in the 1D case, if the scheme is stable enough and converges to a limit (in a sense to be defined), this limit satisfies a weak form of (22c) which implies the correct Rankine-Hugoniot conditions. Then, we perform numerical tests which substantiate this analysis. Two different time discretizations are proposed in [29]: first, a fully implicit scheme (a solution to which may be rather difficult to obtain in practice) and, second, a pressure correction scheme (the algorithm indeed used in the tests presented here); we only present here the latter algorithm.

#### 4.1. The discrete kinetic energy balance equation and the corrective source terms

We derive here a slightly different discrete kinetic energy balance than in Section 2.2.2. Our starting point, however, is still the velocity prediction step which we multiply by the corresponding unknown, *i.e.* Equation (20), which now reads, since, in the present algorithm, we omit the pressure renormalization step:

$$\frac{|D_\sigma|}{\delta t} (\rho_\sigma^n \tilde{\mathbf{u}}_{\sigma,i}^{n+1} - \rho_\sigma^{n-1} \mathbf{u}_{\sigma,i}^n) \tilde{\mathbf{u}}_{\sigma,i}^{n+1} + \sum_{\varepsilon \in \tilde{\mathcal{E}}(D_\sigma)} F_{\sigma,\varepsilon}^n \tilde{\mathbf{u}}_{\varepsilon,i}^{n+1} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} + |D_\sigma| (\nabla p^n)_{\sigma,i} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} = 0.$$

The next step is to multiply the velocity correction equation by  $\tilde{\mathbf{u}}_{\sigma,i}^{n+1}$  and use the identity  $2a(a-b) = a^2 + (a-b)^2 - b^2$  to get:

$$\frac{1}{2} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^n (\tilde{\mathbf{u}}_{\sigma,i}^n)^2 \right] + |D_\sigma| \left[ (\nabla p^{n+1})_{\sigma,i} - (\nabla p^n)_{\sigma,i} \right] \tilde{\mathbf{u}}_{\sigma,i}^{n+1} + \frac{|D_\sigma|}{2\delta t} \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1} - \tilde{\mathbf{u}}_{\sigma,i}^{n+1})^2 = 0.$$

Invoking Lemma A.2 for the first two terms of the first of these relations and summing with the second one yields:

$$\frac{1}{2} \frac{|D_\sigma|}{\delta t} \left[ \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1})^2 - \rho_\sigma^{n-1} (\mathbf{u}_{\sigma,i}^n)^2 \right] + \frac{1}{2} \sum_{\varepsilon = D_\sigma | D_{\sigma'}} F_{\sigma,\varepsilon}^n \tilde{\mathbf{u}}_{\sigma,i}^{n+1} \tilde{\mathbf{u}}_{\sigma',i}^{n+1} + |D_\sigma| (\nabla p^{n+1})_{\sigma,i} \tilde{\mathbf{u}}_{\sigma,i}^{n+1} = R_{\sigma,i}^{n+1}, \quad (27)$$

with:

$$R_{\sigma,i}^{n+1} = \frac{|D_\sigma|}{2\delta t} \rho_\sigma^n (\mathbf{u}_{\sigma,i}^{n+1} - \tilde{\mathbf{u}}_{\sigma,i}^{n+1})^2 - \frac{|D_\sigma|}{2\delta t} \rho_\sigma^{n-1} (\tilde{\mathbf{u}}_{\sigma,i}^{n+1} - \mathbf{u}_{\sigma,i}^n)^2 - \delta^{\text{up}} \left[ \sum_{\varepsilon = D_\sigma | D_{\sigma'}} \frac{1}{2} |F_{\sigma,\varepsilon}^n| (\tilde{\mathbf{u}}_{\sigma,i}^{n+1} - \tilde{\mathbf{u}}_{\sigma',i}^{n+1}) \right] \tilde{\mathbf{u}}_{\sigma,i}^{n+1}.$$

We recognize at the left-hand side a conservative discrete kinetic energy balance. The next step is now to deal with the residual terms at the right-hand side, or, more precisely speaking, to somewhat compensate them

by some source term which we introduced in the internal energy balance. Let us denote by  $S_K^{n+1}$  the source term in the balance over the cell  $K$ . We choose:

$$\begin{aligned} \forall K \in \mathcal{M}, \\ S_K^{n+1} = \frac{1}{2} \sum_{\sigma \in \mathcal{E}(K)} \frac{|D_{K,\sigma}|}{\delta t} \rho_K^{n-1} (\tilde{\mathbf{u}}_\sigma^{n+1} - \mathbf{u}_\sigma^n)^2 - \frac{1}{2} \sum_{\sigma \in \mathcal{E}(K)} \frac{|D_{K,\sigma}|}{\delta t} \rho_K^n (\mathbf{u}_\sigma^{n+1} - \tilde{\mathbf{u}}_\sigma^{n+1})^2 \\ + \delta^{\text{up}} \sum_{\substack{\varepsilon \cap \bar{K} \neq \emptyset, \\ \varepsilon = D_\sigma | D_{\sigma'} }} \alpha_{K,\varepsilon} \frac{|F_{\sigma,\varepsilon}^{n+1}|}{2} (\mathbf{u}_\sigma^{n+1} - \mathbf{u}_{\sigma'}^{n+1})^2. \end{aligned} \quad (28)$$

The coefficient  $\alpha_{K,\varepsilon}$  is fixed to 1 if the face  $\varepsilon$  is included in  $K$ , and this is the only situation to consider for the RT and CR discretization. For the MAC scheme, some dual edges are included in the primal cells, whereas some lie on their boundary; for  $\varepsilon$  on a cell boundary, we denote by  $\mathcal{N}_\varepsilon$  the set of cells  $M$  such that  $\bar{M} \cap \varepsilon \neq \emptyset$  (the cardinal of this set is always 4), and compute  $\alpha_{K,\varepsilon}$  by:

$$\alpha_{K,\varepsilon} = \frac{|K|}{\sum_{M \in \mathcal{N}_\varepsilon} |M|}.$$

For a uniform grid, this formula yields  $\alpha_{K,\varepsilon} = 1/4$ .

The expression of the terms  $(S_K)_{K \in \mathcal{M}}$  is justified by the passage to the limit in the scheme (for a one-dimensional problem) performed in Section 4.2. Let us just here remark that:

$$\sum_{K \in \mathcal{M}} S_K^{n+1} + \sum_{\mathcal{E}, i} R_{\sigma,i}^{n+1} = 0,$$

which shows that the introduction of this term allows to recover the total energy balance over the whole computational domain  $\Omega$ . Note however that, the term  $S_K^{n+1}$  may be negative, which we have indeed observed in computations, and so the above proof of the positivity of the internal energy is not valid here; however, even in very severe cases (as, for instance, Test 3 of [55, chapter 4]), at least with a reasonable time step, we still obtained  $e > 0$ .

**Remark 4.1** (Form of the corrective source terms). Comparing with the source term of the continuous internal energy balance (23), it is easy to identify in the last part of  $S_K$  the viscous dissipation associated to the numerical diffusion introduced by the upwinding. In fact, this analogy also holds for the first two terms: they are dissipation and antidissipation terms associated to the diffusion and antidiffusion introduced by the semi-implicit time discretization.

Note by the way that only a dissipation term is obtained for the implicit case (*i.e.* the corrective terms  $S_K^{n+1}$  are non-negative, see Chapter 4), and thus, for this time discretization, the positivity of the internal energy is ensured.

**Remark 4.2** (On the necessity of the corrective source terms). Let us consider a sequence of discretizations  $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m \in \mathbb{N}}$ , the space and time steps of which tend to zero, an associated sequence of discrete velocities  $(\mathbf{u}^{(m)})_{m \in \mathbb{N}}$ , and the corresponding sequence of (piecewise constant functions associated to the) corrective term  $(S^{(m)})_{m \in \mathbb{N}}$ . It may be checked that  $S^{(m)}$  tends to zero in  $L^1(\Omega \times (0, T))$  as soon as the time and space derivatives of the functions  $(\mathbf{u}^{(m)})_{m \in \mathbb{N}}$  are bounded in a strong enough norm, and in particular stronger than the BV norm (for instance, suppose that the jumps between two consecutive time steps and adjacent cells are bounded by  $\delta t$  and  $h$  respectively), *i.e.* everywhere the solution is regular. On the opposite, for a sequence  $(\mathbf{u}^{(m)})_{m \in \mathbb{N}}$  obtained by projecting a discontinuous function  $\mathbf{u}$ ,  $S^{(m)}$  does not tend to zero.

## 4.2. Passing to the limit in the scheme

As for the barotropic equations, we now pass to the limit in the scheme.

We suppose given a sequence of meshes and time steps  $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m \in \mathbb{N}}$ , such that the time step and the size  $h^{(m)}$  of the mesh  $\mathcal{M}^{(m)}$ , defined by:

$$h^{(m)} = \sup_{K \in \mathcal{M}^{(m)}} \text{diam}(K),$$

tend to zero as  $m \rightarrow \infty$ .

Let  $\rho^{(m)}$ ,  $p^{(m)}$ ,  $e^{(m)}$ ,  $\tilde{u}^{(m)}$  and  $u^{(m)}$  be the associated solution of the pressure correction scheme (25) with the mesh  $\mathcal{M}^{(m)}$  and the time step  $\delta t^{(m)}$  (or, more precisely speaking, as in the barotropic case, a 1D version of the scheme). To the discrete unknowns, we associate piecewise constant functions on time intervals and on primal or dual meshes:

$$\begin{aligned} \rho^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{K \in \mathcal{M}} (\rho^{(m)})_K^n \mathcal{X}_K \mathcal{X}_{(n, n+1)}, & p^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{K \in \mathcal{M}} (p^{(m)})_K^n \mathcal{X}_K \mathcal{X}_{(n, n+1)}, \\ e^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{K \in \mathcal{M}} (e^{(m)})_K^n \mathcal{X}_K \mathcal{X}_{(n, n+1)}, & u^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{\sigma \in \mathcal{E}} (u^{(m)})_\sigma^n \mathcal{X}_{D_\sigma} \mathcal{X}_{(n, n+1)}, \\ \tilde{u}^{(m)}(x, t) &= \sum_{n=0}^{N-1} \sum_{\sigma \in \mathcal{E}} (\tilde{u}^{(m)})_\sigma^n \mathcal{X}_{D_\sigma} \mathcal{X}_{(n, n+1)}. \end{aligned}$$

We suppose that the sequence of discrete solutions  $(\rho^{(m)}, p^{(m)}, e^{(m)}, u^{(m)}, \tilde{u}^{(m)})_{m \in \mathbb{N}}$  is uniformly bounded in  $L^\infty((0, T) \times \Omega)$ , i.e.:

$$|(\rho^{(m)})_K^n| + |(p^{(m)})_K^n| + |(e^{(m)})_K^n| \leq C, \quad \forall K \in \mathcal{M}^{(m)}, \text{ for } 0 \leq n \leq N^{(m)}, \forall m \in \mathbb{N}, \quad (29)$$

and:

$$|(u^{(m)})_\sigma^n| + |(\tilde{u}^{(m)})_\sigma^n| \leq C, \quad \forall \sigma \in \mathcal{E}^{(m)}, \text{ for } 0 \leq n \leq N^{(m)}, \forall m \in \mathbb{N}. \quad (30)$$

In addition, we also suppose the following uniform control on the translates in space and time:

$$\|\rho^{(m)}\|_{T, x, BV} + \|e^{(m)}\|_{T, x, BV} + \|u^{(m)}\|_{T, x, BV} + \|\tilde{u}^{(m)}\|_{T, x, BV} \leq C, \quad \forall m \in \mathbb{N}, \quad (31)$$

and:

$$\|\rho^{(m)}\|_{T, t, BV} + \|u^{(m)}\|_{T, t, BV} \leq C, \quad \forall m \in \mathbb{N}. \quad (32)$$

As in the barotropic case, we are not able to prove the estimates (29)–(32) for the solutions of the scheme, but such inequalities are satisfied by the "interpolation" of the solution to a Riemann problem, and are observed, with a limited sequence of meshes and time steps, in computations.

A weak solution to the continuous problem satisfies, for any  $\varphi \in C_c^\infty([0, T] \times \Omega)$ :

$$-\int_{\Omega \times (0, T)} [\rho \partial_t \varphi + \rho u \partial_x \varphi] dx dt - \int_{\Omega} \rho(x, 0) \varphi(x, 0) dx = 0, \quad (33a)$$

$$-\int_{\Omega \times (0, T)} [\rho u \partial_t \varphi + (\rho u^2 + p) \partial_x \varphi] dx dt - \int_{\Omega} \rho(x, 0) u(x, 0) \varphi(x, 0) dx = 0, \quad (33b)$$

$$-\int_{\Omega \times (0, T)} [\rho E \partial_t \varphi + (\rho E + p) u \partial_x \varphi] dx dt - \int_{\Omega} \rho(x, 0) E(x, 0) \varphi(x, 0) dx = 0, \quad (33c)$$

$$\rho = \varrho(p, e), \quad E = \frac{1}{2} u^2 + e. \quad (33d)$$

Once again, since the test function  $\varphi$  vanishes at the boundary, these relations do not imply anything about the boundary conditions, but imply the Rankine-Hugoniot conditions. The scheme consistency result that we are seeking for is stated in the following theorem.

**Theorem 4.3.** *Let  $\Omega$  be an open bounded interval of  $\mathbb{R}$ . Let  $(\mathcal{M}^{(m)}, \delta t^{(m)})_{m \in \mathbb{N}}$  be a sequence of meshes and time steps, such that  $h^{(m)}$  and  $\delta t^{(m)}$  tend to zero as  $m$  tends to infinity. Let  $(\rho^{(m)}, p^{(m)}, e^{(m)}, u^{(m)}, \tilde{u}^{(m)})_{m \in \mathbb{N}}$  be the corresponding sequence of solutions. We suppose that this sequence satisfies (29)–(32) and converges in  $L^r((0, T) \times \Omega)^5$ , for  $1 \leq r < \infty$ , to a limit  $(\bar{\rho}, \bar{p}, \bar{e}, \bar{u}, \bar{\tilde{u}}) \in L^\infty((0, T) \times \Omega)^5$ .*

*Then  $\bar{\tilde{u}} = \bar{u}$  and the limit  $(\bar{\rho}, \bar{p}, \bar{e}, \bar{u})$  satisfies the system (33).*

### 4.3. Numerical tests

We now assess the behaviour of the scheme on a one dimensional Riemann problem. We choose initial conditions such that the structure of the solution consists in two shock waves, separated by the contact discontinuity, with sufficiently strong shocks to allow to easily discriminate between convergence to the correct weak solution or not. These initial conditions are those proposed in [55, chapter 4], for the test referred to as Test 5:

$$\text{left state: } \begin{bmatrix} \rho_L \\ u_L \\ p_L \end{bmatrix} = \begin{bmatrix} 5.99924 \\ 19.5975 \\ 460.894 \end{bmatrix} \quad \text{right state: } \begin{bmatrix} \rho_R \\ u_R \\ p_R \end{bmatrix} = \begin{bmatrix} 5.99242 \\ -6.19633 \\ 46.0950 \end{bmatrix}$$

The problem is posed over  $\Omega = (-0.5, 0.5)$ , and the discontinuity is initially located at  $x = 0$ .

Since numerical experiments addressing barotropic flows (see Section 2.2.4) show that, at least for one dimensional computations, it is not necessary to use upwinding in the momentum balance equation, we only employ a centered approximation of the velocity at the dual faces.

The density fields obtained with  $h = 1/2000$  (or a number of cells  $n = 2000$ ) at  $t = 0.035$ , with and without assembling the corrective source term in the internal energy balance, are shown together with the analytical solution on Figure 4. The density and the pressure obtained, still with and without corrective terms, for various meshes, are plotted on Figure 5 and 6 respectively. For these computations, we take  $\delta t = h/20$ , which yields a CFL number, with respect to the material velocity only, close to one. The first conclusion is that both schemes seem to converge, but the corrective term is necessary to obtain the correct solution. In this case, for instance, we obtain the correct intermediate state for the pressure and velocity up to four digits in the essential part of the corresponding zone:

$$\begin{array}{l} \text{(analytical) intermediate state: } \begin{bmatrix} p^* \\ u^* \end{bmatrix} = \begin{bmatrix} 1691.65 \\ 8.68977 \end{bmatrix} \text{ for } x \in (0.028, 0.428) \\ \\ \text{numerical results: } \begin{array}{l} \left| \begin{array}{l} p \in (1691.6, 1691.8) \\ u \in (8.689, 8.690) \end{array} \right. \text{ for } x \in (0.032, 0.417) \end{array} \end{array}$$

One can check that the solution obtained without the corrective term is not a weak solution to the Euler system.

We also observe that the scheme is rather diffusive, specially at the contact discontinuity, where the beneficial compressive effect of the shocks does not apply.

## 5. CONCLUSION AND PERSPECTIVES

We developed a class of schemes for barotropic and non-barotropic flows, based on staggered space discretizations and on a fractional time-stepping technique falling in the class of pressure correction methods. Upwinding is performed in an equation-by-equation way, and only with respect to the material velocity; for non-barotropic

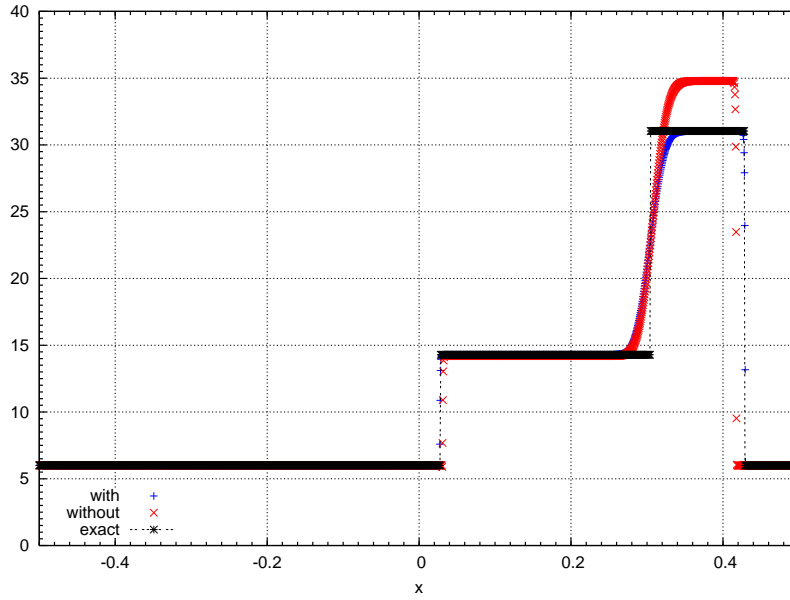


FIGURE 4. Test 5 of [55, chapter 4] - Density obtained with  $n = 2000$  cells, with and without corrective source terms, and analytical solution.

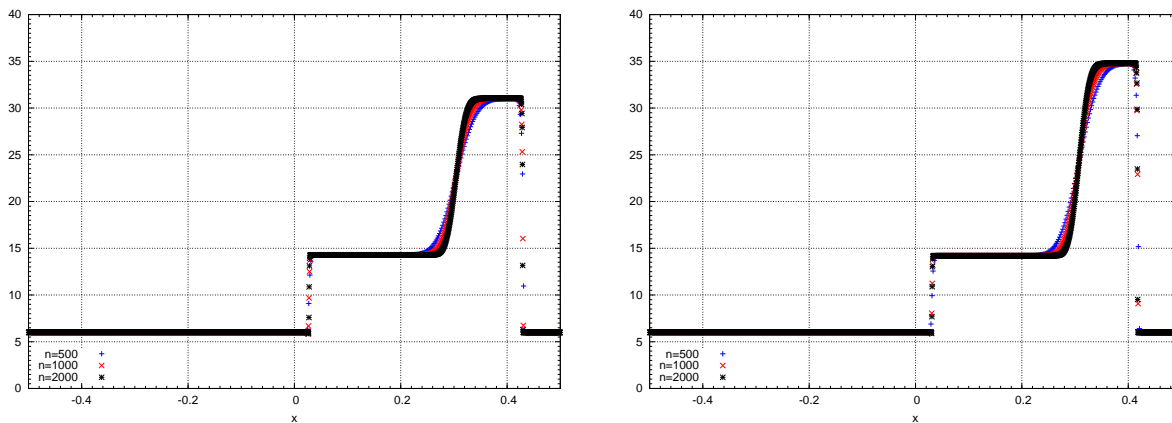


FIGURE 5. Test 5 of [55, chapter 4] - Density obtained with various meshes, with (left) and without (right) corrective source terms.

equations, the energy equation is the internal energy balance. All of these characteristics ensure that the schemes boil down to usual incompressible flow solvers for a vanishing Mach number; therefore they are hoped to be stable and accurate in the whole incompressible to compressible range. Numerical tests performed here focus on compressible flows, and assess the fact that weak solution to inviscid problems are correctly computed; they are supported by theoretical arguments. These tests will be continued, addressing complex multi-dimensional geometries.

From an algorithmic point of view, let us first mention that, for high Mach number flows, explicit versions of these schemes are now under development [47]; this would provide efficient algorithms (in particular, with

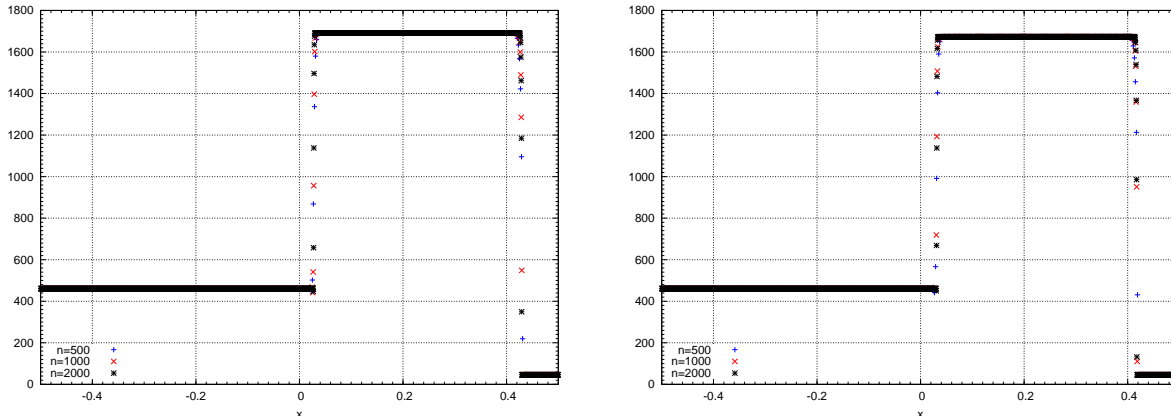


FIGURE 6. Test 5 of [55, chapter 4] - pressure obtained with various meshes, with (left) and without (right) corrective source terms.

an immediate construction of the fluxes at the cell faces), well suited to fast transient regimes, and offering, if necessary, the possibility of a partial implicitation without loosing any stability features (by the schemes studied in this work). In explicit schemes, less diffusive space discretizations, such as MUSCL-like or adaptative numerical viscosity [21, 22] techniques, are easy to implement; this will be done in a near future.

A lot of theoretical questions are suggested by the present study. First, the passage to the limit in the schemes in the multi-dimensional case raises only technical problems, which should not be so difficult to fix. A more intricate question is that of the boundary conditions (see *eg.* some numerical experiments described in [37]): the decoupling of pressure correction schemes is known to produce inherent spurious boundary conditions, the effect of which is extensively discussed for incompressible flows; for compressible problems, this question seems to remain largely open, and should deserve to be studied in the future. We did not prove in this work that the solutions obtained for non-barotropic Euler equations, if they converge, tend to the entropy weak solution; this is another issue to be addressed in the near future. Yet another topic is to analyse more indepth the behaviour of the proposed schemes in the low Mach number regime. In particular, since these algorithms satisfy stability estimates, it seems possible, at least with a fixed mesh and using the norm equivalence property of finite dimensional problems, to pass to the limit for a vanishing Mach number. This is interesting both theoretically and from an engineering point of view, to get some insight in what physical model is indeed solved by the code in such situations. Last but not least, we performed here some parts of the convergence analysis; this should be continued as far as possible, in particular for barotropic viscous flows.

## A. SOME RESULTS ASSOCIATED TO FINITE VOLUME CONVECTION OPERATORS

We gather in this section some results concerning the discretization by the finite volume method of two convection operators:

- the first one reads, at the continuous level,  $\rho \rightarrow \mathcal{C}(\rho) = \partial_t \rho + \text{div}(\rho \mathbf{u})$ , where  $\mathbf{u}$  stands for a given velocity field, which is not assumed to satisfy any divergence constraint,
- the second one is  $z \rightarrow \mathcal{C}_\rho(z) = \partial_t(\rho z) + \text{div}(\rho z \mathbf{u})$ , where  $\rho$  and  $\mathbf{u}$  stands for two given scalar and vector fields, which are supposed to satisfy  $\partial_t \rho + \text{div}(\rho \mathbf{u}) = 0$ .

Multiplying these operators by functions depending on the unknown is currently used to obtain convection operators acting over different variables, possibly with residual terms: one may think, for instance, to the theory of renormalized solutions (for the first one), or, in mechanics, to the derivation of the so-called kinetic

energy transport identity (for the second one). The results provided in this section are discrete variants of such relations.

We begin with a property of  $\mathcal{C}$ , which, at the continuous level, may be formally obtained as follows. Let  $\psi$  be a regular function from  $(0, +\infty)$  to  $\mathbb{R}$ ; then:

$$\psi'(\rho) \mathcal{C}(\rho) = \psi'(\rho) \partial_t(\rho) + \psi'(\rho) \mathbf{u} \cdot \nabla \rho + \psi'(\rho) \rho \operatorname{div} \mathbf{u} = \partial_t(\psi(\rho)) + \mathbf{u} \cdot \nabla \psi(\rho) + \rho \psi'(\rho) \operatorname{div} \mathbf{u},$$

so adding and subtracting  $\psi(\rho) \operatorname{div} \mathbf{u}$  yields:

$$\psi'(\rho) \mathcal{C}(\rho) = \partial_t(\psi(\rho)) + \operatorname{div}(\psi(\rho) \mathbf{u}) + (\rho \psi'(\rho) - \psi(\rho)) \operatorname{div} \mathbf{u}. \quad (34)$$

Obtaining a proof of this last identity, in a weak sense and with minimal regularity assumptions for  $\rho$  and  $\mathbf{u}$  and increasing properties of  $\psi$  is the object of the theory of renormalized solutions. The following lemma states a discrete analogue to (34).

**Lemma A.1.** *Let  $K \in \mathcal{M}$ . Let  $\rho_K^*$  and  $\rho_K$  be two positive real numbers. For  $\sigma \in \mathcal{E}(K)$ , let  $F_\sigma$  be a quantity associated to the face  $\sigma$  and the control volume  $K$ , defined by*

$$\forall \sigma \in \mathcal{E}(K), \quad F_\sigma = \rho_\sigma V_\sigma.$$

where  $\rho_\sigma$  and  $V_\sigma$  are a positive real number and a real number respectively, both associated to the edge  $\sigma$ . Let  $\psi$  be a twice continuously differentiable function, defined over  $(0, +\infty)$ .

Then the following identity holds:

$$\begin{aligned} \left[ \frac{|K|}{\delta t} (\rho_K - \rho_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_\sigma \right] \psi'(\rho_K) &= \frac{|K|}{\delta t} [\psi(\rho_K) - \psi(\rho_K^*)] + \sum_{\sigma \in \mathcal{E}(K)} \psi(\rho_\sigma) V_\sigma \\ &+ \left[ \rho_K \psi'(\rho_K) - \psi(\rho_K) \right] \sum_{\sigma \in \mathcal{E}(K)} V_\sigma + R_{\sigma, \text{dt}} \end{aligned} \quad (35)$$

where

$$R_{\sigma, \text{dt}} = \frac{1}{2} \frac{|K|}{\delta t} \psi''(\bar{\rho}_K) (\rho_K - \rho_K^*)^2 - \frac{1}{2} \sum_{\sigma \in \mathcal{E}(K)} V_\sigma \psi''(\bar{\rho}_\sigma) (\rho_\sigma - \rho_K)^2, \quad (36)$$

and,  $\forall \sigma \in \mathcal{E}(K)$ ,  $\bar{\rho}_K \in [\min(\rho_K, \rho_K^*), \max(\rho_K, \rho_K^*)]$  and  $\bar{\rho}_\sigma \in [\min(\rho_\sigma, \rho_K), \max(\rho_\sigma, \rho_K)]$ . If we now suppose that the function  $\psi$  is once differentiable and convex, and that  $\rho_\sigma = \rho_K$  as soon as  $V_\sigma \geq 0$ , then expression (36) is no-more valid but the residual  $R_{\sigma, \text{dt}}$  is non-negative.

We now turn to the second operator, for which we have, at the continuous level and formally, using twice the assumption  $\partial_t \rho + \operatorname{div}(\rho \mathbf{u}) = 0$ :

$$\begin{aligned} \psi'(z) \mathcal{C}_\rho(z) &= \psi'(z) [\partial_t(\rho z) + \operatorname{div}(\rho z \mathbf{u})] = \psi'(z) \rho [\partial_t z + \mathbf{u} \cdot \nabla z] \\ &= \rho [\partial_t \psi(z) + \mathbf{u} \cdot \nabla \psi(z)] = \partial_t(\rho \psi(z)) + \operatorname{div}(\rho \psi(z) \mathbf{u}). \end{aligned}$$

Taking for  $z$  a component of the velocity field and  $\psi(s) = s^2/2$ , this relation is the central argument used to derive the kinetic energy balance. The following lemma states a discrete counterpart of this identity.

**Lemma A.2.** Let  $K \in \mathcal{M}$ . Let  $\rho_K^*$  and  $\rho_K$  be two positive real numbers. For  $\sigma \in \mathcal{E}(K)$ , let  $F_\sigma$  be a quantity associated to the face  $\sigma$ , such that the following identity holds:

$$\frac{|K|}{\delta t} (\rho_K - \rho_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_\sigma = 0. \quad (37)$$

Let  $u_K^*$  and  $u_K$  be two real numbers, and, to each  $\sigma \in \mathcal{E}(K)$ , we associate a real number  $u_\sigma$ . Let  $\psi$  be a twice continuously differentiable function, defined over  $(0, +\infty)$ . Then the following relation holds:

$$\begin{aligned} \left[ \frac{|K|}{\delta t} (\rho_K u_K - \rho_K^* u_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_\sigma u_\sigma \right] \psi'(u_K) \\ = \frac{|K|}{\delta t} [\rho_K \psi(u_K) - \rho_K^* \psi(u_K^*)] + \sum_{\sigma \in \mathcal{E}(K)} F_\sigma \psi(u_\sigma) + R_{K,\delta t} \end{aligned} \quad (38)$$

where:

$$R_{K,\delta t} = \frac{1}{2} \frac{|K|}{\delta t} \rho_K^* \psi''(\bar{u}_K) (u_K - u_K^*)^2 - \frac{1}{2} \sum_{\sigma \in \mathcal{E}(K)} F_\sigma \psi''(\bar{u}_\sigma) (u_\sigma - u_K)^2, \quad (39)$$

with,  $\bar{u}_K \in [\min(u_K, u_K^*), \max(u_K, u_K^*)]$  and,  $\forall \sigma \in \mathcal{E}(K)$ ,  $\bar{u}_\sigma \in [\min(u_\sigma, u_K), \max(u_\sigma, u_K)]$ .

If we now suppose that the function  $\psi$  is once continuously differentiable and convex, and that  $u_\sigma = u_K$  as soon as  $F_\sigma \geq 0$ , then expression (39) is no more valid but the residual  $R_{\sigma,\delta t}$  is non-negative.

If we now take for  $\psi$  the function  $\psi(s) = s^2/2$  and write,  $\forall \sigma \in \mathcal{E}(K)$ ,  $u_\sigma = (u_K + u_{K|\cdot})/2$  (or, in other words, define  $u_{K|\cdot}$  as  $u_{K|\cdot} = 2u_\sigma - u_K$ ), we get the following identity:

$$\begin{aligned} \left[ \frac{|K|}{\delta t} (\rho_K u_K - \rho_K^* u_K^*) + \sum_{\sigma \in \mathcal{E}(K)} F_\sigma u_\sigma \right] u_K \\ = \frac{1}{2} \frac{|K|}{\delta t} [\rho_K u_K^2 - \rho_K^* (u_K^*)^2] + \sum_{\sigma \in \mathcal{E}(K)} F_\sigma u_K u_{K|\cdot} + R_{K,\delta t}, \end{aligned} \quad (40)$$

with  $R_{K,\delta t} = \frac{1}{2} \frac{|K|}{\delta t} \rho_K^* (u_K - u_K^*)^2$ .

## REFERENCES

- [1] G. Ansanay-Alex, F. Babik, J.-C. Latché, and D. Vola. An L2-stable approximation of the Navier-Stokes convection operator for low-order non-conforming finite elements. *International Journal for Numerical Methods in Fluids*, online (2010).
- [2] R. Babik, R. Herbin, W. Kheriji, and J.-C. Latché. Discretization of the viscous dissipation term with the MAC scheme. In *Finite Volumes for Complex Applications VI - Problems and Perspectives - Prague, Czech Republic*, 2011.
- [3] H. Bijl and P. Wesseling. A unified method for computing incompressible and compressible flows in boundary-fitted coordinates. *Journal of Computational Physics*, 141:153–173, 1998.
- [4] F. Bouchut. *Nonlinear Stability of finite volume methods for hyperbolic conservation laws*. Birkhauser, 2004.
- [5] V. Casulli and D. Greenspan. Pressure method for the numerical solution of transient, compressible fluid flows. *International Journal for Numerical Methods in Fluids*, 4:1001–1012, 1984.
- [6] A.J. Chorin. Numerical solution of the Navier-Stokes equations. *Mathematics of Computation*, 22:745–762, 1968.
- [7] P. G. Ciarlet. Handbook of numerical analysis volume II : Finite elements methods – Basic error estimates for elliptic problems. In P. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Volume II*, pages 17–351. North Holland, 1991.
- [8] P. Colella and K. Pao. A projection method for low speed flows. *Journal of Computational Physics*, 149:245–269, 1999.
- [9] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations I. *Revue Française d'Automatique, Informatique et Recherche Opérationnelle (R.A.I.R.O.)*, R-3:33–75, 1973.

- [10] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number. *Journal of Computational Physics*, 229:978–1016, 2010.
- [11] I. Demirdžić, Ž. Lilek, and M. Perić. A collocated finite volume method for predicting flows at all speed. *International Journal for Numerical Methods in Fluids*, 16:1029–1050, 1993.
- [12] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. In P. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Volume VII*, pages 713–1020. North Holland, 2000.
- [13] R. Eymard, T. Gallouët, R. Herbin, and J.-C. Latché. Convergence of the MAC scheme for the compressible Stokes equations. *SIAM Journal on Numerical Analysis*, 48:2218–2246, 2010.
- [14] R. Eymard, T. Gallouët, R. Herbin, and J.-C. Latché. A convergent finite element-finite volume scheme for the compressible Stokes problem. Part II: the isentropic case. *Mathematics of Computation*, 79:649–675, 2010.
- [15] E. Feireisl. Dynamics of viscous compressible flows. volume 26 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, 2004.
- [16] T. Gallouët, L. Gastaldo, R. Herbin, and J.-C. Latché. An unconditionally stable pressure correction scheme for compressible barotropic Navier-Stokes equations. *Mathematical Modelling and Numerical Analysis*, 42:303–331, 2008.
- [17] T. Gallouët, R. Herbin, and J.-C. Latché. A convergent finite element-finite volume scheme for the compressible Stokes problem. Part I: the isothermal case. *Mathematics of Computation*, 267:1333–1352, 2009.
- [18] L. Gastaldo, R. Herbin, W. Kheriji, C. Lapuerta, and J.-C. Latché. Staggered discretizations, pressure correction schemes and all speed barotropic flows. In *Finite Volumes for Complex Applications VI - Problems and Perspectives - Prague, Czech Republic*, volume 2, pages 39–56, 2011.
- [19] J.-L. Guermond and L. Quartapelle. A projection FEM for variable density incompressible flows. *Journal of Computational Physics*, 165:167–188, 2000.
- [20] J.L. Guermond, P. Mineev, and J. Shen. An overview of projection methods for incompressible flows. *Computer Methods in Applied Mechanics and Engineering*, 195:6011–6045, 2006.
- [21] J.L. Guermond and R. Pasquetti. Entropy-based nonlinear viscosity for Fourier approximations of conservation laws. *Comptes Rendus de l'Académie des Sciences de Paris – Série I – Analyse Numérique*, 346:801–806, 2008.
- [22] J.L. Guermond, R. Pasquetti, and B. Popov. Entropy viscosity method for nonlinear conservation laws. *Journal of Computational Physics*, 230:4248–4267, 2011.
- [23] H. Guillard and A. Murrone. On the behavior of upwind schemes in the low Mach number limit: II. Godunov type schemes. *Computer & Fluids*, 33:655–675, 2004.
- [24] H. Guillard and C. Viozat. On the behavior of upwind schemes in the low Mach number limit. *Computer & Fluids*, 28:63–86, 1999.
- [25] F.H. Harlow and A.A. Amsden. Numerical calculation of almost incompressible flow. *Journal of Computational Physics*, 3:80–93, 1968.
- [26] F.H. Harlow and A.A. Amsden. A numerical fluid dynamics calculation method for all flow speeds. *Journal of Computational Physics*, 8:197–213, 1971.
- [27] F.H. Harlow and J.E. Welsh. Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface. *Physics of Fluids*, 8:2182–2189, 1965.
- [28] R. Herbin, W. Kheriji, and J.-C. Latché. Consistent staggered schemes for compressible flows – Part I: barotropic Navier-Stokes equations. *submitted*, 2011.
- [29] R. Herbin, W. Kheriji, and J.-C. Latché. Consistent staggered schemes for compressible flows – Part II: Euler equations. *submitted*, 2011.
- [30] R. Herbin, W. Kheriji, and J.-C. Latché. An unconditionally stable pressure correction scheme for the compressible Navier-Stokes equations. *submitted*, 2011.
- [31] R. Herbin and J.-C. Latché. Kinetic energy control in the MAC discretization of the compressible Navier-Stokes equations. *International Journal of Finite Volumes*, 7, 2010.
- [32] ISIS. A CFD computer code for the simulation of reactive turbulent flows.  
<https://gforge.irsn.fr/gf/project/isis>.
- [33] R.I. Issa. Solution of the implicitly discretised fluid flow equations by operator splitting. *Journal of Computational Physics*, 62:40–65, 1985.
- [34] R.I. Issa, A.D. Gosman, and A.P. Watkins. The computation of compressible and incompressible recirculating flows by a non-iterative implicit scheme. *Journal of Computational Physics*, 62:66–82, 1986.
- [35] R.I. Issa and M.H. Javareshkian. Pressure-based compressible calculation method utilizing total variation diminishing schemes. *AIAA Journal*, 36:1652–1657, 1998.
- [36] K.C. Karki and S.V. Patankar. Pressure based calculation procedure for viscous flows at all speeds in arbitrary configurations. *AIAA Journal*, 27:1167–1174, 1989.
- [37] W. Kheriji, R. Herbin, and J.-C. Latché. Pressure correction staggered schemes for barotropic monophasic and two-phase flows. *submitted*, 2011.
- [38] M.H. Kobayashi and J.C.F. Pereira. Characteristic-based pressure correction at all speeds. *AIAA Journal*, 34:272–280, 1996.

- [39] B. Larrouturou. How to preserve the mass fractions positivity when computing compressible multi-component flows. *Journal of Computational Physics*, 95:59–84, 1991.
- [40] P.-L. Lions. Mathematical topics in fluid mechanics – volume 2 – compressible models. volume 10 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, 1998.
- [41] M.-S. Liou. A sequel to AUSM, part II:AUSM+–up. *Journal of Computational Physics*, 214:137–170, 2006.
- [42] M.-S. Liou and C.J. Steffen. A new flux splitting scheme. *Journal of Computational Physics*, 107:23–39, 1993.
- [43] A. Majda and J. Sethian. The derivation and numerical solution of the equations for zero Mach number solution. *Combustion Science and Techniques*, 42:185–205, 1985.
- [44] M. Marion and R. Temam. Navier-Stokes equations: Theory and approximation. In P. Ciarlet and J.L. Lions, editors, *Handbook of Numerical Analysis, Volume VI*. North Holland, 1998.
- [45] F. Moukalled and M. Darwish. A high-resolution pressure-based algorithm for fluid flow at all speeds. *Journal of Computational Physics*, 168:101–133, 2001.
- [46] K. Nerinckx, J. Vierendeels, and E. Dick. A Mach-uniform algorithm: coupled versus segregated approach. *Journal of Computational Physics*, 224:314–331, 2007.
- [47] T.T. Nguyen, R. Herbin, and J.-C. Latché. An explicit staggered scheme for euler equations. *in preparation*, 2011.
- [48] P. Nithiarasu, R. Codina, and O.C. Zienkiewicz. The Characteristic-Based Split (CBS) scheme – A unified approach to fluid dynamics. *International Journal for Numerical Methods in Engineering*, 66:1514–1546, 2006.
- [49] A. Novotný and I. Straškraba. Introduction to the mathematical theory of compressible flow. volume 27 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, 2004.
- [50] G. Patnaik, R.H. Guirguis, J.P. Boris, and E.S. Oran. A barely implicit correction for flux-corrected transport. *Journal of Computational Physics*, 71:1–20, 1987.
- [51] PELICANS. Collaborative development environment.  
<https://gforge.irsn.fr/gf/project/pelicans>.
- [52] E.S. Politis and K.C. Giannakoglou. A pressure-based algorithm for high-speed turbomachinery flows. *International Journal for Numerical Methods in Fluids*, 25:63–80, 1997.
- [53] R. Rannacher and S. Turek. Simple nonconforming quadrilateral Stokes element. *Numerical Methods for Partial Differential Equations*, 8:97–111, 1992.
- [54] R. Temam. Sur l’approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires II. *Arch. Rat. Mech. Anal.*, 33:377–385, 1969.
- [55] E. Toro. *Riemann solvers and numerical methods for fluid dynamics – A practical introduction (third edition)*. Springer, 2009.
- [56] D.R. Van der Heul, C. Vuik, and P. Wesseling. Stability analysis of segregated solution methods for compressible flow. *Applied Numerical Mathematics*, 38:257–274, 2001.
- [57] D.R. Van der Heul, C. Vuik, and P. Wesseling. A conservative pressure-correction method for flow at all speeds. *Computer & Fluids*, 32:1113–1132, 2003.
- [58] J.P. Van Dordmaal, G.D. Raithby, and B.H. McDonald. The segregated approach to predicting viscous compressible fluid flows. *Transactions of the ASME*, 109:268–277, 1987.
- [59] D. Vidović, A. Segal, and P. Wesseling. A superlinearly convergent Mach-uniform finite volume method for the Euler equations on staggered unstructured grids. *Journal of Computational Physics*, 217:277–294, 2006.
- [60] C. Wall, C.D. Pierce, and P. Moin. A semi-implicit method for resolution of acoustic waves in low Mach number flows. *Journal of Computational Physics*, 181:545–563, 2002.
- [61] I. Wenneker, A. Segal, and P. Wesseling. A Mach-uniform unstructured staggered grid method. *International Journal for Numerical Methods in Fluids*, 40:1209–1235, 2002.
- [62] P. Wesseling. Principles of computational fluid dynamics. volume 29 of *Springer Series in Computational Mathematics*. Springer, 2001.
- [63] O.C. Zienkiewicz and R. Codina. A general algorithm for compressible and incompressible flow – Part I. The split characteristic-based scheme. *International Journal for Numerical Methods in Fluids*, 20:869–885, 1995.