



Long time behavior of finite volume discretization of symmetrizable linear hyperbolic systems

Jonathan Jung, Vincent Perrier

► To cite this version:

Jonathan Jung, Vincent Perrier. Long time behavior of finite volume discretization of symmetrizable linear hyperbolic systems. IMA Journal of Numerical Analysis, 2023, 43 (1), pp.326-356. 10.1093/imanum/drab092 . hal-03503847

HAL Id: hal-03503847

<https://hal.science/hal-03503847>

Submitted on 28 Dec 2021

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.

Long time behavior of finite volume discretization of symmetrizable linear hyperbolic systems

Jonathan Jung^{*} and Vincent Perrier[†]

December 28, 2021

Contents

1	Introduction	1
2	Discrete long time behaviour for a linear first order hyperbolic system	2
2.1	Finite volume space-discretization of hyperbolic linear systems	2
2.2	Reminder on the long time behavior of linear differential systems	3
2.3	Long time limit of finite volume schemes for symmetrizable hyperbolic systems . .	5
3	Discrete long time behaviour for the wave system	9
3.1	Verification of hypothesis of Proposition 3 and Proposition 4 for the boundary conditions and some stabilizations	10
3.2	Verification of hypothesis of Proposition 1	12
3.3	Long time limit of the discrete wave system	13
3.4	Pressure of the long time limit	13
4	Numerical experiments on the long time behavior of the wave system	15
5	Conclusion	17

Abstract

This article is dedicated to the long time behavior of a finite volume approximation of general symmetrizable linear hyperbolic system on a bounded domain. In the continuous case, this problem is very difficult, and the ω -limit set (namely the set of all the possible long time limits) may be large and complicated to depict if no dissipation is introduced. In this article, we prove that in general, with a stable finite volume scheme, the discrete solution converges to a steady state when the time goes to infinity. This property is a direct consequence of the numerical dissipation mechanisms used for stabilizing the discretization. We apply this result for determining the long time limit for several stabilizations of the wave system, and perform a formal link with the low Mach number problem of the nonlinear Euler system. Numerical experiments with the wave system are performed for confirming the theoretical results obtained.

1 Introduction

In this article, we are interested in the long time behavior of symmetrizable linear systems on a bounded domain. In the continuous case, when the system is not dissipative, this problem may be very complicated: for example, if the wave system is considered on the domain $[0, 1]$ with wall boundary conditions in 0 and 1, an initial condition will travel and reflect forever without dissipating, leading to a large and complicated ω -limit set. The aim of this article is to investigate this problem, but at the discrete case, with a stable finite volume discretization.

The long time asymptotic behavior of finite volume approximations has been investigated in several contexts, such as coagulation-fragmentation model [12], drift-diffusion equation [7], nonlinear diffusion equation [9], convection-diffusion equations [8, 5], see also [6] for a review. All

^{*}LMAP, E2S UPPA, CNRS, Pau, France and Cagire team, Inria Bordeaux Sud-Ouest, France jonathan.jung@univ-pau.fr.

[†]Cagire team, Inria Bordeaux Sud-Ouest, France and LMAP, E2S UPPA, CNRS, Pau, France, vincent.perrier@inria.fr.

these studies rely on the entropy method [2], more precisely, all the continuous systems that have been studied in these articles have an entropy that is dissipated, and the numerical schemes are proven to dissipate also these entropies, leading to an exponential decay of the relative entropy. In our case, the situation is slightly different: we are considering symmetrizable first order systems, which means that they have an entropy, but the continuous system does not have any dissipation mechanism of this entropy. Surprisingly, it is however possible to prove the existence of a long time limit for the numerical scheme thanks to the numerical dissipation.

This article is organized as follows: in Section 2, some tight sufficient conditions are given for ensuring a long time limit for the finite volume discretization of a general first order symmetrizable hyperbolic system. Then in Section 3, we prove that these existence conditions can be applied to the different known discretizations of the wave system. Last, in Section 4, numerical results on the long time behavior of the wave system are reported. The article finishes with a conclusion including a formal link between the low Mach number problem and the long time limit problem of the wave system.

2 Discrete long time behaviour for a linear first order hyperbolic system

In this section, we are interested in the long time behavior of the finite volume discretization of the following symmetrizable system

$$\partial_\tau \mathbf{U} + \sum_{k=1}^d A_k \partial_{\mathbf{x}_k} \mathbf{U} = 0, \quad (1)$$

where d is the dimension, \mathbf{U} is a vector of dimension n_{var} , and the A_k are d matrices of dimension $n_{var} \times n_{var}$. For any vector \mathbf{n} of \mathbb{R}^d , we define

$$A(\mathbf{n}) := \sum_{k=1}^d A_k \mathbf{n}_k.$$

We suppose that (1) is symmetrizable: a positive definite matrix S exists such that for all \mathbf{n} , $SA(\mathbf{n})$ is symmetric. In Section 2.1, the finite volume discretization of (1) is described.

Once discretized in space, the semi-discrete (discrete in space, continuous in time) numerical scheme can be formally written as

$$\mathcal{U}'(\tau) = \mathcal{B} - \mathcal{A}\mathcal{U}, \quad (2)$$

where \mathcal{A} is the matrix of space discretization and \mathcal{B} depends on the boundary conditions. The long time limit problem for this linear, finite-dimensional system is known, and depends on the spectral structure of \mathcal{A} . The different cases are summarized for example in [14], and we recall it in Section 2.2. Last in Section 2.3, sufficient conditions on the discretization described in Section 2.1 are derived for ensuring existence of the long time limit described in Section 2.2.

2.1 Finite volume space-discretization of hyperbolic linear systems

The domain Ω is supposed to be bounded, and \mathcal{T}_h denotes a mesh of this domain, on which a cell-centered finite volume discretization is applied. For a given cell i of \mathcal{T}_h , we denote by $|\Omega_i|$ its measure, by $\mathcal{V}(i)$ all the neighboring sides of the cell i , $\mathcal{V}_{\text{int}}(i)$ the set of the neighbors of the cell i . We suppose that we have two types of boundary conditions on $\partial\Omega$:

- *Wall boundary conditions*, in which no external value is imposed, and in which the boundary flux depends only on the internal value. The sides of the cell i on which such boundary conditions are imposed are denoted by $\mathcal{V}_{\text{wall}}(i)$.
- *Steger-Warming boundary conditions*, in which a full state \mathbf{U}_b is weakly imposed. The sides of the cell i on which such boundary conditions are imposed are denoted by $\mathcal{V}_{\text{SW}}(i)$.

We denote by Γ_{ij} the side j of the cell i , and by $|\Gamma_{ij}|$ its measure. We consider the following finite volume numerical scheme of (1)

$$\begin{aligned} \partial_\tau \mathbf{U}_i + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| F(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij}) \\ + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| F^{\text{wall}}(\mathbf{U}_i, \mathbf{n}_{ij}) \\ + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| F^{\text{SW}}(\mathbf{U}_i, \mathbf{U}_b, \mathbf{n}_{ij}) = 0, \end{aligned} \quad (3a)$$

where the numerical flux $F(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij})$ reads

$$F(\mathbf{U}_i, \mathbf{U}_j, \mathbf{n}_{ij}) = A(\mathbf{n}_{ij}) \{ \mathbf{U} \}_{ij} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_i),$$

where \mathbf{n}_{ij} is the unit vector going from i to j , and where the symbol $\{ \cdot \}_{ij}$ denotes the average of the left and right values

$$\{ \mathbf{U} \}_{ij} = \frac{\mathbf{U}_i + \mathbf{U}_j}{2}.$$

Conservativity is ensured provided $D(\mathbf{n}_{ij}) = D(\mathbf{n}_{ji})$. We now detail the two types of boundary conditions that are used in (3a).

Steger-Warming boundary conditions. In this case, we wish to impose weakly a state \mathbf{U}_b . The matrix $A(\mathbf{n})$ can be diagonalized in \mathbb{R}

$$\exists Q \in \text{GL}_{n_{var}}(\mathbb{R}) \quad A(\mathbf{n}) = Q^{-1}(\mathbf{n}) \Lambda(\mathbf{n}) Q(\mathbf{n}),$$

where $\Lambda(\mathbf{n})$ is a diagonal matrix. Denoting by $\Lambda^+(\mathbf{n})$ (resp. $\Lambda^-(\mathbf{n})$) the diagonal matrix such that $\Lambda_{i,i}^+(\mathbf{n}) = \max(\Lambda_{i,i}(\mathbf{n}), 0)$ (resp. $\Lambda_{i,i}^-(\mathbf{n}) = \min(\Lambda_{i,i}(\mathbf{n}), 0)$), we define

$$A^+(\mathbf{n}) := Q^{-1}(\mathbf{n}) \Lambda^+(\mathbf{n}) Q(\mathbf{n}) \quad A^-(\mathbf{n}) := Q^{-1}(\mathbf{n}) \Lambda^-(\mathbf{n}) Q(\mathbf{n}).$$

Given a boundary state \mathbf{U}_b , and \mathbf{n} the outgoing normal, the Steger-Warming boundary flux is defined as

$$F^{\text{SW}}(\mathbf{U}_i, \mathbf{U}_b, \mathbf{n}) = A^+(\mathbf{n}) \mathbf{U}_i + A^-(\mathbf{n}) \mathbf{U}_b.$$

This boundary condition can be reformulated as

$$\begin{aligned} F^{\text{SW}}(\mathbf{U}_i, \mathbf{U}_b, \mathbf{n}) &= A(\mathbf{n}) \mathbf{U}_b + A^+(\mathbf{n}) (\mathbf{U}_i - \mathbf{U}_b) \\ &= A(\mathbf{n}) \frac{\mathbf{U}_i + \mathbf{U}_b}{2} - \frac{1}{2} (A^+(\mathbf{n}) - A^-(\mathbf{n})) (\mathbf{U}_b - \mathbf{U}_i). \end{aligned} \quad (3b)$$

Wall boundary conditions. In this case, the boundary flux depends only on the interior value, but we require some dissipation in some sense. The boundary flux is

$$F^{\text{wall}}(\mathbf{U}_i, \mathbf{n}) = A(\mathbf{n}) \frac{\mathbf{U}_i + \tilde{\mathbf{U}}_i}{2} - \frac{1}{2} (A^+(\mathbf{n}) - A^-(\mathbf{n})) (\tilde{\mathbf{U}}_i - \mathbf{U}_i)$$

where $\tilde{\mathbf{U}}_i$ is obtained from a transformation of \mathbf{U}_i : $\tilde{\mathbf{U}}_i = (\mathbf{I}_n - P_{\text{wall}}(\mathbf{n})) \mathbf{U}_i$ where $P_{\text{wall}}(\mathbf{n})$ is positive semi-definite. The flux can be rewritten as

$$F^{\text{wall}}(\mathbf{U}_i, \mathbf{n}) = A(\mathbf{n}) \frac{\mathbf{U}_i}{2} + \frac{1}{2} (A(\mathbf{n}) - 2A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n})) \mathbf{U}_i. \quad (3c)$$

2.2 Reminder on the long time behavior of linear differential systems

2.2.1 Continuous differential system

In this subsection, we are interested in the long time behavior of (2). We denote by N the size of the system. The solution of this system reads

$$\mathbf{u}(\tau) = \exp(-\tau \mathcal{A}) \mathbf{u}_0 + \int_0^\tau \exp(-(\tau - t) \mathcal{A}) \mathbf{B} dt.$$

We are first interested in the long time limit of the homogeneous problem,

$$\mathbf{u}(\tau) = \exp(-\tau \mathcal{A}) \mathbf{u}_0. \quad (4)$$

\mathbb{C}^N can be decomposed as the sum of generalized eigenspace of \mathcal{A} . The operator \mathcal{A} restricted to a given generalized eigenspace E_λ associated with the eigenvalue λ is equal to $\lambda \mathbf{I}_d + N_\lambda$, where N_λ is a nilpotent matrix. Denoting by l the nilpotent index of N_λ , the exponential of the operator on E_λ can be written as the following finite sum

$$\exp(-\tau \mathcal{A})|_{E_\lambda} = \exp(-\lambda \tau) \sum_{k=0}^{l-1} \frac{(-1)^k \tau^k}{k!} N_\lambda^k.$$

Then (4) is bounded for $\tau > 0$ if and only if either $\text{Re}(\lambda) > 0$, or $\text{Re}(\lambda) = 0$ and $N_\lambda = 0$. It has a finite limit if and only if the only eigenvalue such that $\text{Re}(\lambda) = 0$ is 0, i.e. \mathcal{A} has no purely imaginary eigenvalue. Noting that $N_0 = 0$ is equivalent to $\text{range } \mathcal{A} \cap \ker \mathcal{A} = \{0\}$, we define the following set of hypothesis

Hypothesis 1. Suppose that \mathcal{A} follows the following conditions

- [1] $\text{range } \mathcal{A} \cap \ker \mathcal{A} = \{0\}$.
- [2] All the eigenvalues λ of \mathcal{A} are such that $\text{Re}(\lambda) \geq 0$.
- [3] 0 is the only eigenvalue of \mathcal{A} such that $\text{Re}(\lambda) = 0$.

We can then prove that

Proposition 1 (Long time limit of (2)). Suppose that Hypothesis 1 hold. Then

- A given solution of (2) is bounded for $\tau \geq 0$, if and only if $\mathcal{B} \in \text{range } \mathcal{A}$.
- Denoting by \mathcal{U}^∞ the unique vector of $\text{range } \mathcal{A}$ such that $\mathcal{B} = \mathcal{A}\mathcal{U}^\infty$, (2) with initial data \mathcal{U}_0 converges towards $\mathcal{U}^\infty + \mathbb{P}(\mathcal{U}_0)$, where \mathbb{P} is the projection on $\ker \mathcal{A}$ along $\text{range } \mathcal{A}$.
- The ω -limit set of the whole space is an affine space, equal to $\mathcal{U}^\infty + \ker \mathcal{A}$.

Proof. Under Hypothesis 1, any vector \mathcal{B} can be decomposed into $\mathcal{B}_1 + \mathcal{B}_2$, with $\mathcal{B}_1 \in \text{range } \mathcal{A}$ and $\mathcal{B}_2 \in \ker \mathcal{A}$. Then

$$\begin{aligned} \int_0^\tau \exp(-(\tau-t)\mathcal{A}) \mathcal{B} dt &= \int_0^\tau \exp(-(\tau-t)\mathcal{A}) \mathcal{B}_1 dt + \int_0^\tau \exp(-(\tau-t)\mathcal{A}) \mathcal{B}_2 dt \\ &= \int_0^\tau \exp(-(\tau-t)\mathcal{A}) \mathcal{B}_1 dt + \tau \mathcal{B}_2. \end{aligned}$$

$\mathcal{A}|_{\text{range } \mathcal{A}}$ has eigenvalues with strictly positive real part, so that $\int_0^\tau \exp(-(\tau-t)\mathcal{A}) \mathcal{B}_1 dt \rightarrow 0$ when $\tau \rightarrow \infty$, whereas $\|\tau \mathcal{B}_2\| \rightarrow \infty$ when $\tau \rightarrow \infty$, except if $\mathcal{B}_2 = 0$. We conclude that (2) has a long time limit if and only if Hypothesis 1 holds and $\mathcal{B} \in \text{range } \mathcal{A}$. We remark that

$$\exp(-\tau\mathcal{A})\mathcal{U}_0 \rightarrow \mathbb{P}(\mathcal{U}_0).$$

Moreover, if $\mathcal{B} \in \text{range } \mathcal{A}$, we denote by \mathcal{B}'_1 an element such that $\mathcal{B} = \mathcal{A}\mathcal{B}'_1$. Then

$$\begin{aligned} \int_0^\tau \exp(-(\tau-t)\mathcal{A}) \mathcal{A}\mathcal{B}'_1 dt &= [\exp(-(\tau-t)\mathcal{A}) \mathcal{B}'_1]_0^\tau \\ &= \mathcal{B}'_1 - \exp(-\tau\mathcal{A}) \mathcal{B}'_1 \end{aligned}$$

which converges towards $\mathcal{B}'_1 - \mathbb{P}(\mathcal{B}'_1) = \mathcal{U}^\infty$. We conclude that $\mathcal{U}(\tau) \rightarrow \mathcal{U}^\infty + \mathbb{P}(\mathcal{U}_0)$. The last point of Proposition 1 is straightforward. \square

2.2.2 Time discretized system

From a practical point of view, the system (2) will be discretized in time. The aim of this subsection is to study the long time behaviour of the fully discrete dynamical system obtained with a forward or backward Euler time discretization. This behaviour is summarized in the following proposition

Proposition 2 (Long time behaviour of the time-discretized system). We suppose that Hypothesis 1 hold, and take the same notations as Proposition 1. For a given \mathcal{U}_0 and if $\mathcal{B} \in \text{range } \mathcal{A}$, we consider (2) discretized with either the forward or the backward Euler method, with the time step δt . We denote by \mathcal{U}^n the serie obtained after time discretization. Then

- with the forward Euler method, \mathcal{U}^n converges towards $\mathcal{U}^\infty + \mathbb{P}(\mathcal{U}_0)$ in infinite time under the CFL (for Courant Friedrichs Lewy) condition

$$\forall \lambda \in \text{Sp}(\mathcal{A}), \lambda \neq 0 \quad \delta t < \frac{1}{|\lambda|}.$$

- with the backward Euler method, \mathcal{U}^n converges towards $\mathcal{U}^\infty + \mathbb{P}(\mathcal{U}_0)$ in infinite time unconditionally.

Proof. We first consider the discrete system with the forward Euler method

$$\begin{cases} \mathbf{u}^{n+1} = \delta t \mathbf{B} + (\mathbf{I}_d - \delta t \mathbf{A}) \mathbf{u}^n \\ \mathbf{u}^0 = \mathbf{u}_0 \end{cases}$$

We define \mathbf{v}^n as

$$\mathbf{v}^n := \mathbf{u}^n - (\mathbf{u}^\infty + \mathbb{P}(\mathbf{u}_0)).$$

\mathbf{v}^n is then such that $\mathbb{P}(\mathbf{v}^0) = 0$ and

$$\mathbf{v}^{n+1} = (\mathbf{I}_d - \delta t \mathbf{A}) \mathbf{v}^n,$$

so that

$$\mathbf{v}^n = (\mathbf{I}_d - \delta t \mathbf{A})^n \mathbf{v}^0.$$

Any eigenspace E_λ of \mathbf{A} associated with the eigenvalue λ is stable by $(\mathbf{I}_d - \delta t \mathbf{A})^n$. On this generalized eigenspace, \mathbf{A} is equal to $\lambda \mathbf{I}_d + N_\lambda$, where N_λ is a nilpotent matrix of index l . Then, on E_λ , we have

$$(\mathbf{I}_d - \delta t \mathbf{A})^n = (\mathbf{I}_d - \delta t (\lambda \mathbf{I}_d + N_\lambda))^n = \sum_{k=0}^l \frac{n(n-1) \dots (n-k+1)}{k!} (1 - \delta t \lambda)^{n-k} (-\delta t N_\lambda)^k.$$

if $\lambda \neq 0$, the above matrices all converge to 0 under the CFL condition. For $\lambda = 0$, $N_\lambda = 0$, and $(\mathbf{I}_d - \delta t \mathbf{A})^n$ is the identity matrix. However, as $\mathbb{P}(\mathbf{v}_0) = 0$, we conclude that $\mathbf{v}^n \rightarrow 0$ when $n \rightarrow \infty$. We now consider the backward Euler method. Then the same \mathbf{v}^n follows

$$\mathbf{v}^n = (\mathbf{I}_d + \delta t \mathbf{A})^{-n} \mathbf{v}_0.$$

Note that the above invert makes sense because the eigenvalues of \mathbf{A} have positive real part. Each generalized eigenspace is stable for \mathbf{A} , but also for $(\mathbf{I}_d + \delta t \mathbf{A})^{-1}$ and so for $(\mathbf{I}_d + \delta t \mathbf{A})^{-n}$. On the generalized eigenspace E_λ , \mathbf{A} is equal to $\lambda \mathbf{I}_d + N_\lambda$, so that

$$(\mathbf{I}_d + \delta t \mathbf{A})^{-n} = ((1 + \lambda \delta t) \mathbf{I}_d + \delta t N_\lambda)^{-n} = \sum_{k=0}^l \frac{(n+k-1)(n+k-2) \dots n}{k!} (1 + \lambda \delta t)^{-n-k} (-\delta t)^k N_\lambda^k$$

In each term of this sum, the leading term for $\lambda \neq 0$ is $(1 + \lambda \delta t)^{-n}$, which is going exponentially to 0. For $\lambda = 0$, the matrix is the identity matrix, but as seen before, $\mathbb{P}(\mathbf{v}_0) = 0$. We conclude that $\mathbf{v}^n \rightarrow 0$ when $n \rightarrow \infty$, which ends the proof for the backward Euler scheme. \square

Remark 1. For the forward Euler method, the convergence is exponential, driven by $n^{l_{\min}-1} (1 - \delta t \lambda_{\min})^n$ where λ_{\min} is the eigenvalue with the smallest nonzero real part, and l_{\min} the matching nilpotent index. For the backward Euler scheme, the convergence is exponential, and the slowest component is behaving like $n^{l_{\min}-1} (1 + \lambda_{\min} \delta t)^{-n}$.

2.3 Long time limit of finite volume schemes for symmetrizable hyperbolic systems

Now, we are interested in applying Proposition 1 to the finite volume discretization (3) of (1). In this case, the size N of the matrix \mathbf{A} and of the vectors \mathbf{u} and \mathbf{B} is equal to $N = n_{var} \times N_{Cell}$, where N_{Cell} is the number of cells of the mesh \mathcal{T}_h . We denote by \mathcal{F}_{int} the set of the interior faces, by $\mathcal{F}_{\text{wall}}$ the set of the boundary faces on which a wall boundary condition is imposed, and by \mathcal{F}_{SW} the set of the boundary faces on which a Steger-Warming boundary condition is imposed.

We denote by \mathcal{S} the block-diagonal matrix where each block is $|\Omega_i|S$ where S is the symmetrizer of the continuous system. In a more general manner, we will denote by bold capital letters (e.g. \mathbf{X}) the vectors of size n_{var} , and by typewriter font (e.g. \mathcal{X}) the vectors of size N .

For a real matrix M , we denote by $\mathcal{S}[M]$ its symmetric part and by $\mathcal{A}[M]$ its skew-symmetric part

$$\mathcal{S}[M] = \frac{M + M^T}{2}, \quad \mathcal{A}[M] = \frac{M - M^T}{2}$$

so that $M = (\mathcal{S} + \mathcal{A})[M]$.

Proposition 3. If the scheme defined by (3) is

- conservative: $D(\mathbf{n}) = D(-\mathbf{n})$,

- *dissipative* :

$$\forall \mathbf{U} \in \mathbb{R}^{n_{var}}, \quad \mathbf{U}^T S D(\mathbf{n}) \mathbf{U} \geq 0,$$

- *dissipative at the wall boundary condition* :

$$\forall \mathbf{U} \in \mathbb{R}^{n_{var}}, \quad \mathbf{U}^T S (A(\mathbf{n}) - 2A^-(\mathbf{n})P_{wall}(\mathbf{n})) \mathbf{U} \geq 0,$$

- *dissipative at the Steger-Warming boundary condition* :

$$\forall \mathbf{U} \in \mathbb{R}^{n_{var}}, \quad \mathbf{U}^T S (A^+(\mathbf{n}) - A^-(\mathbf{n})) \mathbf{U} \geq 0,$$

then points [1](#) and [2](#) of Hypothesis 1 hold. Moreover, $\ker \mathcal{A}$ and $\text{range } \mathcal{A}$ are S -orthogonal.

Proof. The scheme (3) can be put in the form (2) with

$$\begin{aligned} (\mathcal{A}\mathbf{U})_i := & \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \{ \mathbf{U} \}_{ij} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_i) \right) \\ & + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i}{2} + \frac{1}{2} (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij})P_{\text{wall}}(\mathbf{n}_{ij})) \mathbf{U}_i \right) \\ & + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i}{2} + \frac{1}{2} (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) \mathbf{U}_i \right). \end{aligned}$$

As $\sum_{j \in \mathcal{V}(i)} |\Gamma_{ij}| \mathbf{n}_{ij} = 0$, we have also

$$\sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \frac{A(\mathbf{n}_{ij}) \mathbf{U}_i}{2} + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \frac{A(\mathbf{n}_{ij}) \mathbf{U}_i}{2} + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \frac{A(\mathbf{n}_{ij}) \mathbf{U}_i}{2} = 0.$$

Thus,

$$\begin{aligned} (\mathcal{A}\mathbf{U})_i = & \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_j}{2} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_i) \right) \\ & + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \frac{1}{2} (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij})P_{\text{wall}}(\mathbf{n}_{ij})) \mathbf{U}_i \\ & + \frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \frac{1}{2} (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) \mathbf{U}_i. \end{aligned}$$

And now we can calculate for \mathcal{X} and \mathcal{Y} in \mathbb{C}^N

$$\begin{aligned} \bar{\mathcal{Y}}^T S \mathcal{A} \mathcal{X} = & \sum_{i \in \mathcal{M}} \left\{ \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(\frac{1}{2} \bar{\mathcal{Y}}_i^T S A(\mathbf{n}_{ij}) \mathbf{X}_j - \bar{\mathcal{Y}}_i^T S D(\mathbf{n}_{ij}) (\mathbf{X}_j - \mathbf{X}_i) \right) \right. \\ & + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \frac{1}{2} \bar{\mathcal{Y}}_i^T S (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij})P_{\text{wall}}(\mathbf{n}_{ij})) \mathbf{X}_i \\ & \left. + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \frac{1}{2} \bar{\mathcal{Y}}_i^T S (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) \mathbf{X}_i \right\}. \end{aligned}$$

We switch to a sum on all the faces of the mesh

$$\begin{aligned} \bar{\mathcal{Y}}^T S \mathcal{A} \mathcal{X} = & \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| \left(\frac{1}{2} \bar{\mathcal{Y}}_L^T S A(\mathbf{n}_f) \mathbf{X}_R - \bar{\mathcal{Y}}_L^T S D(\mathbf{n}_f) (\mathbf{X}_R - \mathbf{X}_L) \right. \\ & + \frac{1}{2} \bar{\mathcal{Y}}_R^T S A(-\mathbf{n}_f) \mathbf{X}_L - \bar{\mathcal{Y}}_R^T S D(-\mathbf{n}_f) (\mathbf{X}_L - \mathbf{X}_R) \Big) \\ & + \sum_{f \in \mathcal{F}_{\text{wall}}} |\Gamma_f| \frac{1}{2} \bar{\mathcal{Y}}_L^T S (A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f)) \mathbf{X}_L \\ & + \sum_{f \in \mathcal{F}_{\text{SW}}} |\Gamma_f| \frac{1}{2} \bar{\mathcal{Y}}_L^T S (A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f)) \mathbf{X}_L, \end{aligned}$$

where \mathbf{X}_L and \mathbf{X}_R are the left and right values relative to the a given face f . As $A(-\mathbf{n}) = -A(\mathbf{n})$, and as $D(\mathbf{n}) = D(-\mathbf{n})$, it can be rewritten as

$$\begin{aligned} \bar{\mathcal{Y}}^T S \mathcal{A} \mathcal{X} = & \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| \left(\frac{1}{2} \left(\bar{\mathcal{Y}}_L^T S A(\mathbf{n}_f) \mathbf{X}_R - \bar{\mathcal{Y}}_R^T S A(\mathbf{n}_f) \mathbf{X}_L \right) \right. \\ & + \left. \overline{(\mathbf{Y}_R - \mathbf{Y}_L)}^T S D(\mathbf{n}_f) (\mathbf{X}_R - \mathbf{X}_L) \right) \\ & + \sum_{f \in \mathcal{F}_{\text{wall}}} |\Gamma_f| \frac{1}{2} \bar{\mathcal{Y}}_L^T S (A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f)) \mathbf{X}_L \\ & + \sum_{f \in \mathcal{F}_{\text{SW}}} |\Gamma_f| \frac{1}{2} \bar{\mathcal{Y}}_L^T S (A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f)) \mathbf{X}_L. \end{aligned} \tag{5}$$

We can now prove that [1] and [2] of Hypothesis 1 hold.

Point [1] of Hypothesis 1. Suppose that $\mathbf{v} \in \ker \mathcal{A} \cap \text{range } \mathcal{A}$. Then, it exists $\mathbf{u} \in \mathbb{R}^N$ such that $\mathbf{v} = \mathcal{A}\mathbf{u}$. We are going to prove that $\mathbf{v}^T \mathcal{S} \mathcal{A} \mathbf{u} = -\mathbf{u}^T \mathcal{S} \mathcal{A} \mathbf{v}$ so that

$$\mathbf{v}^T \mathcal{S} \mathbf{v} = \mathbf{v}^T \mathcal{S} \mathcal{A} \mathbf{u} = -\mathbf{u}^T \mathcal{S} \mathcal{A} \mathbf{v} = 0$$

and then $\mathbf{v} = 0$. Using (5) and writting matrices $SD(\mathbf{n}_f)$, $S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))$ and $S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))$ as the sum of their symmetric part and skew-symmetric part, we have

$$\begin{aligned} \mathbf{v}^T \mathcal{S} \mathcal{A} \mathbf{u} = & \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| \left(\frac{1}{2} (\mathbf{V}_L^T \mathcal{S} A(\mathbf{n}_f) \mathbf{U}_R - \mathbf{V}_R^T \mathcal{S} A(\mathbf{n}_f) \mathbf{U}_L) \right. \\ & + (\mathbf{V}_R - \mathbf{V}_L)^T (\mathcal{S} + \mathcal{A}) [SD(\mathbf{n}_f)] (\mathbf{U}_R - \mathbf{U}_L) \Big) \\ & + \sum_{f \in \mathcal{F}_{\text{wall}}} |\Gamma_f| \frac{1}{2} \mathbf{V}_L^T (\mathcal{S} + \mathcal{A}) [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))] \mathbf{U}_L \\ & + \sum_{f \in \mathcal{F}_{\text{SW}}} |\Gamma_f| \frac{1}{2} \mathbf{V}_L^T (\mathcal{S} + \mathcal{A}) [S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))] \mathbf{U}_L \end{aligned} \quad (6)$$

where $\mathbf{V}_L^T \mathcal{S} A(\mathbf{n}_f) \mathbf{U}_R - \mathbf{V}_R^T \mathcal{S} A(\mathbf{n}_f) \mathbf{U}_L = -(\mathbf{U}_L^T \mathcal{S} A(\mathbf{n}_f) \mathbf{V}_R - \mathbf{U}_R^T \mathcal{S} A(\mathbf{n}_f) \mathbf{V}_L)$ because $\mathcal{S} A(\mathbf{n}_f)$ is symmetric. Moreover, for any matrix M , any vector $\mathbf{U} \in \ker \mathcal{S}[M]$ and any $\mathbf{V} \in \mathbb{R}^N$, we have $\mathcal{S}[M]\mathbf{U} = 0 = \mathbf{U}^T \mathcal{S}[M]$ so that

$$\begin{aligned} \mathbf{U}^T M \mathbf{V} &= \mathbf{U}^T (\mathcal{S} + \mathcal{A}) [M] \mathbf{V} = \mathbf{U}^T \mathcal{A} [M] \mathbf{V} = -\mathbf{V}^T \mathcal{A} [M] \mathbf{U} \\ &= -(\mathbf{V}^T \mathcal{S} [M] \mathbf{U} + \mathbf{V}^T \mathcal{A} [M] \mathbf{U}) \\ &= -\mathbf{V}^T M \mathbf{U}. \end{aligned}$$

Then, to get that $\mathbf{v}^T \mathcal{S} \mathcal{A} \mathbf{u} = -\mathbf{u}^T \mathcal{S} \mathcal{A} \mathbf{v}$ with (6), it is sufficient to prove that the symmetric part of (6) vanishes. However, since $\mathbf{v} \in \ker \mathcal{A}$, we have

$$\begin{aligned} 0 = \mathbf{v}^T \mathcal{S} \mathcal{A} \mathbf{v} = & \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| \left(\frac{1}{2} (\mathbf{V}_L^T \mathcal{S} A(\mathbf{n}_f) \mathbf{V}_R - \mathbf{V}_R^T \mathcal{S} A(\mathbf{n}_f) \mathbf{V}_L) \right. \\ & + (\mathbf{V}_R - \mathbf{V}_L)^T \mathcal{S} [SD(\mathbf{n}_f)] (\mathbf{V}_R - \mathbf{V}_L) \Big) \\ & + \sum_{f \in \mathcal{F}_{\text{wall}}} |\Gamma_f| \frac{1}{2} \mathbf{V}_L^T \mathcal{S} [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))] \mathbf{V}_L \\ & + \sum_{f \in \mathcal{F}_{\text{SW}}} |\Gamma_f| \frac{1}{2} \mathbf{V}_L^T \mathcal{S} [S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))] \mathbf{V}_L \end{aligned}$$

where $\mathbf{U}_L^T \mathcal{S} A(\mathbf{n}_f) \mathbf{V}_R - \mathbf{V}_R^T \mathcal{S} A(\mathbf{n}_f) \mathbf{U}_L = 0$ because $\mathcal{S} A(\mathbf{n}_f)$ is a symmetric matrix. As $SD(\mathbf{n}_f)$, $S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))$, and $S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))$ are positive semi-definite real matrix, their symmetric part are positive semi-definite and the right hand side is positive, so that all terms in the right hand side are null. We recall that for a symmetric, positive semi-definite matrix, the isotropic cone is equal to its kernel. Then, we have on all interior faces,

$$\mathbf{V}_R - \mathbf{V}_L \in \ker \mathcal{S} [SD(\mathbf{n}_f)],$$

on all neighboring cells of a wall,

$$\mathbf{V}_L \in \ker \mathcal{S} [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))],$$

and on all neighboring cells of a Steger-Warming boundary condition,

$$\mathbf{V}_L \in \ker \mathcal{S} [S(A^-(\mathbf{n}_f) - A^+(\mathbf{n}_f))].$$

Then, using (6) we directly get that $\mathbf{v}^T \mathcal{S} \mathcal{A} \mathbf{u} = -\mathbf{u}^T \mathcal{S} \mathcal{A} \mathbf{v}$ so that $\mathbf{v}^T \mathcal{S} \mathbf{v} = 0$ and $\mathbf{v} = 0$. We conclude that

$$\ker \mathcal{A} \cap \text{range } \mathcal{A} = \{0\}.$$

Point [2] of Hypothesis 1. We denote by λ an eigenvalue of \mathcal{A} , and by $\mathbf{x} \in \mathbb{C}^N$ an eigenvector associated with λ . Then

$$\overline{\mathbf{x}}^T \mathcal{S} \mathcal{A} \mathbf{x} = \lambda \overline{\mathbf{x}}^T \mathcal{S} \mathbf{x}.$$

As $SA(\mathbf{n}_f)$ is symmetric and real, $\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R - \bar{\mathbf{X}}_R^T SA(\mathbf{n}_f) \mathbf{X}_L = 2i \operatorname{Im}(\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R)$. Adding (5) evaluated with $\mathbf{y} = \mathbf{x}$, with its transconjugate gives

$$\begin{aligned} \operatorname{Re}(\lambda) \bar{\mathbf{x}}^T \mathcal{S} \mathbf{x} &= \operatorname{Re}(\lambda \bar{\mathbf{x}}^T \mathcal{S} \mathbf{x}) = \operatorname{Re}(\bar{\mathbf{x}}^T \mathcal{S} A \mathbf{x}) \\ &= \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| \left(\overline{(\mathbf{X}_R - \mathbf{X}_L)}^T \mathcal{S} [SD(\mathbf{n}_f)] (\mathbf{X}_R - \mathbf{X}_L) \right) \\ &\quad + \sum_{f \in \mathcal{F}_{\text{wall}}} |\Gamma_f| \frac{1}{2} \bar{\mathbf{X}}_L^T \mathcal{S} [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))] \mathbf{X}_L \\ &\quad + \sum_{f \in \mathcal{F}_{\text{SW}}} |\Gamma_f| \frac{1}{2} \bar{\mathbf{X}}_L^T \mathcal{S} [S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))] \mathbf{X}_L, \end{aligned} \quad (7)$$

because \mathcal{S} is a symmetric and real matrix. As $\mathcal{S} [SD(\mathbf{n}_f)]$, $\mathcal{S} [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))]$, and $\mathcal{S} [S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))]$ are positive semi-definite real matrix, the right hand side is positive semi-definite, so that $\operatorname{Re}(\lambda) \bar{\mathbf{x}}^T \mathcal{S} \mathbf{x} \geq 0$. Since $\mathbf{x} \neq 0$, $\bar{\mathbf{x}}^T \mathcal{S} \mathbf{x} > 0$, then $\operatorname{Re}(\lambda) \geq 0$, and point [2] of Hypothesis 1 holds. \square

Proposition 4. *If the scheme defined by (3) satisfies the hypothesis of Proposition 3 and if*

- *for all $\mathbf{x} \in \mathbb{C}^N$ such that*

$$\begin{cases} \forall f \in \mathcal{F}_{\text{int}}, & \mathbf{X}_R - \mathbf{X}_L \in \ker \mathcal{S} [SD(\mathbf{n}_f)] \\ \forall f \in \mathcal{F}_{\text{wall}}, & \mathbf{X}_L \in \ker \mathcal{S} [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))] \\ \forall f \in \mathcal{F}_{\text{SW}}, & \mathbf{X}_L \in \ker \mathcal{S} [S(A^-(\mathbf{n}_f) - A^+(\mathbf{n}_f))] \end{cases} \quad (8)$$

we have

$$\forall f \in \mathcal{F}_{\text{int}}, \quad \operatorname{Im}(\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R) = 0$$

- *the kernel of $\mathcal{S}[M]$ is $\mathcal{A}[M]$ -orthogonal, for the following matrices M*

$$\begin{cases} SD(\mathbf{n}) \\ S(A(\mathbf{n}) - 2A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n})) \\ S(A^+(\mathbf{n}) - A^-(\mathbf{n})) \end{cases}$$

then point [3] of Hypothesis 1 holds.

Proof. We assume that λ is an eigenvalue of \mathcal{A} such that $\operatorname{Re}(\lambda) = 0$ and we want prove that $\lambda = 0$. We denote by \mathbf{x} an eigenvector associated with λ . Since for all real matrix M and all vector $\mathbf{x} \in \mathbb{C}^N$, $\bar{\mathbf{x}}^T \mathcal{S} [M] \mathbf{x}$ is real, using (5) with $\mathbf{y} = \mathbf{x}$ and taking the imaginary part (as we take the real part in (7)) gives

$$\begin{aligned} i \operatorname{Im}(\lambda) \bar{\mathbf{x}}^T \mathcal{S} \mathbf{x} &= i \operatorname{Im}(\lambda \bar{\mathbf{x}}^T \mathcal{S} \mathbf{x}) \\ &= \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| \left(i \operatorname{Im}(\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R) \right. \\ &\quad \left. + \overline{(\mathbf{X}_R - \mathbf{X}_L)}^T \mathcal{A} [SD(\mathbf{n}_f)] (\mathbf{X}_R - \mathbf{X}_L) \right) \\ &\quad + \sum_{f \in \mathcal{F}_{\text{wall}}} |\Gamma_f| \frac{1}{2} \bar{\mathbf{X}}_L^T \mathcal{A} [S(A(\mathbf{n}_f) - 2A^-(\mathbf{n}_f)P_{\text{wall}}(\mathbf{n}_f))] \mathbf{X}_L \\ &\quad + \sum_{f \in \mathcal{F}_{\text{SW}}} |\Gamma_f| \frac{1}{2} \bar{\mathbf{X}}_L^T \mathcal{A} [S(A^+(\mathbf{n}_f) - A^-(\mathbf{n}_f))] \mathbf{X}_L. \end{aligned}$$

Since $\operatorname{Re}(\lambda) = 0$, (7) gives that \mathbf{x} satisfies all the conditions (8) and so are their real and imaginary parts. Then, using the hypothesis of Proposition 4, we get

$$\begin{aligned} \operatorname{Im}(\lambda) \bar{\mathbf{x}}^T \mathcal{S} \mathbf{x} &= \sum_{f \in \mathcal{F}_{\text{int}}} \operatorname{Im}(\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R) \\ &= 0 \end{aligned}$$

where $\mathbf{x} \neq 0$, so that $\operatorname{Im}(\lambda) = 0$ and the only eigenvalue that has a zero real part is 0. \square

Note that the first item of Proposition 4 holds under the stronger following hypothesis

$$\mathbf{X}_R - \mathbf{X}_L \in \ker \mathcal{S} [SD(\mathbf{n}_f)] \implies \operatorname{Im}(\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R) = 0. \quad (9)$$

Remark 2. *The convergence rate to the long time limit is driven by the spectral gap of the finite volume discretization (3). We point out that in our case, only the numerical dissipation is responsible for the existence of the long time limit. We therefore cannot rely on any property of the continuous system for having an estimate of the spectral gap.*

Evaluating numerically the spectral gap is also a difficult problem as the matrix of the discretization of (3) is not symmetric.

In this section, general hypothesis were derived for ensuring the existence of a discrete long time limit for a first order linear hyperbolic system. Next section is dedicated to the application of these propositions to the first order wave system.

3 Discrete long time behaviour for the wave system

The aim of this section is to prove that the results of the previous section can be applied in different numerical schemes for the first order wave system

$$\begin{cases} \partial_\tau p + \frac{1}{\rho_0} \operatorname{div}_{\mathbf{x}} \mathbf{u} = 0, \\ \partial_\tau \mathbf{u} + \kappa_0 \nabla p = 0, \end{cases} \quad (10)$$

where κ_0 and ρ_0 are two strictly positive real numbers. Of course, the system (10) is hyperbolic, and its symmetrizer is

$$S = \begin{pmatrix} \rho_0 \kappa_0 & 0 \\ 0 & \mathbf{I}_d \end{pmatrix}.$$

The wave velocity c_0 of (10) is defined as

$$c_0 = \sqrt{\frac{\kappa_0}{\rho_0}}.$$

The scheme (3) is applied to (10). The matrix $P_{\text{wall}}(\mathbf{n})$ of (3c) is

$$P_{\text{wall}}(\mathbf{n}) = \begin{pmatrix} 0 & 0 \\ 0 & 2\mathbf{n}\mathbf{n}^T \end{pmatrix}. \quad (11)$$

Moreover, we will consider the following stabilization matrices $D(\mathbf{n})$ for the numerical scheme (3a) of the wave system

- The Rusanov flux

$$D(\mathbf{n}) = \frac{c_0}{2} \mathbf{I}_d, \quad (12a)$$

- The Godunov flux

$$D(\mathbf{n}) = \frac{c_0}{2} \begin{pmatrix} 1 & 0 \\ 0 & \mathbf{n}\mathbf{n}^T \end{pmatrix}, \quad (12b)$$

- The Godunov flux with a centered discretization of the pressure gradient

$$D(\mathbf{n}) = \frac{c_0}{2} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad (12c)$$

- The fix of [4] (that we will call LMAAP, for Low Mach Acoustic Accuracy Preserving) which can be seen as an extension of [1] on unstructured meshes

$$D(\mathbf{n}) = \begin{pmatrix} c_0 & \frac{\varepsilon}{2\rho_0} \mathbf{C}_{21}^T \\ -\frac{\varepsilon\kappa_0}{2} \mathbf{C}_{21} & 0 \end{pmatrix}, \quad (12d)$$

where $\varepsilon = \pm 1$, $\mathbf{C}_{21} = \frac{1}{\sqrt{d}} \mathbf{1}_d$ and $\mathbf{1}_d = (1, \dots, 1)^T$.

We first prove that hypothesis of Proposition 3 and Proposition 4 hold for the numerical scheme in Section 3.1. Then we check that all the hypothesis of Proposition 1 hold in Section 3.2.

3.1 Verification of hypothesis of Proposition 3 and Proposition 4 for the boundary conditions and some stabilizations

In this whole section, for being consistent with the notations of (10), we will denote by (p, \mathbf{u}) the components of a vector \mathbf{U} . Hypothesis of Proposition 3 and Proposition 4 will be checked for each of the boundary conditions and stabilizations of (12).

3.1.1 Steger-Warming boundary condition

We first prove that the matrix $S(A^+(\mathbf{n}) - A^-(\mathbf{n}))$ is positive semi-definite. We denote by $S^{1/2}$ a symmetric positive definite square root of the symmetrizer S (which exists because S is symmetric, positive definite). As $SA(\mathbf{n})$ is symmetric, so is $S^{1/2}A(\mathbf{n})S^{-1/2} = S^{-1/2}SA(\mathbf{n})S^{-1/2}$. Therefore, an orthogonal matrix Q exists such that

$$S^{1/2}A(\mathbf{n})S^{-1/2} = Q^T \Lambda Q$$

where Λ is a diagonal matrix. This can be rewritten as

$$A(\mathbf{n}) = (Q S^{1/2})^{-1} \Lambda Q S^{1/2}$$

so that A^+ and A^- can be written

$$A^+ = (Q S^{1/2})^{-1} \Lambda^+ Q S^{1/2} \quad A^- = (Q S^{1/2})^{-1} \Lambda^- Q S^{1/2}.$$

And we have

$$\begin{aligned} \mathbf{U}^T S(A^+(\mathbf{n}) - A^-(\mathbf{n})) \mathbf{U} &= \mathbf{U}^T S^{1/2} Q^T (\Lambda^+ - \Lambda^-) Q S^{1/2} \mathbf{U} \\ &= (Q S^{1/2} \mathbf{U})^T (\Lambda^+ - \Lambda^-) Q S^{1/2} \mathbf{U} \\ &\geq 0 \end{aligned} \tag{13}$$

then the matrix is positive semi-definite.

As a consequence of (13), $S(A^+(\mathbf{n}) - A^-(\mathbf{n}))$ is a symmetric matrix and $\mathcal{A}[S(A^+(\mathbf{n}) - A^-(\mathbf{n}))] = 0$. Then, the kernel of $\mathcal{S}[S(A^+(\mathbf{n}) - A^-(\mathbf{n}))]$ is $\mathcal{A}[S(A^+(\mathbf{n}) - A^-(\mathbf{n}))]$ -orthogonal.

3.1.2 Wall boundary conditions

We first prove that the matrix $S(A(\mathbf{n}) - 2A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n}))$ is positive semi-definite. For the wave system, we have

$$A^-(\mathbf{n}) = \begin{pmatrix} -\frac{c_0}{2} & \frac{1}{2} \mathbf{n}^T \\ \frac{\kappa_0}{2} \mathbf{n} & -\frac{2\rho_0}{2} \mathbf{n} \mathbf{n}^T \end{pmatrix} \quad A^+(\mathbf{n}) = \begin{pmatrix} \frac{c_0}{2} & \frac{1}{2} \mathbf{n}^T \\ \frac{\kappa_0}{2} \mathbf{n} & \frac{2\rho_0}{2} \mathbf{n} \mathbf{n}^T \end{pmatrix}$$

and P_{wall} was given in (11). This gives

$$A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n}) = \begin{pmatrix} 0 & \frac{1}{2} \mathbf{n}^T \\ 0 & -c_0 \mathbf{n} \mathbf{n}^T \end{pmatrix},$$

so that

$$S(A(\mathbf{n}) - 2A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n})) = \begin{pmatrix} 0 & -\kappa_0 \mathbf{n}^T \\ \kappa_0 \mathbf{n} & 2c_0 \mathbf{n} \mathbf{n}^T \end{pmatrix}.$$

Then for $\mathbf{U} = (p, \mathbf{u})^T \in \mathbb{R}^{n_{\text{var}}}$

$$\mathbf{U}^T S(A(\mathbf{n}) - 2A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n})) \mathbf{U} = 2c_0(\mathbf{u}^T \mathbf{n})^2 \geq 0$$

so that the matrix is positive semi-definite.

We denote by M the matrix $S(A(\mathbf{n}) - 2A^-(\mathbf{n})P_{\text{wall}}(\mathbf{n}))$. Its symmetric and skew-symmetric parts are

$$\mathcal{S}[M] = \begin{pmatrix} 0 & 0 \\ 0 & 2c_0 \mathbf{n} \mathbf{n}^T \end{pmatrix} \quad \mathcal{A}[M] = \begin{pmatrix} 0 & -\kappa_0 \mathbf{n}^T \\ \kappa_0 \mathbf{n} & 0 \end{pmatrix}.$$

The kernel of $\mathcal{S}[M]$ is the set such that $\mathbf{u}^T \mathbf{n} = 0$. If $\mathbf{U}_1, \mathbf{U}_2 \in \ker \mathcal{S}[M]$, with $\mathbf{U}_1 = (p_1, \mathbf{u}_1)^T \in \mathbb{R}^{n_{\text{var}}}$ and $\mathbf{U}_2 = (p_2, \mathbf{u}_2)^T \in \mathbb{R}^{n_{\text{var}}}$, we have

$$\mathbf{U}_1^T \mathcal{A}[M] \mathbf{U}_2 = -\kappa_0 p_1 \mathbf{u}_2^T \mathbf{n} + \kappa_0 p_2 \mathbf{u}_1^T \mathbf{n} = 0$$

because $\mathbf{u}_1^T \mathbf{n} = \mathbf{u}_2^T \mathbf{n} = 0$. Then, the kernel of $\mathcal{S}[M]$ is $\mathcal{A}[M]$ -orthogonal.

3.1.3 Rusanov scheme

With the Rusanov flux (12a), the matrix

$$M = SD(\mathbf{n}) = \frac{c_0}{2} S,$$

is symmetric, positive definite and satisfies $D(-\mathbf{n}) = D(\mathbf{n})$ then the hypothesis of Proposition 3 hold. Its symmetric and skew-symmetric parts are

$$\mathcal{S}[M] = M = \frac{1}{2} \begin{pmatrix} \kappa_0 \rho_0 c_0 & 0 \\ 0 & c_0 \mathbf{I}_d \end{pmatrix} \quad \mathcal{A}[M] = 0.$$

Then, the kernel of $\mathcal{S}[M]$ is $\mathcal{A}[M]$ -orthogonal. Moreover, the kernel of $\mathcal{S}[M]$ is the set such that $p = 0$, and $\mathbf{u} = 0$.

For $\mathbf{X}_L = (p_L, \mathbf{u}_L)^T$ and $\mathbf{X}_R = (p_R, \mathbf{u}_R)^T$ in $\mathbb{C}^{n_{var}}$ such that $\mathbf{X}_R - \mathbf{X}_L \in \ker \mathcal{S}[M]$, we have $p_L = p_R$ and $\mathbf{u}_L = \mathbf{u}_R$, and so

$$\begin{aligned} \bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R &= \kappa_0 (\bar{p}_L \mathbf{u}_R \cdot \mathbf{n} + p_R \bar{\mathbf{u}}_L \cdot \mathbf{n}) = \kappa_0 (\bar{p}_L \mathbf{u}_L \cdot \mathbf{n} + p_L \bar{\mathbf{u}}_L \cdot \mathbf{n}) \\ &= \kappa_0 (\bar{p}_L \mathbf{u}_L \cdot \mathbf{n} + \overline{\bar{p}_L \mathbf{u}_L \cdot \mathbf{n}}) \in \mathbb{R} \end{aligned}$$

so that (9) holds, and thus hypothesis of Proposition 4 hold.

3.1.4 Godunov scheme

The Godunov flux (12b) for the wave system is deduced from the more general expression of the Godunov flux

$$F(\mathbf{U}_L, \mathbf{U}_R, \mathbf{n}) = A^+(\mathbf{n})\mathbf{U}_L + A^-(\mathbf{n})\mathbf{U}_R.$$

As for the Steger-Warming boundary condition, for the Godunov scheme we are interested in the following matrix

$$M = SD(\mathbf{n}) = \frac{1}{2} S(A^+(\mathbf{n}) - A^-(\mathbf{n})).$$

Then, the hypothesis of Proposition 3 hold and the kernel of $\mathcal{S}[M]$ is $\mathcal{A}[M]$ -orthogonal (see Section 3.1.1). Moreover, in the case of the wave system (12b), we have

$$M = SD(\mathbf{n}) = \frac{1}{2} \begin{pmatrix} \rho_0 \kappa_0 & 0 \\ 0 & \mathbf{I}_d \end{pmatrix} \begin{pmatrix} c_0 & 0 \\ 0 & c_0 \mathbf{nn}^T \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \kappa_0 \rho_0 c_0 & 0 \\ 0 & c_0 \mathbf{nn}^T \end{pmatrix}$$

so that the kernel of $\mathcal{S}[M]$ is the set such that $p = 0$, and $\mathbf{u}^T \mathbf{n} = 0$.

For $\mathbf{X}_L = (p_L, \mathbf{u}_L)^T$ and $\mathbf{X}_R = (p_R, \mathbf{u}_R)^T$ in $\mathbb{C}^{n_{var}}$ such that $\mathbf{X}_R - \mathbf{X}_L \in \ker \mathcal{S}[M]$, we have $p_L = p_R$ and $\mathbf{u}_L^T \mathbf{n} = \mathbf{u}_R^T \mathbf{n}$, and so

$$\begin{aligned} \bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R &= \kappa_0 (\bar{p}_L \mathbf{u}_R \cdot \mathbf{n} + p_R \bar{\mathbf{u}}_L \cdot \mathbf{n}) = \kappa_0 (\bar{p}_L \mathbf{u}_L \cdot \mathbf{n} + p_L \bar{\mathbf{u}}_L \cdot \mathbf{n}) \\ &= \kappa_0 (\bar{p}_L \mathbf{u}_L \cdot \mathbf{n} + \overline{\bar{p}_L \mathbf{u}_L \cdot \mathbf{n}}) \in \mathbb{R} \end{aligned}$$

so that (9) holds, and thus hypothesis of Proposition 4 hold.

3.1.5 Godunov scheme with a centered discretization of the pressure gradient

For the stabilization (12c), we have

$$M = SD(\mathbf{n}) = \frac{1}{2} \begin{pmatrix} \rho_0 \kappa_0 & 0 \\ 0 & \mathbf{I}_d \end{pmatrix} \begin{pmatrix} c_0 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \kappa_0 \rho_0 c_0 & 0 \\ 0 & 0 \end{pmatrix}$$

which is positive semi-definite. Then, the hypothesis of Proposition 3 hold.

The symmetric and skew-symmetric parts of M are

$$\mathcal{S}[M] = \frac{1}{2} \begin{pmatrix} \kappa_0 \rho_0 c_0 & 0 \\ 0 & 0 \end{pmatrix} \quad \mathcal{A}[M] = 0$$

so that the kernel of $\mathcal{S}[M]$ is $\mathcal{A}[M]$ -orthogonal. Moreover, the kernel of $\mathcal{S}[M]$ is the set such that $p = 0$.

For this flux, (9) does not hold, and thus we have to prove that the first item of Proposition 4 holds. For $\mathbf{X}_L = (p_L, \mathbf{u}_L)^T$ and $\mathbf{X}_R = (p_R, \mathbf{u}_R)^T$ in $\mathbb{C}^{n_{var}}$ such that $\mathbf{X}_R - \mathbf{X}_L \in \ker \mathcal{S}[M]$, we have $p_L = p_R$ and so

$$\begin{aligned} \bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R &= \kappa_0 (\bar{p}_L \mathbf{u}_R \cdot \mathbf{n} + p_R \bar{\mathbf{u}}_L \cdot \mathbf{n}) = \kappa_0 (\bar{p}_L \mathbf{u}_R \cdot \mathbf{n} + p_L \bar{\mathbf{u}}_L \cdot \mathbf{n}) \\ &= \kappa_0 (\bar{p}_L \mathbf{u}_R \cdot \mathbf{n} + \bar{p}_L \mathbf{u}_R \cdot \mathbf{n}) + \kappa_0 p_L (\bar{\mathbf{u}}_L - \bar{\mathbf{u}}_R) \cdot \mathbf{n} \end{aligned}$$

The kernel of $\mathcal{S}[M]$ is the set such that $p = 0$. It means that for all interior faces $p_L = p_R$. Moreover, since $\ker \mathcal{S}[S(A^+(\mathbf{n}) - A^-(\mathbf{n}))]$ is the set such that $p = 0$ and $\mathbf{u}^T \mathbf{n} = 0$, we have for all neighboring cells to a Steger-Warming boundary condition $p_L = 0$. Then, by propagation, we have for all cells $p_L = 0$. Then,

$$\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R = \kappa_0 (\bar{p}_L \mathbf{u}_R \cdot \mathbf{n} + p_R \bar{\mathbf{u}}_L \cdot \mathbf{n}) = 0 \in \mathbb{R}$$

and the hypothesis of Proposition 4 hold.

3.1.6 Scheme with the LMAAP stabilization of [4]

For the stabilization (12d), we are interested in the following matrix

$$M = SD(\mathbf{n}) = \begin{pmatrix} \rho_0 \kappa_0 c_0 & \frac{\varepsilon \kappa_0}{2} \mathbf{C}_{21}^T \\ -\frac{\varepsilon \kappa_0}{2} \mathbf{C}_{21} & 0 \end{pmatrix}.$$

Then, for $\mathbf{U} = (p, \mathbf{u})^T \in \mathbb{R}^{n_{var}}$,

$$\mathbf{U}^T M \mathbf{U} = \rho_0 \kappa_0 c_0 p^2 \geq 0$$

and the hypothesis of Proposition 3 hold.

The symmetric and skew-symmetric parts of M are

$$\mathcal{S}[M] = \begin{pmatrix} \rho_0 \kappa_0 c_0 & 0 \\ 0 & 0 \end{pmatrix} \quad \mathcal{A}[M] = \begin{pmatrix} 0 & \frac{\varepsilon \kappa_0}{2} \mathbf{C}_{21}^T \\ -\frac{\varepsilon \kappa_0}{2} \mathbf{C}_{21} & 0 \end{pmatrix}$$

and the kernel of $\mathcal{S}[M]$ is the set such that $p = 0$. If $\mathbf{U}_1, \mathbf{U}_2 \in \ker \mathcal{S}[M]$ with $\mathbf{U}_1 = (p_1, \mathbf{u}_1)^T$ and $\mathbf{U}_2 = (p_2, \mathbf{u}_2)^T$ in $\mathbb{R}^{n_{var}}$, we have

$$\mathbf{U}_1^T \mathcal{A}[M] \mathbf{U}_2 = \frac{\varepsilon \kappa_0}{2} (p_1 \mathbf{C}_{21}^T \mathbf{u}_2 - p_2 \mathbf{u}_1^T \mathbf{C}_{21}) = 0$$

because $p_1 = p_2 = 0$. Then $\ker \mathcal{S}[M]$ is $\mathcal{A}[M]$ -orthogonal.

The proof that the imaginary part of $\bar{\mathbf{X}}_L^T SA(\mathbf{n}_f) \mathbf{X}_R$ is null is the same as in Section 3.1.5.

3.2 Verification of hypothesis of Proposition 1

For ensuring that hypothesis of Proposition 1 hold, it remains to check that $\mathcal{B} \in \text{range } \mathcal{A}$. The right hand side is obtained by evaluating (3) with $\mathcal{U} = 0$

$$(\mathcal{B})_i = -\frac{1}{|\Omega_i|} \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}) \frac{\mathbf{U}_b}{2} - \frac{1}{2} (A^+(\mathbf{n}) - A^-(\mathbf{n})) \mathbf{U}_b \right).$$

For any \mathcal{U} , we find

$$\mathcal{U}^T \mathcal{S} \mathcal{B} = -\sum_i \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(\mathbf{U}_i^T SA(\mathbf{n}) \frac{\mathbf{U}_b}{2} - \frac{1}{2} \mathbf{U}_i^T S(A^+(\mathbf{n}) - A^-(\mathbf{n})) \mathbf{U}_b \right).$$

Suppose that $\mathcal{U} \in \ker \mathcal{A}$. Then, as seen in the proof of Proposition 3, all the \mathbf{U}_i having a neighboring Steger-Warming boundary condition follow

$$\mathbf{U}_i^T S(A^+(\mathbf{n}) - A^-(\mathbf{n})) \mathbf{U}_i = \mathbf{U}_i^T \mathcal{S}[S(A^+(\mathbf{n}) - A^-(\mathbf{n}))] \mathbf{U}_i = 0.$$

As a consequence, \mathbf{U}_i is in the kernel of $S(A^+(\mathbf{n}) - A^-(\mathbf{n}))$, but also in the kernel of $S(A^+(\mathbf{n}) + A^-(\mathbf{n})) = A(\mathbf{n})$ (which are the same). Therefore, $\mathbf{U}_i^T SA(\mathbf{n}) = 0$, and $\mathbf{U}_i^T S(A^+(\mathbf{n}) - A^-(\mathbf{n})) = 0$. This leads to $\mathcal{U}^T \mathcal{S} \mathcal{B} = 0$ for any $\mathcal{U} \in \ker \mathcal{A}$, so that \mathcal{B} is in the \mathcal{S} -orthogonal of $\ker \mathcal{A}$, which was previously proved to match with $\text{range } \mathcal{A}$.

3.3 Long time limit of the discrete wave system

In this section, we proved that for very different discretization of the wave system (10), the long time limit of (3) exists if the system has Steger-Warming and wall boundary conditions.

We have proved in Section 3.1 that all the hypothesis of Proposition 3 and Proposition 4 hold for the finite volume scheme (3) obtained with the Rusanov flux (12a), the Godunov flux (12b), the Godunov flux with a centered discretization of the pressure gradient (12c) or the LMAAP flux (12d). Then, points [1], [2] and [3] of Hypothesis 1 hold. Moreover, in Section 3.2 we proved that $\mathcal{B} \in \text{range } \mathcal{A}$, then by applying Proposition 1, we get that all these numerical schemes (3) converge when τ tends to infinity to

$$\mathcal{U}^\infty + \mathbb{P}(\mathcal{U}_0)$$

where \mathcal{U}^∞ the unique vector of range \mathcal{A} such that $\mathcal{B} = \mathcal{A}\mathcal{U}^\infty$ and $\mathbb{P}(\mathcal{U}_0)$ is the projection on $\ker \mathcal{A}$ of the initial condition \mathcal{U}_0 .

The state \mathcal{U}^∞ can of course change depending on the numerical flux used.

3.4 Pressure of the long time limit

In this section, we wish to go further and check whether the long time limit of (3) shares some properties with the long time limit of (10) when it exists.

More precisely, we are interested in the behavior of the long time limit of the pressure. If (10) has a long time limit, and if \mathbf{U}_b is such that $p_b = 0$ along all the boundary sides on which a state is imposed, then the long time pressure is uniformly equal to 0. This property is known to also hold at the discrete level with the Godunov scheme on triangular meshes, this is a useful result in the low Mach number context, see [17, 13, 11]. On general meshes, using the result of Section 3.3, we can prove the following

Proposition 5. *We denote by $D_{12}(\mathbf{n})$ the coefficients (1, 2) to (1, d + 1) of $D(\mathbf{n})$. For a given \mathbf{u}_b , we define $\mathcal{K}_{\mathbf{u}_b}$ the set*

$$\mathcal{K}_{\mathbf{u}_b} = \left\{ \mathbf{u} \in (\mathcal{T}_h)^d, \quad \forall i \quad \left\{ \begin{array}{l} \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(\frac{\{\mathbf{u}\}_{ij} \cdot \mathbf{n}_{ij}}{2\rho_0} - D_{12}(\mathbf{n}_{ij}) \cdot (\mathbf{u}_j - \mathbf{u}_i) \right) \\ \quad + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \frac{(\mathbf{u}_i + \mathbf{u}_b) \cdot \mathbf{n}_{ij}}{2\rho_0} = 0 \\ \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \mathbf{u}_i \cdot \mathbf{n}_{ij} = 0 \end{array} \right\} \right\}. \quad (14)$$

Then for the Godunov flux with a centered discretization of the pressure gradient (12c) and for the LMAAP flux (12d), the set $\mathcal{K}_{\mathbf{u}_b}$ is nonempty if and only if the long time limit of (3) with $\mathbf{U}_b = (0, \mathbf{u}_b)$ applied to the system (10) has a uniform pressure, equal to 0.

Proof. If the long time limit of (3) with $\mathbf{U}_b = (0, \mathbf{u}_b)$ applied to the system (10) has a uniform pressure, equal to 0, then its limit is an element of $\mathcal{K}_{\mathbf{u}_b}$, and thus, this set is nonempty.

We suppose now that $\mathcal{K}_{\mathbf{u}_b}$ is not empty. We denote by $p^\infty = 0$, and by \mathbf{U}_i^∞ the vector $(p^\infty, \mathbf{u}_i^\infty)$, where \mathbf{u}_i^∞ is an element of $\mathcal{K}_{\mathbf{u}_b}$. Then

$$\begin{aligned} & \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \{\mathbf{U}^\infty\}_{ij} - D(\mathbf{n}_{ij}) (\mathbf{U}_j^\infty - \mathbf{U}_i^\infty) \right) \\ & + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i^\infty}{2} + \frac{1}{2} (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij}) P_{\text{wall}}(\mathbf{n}_{ij})) \mathbf{U}_i^\infty \right) \\ & + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i^\infty + \mathbf{u}_b}{2} - \frac{1}{2} (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) (\mathbf{u}_b - \mathbf{U}_i^\infty) \right) = 0, \end{aligned}$$

and we recall that the finite volume scheme for \mathbf{U}_i is

$$\begin{aligned} & |\Omega_i| \partial_\tau \mathbf{U}_i + \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \{\mathbf{U}\}_{ij} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_i) \right) \\ & + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i}{2} + \frac{1}{2} (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij}) P_{\text{wall}}(\mathbf{n}_{ij})) \mathbf{U}_i \right) \\ & + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i + \mathbf{u}_b}{2} - \frac{1}{2} (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) (\mathbf{u}_b - \mathbf{U}_i) \right) = 0. \end{aligned}$$

This leads to the following equation on the difference

$$\begin{aligned} |\Omega_i| \partial_\tau (\mathbf{U}_i - \mathbf{U}_i^\infty) + \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \left\{ \mathbf{U} - \mathbf{U}^\infty \right\}_{ij} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_j^\infty - \mathbf{U}_i + \mathbf{U}_i^\infty) \right) \\ + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i - \mathbf{U}_i^\infty}{2} + \frac{1}{2} (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij}) P_{\text{wall}}(\mathbf{n}_{ij})) (\mathbf{U}_i - \mathbf{U}_i^\infty) \right) \\ + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_i - \mathbf{U}_i^\infty}{2} - \frac{1}{2} (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) (\mathbf{U}_i^\infty - \mathbf{U}_i) \right) = 0. \end{aligned}$$

As in the proof of Proposition 3, we use the following identity

$$\sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \frac{A(\mathbf{n}_{ij}) (\mathbf{U}_i - \mathbf{U}_i^\infty)}{2} + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \frac{A(\mathbf{n}_{ij}) (\mathbf{U}_i - \mathbf{U}_i^\infty)}{2} + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \frac{A(\mathbf{n}_{ij}) (\mathbf{U}_i - \mathbf{U}_i^\infty)}{2} = 0,$$

for simplifying the equation on the difference as follows

$$\begin{aligned} |\Omega_i| \partial_\tau (\mathbf{U}_i - \mathbf{U}_i^\infty) + \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_j - \mathbf{U}_j^\infty}{2} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_j^\infty - \mathbf{U}_i + \mathbf{U}_i^\infty) \right) \\ + \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| \left(\frac{1}{2} (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij}) P_{\text{wall}}(\mathbf{n}_{ij})) (\mathbf{U}_i - \mathbf{U}_i^\infty) \right) \\ + \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| \left(-\frac{1}{2} (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) (\mathbf{U}_i^\infty - \mathbf{U}_i) \right) = 0. \end{aligned}$$

Now, we multiply on the left by $(\mathbf{U}_i - \mathbf{U}_i^\infty)^T S$, and sum on i for finding the following equation on the evolution of the volumic integral of the relative entropy

$$\begin{aligned} \sum_i |\Omega_i| \partial_\tau \left(\frac{(\mathbf{U}_i - \mathbf{U}_i^\infty)^T S (\mathbf{U}_i - \mathbf{U}_i^\infty)}{2} \right) \\ + \sum_i \sum_{j \in \mathcal{V}_{\text{int}}(i)} |\Gamma_{ij}| (\mathbf{U}_i - \mathbf{U}_i^\infty)^T S \left(A(\mathbf{n}_{ij}) \frac{\mathbf{U}_j - \mathbf{U}_j^\infty}{2} - D(\mathbf{n}_{ij}) (\mathbf{U}_j - \mathbf{U}_j^\infty - \mathbf{U}_i + \mathbf{U}_i^\infty) \right) \\ + \frac{1}{2} \sum_i \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| (\mathbf{U}_i - \mathbf{U}_i^\infty)^T S (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij}) P_{\text{wall}}(\mathbf{n}_{ij})) (\mathbf{U}_i - \mathbf{U}_i^\infty) \\ + \frac{1}{2} \sum_i \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| (\mathbf{U}_i - \mathbf{U}_i^\infty)^T S (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) (\mathbf{U}_i - \mathbf{U}_i^\infty) = 0. \end{aligned}$$

We now switch the sum over i and $\mathcal{V}_{\text{int}}(i)$ into a sum on the interior faces, and use the fact that $A(\mathbf{n}_{ij}) = -A(-\mathbf{n}_{ij})$, and that $D(\mathbf{n}_{ij}) = D(-\mathbf{n}_{ij})$, which gives the following equation on the evolution of the volumic integral of the relative entropy

$$\begin{aligned} \sum_i |\Omega_i| \partial_\tau \left(\frac{(\mathbf{U}_i - \mathbf{U}_i^\infty)^T S (\mathbf{U}_i - \mathbf{U}_i^\infty)}{2} \right) \\ + \sum_{f \in \mathcal{F}_{\text{int}}} |\Gamma_f| (\mathbf{U}_L - \mathbf{U}_L^\infty - \mathbf{U}_R + \mathbf{U}_R^\infty)^T S D(\mathbf{n}_f) (\mathbf{U}_L - \mathbf{U}_L^\infty - \mathbf{U}_R + \mathbf{U}_R^\infty) \\ + \frac{1}{2} \sum_i \sum_{j \in \mathcal{V}_{\text{wall}}(i)} |\Gamma_{ij}| (\mathbf{U}_i - \mathbf{U}_i^\infty)^T S (A(\mathbf{n}_{ij}) - 2A^-(\mathbf{n}_{ij}) P_{\text{wall}}(\mathbf{n}_{ij})) (\mathbf{U}_i - \mathbf{U}_i^\infty) \\ + \frac{1}{2} \sum_i \sum_{j \in \mathcal{V}_{\text{SW}}(i)} |\Gamma_{ij}| (\mathbf{U}_i - \mathbf{U}_i^\infty)^T S (A^+(\mathbf{n}_{ij}) - A^-(\mathbf{n}_{ij})) (\mathbf{U}_i - \mathbf{U}_i^\infty) = 0. \end{aligned}$$

Taking the limit when $\tau \rightarrow \infty$, the time derivative vanishes, and there remain a sum of positive terms that are 0. The first term leads to uniform pressure for stabilizations (12c) and (12d) because their symmetric part contains only the top left diagonal term. Using the equation on the Steger-Warming boundary condition gives $p = p^\infty = 0$. \square

Remark 3. As the set $\mathcal{K}_{\mathbf{u}_b}$ depends on the numerical flux, we denote it more precisely as $\mathcal{K}_{\mathbf{u}_b}^{(12c)}$ for the Godunov scheme with centered pressure. We now give a hint on how Proposition 5 can be used for proving uniformity of the pressure. Suppose that we deal with a triangular mesh. Then a known result states that with the Godunov scheme (12b), the long time limit has a uniform zero pressure see e.g. [17, 13, 11]. Moreover, this limit is such that the jump of $\mathbf{u} \cdot \mathbf{n}$ is also zero at the interior faces of the mesh and at the wall boundaries, and equal to $\mathbf{u}_b \cdot \mathbf{n}$ at the Steger-Warming boundary faces. This means that this limit lies also in $\mathcal{K}_{\mathbf{u}_b}^{(12c)}$, which is then nonempty. Using Proposition 5, the long time pressure for the Godunov scheme with centered pressure (12c) is uniformly zero.

4 Numerical experiments on the long time behavior of the wave system

The aim of this section is to perform numerical tests on the long time limit of (10) for the wave system with the different stabilizations that were described in (12) on triangular and quadrangular meshes. For this, we consider the problem of the scattering of the wave system by a cylinder of radius r_0 . The domain Ω is an annulus $[r_0, r_1] \times [0, 2\pi[$. Here, we use $r_0 = 0.5$ and $r_1 = 5.5$. We consider a wall boundary condition on the internal circle of radius r_0 and a Steger-Warming boundary condition (with $\mathbf{U}_b = (p_b, \mathbf{u}_b)^T$) on the external circle of radius r_1 . The boundary conditions are such that

$$p_b = 0, \quad \mathbf{u}_b = (1, 0)^T$$

and the initial data are uniform and set equal to

$$p_0 = 0, \quad \mathbf{u}_0 = (0, 0)^T.$$

All simulations were run with $\rho_0 = 1$, $\kappa_0 = 1$ and an Euler explicit time stepping.

This numerical test is useful because the exact steady solution can be computed. It satisfies $\nabla p = 0$, $\text{div}_{\mathbf{x}} \mathbf{u} = 0$ and $\text{curl}_{\mathbf{x}} \mathbf{u} = 0$. Indeed, we have $\partial_\tau(\text{curl}_{\mathbf{x}} \mathbf{u}) = 0$, then if at initial time $\text{curl}_{\mathbf{x}} \mathbf{u} = 0$, it is the case at any time. Using that $\mathbf{u} \cdot \mathbf{n} = 0$ on the internal circle of radius r_0 and that $\mathbf{u} \cdot \mathbf{n} = \mathbf{u}_b \cdot \mathbf{n} = (1, 0)^T \cdot \mathbf{n}$ on the external circle of radius r_1 , the exact steady state can be proven to follow

$$\begin{cases} p_{\text{exact}}(r, \theta) = 0, \\ \mathbf{u}_{\text{exact}}(r, \theta) = \frac{r_1^2}{r_1^2 - r_0^2} \begin{pmatrix} 1 - \frac{r_0^2}{r^2} \cos(2\theta) \\ -\frac{r_0^2}{r^2} \sin(2\theta) \end{pmatrix}. \end{cases} \quad (15)$$

We first aim at illustrating the convergence towards a steady state. As we saw in Proposition 1, Proposition 2 and Section 3, all the presented schemes have a long time limit.

This test is performed on triangular and quadrangular mesh. The quadrangular mesh is obtained by discretizing the annulus $[r_0, r_1] \times [0, 2\pi[$ with a resolution of $n_r = 50$ in the radial direction and $n_\theta = 160$ in the orthoradial direction. This mesh contains 8 000 quadrangular cells. The triangular mesh is obtained from a quadrangular mesh produced with $n_r = 25$ and $n_\theta = 80$ by dividing each quadrangle into two triangles. This mesh contains 4 000 triangular cells. On Figure 1, the pressure residual

$$\max_i |\Omega_i| \left| \frac{p_i^{n+1} - p_i^n}{\delta t} \right|$$

and the velocities residuals are plotted as a function of the time. Even if the time required is very different for the different stabilizations, all these schemes allow to reach a steady state with an exponential convergence rate with respect to the time, which is in agreement with the theoretical results of Section 2.2 and Section 3. On Figure 2, the influence of the CFL number on the convergence rate is studied. The residuals obtained with respect to the number of time iterations are plotted for the Godunov scheme on quadrangular mesh for various CFL number. The results are in agreement with Remark 1, the convergence rate is proportional to the time step δt .

On Figure 3, the influence of the space discretization is studied. The residuals obtained with respect to the number of time iterations are plotted for the Godunov scheme on different quadrangular meshes for a fixed time step $\delta t = 0.002$. Quadrangular meshes are obtained using a resolution n_r in the radial direction and n_θ in the orthoradial direction of $(n_r, n_\theta) = (6, 20)$, $(n_r, n_\theta) = (13, 40)$, $(n_r, n_\theta) = (25, 80)$, $(n_r, n_\theta) = (50, 160)$ and $(n_r, n_\theta) = (100, 320)$ and contain respectively 120, 520, 2000, 8000 and 32000 cells. The space step h is respectively equal to 1.14, 0.57, 0.29, 0.15 and 0.07. The numerical diffusion of the scheme allows to prove the long time limit of the numerical solution (see Section 3). We observe that the finer the mesh is (i.e. the less diffusive the scheme is), the longer the convergence to the long time limit is.

Similar results can be obtained with triangular mesh and the different stabilizations of (12). In the remainder of the paper, all the simulations were performed with a CFL number of 0.4.

We want now to study the convergence towards the exact solution (15). This is assessed in two steps: first, the variables of \mathcal{U}^∞ are shown for the different numerical schemes on the previous meshes. The pressure obtained is shown in Figure 4 for quadrangles, in Figure 5 for triangles, whereas the velocity norm obtained is shown in Figure 6 for quadrangles and in Figure 7 for triangles. At the top left of each figure the minimum and maximum value of the field on the domain are displayed. The second step for assessing the convergence towards the analytical solution

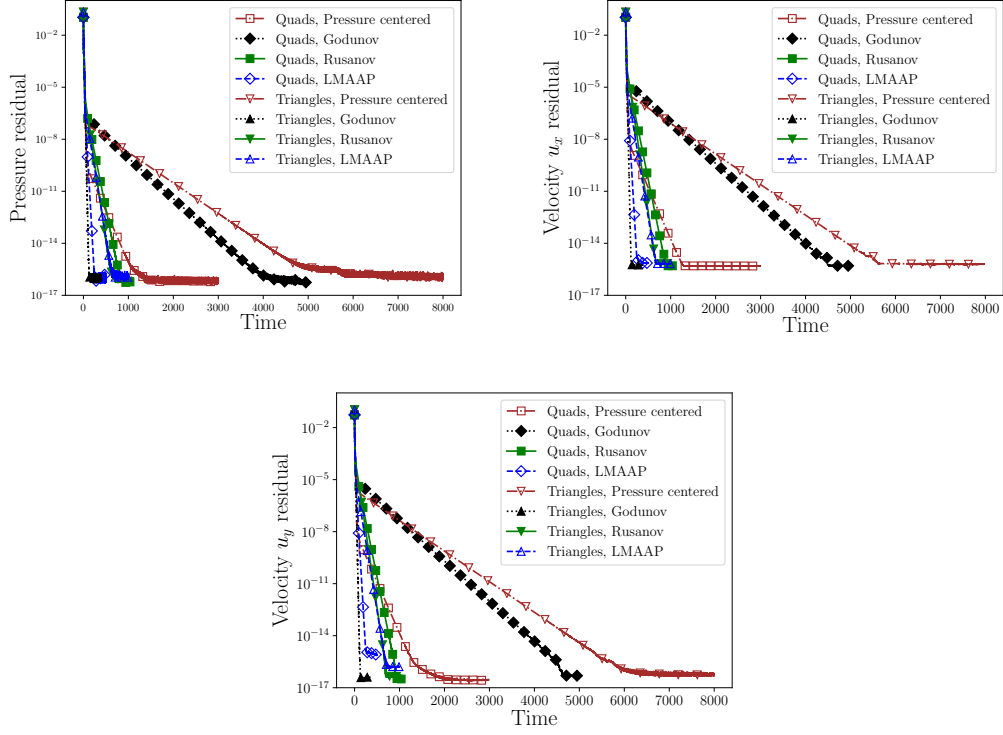


Figure 1: Residual obtained with a CFL number of 0.4 with quadrangular and triangular mesh on the pressure p and the velocities u_x and u_y as a function of the time.

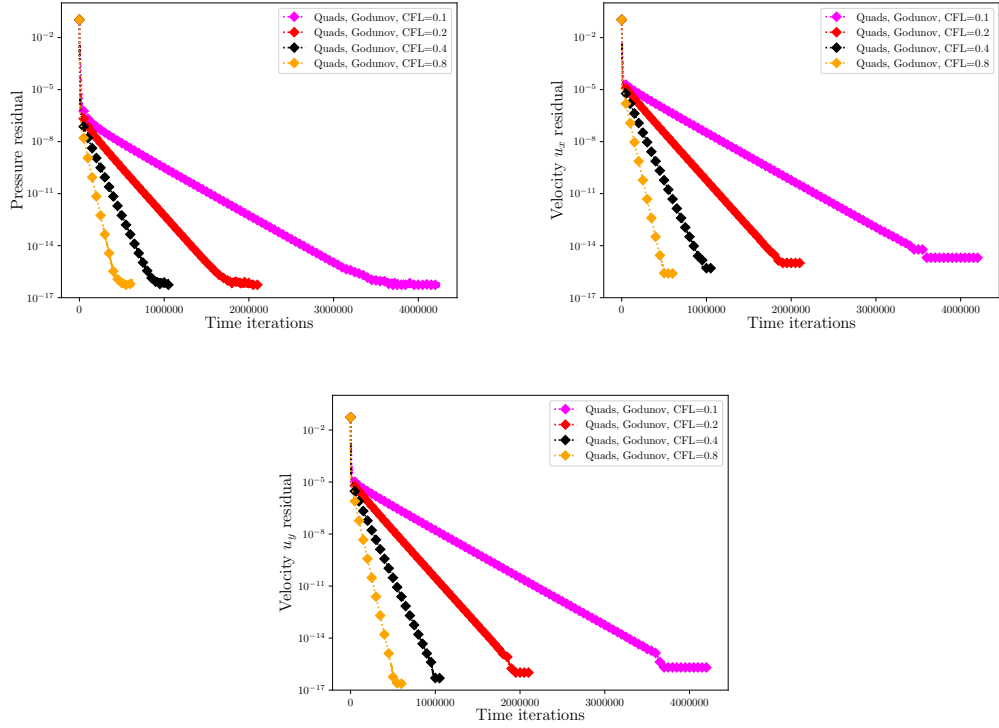


Figure 2: Residual obtained for different CFL numbers with quadrangular mesh on the pressure p and the velocities u_x and u_y as a function of the number of time iterations.

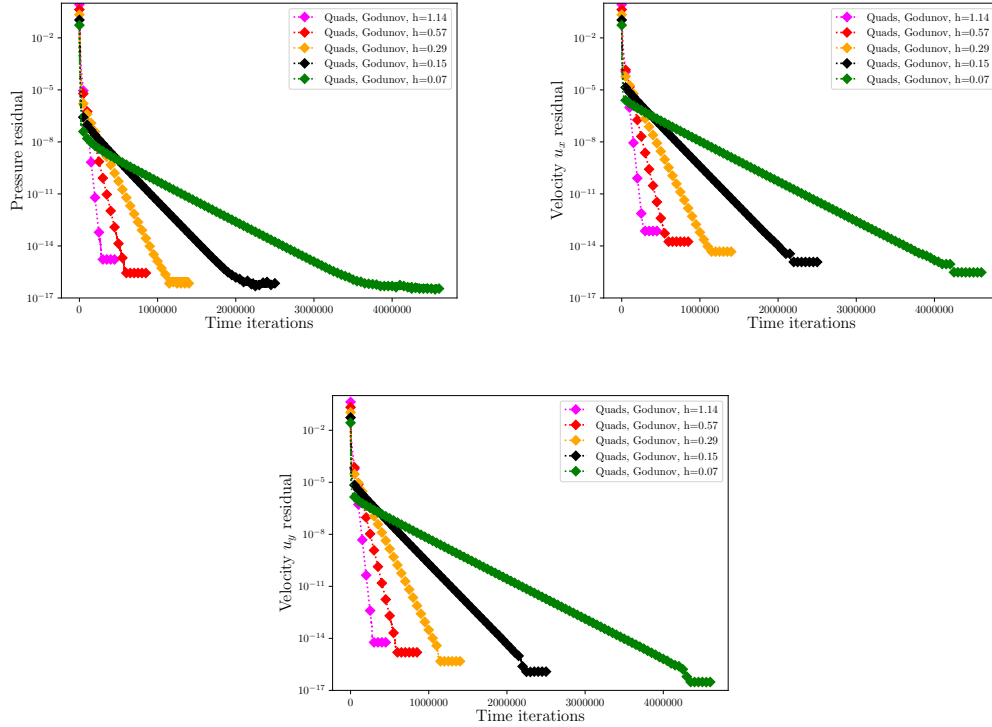


Figure 3: Residual obtained for different quadrangular meshes at fixed time step $\delta t = 0.002$ on the pressure p and the velocities u_x and u_y as a function of the number of time iterations.

consists in performing a mesh convergence study. In the triangular and quadrangular cases, five different meshes are used. Quadrangular meshes are the same as on Figure 3. Triangular meshes are obtained from the quadrangular meshes by dividing each quadrangle into two triangles and contain respectively 240, 1 040, 4 000, 16 000 and 64 000 cells. The final computational time depends on the mesh and the scheme. It is fixed such that all the residuals are converged. Results are shown in Figure 8.

From the results of Figures 4, 5, 6, 7 and 8, we remark that the stabilizations of the wave system of (12) can be divided into two families

- The ones for which the long time limit of the discrete scheme has a uniform pressure equal to 0, namely the pressure centered stabilization (12c), the LMAAP stabilization (12d), and also the Godunov scheme (12b) on triangles. These schemes are such that the velocity of the long time limit converges towards (15). For these schemes, the pressure of the long time limit observed in Figure 4, Figure 5 and Figure 8 is uniform and equal to 0. Moreover, the velocity shown in Figure 6 and Figure 7 is close of the exact solution. Figure 8 shows that the velocity is converging towards the exact solution with a rate of 1 for the pressure centered stabilization and the Godunov scheme on triangles, and $1/2$ for the LMAAP stabilization.
- The ones for which the long time limit does not have a uniform pressure, namely on one hand the Rusanov scheme and on the other hand the Godunov scheme on quadrangular meshes. For these schemes, the pressure of the long time limit found in Figure 4 and Figure 5 is not 0, but in Figure 8 a convergence rate of 1 of the pressure towards 0 is observed. Figure 6 and Figure 7 show that the velocity is far from the exact solution, and Figure 8 shows that the velocity does not converge towards the exact solution.

5 Conclusion

In this article, the long time limit of finite volume discretizations of hyperbolic systems was addressed. We proved in Section 2 that provided the numerical flux is sufficiently dissipative, the long time limit always exist. We proved also that in general, the ω -limit set of the whole space,

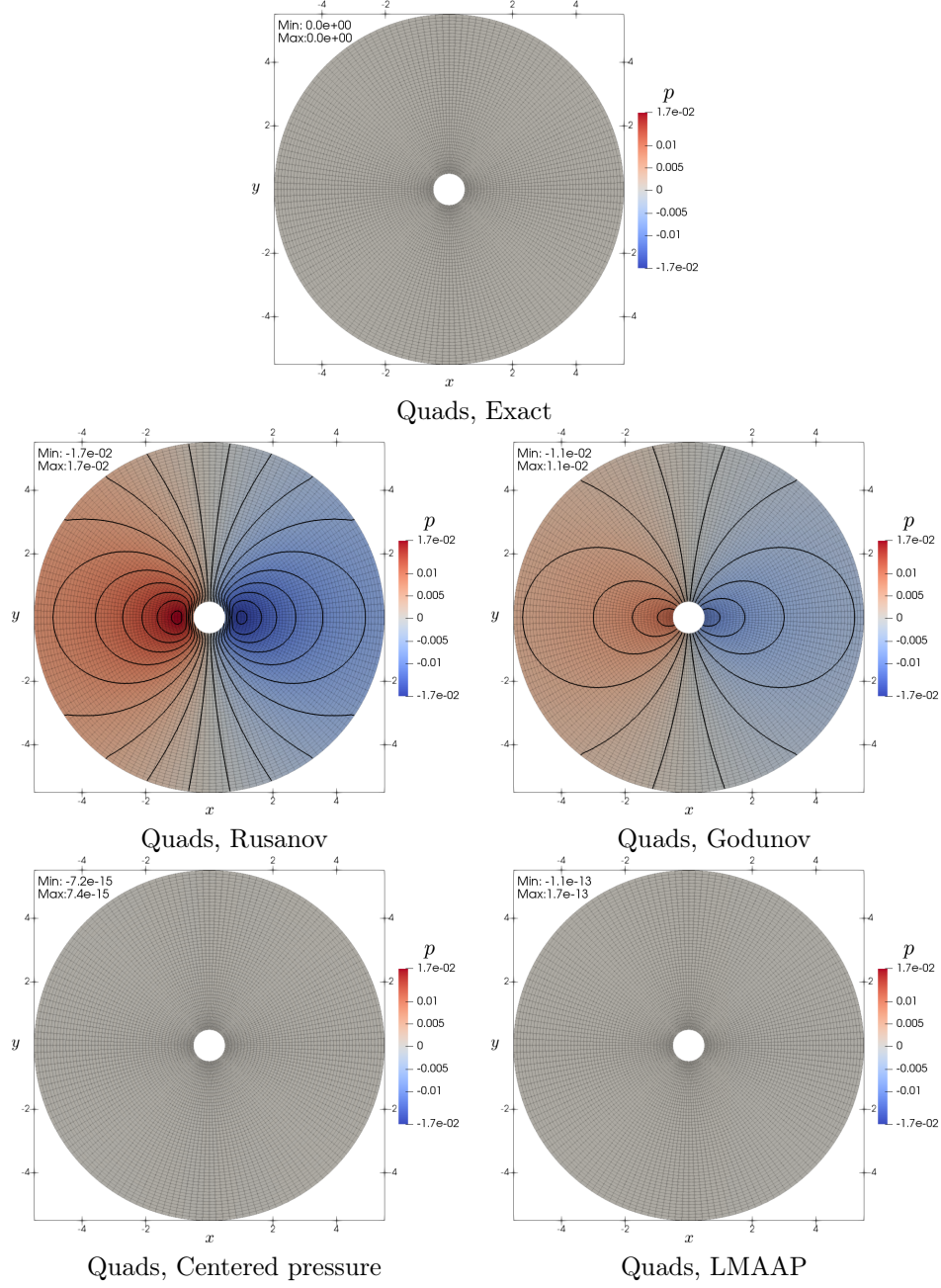


Figure 4: Pressure field of the long time limit state obtained with the different stabilizations on quadrangular mesh. Twenty equally reparted contours between -0.017 and 0.017 are plotted.

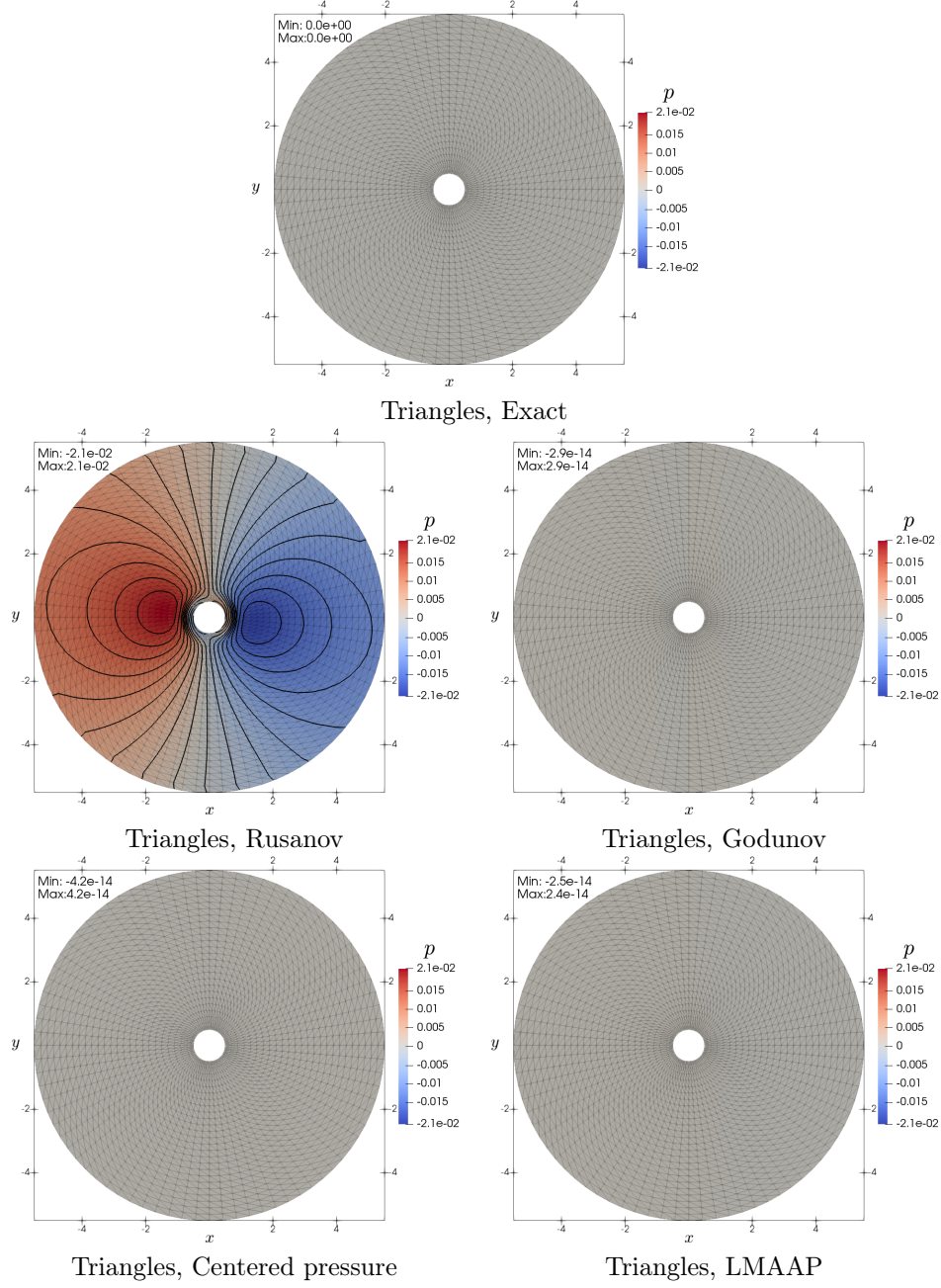


Figure 5: Pressure field of the long time limit state obtained with the different stabilizations on triangular mesh. Twenty equally reparted contours between -0.021 and 0.021 are plotted.

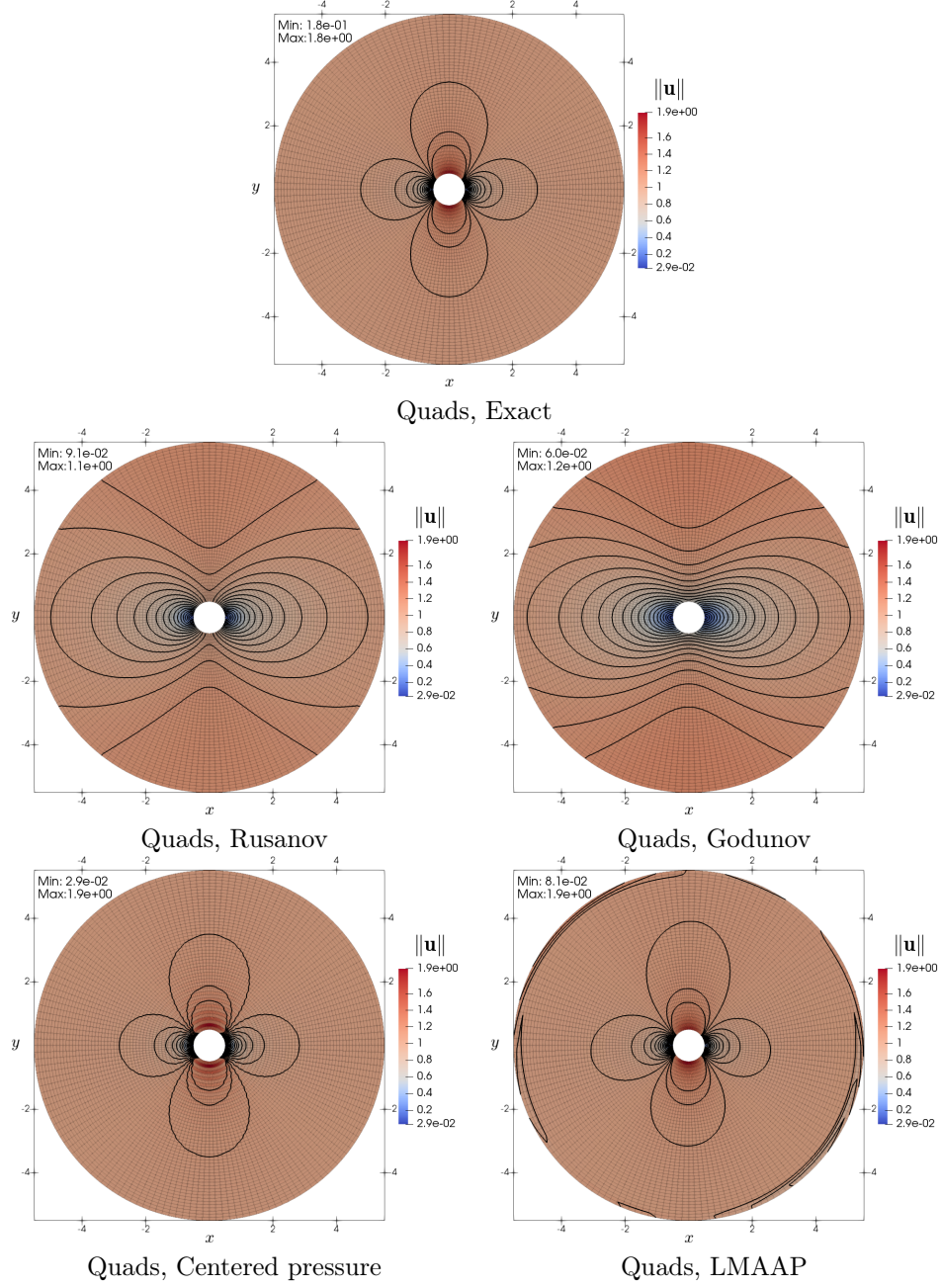


Figure 6: Velocity norm of the long time limit state obtained with the different stabilizations on quadrangular mesh. Twenty equally reparted contours between 0.1 and 1.14 are plotted.

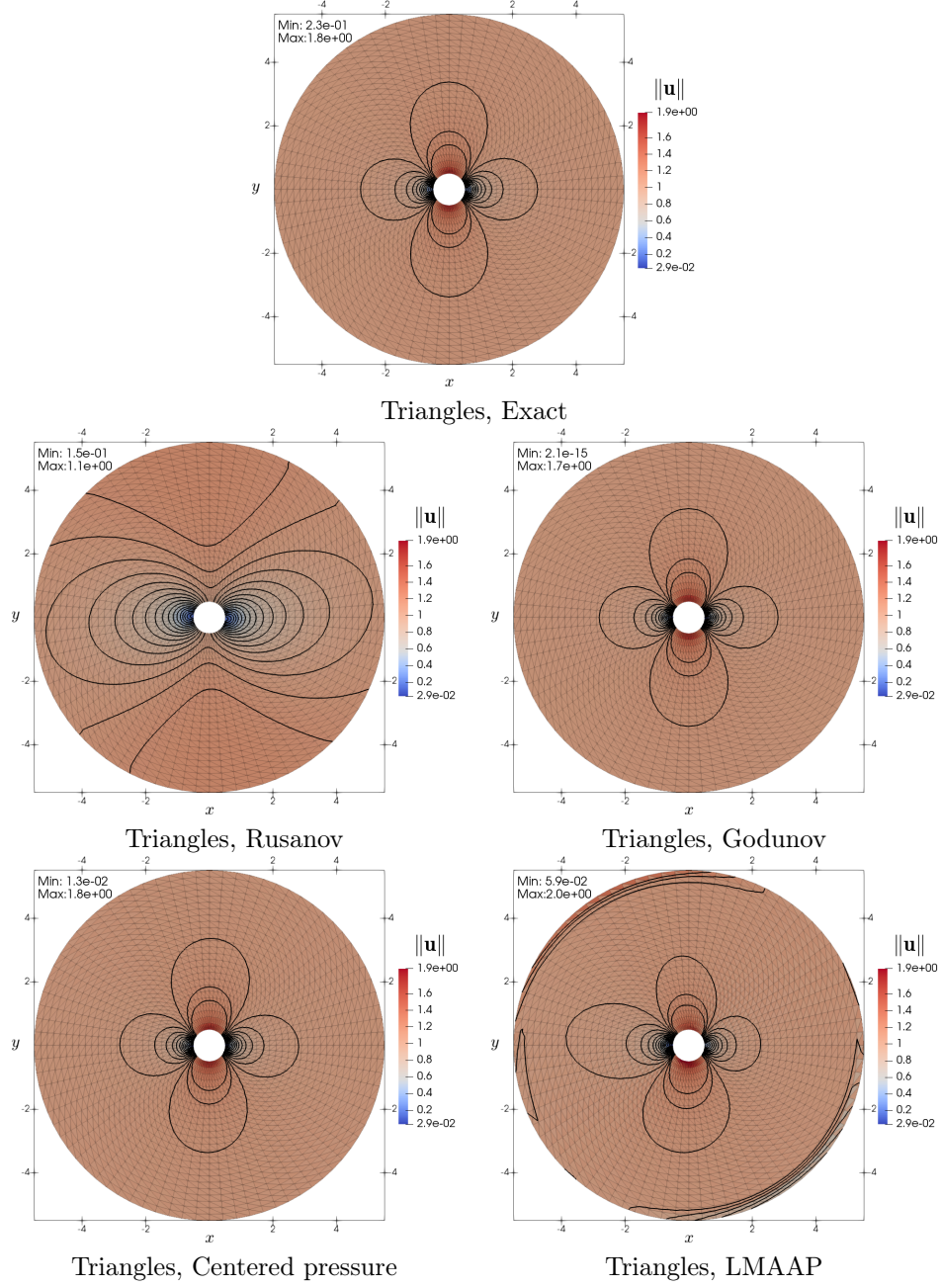


Figure 7: Velocity norm of the long time limit state obtained with the different stabilizations on triangular mesh. Twenty equally reparted contours between 0.1 and 1.14 are plotted.

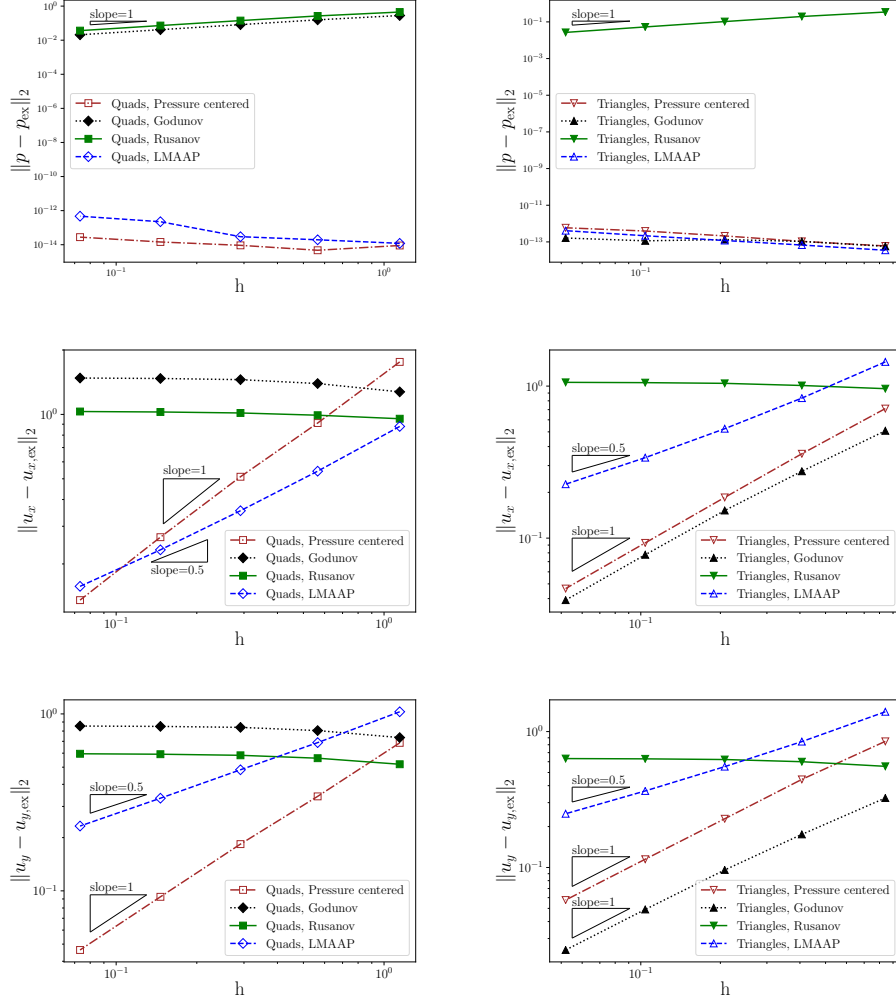


Figure 8: L_2 norm of the error between the exact solution and the long time limit state obtained with the different stabilizations. Results are shown for the pressure p (top) and the velocities u_x (middle) and u_y (bottom) on quadrangular (left) and triangular meshes (right). Five different meshes are used in both cases.

namely the set of all the possible long time limit for all the possible initial conditions, is an affine space.

Then the results of Section 2 were applied to the wave system in Section 3, and numerically tested in Section 4. In these two sections, the uniformity of the pressure of the long time limit was discussed.

Last, we would like to point out an interesting link between the behavior of the numerical fluxes at low Mach number and the long time limit of the matching stabilization of the wave system. In [4, Proposition 2.3], the double scale asymptotic expansion of classical numerical flux and low Mach number fixes was performed, leading to a stabilization of the wave system. It is worth noting the exact matching between the low Mach number fluxes that are low Mach number accurate and the uniformity of the pressure of the long time limit for the matching stabilization of the wave system:

- The numerical fluxes that are low Mach number accurate match with a discretization of the wave system for which the long time limit ensures a uniform pressure. In this case, the velocity of long time limit converges towards the analytical solution, namely a harmonic velocity.
- The numerical fluxes that are not low Mach accurate match with a discretization of the wave system for which the long time limit does not ensure a uniform pressure. In case, the velocity of long time limit does not converge towards the analytical solution.

Especially, a lot of recently proposed low Mach fixes [3, 10, 16, 15] match with the stabilization (12c) of the wave system when the double time scale asymptotic expansion is performed. Remark 3 proves that for all these schemes, the long time limit pressure of the matching wave stabilization on triangular meshes is uniform. A more thorough study of the set $\mathcal{K}_{\mathbf{u}_b}$ defined in (14) is necessary for concluding on the other low Mach number fixes.

Acknowledgment: Experiments presented in this article were carried out using the PlaFRIM experimental testbed, supported by Inria, CNRS (LABRI and IMB), Université de Bordeaux, Bordeaux INP and Conseil Régional d'Aquitaine (see <https://www.plafrim.fr/>).

References

- [1] M. Ainsworth, P. Monk, and W. Muniz. Dispersive and dissipative properties of discontinuous Galerkin finite element methods for the second-order wave equation. *Journal of Scientific Computing*, 27(1):5–40, Jun 2006.
- [2] A. Arnold, J. A. Carrillo, L. Desvillettes, J. Dolbeault, A. Jüngel, C. Lederman, P. A. Markowich, G. Toscani, and C. Villani. Entropies and equilibria of many-particle systems: an essay on recent research. In *Nonlinear Differential Equation Models*, pages 35–43. Springer, Vienna, 2004.
- [3] P. Birken and A. Meister. Stability of preconditioned finite volume schemes at low Mach numbers. *BIT Numerical Mathematics*, 45(3):463–480, 2005.
- [4] P. Bruel, S. Delmas, J. Jung, and V. Perrier. A low Mach correction able to deal with low Mach acoustics. *Journal of Computational Physics*, 378:723–759, 2019.
- [5] C. Cancès, C. Chainais-Hillairet, M. Herda, and S. Krell. Large time behavior of nonlinear finite volume schemes for convection-diffusion equations. *SIAM Journal on Numerical Analysis*, 58(5):2544–2571, 2020.
- [6] C. Chainais-Hillairet. Entropy method and asymptotic behaviours of finite volume schemes. In *Finite Volumes for Complex Applications VII-Methods and Theoretical Aspects*, volume 77, pages 17–35. Springer Proceedings in Mathematics & Statistics, Springer Cham, 2014.
- [7] C. Chainais-Hillairet and F. Filbet. Asymptotic behaviour of a finite-volume scheme for the transient drift-diffusion model. *IMA journal of numerical analysis*, 27(4):689–716, 2007.
- [8] C. Chainais-Hillairet and M. Herda. Large-time behaviour of a family of finite volume schemes for boundary-driven convection–diffusion equations. *IMA Journal of Numerical Analysis*, 40(4):2473–2504, 2020.

- [9] C. Chainais-Hillairet, A. Jüngel, and S. Schuchnigg. Entropy-dissipative discretization of non-linear diffusion equations and discrete Beckner inequalities. *ESAIM: Mathematical Modelling and Numerical Analysis*, 50(1):135–162, 2016.
- [10] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number. *Journal of Computational Physics*, 4(229):978–1016, 2010.
- [11] S. Dellacherie, P. Omnes, and F. Rieper. The influence of cell geometry on the Godunov scheme applied to the linear wave equation. *Journal of Computational Physics*, 229(14):5315–5338, 2010.
- [12] F. Filbet. An asymptotically stable scheme for diffusive coagulation-fragmentation models. *Communications in Mathematical Sciences*, 6(2):257–280, 2008.
- [13] H. Guillard. On the behavior of upwind schemes in the low Mach number limit. IV: P0 approximation on triangular and tetrahedral cells. *Computers & Fluids*, 38(10):1969–1972, 2009.
- [14] E. Hainry. Computing omega-limit sets in linear dynamical systems. In *International Conference on Unconventional Computation*, pages 83–95. Springer, 2008.
- [15] E. Labourasse. A low-Mach correction for multi-dimensional finite volume shock capturing schemes with application in lagrangian frame. *Computers & Fluids*, 179:372–393, 2019.
- [16] F. Rieper. A low-Mach number fix for Roe’s approximate Riemann solver. *Journal of Computational Physics*, 230(13):5263–5287, 2011.
- [17] F. Rieper and G. Bader. The influence of cell geometry on the accuracy of upwind schemes in the low Mach number regime. *Journal of Computational Physics*, 228(8):2918–2933, 2009.