Second-order entropy satisfying BGK-FVS schemes for incompressible Navier-Stokes equations

François Bouchut, Yann Jobic, Roberto Natalini,
René Occelli & Vincent Pavan

Second-order entropy satisfying BGK-FVS schemes for incompressible Navier-Stokes equations

François Bouchut [1]
Yann Jobic [2]
Roberto Natalini [3]
René Occelli [4]
Vincent Pavan [5]

[1] Université Paris-Est, Laboratoire d'Analyse et de Mathématiques Appliquées (UMR 8050),
CNRS, UPEM, UPEC, F-77454, Marne-la-Vallée, France
E-mail address: francois.bouchut@u-pem.fr
[2] IUSTI, 5 rue Enrico Fermi, 13453 Marseille Cedex 13, France
E-mail address: yann.jobic@univ-amu.fr
[3] CNR, Istituto per le Applicazioni del Calcolo "M. Picone", via dei Taurini 19, I-00185,
Roma, Italy
E-mail address: roberto.natalini@cnr.it
[4] IUSTI, 5 rue Enrico Fermi, 13453 Marseille Cedex 13, France
E-mail address: rene.occelli@univ-amu.fr
[5] IUSTI, 5 rue Enrico Fermi, 13453 Marseille Cedex 13, France
E-mail address: vincent.pavan@univ-amu.fr.

Abstract. Kinetic BGK numerical schemes for the approximation of incompressible Navier-Stokes equations are derived via classical discrete velocity vector BGK approximations, but applied to an inviscid compressible gas dynamics system with small Mach number parameter, according to the approach of Carfora and Natalini (2008). As the Mach number, the grid size and the timestep tend to zero, the low Mach number limit and the time-space convergence of the scheme are achieved simultaneously, and the numerical viscosity tends to the physical viscosity of the Navier-Stokes system. The method is analyzed and formulated as an explicit finite volume/difference flux vector splitting (FVS) scheme over a Cartesian mesh. It is close in spirit to lattice Boltzmann schemes, but it has several advantages. The first is that the scheme is expressed only in terms of momentum and mass compressible variables. It is therefore very easy to implement, and several types of boundary conditions are straightforward to apply. The second advantage is that the scheme satisfies a discrete entropy inequality, under a CFL condition of parabolic type and a subcharacteristic stability condition involving a cell Reynolds number that ensures that diffusion dominates advection at the level of the grid size. This ensures the robustness of the method, with explicit uniform bounds on the approximate solution. Moreover the scheme is proved to be second-order accurate in space if the parameters are well chosen, this is the case in particular for the Lax-Friedrichs scheme with Mach number proportional to the grid size. The scheme falls then into the class of artificial compressibility methods, the novelty being its exceptionally good theoretical properties. We show the efficiency of the method in terms of accuracy and robustness on a variety of classical two-dimensional benchmark tests. The method is finally applied in three dimensions to compute the permeability of a porous medium defined by a complex idealized Kelvin-like cell. Relations between our scheme and compressible low Mach number schemes are discussed.

2010 Mathematics Subject Classification. 65M12, 65X68.

Keywords. incompressible Navier-Stokes equations, vector BGK schemes, flux vector splitting, low Mach number limit, discrete entropy inequality, cell Reynolds number, lattice Boltzmann schemes .

1. Introduction

1.1. Navier-Stokes equations and high performance computing

The development of hardware components like GPU, CPU clusters, infiniband connections etc. . . has renewed all the paradigms of scientific computing. Challenges such as real time and/or real scale simulations, especially on industrial and natural systems, require the ability of using the computing power at their maximum level. High Performance Computing (HPC) is thus a current important topic for research and development, especially for computational physics which is known to be a greedy client. HPC involves of course purely computational aspects, like parallel or processor optimizations. But it also requires selecting relevant numerical methods, especially as regards to their simplicity, robustness and their capacity to run efficiently on parallel machines. About fifteen years ago, the achievement of the Lattice Boltzmann Methods (LBM) has brought important progresses in that direction. Such methods are both general, explicit (thus tailor made for parallel structures) and very simple to program. They have reached a high degree of maturity and are now used pretty heavily in aerodynamic flow computations. However, several issues arising in the use of LBM remain unsolved. Such an issue is the limitations that comes in applying boundary conditions, that may lead to difficulties for computations in confined structures [84]. Another issue is the question of stability and the optimal choice of the parameters in order to minimize numerical diffusion while maintaining the robustness. Additionally, the possibility to use GPU for scientific computations have led people to try decreasing the RAM, something in which the LBM are practically limited, hence letting space for new algorithms [6]. In the present paper we propose an approach that aims to make progresses on these points.

1.2. The kinetic BGK method

The kinetic method for the simulation of compressible flows has been introduced in the eighties in [73, 74, 40, 41, 59, 70, 42, 2]. The schemes were derived originally by imitation of the hydrodynamic limit from the Boltzmann equation to the compressible Euler equations [9, 10], and are called either *kinetic schemes, Boltzmann schemes*, or *BGK schemes*. A strong property of these schemes is that they are naturally compatible with the decrease of an entropy, this property ensuring robustness. The mathematical understanding of such approximations has been developed in parallel to numerical methods, starting from scalar equations in [23, 24, 71]. The general structure of kinetic schemes that enables entropy compatibility for systems of conservation laws has later been established in [15], and their equivalence with flux vector splitting schemes was proved in [16]. A sharp direct analysis of FVS schemes in the multi-dimensional setting was then provided in [19]. More general *relaxation approximations* for systems were proposed in [31]. These enable to avoid the problem of excessive diffusion on contact discontinuities that are inherent to kinetic methods [48, 44]. Relaxation approximations that well resolve contact discontinuities have been then developed, making the link to approximate Riemann solvers [18, 17, 21]. The kinetic method is nevertheless useful when approximate Riemann solvers are not available, which is the case for example in large systems as in [7]. The use of kinetic schemes with a finite number of velocities (velocities on a lattice in several dimensions), or *lattice Boltzmann schemes* have been considered for long, e.g. [8, 61, 49, 82, 42]. The application of the kinetic method to diffusive problems was developed later [46, 51, 52, 20, 4, 38].

For incompressible Navier-Stokes equations, the kinetic method with a finite number of velocities (or lattice Boltzmann method) has been intensively used [81]. In this case the issue of contact discontinuities no longer exists. The difficulty however is that in comparison to compressible flows, the hydrodynamic limit is superimposed to the *low Mach number limit* (also called *acoustic limit*) [11]. It raises then a fundamental difficulty to analyze the accuracy of the scheme [56, 45]. Tools for improving the accuracy are necessary, and include in particular overrelaxed time integration [43], that

unfortunately destroys the entropy compatibility. The use of special velocity lattices [32] and particular Maxwellians that are not entropy compatible also contribute to the loss of robustness. Indeed, as was shown in [86], usual LBM (as opposed to entropic LBM, see for instance [53, 83]) are not able to manage with any entropy theorem, and the study of stability must be performed with additional difficulties.

In this paper we propose to come back to the original kinetic BGK approach applied to incompressible Navier-Stokes equations, with the leading idea to keep the entropy compatibility property. By analyzing the simultaneous hydrodynamic and low Mach number limits, we establish the properties for having consistency, and second-order accuracy. Our work takes its roots in [29] where finitely many vector Maxwellians are used, in accordance with [16]. Even if we start from kinetic considerations, we write the scheme under the form of a simple explicit finite volume/difference flux vector splitting scheme over a Cartesian mesh and written on the macroscopic moments themselves, thus finally avoiding the kinetic aspects. This writing enables to use the sharp stability analysis of [19] for FVS schemes. Moreover it enables to analyze the accuracy more simply than at the kinetic level. We end up with a BGK-FVS method that is second-order accurate without any special choice of velocities and with a standard forward Euler time stepping, indeed it is just the Lax-Friedrichs scheme applied to a scaled compressible isentropic system. It satisfies a discrete entropy inequality under a CFL condition of parabolic type involving only the viscosity, and a subcharacteristic condition that says mainly that a cell Reynolds number associated to the grid size must be less than one. This means that diffusion dominates advection at the level of the grid size.

The important properties of the scheme are its robustness with explicit uniform bounds on the approximate solution, second-order accuracy, generality and easiness of implementation, and its low memory consumption.

The paper is organized as follows. In Section 2 we recall hydrodynamic limits from kinetic models, and in particular the kinetic method for hyperbolic systems of conservation laws and its equivalence with flux vector splitting schemes. In Section 3 we explain the parabolic scaling and the low Mach number limit. In Section 4 we analyze BGK-FVS schemes in the context of incompressible Navier-Stokes equations. The practical settings are described in Section 5, and numerical tests are performed in Section 6. A discussion is proposed in Section 7, and a conclusion is provided in Section 8. A list of the main parameters used in the paper and their significance is provided in A.

2. **Hydrodynamic limit from kinetic models**

2.1. **From Boltzmann to incompressible Navier-Stokes**

Whereas the Boltzmann equation was basically established to model gases, hence compressible fluids, it has been observed for a long time [9] that formal expansions on the latter could lead to various formulations of the Euler or Navier Stokes Equations (NSE in the sequel), including compressible or incompressible NSE (INSE in the sequel), or the Stokes approximation for instance, see [76]. What was first essentially a theoretical result, bridging microscopic to macroscopic point of views, has then turned out to be a model for developing numerical schemes for INSE, as well as many other transport equations of hyperbolic or/and parabolic type.

To recall the scaling that enables to derive INSE from the Boltzmann equation, let us start from the rescaled Boltzmann equation

$$\partial_t f + \frac{\mathbf{v}}{\epsilon} \cdot \nabla_{\mathbf{x}} f = \frac{1}{\epsilon^{q+1}} Q(f, f), \qquad (2.1)$$

where $q \geq 1$ is an integer, $f(t, \mathbf{x}, \mathbf{v})$ is the particle distribution function, t is the time, $\mathbf{x} \in \mathbb{R}^D$ is the space position, $\mathbf{v} \in \mathbb{R}^D$ is the velocity, $Q(f, f)$ is the collision term, and ϵ is a small parameter (related

to the so called Knundsen number). We look for asymptotic solutions to this equation as

$$f(t, \mathbf{x}, \mathbf{v}) = \mathcal{M}_{\overline{\rho}, \mathbf{0}, \overline{T}}(\mathbf{v}) \left(1 + \epsilon^r g(t, \mathbf{x}, \mathbf{v}) + \cdots \right), \tag{2.2}$$

with $r \geq 1$ an integer to be specified and where $\mathcal{M}_{\overline{\rho}, \mathbf{0}, \overline{T}}(\mathbf{v})$ is the *absolute* Maxwellian equilibrium, defined by

$$\mathcal{M}_{\overline{\rho}, \mathbf{0}, \overline{T}}(\mathbf{v}) = \frac{\overline{\rho}}{\left(2\pi k_B \overline{T}/m\right)^{D/2}} \exp\left(-\frac{m|\mathbf{v}|^2}{2 k_B \overline{T}}\right). \tag{2.3}$$

It satisfies $Q(\mathcal{M}, \mathcal{M})(\mathbf{v}) = 0$, and its moments $\overline{\rho}, \mathbf{0}, \frac{D}{2}\overline{\rho}k_B\overline{T}/m$ against $(1, \mathbf{v}, |\mathbf{v}|^2/2)$ do not depend on space or time. Introducing then the notations

$$\left(\overline{\rho} + \epsilon^r \widetilde{\rho}(t, \mathbf{x}), \epsilon^r \overline{\rho} \mathbf{u}(t, \mathbf{x}), \frac{D}{2}\overline{\rho}\frac{k_B \overline{T}}{m} + \epsilon^r \frac{D}{2}\overline{\rho}\frac{k_B}{m}\widetilde{T}(t, \mathbf{x})\right) := \int \left(1, \mathbf{v}, \frac{|\mathbf{v}|^2}{2}\right) f(t, \mathbf{x}, \mathbf{v}) d\mathbf{v}, \tag{2.4}$$

depending on the pair (q, r) of integers, the moments $\widetilde{\rho}(t, \mathbf{x}), \mathbf{u}(t, \mathbf{x}), \widetilde{T}(t, \mathbf{x})$ satisfy at first order in ϵ various approximations of the INSE as follows:

(1) For any pair (q, r) there holds both incompressibility and the Boussinesq approximation, that is

$$\nabla_{\mathbf{x}} \cdot \mathbf{u} = 0, \tag{2.5}$$

$$\nabla_{\mathbf{x}} \left(\frac{\widetilde{\rho}}{\overline{\rho}} + \frac{\widetilde{T}}{\overline{\overline{T}}}\right) = \mathbf{0}. \tag{2.6}$$

(2) Depending on (q, r) we have the following various momentum and temperature equations

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (\mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) = \nu \Delta_{\mathbf{x}} \mathbf{u}, \quad \partial_t \widetilde{T} + \nabla_{\mathbf{x}} \cdot \left(\widetilde{T}\mathbf{u}\right) = \kappa \Delta_{\mathbf{x}} \widetilde{T}, \quad q = 1, r = 1, \tag{2.7}$$

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (p\mathbf{I}) = \nu \Delta_{\mathbf{x}} \mathbf{u}, \quad \partial_t \widetilde{T} = \kappa \Delta_{\mathbf{x}} \widetilde{T}, \quad q = 1, r > 1, \tag{2.8}$$

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (\mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) = \mathbf{0}, \quad \partial_t \widetilde{T} + \nabla_{\mathbf{x}} \cdot \left(\widetilde{T}\mathbf{u}\right) = 0, \quad q > 1, r = 1, \tag{2.9}$$

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (p\mathbf{I}) = \mathbf{0}, \quad \partial_t \widetilde{T} = 0, \quad q > 1, r > 1. \tag{2.10}$$

As a consequence of these calculations, one realizes that it is possible to get solutions to the previous incompressible macroscopic equations from the resolution of the rescaled Boltzmann Equation. In particular, the case $q = 1$, $r = 1$ enables to recover INSE. In practice it is however illusory to do that directly since the Boltzmann equation is much more difficult to solve than INSE. Moreover, the viscosity coefficients ν, κ are not arbitrary since they derive from the specific shape of the collision operator $Q(f, f)$, which fails to model liquids.

2.2. **Kinetic relaxation to compressible models**

The kinetic relaxation method has been developed a lot to approximate a system of conservation laws

$$\frac{\partial \mathbf{W}}{\partial t} + \sum_{j=1}^{D} \frac{\partial}{\partial x_j} \mathbf{F}_j(\mathbf{W}) = \mathbf{0}, \tag{2.11}$$

with unknown $\mathbf{W} := (W_1, \cdots, W_p)$. Compressible flow models enter in this class, since they do not have the steady equation $\nabla_{\mathbf{x}} \cdot \mathbf{u} = 0$. The kinetic vector BGK models described in [15] for approximating (2.11) are designed as follows. They generalize the Boltzmann equation to a kinetic vector

4

equation with unknown $\mathbf{f}(t, \mathbf{x}, \xi) \in \mathbb{R}^p$,

$$\partial_t \mathbf{f} + \mathbf{v}(\xi) \cdot \nabla_\mathbf{x} \mathbf{f} = \frac{1}{\epsilon} (\mathbf{M}[\mathbf{W}] - \mathbf{f}), \tag{2.12}$$

$$\mathbf{W}(t, \mathbf{x}) = \int_\Xi \mathbf{f}(t, \mathbf{x}, \xi) d\xi. \tag{2.13}$$

In (2.12), (2.13), $\xi \in \Xi$ is a kinetic variable that belongs to some space Ξ,

$$\Xi \quad \text{is a measure space with measure } d\xi. \tag{2.14}$$

It can be chosen arbitrarily, and main examples are $\Xi = \mathbb{R}^D$ (the case of the original Boltzmann equation) or $\Xi = \mathbb{R}$ with the Lebesgue measure, or Ξ a finite set with the counting measure. The function $\mathbf{v}(\xi) \in \mathbb{R}^D$ is the associated velocity, and $\mathbf{M}$ is the Maxwellian equilibrium, which to a set of moments $\mathbf{W}$ associates a distribution $\mathbf{M}[\mathbf{W}](\xi)$ with values in $\mathbb{R}^p$,

$$\mathbf{W} \to \mathbf{M}[\mathbf{W}](\xi). \tag{2.15}$$

In the case of the Boltzmann equation one has $\Xi = \mathbb{R}^D$, $\mathbf{v}(\xi) = \xi$, $\mathbf{M}[\mathbf{W}](\xi) = (1, \xi, |\xi|^2/2) \mathcal{M}(\mathbf{W}, \xi)$ with $\mathcal{M}$ defined by (2.3), $\mathbf{f} = (1, \xi, |\xi|^2/2)f$, thus the BGK equation (2.12) is obtained by the multiplication of the Boltzmann equation (2.1) by the vector $(1, \xi, |\xi|^2/2)$, owing that we replaced the Boltzmann collision term $Q(f, f)$ by the BGK relaxation term in the right-hand side, which is much easier to handle since it involves only the moment $\mathbf{W}$ of $\mathbf{f}$ defined as (2.13).

For the kinetic BGK model (2.12), (2.13) to be consistent with the system (2.11) in the limit $\epsilon \to 0$, the Maxwellian has to satisfy the moment relations

$$\int \mathbf{M}[\mathbf{W}](\xi) \, d\xi = \mathbf{W}, \qquad \int \mathbf{v}_j(\xi) \, \mathbf{M}[\mathbf{W}](\xi) \, d\xi = \mathbf{F}_j(\mathbf{W}), \tag{2.16}$$

for all $\mathbf{W}$. Indeed, by integration in ξ, the solution to (2.12) satisfies

$$\partial_t \int \mathbf{f}(t, \mathbf{x}, \xi) \, d\xi + \sum_{j=1}^D \frac{\partial}{\partial x_j} \int \mathbf{v}_j(\xi) \, \mathbf{f}(t, \mathbf{x}, \xi) \, d\xi = \mathbf{0}, \tag{2.17}$$

because according to the first relation of (2.16), the right-hand side of (2.12) has vanishing integral with respect to ξ. Then as $\epsilon \to 0$, the equation (2.12) forces $\mathbf{f}$ to be Maxwellian, $\mathbf{f}(t, \mathbf{x}, \xi) \simeq \mathbf{M}[\mathbf{W}(t, \mathbf{x})](\xi)$, thus (2.17) reduces to the system (2.11) because of the moment relations (2.16).

Introducing a new kinetic variable ξ may look to readers unused to kinetic approximations as adding difficulties. This fact is largely counterbalanced by the fact that we have a linear transport term $\mathbf{v} \cdot \nabla_\mathbf{x} \mathbf{f}$ that leads to a very simple solution to the BGK equation (2.12). The larger space of unknowns can be limited by taking for Ξ a relatively small finite set, leading to finitely many velocities $\mathbf{v}(\xi)$, or by writing only the moments (2.17) of the kinetic equation, that eliminate the kinetic variable.

The main difficulty in choosing the kinetic model is to design the Maxwellian such that the kinetic equation preserves good properties of the original system (2.11). In particular, it is desirable that the kinetic model possesses a *kinetic entropy*.

2.3. Entropy compatible Maxwellians

We recall that an *entropy* for the system (2.11) is by definition a real valued convex function $\mathbf{W} \mapsto \eta(\mathbf{W})$ that satisfies

$$\forall j = 1, \ldots, D, \ \forall \mathbf{W}, \quad \left(\mathbf{F}_j'(\mathbf{W}) \right)^T \eta''(\mathbf{W}) \ \text{is symmetric}, \tag{2.18}$$

where the prime symbol means differentiation with respect to the variable $\mathbf{W}$. This symmetry has to be understood either in the sense of bilinear forms [15, Appendix], or simply in the sense of $p \times p$

matrices if one writes everything in the canonical basis. The condition (2.18) is equivalent to the existence of scalar valued functions $\vartheta_j(\mathbf{W})$ (called *entropy fluxes*) such that

$$\forall j = 1, \ldots, D, \ \forall \mathbf{W}, \quad \vartheta_j'(\mathbf{W}) = \eta'(\mathbf{W})\mathbf{F}_j'(\mathbf{W}), \tag{2.19}$$

Then an *entropy solution* to (2.11) is a weak solution $\mathbf{W}$ that satisfies

$$\partial_t(\eta(\mathbf{W})) + \sum_{j=1}^{D} \frac{\partial}{\partial x_j}(\vartheta_j(\mathbf{W})) \leq 0. \tag{2.20}$$

We shall call *monotone* relatively to a convex function $\eta(\mathbf{W})$ any vector function $\mathbf{W} \mapsto \mathbf{M}(\mathbf{W}) \in \mathbb{R}^p$ such that there holds

$$\forall \mathbf{W}, \quad (\mathbf{M}'(\mathbf{W}))^T \eta''(\mathbf{W}) \ \text{ is symmetric nonnegative}, \tag{2.21}$$

again in the sense of bilinear forms or of $p \times p$ matrices. To be more explicit, if we denote the components of $\mathbf{M}$ by $\mathbf{M} = (\mathcal{M}_1, \cdots, \mathcal{M}_p)$, then the coefficients of the matrix $\mathbf{M}'$ are $\mathbf{M}'_{rs}(\mathbf{W}) = \frac{\partial \mathcal{M}_r}{\partial W_s}(\mathbf{W})$, for all $r, s \in [1, p]$.

Theorem 2.1 ([15] Entropy compatibility of kinetic BGK models)**.** *Assume that η is a convex entropy for the system (2.11), and that Maxwellians $\mathbf{M}[\mathbf{W}](\xi)$ are given satisfying the moment relations (2.16). Assume moreover some technical regularity and topological conditions stated in [15, Theorem 2.1]. Then the Maxwellians satisfy*

$$\textit{for a.e. } \xi \in \Xi, \quad \mathbf{W} \mapsto \mathbf{M}[\mathbf{W}](\xi) \textit{ is monotone relatively to } \eta \tag{2.22}$$

if and only if there exists a kinetic entropy $H(\mathbf{f}, \xi)$ satisfying the three following properties.

(1) *For a.e. $\xi \in \Xi$ the function $\mathbf{f} \mapsto H(\mathbf{f}, \xi)$ is convex with respect to $\mathbf{f}$.*

(2) *For any function $\xi \mapsto \mathbf{f}(\xi)$, denoting*

$$\mathcal{H}(\mathbf{f}) := \int H(\mathbf{f}(\xi), \xi) \, d\xi, \tag{2.23}$$

one has the entropy minimization principle

$$\mathbf{W} = \int \mathbf{f}(\xi) \, d\xi \Rightarrow \ \mathcal{H}(\mathbf{f}) \geq \mathcal{H}(\mathbf{M}[\mathbf{W}](.)). \tag{2.24}$$

(3) *There exists a constant $\bar{c}_\eta$ such that for any $\mathbf{W}$ there holds*

$$\mathcal{H}(\mathbf{M}[\mathbf{W}](.)) = \eta(\mathbf{W}) + \bar{c}_\eta. \tag{2.25}$$

Then, the scalar function

$$G(\mathbf{W}, \xi) := H(\mathbf{M}[\mathbf{W}](\xi), \xi) \tag{2.26}$$

satisfies

$$G'(\mathbf{W}, \xi) = \eta'(\mathbf{W})\mathbf{M}'[\mathbf{W}](\xi), \tag{2.27}$$

where again, prime denotes differentiation with respect to $\mathbf{W}$, and there exist constants $\bar{c}_{\eta,j}$ such that for all $j = 1, \ldots, D$ and all $\mathbf{W}$,

$$\int \mathbf{v}_j(\xi)G(\mathbf{W}, \xi) \, d\xi = \vartheta_j(\mathbf{W}) + \bar{c}_{\eta,j}, \tag{2.28}$$

with ϑ_j the entropy fluxes associated to η, defined by (2.19). The kinetic entropy H has also the following property,

$$\forall \mathbf{W}, \ \textit{a.e. } \xi, \qquad \partial_\mathbf{f} H(\mathbf{M}[\mathbf{W}](\xi), \xi) \ni \eta'(\mathbf{W}), \tag{2.29}$$

where $\partial_\mathbf{f} H$ stands for the subdifferential of the convex function H with respect to $\mathbf{f}$.

The interest of the kinetic entropy H is that multiplying the BGK model (2.12) by $\partial_{\mathbf{f}} H(\mathbf{f}(t,\mathbf{x},\xi),\xi)$ and integrating with respect to ξ one gets according to the properties 1 and 2 above

$$\partial_t \int H(\mathbf{f}(t,\mathbf{x},\xi),\xi)\,d\xi + \sum_{j=1}^D \frac{\partial}{\partial x_j} \int \mathbf{v}_j(\xi) H(\mathbf{f}(t,\mathbf{x},\xi),\xi)\,d\xi \leq 0. \tag{2.30}$$

In the hydrodynamic limit when $\epsilon \to 0$, $\mathbf{f}$ becomes Maxwellian, and this inequality reduces to the macroscopic entropy inequality (2.20) according to the integral relations (2.25) and (2.28). Such entropy compatible hydrodynamic limits were established in particular in [13, 14] for isentropic gas dynamics.

A simple criterion for getting the nonnegativity in (2.21) is as follows.

Lemma 2.2 ([15] Characterization of monotonicity via eigenvalues). *Assume that $\eta(\mathbf{W})$ is a strictly convex function, and consider a vector function $\mathbf{W} \mapsto \mathbf{M}(\mathbf{W}) \in \mathbb{R}^p$ such that there holds*

$$\forall \mathbf{W}, \quad \left(\mathbf{M}'(\mathbf{W})\right)^T \eta''(\mathbf{W}) \quad \text{is symmetric.} \tag{2.31}$$

Then for all $\mathbf{W}$ the matrix $\mathbf{M}'(\mathbf{W})$ is diagonalizable with real eigenvalues. Moreover, the nonnegatvity required in (2.21) of the matrices in (2.31) is equivalent to the property that the eigenvalues of $\mathbf{M}'(\mathbf{W})$ are nonnegative.

The choice of Maxwellians satisfying the entropy compatibility condition (2.22) is an important step to make a significant break with respect to the usual LBM, that are rarely analyzed in terms of entropy. In order to achieve the moment consistency relations (2.16) together with the entropy compatibility (2.22), a very simple choice of Maxwellians is the one proposed in [3, 15], which is to write a linear decomposition as

$$\forall \mathbf{W}, \quad \text{a.e. } \xi \in \Xi, \qquad \mathbf{M}[\mathbf{W}](\xi) = a(\xi)\mathbf{W} + \sum_{j=1}^D b_j(\xi)\mathbf{F}_j(\mathbf{W}), \tag{2.32}$$

where the real coefficients $a(\xi)$, $b_j(\xi)$ satisfy for $i,j = 1,\ldots,D$

$$\int a(\xi)\,d\xi = 1, \quad \int a(\xi)\mathbf{v}_j(\xi)\,d\xi = 0, \tag{2.33}$$

$$\int b_j(\xi)\,d\xi = 0, \quad \int b_i(\xi)\mathbf{v}_j(\xi)\,d\xi = \delta_{ij}. \tag{2.34}$$

These relations ensure obviously the moment relations (2.16). The symmetry condition (2.31) is automatically satisfied because of the characterization (2.18) for η to be an entropy. The function G can then be computed via the differential relation (2.27), that yields

$$G(\mathbf{W},\xi) = a(\xi)\eta(\mathbf{W}) + \sum_{j=1}^D b_j(\xi)\vartheta_j(\mathbf{W}). \tag{2.35}$$

According to Lemma 2.2, the entropy compatibility condition (2.22) then reduces to

$$\forall \mathbf{W}, \quad \text{a.e. } \xi \in \Xi, \qquad a(\xi)\mathbf{I} + \sum_{j=1}^D b_j(\xi)\mathbf{F}_j'(\mathbf{W}) \quad \text{has nonnegative eigenvalues,} \tag{2.36}$$

knowing in advance that these eigenvalues are real. An important aspect of (2.36) is that it does not involve explicitly the entropy η, but just the nonlinearities of the system (2.11). This condition can be understood as a *subcharacteristic condition*, that relates the eigenvalues of the system to the coefficients $a(\xi)$ and $b_j(\xi)$ of the method. Similar generalized stability and subcharacteristic conditions for relaxation systems are stated in [31, 18]. Once the conditions (2.33), (2.34), (2.36) are satisfied, one can compute the kinetic entropy $H(\mathbf{f},\xi)$ by the relation (2.26), since G is given by (2.35) and at

7

fixed ξ one could in principle invert the relation $\mathbf{W} \mapsto \mathbf{M}[\mathbf{W}](\xi)$, although there is very little chance that an explicit formula could be available. Nevertheless within the flux vector splitting approach that is explained in the next sections, the formulation of the entropy inequality does not need the explicit knowledge of the kinetic entropy H, nor the inversion of $\mathbf{M}[\mathbf{W}]$.

Finally, for the choice of the set Ξ, one can take a finite set, and for the measure $d\xi$ the counting measure. This means that $\int d\xi$ simply means the sum over all elements $\xi \in \Xi$. Then the velocity function $\mathbf{v}(\xi)$ can be understood as just a set of finitely many velocities $(\mathbf{v}(\xi))_{\xi \in \Xi}$. The minimal number of velocities (i.e. the number of elements in Ξ) necessary to have solutions $a(\xi)$, $b_j(\xi)$ to (2.33), (2.34) is $D + 1$, but for convenience one can take more, as we shall see.

2.4. **Time-space discretization and flux vector splitting**

For practical implementation, the kinetic BGK model (2.12), (2.13) is always replaced by the so called *transport-projection* algorithm, which means that the BGK collision term in the right-hand side of (2.12) is replaced by a projection to Maxwellian states every timestep. This means to solve in successive time intervals

$$\partial_t \mathbf{f} + \mathbf{v}(\xi) \cdot \nabla_{\mathbf{x}} \mathbf{f} = \mathbf{0}, \quad \text{for } t_n < t < t_{n+1}, \tag{2.37}$$

with Maxwellian initial data

$$\mathbf{f}(t_n, \mathbf{x}, \xi) = \mathbf{f}^n(\mathbf{x}, \xi) = \mathbf{M}[\mathbf{W}^n(\mathbf{x})](\xi), \tag{2.38}$$

$$\mathbf{W}^n(\mathbf{x}) = \int_{\Xi} \mathbf{f}(t_n - 0, \mathbf{x}, \xi) \, d\xi. \tag{2.39}$$

As proved in [16], the previous algorithm means to replace ϵ in (2.12) by the timestep $\Delta t = t_{n+1} - t_n$. This is how time discretization is performed. Note that then the solution to (2.37) is given by $\mathbf{f}(t, \mathbf{x}, \xi) = \mathbf{f}^n(\mathbf{x} - (t - t_n)\mathbf{v}(\xi), \xi)$, for $t_n < t < t_{n+1}$.

About the treatment of the kinetic variable, most lattice Boltzmann methods keep the variable ξ (in a finite set), leading to finitely many transport equations (2.37) that are solved by a numerical scheme in time and space. It gives as many variables to store as the number of ξs. Here we rather adopt the flux vector splitting formulation (FVS for short), that enables to write everything in terms of the integral $\mathbf{W}(t, \mathbf{x}) = \int \mathbf{f}(t, \mathbf{x}, \xi) \, d\xi$.

It was proved in [16] that when space discretization is performed in one dimension, the kinetic transport-projection algorithm (2.37)-(2.39) leads to a FVS method. Conversely, any FVS method can be interpreted as a kinetic transport-projection algorithm. The kinetic entropy compatibility property (2.22) can then be expressed directly at the level of the flux vector splitting, characterizing its entropy property.

We now show that in the multidimensional setting $\mathbf{x} \in \mathbb{R}^D$, the kinetic transport-projection method (2.37)-(2.39) leads to a multidimensional FVS method when discretized in space, generalizing [16]. To simplify we consider only the case of $2D + 1$ velocities defined as

$$\xi \in \Xi = \{0, 1, \dots, 2D\}, \quad d\xi \text{ is the counting measure}, \tag{2.40}$$

$$\mathbf{v}(0) = \mathbf{0}, \ \mathbf{v}(1) = \lambda \mathbf{e}_1, \ \mathbf{v}(2) = -\lambda \mathbf{e}_1, \quad \dots \quad , \mathbf{v}(2D-1) = \lambda \mathbf{e}_D, \ \mathbf{v}(2D) = -\lambda \mathbf{e}_D, \tag{2.41}$$

where $\mathbf{e}_j$ denotes the jth basis vector for $j = 1, \dots, D$, and $\lambda > 0$ is a parameter. There are two velocities per direction, a representation of these velocities in two and three dimensions is proposed on Figures 2.1 and 2.2. We recall that *the counting measure* is the measure for which an integral in $d\xi$ means simply the sum over elements $\xi \in \Xi$.
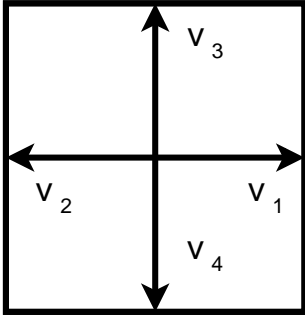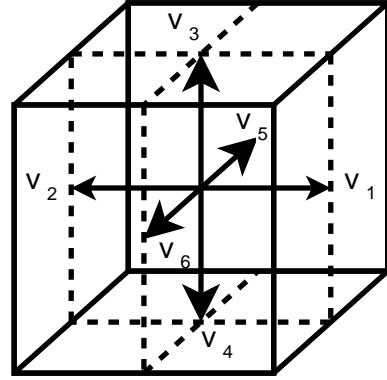
FIGURE 2.1. Discrete velocities in 2 dimensions

FIGURE 2.2. Discrete velocities in 3 dimensions

We consider for each direction $j = 1, \ldots, D$ a splitting of the fluxes $\mathbf{F}_j$ from (2.11), that is a decomposition

$$\mathbf{F}_j(\mathbf{W}) = \mathbf{F}_j^+(\mathbf{W}) + \mathbf{F}_j^-(\mathbf{W}). \tag{2.42}$$

Then we define the Maxwellians $\mathbf{M}[\mathbf{W}](\xi)$ by

$$\mathbf{M}[\mathbf{W}](0) = \mathbf{W} - \sum_{j=1}^{D} \frac{\mathbf{F}_j^+(\mathbf{W}) - \mathbf{F}_j^-(\mathbf{W})}{\lambda}, \tag{2.43}$$

$$\mathbf{M}[\mathbf{W}](2j-1) = \frac{\mathbf{F}_j^+(\mathbf{W})}{\lambda}, \quad \mathbf{M}[\mathbf{W}](2j) = -\frac{\mathbf{F}_j^-(\mathbf{W})}{\lambda}, \qquad \text{for } j = 1, \ldots, D. \tag{2.44}$$

With these definitions (2.40)-(2.44) it is straightforward to check that the moment relations (2.16) hold.

In order to write the scheme, we consider a Cartesian mesh of cells $C_{i_1,\ldots,i_D} = (x_{i_1-1/2}^1, x_{i_1+1/2}^1) \times \ldots \times (x_{i_D-1/2}^D, x_{i_D+1/2}^D) \subset \mathbb{R}^D$, indexed by $\mathbf{i} = (i_1, \ldots, i_D) \in \mathbb{Z}^D$. We start with an initial data $\mathbf{W}^n(x)$ constant in each cell with associated value $\mathbf{W}_{\mathbf{i}}^n$. Then we solve the kinetic transport equation (2.37) with Maxwellian initial data (2.38) associated to $\mathbf{W}^n(x)$. Finally the updated values $\mathbf{W}_{\mathbf{i}}^{n+1}$ are obtained by averaging (2.39) over each cell, which gives

$$\mathbf{W}_{\mathbf{i}}^{n+1} = \fint_{C_{\mathbf{i}}} \int_{\Xi} \mathbf{f}(t_{n+1} - 0, \mathbf{x}, \xi) \, d\xi d\mathbf{x}, \tag{2.45}$$

where the bar integral means that the integral is normalized by the volume. A numerical flux in the direction j is associated to a location $(i_1, \ldots, i_j + 1/2, \ldots, i_D) = \mathbf{i} + \mathbf{e}_j/2$, with $\mathbf{e}_j$ the jth basis vector of $\mathbb{R}^D$.

Proposition 2.3 (Expression of the transport-projection method as a flux vector splitting scheme)**.** *Consider a multidimensional system of conservation laws (2.11), and for each direction j a decomposition of the flux $\mathbf{F}_j$ as (2.42). Define the velocities by (2.40), (2.41) and the Maxwellian by (2.43), (2.44). Then under the CFL condition*

$$\lambda \Delta t \leq \min(x_{i_j+1/2}^j - x_{i_j-1/2}^j, x_{i_j+3/2}^j - x_{i_j+1/2}^j), \tag{2.46}$$

the kinetic transport-projection method (2.37)-(2.39) starting with piecewise constant data over a Cartesian mesh and followed by the piecewise constant projection (2.45) can be written as the flux

9

vector splitting scheme

$$\mathbf{W_i}^{n+1} - \mathbf{W_i}^n + \sum_{j=1}^{D} \frac{\Delta t}{x_{i_j+1/2}^j - x_{i_j-1/2}^j}(\mathbf{F}_{\mathbf{i+e}_j/2}^n - \mathbf{F}_{\mathbf{i-e}_j/2}^n) = \mathbf{0}, \tag{2.47}$$

$$\mathbf{F}_{\mathbf{i+e}_j/2}^n = \mathbf{F}_j^+(\mathbf{W_i}^n) + \mathbf{F}_j^-(\mathbf{W_{i+e}}_j^n). \tag{2.48}$$

Proof. Integrating (2.37) in $t, \mathbf{x}, \xi$ over $(t_n, t_{n+1}) \times C_\mathbf{i} \times \Xi$ and taking into account (2.38) yields (2.47) with the numerical fluxes

$$\mathbf{F}_{\mathbf{i+e}_j/2}^n = \frac{1}{\Delta t}\int_{t_n}^{t_{n+1}} \fint_{C_{i_1,\dots,i_{j-1},i_{j+1},\dots,i_D}} \int_\Xi \mathbf{v}_j(\xi)\mathbf{f}(t, x_1, \dots, x_{j-1}, x_{i_j+1/2}^j, x_{j+1}, \dots, x_D, \xi)$$
$$\times dt dx_1 \dots dx_{j-1} dx_{j+1} \dots dx_D d\xi, \tag{2.49}$$

where $C_{i_1,\dots,i_{j-1},i_{j+1},\dots,i_D} \subset \mathbb{R}^{D-1}$ stands for the rectangle where the interval $(x_{i_j-1/2}^j, x_{i_j+1/2}^j)$ is missing. Then we write that the solution $\mathbf{f}$ to (2.37) is given for $t_n < t < t_{n+1}$ by $\mathbf{f}(t, \mathbf{x}, \xi) = \mathbf{f}^n(\mathbf{x} - (t-t_n)\mathbf{v}(\xi), \xi) = \mathbf{M}[\mathbf{W}^n(\mathbf{x} - (t-t_n)\mathbf{v}(\xi))](\xi)$. In computing the integral (2.49), only the values of ξ satisfying $\mathbf{v}_j(\xi) \neq 0$ are involved. According to (2.41), this means that only the values $\xi = 2j-1$ and $\xi = 2j$ have to be considered. Then $\mathbf{v}(\xi) = \pm\lambda\mathbf{e}_j$, thus with the values of $t, \mathbf{x}, \xi$ involved in (2.49) one has $\mathbf{x} - (t-t_n)\mathbf{v}(\xi) \in C_{i_1,\dots,i_j,\dots,i_D}$ if $\xi = 2j-1$, or $\mathbf{x} - (t-t_n)\mathbf{v}(\xi) \in C_{i_1,\dots,i_j+1,\dots,i_D}$ if $\xi = 2j$, provided that the CFL condition (2.46) holds. It follows that we can compute the integral (2.49),

$$\mathbf{F}_{\mathbf{i+e}_j/2}^n = \mathbf{v}_j(2j-1)\mathbf{M}[\mathbf{W_i}^n](2j-1) + \mathbf{v}_j(2j)\mathbf{M}[\mathbf{W_{i+e}}_j^n](2j) = \mathbf{F}_j^+(\mathbf{W_i}^n) + \mathbf{F}_j^-(\mathbf{W_{i+e}}_j^n), \tag{2.50}$$

which proves the claim. ∎

The denomination *flux vector splitting* (FVS) means that the numerical flux takes the form (2.48), i.e. it is the sum of a term $\mathbf{F}_j^+$ depending on the left unknown and a term $\mathbf{F}_j^-$ depending on the right unknown. Then in the update formula (2.47), the new value $\mathbf{W_i}^{n+1}$ depends on the old value $\mathbf{W_i}^n$ and on the values $\mathbf{W_{i\pm e}}_j^n$ at its neighbor cells in the directions of the discrete velocities, as illustrated on Figures 2.1 and 2.2.

Corollary 2.4 (Equivalence between kinetic transport-projection and flux vector splitting)**.** *Any flux vector splitting scheme (2.47), (2.48) associated to given $\mathbf{F}_j^\pm$ satisfying (2.42) can be interpreted as a kinetic scheme with $2D+1$ velocities defined by (2.40), (2.41) and Maxwellians defined by (2.43), (2.44). Conversely, any kinetic scheme with the $2D+1$ velocities defined by (2.40), (2.41) gives rise to a flux vector splitting scheme.*

Proof. The first statement is given by Proposition 2.3. For the converse, we just have to remark that the moment conditions (2.16) imply that the Maxwellians necessarily take the form (2.43), (2.44) with (2.42). ∎

In the context of Proposition 2.3, writing the entropy compatibility condition (2.22) on the Maxwellians (2.43), (2.44) gives the conditions that

$$\mathbf{F}_j^+ \quad \text{and} \ -\mathbf{F}_j^- \quad \text{are monotone respectively to } \eta \text{ for all } j = 1, \dots, D, \tag{2.51}$$

$$\mathbf{W} - \sum_{j=1}^{D} \frac{\mathbf{F}_j^+(\mathbf{W}) - \mathbf{F}_j^-(\mathbf{W})}{\lambda} \quad \text{is monotone respectively to } \eta. \tag{2.52}$$

We can slightly simplify the condition (2.52) and give the following entropy stability result.

Proposition 2.5 (Entropy inequality for a FVS scheme). *Assume that η is a strictly convex entropy for the system of conservation laws (2.11), and consider a Cartesian mesh with constant size Δx in all directions. Then the flux vector splitting scheme (2.47), (2.48) associated to some given split fluxes $\mathbf{F}_j^{\pm}$ satisfying (2.42) satisfies a discrete entropy inequality*

$$\eta(\mathbf{W}_{\mathbf{i}}^{n+1}) - \eta(\mathbf{W}_{\mathbf{i}}^{n}) + \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^{n} - \vartheta_{\mathbf{i}-\mathbf{e}_j/2}^{n}) \leq 0, \tag{2.53}$$

under the conditions

$$\mathbf{F}_j^+ \quad and \quad -\mathbf{F}_j^- \quad are \ monotone \ respectively \ to \ \eta \ for \ all \ j = 1, \ldots, D, \tag{2.54}$$

$$the \ eigenvalues \ of \quad \sum_{j=1}^{D} (\mathbf{F}_j^+ - \mathbf{F}_j^-)' \quad are \ less \ than \ or \ equal \ to \ \lambda, \tag{2.55}$$

$$\lambda \Delta t \leq \Delta x. \tag{2.56}$$

Moreover, the numerical entropy fluxes take the form

$$\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^{n} = \vartheta_j^+(\mathbf{W}_{\mathbf{i}}^{n}) + \vartheta_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^{n}), \tag{2.57}$$

with

$$(\vartheta_j^+)' = \eta'(\mathbf{F}_j^+)', \qquad (\vartheta_j^-)' = \eta'(\mathbf{F}_j^-)'. \tag{2.58}$$

Proof. With Lemma 2.2, condition (2.52) simplifies to (2.55). Since the cell length Δx is constant, (2.46) reduces to (2.56). Next, under the two conditions (2.51), (2.52) we can apply Theorem 2.1 thus there exists a kinetic entropy $H(\mathbf{f}, \xi)$. Since $\mathbf{f}$ satisfies the kinetic equation (2.37) for $t_n < t < t_{n+1}$, one has also

$$\partial_t(H(\mathbf{f}, \xi)) + \mathbf{v}(\xi) \cdot \nabla_{\mathbf{x}}(H(\mathbf{f}, \xi)) = 0, \quad \text{for } t_n < t < t_{n+1}. \tag{2.59}$$

Integrating this in $t, \mathbf{x}, \xi$ over $(t_n, t_{n+1}) \times C_{\mathbf{i}} \times \Xi$ as in the proof of Proposition 2.3 yields

$$\fint_{C_{\mathbf{i}}} \int_{\Xi} H(\mathbf{f}(t_{n+1} - 0, \mathbf{x}, \xi), \xi) \, d\xi d\mathbf{x} - \fint_{C_{\mathbf{i}}} \int_{\Xi} H(\mathbf{f}(t_n, \mathbf{x}, \xi), \xi) \, d\xi d\mathbf{x}$$
$$+ \sum_{j=1}^{D} \frac{\Delta t}{x_{i_j+1/2}^j - x_{i_j-1/2}^j} (\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^{n} - \vartheta_{\mathbf{i}-\mathbf{e}_j/2}^{n}) = 0, \tag{2.60}$$

for some numerical entropy fluxes $\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^{n}$ that can be expressed as

$$\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^{n} = \mathbf{v}_j(2j-1) H(\mathbf{M}[\mathbf{W}_{\mathbf{i}}^{n}](2j-1), 2j-1) + \mathbf{v}_j(2j) H(\mathbf{M}[\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^{n}](2j), 2j)$$
$$\equiv \vartheta_j^+(\mathbf{W}_{\mathbf{i}}^{n}) + \vartheta_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^{n}). \tag{2.61}$$

According to (2.26), (2.27), the functions $\vartheta_j^{\pm}$ satisfy (2.58). Then by (2.38) $\mathbf{f}(t_n, \mathbf{x}, \xi) = \mathbf{M}[\mathbf{W}^n(\mathbf{x})](\xi)$, while by (2.39) $\int \mathbf{f}(t_{n+1} - 0, \mathbf{x}, \xi) \, d\xi = \mathbf{W}^{n+1-}(\mathbf{x})$. Using (2.23), (2.24), (2.25) we deduce that

$$\int_{\Xi} H(\mathbf{f}(t_n, \mathbf{x}, \xi), \xi) \, d\xi = \eta(\mathbf{W}^n(\mathbf{x})) + \bar{c}_\eta, \qquad \int_{\Xi} H(\mathbf{f}(t_{n+1} - 0, \mathbf{x}, \xi), \xi) \, d\xi \geq \eta(\mathbf{W}^{n+1-}(\mathbf{x})) + \bar{c}_\eta. \tag{2.62}$$

Finally, using that $\mathbf{W}^n(\mathbf{x})$ is piecewise constant, that $\mathbf{W}_{\mathbf{i}}^{n+1} = \fint \mathbf{W}^{n+1-}(\mathbf{x}) \, d\mathbf{x}$ according to (2.45), averaging (2.62) over $\mathbf{x} \in C_{\mathbf{i}}$ yields with the help of Jensen's inequality

$$\fint_{C_{\mathbf{i}}} \int_{\Xi} H(\mathbf{f}(t_n, \mathbf{x}, \xi), \xi) \, d\xi d\mathbf{x} = \eta(\mathbf{W}_{\mathbf{i}}^{n}) + \bar{c}_\eta, \qquad \fint_{C_{\mathbf{i}}} \int_{\Xi} H(\mathbf{f}(t_{n+1} - 0, \mathbf{x}, \xi), \xi) \, d\xi d\mathbf{x} \geq \eta(\mathbf{W}_{\mathbf{i}}^{n+1}) + \bar{c}_\eta. \tag{2.63}$$

Plugging this in (2.60) finally yields (2.53). ∎

We can remark that the parameter λ is somehow artificial since it does not appear in the scheme formulation (2.47), (2.48). The consequence is that one can indeed eliminate it from (2.55), (2.56), giving the sharp CFL condition

$$\text{the eigenvalues of} \quad \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\mathbf{F}_j^+ - \mathbf{F}_j^-)' \quad \text{are less than or equal to 1.} \tag{2.64}$$

2.5. **The multidimensional Lax-Friedrichs scheme**

A particularly simple choice of flux vector splitting scheme for which the analysis of the previous subsection applies is the Lax-Friedrichs scheme, for which one chooses

$$\mathbf{F}_j^+(\mathbf{W}) = \frac{c\mathbf{W} + \mathbf{F}_j(\mathbf{W})}{2}, \qquad \mathbf{F}_j^-(\mathbf{W}) = \frac{-c\mathbf{W} + \mathbf{F}_j(\mathbf{W})}{2}, \tag{2.65}$$

for some constant $c > 0$ that measures the numerical viscosity, and that has to be chosen in relation with stability issues. Then the consistency (2.42) obviously holds, and the Maxwellian $\mathbf{M}[\mathbf{W}](0)$ from (2.43) that is involved in (2.52) is just

$$\mathbf{M}[\mathbf{W}](0) = (1 - \frac{Dc}{\lambda})\mathbf{W}. \tag{2.66}$$

Writing the entropy compatibility conditions (2.51), (2.52), we can use Lemma 2.2. Since the symmetry condition (2.31) is automatically satisfied, the entropy stability reduces to the subcharacteristic condition

$$\text{the eigenvalues of } \mathbf{F}_j{}' \text{ have absolute values less than or equal to } c, \tag{2.67}$$

and the CFL condition

$$\frac{Dc}{\lambda} \leq 1, \qquad \frac{\lambda \Delta t}{\Delta x} \leq 1, \tag{2.68}$$

that can be simplified (by eliminating the parameter λ as in (2.64)) to

$$D\frac{c\Delta t}{\Delta x} \leq 1. \tag{2.69}$$

The numerical entropy fluxes (2.57) can be explicited since (2.58) gives with (2.65)

$$\vartheta_j^+(\mathbf{W}) = \frac{c\,\eta(\mathbf{W}) + \vartheta_j(\mathbf{W})}{2}, \qquad \vartheta_j^-(\mathbf{W}) = \frac{-c\,\eta(\mathbf{W}) + \vartheta_j(\mathbf{W})}{2}, \tag{2.70}$$

with $\vartheta_j(\mathbf{W})$ the entropy fluxes defined by (2.19).

We can remark that with the Lax-Friedrichs choice (2.65), the Maxwellians (2.43), (2.44) take the particular form (2.32) with the coefficients

$$a(\xi) = \begin{cases} 1 - \dfrac{Dc}{\lambda} & \text{if } \xi = 0, \\ \dfrac{c}{2\lambda} & \text{otherwise.} \end{cases} \qquad b_j(\xi) = \begin{cases} \dfrac{1}{2\lambda} & \text{if } \xi = 2j - 1, \\ \dfrac{-1}{2\lambda} & \text{if } \xi = 2j, \\ 0 & \text{otherwise.} \end{cases} \tag{2.71}$$

These satisfy the moment relations (2.33), (2.34). The particular choice $\lambda = Dc$ gives $\mathbf{M}[\mathbf{W}](0) = \mathbf{0}$ in (2.66). Thus in this case the model is indeed a $2D$ velocity model (instead of $2D + 1$). Notice that in any case, as in (2.47), (2.48), (2.64), the parameter λ disappears at the level of integrals in ξ, and it remains only the parameter c.

2.6. **Direct entropy analysis of flux vector splitting schemes**

As stated in Proposition 2.3 the kinetic approach leads at end to a FVS scheme, for which the stability can be analyzed (Proposition 2.5). However the interpretation of the FVS scheme (2.47), (2.48) as a kinetic scheme involves the dummy parameter λ (kinetic velocity in (2.41)), and the kinetic variable ξ is also somehow artificial since they do no longer appear in the formulas (2.47), (2.48). A more direct way to analyse a FVS scheme as (2.47), (2.48) in the multidimensional context is provided in [19], and we describe it here. The formulation has the advantage to completely avoid any kinetic variable, and moreover gives a precise meaning to the "technical regularity and topological conditions" involved in Theorem 2.1. Thus it improves the results of Proposition 2.5. The sharp characterization of entropy compatibility of FVS schemes can be written as follows according to [19].

The result states the entropy inequality, for a scheme that can be written under the form of updated values computed at a particular location, as

$$\mathbf{W}^{update} = \sum_{J \in \mathbf{J}} \varphi_J(\mathbf{W}_J), \tag{2.72}$$

where $\mathbf{J}$ represents the stencil of dependency, $\mathbf{W}_J$ (for $J \in \mathbf{J}$) are the values of the unknown on which the update depend on, and φ_J are nonlinear functions.

Proposition 2.6 ([19] Direct entropy analysis of VFS schemes)**.** *Let $\mathcal{W} \subset \mathbb{R}^p$ be an open convex set, and let $\varphi_J : \mathcal{W} \to \mathbb{R}^p$ be of class C^2, $J \in \mathbf{J}$, where $\mathbf{J}$ is a finite set. Let $\eta : \mathcal{W} \to \mathbb{R}$ be of class C^2, convex, and $q_J : \mathcal{W} \to \mathbb{R}$ be such that*

$$q_J'(\mathbf{W}) = \eta'(\mathbf{W})\varphi_J'(\mathbf{W}) \qquad \text{for all } \mathbf{W} \in \mathcal{W}, J \in \mathbf{J}, \tag{2.73}$$

normalized so that

$$\sum_{J \in \mathbf{J}} q_J(0) = \eta(0). \tag{2.74}$$

Assume that each φ_J is η-dissipative in a (not necessarily convex) subset $\mathcal{W}_{stab}$ of $\mathcal{W}$, that is,

$$\mathcal{D}_J(\mathbf{W}, \widehat{\mathbf{W}}) := q_J(\mathbf{W}) - q_J(\widehat{\mathbf{W}}) - \eta'(\mathbf{W})(\varphi_J(\mathbf{W}) - \varphi_J(\widehat{\mathbf{W}})) \leq 0, \qquad \mathbf{W}, \widehat{\mathbf{W}} \in \mathcal{W}_{stab}. \tag{2.75}$$

Suppose further that

$$\sum_{J \in \mathbf{J}} \varphi_J(\mathbf{W}) = \mathbf{W}, \qquad \text{for all } \mathbf{W} \in \mathcal{W}. \tag{2.76}$$

Then, if $\mathbf{W}_J \in \mathcal{W}_{stab}$ for all $J \in \mathbf{J}$ and $\sum_{J \in \mathbf{J}} \varphi_J(\mathbf{W}_J) \in \mathcal{W}_{stab}$, we have

$$\eta\left(\sum_{J \in \mathbf{J}} \varphi_J(\mathbf{W}_J)\right) \leq \sum_{J \in \mathbf{J}} q_J(\mathbf{W}_J). \tag{2.77}$$

For the understanding of this proposition, some comments are in order. First, the assumption (2.76) is very natural for a system of conservation laws (2.11), it states that if we start from constant initial data, the scheme gives the same constant as updates. Then, the condition that there exist q_J satisfying (2.73) is equivalent, as in (2.18), (2.19), to the property that $(\varphi_J')^T \eta''$ is symmetric. Finally, the dissipation condition (2.75) means that $\mathbf{W} \mapsto \varphi_J(\mathbf{W})$ is monotone relatively to η (definition (2.21)) plus some technical/topological assumptions (see [19]). Additionally, we have to mention that within the framework of Proposition 2.6 one can also establish that the scheme preserves some *invariant domains* for the system (2.11), if they exist [19].

With Proposition 2.6 we can analyze the flux vector splitting scheme (2.47), (2.48) without involving any kinetic model. This is done as follows. The first step is to express the numerical fluxes into the

update, giving

$$
\begin{aligned}
\mathbf{W}_{\mathbf{i}}^{n+1} &= \mathbf{W}_{\mathbf{i}}^{n} - \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\mathbf{F}_j^+(\mathbf{W}_{\mathbf{i}}^n) + \mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n) - \mathbf{F}_j^+(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n) - \mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}}^n)) \\
&= \left(\mathbf{W}_{\mathbf{i}}^{n} - \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\mathbf{F}_j^+(\mathbf{W}_{\mathbf{i}}^n) - \mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}}^n)) \right) - \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n) - \mathbf{F}_j^+(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n)).
\end{aligned}
\tag{2.78}
$$

In this form the update $\mathbf{W}_{\mathbf{i}}^{n+1}$ is as in (2.72) the sum of nonlinear functions of the $2D+1$ unknowns $\mathbf{W}_{\mathbf{i}}^n$ and $\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n$, $\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n$ for $j = 1, \ldots, D$. Thus we can apply Proposition 2.6. Condition (2.76) holds obviously, thus the entropy condition is that all the nonlinear functions φ_J are $\eta-$dissipative. We therefore recover (2.51), (2.52) (recall that the $\eta-$dissipativity is a technically refined analogue of the monotonicity relatively to η, see [19]), in the case $\lambda = \Delta x / \Delta t$. For general λ satisfying (2.55), (2.56), denoting $\mathbf{S} = \sum_j (\mathbf{F}_j^+ - \mathbf{F}_j^-)$, one can write

$$
\mathbf{W} - \frac{\Delta t}{\Delta x} \mathbf{S}(\mathbf{W}) = (1 - \frac{\lambda \Delta t}{\Delta x}) \mathbf{W} + \frac{\lambda \Delta t}{\Delta x} \left(\mathbf{W} - \frac{\mathbf{S}(\mathbf{W})}{\lambda} \right),
\tag{2.79}
$$

thus again we recover the sufficient conditions (2.51), (2.52) under the CFL condition (2.56). Finally, according to (2.77) the entropy inequality writes

$$
\eta(\mathbf{W}_{\mathbf{i}}^{n+1}) \le \eta(\mathbf{W}_{\mathbf{i}}^{n}) - \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\vartheta_j^+(\mathbf{W}_{\mathbf{i}}^n) - \vartheta_j^-(\mathbf{W}_{\mathbf{i}}^n)) - \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} (\vartheta_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n) - \vartheta_j^+(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n)),
\tag{2.80}
$$

where $\vartheta_j^\pm$ satisfy (2.58). The inequality (2.80) can then be written also as (2.53), with the numerical entropy fluxes (2.57). Therefore we recover Proposition 2.5 without using the kinetic approach, and with the precise notion of η-dissipativity instead of monotonicity respectively to η.

We can give some slight generalization of Proposition 2.6 as follows.

Proposition 2.7. *Proposition 2.6 is still valid if we replace the property* $\mathbf{W}, \widehat{\mathbf{W}} \in \mathcal{W}_{stab}$ *in (2.75) by* $(\mathbf{W}, \widehat{\mathbf{W}}) \in \mathcal{P}_{stab}$, *for some set* $\mathcal{P}_{stab} \subset \mathcal{W} \times \mathcal{W}$, *and the requirement that* $\mathbf{W}_J$, $\sum_{J \in \mathbf{J}} \varphi_J(\mathbf{W}_J)$ *lie in* $\mathcal{W}_{stab}$ *by the requirement that* $(\sum_{J' \in \mathbf{J}} \varphi_{J'}(\mathbf{W}_{J'}), \mathbf{W}_J) \in \mathcal{P}_{stab}$ *for all* $J \in \mathbf{J}$.

Proof. The set $\mathcal{P}_{stab}$ is not necessarily symmetric in its two variables, but anyway the proof is identical to that of [19, Lemma 2.2]. ■

The set $\mathcal{P}_{stab}$ represents a set of states where the scheme is stable. It has to be chosen as large as possible while satisfying the dissipation property (2.75). An example is given in (4.27) with a choice depending on the parameter c.

3. Parabolic scaling and low Mach number limit

3.1. Parabolic scaling

The origin of hyperbolic/parabolic limits from kinetic equations lies in the scaled Boltzmann equation (2.1). We can consider similarly the two-scale kinetic BGK model

$$
\partial_t \mathbf{f} + \frac{1}{\epsilon_{\mathrm{adv}}} \mathbf{v}(\xi) \cdot \nabla_{\mathbf{x}} \mathbf{f} = \frac{1}{\epsilon_{coll}} (\mathbf{M}[\mathbf{W}] - \mathbf{f}),
\tag{3.1}
$$

where $\mathbf{W}$ is still defined by (2.13), i.e. $\mathbf{W} = \int \mathbf{f} d\xi$, and ϵ_{adv}, ϵ_{coll} both tend to zero. As explained in Subsection 2.2, the hydrodynamic limit $\epsilon_{coll} \to 0$ leads to the system of conservation laws (2.11) with

fluxes $\mathbf{F}_j^{\epsilon_{\mathrm{adv}}}$ given by the integrals in (2.16) where the scaling by ϵ_{adv} in the velocity has to be taken into account, giving thus

$$\int \frac{\mathbf{v}_j(\xi)}{\epsilon_{\mathrm{adv}}}\,\mathbf{M}[\mathbf{W}](\xi)\,d\xi = \mathbf{F}_j^{\epsilon_{\mathrm{adv}}}(\mathbf{W}). \tag{3.2}$$

Therefore, performing first the limit $\epsilon_{coll} \to 0$ and then the limit $\epsilon_{\mathrm{adv}} \to 0$ does not lead to anything interesting since $\mathbf{F}_j^{\epsilon_{\mathrm{adv}}}(\mathbf{W})$ will either blow up or vanish identically if $\int \mathbf{v}_j \mathbf{M}\,d\xi = 0$.

On the contrary if we consider the simultaneous limit when $\epsilon_{\mathrm{adv}} = \sqrt{\epsilon_{coll}} \to 0$ (corresponding to $q = 1$ in the scaled Boltzmann equation (2.1)), and assume that the Maxwellian satisfies the relations

$$\int \mathbf{v}_j(\xi)\,\mathbf{M}[\mathbf{W}](\xi)\,d\xi = \mathbf{0}, \qquad \int \mathbf{v}_j(\xi)\mathbf{v}_k(\xi)\,\mathbf{M}[\mathbf{W}](\xi)\,d\xi = \delta_{jk}\mathbf{B}(\mathbf{W}), \tag{3.3}$$

for some nonlinear function $\mathbf{B}$, then as proved in [20, 4] the limit equation on $\mathbf{W} = \int \mathbf{f}\,d\xi$ when $\mathbf{f}$ solves (3.1) is the nonlinear parabolic equation

$$\partial_t \mathbf{W} - \Delta_{\mathbf{x}}(\mathbf{B}(\mathbf{W})) = \mathbf{0}. \tag{3.4}$$

By putting some adequate dependency of $\mathbf{v}(\xi)$ in ϵ_{coll}, it is also possible to get a mixed hyperbolic-parabolic system [20, 4].

When going to time-space discretization following Subsection 2.4, ϵ_{coll} becomes the timestep Δt, and somehow ϵ_{adv} is replaced by the space grid size Δx. The relation $\epsilon_{\mathrm{adv}} \sim \sqrt{\epsilon_{coll}}$ becomes then a parabolic CFL condtion saying that Δt should scale as Δx^2, a fact that will be analyzed further on in the case of Navier-Stokes equations. However, discrete effects in the singular hydrodynamic limit lead to the important fact that the consistency, which is related to moment relations, has to be examined at the discrete level and cannot be deduced from the continuous approach. This is what happens in [16, section 6] and in [4, 19, 29]. In this context not only the Laplacian but also cross derivatives can be handled, see [19]. In the case of incompressible Navier-Stokes equations we have also to analyze the scheme at the discrete level, and we indeed use the discrete form of flux vector splitting provided by (2.47), (2.48).

3.2. **Low Mach number limit**

A key point that enables our analysis is that the simultaneous limit as $\epsilon_{adv}, \epsilon_{coll} \to 0$ with $\epsilon_{\mathrm{adv}} \sim \sqrt{\epsilon_{coll}}$ of the kinetic equation (3.1) can be replaced by successive limits of first the hydrodynamic limit $\epsilon_{coll} \to 0$, then the limit $\epsilon_{\mathrm{adv}} \to 0$ but in which *the variable is scaled with respect to ϵ_{adv}*, in the spirit of (2.4). The limit $\epsilon_{coll} \to 0$ at fixed ϵ_{adv} leads to a system of conservation laws, and the one that is our concern is the classical system of compressible isentropic gas dynamics

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot \mathbf{m} = 0, \tag{3.5}$$

$$\partial_t \mathbf{m} + \nabla_{\mathbf{x}} \cdot \left(\frac{\mathbf{m} \otimes \mathbf{m}}{\rho} + P(\rho)\mathbf{I} \right) = 0, \tag{3.6}$$

where the pressure law $P(\rho)$ satisfies $dP/d\rho > 0$ (in particular, $P(\rho) = \kappa \rho^\gamma$ with $\gamma \geq 1$ for a polytropic gas). In (3.5), (3.6), $\rho \geq 0$ is the density, and $\mathbf{m} \in \mathbb{R}^D$ the momentum. The system (3.5), (3.6) is thus of the form of the general system of conservation laws (2.11), with

$$\mathbf{W} = \begin{pmatrix} \rho \\ \mathbf{m} \end{pmatrix}, \qquad \mathbf{F}_j(\mathbf{W}) = \begin{pmatrix} \mathbf{m}_j \\ \frac{\mathbf{m}_j \mathbf{m}}{\rho} + P(\rho)\mathbf{e}_j \end{pmatrix}, \tag{3.7}$$

where $\mathbf{e}_j \in \mathbb{R}^D$ is the jth basis vector. As is well-known, the energy

$$\eta(\rho, \mathbf{m}) = \frac{1}{2}\frac{|\mathbf{m}|^2}{\rho} + \rho e(\rho), \quad \text{with} \quad \frac{de}{d\rho} = \frac{P(\rho)}{\rho^2}, \tag{3.8}$$

15

is a strictly convex entropy for the system, with entropy fluxes

$$\vartheta_j(\rho, \mathbf{m}) = \left(\frac{1}{2} \frac{|\mathbf{m}|^2}{\rho} + \rho e(\rho) + P(\rho) \right) \frac{\mathbf{m}_j}{\rho}. \tag{3.9}$$

What is called the Mach number is $|\mathbf{m}|/(\rho \sqrt{P'(\rho)})$.

We now introduce the scaling with respect to ϵ_{adv}. We shall denote now $\epsilon := \epsilon_{adv}$, (there is no more ϵ_{coll} since it has been sent to 0). Note that this notation is different from that in Subsection 2.2, where ϵ was ϵ_{coll} For the reader's convenience, the different meanings of the parameter ϵ are mentioned in the appendix. We consider that the unkown $\mathbf{W} = (\rho, \mathbf{m})$ takes the form

$$\rho = \overline{\rho} + O(\epsilon^2), \qquad \mathbf{m} = \epsilon \rho \mathbf{u}, \quad \text{with } \mathbf{u} = O(1), \tag{3.10}$$

where $\overline{\rho} > 0$ is a constant. Then writing the isentropic system (3.5), (3.6) with (3.10) and the factor $1/\epsilon$ gives

$$\partial_t \rho + \frac{1}{\epsilon} \nabla_{\mathbf{x}} \cdot (\epsilon \rho \mathbf{u}) = 0, \tag{3.11}$$

$$\partial_t(\epsilon \rho \mathbf{u}) + \frac{1}{\epsilon} \nabla_{\mathbf{x}} \cdot \left(\frac{\epsilon \rho \mathbf{u} \otimes \epsilon \rho \mathbf{u}}{\rho} + P(\rho)\mathbf{I} \right) = 0. \tag{3.12}$$

This can be written more simply as

$$\partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho \mathbf{u}) = 0, \tag{3.13}$$

$$\partial_t(\rho \mathbf{u}) + \nabla_{\mathbf{x}} \cdot \left(\rho \mathbf{u} \otimes \mathbf{u} + \frac{P(\rho)}{\epsilon^2} \mathbf{I} \right) = 0. \tag{3.14}$$

Hence we see that the above scaling is equivalent to putting a factor $1/\epsilon^2$ in front of the pressure, and this means that the Mach number is of the order ϵ, justifying the denomination *low Mach number limit*. Writing the entropy inequality for this system, integrating in space and subtracting a conserved quantity leads to

$$\frac{d}{dt} \int \left(\rho \frac{|\mathbf{u}|^2}{2} + \frac{\rho e(\rho) - \overline{\rho} e(\overline{\rho}) - (e(\overline{\rho}) + P(\overline{\rho})/\overline{\rho})(\rho - \overline{\rho})}{\epsilon^2} \right) d\mathbf{x} \leq 0, \tag{3.15}$$

where we neglected the boundary terms and assumed that $\rho \to \overline{\rho}$ at infinity. Since the function $\rho e(\rho)$ is convex with respect to ρ, with derivative $e(\rho) + P(\rho)/\rho$, the term in $1/\epsilon^2$ in the integral is nonnegative. The relative entropy estimate (3.15) thus gives that $\mathbf{u}$ and $(\rho - \overline{\rho})/\epsilon$ remain bounded if they initially are. But then (3.14) yields that $\nabla_{\mathbf{x}}(P(\rho) - P(\overline{\rho}))/\epsilon^2$ is bounded (in distribution sense at least), which leads finally to $\rho = \overline{\rho} + O(\epsilon^2)$ if all data are smooth, i.e. (3.10) holds. Indeed the fact that there is no first-order term in ϵ in the expansion of ρ is related to the fact that in the expansion (2.4) with $r = 1$, $q = 1$, the first-order term $\widetilde{\rho}$ has to be constant according to (2.6) since the temperature $\widetilde{T}$ is constant in view of the second equation in (2.7).

Finally, writing

$$\rho \to \overline{\rho}, \qquad \frac{P(\rho) - P(\overline{\rho})}{\epsilon^2} \to \overline{\rho} p, \tag{3.16}$$

passing to the limit in (3.13), (3.14) as $\epsilon \to 0$ yields

$$\nabla_{\mathbf{x}} \cdot \mathbf{u} = 0, \tag{3.17}$$

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (\mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) = \mathbf{0}, \tag{3.18}$$

that is the inviscid incompressible Navier-Stokes equations (or incompressible Euler equations). Note that the second-order approximation $\rho = \overline{\rho} + O(\epsilon^2)$ enables indeed to get from (3.13) the order of convergence in the incompressibility condition, that is $\nabla_{\mathbf{x}} \cdot \mathbf{u} = O(\epsilon^2)$. The lack of viscosity in (3.18) is here due to the fact that collisions are no longer involved in the previous low Mach number limit $\epsilon \to 0$ (the hydrodynamic limit $\epsilon_{coll} \to 0$ has been performed first, so as to consider only the macroscopic

isentropic model). The low Mach number limit is justified rigorously in [54, 77, 58]. The case when viscosity is present in the isentropic system is also established in [58, 25]. More recent results are in [1, 26].

4. Flux vector splitting schemes for incompressible Navier-Stokes equations

In Section 2 we have recalled and developed the consistency and stability analysis of kinetic/FVS schemes. We would like now to use the approach while doing at the same time the low Mach number limit recalled in Section 3, so as to obtain a scheme for the Navier-Stokes equations with given viscosity $\nu > 0$. Doing that, the Mach number ϵ becomes a numerical parameter that needs to be adjusted in accordance with accuracy and stability properties. In particular, the given viscosity ν of INSE has to be recovered, whereas it does not appear in the continuous case described above. This is possible because of the numerical viscosity of the scheme. This approach to approximate solutions to the incompressible Navier-Stokes equations by the numerical solutions to a compressible system is the basis of the so called *artificial compressibility method*, [78, 79, 65, 66]. Here we obtain such a method, but with a particular structure that enables us to prove strong stability properties. Notice that our aim here is not to obtain a scheme for the compressible Euler equations with small Mach number. This issue could be considered also (see Subsection 7.3), but the method adopted in this paper would be of very low accuracy in this context where there must be no viscosity in the limit.

As explained in Subsection 3.1, the simultaneous limit when $\epsilon_{adv} = \sqrt{\epsilon_{coll}} \to 0$ of the two-scale BGK equation (3.1) is a hyperbolic/parabolic system on $\mathbf{W} = \int \mathbf{f} d\xi$. In this system the second-order terms cannot vanish, as can be seen on (3.3) and using the monotonicity of the Maxwellian. This is the same in the scaled Boltzmann equation (2.1) with $q = 1$, where the limit system on the moments (2.7) has viscosity coefficients depending on the collision operator.

In the time-space discrete case where the collision term is replaced by the projection onto Maxwellians at discrete times, the second-order viscous terms in the limit equations on the moments derive indeed from the numerical viscosity of the hydrodynamic scheme, see [16, Section 6], [4, 19, 29].

In order to get incompressible limit systems on the moments, the variable $\mathbf{W}$ has to be scaled, as is the case in (2.4) for the Boltzmann equation. We are now going to consider time-space discrete schemes for the incompressible Navier-Stokes equations. Then the idea is to replace the kinetic equation by the FVS formulation on the moments, that follows from Proposition 2.3. In this formulation, the consistency (2.42) has to be written with the fluxes $\mathbf{F}_j$ of the compressible isentropic gas dynamics system (3.5), (3.6), so that the scaling mechanism on the variable $\mathbf{W}$ acts as in the low Mach number limit of Subsection 3.2. With this approach we bypass the kinetic formalism, and it has the great advantage to simplify a lot the accuracy analysis.

4.1. Flux vector splitting schemes and low Mach number asymptotics

In this subsection we consider scaled general flux vector splitting schemes

$$\mathbf{W_i}^{n+1} - \mathbf{W_i}^n + \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^{D} (\mathbf{F}_{\mathbf{i}+\mathbf{e}_j/2}^n - \mathbf{F}_{\mathbf{i}-\mathbf{e}_j/2}^n) = \mathbf{0}, \tag{4.1}$$

$$\mathbf{F}_{\mathbf{i}+\mathbf{e}_j/2}^n = \mathbf{F}_j^+(\mathbf{W_i}^n) + \mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n), \tag{4.2}$$

where as in Subsection 2.4 the index n relates to the time t_n with $t_{n+1} - t_n = \Delta t$, and the multi-index $\mathbf{i} = (i_1, \ldots, i_D) \in \mathbb{Z}^D$ relates to the cell $C_\mathbf{i} \subset \mathbb{R}^D$ of a Cartesian mesh in $\mathbb{R}^D$ of constant length Δx in all directions. The flux location is $\mathbf{i} + \mathbf{e}_j/2 = (i_1, \ldots, i_j + 1/2, \ldots, i_D)$, where again $\mathbf{e}_j$ is the jth basis vector. The parameter $\epsilon > 0$ needs to be determined by consistency relations, but is anyway of the

order of Δx. The functions $\mathbf{F}_j^{\pm}$ must satisfy for all $j = 1, \ldots, D$

$$\mathbf{F}_j^+(\mathbf{W}) + \mathbf{F}_j^-(\mathbf{W}) = \mathbf{F}_j(\mathbf{W}). \tag{4.3}$$

In order to illustrate the possibilities of the approach, we can consider the case $\mathbf{F}_j = 0$, and $\mathbf{F}_j^+(\mathbf{W}) = -\mathbf{F}_j^-(\mathbf{W}) = \mathbf{B}(\mathbf{W})$. Then we obtain from (4.2) the expansion $\mathbf{F}_{\mathbf{i}+\mathbf{e}_j/2}^n \simeq -\Delta x \, \partial_j(\mathbf{B}(\mathbf{W}))$, thus the scheme is consistent with (3.4) for $\epsilon = \Delta x$. If $\mathbf{B}$ and $\mathbf{I} - 2D(\Delta t/\Delta x^2)\mathbf{B}$ are η−dissipative, Proposition 2.6 yields the entropy stability of the scheme, which is nothing else than the usual 3-point scheme for diffusion.

For the Navier-Stokes system we take now for $\mathbf{F}_j$ the flux of the isentropic gas dynamics system, given by (3.7), and the variable $\mathbf{W}$ is written as

$$\mathbf{W} = (\rho, \mathbf{m}) = (\rho, \epsilon \rho \mathbf{u}). \tag{4.4}$$

Proposition 4.1 (Stability of the FVS scheme for INSE). *Assume a decomposition of the isentropic fluxes (3.7) as (4.3), with the property that*

$$\mathbf{F}_j^+(\mathbf{W}), \quad -\mathbf{F}_j^-(\mathbf{W}), \quad \mathbf{W} - \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^D (\mathbf{F}_j^+(\mathbf{W}) - \mathbf{F}_j^-(\mathbf{W})) \quad \textit{are } \eta - \textit{dissipative in } \mathcal{W}_{stab} \tag{4.5}$$

for some set $\mathcal{W}_{stab}$, in the sense of (2.75) where η is the entropy (3.8) of the isentropic system. Then the scheme (4.1), (4.2) satisfies the discrete entropy inequality

$$\eta(\mathbf{W}_{\mathbf{i}}^{n+1}) - \eta(\mathbf{W}_{\mathbf{i}}^n) + \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^D (\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^n - \vartheta_{\mathbf{i}-\mathbf{e}_j/2}^n) \le 0, \tag{4.6}$$

with the numerical entropy fluxes

$$\vartheta_{\mathbf{i}+\mathbf{e}_j/2}^n = \vartheta_j^+(\mathbf{W}_{\mathbf{i}}^n) + \vartheta_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n), \tag{4.7}$$

with

$$(\vartheta_j^+)' = \eta'(\mathbf{F}_j^+)', \qquad (\vartheta_j^-)' = \eta'(\mathbf{F}_j^-)'. \tag{4.8}$$

It holds as soon as all the values $\mathbf{W}_{\mathbf{i}}^n$ and $\mathbf{W}_{\mathbf{i}}^{n+1}$ at time t_n and t_{n+1} remain in $\mathcal{W}_{stab}$. In particular, summing (4.6) over the cells gives the a priori estimate analogous to (3.15)

$$\sum_{\mathbf{i}} (\Delta x)^D E_{\mathbf{i}}^{n+1} \le \sum_{\mathbf{i}} (\Delta x)^D E_{\mathbf{i}}^n + \textit{ boundary terms}, \tag{4.9}$$

with the relative energy defined by

$$\begin{aligned} E_{\mathbf{i}}^n &= \frac{\eta(\mathbf{W}_{\mathbf{i}}^n) - \eta(\bar{\rho}, 0) - \eta'(\bar{\rho}, 0)(\mathbf{W}_{\mathbf{i}}^n - (\bar{\rho}, 0))}{\epsilon^2} \\ &= \rho_{\mathbf{i}}^n \frac{|\mathbf{u}_{\mathbf{i}}^n|^2}{2} + \frac{\rho_{\mathbf{i}}^n e(\rho_{\mathbf{i}}^n) - \bar{\rho} e(\bar{\rho}) - (e(\bar{\rho}) + P(\bar{\rho})/\bar{\rho})(\rho_{\mathbf{i}}^n - \bar{\rho})}{\epsilon^2}. \end{aligned} \tag{4.10}$$

Proof. The scheme can be written as in (2.78),

$$\mathbf{W}_{\mathbf{i}}^{n+1} = \left(\mathbf{W}_{\mathbf{i}}^n - \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^D (\mathbf{F}_j^+(\mathbf{W}_{\mathbf{i}}^n) - \mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}}^n)) \right) + \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^D (\mathbf{F}_j^+(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n) - \mathbf{F}_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n)). \tag{4.11}$$

Applying Proposition 2.6 we get the inequality

$$\eta(\mathbf{W}_{\mathbf{i}}^{n+1}) \le \eta(\mathbf{W}_{\mathbf{i}}^n) - \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^D (\vartheta_j^+(\mathbf{W}_{\mathbf{i}}^n) - \vartheta_j^-(\mathbf{W}_{\mathbf{i}}^n)) + \frac{\Delta t}{\epsilon \Delta x} \sum_{j=1}^D (\vartheta_j^+(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n) - \vartheta_j^-(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n)), \tag{4.12}$$

with the definition (4.8) of $\vartheta_j^{\pm}$. It can then be recast under the form (4.6), (4.7). ∎

The inequality (4.9) ensures in particular that $\mathbf{u}$ remains bounded and that $\rho - \overline{\rho} = O(\epsilon)$ as $\epsilon \to 0$, because we can take $\rho = \overline{\rho}$ at initial time, giving a bounded relative energy (but however $\rho - \overline{\rho}$ should be $O(\epsilon^2)$ in distribution sense, according to the low Mach number limit explained in Subsection 3.2). We now state a result on the accuracy of the scheme. It enables to relate the Navier-Stokes viscosity to the numerical viscosity of the FVS scheme.

Theorem 4.2 (Consistency and accuracy of the FVS scheme for INSE). *Given a decomposition of the isentropic fluxes (3.7) as (4.3), the scheme (4.1), (4.2) with the definition of the velocity $\mathbf{u}$ as (4.4), is*

(1) *Consistent with the incompressible Navier-Stokes equations*

$$\nabla_{\mathbf{x}} \cdot \mathbf{u} = 0, \tag{4.13}$$

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (\mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) - \nu \Delta_{\mathbf{x}} \mathbf{u} = \mathbf{0}, \tag{4.14}$$

with given viscosity $\nu > 0$, if for all $j = 1, \ldots, D$

$$\frac{\Delta x}{\epsilon} \partial_{\mathbf{m}} (\mathbf{F}_j^+ - \mathbf{F}_j^-)_{\mathbf{m}} (\overline{\rho}, 0) = 2\nu \, \mathbf{I}, \tag{4.15}$$

where the index $\mathbf{m}$ denotes the second component,

(2) *Second-order accurate in space if furthermore*

$$\partial_{\mathbf{m}} (\mathbf{F}_j^+ - \mathbf{F}_j^-)_\rho (\overline{\rho}, 0) = \mathbf{0}, \qquad \partial_\rho (\mathbf{F}_j^+ - \mathbf{F}_j^-)_{\mathbf{m}} (\overline{\rho}, 0) = \mathbf{0}, \qquad \partial_{\mathbf{mm}}^2 (\mathbf{F}_j^+ - \mathbf{F}_j^-)_{\mathbf{m}} (\overline{\rho}, 0) = \mathbf{0}, \tag{4.16}$$

where similarly the index ρ denotes the first component.

Notice that the stated first or second-order accuracy is with respect to both equations (4.13), (4.14), thus the incompressible constraint (4.13) is resolved with the same accuracy as (4.14). **Proof.** The relation (4.3) linking $\mathbf{F}_j^+$, $\mathbf{F}_j^-$ and the isentropic flux $\mathbf{F}_j$ given by (3.7) can be written

$$\mathbf{F}_j^+(\mathbf{W}) + \mathbf{F}_j^-(\mathbf{W}) = \begin{pmatrix} \mathbf{m}_j \\ \frac{\mathbf{m}_j \mathbf{m}}{\rho} + P(\rho)\mathbf{e}_j \end{pmatrix}, \tag{4.17}$$

with $\mathbf{W} = (\rho, \mathbf{m})$. We deduce that $\partial_{\mathbf{m}} (\mathbf{F}_j^+ + \mathbf{F}_j^-)_{\mathbf{m}} (\overline{\rho}, 0) = \mathbf{0}$. Thus the statement (4.15) can be written equivalently

$$\frac{\Delta x}{\epsilon} \partial_{\mathbf{m}} (\mathbf{F}_j^+)_{\mathbf{m}} (\overline{\rho}, 0) = \nu \, \mathbf{I}. \tag{4.18}$$

The idea of the proof of consistency is to follow (3.11)-(3.18), with the orders of magnitude (3.10). Therefore, it is enough to prove that the numerical fluxes in (4.1), divided by ϵ for the first component, and by ϵ^2 for the second component, are consistent with the fluxes in (3.13), (3.14). In doing so we look at the fluxes in each direction j separately. Thus for simplifying the notations we just indicate the index i related to this direction. We also remove the time index n. Then we have $\mathbf{F}_{i+1/2} =$

$\mathbf{F}_j^+(\mathbf{W}_i) + \mathbf{F}_j^-(\mathbf{W}_{i+1})$, and we compute with (4.17) and (4.4)

$$
\begin{aligned}
&\frac{1}{\epsilon}\mathbf{F}_{\rho,i+1/2}\\
&= \frac{1}{\epsilon}\left(\mathbf{F}_{j,\rho}^+(\mathbf{W}_i) + \mathbf{F}_{j,\rho}^-(\mathbf{W}_{i+1})\right)\\
&= \frac{(\mathbf{m}_{i+1})_j}{\epsilon} + \frac{1}{\epsilon}\left(\mathbf{F}_{j,\rho}^+(\mathbf{W}_i) - \mathbf{F}_{j,\rho}^+(\mathbf{W}_{i+1})\right) = \frac{(\mathbf{m}_i)_j}{\epsilon} + \frac{1}{\epsilon}\left(\mathbf{F}_{j,\rho}^-(\mathbf{W}_{i+1}) - \mathbf{F}_{j,\rho}^-(\mathbf{W}_i)\right)\\
&= \frac{\rho_i(\mathbf{u}_i)_j + \rho_{i+1}(\mathbf{u}_{i+1})_j}{2} + \frac{1}{2\epsilon}\left((\mathbf{F}_{j,\rho}^+ - \mathbf{F}_{j,\rho}^-)(\mathbf{W}_i) - (\mathbf{F}_{j,\rho}^+ - \mathbf{F}_{j,\rho}^-)(\mathbf{W}_{i+1})\right)\\
&= (\rho\mathbf{u}_j)(x_{i+1/2}) + \frac{1}{2\epsilon}\left(\partial_\rho(\mathbf{F}_{j,\rho}^+ - \mathbf{F}_{j,\rho}^-)(\overline{\rho},0)(\rho_i - \rho_{i+1}) + \partial_\mathbf{m}(\mathbf{F}_{j,\rho}^+ - \mathbf{F}_{j,\rho}^-)(\overline{\rho},0)(\epsilon\rho_i\mathbf{u}_i - \epsilon\rho_{i+1}\mathbf{u}_{i+1})\right)\\
&\hspace{10cm} + O(\Delta x^2)\\
&= (\rho\mathbf{u}_j)(x_{i+1/2}) - \frac{\overline{\rho}\Delta x}{2}\partial_\mathbf{m}(\mathbf{F}_{j,\rho}^+ - \mathbf{F}_{j,\rho}^-)(\overline{\rho},0)\partial_j\mathbf{u} + O(\Delta x^2),
\end{aligned}
$$
(4.19)

because ϵ is of the order of Δx, and where we used that $\rho_i - \rho_{i+1} = O(\epsilon^2\Delta x)$ because of (3.10). This approximation should be true for smooth solutions to the INSE because we expect then that $\rho - \overline{\rho} = \epsilon^2\widetilde{\rho}$ for some $\widetilde{\rho}$ uniformly smooth with respect to ϵ. We deduce from (4.19) the consistency of the first component of the fluxes with (3.13), and the second-order accuracy in the case $\partial_\mathbf{m}(\mathbf{F}_{j,\rho}^+ - \mathbf{F}_{j,\rho}^-)(\overline{\rho},0) = \mathbf{0}$. We next write the second component of the fluxes in a similar way

$$
\begin{aligned}
&\frac{1}{\epsilon^2}\mathbf{F}_{\mathbf{m},i+1/2}\\
&= \frac{1}{\epsilon^2}\left(\mathbf{F}_{j,\mathbf{m}}^+(\mathbf{W}_i) + \mathbf{F}_{j,\mathbf{m}}^-(\mathbf{W}_{i+1})\right)\\
&= \frac{\rho_i(\mathbf{u}_i)_j\,\mathbf{u}_i + \rho_{i+1}(\mathbf{u}_{i+1})_j\,\mathbf{u}_{i+1}}{2} + \frac{P(\rho_i) + P(\rho_{i+1})}{2\epsilon^2}\mathbf{e}_j\\
&\quad + \frac{1}{2\epsilon^2}\left((\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\mathbf{W}_i) - (\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\mathbf{W}_{i+1})\right).
\end{aligned}
$$
(4.20)

The first term gives obviously a second-order approximation of $\rho\mathbf{u}_j\mathbf{u}$. For the second term, we write

$$
\begin{aligned}
\frac{P(\rho_i) + P(\rho_{i+1})}{2\epsilon^2} &= \frac{P(\overline{\rho})}{\epsilon^2} + \frac{P(\rho_i) - P(\overline{\rho}) + P(\rho_{i+1}) - P(\overline{\rho})}{2\epsilon^2}\\
&= \frac{P(\overline{\rho})}{\epsilon^2} + P'(\overline{\rho})\frac{\rho_i - \overline{\rho} + \rho_{i+1} - \overline{\rho}}{2\epsilon^2} + O(\epsilon^2)\\
&= \frac{P(\overline{\rho})}{\epsilon^2} + P'(\overline{\rho})\widetilde{\rho}(x_{i+1/2}) + O(\Delta x^2),
\end{aligned}
$$
(4.21)

where we used again the ansatz $\rho - \overline{\rho} = \epsilon^2\widetilde{\rho}$, and the fact that ϵ is of the order of Δx. Setting $p = P'(\overline{\rho})\widetilde{\rho}/\overline{\rho}$, we get thus that the second term in the right-hand side of (4.20) is a second-order accurate approximation of the pressure term in INSE (the term $P(\overline{\rho})/\epsilon^2$ is a constant that disappears when taking the flux difference in (4.1)). The approximation is indeed the same as in the continuous case (3.16). Finally for the last term in the right-hand side of (4.20) we write

$$
\begin{aligned}
&(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\mathbf{W}_i) - (\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\mathbf{W}_{i+1})\\
&= \partial_\rho(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\overline{\rho},0)(\rho_i - \rho_{i+1}) + \partial_\mathbf{m}(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\overline{\rho},0)(\epsilon\rho_i\mathbf{u}_i - \epsilon\rho_{i+1}\mathbf{u}_{i+1})\\
&\quad + \frac{\epsilon^2\rho_i^2}{2}\partial_{\mathbf{mm}}^2(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\overline{\rho},0)\mathbf{u}_i \otimes \mathbf{u}_i - \frac{\epsilon^2\rho_{i+1}^2}{2}\partial_{\mathbf{mm}}^2(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\overline{\rho},0)\mathbf{u}_{i+1} \otimes \mathbf{u}_{i+1} + O(\Delta x^4).
\end{aligned}
$$
(4.22)

We can expand the factor in the main term as

$$
\epsilon\rho_i\mathbf{u}_i - \epsilon\rho_{i+1}\mathbf{u}_{i+1} = \epsilon\overline{\rho}(\mathbf{u}_i - \mathbf{u}_{i+1}) + \epsilon^3(\widetilde{\rho}_i\mathbf{u}_i - \widetilde{\rho}_{i+1}\mathbf{u}_{i+1}) = -\epsilon\overline{\rho}\Delta x(\partial_j\mathbf{u})(x_{i+1/2}) + O(\Delta x^4). \tag{4.23}
$$

Therefore we get with (4.15)

$$
\begin{aligned}
\frac{1}{2\epsilon^2}\partial_{\mathbf{m}}(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\overline{\rho},0)(\epsilon\rho_i\mathbf{u}_i - \epsilon\rho_{i+1}\mathbf{u}_{i+1}) & \\
= -\overline{\rho}\frac{\Delta x}{2\epsilon}\partial_{\mathbf{m}}(\mathbf{F}_{j,\mathbf{m}}^+ - \mathbf{F}_{j,\mathbf{m}}^-)(\overline{\rho},0)(\partial_j\mathbf{u})(x_{i+1/2}) + O(\Delta x^2) & \\
= -\overline{\rho}\,\nu\,(\partial_j\mathbf{u})(x_{i+1/2}) + O(\Delta x^2),
\end{aligned}
\tag{4.24}
$$

which means the second-order accurate consistency of this term with the viscous term in (4.14). The other terms in the right-hand side of (4.22) are $O(\Delta x^3)$, thus after division by $2\epsilon^2$ to evaluate the last term in (4.20) they do not contribute to the consistency since they give $O(\Delta x)$. Finally when the relations (4.16) hold these terms vanish, leading to the second-order accuracy of the full scheme. ■

4.2. The Lax-Friedrichs scheme for incompressible Navier-Stokes

With the analysis of the previous subsection we are now able to state our main result on the Lax-Friedrichs scheme, which is a particular flux vector splitting scheme. The result involves a sharp analysis of the η-dissipativity condition, that gives a more precise statement than monotonicity would give.

Theorem 4.3 (Lax-Friedrichs scheme for incompressible Navier-Stokes). *Consider the decomposition of the isentropic fluxes $\mathbf{F}_j$ defined in (3.7) as the sum of $\mathbf{F}_j^+$, $\mathbf{F}_j^-$ given by the Lax-Friedrichs decomposition (2.65) with parameter $c > 0$. Then the flux vector splitting scheme (4.1), (4.2) with the definition of the velocity $\mathbf{u}$ as (4.4) gives a second-order accurate approximation of the incompressible Navier-Stokes equations (4.13), (4.14) with given viscosity $\nu > 0$ if*

$$
\epsilon = \frac{c\Delta x}{2\nu}.
\tag{4.25}
$$

Moreover, the density ρ at time t_{n+1} is nonnegative under the conditions that the data at time t_n satisfy $\rho \geq 0$, $|\mathbf{m}_j|/\rho \leq c$, and the CFL condition

$$
2D\frac{\nu\Delta t}{\Delta x^2} \leq 1.
\tag{4.26}
$$

Then, assuming that $\rho \mapsto \rho\sqrt{P'(\rho)}$ is nondecreasing, and defining for any $c > 0$

$$
\mathcal{P}_{stab}^c = \Big\{ (\mathbf{W}, \widehat{\mathbf{W}}) \ with \ \mathbf{W} = (\rho, \mathbf{m}), \widehat{\mathbf{W}} = (\widehat{\rho}, \widehat{\mathbf{m}}), \ such \ that
$$
$$
\max(\frac{|\mathbf{m}_j|}{\rho}, \frac{|\widehat{\mathbf{m}}_j|}{\widehat{\rho}}) + \frac{\max(\rho\sqrt{P'(\rho)}, \widehat{\rho}\sqrt{P'(\widehat{\rho})})}{\widehat{\rho}} \leq c \ \ for \ j = 1,\dots,D \Big\},
\tag{4.27}
$$

the η-dissipativity properties of (4.5) (in the modified sense defined in Proposition 2.7) are satisfied under the CFL condition (4.26). It follows that the scheme satisfies the discrete entropy inequality (4.6) with the entropy (3.8) of the isentropic system and the numerical entropy fluxes (4.7), (2.70), (3.9) as long as the couples of data $(\mathbf{W}_{\mathbf{i}}^{n+1}, \mathbf{W}_{\mathbf{i'}}^n)$ remain in $\mathcal{P}_{stab}^c$ for all neighbors $\mathbf{i'}$ of $\mathbf{i}$, i.e. $\mathbf{i'} = \mathbf{i}$ or $\mathbf{i'} = \mathbf{i}\pm\mathbf{e}_j$ for some j.

Proof. With the Lax-Friedrichs decomposition (2.65) we have $\mathbf{F}_j^+(\mathbf{W}) - \mathbf{F}_j^-(\mathbf{W}) = c\mathbf{W}$. Thus, applying Theorem 4.2, the relations (4.15), (4.16) are satisfied using the relation (4.25), and this concludes the first part of the theorem.
Concerning the nonnegativity of ρ, we write the scheme as (4.11). We use again that $\mathbf{F}_j^+(\mathbf{W}) - \mathbf{F}_j^-(\mathbf{W}) = c\mathbf{W}$, which implies that the first term in (4.11) is $(1 - Dc\Delta t/(\epsilon\Delta x))\mathbf{W}_{\mathbf{i}}^n$. Taking the first

component of (4.11) we get that $\rho_{\mathbf{i}}^n \geq 0$ as soon as the CFL condition

$$D\frac{c\Delta t}{\epsilon\Delta x} \leq 1 \tag{4.28}$$

holds, and the first components of $\mathbf{F}_j^+$ and $-\mathbf{F}_j^-$ are nonnegative. With (4.25), the CFL condition (4.28) identifies with (4.26). According to (2.65), the nonnegativity of the first components of $\mathbf{F}_j^+$ and $-\mathbf{F}_j^-$ writes $c\rho \pm \mathbf{m}_j \geq 0$, which concludes the nonnegativity statement.

Concerning the entropy part of the theorem we only have to prove (4.5) (in the modified sense defined in Proposition 2.7) with the choice (4.27) of $\mathcal{P}_{stab}$ since then we can apply Proposition 4.1. We notice that the definition (4.27) implies the subcharacteristic condition (2.67), because $\mathbf{m}_j/\rho \pm \sqrt{P'(\rho)}$ are the eigenvalues of $\mathbf{F}_j'$. Thus the monotonicity of the quantities in (4.5) (nonnegativity of the eigenvalues of the derivative with respect to $\mathbf{W}$) is ensured under the CFL condition (4.28), that is (4.26) with the relation (4.25). Therefore, it remains to justify that not only the quantities in (4.5) are monotone, but they are also η-dissipative in $\mathcal{P}_{stab}^c$.

We recall that η is the entropy for the isentropic system, given by (3.8). The η-dissipation $\mathcal{D}[\varphi](\mathbf{W}, \widehat{\mathbf{W}})$ of a function $\mathbf{W} \mapsto \varphi(\mathbf{W})$ is defined by (2.75), with $q[\varphi]$ defined by (2.73). We thus have to prove that $\mathcal{D}[\varphi](\mathbf{W}, \widehat{\mathbf{W}}) \leq 0$ for all $(\mathbf{W}, \widehat{\mathbf{W}}) \in \mathcal{P}_{stab}^c$, for φ any of the functions in (4.5). Since $\mathcal{D}[]$ is linear in its argument, taking into account the CFL condition it means to prove it for all φ among the functions $\mathbf{W} \mapsto c\mathbf{W} + \mathbf{F}_j(\mathbf{W})$, $\mathbf{W} \mapsto c\mathbf{W} - \mathbf{F}_j(\mathbf{W})$ and $\mathbf{W} \mapsto \mathbf{W}$. We first compute

$$\begin{aligned}
\mathcal{D}[\mathbf{W} \mapsto \mathbf{W}](\mathbf{W}, \widehat{\mathbf{W}}) &= \eta(\mathbf{W}) - \eta(\widehat{\mathbf{W}}) - \eta'(\mathbf{W})(\mathbf{W} - \widehat{\mathbf{W}}) \\
&= \frac{1}{2}\frac{|\mathbf{m}|^2}{\rho} + \rho e(\rho) - \frac{1}{2}\frac{|\widehat{\mathbf{m}}|^2}{\widehat{\rho}} - \widehat{\rho}e(\widehat{\rho}) - \left(e(\rho) + \frac{P(\rho)}{\rho} - \frac{|\mathbf{m}|^2}{2\rho^2}\right)(\rho - \widehat{\rho}) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad - \frac{\mathbf{m}}{\rho} \cdot (\mathbf{m} - \widehat{\mathbf{m}}) \\
&= \widehat{\rho}(e(\rho) - e(\widehat{\rho})) - \frac{P(\rho)}{\rho}(\rho - \widehat{\rho}) - \frac{\widehat{\rho}}{2}\left|\frac{\mathbf{m}}{\rho} - \frac{\widehat{\mathbf{m}}}{\widehat{\rho}}\right|^2 .
\end{aligned} \tag{4.29}$$

This quantity is nonpositive (for the first part, differentiate with respect to $\widehat{\rho}$ and use that $P' > 0$). We then compute using the value of the entropy fluxes (3.9)

$$\begin{aligned}
\mathcal{D}[\mathbf{F}_j](\mathbf{W}, \widehat{\mathbf{W}}) &= \vartheta_j(\mathbf{W}) - \vartheta_j(\widehat{\mathbf{W}}) - \eta'(\mathbf{W})(\mathbf{F}_j(\mathbf{W}) - \mathbf{F}_j(\widehat{\mathbf{W}})) \\
&= \left(\frac{|\mathbf{m}|^2}{2\rho} + \rho e(\rho) + P(\rho)\right)\frac{\mathbf{m}_j}{\rho} - \left(\frac{|\widehat{\mathbf{m}}|^2}{2\widehat{\rho}} + \widehat{\rho}e(\widehat{\rho}) + P(\widehat{\rho})\right)\frac{\widehat{\mathbf{m}}_j}{\widehat{\rho}} \\
&\quad - \left(e(\rho) + \frac{P(\rho)}{\rho} - \frac{|\mathbf{m}|^2}{2\rho^2}\right)(\mathbf{m}_j - \widehat{\mathbf{m}}_j) - \frac{\mathbf{m}}{\rho} \cdot \left(\frac{\mathbf{m}\mathbf{m}_j}{\rho} + P(\rho)\mathbf{e}_j - \frac{\widehat{\mathbf{m}}\widehat{\mathbf{m}}_j}{\widehat{\rho}} - P(\widehat{\rho})\mathbf{e}_j\right) \\
&= (P(\widehat{\rho}) - P(\rho))\frac{\mathbf{m}_j}{\rho} + \left(e(\rho) + \frac{P(\rho)}{\rho} - e(\widehat{\rho}) - \frac{P(\widehat{\rho})}{\widehat{\rho}} - \frac{1}{2}\left|\frac{\mathbf{m}}{\rho} - \frac{\widehat{\mathbf{m}}}{\widehat{\rho}}\right|^2\right)\widehat{\mathbf{m}}_j \\
&= (P(\widehat{\rho}) - P(\rho))\left(\frac{\mathbf{m}_j}{\rho} - \frac{\widehat{\mathbf{m}}_j}{\widehat{\rho}}\right) + \left(\widehat{\rho}(e(\rho) - e(\widehat{\rho})) - \frac{P(\rho)}{\rho}(\rho - \widehat{\rho}) - \frac{\widehat{\rho}}{2}\left|\frac{\mathbf{m}}{\rho} - \frac{\widehat{\mathbf{m}}}{\widehat{\rho}}\right|^2\right)\frac{\widehat{\mathbf{m}}_j}{\widehat{\rho}}
\end{aligned} \tag{4.30}$$

Denoting by $\mathbf{I}$d the mapping $\mathbf{W} \mapsto \mathbf{W}$, we have therefore

$$\mathcal{D}[c\mathbf{I}\text{d} \pm \mathbf{F}_j](\mathbf{W}, \widehat{\mathbf{W}}) = \left(c \pm \frac{\widehat{\mathbf{m}}_j}{\widehat{\rho}}\right)\mathcal{D}[\mathbf{I}\text{d}](\mathbf{W}, \widehat{\mathbf{W}}) \pm (P(\widehat{\rho}) - P(\rho))\left(\frac{\mathbf{m}_j}{\rho} - \frac{\widehat{\mathbf{m}}_j}{\widehat{\rho}}\right). \tag{4.31}$$

Let us denote

$$m = \max(\rho\sqrt{P'(\rho)}, \widehat{\rho}\sqrt{P'(\widehat{\rho})}). \tag{4.32}$$

Since $(\mathbf{W}, \widehat{\mathbf{W}}) \in \mathcal{P}_{stab}^c$, one has $c \pm \widehat{\mathbf{m}}_j/\widehat{\rho} \geq m/\widehat{\rho}$. Thus we can estimate (4.31) as

$$\mathcal{D}[c\mathbf{Id} \pm \mathbf{F}_j](\mathbf{W}, \widehat{\mathbf{W}}) \leq \frac{m}{\widehat{\rho}} \mathcal{D}[\mathbf{Id}](\mathbf{W}, \widehat{\mathbf{W}}) + \frac{m}{2}\Big(\frac{\mathbf{m}_j}{\rho} - \frac{\widehat{\mathbf{m}}_j}{\widehat{\rho}}\Big)^2 + \frac{(P(\widehat{\rho}) - P(\rho))^2}{2m}$$

$$\leq m\Big(e(\rho) - e(\widehat{\rho}) - P(\rho)(\frac{1}{\widehat{\rho}} - \frac{1}{\rho}) + \frac{(P(\widehat{\rho}) - P(\rho))^2}{2m^2}\Big). \tag{4.33}$$

By the assumption of monotonicity of $\rho \mapsto \rho\sqrt{P'(\rho)}$, one has $m = \sup \rho'\sqrt{P'(\rho')}$ over all ρ' in the interval $[\rho, \widehat{\rho}]$. Then according to [21, Lemma 3.3], the right-hand side of (4.33) is nonpositive, concluding the desired inequality $\mathcal{D}[c\mathbf{Id} \pm \mathbf{F}_j](\mathbf{W}, \widehat{\mathbf{W}}) \leq 0$ for all $(\mathbf{W}, \widehat{\mathbf{W}}) \in \mathcal{P}_{stab}^c$. ∎

4.3. **The stability condition for the Lax-Friedrichs scheme**

The stability of the Lax-Friedrichs scheme for the incompressible Navier-Stokes equations is governed by Theorem 4.3. In the statement of this theorem, it is required that the couples of data $(\mathbf{W}_{\mathbf{i}}^{n+1}, \mathbf{W}_{\mathbf{i'}}^n)$ remain in $\mathcal{P}_{stab}^c$ for all $\mathbf{i'}$ neighbor of $\mathbf{i}$. In order to check this property, it would be necessary to evaluate $\mathbf{W}_{\mathbf{i}}^{n+1}$. This is possible in principle with the update formula (4.11), this would lead to an estimate of $\mathbf{W}_{\mathbf{i}}^{n+1}$ in terms of $\mathbf{W}_{\mathbf{i}}^n$ and the differences $\mathbf{W}_{\mathbf{i'}}^n - \mathbf{W}_{\mathbf{i}}^n$ for $\mathbf{i'}$ neighbor of $\mathbf{i}$. Then this estimate has to be involved in the definition of $\mathcal{P}_{stab}^c$. However, in order to simplify, we shall now just replace $\mathbf{W}_{\mathbf{i}}^{n+1}$ by $\mathbf{W}_{\mathbf{i}}^n$. Thus we just require that the couples of neighbors $(\mathbf{W}_{\mathbf{i}}^n, \mathbf{W}_{\mathbf{i'}}^n)$ remain in $\mathcal{P}_{stab}^c$. Looking at the definition (4.27) of $\mathcal{P}_{stab}^c$, it implies in particular that the density must be away from 0, because of the ratio $\rho/\widehat{\rho}$ that appears in the case $\rho > \widehat{\rho}$. Apart from this fact, if we neglect again the difference $\mathbf{W}_{\mathbf{i'}}^n - \mathbf{W}_{\mathbf{i}}^n$, the property to be in $\mathcal{P}_{stab}^c$ reduces simply to the subcharacteristic condition

$$\frac{|\mathbf{m}_j|}{\rho} + \sqrt{P'(\rho)} \leq c \ \text{ for } j = 1, \dots, D, \tag{4.34}$$

with

$$\mathbf{W} = (\rho, \mathbf{m}) = (\rho, \epsilon\rho\mathbf{u}). \tag{4.35}$$

But since the true variable is $\mathbf{u}$, replacing $\mathbf{m}$ by $\epsilon\rho\mathbf{u}$ in (4.34) yields the effective subcharacteristic stability condition

$$\epsilon|\mathbf{u}_j| + \sqrt{P'(\rho)} \leq c \ \text{ for } j = 1, \dots, D. \tag{4.36}$$

Taking the supremum over all the values $\mathbf{i} \in \mathbb{Z}^D$, and knowing that ϵ and c are related by (4.25), it gives

$$c\frac{\Delta x}{2\nu} \sup_{\mathbf{i}} \max_{j=1,\dots,D} |(\mathbf{u}_j)_{\mathbf{i}}^n| + \sup \sqrt{P'(\rho_{\mathbf{i}}^n)} \leq c. \tag{4.37}$$

Therefore we need the inequality

$$\frac{\Delta x}{2\nu} \sup_{\mathbf{i}} \max_{j=1,\dots,D} |(\mathbf{u}_j)_{\mathbf{i}}^n| < 1, \tag{4.38}$$

and then (4.37) can be rewritten

$$\sup_{\mathbf{i}} \sqrt{P'(\rho_{\mathbf{i}}^n)} \leq c\Big(1 - \frac{\Delta x}{2\nu} \sup_{\mathbf{i}} \max_{j=1,\dots,D} |(\mathbf{u}_j)_{\mathbf{i}}^n|\Big). \tag{4.39}$$

The left-hand side of (4.38) can be interpreted as a *discrete cell Reynolds number*, related to the size Δx of the grid. Thus the condition says that at the level of the grid, the advection term $\mathbf{u} \cdot \nabla\mathbf{u}$ of the Navier-Stokes equations is dominated by the viscous term $\nu\Delta\mathbf{u}$. This is the reason why the CFL condition (4.26) involves only the viscosity, and not the velocity $\mathbf{u}$. The cell Reynolds stability

23

condition (4.38) thus plays an important role in the stability of the method. It needs a sufficiently small grid size Δx to be satisfied.

Taking into account that the sound speed in the ϵ-scaled isentropic system (3.13), (3.14) is of the order of $1/\epsilon$, and using the relation (4.25), the discrete cell Reynolds number (left-hand side of (4.38)) can also be interpreted as a numerical Mach number, that tends to zero with the space step Δx.

5. **Numerical settings**

In this section we explain how we implement the BGK-FVS method described in the previous sections, and more particularly in Theorem 4.3. The aim is to solve the classical incompressible Navier-Stokes equations

$$
\begin{aligned}
\nabla_{\mathbf{x}} \cdot \mathbf{u} &= 0, \\
\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot (\mathbf{u} \otimes \mathbf{u} + p\mathbf{I}) - \nu \Delta_{\mathbf{x}} \mathbf{u} &= \mathbf{0},
\end{aligned}
\tag{5.1}
$$

where $t > 0$ is the time, $\mathbf{x} \in \mathbb{R}^D$ the position vector, $\mathbf{u} \in \mathbb{R}^d$ the velocity vector, $\nu > 0$ the kinematic viscosity, and p the kinematic pressure. The associated energy equation is

$$
\partial_t \left(\frac{1}{2} |\mathbf{u}|^2 \right) + \nabla_{\mathbf{x}} \cdot \left(\left(\frac{1}{2} |\mathbf{u}|^2 + p \right) \mathbf{u} - \nu \nabla_{\mathbf{x}} \left(\frac{1}{2} |\mathbf{u}|^2 \right) \right) = -\nu |\nabla_{\mathbf{x}} \mathbf{u}|^2.
\tag{5.2}
$$

5.1. **Discrete scheme and scaling laws**

The space discretization is performed by using a uniform Cartesian grid of cell length Δx. The cell vector index is denoted by $\mathbf{i} = (i_1, \ldots, i_D) \in \mathbb{Z}^D$. The time is indexed by n, corresponding to the time t_n, with the relation $t_{n+1} - t_n = \Delta t$, with Δt the timestep.

The scheme uses a $2D$ points stencil including the neighbors of the cell in the directions of the discrete velocities (2.41), as illustrated on Figures 2.1 and 2.2. It takes the form of a flux vector splitting scheme (FVS) written in a variable $\mathbf{W} = (\rho, \mathbf{m})$, with $\rho > 0$ a density and $\mathbf{m} \in \mathbb{R}^D$ a momentum. It can be written as (4.1), (4.2) or (4.11) with $\mathbf{F}_j^+$, $\mathbf{F}_j^-$ given by the Lax-Friedrichs decomposition (2.65) with $\mathbf{F}_j(\mathbf{W})$ the isentropic flux (3.7) associated to an increasing pressure law $\rho \mapsto P(\rho)$, and $\epsilon > 0$ (magnitude of the Mach number), $c > 0$ (isentropic speed) are parameters related by

$$
\epsilon = \frac{c \Delta x}{2\nu}.
\tag{5.3}
$$

Using these definitions, the scheme can be written

$$
\mathbf{W}_{\mathbf{i}}^{n+1} = \left(1 - 2D \frac{\nu \Delta t}{\Delta x^2} \right) \mathbf{W}_{\mathbf{i}}^n + \frac{\nu \Delta t}{\Delta x^2} \sum_{j=1}^{D} \left(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n + \frac{1}{c} \mathbf{F}_j(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n) + \mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n - \frac{1}{c} \mathbf{F}_j(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n) \right),
\tag{5.4}
$$

with $\mathbf{e}_j$ the jth basis vector, i.e. $\mathbf{i} \pm \mathbf{e}_j = (i_1, \ldots, i_j \pm 1, \ldots, i_D)$. The moment variable $\mathbf{W}$ is related to the velocity $\mathbf{u}$ by

$$
\mathbf{W} = (\rho, \mathbf{m}) = (\rho, \epsilon \rho \mathbf{u}).
\tag{5.5}
$$

We can simplify the scheme (5.4) by expressing $\mathbf{W}$ in terms of ρ and $\mathbf{u}$ according to (5.5). Taking into account the relation (5.3) between ϵ and c and the expression of the fluxes (3.7), it yields

$$
\rho_{\mathbf{i}}^{n+1} = \left(1 - 2D \frac{\nu \Delta t}{\Delta x^2} \right) \rho_{\mathbf{i}}^n + \frac{\nu \Delta t}{\Delta x^2} \sum_{j=1}^{D} \left(\rho_{\mathbf{i}-\mathbf{e}_j}^n + \frac{\Delta x}{2\nu} (\rho \mathbf{u}_j)_{\mathbf{i}-\mathbf{e}_j}^n + \rho_{\mathbf{i}+\mathbf{e}_j}^n - \frac{\Delta x}{2\nu} (\rho \mathbf{u}_j)_{\mathbf{i}+\mathbf{e}_j}^n \right),
$$

$$
\begin{aligned}
(\rho \mathbf{u})_{\mathbf{i}}^{n+1} = \left(1 - 2D \frac{\nu \Delta t}{\Delta x^2} \right) (\rho \mathbf{u})_{\mathbf{i}}^n + \frac{\nu \Delta t}{\Delta x^2} \sum_{j=1}^{D} \Big(&(\rho \mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n + \frac{\Delta x}{2\nu} (\rho \mathbf{u}_j \mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n + \frac{2\nu}{\Delta x} \frac{P(\rho_{\mathbf{i}-\mathbf{e}_j}^n)}{c^2} \mathbf{e}_j \\
&+ (\rho \mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n - \frac{\Delta x}{2\nu} (\rho \mathbf{u}_j \mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n - \frac{2\nu}{\Delta x} \frac{P(\rho_{\mathbf{i}+\mathbf{e}_j}^n)}{c^2} \mathbf{e}_j \Big).
\end{aligned}
\tag{5.6}
$$

According to the expansions in the proof of Theorem 4.2, a second-order approximation of the pressure p of INSE is

$$p = \frac{P(\rho) - P(\overline{\rho})}{\overline{\rho}\epsilon^2} = \left(\frac{2\nu}{\Delta x}\right)^2 \frac{P(\rho) - P(\overline{\rho})}{c^2 \overline{\rho}}. \tag{5.7}$$

The scheme (5.6) can be also written in conservative form related to (4.1), as

$$\rho_{\mathbf{i}}^{n+1} - \rho_{\mathbf{i}}^n + \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} \left(\mathcal{F}_j^{\rho}((\rho, \rho\mathbf{u})_{\mathbf{i}}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n) - \mathcal{F}_j^{\rho}((\rho, \rho\mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}}^n) \right) = 0,$$

$$(\rho\mathbf{u})_{\mathbf{i}}^{n+1} - (\rho\mathbf{u})_{\mathbf{i}}^n + \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} \left(\mathcal{F}_j^{\rho\mathbf{u}}((\rho, \rho\mathbf{u})_{\mathbf{i}}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n) - \mathcal{F}_j^{\rho\mathbf{u}}((\rho, \rho\mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}}^n) \right) = \mathbf{0},$$
$$\tag{5.8}$$

with

$$\mathcal{F}_j^{\rho}((\rho, \rho\mathbf{u})_l, (\rho, \rho\mathbf{u})_r) = -\nu\frac{\rho_r - \rho_l}{\Delta x} + \frac{(\rho\mathbf{u}_j)_l + (\rho\mathbf{u}_j)_r}{2},$$

$$\mathcal{F}_j^{\rho\mathbf{u}}((\rho, \rho\mathbf{u})_l, (\rho, \rho\mathbf{u})_r) = -\nu\frac{(\rho\mathbf{u})_r - (\rho\mathbf{u})_l}{\Delta x} + \frac{(\rho\mathbf{u}_j\mathbf{u})_l + (\rho\mathbf{u}_j\mathbf{u})_r}{2} \tag{5.9}$$

$$+ \left(\frac{2\nu}{\Delta x}\right)^2 \frac{1}{c^2} \left(\frac{P(\rho_l) + P(\rho_r)}{2} - P(\overline{\rho})\right)\mathbf{e}_j.$$

It is easy to see on (5.8), (5.9) that the scheme is second-order accurate in space, since the derivatives are discretized by centered differences. The relation (5.3) ensures that the scaled numerical viscosity of the scheme becomes the physical viscosity ν.

The formulation (5.6) is extremely good for getting fast computer execution with the least possible operations, especially for parallel computers. In comparison to LBM we do not store distribution function values, but only the $D+1$ moments. The memory footprint is then quite low compared to LBM that use commonly 9 discrete velocities in two dimensions.

The scheme (5.8), (5.9) falls indeed into the class of artificial compressibility methods, that are described for example in [78, 79, 65, 66]. The one we propose is the only one that has a discrete entropy inequality proved for all data.

The discrete entropy inequality from Theorem 4.3 can be written in the same spirit as (5.4),

$$\eta(\mathbf{W}_{\mathbf{i}}^{n+1}) \le \left(1 - 2D\frac{\nu\Delta t}{\Delta x^2}\right)\eta(\mathbf{W}_{\mathbf{i}}^n) + \frac{\nu\Delta t}{\Delta x^2} \sum_{j=1}^{D} \left(\eta(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n) + \frac{1}{c}\vartheta_j(\mathbf{W}_{\mathbf{i}-\mathbf{e}_j}^n) + \eta(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n) - \frac{1}{c}\vartheta_j(\mathbf{W}_{\mathbf{i}+\mathbf{e}_j}^n) \right),$$
$$\tag{5.10}$$

with the entropy η given in (3.8) and entropy fluxes ϑ_j given in (3.9). It can be rewritten in the conservative form (4.6) as

$$\eta(\mathbf{W}_{\mathbf{i}}^{n+1}) - \eta(\mathbf{W}_{\mathbf{i}}^n) + \frac{\Delta t}{\Delta x} \sum_{j=1}^{D} \left(\mathcal{G}_j((\rho, \rho\mathbf{u})_{\mathbf{i}}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n) - \mathcal{G}_j((\rho, \rho\mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}}^n) \right) \le 0, \tag{5.11}$$

with

$$\mathcal{G}_j((\rho, \rho\mathbf{u})_l, (\rho, \rho\mathbf{u})_r) = -\nu\frac{\eta(\mathbf{W}_r) - \eta(\mathbf{W}_l)}{\Delta x} + \frac{\nu}{c\Delta x}(\vartheta_j(\mathbf{W}_r) + \vartheta_j(\mathbf{W}_l)). \tag{5.12}$$

Still with the relation (5.5) between $\mathbf{W}$ and $(\rho, \rho\mathbf{u})$ we define then the relative entropy

$$\eta_R(\rho, \rho\mathbf{u}) = \frac{\eta(\mathbf{W}) - \overline{\rho}e(\overline{\rho}) - (e(\overline{\rho}) + P(\overline{\rho})/\overline{\rho})(\rho - \overline{\rho})}{\epsilon^2}$$

$$= \rho\frac{|\mathbf{u}|^2}{2} + \left(\frac{2\nu}{\Delta x}\right)^2 \frac{\rho e(\rho) - \overline{\rho}e(\overline{\rho}) - (e(\overline{\rho}) + P(\overline{\rho})/\overline{\rho})(\rho - \overline{\rho})}{c^2}. \tag{5.13}$$

Taking into account the mass conservation in (5.8), the inequality (5.11) yields

$$\text{Diss}_{\mathbf{i}}^n \equiv \frac{\eta_R((\rho, \rho\mathbf{u})_{\mathbf{i}}^{n+1}) - \eta_R((\rho, \rho\mathbf{u})_{\mathbf{i}}^n)}{\Delta t} + \frac{1}{\Delta x}\sum_{j=1}^{D}\Big(\mathcal{G}_j^R((\rho, \rho\mathbf{u})_{\mathbf{i}}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n) \\ -\mathcal{G}_j^R((\rho, \rho\mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}}^n)\Big) \le 0, \tag{5.14}$$

where with (5.12), (5.9) and (3.9),

$$\begin{aligned}\mathcal{G}_j^R((\rho, \rho\mathbf{u})_l, (\rho, \rho\mathbf{u})_r) &= \frac{1}{\epsilon^2}\Big(\mathcal{G}_j((\rho, \rho\mathbf{u})_l, (\rho, \rho\mathbf{u})_r) - (e(\overline{\rho}) + \frac{P(\overline{\rho})}{\overline{\rho}})\mathcal{F}_j^\rho((\rho, \rho\mathbf{u})_l, (\rho, \rho\mathbf{u})_r)\Big)\\ &= -\nu\frac{\eta_R((\rho, \rho\mathbf{u})_r) - \eta_R((\rho, \rho\mathbf{u})_l)}{\Delta x} + \frac{\nu}{c\epsilon^2\Delta x}(\vartheta_j(\mathbf{W}_r) + \vartheta_j(\mathbf{W}_l))\\ &\quad - \frac{1}{\epsilon^2}(e(\overline{\rho}) + \frac{P(\overline{\rho})}{\overline{\rho}})\frac{(\rho\mathbf{u}_j)_l + (\rho\mathbf{u}_j)_r}{2}\\ &= -\nu\frac{\eta_R((\rho, \rho\mathbf{u})_r) - \eta_R((\rho, \rho\mathbf{u})_l)}{\Delta x} + \frac{1}{4}(\rho|\mathbf{u}|^2\mathbf{u}_j)_l + \frac{1}{4}(\rho|\mathbf{u}|^2\mathbf{u}_j)_r\\ &\quad + \frac{1}{2}\Big(\frac{2\nu}{c\Delta x}\Big)^2\Big((e(\rho) + \frac{P(\rho)}{\rho} - e(\overline{\rho}) - \frac{P(\overline{\rho})}{\overline{\rho}})\rho\mathbf{u}_j\Big)_l\\ &\quad + \frac{1}{2}\Big(\frac{2\nu}{c\Delta x}\Big)^2\Big((e(\rho) + \frac{P(\rho)}{\rho} - e(\overline{\rho}) - \frac{P(\overline{\rho})}{\overline{\rho}})\rho\mathbf{u}_j\Big)_r.\end{aligned} \tag{5.15}$$

The entropy inequality (5.14) is the discrete form of (3.15), and is consistent at second-order with (5.2) where the right-hand side is neglected. As in the continuous case, neglecting the boundary terms it implies by summing up over the cells and timesteps that $\sum_{\mathbf{i}} \Delta x^D \eta_R((\rho, \rho\mathbf{u})_{\mathbf{i}}^n) \le \sum_{\mathbf{i}} \Delta x^D \eta_R((\rho, \rho\mathbf{u})_{\mathbf{i}}^0$. With (5.13) we obtain that $\rho|\mathbf{u}|^2$ is bounded in L^1, and (since $\rho e(\rho) - \overline{\rho}e(\overline{\rho}) - (e(\overline{\rho}) + P(\overline{\rho})/\overline{\rho})(\rho - \overline{\rho}) \simeq P'(\overline{\rho})(\rho - \overline{\rho})^2/2\overline{\rho}$ as $\rho \to \overline{\rho}$) that $\rho - \overline{\rho}$ is of the order of Δx in L^2. Note that this holds for all weak solutions, whereas the stronger convergence (3.10) i.e. $\rho - \overline{\rho} = O(\Delta x^2)$ is true only for regular enough solutions.

5.2. **Stability and parameter settings**

The originality of our BGK-FVS method with respect to finite element methods is that we have a scalar variable ρ which is related to the pressure, but that obeys an evolution equation. Both unknowns ρ and $\rho\mathbf{u}$ are treated explicitly, and we have to respect the CFL condition

$$2D\frac{\nu\Delta t}{\Delta x^2} \le 1. \tag{5.16}$$

In practice we take equality in (5.16). Note that because of the CFL condition, Δt is of the order of Δx^2, thus the first-order Euler time stepping is enough accurate, since a first-order time error in Δt is of the same order of magnitude as a second-order space error Δx^2. In practice, Δt is chosen by taking equality in (5.16).

The difference with lattice Boltzmann schemes is that in our scheme the kinetic variable has been eliminated, keeping only the moments $\rho, \mathbf{m}$ of the distribution function. The space discretization is extremely simple (Lax-Friedrichs scheme, that can be interpreted as a kinetic scheme with $2D$ velocities) but achieves second-order accuracy. The moment $\mathbf{W}$ is a perturbation of the equilibrium value $(\overline{\rho}, 0)$, where $\overline{\rho} > 0$ is a parameter. The initialization is performed by setting $\mathbf{u}$ to $\mathbf{u}^0$, with $\mathbf{u}^0$ the initial data for INSE. We take $\mathbf{u}^0 = \mathbf{0}$ for a steady problem. The initial density is taken either $\rho^0 = \overline{\rho}$, or defined through the pressure via (5.7) if there is a guess of the initial pressure.

The discrete entropy inequality (5.10) or (5.14) holds under the CFL condition (5.16), the cell Reynolds condition

$$\mathrm{Re}_{\mathrm{cell}} \equiv \frac{\Delta x}{2\nu} \sup_{\mathbf{i}} \max_{j=1,\dots,D} |(\mathbf{u}_j)_{\mathbf{i}}^n| < 1, \tag{5.17}$$

and the subcharacteristic condition

$$\frac{\sup_{\mathbf{i}} \sqrt{P'(\rho_{\mathbf{i}}^n)}}{1 - \frac{\Delta x}{2\nu} \sup_{\mathbf{i}} \max_{j=1,\dots,D} |(\mathbf{u}_j)_{\mathbf{i}}^n|} \leq c, \tag{5.18}$$

where the suprema are taken over all the values $\mathbf{i} \in \mathbb{Z}^D$. We recall that the cell Reynolds number $\mathrm{Re}_{\mathrm{cell}}$ can also be interpreted as a numerical Mach number. One can indeed check directly on (5.6) the monotonicity with respect to each variable $(\rho, \rho\mathbf{u})_{\mathbf{i}}^n$, $(\rho, \rho\mathbf{u})_{\mathbf{i}-\mathbf{e}_j}^n$, $(\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n$ under the conditions (5.16), (5.17), (5.18).

Since ρ is close to $\overline{\rho}$, only the linearization of the pressure law $P(\rho)$ around $\overline{\rho}$ matters. Therefore we make the most simple choice of isothermal law

$$P(\rho) = c_s^2\, \rho, \tag{5.19}$$

for some sound speed $c_s > 0$. It gives $e(\rho) = c_s^2 \log(\rho/\overline{\rho})$. Then (5.18) gives a limitation on c_s/c. With the choice (5.19), all terms in the scheme (5.6) are proportional to ρ. Thus we can factorize the constant $\overline{\rho}$ everywhere, so that its choice does not modify the scheme. In practice we take $\overline{\rho} = 1$.

On the formulas (5.6), (5.7), and also on (5.9), we observe with the choice of P given by (5.19) that the scheme depends only on the ratio c_s/c, and not on c_s and c separately. Thus to finish parameter settings we just have to give the value of c_s/c. This is done by computing an evaluation u_{max} of the maximum value over all $t, \mathbf{x}$ of $\max_{j=1,\dots,D} |\mathbf{u}_j(t, \mathbf{x})|$. Then to satisfy (5.18) and taking into account (5.19) we take simply

$$\frac{c_s}{c} = 1 - \frac{\Delta x}{2\nu} u_{max}. \tag{5.20}$$

The value u_{max} can be computed for example by considering only the initial data. In cases when it is not easy to have a good value for u_{max}, another method is to take just $c_s/c = 1$, see Subsubsection 6.1.3. Then the subcharacteristic condition (5.18) is violated, and in general it can be violated when u_{max} is not well chosen. Our experience shows (see Section 6) that in fact it is not mandatory to satisfy the subcharacteristic condition in practice. When $c_s/c = 1$ the overshoot in the subcharacteristic condition (5.18) is directly related to the value of the cell Reynolds number $\mathrm{Re}_{\mathrm{cell}}$ defined in (5.17), the larger $\mathrm{Re}_{\mathrm{cell}}$ is, the more the subchracteristic condition is overpassed. Related diagnostics are defined in Subsection 5.4.

5.3. **Boundary conditions**

As the scheme is using only the moments, and not discrete velocities as in LBM, we can set boundary values as in finite difference methods. They are applied as follows, at formal second-order of accuracy.

5.3.1. *Dirichley boundary conditions*

For applying Dirichley boundary conditions we use the standard *ghost cell method*, which means to consider an extra cell outside the domain, to which we assign some well chosen value. This means the following in 1d. Consider two cells identified by their centers x_1, x_2, with $x_2 - x_1 = \Delta x$, and with a boundary located at $x_1 - \Delta x/2$ (see Figure 5.1).
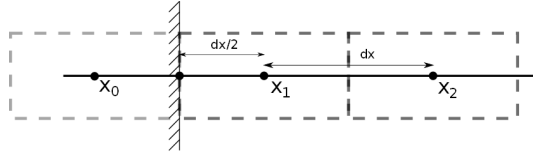
27

FIGURE 5.1. One-dimensional picture of the boundary, the unknowns are represented by dots.

Then we consider a ghost cell with center $x_0 = x_1 - \Delta x$. The scheme (5.4) or (5.6) is then applied in order to get the new value $\mathbf{W}_1^{n+1}$ at x_1, thus involving the values $\mathbf{W}_1^n$, $\mathbf{W}_2^n$, $\mathbf{W}_0^n$. The ghost value $\mathbf{W}_0^n$ is set in order to get second-order accuracy, which is $(\mathbf{W}_0^n + \mathbf{W}_1^n)/2 = \mathbf{W}_{des}$, where $\mathbf{W}_{des}$ is the desired value at the boundary. This gives the ghost value

$$\mathbf{W}_0^n = 2\mathbf{W}_{des} - \mathbf{W}_1^n. \tag{5.21}$$

This algorithm can be applied either in the variable $\mathbf{W}$, or in the variable $(\rho, \rho\mathbf{u})$ for the formulation (5.6). The value of ρ is converted by (5.7) to a pressure value.

For a velocity inlet condition, we give the velocity value $\mathbf{u}_{des}$ at the boundary, and apply the formula (5.21). The pressure at the boundary is not known, thus we simply set its value by extrapolation from the values in the domain, which means setting a ghost value $\mathbf{W}_0^n = 2\mathbf{W}_1^n - \mathbf{W}_2^n$. Thus we define

$$\rho_0^n = 2\rho_1^n - \rho_2^n, \qquad \rho_0^n \mathbf{u}_0^n = 2\overline{\rho}\,\mathbf{u}_{des} - \rho_1^n \mathbf{u}_1^n, \tag{5.22}$$

which means indeed to impose the momentum instead of the velocity. The no-slip condition corresponds to the case $\mathbf{u}_{des} = \mathbf{0}$.

5.3.2. *Neumann boundary conditions*

In order to apply Neumann boundary conditions $\mathbf{n} \cdot \nabla_{\mathbf{x}}\mathbf{W} = 0$ where $\mathbf{n}$ is the normal at the boundary, we use the ghost cell method as above, where we set the ghost value as $\mathbf{W}_0^n = \mathbf{W}_1^n$. Combinations of Dirichlet and Neumann are possible also, like setting Dirichlet conditions for some components of $\mathbf{W}$, and Neumann to the other components. In this way we can treat the slip conditions, which correspond to setting zero Dirichlet condition to the normal velocity $\mathbf{u} \cdot \mathbf{n}$, and Neumann conditions to the other components of the velocity (the pressure is extrapolated).

5.3.3. *Imposed normal flux*

For imposing the normal flux on the boundary, we use that the scheme is conservative, i.e. it takes the form (5.8). Then we just replace the numerical flux $\mathcal{F}_j^{\rho\mathbf{u}}((\rho, \rho\mathbf{u})_{\mathbf{i}}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n)$ in (5.8), where j corresponds to the direction normal to the boundary (i.e. $\mathbf{n} = \mathbf{e}_j$), by the desired value.

5.3.4. *Pressure outlet*

We have tested two kind of pressure outlet formulations. The first is very naive, we compute by extrapolation the values of the momentum, and the pressure is imposed as $p = p_o$. The second one is known as the pressure drop problem (e.g. [60]), and consists in solving this equation at the boundary:

$$-\nu\nabla_{\mathbf{x}}\mathbf{u}\,\mathbf{n} + p\mathbf{n} = p_o\mathbf{n}, \tag{5.23}$$

with p_o the imposed pressure value, $\mathbf{n}$ the normal at the boundary. This condition is implemented by imposing the normal flux, as explained in the previous subsection. Indeed we keep the transport term in $\rho\mathbf{u}_j\mathbf{u}$ in the formula (5.9) of $\mathcal{F}_j^{\rho\mathbf{u}}((\rho, \rho\mathbf{u})_{\mathbf{i}}^n, (\rho, \rho\mathbf{u})_{\mathbf{i}+\mathbf{e}_j}^n)$, and replace the other terms by $\overline{\rho}p_o\mathbf{e}_j$. The

transport terms in the density and momentum fluxes are computed by using extrapolated values of ρ and $\rho\mathbf{u}$.

Both formulations give similar results.

5.4. **Diagnostics**

In order to evaluate the performance of the method or to study the behavior of various quantities, we shall use some quantitative indicators. Those will be classified in three categories.

- Entropy indicators. The first one is the local dissipation $\text{Diss}_\mathbf{i}^n$ of the relative entropy, which is defined by (5.14). It is an approximation of the left-hand side of (5.2), hence also an approximation of the right-hand side, the dissipation $-\nu\left|\nabla_\mathbf{x}\mathbf{u}\right|^2$. The local dissipation $\text{Diss}_\mathbf{i}^n$ can thus be compared to a discrete evaluation $-\nu(|\nabla(\rho\mathbf{u})|_\mathbf{i}^n)^2$ of this dissipation, with

$$|\nabla(\rho\mathbf{u})|_\mathbf{i}^n = \left(\sum_{j'=1}^D \sum_{j=1}^D \left|\frac{(\rho\mathbf{u}_{j'})_{\mathbf{i}+\mathbf{e}_j}^n - (\rho\mathbf{u}_{j'})_{\mathbf{i}-\mathbf{e}_j}^n}{2\Delta x}\right|^2\right)^{1/2}. \tag{5.24}$$

We compute thus the relative L^1 error

$$\text{ErrDiss}^n = \frac{\sum_\mathbf{i}\left|\text{Diss}_\mathbf{i}^n + \nu(|\nabla(\rho\mathbf{u})|_\mathbf{i}^n)^2\right|}{\sum_\mathbf{i}|\text{Diss}_\mathbf{i}^n|}. \tag{5.25}$$

The second entropy indicator is the maximum value of $\text{Diss}_\mathbf{i}^n$,

$$\text{EntMax}^n = \max_\mathbf{i} \text{Diss}_\mathbf{i}^n. \tag{5.26}$$

According to (5.14) it should always be nonpositive, and checking this property is a way to verify the validity of our stability analysis.

- Incompressibility indicators. We compute the local discrete divergence of the velocity field and the associated time dependent relative $L1$ error, as

$$\text{Div}_\mathbf{i}^n = \sum_{j=1}^D \frac{(\rho\mathbf{u}_j)_{\mathbf{i}+\mathbf{e}_j}^n - (\rho\mathbf{u}_j)_{\mathbf{i}-\mathbf{e}_j}^n}{2\Delta x}, \quad \text{ErrDiv}^n = \frac{\sum_\mathbf{i}|\text{Div}_\mathbf{i}^n|}{\sum_\mathbf{i}|\nabla(\rho\mathbf{u})|_\mathbf{i}^n}. \tag{5.27}$$

- Stiffness indicator. This indicator is just the cell Reynolds number Re_cell defined in (5.17), related to the grid size, the viscosity ν and the velocity. It needs to be less than one for stability, and getting close to this value indicates a stiff problem for which it is numerically difficult to find the solution. The cell Reynolds number Re_cell can also be interpreted as an accuracy indicator since it can be written $\text{Re}_\text{cell} = \text{Re} \times \Delta x/L$. It is thus proportional to the cell size Δx, but penalized by the Reynolds number. Note that when taking $c_s/c = 1$ instead of the theoretical value (5.20), the subcharacteristic condition (5.18) is violated, and the relative overshoot in this subcharacteristic condition is only related to the value of Re_cell, the larger Re_cell is, the more the subchracteristic condition is overpassed (see the comments after (5.20)).

When having an analytical solution, we compute the relative $L1$ norm (or $L2$ norm) error. Denoting by $\mathbf{u}$ the computed solution and $\mathbf{u}_{ex}$ the analytical one, this means that we define

$$\text{Err} = \frac{||\mathbf{u} - \mathbf{u}_{ex}||_1}{||\mathbf{u}_{ex}||_1}. \tag{5.28}$$

We eventually also compute the relative $L2$ norm gradient error, called ErrGrad, which is

$$\text{ErrGrad} = \frac{||\nabla\mathbf{u} - \nabla\mathbf{u}_{ex}||_2}{||\nabla\mathbf{u}_{ex}||_2}. \tag{5.29}$$

6. **Numerical results**

In this section we show the accuracy of the BGK-FVS method, implemented as explained in Section 5. We consider two-dimensional classical benchmarks which have an analytical solution or some strong experimental references. We finally apply our scheme to a three-dimensional device in Subsection 6.2.3. There are two different classes of test cases. The first class consists in validation test cases. Those have analytical solutions, which are suitable for checking the order of accuracy of the scheme. We consider the Taylor-Green vortex and the Poiseuille flow. The second class consists in assessment test cases. Those are more physical, with lots of experiments and numerical data. We consider the backward-facing step as steady problem, and the laminar flow past a square cylinder as unsteady one.

   The number of grid points per direction is denoted by N. We analyze the numerical results in terms of the diagnostics defined in the previous subsection. The subcharacteristic condition (5.18) involving the cell Reynolds number defined in (5.17) is pretty restrictive, since the supremum on velocities is over all computational cells. However, as we show, violating this condition on a few cells does not make the whole simulation break down, even if we can then foresee some impacts on the accuracy of the simulation.

6.1. **Validation tests**

The setting of parameter values is described in Subsection 5.2. We use the formula (5.20) to compute c_s/c. This means that we first estimate the maximum $u_{\max}$ of $\mathbf{u}$ that is involved in (5.18). Then we define c_s/c with (5.20). Another method to take $c_s/c = 1$, it is evaluated in Subsubsection 6.1.3.

6.1.1. *Taylor-Green vortex*

This simple test case is commonly used to determine the stability and accuracy of a scheme without the artifacts of boundary treatment, since periodic boundary conditions are used, for velocity and pressure. It consists in an analytical time evolution of a decaying vortex described by

$$\begin{cases} u_x(t,x,y) = -\exp\left(-\nu t(w_1^2 + w_2^2)\right)\cos(w_1 x)\sin(w_2 y), \\ u_y(t,x,y) = \frac{w_1}{w_2}\exp\left(-\nu t(w_1^2 + w_2^2)\right)\sin(w_1 x)\cos(w_2 y), \\ p(t,x,y) = -\frac{1}{4}\exp\left(-2\nu t(w_1^2 + w_2^2)\right)\left(\cos(2w_1 x) + \frac{w_1^2}{w_2^2}\cos(2w_2 y)\right). \end{cases} \tag{6.1}$$

The computational domain is the square $[-\pi, \pi]^2$. We fix the Reynolds number at $Re \equiv U_{max}L/\nu = 100$, with $u_{max} = 1.6$ and $L = 2\pi$. The integer parameters w_1, w_2 define the numbers of vortices in the x, y directions. In order to avoid symmetries we take $w_1 = 3$, $w_2 = 2$. The initialization of (u_x, u_y, p) is done via equation (6.1) at t=0.

   The L^1 relative error is shown on Figure 6.1. On these plots we see clearly the appearance of time oscillations related to sound waves induced by the numerical resolution via a compressible system with variable ρ. These oscillations have growing frequency as Δx decreases. This is due to the choice of the Mach number ϵ being proportional to Δx as (5.3), i.e. $\epsilon = c\Delta x/(2\nu)$. Indeed according to the scaled compressible system (3.13), (3.14) with pressure law (5.19), the typical form of the sound waves is $\exp(i(\omega t + \mathbf{k} \cdot \mathbf{x}))$, with $\mathbf{k}$ the wavevector, ω the pulsation given by

$$\omega = \pm\frac{|\mathbf{k}|c_s}{\epsilon} = \pm|\mathbf{k}|\frac{c_s}{c}\frac{2\nu}{\Delta x}, \tag{6.2}$$

which is inversely proportional to Δx. The oscillation period is $2\pi/\omega$, proportional to Δx. The value of $|\mathbf{k}|$ can be computed by $|\mathbf{k}| = 2\pi/\lambda$ with λ the wavelength, i.e. the characteristic length of the problem. Here we can take $|\mathbf{k}| = w_1$ for a wave in the x direction, $|\mathbf{k}| = w_2$ for a wave in the y

(A) Relative $L1$ error for u_x vs time

(B) Relative $L1$ error for u_y vs time

(C) Relative $L1$ error for p vs time

(D) Relative $L1$ error for *prim p* vs time

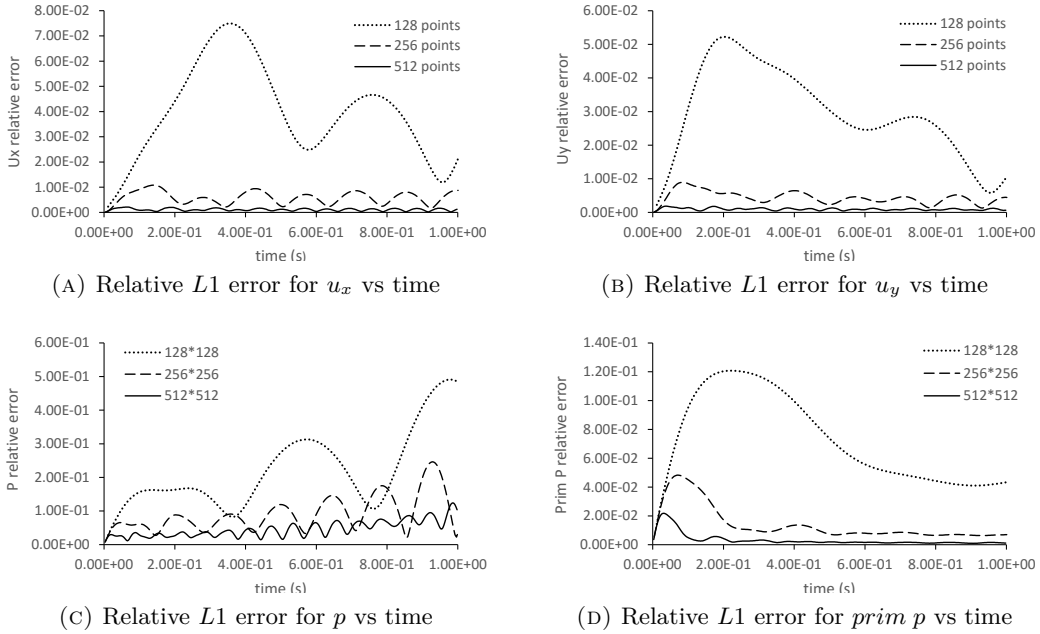FIGURE 6.1. Relative $L1$ errors vs time for pressure and velocity for different grid sizes. The quantity *primp* is the primitive of the pressure with respect to time, i.e. $\int^t p(\tau, x, y)\, d\tau$, computed by finite differences in time.

direction. The latter gives the observed frequencies on Figure 6.1. We refer to [25] for the analysis of fast sound waves.

Indeed kinetic methods have the same fast oscillatory behavior ([12, 62]). In order to take a relevant value for the considered error, a mean error over time is computed. The results are shown on Table 6.1, together with the values of the ratio c_s/c and of $\mathrm{Re}_{\mathrm{cell}}$.

N	c_s/c	$\mathrm{Re}_{\mathrm{cell}}$	Err u_x	Err u_y	Err p	Err *prim p*
128	0.583	0.416	3.94E-02	2.94E-02	2.20E-01	7.29E-02
256	0.792	0.208	5.84E-03	4.39E-03	9.14E-02	1.40E-02
512	0.895	0.104	1.10E-03	9.58E-04	4.61E-02	3.25E-03

TABLE 6.1. Mean values over time of relative $L1$ errors for u_x, u_y, p and its primitive in time *prim p* for different grid sizes. N denotes the number of grid points per direction.

We observe that as the mesh is refined, $\mathrm{Re}_{\mathrm{cell}}$ becomes smaller, and the ratio c_s/c tends to 1. A second-order accuracy is observed for the velocity, and first-order accuracy for the pressure, as shown on Figure 6.2b. Thus we obtain the same rates as for a $P2/P1$ finite element method. The numerical solution at time $t = 1$ is shown on Figure 6.2a.

(A) Pressure and velocity fields, colored as magnitude

(B) Relative $L1$ errors for velocity and pressure fields. The scheme is second-order for the velocity, and first-order for the pressure.
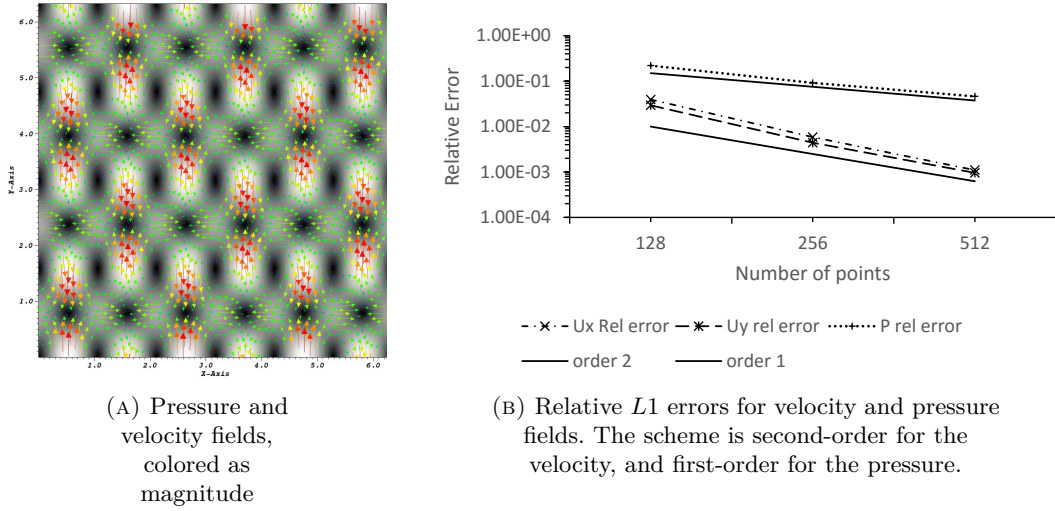
FIGURE 6.2. Numerical solution at $t = 1$; and mean relative $L1$ errors vs grid

The convergence on the pressure should be however second-order accurate if measured in a weaker norm such as H^{-1}. This property can be tested by computing the relative $L1$ error on the primitive in time of the pressure, with respect to the primitive in time of the exact pressure. The result is shown on Table 6.1 and indicates an order of accuracy of 1.85, thus improving the first-order accuracy observed on p. We can compute also the relative $L2$ error for $\nabla \mathbf{u}$, which is ErrGrad defined by (5.29). The results are in Table 6.2 for $t = 1$. We get a (higher than) second-order accuracy on the gradient, which is thus even better than the pressure. Note that the periodic context is the most oscillatory situation since the sound waves are not damped nor dispersed, see the recent results in [26]. Our method shows to work nicely in this case.

N	ErrGrad	Order ErrGrad
128	4.72e-02	-
256	6.86e-03	2.62
512	1.42e-03	2.2

TABLE 6.2. Relative $L2$ error for $\nabla \mathbf{u}$ at $t = 1$ for the Taylor-Green vortex, for different grid sizes.

Next, we check that the entropy inequality (5.14) is satisfied, by plotting the maximal dissipation EntMax defined in (5.26). Figure 6.3 shows that it is always negative, as expected. It indeed tends to zero (at second order) with mesh refinement, because the right-hand side of (5.2) vanishes at some locations in the domain, that correspond to the maximum possible values of $\mathrm{Diss}_{\mathbf{i}}^{n}$.

N	ErrDiv	Order ErrDiv	ErrDiss	Order ErrDiss
128	2.02e-2	-	5.05e-02	-
256	3.00e-3	2.59	8.04e-03	2.50
512	6.32e-4	2.17	1.73e-03	2.15

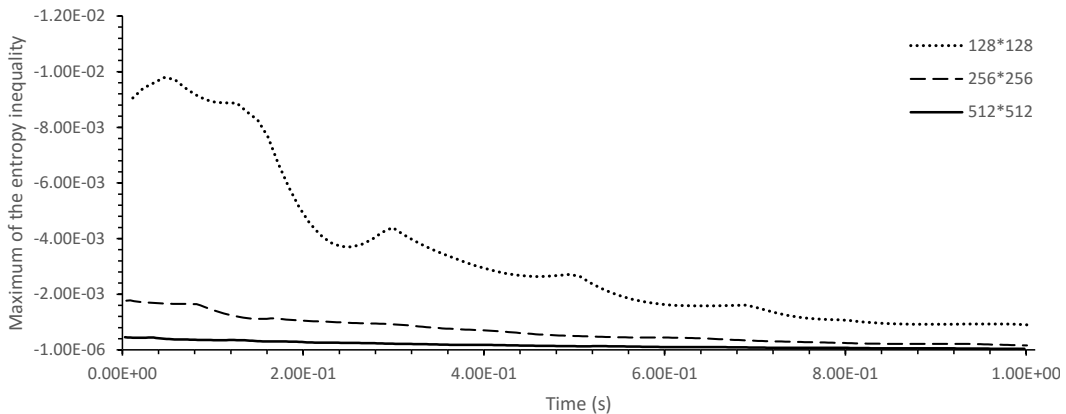TABLE 6.3. Divergence and dissipation relative errors in $L1$ norm at $t = 1$ for the Taylor-Green vortex

FIGURE 6.3. Maximum EntMax of the entropy dissipation vs time. It remains negative.

We finally check the order of accuracy on the indicators ErrDiv defined in (5.27) (measuring the error on $\nabla_{\mathbf{x}} \cdot \mathbf{u}$) and ErrDiss defined in (5.25) (measuring the error on the dissipation). We observe a better than second-order convergence on the divergence and entropy dissipation, as seen on Table 6.3.

6.1.2. *Poiseuille flow*

We consider the 2D Poiseuille flow in a rectangular channel of dimensions $L * H$. It is commonly used to test the accuracy of boundary conditions. The exact solution is given for $y \in [0, H]$ by its x component of velocity (the y component vanishes) and its pressure,

$$u_{ex}(y) = \frac{4\,u_{max}}{H^2}\,y\,(H - y), \quad p_{ex}(x) = \frac{8\nu\,u_{max}}{H^2}(x_r - x). \tag{6.3}$$

The boundary conditions are taken as follows: we impose zero velocity at the lower and upper walls, at the inlet (left side) we impose the exact parabolic velocity profile, and at the outlet (right side) we set a vanishing pressure, meaning that $p_o = 0$ in the boundary condition of Subsection 5.3.4. We take H as characteristic length, and the Reynolds number is defined by $Re = u_{max}H/\nu$. The value of the parameters are $Re = 100$, $H = 1$, $L = 1$, $u_{max} = 1$. The simulation is stopped when the relative difference of solutions between two timesteps is small enough, meaning that the numerical solution is steady.

The computed errors are shown on Table 6.4. The discrete relative entropy inequality is satisfied since EntMax remains negative. Figure 6.4 shows that we recover second-order accuracy for both the velocity and the pressure. This is due to the steadiness of the solution that eliminates sound waves, that are indeed damped by boundary conditions.

N	c_s/c	Re_{cell}	Err u_x	Err p	EntMax
160	0.6875	0.3125	1.15e-02	1.29e-01	-1.91e-05
320	0.8438	0.1563	1.75e-03	1.89e-02	-4.70e-06
640	0.9219	0.0781	3.64e-04	3.87e-03	-1.17e-06
1280	0.9609	0.0390	8.41e-05	8.86e-04	-2.93e-07

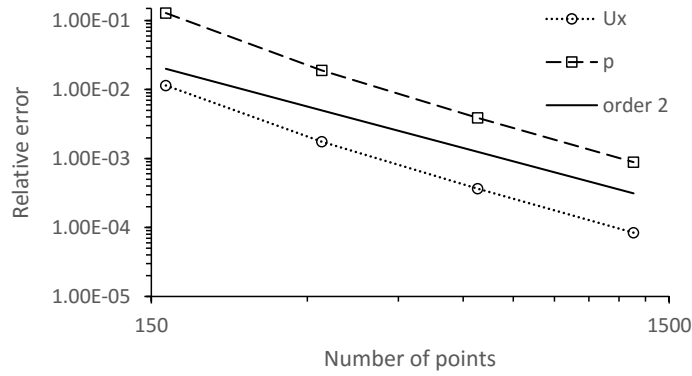TABLE 6.4. Numerical parameters, relative $L1$ errors for u_x and p and EntMax defined as (5.26)

FIGURE 6.4. Accuracy of the Poiseuille flow computation. Second-order is achieved for both the velocity and the pressure

Next, Figure 6.5 and Table 6.5 show the divergence and dissipation relative $L1$ errors, as well as gradient relative $L2$ error. The divergence and dissipation errors tend toward a fixed value as time tends to infinity, and indicate second-order accuracy. This is also the case for the gradient relative error.
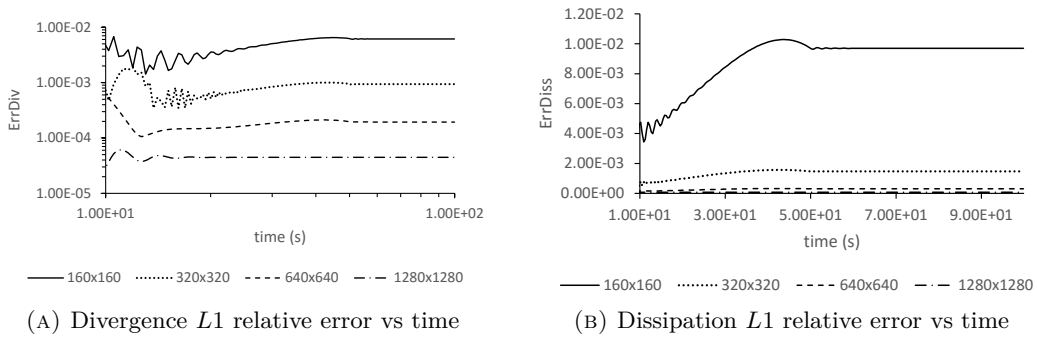

(A) Divergence $L1$ relative error vs time

(B) Dissipation $L1$ relative error vs time

FIGURE 6.5. Divergence and dissipation relative $L1$ errors vs time

34

N	ErrDiv	Order ErrDiv	ErrDiss	Order ErrDiss	ErrGrad	Order ErrGrad
160	6.15e-03	-	9.70e-03	-	1.59e-02	-
320	9.37e-04	2.56	1.47E-03	2.57	2.43e-03	2.56
640	1.94e-04	2.20	3.04e-04	2.20	5.06e-04	2.19
1280	4.47e-05	2.09	7.00e-05	2.09	1.17e-04	2.08

TABLE 6.5. Divergence and dissipation relative $L1$ errors, $L2$ gradient error, at $t = 100$

6.1.3. *Effect of the choice of c_s/c*

Evaluating $u_{\max}$ could be an issue in real geometry simulations. Then the ratio c_s/c cannot be defined by (5.20), and another strategy needs to be used. The value of c_s/c controls somehow the numerical diffusion of the scheme. The larger c_s/c is, the less diffusive is the scheme. Taking a too large value of c_s/c will however give instability. In fact, following (6.2), $c_s/\epsilon = (c_s/c)(2\nu/\Delta x)$ is the sound speed. Thus taking a large c_s/c will better realize the low Mach number limit. Indeed according to (5.7), $\rho - \overline{\rho}$ is proportional to $(c/c_s)^2$. Taking c_s/c large, we force ρ to be close to $\overline{\rho}$ and $\nabla_{\mathbf{x}} \cdot \mathbf{u}$ to be small, because of the mass equation. Moreover we also need ρ to be close to $\overline{\rho}$ for the momentum equation. It is thus tempting to force the highest value for the ratio c_s/c, provided that the scheme remains stable. In the following tests, we use

$$\frac{c_s}{c} = 1. \tag{6.4}$$

With this formula, we do not need to compute $u_{\max}$ in advance. We investigate below the violation of the subcharacteristic condition (5.18) with this choice, and analyse its impact on the quality of the computed solution on the two previous tests.

We first consider the Taylor-Green vortex. As seen on Figure 6.6, the value $c_s/c = 1$ gives the same orders of accuracy as with the reference choice (5.20), namely second-order accuracy for velocity, and first-order for pressure. However the error is slightly reduced for the velocity, and significantly reduced for the pressure. No instability is observed and the entropy inequality is satisfied (EntMax is negative) even if the overshoot in the subcharacteristic condition is quite large when $\mathrm{Re}_{\mathrm{cell}}$ grows, see Table 6.6. We have a converged solution for a mesh of 64 points, which could not be obtained with the reference choice of c_s/c. We observe on Table 6.7 that the ErrDiss, ErrDiv and ErrGrad indicators are all second-order accurate.

(A) Relative error on $\mathbf{u}$

(B) Relative error on p
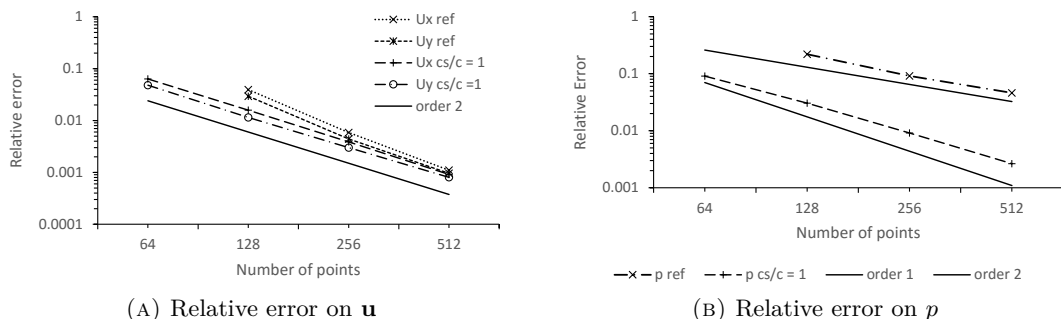
FIGURE 6.6. Relative $L1$ errors Err u_x, Err u_y and Err p for the Taylor-Green vortex test case with the choice $c_s/c = 1$. The error on pressure is significantly reduced.

N	Re$_{\text{cell}}$	Err u_x	EntMax
64	0.833	6.32e-02	-2.07e-03
128	0.416	9.71e-03	-6.76e-04
256	0.21	2.32e-03	-1.69e-04
512	0.1	5.40e-04	-3.94e-05

TABLE 6.6. Indicator values for the Taylor-Green vortex test case with the choice $c_s/c = 1$.

N	ErrDiv	Order ErrDiv	ErrDiss	Order ErrDiss	ErrGrad	Order ErrGrad
64	0.03	-	9.41e-02	-	7.49e-02	-
128	7.98e-03	2.03	2.44e-02	1.97	1.85e-02	2.01
256	2.06e-03	1.97	6.13e-03	1.99	4.69e-03	1.99
512	5.33e-04	1.96	1.54e-03	2.00	1.20e-03	1.98

TABLE 6.7. Divergence and dissipation relative $L1$ errors, $L2$ gradient error, at $t = 1$ for the Taylor-Green vortex test case with $c_s/c = 1$

We proceed then to the Poiseuille flow test case, with choosing $c_s/c = 1$. Figure 6.7 shows the accuracy of the computation, which is still second-order for velocity and pressure but with slightly reduced errors. As for the previous test case, the choice $c_s/c = 1$ leads to a converged solution for a coarser grid. Table 6.8 shows that the entropy inequality still holds, even if an overshoot in the subcharacteristic condition exists. We keep second-order accuracy for ErrDiv, ErrDiss and ErrGrad, as shown on Table 6.9.
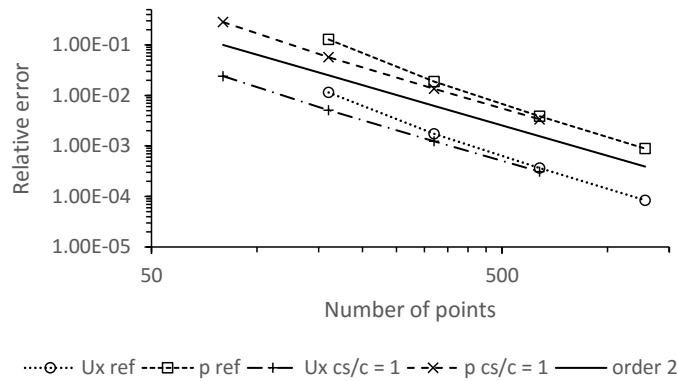

FIGURE 6.7. Relative $L1$ errors for the Poiseuille flow with $c_s/c = 1$

N	$\mathrm{Re}_{\mathrm{cell}}$	Err u_x	EntMax
80	0.64	2.42e-02	-7.00e-05
160	0.314	5.13e-03	-1.76e-05
320	0.156	1.24e-03	-4.50e-06
640	0.078	3.10e-04	-1.15e-06

TABLE 6.8. Cell Reynolds number for the Poiseuille flow with $c_s/c = 1$, $L1$ relative errors of u_x, and EntMax

N	ErrDiv	Order ErrDiv	ErrDiss	Order ErrDiss	ErrGrad	Order ErrGrad
80	0.013	-	2.10e-02	-	3.20e-02	-
160	2.75e-03	2.17	4.36e-03	2.19	7.01e-03	2.15
320	6.65e-04	2.03	1.05e-03	2.04	1.72e-03	2.02
640	1.65e-04	2.00	2.60e-04	2.00	4.29e-04	2.00

TABLE 6.9. Divergence and dissipation relative $L1$ errors, $L2$ gradient error, at $t = 100$ for the Poiseuille test case with $c_s/c = 1$

The conclusion of this subsection is that the choice $c_s/c = 1$ is interesting. On the presented test cases of Taylor-Green vortex and Poiseuille flow, the entropy inequality (5.14) is satisfied even if the subcharacteristic condition (5.18) is not. This choice of c_s/c gives much smaller errors than the theoretical choice (5.20), and thus allows a coarser mesh. Further on we shall use this choice $c_s/c = 1$.

6.2. **Assessments tests**

In this subsection we evaluate our BGK-FVS scheme on some classical fluid mechanics benchmarks. We make the choice (6.4) i.e. $c_s/c = 1$ instead of (5.20).

6.2.1. *Backward-facing step*

The laminar flow around a backward-facing step in a 2D channel has been extensively studied in the literature since the experiments of [5]. The sudden enlargement of the section causes a reverse pressure gradient which leads to a separation of the flow into several zones, with the appearance of a recirculation behind the step. When the Reynolds number increases, a second recirculation is observed close to the top lid. It is important to ensure that the scheme reproduces well these features. The geometry is chosen in accordance with the experimental setup of Armaly et al. sketched on Figure 6.8, with a slight difference: the expansion ratio is defined by $H/h_i = 2$, which is the ratio between the channel height H downstream and the channel height h_i upstream. We choose this expansion ratio for the numbers of numerical cells in both the upstream and downstream channels to be integers. The study corresponding to this ratio can be found in [47]. On the converse, the experimental value 1.9423 of the ratio would not give integers, and would not be accurate for coarse grids, implying some unquantified errors.

The boundary conditions are set as follows. The top and bottom part of the channel are considered as walls (with vanishing velocities in all directions). A Poiseuille profile is imposed at the inflow and

a free outflow condition is used on the right side, as in the above Poiseuille test case. We measure the position x_1 at the end of the first bottom recirculation, the position x_2 corresponding to the beginning of the top lid recirculation, and x_3 corresponding to the end of it, as shown on Figure 6.8.
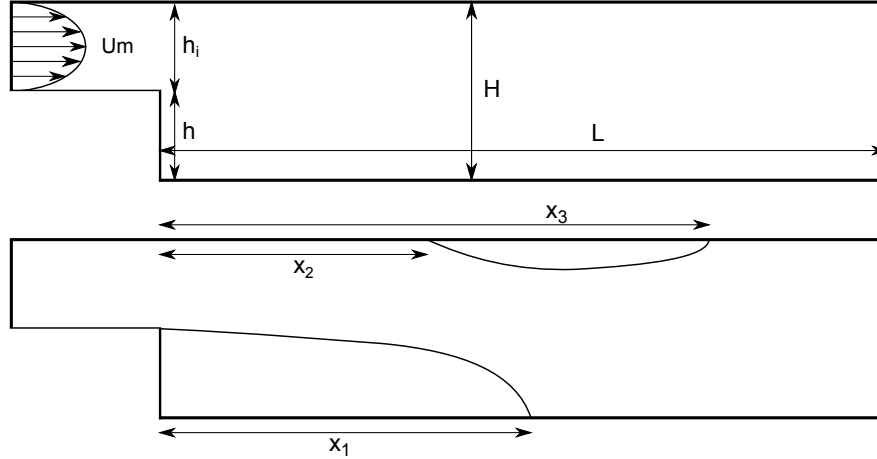

FIGURE 6.8. Domain and measured quantities

The definition of the Reynolds number which is used in this study and by the literature is $Re_D = u_m D/\nu$, where u_m is two-thirds of the maximum inlet velocity (equivalent to the average inlet velocity), D is the hydraulic diameter of the inlet channel, which is twice its height ($D = 2h_i$), and ν is the kinematic viscosity. We present the results with $H = 0.5$, $L = 46$ and $u_m = 1$.

We perform a mesh refinement study for Re = 100, with 3 different grids, dividing by 2 the space increment Δx in both directions from one mesh to the other. We measure the position x_1, the results are shown on Table 6.10. In order to evaluate the asymptotic limit value corresponding to the mesh convergence, we use the well-known Richardson extrapolation (see [63] for more explanations). We obtain second-order accuracy for the position x_1. Moreover, ErrDiv and ErrDiss are second-order accurate and the entropy inequality is satisfied for all grids, as seen on Table 6.11.

Grid level	Δx	x_1	Error	Error (%)
Richardson ext.	-	2.915429	-	-
3	0.002	2.918	0.002571	0.088
2	0.004	2.924	0.0086	0.29
1	0.008	2.944	0.0286	0.98

TABLE 6.10. Mesh refinement study for $Re = 100$

Grid level	EntMax	ErrDiv	Order ErrDiv	ErrDiss	Order ErrDiss
3	-2.83e-11	5.94e-05	2.01	4.83e-04	2.06
2	-3.76e-10	2.37e-04	2.01	2.05e-03	1.96
1	-9.58e-09	9.61e-04	-	7.88e-03	-

TABLE 6.11. Entropy indicator EntMax, divergence and dissipation relative $L1$ errors

In order to study the flow pattern, we investigate Reynolds numbers between 100 and 800. Figure 6.9 shows the streamlines for these high Reynolds numbers. There are two different behaviors. When $Re_D \leq 400$ there is only one recirculation behind the step, and our results perfectly match the literature. A precise evaluation is presented on Figure 6.10a. A difference of 2.2% is observed on the position x_1 for Re = 800. When $Re_D \geq 400$, the top lid recirculation appears, and the lower reattachment point continues to move away from the step. The computed attachment points are shown on Table 6.12. The positions x_2 and x_3 are also well predicted, see Figure 6.10b. The relative error, taking the value of the literature as a reference, is very small for x_2 (less than 1%). For x_3, it grows as Re grows, and ends at 4% for Re = 800. As the same finer mesh is used for every Reynolds number greater than 400, the cell Reynolds number grows as the Reynolds number grows, and therefore the error grows as the problem is stiffer. We see on Table 6.12 that a value of Re_{cell} greater than 0.4 leads to an error growth from 2.5% to 4.3%. This test case shows that a value of Re_{cell} greater than 0.4 indicates a loss of accuracy.

Re	Re_{cell}	x_1	x_2	x_3	Err x_3 (%)
100	0.075	2.918	-	-	-
200	0.15	5.008	-	-	-
300	0.225	6.824	-	-	-
400	0.3	8.366	7.65	10.258	2.1
500	0.375	9.594	7.992	13.484	2.3
600	0.45	10.554	8.484	16.308	2.9
700	0.525	11.366	8.992	18.942	3.5
800	0.6	12.112	9.514	21.496	4.3

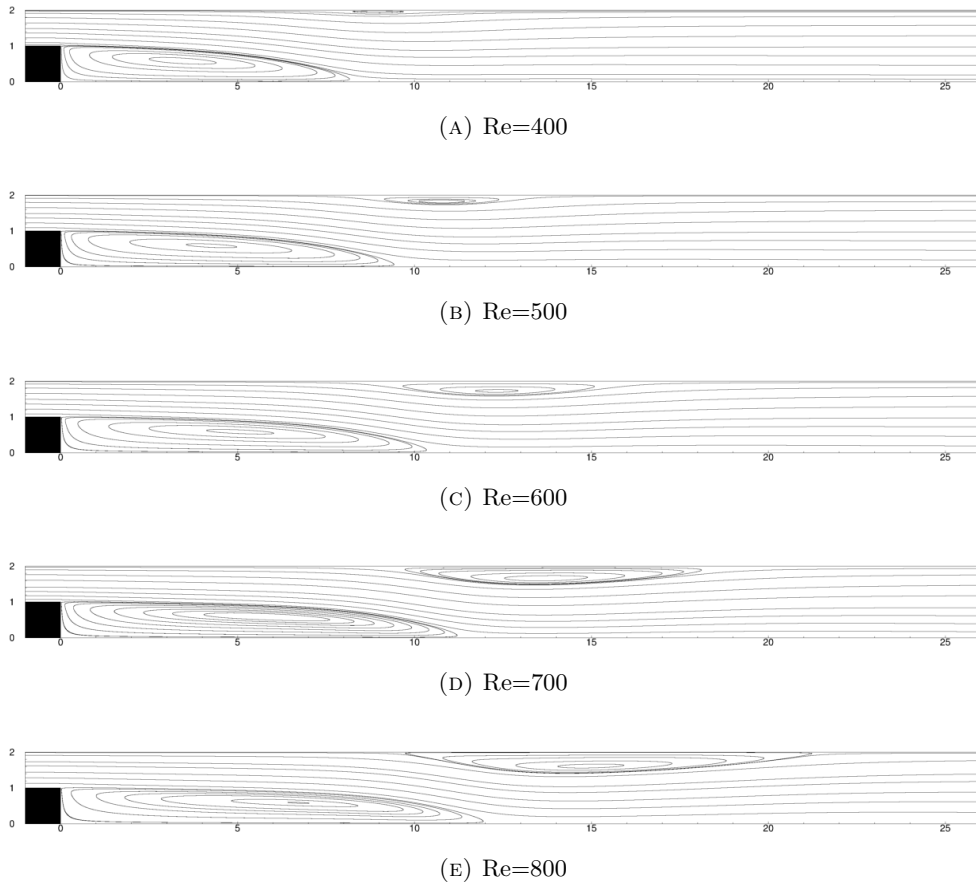TABLE 6.12. Numerical results of the Backward-facing step benchmark, $\Delta x = 0.002$ for all those results.

(A) Re=400

(B) Re=500

(C) Re=600

(D) Re=700

(E) Re=800

Figure 6.9. Streamlines for high Reynolds numbers for the backward-facing step test case

(A) Position x_1 compared to the literature

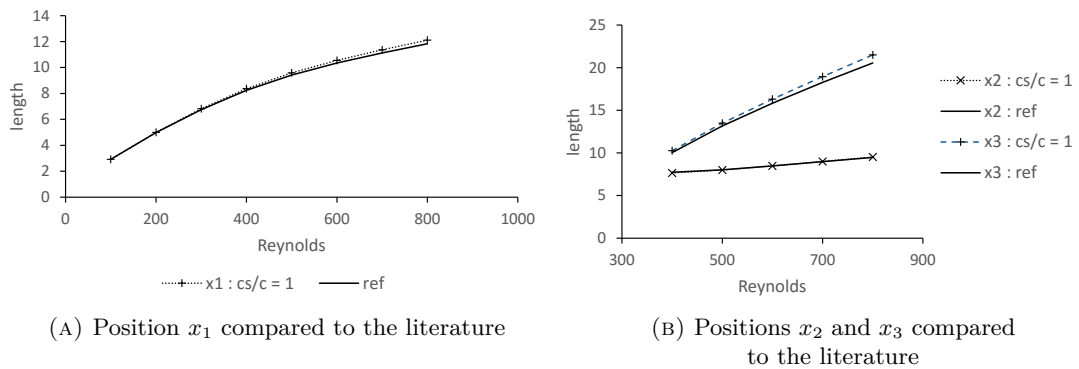(B) Positions x_2 and x_3 compared to the literature

Figure 6.10. Measured lengths compared to the literature

Our conclusions for this test case are that we recover correctly the two recirculations at high Reynolds number, and the position of the attachment points are well located.

6.2.2. *Laminar flow past a square cylinder*

This test case concerns a laminar flow around a square cylinder between two planes distant of H. An important parameter is the blocking factor β, defined as the ratio between the side D of the square and H. It is fixed at $\beta = 0.125$, in accordance with [27]. The domain is shown on Figure 6.11. The inlet is a Poiseuille velocity profile with a maximum velocity of $1\,\mathrm{m\,s^{-1}}$. We impose no-slip boundary conditions on the top and bottom planes, and a free outlet on the right (transparent boundary condition). This last condition is very important (see [80]), not resolving it well may radically change the results. Applying a good treatment of the free outlet allows to have the smallest effect on the flow near the boundaries and the obstacle, and therefore to save some computational time by truncating the domain on the right without any noticeable change. The analysis of transparent boundary conditions can be found in [28] and [22, Chapter 7].
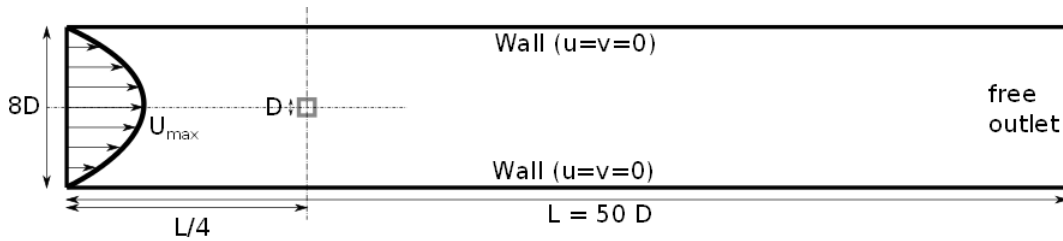

FIGURE 6.11. Description of the domain

Here we use a Convective Boundary Condition (CBC) ([69, 67, 72]). The CBC consists in solving at the boundary $\partial_t \mathbf{u} + u_c \partial_x \mathbf{u} = 0$, where u_c is considered as u_m, as recommended in [80], discretized with finite differences at second order in space, first order in time, implicitly. We initialize the flow fields at zero for velocity and pressure, and take $D = 1$ and $u_{max} = 1$.

The flow pattern has been studied for a Reynolds number between 20 and 133. For $Re < 60$, the flow separates at the trailing edge of the square cylinder, and a steady recirculation behind the obstacle is observed. It is perfectly symmetric, and there is no vortex shedding. The length L_r of this recirculation follows a linear law ([27]),

$$L_r/D = -0.065 + 0.0554 Re. \tag{6.5}$$

As can be seen on Figure 6.12, the length of the recirculation region is correctly predicted compared to this law, on the grid (8000x1280). When $Re \geq 60$, the flow becomes unsteady, with some vortex shedding that becomes periodic in time with frequency f.
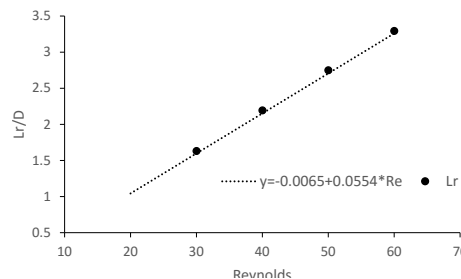

FIGURE 6.12. Recirculation length

For a quantitative comparison, we define three dimensionless numbers. The first one is the Strouhal number St, related to the frequency f of the vortex shedding. It is defined by St $= fD/u_{max}$, with

u_{max} the maximal velocity at the inflow. The frequency f is computed by a spectral analysis over the values of the velocity u_x at one cell past the square cylinder. The other quantities are the drag and lift numbers C_d and C_l. These are defined via drag and lift forces respectively, defined as follows. We denote the symmetric stress tensor by

$$\sigma = -p\mathbf{I} + \nu \left(\nabla \mathbf{u} + (\nabla \mathbf{u})^t \right), \tag{6.6}$$

with p the pressure. Then the drag and lift forces are, with x the direction parallel to the flow and y the perpendicular direction, $\mathbf{n}$ the external unit normal to the square cylinder, and Γ the contour of the square:

$$F_d = \int_\Gamma \sigma \mathbf{n} \, dl \cdot \mathbf{e}_x, \qquad F_l = \int_\Gamma \sigma \mathbf{n} \, dl \cdot \mathbf{e}_y. \tag{6.7}$$

Finally, C_d and C_l are defined by normalizing those forces by the kinetic energy,

$$C_d = \frac{F_d}{\frac{1}{2}u_{max}^2}, \qquad C_l = \frac{F_l}{\frac{1}{2}u_{max}^2}. \tag{6.8}$$

The convergence with mesh refinement is achieved for an unsteady problem at Re = 65, see Tables 6.13 and 6.14. The order of accuracy observed on the ErrDiv indicator is 2. The order of accuracy for C_d and C_l (which are related to the pressure) is almost 3 (2.9), taking the Richardson extrapolation as the reference solution. Moreover for the Strouhal number (related to velocity), we observe an order 7. Those particularly good orders are specific to this test case.

Grid level	$\Delta x/D$	Mean C_d	Mean C_d Error (%)	Max C_l	St	St Error (%)
Richardson ext.	-	1.516	-	0.0759	0.123198	-
3	0.003125	1.52	0.26	0.076	0.1232	1.6e-3
2	0.00625	1.55	2.24	0.08	0.1233	0.08
1	0.0125	1.80	18.7	0.322	0.1281	3.98

TABLE 6.13. Computed values for different meshes at $Re = 65$.

Grid level	$\mathrm{Re}_{\mathrm{cell}}$	EntMax	ErrDiss	ErrDiv
3	0.12	-1.72e-09	0.0118	2.29e-04
2	0.246	-6.77e-09	0.0249	8.16e-04
1	0.522	-7.62e-08	0.0664	0.0102

TABLE 6.14. Indicators for Re = 65

Concerning the entropy indicators, EntMax is always negative, meaning that the entropy inequality is satisfied for all the cells of the computational domain. There is almost a second-order accuracy for ErrDiss (1.88). There is a particularly good order of accuracy for grids 1 to 2, of 3.5. This can be explained from the fact that the cell Reynolds number is 0.52 for grid 1, which is quite high. Recalling the conclusion of the previous test case, we may conclude that grid 1 has unsufficient accuracy.

The Von Karman streets begin at Reynolds number greater than 60, see Figure 6.14. We have chosen the grid level 2 because it makes a good balance between accuracy and computational time. The indicators are shown on Tables 6.15 and 6.16. As can be seen on Figure 6.13b, the Strouhal numbers are perfectly recovered until Re reaches 100. After that, we observe a degradation in the

accuracy of the solution, with 8% of relative error compared to FVM. Past that Reynolds number, the cell Reynolds number exceeds 0.4, indicating an unacceptable loss of accuracy, and the need of a finer mesh in order to have a better approximation. The global behavior of the drag coefficient is well predicted (Figure 6.13a).

Re	Re_{cell}	L_r	C_d	Mean C_l
30	0.11	1.63	2.09	-6.58E-14
40	0.148	2.19	1.84	-2.39E-15
50	0.185	2.75	1.68	-3.28E-14

TABLE 6.15. Numerical results for the Von Karman benchmark at low Reynolds number

Re	Re_{cell}	C_d	Mean C_l	Max C_l	St
60	0.22	1.58	-4.57e-04	0.0468	0.1196
65	0.24	1.55	5.06e-03	0.0805	0.1236
80	0.31	1.50	4.18e-03	0.150	0.1307
100	0.4	1.47	0.03	0.228	0.1377
130	0.53	1.492	0.048	0.462	0.136

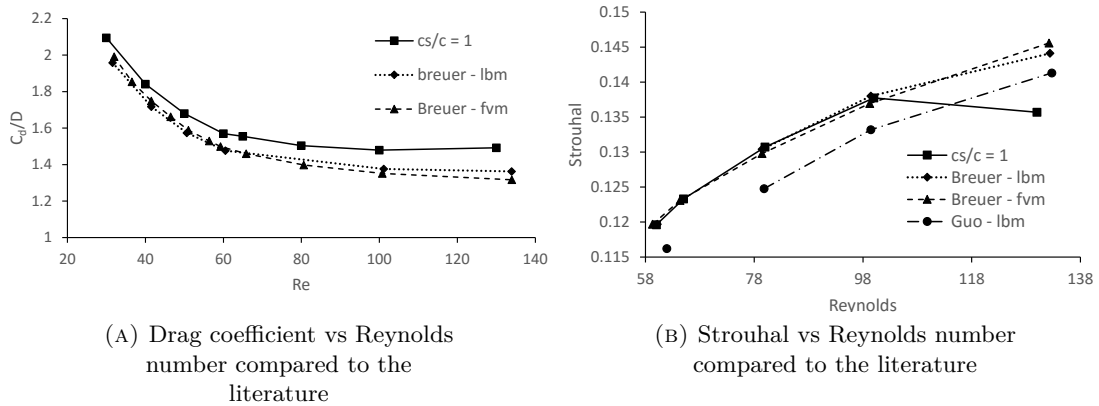TABLE 6.16. Numerical results for the Von Karman benchmark for unsteady solutions

(A) Drag coefficient vs Reynolds number compared to the literature

(B) Strouhal vs Reynolds number compared to the literature

FIGURE 6.13. Drag coefficient and Strouhal number compared to the literature

(A) Re=30

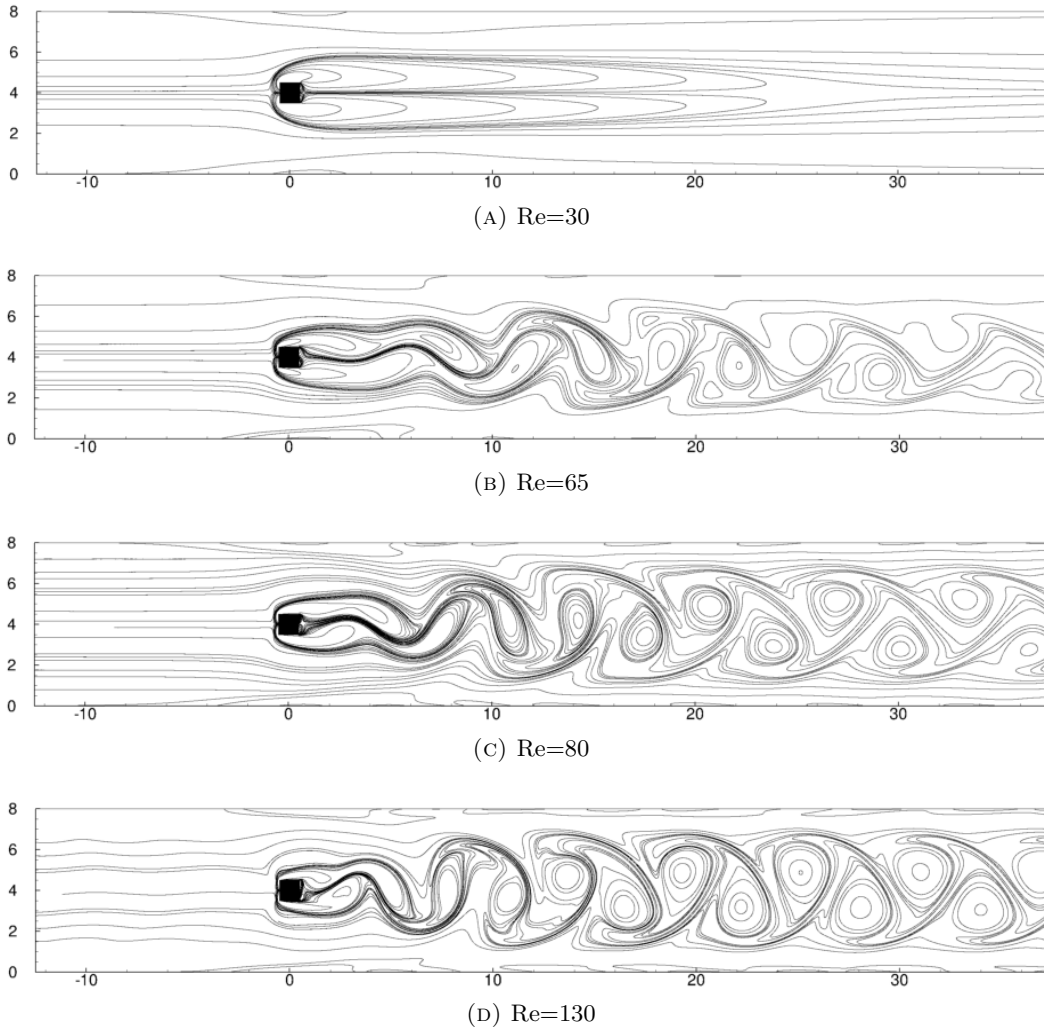(B) Re=65

(C) Re=80

(D) Re=130

FIGURE 6.14. Vorticity contours for different Reynolds numbers for the laminar flow past a square cylinder test case

6.2.3. *3D flow around complex object*

There is no available analytic test case in 3D with complex geometry. In order to evaluate our scheme in such 3D context, were propose here a test case of physical relevance. The numerical results are then evaluated in terms of mean quantities. The test aims to compute intrinsic macroscopic properties of a porous medium. The considered geometry is an idealized 3D Kelvin-like cell (Figure 6.15a), with circular struts. The numerical results are then evaluated in terms of mean quantities. The chosen transport property is the permeability. It characterizes the ability of a porous medium to allow fluids to flow through it, and it is related to the notion of hydraulic conductivity introduced by Darcy. His historical experiment [35] consists in a tube partially filled with sands of height L over a filter of cross-section S. The pressure is measured at two different positions (distance l from each other), from the top to the bottom of the tube containing sands. The upper part of the tube is connected to the mains water supply. His experiments show that the flow velocity is proportional to the pressure head

difference, according to Darcy's formula

$$U = K\Delta P/l, \tag{6.9}$$

with $U = Q/S$ is the filtration velocity, Q is the volume flow rate, S the cross-section in a plane perpendicular to the direction of flowing, and K a proportionality coefficient known as the hydraulic conductivity. The hydraulic conductivity depends on the fluid used. The Darcy law can be generalized for all fluids by introducing the fluid viscosity μ_f, and thereby defining an intrinsic property of the porous medium, the permeability, giving

$$-\bar{\bar{K}}_D \nabla < P > = \mu_f < V >, \tag{6.10}$$

where $\nabla < P >$ is the average pressure gradient, $\bar{\bar{K}}_D$ the permeability tensor (D for Darcy) and $< V >$ the average fluid velocity over all the volume of the foam sample. We consider in those assessments a 1D pressure drop approach. Though the inertial effects can be significant at low Reynolds number in open-cell foams, it is necessary to determine at first the permeability in the Darcy regime for a relevant range of velocities.
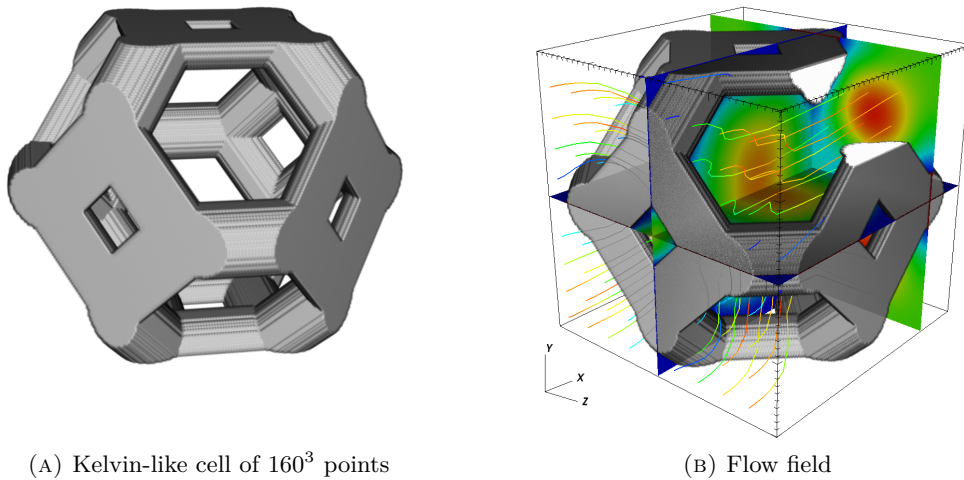

(A) Kelvin-like cell of 160^3 points

(B) Flow field

FIGURE 6.15. (a) Computational domain: Kelvin-like cell of diameter $d_{\text{cell}} = 4mm$. (b) Flow field: velocity magnitude through the foam on two perpendicular planes.

In order to determine the permeability in a given direction, the following boundary conditions are used: prescribed pressure drop between two opposite faces, while other faces are set as symmetry planes as shown on Figure 6.15b. The velocities for the pressure drop faces are set as free.

The mean pressure gradient ($\nabla < P >^f$) is the average pressure difference between the inlet and outlet faces of the medium (see [75]), giving

$$\nabla < P >^f = \frac{\Delta P}{L}, \tag{6.11}$$

where $\Delta P/L$ is the pressure drop applied at the inlet and outlet faces. Using equations (6.10) and (6.11) we can compute the permeability.

The outline of the study is as follows. We first evaluate the influence of mesh resolution. Then we verify the homothetic scaling-law. Finally we investigate the influence of porosity over the permeability and compare the results to literature data obtained using a commercial software: starccm+ ([55]).

Influence of mesh resolution

The mesh convergence is important to obtain results as close as possible to real cases. It must be done carefully for not obtaining a contrary effect. In fact, if the mesh is too thin, it has already been observed that the small scale folding of the surface leads to incorrect results. On the other hand, coarse meshing does not allow the solver to provide accurate results as morphological errors are very important. The first step is to verify the impact of the geometric discretized resolution (voxel size) on porosity (ϵ_0), surface porosity (ϵ_{surf}), specific surface area (a_c) and permeability (K_D) values of the Kelvin-like foam sample of cylindrical strut shape ($\epsilon_0 = 0.80$) as presented on Figure 6.15a.

(A) 40^3 voxels (B) 80^3 voxels (C) 160^3 voxels (D) 320^3 voxels
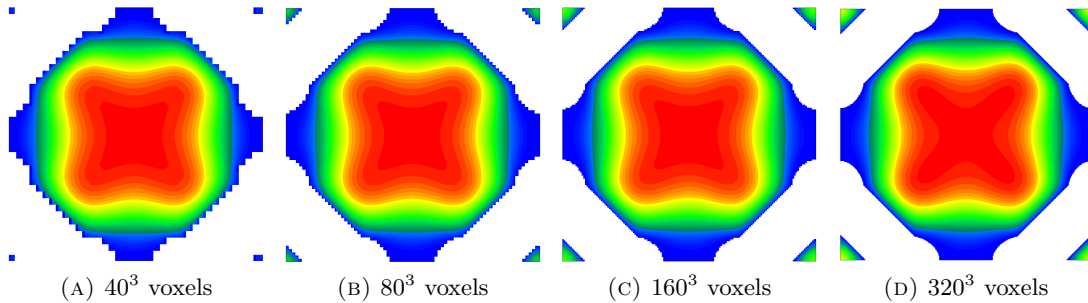
FIGURE 6.16. Velocity field in the middle plan for various resolutions

It can be observed that the morphological parameters start to converge at low resolution and capture precisely the 3D foam structure reported in the works of [55]. On the other hand, the main flow topology is well captured despite the lack of resolution (Figure 6.16). It is a well known phenomenon (see [50, 33, 12]), that comes from the fact that the specific surface area is nearly the same for all resolutions. The calculated permeability converges when improving the resolution. To perform parametric studies, 160^3 mesh cells are chosen to perform optimized numerical simulations without compromising accuracy of the data (see Table 6.17).

Resolution		Morphological properties			Permeability
Voxel size (m)	mesh cells	Porosity (ϵ_0)	Surface Porosity (ϵ_{surf})	Specific surface area (m^{-1})	(m^2)
0.00011	40^3	0.7656	0.5789	869.30	1.3764E-07
5.13e-05	80^3	0.774	0.6358	874.78	1.2729E-07
2.53e-05	160^3	0.7863	0.6735	867.17	1.2710E-07
1.26e-05	320^3	0.7921	0.691	864.58	1.2722E-07

TABLE 6.17. Mesh influence on morphology and permeability to quantify geometric discretization errors

Homothetic scaling-law

Several virtual homothetic foams of different cell diameters/sizes (d_{cell}) for a given porosity (ϵ_0) as well as foam samples of different porosities by changing the strut size for a given cell size have been generated as presented in Table 6.17a. By generating this set of virtual samples, we could assess individually the impact of pore size and porosity on permeability and thus evaluate the reliability of usage of classic description such as Ergun-like approach for such materials. As expected, the permeability varies proportionally to the square of the cell size i.e. $K_D \propto d_{cell}$ (Figure 6.17b). Note that for real samples, porosity and pore shape usually vary with size which makes it difficult to compare directly with experimental data.
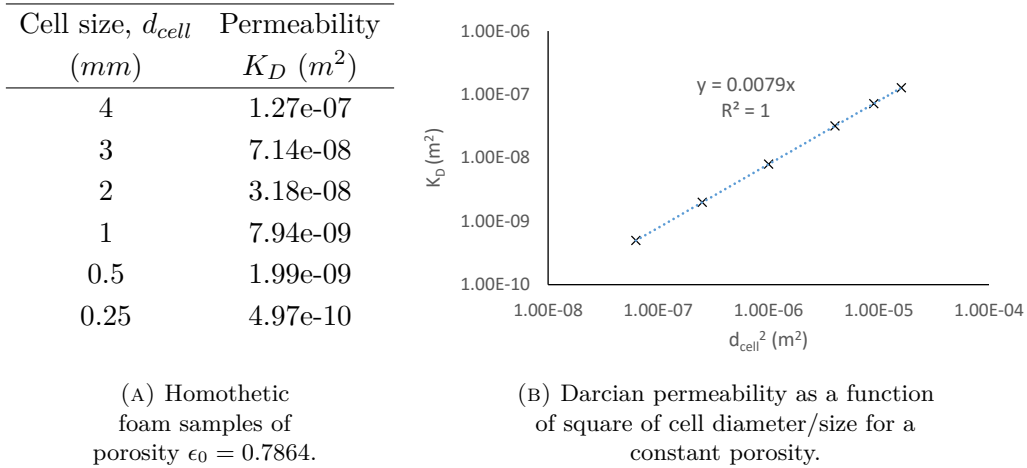
Cell size, d_{cell} (mm)	Permeability K_D (m^2)
4	1.27e-07
3	7.14e-08
2	3.18e-08
1	7.94e-09
0.5	1.99e-09
0.25	4.97e-10

(A) Homothetic foam samples of porosity $\epsilon_0 = 0.7864$.

(B) Darcian permeability as a function of square of cell diameter/size for a constant porosity.

FIGURE 6.17. Homothetic scaling-law results

Influence of porosity at constant cell diameter

The current numerical results obtained from our BGK-FVS method on structured grid are compared against numerical results of Kumar and Topin ([55]) obtained with a commercial software (StarCCM+) on polyhedral meshing. Reference [55] uses a non standard Darcy formalism, by considering the mean pressure gradient on the bulk phase, not on the fluid phase (6.10). Therefore, a scaling factor (surface porosity) has been applied to their results in order to compare the permeabilities. The classical analytical description of permeability with the morphological parameters of the foam structure is based on Ergun-like approach that was originally developed for packed bed of spheres. Most commonly, permeability is linked to two parameters i.e. pore size and porosity for isotropic and commercially available foams.

Porosity ϵ_0	Permeability K_D
0.6306	3.65E-08
0.6828	4.90E-08
0.7358	6.50E-08
0.7864	8.56E-08
0.8373	1.13E-07
0.8897	1.55E-07
0.9429	2.28E-07

TABLE 6.18. Morphological parameters and permeability values for virtual foam samples of constant cell size $d_{cell} = 4mm$
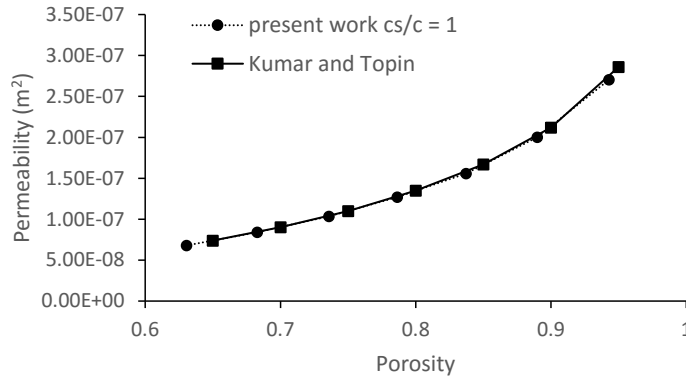

FIGURE 6.18. Assessment and comparison of Darcian permeability as a function of porosity for a given cell size $d_{cell} = 4mm$

From Figure 6.18 it is observed that the current numerical data and rescaled data from the works of [55] coincide perfectly. On the other hand, the offset points corresponding to the proposed work come from the different discretization (including some morphological error) as a function of resolution. The proposed discretization method causes errors at the walls according to the resolution where thin or sharp edges of strut shape of foam structure under voxel resolution are not visible.

7. **Discussion and related problems**

7.1. **Timestep limitation**

The CFL condition (5.16) limits the timestep Δt to be proportional to Δx^2. This condition is natural since we are solving the incompressible viscous Navier-Stokes system (5.1), which is parabolic. It is possible to bypass this limitation by making the viscous terms implicit in the scheme formulation (5.8), (5.9). Then it remains the stability conditions (5.17), (5.18), but no restriction on the timestep. This variant of our scheme amounts to solve a linear system at each timestep (discrete Laplace operator). Such implicit/explicit (IMEX) schemes have been considered for example in [36, 34, 64]. However the appropriate treatment of boundary conditions could become more difficult in such implicit formulation.

The choice of an explicit scheme can be nevertheless appropriate for three-dimensional flows with parallel programming (as in Test 6.2.3), since in this case implicit schemes become extremely costy in terms of computational time.

7.2. **Optimality of the scheme**

Compared to finite element methods, our BGK-FVS scheme has the property to be stable for an explicit use, even if the implicit treatment is possible (see above). It has however some limitations which are that it is so far limited to Cartesian grids, and the size of the grid is limited by the cell Reynolds stability condition (5.17). Compared to lattice Boltzmann methods, our scheme has the advantage to be second-order accurate and to satisfy a discrete entropy inequality. It is interesting also to compare our approach to a simple entropy satisfying penalization approximation, that would be to solve

$$\partial_t \mathbf{u} + \nabla_{\mathbf{x}} \cdot \left(\mathbf{u} \otimes \mathbf{u} - \frac{\nabla_{\mathbf{x}} \cdot \mathbf{u}}{\epsilon} \mathbf{I} \right) - \nu \Delta_{\mathbf{x}} \mathbf{u} = \mathbf{0}. \tag{7.1}$$

An explicit discretization of this equation implies a CFL condition $\Delta t \leq \epsilon \Delta x^2$, while the error in the divergence constraint of this approximation is of the order of ϵ. For a second-order approximation one needs $\epsilon \sim \Delta x^2$, and then the CFL condition is $\Delta t \leq \Delta x^4$, which is much worse than with our approach. Higher-order methods similar in spirit to ours, with artificial compressibility, are proposed in [78, 79], but do not satisfy an entropy inequality.

The property to be at the same time second-order accurate and entropy satisfying is common when discretizing second-order parabolic terms. It is however uncommon for first-order terms. The fact that our scheme has this property is due to the fact that the cell Reynolds stability condition (5.17) ensures that the advection terms are stabilized by the dissipation terms. A similar property holds when solving for example a one-dimensional scalar equation

$$\partial_t u + \partial_x f(u) - \nu \partial_{xx} u = 0 \tag{7.2}$$

by the scheme

$$u_i^{n+1} - u_i^n + \frac{\Delta t}{\Delta x} (\mathcal{F}_{i+1/2}^n - \mathcal{F}_{i-1/2}^n) = 0, \quad \mathcal{F}_{i+1/2}^n = \frac{f(u_i^n) + f(u_{i+1}^n)}{2} - \nu \frac{u_{i+1}^n - u_i^n}{\Delta x}. \tag{7.3}$$

It is second-order accurate, and monotone under the stability conditions $2\nu \Delta t \leq \Delta x^2$, $\Delta x |f'(u)|/(2\nu) \leq 1$. Similar considerations are used in [38].

7.3. **Low Mach number schemes**

When solving the compressible isentropic gas dynamics system considered in Subsection 3.2

$$\begin{aligned} \partial_t \rho + \nabla_{\mathbf{x}} \cdot (\rho \mathbf{u}) &= 0, \\ \partial_t (\rho \mathbf{u}) + \nabla_{\mathbf{x}} \cdot \left(\rho \mathbf{u} \otimes \mathbf{u} + \frac{P(\rho)}{\epsilon_M^2} \mathbf{I} \right) &= 0, \end{aligned} \tag{7.4}$$

where $\epsilon_M > 0$ is the order of magnitude of the Mach number, a classical problem is to derive schemes that are efficient in the compressible regime ϵ_M fixed, $\Delta x \to 0$, and that remain also stable and accurate in the limit $\epsilon_M \to 0$ (asymptotic preserving schemes). This means that when $\epsilon_M \to 0$ we should get a stable and accurate scheme for the incompressible Euler equations (3.17), (3.18). Note the difference with the problem handled in this paper, that has nonzero viscosity. It has been remarked for long that when considering generic scheme for (7.4), the numerical viscosity becomes in the low Mach number limit $\epsilon_M \to 0$ a nonzero unphysical viscosity, hence recovering the Navier-Stokes equations instead of Euler equations. This is even worse than this, the remaining viscosity is in general nonisotropic, even for Cartesian meshes. Several schemes have been proposed to cope with this problem, a general idea

being to remove some numerical viscosity, see [85, 68, 57, 37, 39, 36, 34, 64, 30] and the references therein. This leads to the so called *low Mach number schemes*. However, removing numerical viscosity leads to the fact that these schemes do no longer satisfy a discrete entropy inequality (the monotonicity is lost), even if they can be linearly stable [30].

With our approach we can propose a low Mach number scheme that satisfies a discrete entropy inequality. It consists in the formulation (5.8), (5.9), where the viscosity ν is related to $\epsilon > 0$ and $c > 0$ by (5.3), i.e. $2\nu = c\Delta x/\epsilon$. The constant c is a fixed parameter, and ϵ is taken

$$\epsilon = \max(\epsilon_M, \beta \Delta x^\alpha), \tag{7.5}$$

for some constants $\beta > 0$, $0 < \alpha < 1$. The scheme is monotone and satisfies the discrete entropy inequality (5.11) or (5.14) under the stability conditions (5.16), (5.17), (5.18). The last condition indeed determines the choice of c. It follows that

- If ϵ_M is not small ($\epsilon_M > \beta \Delta x^\alpha$), then $\epsilon = \epsilon_M$, and the scheme is a standard upwind scheme with viscosity η proportional to $\Delta x/\epsilon_M$. The CFL condition (5.16) is a standard hyperbolic CFL condition.

- If ϵ_M is small ($\epsilon_M \leq \beta \Delta x^\alpha$), then $\epsilon = \beta \Delta x^\alpha$. In this case the scheme is an approximation of the incompressible Navier-Stokes equations with viscosity proportional to $\Delta x^{1-\alpha}$. The error in the divergence constraint is of order $\epsilon^2 \sim \Delta x^{2\alpha}$, thus the scheme is finally consistent with the incompressible Euler equations, and the order of accuracy is $\min(1 - \alpha, 2\alpha)$. The CFL condition becomes $D\Delta t \leq \beta \Delta x^{1+\alpha}/c$.

The choice $\alpha = 1$ is the one made in this paper in order to keep a viscosity η that does not tend to zero. On the contrary, making the choice $\alpha = 1/3$ leads to a low Mach number scheme of order of accuracy $2/3$ in the low Mach number limit. To our knowledge it is an open problem to build a low Mach number scheme that satisfies a discrete entropy inequality and that is firsty-order accurate in the low Mach number limit.

Note that the cutoff in the formula (7.5) implies that for small Mach number $\epsilon_M < \beta \Delta x^\alpha$ the scheme no longer depends on ϵ_M. This property is justifed by the fact that the solution to the compressible system (7.4) is close to its incompressible limit up to a distance $\epsilon_M^2 < (\beta \Delta x^\alpha)^2$, which is less than the error induced by a scheme of order 2α. Thus if we consider schemes of order 2α, the value of ϵ_M between 0 and $\beta \Delta x^\alpha$ does not matter, and we can replace ϵ_M by $\beta \Delta x^\alpha$, which gives a scheme consistent with the incompressible limit.

8. **Conclusion**

With the aim of building numerical approximations of solutions to incompressible Navier-Stokes equations, we have developed and analysed BGK kinetic methods with $2D + 1$ velocities as shown in Figures 2.1, 2.2. Generalizing [16] we have proved an equivalence between these methods and explicit flux vector splitting schemes over a Cartesian mesh. We call these schemes BGK-FVS schemes. At the kinetic level, our approach is reminiscent of Lattice Boltzmann schemes, but compared to those, it has two main advantages. The first is that the schemes rely only on the mass and momentum densities rather than on the distribution functions, leading to low memory consumption, easiness of implementation and parallelization, and to straightforward application of boundary conditions. The second advantage of our approach is that by construction it leads to extremely robust schemes since they satisfy a discrete entropy inequality.

We have shown that our BGK-FVS schemes are related to compressible low Mach number schemes, but have the property to be second-order accurate for a good choice of the parameters, in particular the Mach number has to be proportional to the grid size. The parameters are set according to the

theoretical stability conditions, which are a parabolic CFL condition (5.16) that only involves the viscosity, and a stability condition (5.17) involving a cell Reynolds number, that is satisfied provided that the grid size is not too large. Another parameter c_s/c is involved in the method. Taking a small value for c_s/c gives a quite diffusive scheme (still being second-order accurate), whereas taking a large value gives a more accurate scheme. We have obtained a theoretical value of c_s/c given by (5.20) that yields strong stability properties, but our numerical tests show that it is better to take $c_s/c = 1$ (see Subsubsection 6.1.3), thus slightly violating the subcharacteristic condition. Our main scheme built on the Lax-Friedrichs flux vector splitting is written as (5.8), (5.9) with the pressure law (5.19). It falls into the class of artificial compressibility methods, which means that not only the velocity $\mathbf{u}$, but also a density ρ is recorded and updated at each timestep, mimicking at the discrete level the low Mach number limit from compressible to incompressible fluid dynamics. Within the artificial compressibility methods, the novelty here is the exceptionally good theoretical properties of our scheme. In particular, the entropy inequality (5.14) implies, when summed up over all the cells, that $\rho|\mathbf{u}|^2$ remains bounded in L^1 and that $\rho - \overline{\rho}$ remains of the order of Δx in L^2, both uniformly in time. The inequality (5.14) gives at the limit the energy inequality (5.2) associated to incompressible Navier-Stokes equations (5.1).

We have performed several benchmark test cases in two dimensions. Two test cases with analytic solution (Taylor-Green vortex and Poiseuille flow) show the effective second-order accuracy of the velocity and first-order accuracy of the pressure, with indeed second-order accuracy of the primitive in time of the pressure. In the steady case the pressure is itself second-order accurate. These results are optimal since it is known that the low Mach number limit implies the propagation of sound waves with time oscillations that can be not damped, depending on boundary conditions. The approximation of the free divergence constraint on the velocity is second-order accurate and does not grow with time. Two classical benchmark tests without anlytical solutions (Backward-facing step and Laminar flow past a square cylinder) have been performed and show good flow patterns, with second-order accuracy in terms of physical quantities of interest (attachment points, drag and lift coefficients).

These tests enable to set up a practical method to choose the grid size knowing the Reynolds number. One has indeed that the cell Reynolds number $\mathrm{Re}_{\mathrm{cell}}$ defined by (5.17) satisfies $\mathrm{Re}_{\mathrm{cell}} = \frac{1}{2}\mathrm{Re} \times \Delta x/L$, and it has to be at most 0.4 for a good accuracy with the choice $c_s/c = 1$. We have finally applied our BGK-FVS scheme to a three-dimensional test case with a complex geometry. The robustness of the scheme allows us to get reliable results.

The main limitations of our scheme, so far, are that it is restricted to Cartesian grids, the timestep is limited by $\Delta x^2/2D\nu$ (parabolic type CFL condition) and that there is a limitation of the grid size via the cell Reynolds number. For this reason, compared to lattice Boltzmann schemes, our scheme is well suited for moderately large Reynolds numbers of the order of a few hundreds.

Appendix A. **List of main parameters**

In this appendix we provide a list of the main parameters used in the paper with their significance.

- D: space dimension

- ν: given viscosity for the incompressible Navier-Stokes system

- Δt: timestep

- Δx: space step. The mesh is Cartesien with the same length Δx in all directions.

- λ: magnitude of the kinetic velocities defined in (2.41). This parameter disappears in the FVS formulation (2.47), (2.48).

- c: parameter involved in the definition of the Lax-Friedrichs flux vector splitting (2.65), and measuring the numerical viscosity. Our final scheme involves only the ratio c_s/c, that is chosen either as (5.20) or as $c_s/c = 1$.

- c_s: sound speed for the choice of the isothermal law (5.19)

- ϵ
  in Section 2.1: Knundsen number
  in Section 2.2: kinetic relaxation parameter
  in Section 3.2 and further on: Mach number or numerical Mach number. It is chosen in terms of other parameters as $\epsilon = c\Delta x/2\nu$ in Theorem 4.3 and further on.

- $\overline{\rho}$: constant reference density for the low Mach number limit, in Subsection 3.2 and further on.

- $P(\rho)$ and $e(\rho)$: pressure law and associated internal energy used for the compressible isentropic system in Subsection 3.2 and further on. In our numerical tests the pressure law is taken as (5.19) because ρ is close to $\overline{\rho}$ and only the linearization of the pressure law plays a significant role.

- p: pressure for the incompressible Navier-Stokes equations.

Acknowledgements

References

[1] T. Alazard. Low Mach number limit of the full Navier-Stokes equations. *Arch. Rat. Mech. Anal.*, 180:1–73, 2006.

[2] F. Alexander, H. Chen, S. Chen, and G. Doolen. A lattice Boltzmann model for compressible fluids. *Phys. Rev. A*, 46:1967–1970, 1992.

[3] D. Aregba-Driollet and R. Natalini. Discrete kinetic schemes for multidimensional systems of conservation laws. *SIAM J. Numer. Anal.*, 37:1973–2004, 2000.

[4] D. Aregba-Driollet, R. Natalini, and S. Tang. Explicit diffusive kinetic schemes for nonlinear degenerate parabolic systems. *Math. Comput.*, 73:63–94, 2004.

[5] B.F. Armaly, F. Durst, J.C.F. Pereira, and B. Schonung. Experimental and theoretical investigation of backward-facing step flow. *J. Fluid Mech.*, 127:473–496, 1983.

[6] P. Asinari, T. Ohwada, E. Chiavazzo, and A.F. Di Rienzo. Link-wise artificial compressibility method. *J. Comput. Phys.*, 231:5109–5143, 2012.

[7] E. Audusse, M.-O. Bristeau, B. Perthame, and J. Sainte-Marie. A multilayer Saint-Venant system with mass exchanges for shallow water flows. Derivation and numerical validation. *ESAIM: Math. Modelling Numer. Anal.*, 45:169–200, 2011.

[8] C. Levermore B. Boghosian. A cellular automaton for Burgers's equation. *Complex Systems*, 1:17–30, 1987.

[9] C. Bardos, F. Golse, and C.D. Levermore. Fluid dynamic limits of kinetic equations. I. Formal derivations. *J. Stat. Phys.*, 63:323–344, 1991.

[10] C. Bardos, F. Golse, and C.D. Levermore. Fluid dynamic limits of kinetic equations-II Convergence proofs for the Boltzmann-equation. *Comm. Pure Appl. Math.*, 46:667–753, 1993.

[11] C. Bardos, F. Golse, and C.D. Levermore. The acoustic limit for the Boltzmann equation. *Archive Rat. Mech. Anal.*, 153:177–204, 2000.

[12] J. Bernsdorf. *Simulation of complex flows and multi-physics with the lattice-Boltzmann method*. PhD thesis, Univ. Amsterdam, 2008.

[13] F. Berthelin and F. Bouchut. Kinetic invariant domains and relaxation limit from a BGK model to isentropic gas dynamics. *Asymptotic analysis*, 31:153–176, 2002.

[14] F. Berthelin and F. Bouchut. Relaxation to isentropic gas dynamics for a BGK system with single kinetic entropy. *Meth. and Appl. of Analysis*, 9:313–327, 2002.

[15] F. Bouchut. Construction of BGK models with a family of kinetic entropies for a given system of conservation laws. *J. Stat. Phys.*, 95:113–170, 1999.

[16] F. Bouchut. Entropy satisfying flux vector splittings and kinetic BGK models. *Numer. Math.*, 94:623–672, 2003.

[17] F. Bouchut. *Nonlinear stability of finite volume methods for hyperbolic conservation laws, and well-balanced schemes for sources*. Frontiers in Mathematics. Birkhäuser Verlag, Basel, Boston, Berlin, 2004.

[18] F. Bouchut. A reduced stability condition for nonlinear relaxation to conservation laws. *J. Hyp. Diff. Eq.*, 1:149–170, 2004.

[19] F. Bouchut and H. Frid. Finite difference schemes with cross derivatives correctors for multidimensional parabolic systems. *J. Hyp. Diff. Eq.*, 3:27–52, 2006.

[20] F. Bouchut, F.R. Guarguaglini, and R. Natalini. Diffusive BGK approximations for nonlinear multidimensional parabolic equations. *Indiana Univ. Math. J.*, 49:723–749, 2000.

[21] F. Bouchut, C. Klingenberg, and K. Waagan. A multiwave approximate Riemann solver for ideal MHD based on relaxation I - Theoretical framework. *Numer. Math.*, 108:7–42, 2007.

[22] F. Boyer and P. Fabrie. *Mathematical tools for the study of the incompressible Navier-Stokes equations and related models*. Applied Mathematical Sciences. Springer, New York, 2012.

[23] Y. Brenier. Résolution d'équations d'évolution quasilinéaires en dimension N d'espace à l'aide d'équations linéaires en dimension N+1. *J. Diff. Eq.*, 50:375–390, 1983.

[24] Y. Brenier. Averaged multivalued solutions for scalar conservation laws. *SIAM J. Numer. Anal.*, 21:1013–1037, 1984.

[25] D. Bresch, B. Desjardins, E. Grenier, and C.-K. Lin. Low Mach number limit of viscous polytropic flows: formal asymptotics in the periodic case. *Stud. Appl. Math.*, 109:125–149, 2002.

[26] D. Bresch and G. Métivier. Anelastic limits for Euler-type systems. *Appl. Math. Res. Express.*, 2010:119–141, 2010.

[27] M. Breuer, J. Bernsdorf, T. Zeiser, and F. Durst. Accurate computations of the laminar flow past a square cylinder based on two different methods: lattice-Boltzmann and finite-volume. *Intern. J. Heat Fluid Flow*, 21:186–196, 2000.

[28] C.-H. Bruneau and P. Fabrie. New efficient boundary conditions for incompressible Navier-Stokes equations: a well-posedness result. *Math. Model. Numer. Anal.*, 30:815–840, 1996.

[29] M. Carfora and R. Natalini. A discrete kinetic approximation for the incompressible Navier-Stokes equations. *ESAIM: Math. Modelling Numer. Anal.*, 42:93–112, 2008.

[30] C. Chalons, M. Girardin, and S. Kokh. An all-regime Lagrange-Projection like scheme for the gas dynamics equations on unstructured meshes. *Comm. Comput. Phys.*, 20:188–233, 2016.

[31] G.Q. Chen, C.D. Levermore, and T.-P. Liu. Hyperbolic conservation laws with stiff relaxation terms and entropy. *Comm. Pure Appl. Math.*, 47:787–830, 1994.

[32] S. Chikatamarla and I. Karlin. Lattices for the lattice Boltzmann method. *Phys. Rev. E*, 79:046701, 2009.

[33] D.S. Clague, B.D. Kandhai, R. Zhang, and P.M.A. Sloot. Hydraulic permeability of (un)bounded fibrous media using the lattice Boltzmann method. *Phys. Rev. E*, 61:616–625, 2000.

[34] F. Cordier, P. Degond, and A. Kumbaro. An asymptotic-preserving all-speed scheme for the Euler and Navier-Stokes equations. *J. Comput. Phys.*, 231:5685–5704, 2012.

[35] H. Darcy. *Les fontaines publiques de la ville de Dijon.* Dalmont, Paris, 1856.

[36] P. Degond and M. Tang. All speed scheme for the low Mach number limit of the isentropic Euler equations. *Commun. Comput. Phys.*, 10:1–31, 2011.

[37] S. Dellacherie. Analysis of Godunov type schemes applied to the compressible Euler system at low Mach number. *J. Comput. Phys.*, 229:978–1016, 2010.

[38] S. Dellacherie. Construction and analysis of lattice Boltzmann methods applied to a 1d convection-diffusion equation. *Acta Applicandae Mathematicae*, 131:69–140, 2014.

[39] S. Dellacherie and P. Omnes. On the Godunov scheme applied to the variable cross-section linear wave equation. *Springer Proc. Math.*, 4:313–321, 2011.

[40] S.M. Deshpande. Kinetic theory based new upwind methods for inviscid compressible flows. AIAA 24th Aerospace Science Meeting, Jan 6-9 1986, Nevada, USA, AIAA paper 86-0275, 1986.

[41] S.M. Deshpande. On the Maxwellian distribution, symmetric form and entropy conservation for the Euler equations. NASA Langley Research Centre, Hampton, VA, 1986, NASA TP2613, 1986.

[42] D. d'Humieres. Generalized lattice-Boltzmann equations. In *Rarefied Gas Dynamics: Theory and Simulations, vol. 159 of AIAA Progress in Astronautics and Astronautics*, pages 450–458. 1992.

[43] D. d'Humieres, I. Ginzburg, M. Krafczyk, P. Lallemand, and L.S. Luo. Multiple-relaxation-time lattice Boltzmann models in three dimensions. *Phil. Trans. Royal Soc. London series A-Math. Phys. Eng. Sci.*, 360:437–451, 2002.

[44] F. Dubois. Décomposition de flux et discontinuité de contact stationnaire. *C. R. Acad. Sci. Paris Sér. I Math.*, 330:847–850, 2000.

[45] F. Dubois. Equivalent partial differential equations of a lattice Boltzmann scheme. *Computers Math. Applic.*, 55:1441–1449, 2008.

[46] B. Elton, C. Levermore, and G. Rodrigue. Convergence of convective-diffusive lattice Boltzmann methods. *SIAM J. Numer. Anal.*, 32:1327–1354, 1995.

[47] E. Erturk. Numerical solutions of 2-d steady incompressible flow over a backward-facing step, part I: High Reynolds number solutions. *Computers Fluids*, 37:633–655, 2008.

[48] J. Gressier, P. Villedieu, and J.-M. Moschetta. Positivity of flux vector splitting schemes. *J. Comput. Phys.*, 155:199–220, 1999.

[49] F.J. Higuera and J. Jimenez. Boltzmann approach to lattice gas simulations. *Europhys. Letters*, 9:663–668, 1989.

[50] T. Inamuro, M. Yoshino, and F. Ogino. Accuracy of the lattice Boltzmann method for small Knudsen number with finite Reynolds number. *Phys. Fluids*, 9:3535, 1997.

[51] S. Jin, L. Pareschi, and G. Toscani. Diffusive relaxation schemes for multiscale discrete-velocity kinetic equations. *SIAM J. Numer. Anal.*, 35:2405–2439, 1998.

[52] S. Jin, L. Pareschi, and G. Toscani. Uniformly accurate diffusive relaxation schemes for multiscale transport equations. *SIAM J. Numer. Anal.*, 38:913–936, 2000.

[53] I.V. Karlin, A.N. Gorban, S. Succi, and V. Boffi. Maximum entropy principle for lattice kinetic equations. *Phys. Rev. Lett.*, 81:6–9, 1998.

[54] S. Klainerman and A. Majda. Compressible and incompressible fluids. *Comm. Pure Appl. Math.*, 35:629–651, 1982.

[55] P. Kumar and F. Topin. Micro-structural impact of different strut shapes and porosity on hydraulic properties of Kelvin-like metal foams. *Transport in Porous Media*, 105:57–81, 2014.

[56] P. Lallemand and L.-S. Luo. Theory of the lattice Boltzmann method: dispersion, dissipation, isotropy, Galilean invariance, and stability. *Physical Review E*, 61:6546–6562, 2000.

[57] X.-S. Li and C.-W. Gu. An all-speed Roe-type scheme and its asymptotic analysis of low Mach number behaviour. *J. Comput. Phys.*, 227:5144–5159, 2008.

[58] P.-L. Lions and N. Masmoudi. Incompressible limit for a viscous compressible fluid. *J. Math. Pures Appl.*, 77:585–627, 1998.

[59] J.C. Mandal and S.M. Deshpande. Higher order accurate kinetic flux vector splitting method for Euler equations. In *Nonlinear Hyperbolic Equations – Theory, Computation Methods, and Applications*, Notes on Numerical Fluid Mechanics Series 24, pages 384–392. 1989.

[60] B. Maury. *The respiratory system in equations*. Modeling, Simulation and Applications 7. Springer, Milan, 2013.

[61] G.R. McNamara and G. Zanetti. Use of the Boltzmann-equation to simulate lattice-gas automata. *Phys. Rev. Letters*, 61:2332–2335, 1988.

[62] R. Mei, L.S. Luo, P. Lallemand, and D. d'Humieres. Consistent initial conditions for lattice Boltzmann simulations. *Computers Fluids*, 35:855–862, 2006.

[63] X. Nicolas, M. Medale, S. Glockner, and S. Gounand. Benchmark solution for a three-dimensional mixed-convection flow, part 1: reference solutions. *Numer. Heat Transfer B-Fundamentals*, 60:325–345, 2011.

[64] S. Noelle, G. Bispen, K.R. Arun, M. Lukacova-Medvidova, and C.-D. Munz. A weakly asymptotic preserving low Mach number scheme for the Euler equations of gas dynamics. *SIAM J. Sci. Comput.*, 36:B989–B1024, 2014.

[65] T. Ohwada and P. Asinari. Artificial compressibility method revisited: Asymptotic numerical method for incompressible Navier-Stokes equations. *J. Comput. Phys.*, 229:1698–1723, 2010.

[66] T. Ohwada, P. Asinari, and D. Yabusaki. Artificial compressibility method and lattice Boltzmann method: similarities and differences. *Computers Math. Applic.*, 61:3461–3474, 2011.

[67] M.A. Ol'shanskii and V.M. Staroverov. On simulation of outflow boundary conditions in finite difference calculations for incompressible fluid. *Int. J. Numer. Meth. Fluids*, 33:499–534, 2000.

[68] H. Paillére, C. Viozat, A. Kumbaro, and I. Toumi. Comparison of low Mach number models for natural convection problems. *Heat and Mass Transfer*, 36:567–573, 2000.

[69] T.C. Papanastasiou, N. Malamataris, and K. Ellwood. A new outflow boundary-condition. *Int. J. Numer. Meth. Fluids*, 14:587–608, 1992.

[70] B. Perthame. Boltzmann type schemes for gas dynamics and the entropy property. *SIAM J. Numer. Anal.*, 27:1405–1421, 1990.

[71] B. Perthame and E. Tadmor. A kinetic-equation with kinetic entropy functions for scalar conservation laws. *Comm. Math. Phys.*, 136:501–517, 1991.

[72] A. Poux, S. Glockner, E. Ahusborde, and M. Azaiez. Open boundary conditions for the velocity-correction scheme of the Navier-Stokes equations. *Computers Fluids*, 70:29–43, 2012.

[73] D.I. Pullin. Direct simulation methods for compressible inviscid ideal-gas flow. *J. Comput. Phys.*, 34:231–244, 1980.

[74] R.D. Reitz. One-dimensional compressible gas dynamics calculations using the Boltzmann equations. *J. Comput. Phys.*, 42:108–123, 1981.

[75] P. Renard, A. Genty, and F. Stauffer. Laboratory determination of the full permeability tensor. *J. Geophysical Res.*, 106:26443–26452, 2001.

[76] L. Saint-Raymond. *Hydrodynamic limits of the Boltzmann equation.* Lecture Notes in Mathematics 1971. Springer, Berlin, 2009.

[77] S. Schochet. Fast singular limits of hyperbolic PDEs. *J. Diff. Eq.*, 114:476–512, 1994.

[78] A. Shah and L. Yuan. Flux-difference splitting-based upwind compact schemes for the incompressible Navier-Stokes equations. *Int. J. Numer. Meth. Fluids*, 61:552–568, 2009.

[79] A. Shah, L. Yuan, and A. Khan. Upwind compact finite difference scheme for time-accurate solution of the incompressible Navier-Stokes equations. *Applied Math. Comput.*, 215:3201–3213, 2010.

[80] A. Sohankar, C. Norberg, and L. Davidson. Low-Reynolds-number flow around a square cylinder at incidence: study of blockage, onset of vortex shedding and outlet boundary condition. *Intern. J. Numer. Meth. Fluids*, 26:39–56, 1998.

[81] S. Succi. *The lattice Boltzmann equation for fluid dynamics and beyond.* Numerical Mathematics and Scientific Computation. Oxford Science Publications, the Clarendon Press, Oxford University Press, New York, 2001.

[82] S. Succi, E. Foti, and F. Higuera. 3-dimensional flows in complex geometries with the lattice Boltzmann method. *Europhys. Letters*, 10:433–438, 1989.

[83] S. Succi, I.V. Karlin, and H. Chen. Role of the H theorem in lattice Boltzmann hydrodynamic simulations. *Rev. Mod. Phys.*, 74:1203–1220, 2002.

[84] L. Talon, D. Bauer, N. Gland, S. Youssef, H. Auradou, and I. Ginzburg. Assessment of the two relaxation time lattice-Boltzmann scheme to simulate Stokes flow in porous media. *Water Resources Res.*, 48:W04526, 2012.

[85] E. Turkel. Preconditioned methods for solving the incompressible and low speed compressible equations. *J. Comput. Phys.*, 72:277–298, 1987.

[86] W.A. Yong and L.S. Luo. Nonexistence of H theorem for some lattice Boltzmann models. *J. Stat. Phys.*, 121:91–103, 2005.