DISCRETE VELOCITY MODEL AND IMPLICIT SCHEME FOR THE BGK EQUATION OF RAREFIED GAS DYNAMICS

LUC MIEUSSENS

Mathématiques Appliquées de Bordeaux
Université Bordeaux I
33405 Talence Cedex, France
and
CEA-CESTA (DEV/SIS)
BP2 33114 Le Barp, France

We present a numerical method for computing transitional flows as described by the BGK equation of gas kinetic theory. Using the minimum entropy principle to define a discrete equilibrium function, a discrete velocity model of this equation is proposed. This model, like the continuous one, ensures positivity of solutions, conservation of moments, and dissipation of entropy. The discrete velocity model is then discretized in space and time by an explicit finite volume scheme which is proved to satisfy the previous properties. A linearized implicit scheme is then derived to efficiently compute steady-states; this method is then verified with several test cases.

1. Introduction

In a computation of the flow of a rarefied gas, the most commonly used numerical methods are probabilistic, such as the Direct Simulation Monte Carlo (DSMC) method. The main drawback of these methods is the frequent occurrence of noisy results. Numerous parameters (*i.e.* size of cells, number of representative molecules, number of samples) must be adjusted to obtain a good representation of the gas.

Recently, alternate deterministic methods based on discretization of the Boltzmann equation have been proposed in order to obtain higher accuracy (Rogier-Schneider [29], Buet [7]). However, the Boltzmann equation is difficult to solve numerically due to binary collisions. Generally, this leads to a near quadratic cost with respect to the velocity discretization which makes the approach very expensive. The method by Buet in [7] reduces this cost but introduces a type of stochasticity into the method.

In order to correct this problem, it is interesting to consider the simpler BGK

model proposed by Bathnagar, Gross and Krook in [28]

$$\partial_t f + v \cdot \nabla_x f = \frac{1}{\tau}(M[f] - f),$$

in which the binary interactions are modeled with a relaxation toward a local equilibrium $M[f]$. This modeling approach allows for the possibility of a linear cost of the discretization. In addition, the BGK model is known to be very accurate in near-equilibrium regions because it has the same fluid limit as the Boltzmann equation. This model is especially relevant in transitional flows where the classical Navier-Stokes equation fails to describe non-equilibrium phenomena, and where the Boltzmann equation is very expensive. Despite some recent studies [17] which have suggested the relevance of the BGK model for computing transport properties far from equilibrium, its use for extreme rarefied flows is questionable.

It is important to mention that same other relaxation models exist to fit realistic values of the Prandtl number given by a Chapman-Enskog expansion of BGK, but this approach is not investigated here. For more details on this subject, see the ES model (recently proved to be entropic in [2]), the S-model [31], or the BGK model with a velocity-dependent relaxation time [32].

When solving BGK equation, the main difficulty is the velocity discretization. Most numerical methods (like Issautier [19], Aoki-Kanba-Takata [3], Yang-Huang [34]) lack the properties of conservation of mass, momentum and energy, as well as the entropy property. This leads to algorithms which are expensive due to the fine velocity meshes which guarantee robustness, in particular with implicit schemes.

An elegant approach which allows to have conservation and entropy properties with Boltzmann equation approximations is the discrete kinetic theory. It has been widely developed by Gatignol [18] who constructed some discrete velocity models (DVM) of Boltzmann equation. In addition, she demonstrated conservation properties and some other interesting results. These models were initially used to simplify the mathematical study of the Boltzmann equation. However, the recent numerical methods developed by [29,7] also use the DVM with the above mentioned properties.

In this paper, we present a similar approach to that of [29] and [7] for the development of a DVM of the BGK equation. We propose a non-trivial discrete approximation of $M[f]$, defined by a minimum entropy principle. In addition, we advance the work of [12] and prove the existence and uniqueness result for this principle. Our model is thus a conservative and entropy decreasing DVM. To our knowledge, it is the first time that such a model is presented for the BGK equation (it is relevant to note that the conservative method of Perthame-Coron [14] has the above mentioned properties but cannot be viewed as a DVM). So, as opposed to non-conservative methods for which a large number of velocities is needed to recover conservation properties, our model appears to be very economic.

The second aim of this paper is to provide a fast and robust algorithm for steady-state computations. Many numerical schemes have been proposed for unsteady BGK equation (see the references below), but the simulation of stationary flows involves additional difficulties. For these flows, one may either solve the steady BGK equation with an iterative scheme (due to the nonlinearity of the source term), or solve the unsteady equation and then let time evolve to infinity.

In the first method, the nonlinearity may be treated by a Newton algorithm, but it is well known in CFD that such a method is not very robust and may not converge. It may also be treated by a fixed point technique like Aoki, Kanba, Takata in [3] with the scheme $v \cdot \nabla_x f^{n+1} = \frac{1}{\tau^n}(M[f^n] - f^{n+1})$, but this leads to the uncoupling of the gain and loss terms, and may converge very slowly. This can be viewed as a semi-implicit method for the non-stationary equation where only the loss term $-f$ is implicit (because it gives negative distributions). In this paper, we demonstrate that this method is slow.

In the second method, the main issue is that in order to guarantee positivity of distributions, the time step Δt has to satisfy two severe conditions:

- Δt must be smaller than the relaxation time τ. This is very restrictive in transitional regimes for which τ is small.
- A molecule should not cross more than one cell of the mesh during Δt. This condition is known as the classical CFL condition for convection. Thus Δt must be very small in high velocity regimes, such as reentry problems in aerodynamics applications.

These restrictions on the time step imply that a very large number of iterations are then required in order to reach steady state.

The first constraint has been widely studied. For example, Perthame and Coron in [14] proposed a relaxation scheme (where both the exponential form of BGK equation and a splitting method are used) which is stable for any arbitrary small τ. A similar method for other kinetic equations has been given by Gabetta, Pareschi and Toscani in [16]. The same idea is also used by Issautier in [19] with a Lagrangian discretization instead of the splitting method.

The second constraint can be more restrictive than the first one. However, it has received considerable less attention, and has yet to be properly taken care of. A classical solution in CFD is to use an implicit convective term during the time integration (see for instance [35] for Euler equations). But as mentioned by Waluś in [33], the splitting methods are incompatible with implicit procedures because transport and collision processes are greatly decoupled. In addition the Lagrangian discretization of [19] is not well designed for steady applications (the moving mesh is appropriate for internal unsteady flows but obviously not for steady flows).

The method introduced by Yang and Huang in [34] is completely different and appears to be very interesting for steady computations. They propose an implicit scheme based on an explicit finite volume scheme of the BGK equation and use all the CFD techniques for hyperbolic systems. However, they stress that only the

loss term is implicit because the linearization of $M[f]$ is difficult. Their method is therefore a semi-implicit one and has the same drawbacks of slow convergence as the method of [3] cited above. In this paper, we present a similar approach, but we do make use of a fully implicit source term.

Finally, it is worth mentioning some drawbacks of implicit methods such as spurious solutions and bad long-time behavior (also referred to by Bobylev and Struckmeier in [5] and by Jin in [20]). These problems are not investigated here.

The remainder of the paper follows logically. The properties of the BGK equation as well as a short list of notations are shown in Section 2. In Section 3, we present our DVM and prove the existence of a discrete equilibrium. The properties of the model are then described. In Section 4 we construct an explicit finite volume scheme which is proved to be conservative and have dissipated entropy. Then, a linearized implicit scheme is derived from the previous scheme, and an algorithm for solving the large linear systems is presented. In Section 5, some numerical applications are given, and a brief study of the influence of the discretization on the results is shown. Finally, several comparisons with experiments and the DSMC method are presented.

2. BGK equation

Positions x and velocities v of molecules are given in $\mathbb{R}^D$. The integral of any function $g(v)$ is denoted by

$$\langle g \rangle = \int_{\mathbb{R}^D} g(v)\, dv.$$

Classical fluid quantities are obtained by integration of f on velocity space. For instance, density, momentum, and total energy are given by

$$\rho = \langle f \rangle \in \mathbb{R}, \qquad \rho u = \langle v f \rangle \in \mathbb{R}^D, \qquad E = \langle \tfrac{1}{2}|v|^2 f \rangle \in \mathbb{R}. \tag{2.1}$$

These variables are, therefore, the $D+2$ first moments of f with respect to v. Microscopic quantities $1, v, \frac{1}{2}|v|^2$ (the so-called collisional invariants) can be denoted by the vector

$$\boldsymbol{m}(v) = \left(1, v, \tfrac{1}{2}|v|^2\right)^T \in \mathbb{R}^{D+2}. \tag{2.2}$$

By analogy, the $D+2$ first moments of f are denoted by the vector

$$\boldsymbol{\rho} = (\rho, \rho u, E)^T, \tag{2.3}$$

of which components may be numbered from 0 to $D+1$: $\boldsymbol{\rho} = (\rho^{(i)})_{i=0..D+1}$. Definition (2.1) may thus be rewritten in a shorter form

$$\boldsymbol{\rho} = \langle \boldsymbol{m} f \rangle, \tag{2.4}$$

and one then says that $\boldsymbol{\rho}$ is realizable by f, or that f realizes $\boldsymbol{\rho}$. Finally, we define the kinetic entropy of f by

$$H[f] = \langle f \log f \rangle, \tag{2.5}$$

and we recall that the temperature T and the pressure p of the gas are defined by

$$T = \frac{1}{DR}\langle |v-u|^2 f\rangle, \qquad p = \rho RT, \tag{2.6}$$

where R is the gas constant.

The BGK equation (see [28]) is the following relaxation time model

$$\partial_t f + v\cdot\nabla_x f = \frac{1}{\tau}\left(M[f]-f\right), \tag{2.7}$$

where the relaxation time is defined by

$$\tau^{-1} = C\rho T^{1-\omega}, \tag{2.8}$$

and ω is the exponent of the viscosity law of the gas (see [10]). The source term represents the fact that the distribution function tends towards a local equilibrium distribution $M[f]$, defined by

$$M[f] = \exp\left(\boldsymbol{\alpha}\cdot\boldsymbol{m}(v)\right). \tag{2.9}$$

This Gaussian distribution, called Maxwellian, has the same moments as f. The parameters $\boldsymbol{\alpha}$ are functions of t and x and may be expressed in terms of ρ, u, T through the invertible relation (if $\rho, T > 0$)

$$\boldsymbol{\alpha} = \left(\log\left(\frac{\rho}{(2\pi RT)^{D/2}}\right) - \frac{|u|^2}{2RT}, \frac{u}{RT}, -\frac{1}{RT}\right)^T. \tag{2.10}$$

It can easily be seen that M is the unique solution of the following entropy minimization problem (see for instance [27])

$$(\mathcal{P}) \qquad H[M] = \min\left\{H[g],\ g\geq 0\ s.t.\ \langle \boldsymbol{m}g\rangle = \boldsymbol{\rho}\right\}. \tag{2.11}$$

This simply means that the local equilibrium state minimizes the entropy of all the possible states leading to the same macroscopic properties.

With this characterization of the local Maxwellian equilibrium, the following properties of conservation of mass, momentum, energy, and dissipation of entropy may easily be proved, at least formally,

$$\partial_t\langle \boldsymbol{m}f\rangle + \nabla_x\langle \boldsymbol{m}vf\rangle = 0, \tag{2.12}$$

$$\partial_t\langle f\log f\rangle + \nabla_x\langle vf\log f\rangle \leq 0. \tag{2.13}$$

Furthermore, it is possible to check that solutions of (2.7) are nonnegative.

We point out that in a numerical scheme, the preservation of these properties is essential to a robust and economic discretization. Yet, most of the commonly accepted numerical methods (particular, discrete ordinate or finite difference methods) fail to satisfy these properties. We present a discrete velocity model for which the conditions of positivity of solutions, conservation of moments, and dissipation of entropy are fully satisfied.

3. Discrete Velocity Model

Let $\mathcal{K}$ be a set of N multi-indexes of $\mathbb{N}^D$, defined by $\mathcal{K} = \left\{k=(k^{(i)})_{i=1}^D,\ k^{(i)} \leq K^{(i)}\right\}$, where $\{K^{(i)}\}$ are some given bounds. We define a discrete velocity set $\mathcal{V}$ of $\mathbb{R}^D$ by

$$\mathcal{V} = \{v_k = k\Delta v + a, k \in \mathcal{K}\}, \tag{3.14}$$

where a is an arbitrary vector of $\mathbb{R}^D$ and Δv is a scalar [a]. The microscopic quantities on $\mathcal{V}$ are denoted by $\boldsymbol{m}_k = \left(1, v_k, \frac{1}{2}|v_k|^2\right)^T$. The "continuous" velocity distribution f is then replaced by a N-vector $f_{\mathcal{K}}(t,x) = (f_k(t,x))_{k\in\mathcal{K}}$ where each component $f_k(t,x)$ is assumed to be an approximation of $f(t,x,v_k)$. The fluid quantities are thus given as in continuous case, except that integrals on $\mathbb{R}^D$ are replaced by discrete sums on $\mathcal{V}$. That is, setting

$$\langle g\rangle_{\mathcal{K}} = \sum_{k\in\mathcal{K}} g_k \Delta v^D \tag{3.15}$$

for any vector $g \in \mathbb{R}^N$, we can define discrete moments and discrete entropy of $f_{\mathcal{K}}$ by

$$\boldsymbol{\rho}_{\mathcal{K}} = \langle \boldsymbol{m} f_{\mathcal{K}}\rangle_{\mathcal{K}}, \tag{3.16}$$

$$H_{\mathcal{K}} = \langle f_{\mathcal{K}} \log f_{\mathcal{K}}\rangle_{\mathcal{K}}. \tag{3.17}$$

Our discrete velocity BGK model follows as a set of N equations

$$\partial_t f_k + v_k \cdot \nabla_x f_k = \frac{1}{\tau}\left(\mathcal{E}_k - f_k\right), \qquad \forall k \in \mathcal{K} \tag{3.18}$$

and the main problem is to define an approximation $\mathcal{E}_{\mathcal{K}}$ of the Maxwellian equilibrium $M[f]$ such that conservation properties (2.12) and entropy property (2.13) still hold. First we note that the natural approximation (used by Yang and Huang in [34])

$$\mathcal{E}_k = M[f_{\mathcal{K}}](v_k), \qquad \forall k \in \mathcal{K}, \tag{3.19}$$

cannot satisfy these requirements. Instead, we propose to use the discrete version of entropy minimization problem (2.11). Let $\mathcal{E}_{\mathcal{K}}$ be defined by the minimum of discrete entropy, with the constraints that it must have the same moments as f, *i.e.* $\mathcal{E}_{\mathcal{K}}$ is the solution of the following problem $(\mathcal{P}_{\mathcal{K}})$

$$(\mathcal{P}_{\mathcal{K}}) \qquad H_{\mathcal{K}}[\mathcal{E}_{\mathcal{K}}] = \min\left\{H_{\mathcal{K}}[g],\ g \geq 0 \in \mathbb{R}^N \ s.t.\ \langle \boldsymbol{m} g\rangle_{\mathcal{K}} = \boldsymbol{\rho}_{\mathcal{K}}\right\}. \tag{3.20}$$

Obviously, it must be checked that this problem has a unique and easily solvable solution (solving directly $(\mathcal{P}_{\mathcal{K}})$ in $\mathbb{R}^N$ would be numerically expensive).

In the continuous case, the condition $\rho, T > 0$ is sufficient to characterize the solution of (2.11) by the Maxwellian distribution. However, this is not true for the discrete case where explicit computations are not possible. To this end, we have

[a] The results presented in this section remain valid if Δv depends on the direction (i) and on the index k .

then proved that under a natural assumption on $\mathcal{V}$, the discrete equilibrium $\mathcal{E}_\mathcal{K}$ has an exponential form *if, and only if,* a "strict realizability" condition is fulfilled by $\boldsymbol{\rho}_\mathcal{K}$. This result has been announced in [26,24,12,25,13].

Theorem 3.1 Let $\boldsymbol{\rho}$ be a vector in $\mathbb{R}^{D+2}$, such that the set $E_{\boldsymbol{\rho}} = \{g \geq 0 \in \mathbb{R}^N \ s.t. \ \langle \boldsymbol{m} g \rangle_\mathcal{K} = \boldsymbol{\rho}\}$ of nonnegative discrete distributions realizing $\boldsymbol{\rho}$ is not empty. Then, the problem $(\mathcal{P}_\mathcal{K})$ has a unique solution $\mathcal{E}_\mathcal{K}$ called discrete equilibrium. Moreover, we assume that $\mathcal{V}$ is such that microscopic quantities satisfy

$$\{\boldsymbol{m}_k, k \in \mathcal{K}\} \text{ is of rank } D+2. \tag{3.21}$$

Then there exists a unique vector $\boldsymbol{\alpha}$ in $\mathbb{R}^{D+2}$ such that the following exponential characterization holds

$$\mathcal{E}_k = \exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}_k), \qquad \forall k \in \mathcal{K}, \tag{3.22}$$

if and only if $\boldsymbol{\rho}$ is strictly realizable, *i.e.*

$$\exists g \in E_{\boldsymbol{\rho}} \ s.t. \ g > 0. \tag{3.23}$$

Remark 3.1 Due to the above result, the computation of $\mathcal{E}_\mathcal{K}$ does not require the solution of an expensive minimization problem in $\mathbb{R}^N$. Instead, only the computation of the vector $\boldsymbol{\alpha}$ in $\mathbb{R}^{D+2}$ is necessary. This vector $\boldsymbol{\alpha}$ is the unique solution of the nonlinear set of $D+2$ equations

$$\langle \boldsymbol{m} \exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}) \rangle_\mathcal{K} = \boldsymbol{\rho}_\mathcal{K}, \tag{3.24}$$

since $\mathcal{E}_\mathcal{K}$ realizes $\boldsymbol{\rho}_\mathcal{K}$. This set may be solved by a Newton algorithm (see Sec.). This relation has already been used for initial conditions of numerical simulations (see [30,6]). However, to our knowledge, it is the first time that existence and uniqueness is proved.

Remark 3.2 This result is generalized in [12] for discrete equilibrium associated to higher order moments, in order to define discrete version of Levermore's moment closures (see [22] for Levermore's closures and [11] for discrete theory and numerical applications).

Remark 3.3 It will be shown in the next section that condition on $\mathcal{V}$ is not very restrictive. However, condition (3.23) is more restrictive than in the continuous case. The reason of this difference is mainly the number of degrees of freedom: the smaller the number of discrete velocities is, the fewer degrees of freedom there are to realize $\boldsymbol{\rho}$, *i.e.* , to find g in $E_{\boldsymbol{\rho}}$, and therefore the more difficult it is to find a discrete Gaussian distribution in $E_{\boldsymbol{\rho}}$. For instance with $D=1$ and $N=3$, if we set $\mathcal{V} = \{-1, 0, 1\}$ and $\boldsymbol{\rho} = (2, 0, 2)$, then $E_{\boldsymbol{\rho}}$ has only one element $g = (1, 0, 1)$. Therefore discrete equilibrium $\mathcal{E}_\mathcal{K}$ is equal to g, and cannot be an exponential (since

$\mathcal{E}_2 = 0$), whereas $\rho, T > 0$ implies that the continuous equilibrium is a Maxwellian. In the continuous-velocity limit, the number of degrees of freedom becomes infinite, and condition (3.23) is always satisfied provided that the natural conditions $\rho, T > 0$ are fulfilled.

Proof of Theorem 3.1. First, note that the existence and uniqueness of $\mathcal{E}_\mathcal{K}$ is easily obtained since the mapping $H_\mathcal{K}$ is clearly continuous, coercive, and strictly convex on the closed subset $E_{\boldsymbol{\rho}}$ (we do not write the simple proof of these assertions).

The non-trivial part of this theorem is the characterization of $\mathcal{E}_\mathcal{K}$. For that purpose, let us introduce the mapping $J(\boldsymbol{\beta}) = \langle \exp(\boldsymbol{\beta} \cdot \boldsymbol{m}) \rangle_\mathcal{K} - \boldsymbol{\beta} \cdot \boldsymbol{\rho}$ defined on $\mathbb{R}^{D+2}$, and a second minimization problem

$$(\mathcal{P}'_\mathcal{K}) \qquad \boldsymbol{\alpha} \; s.t. \; J(\boldsymbol{\alpha}) = \min\{J(\boldsymbol{\beta}); \boldsymbol{\beta} \in \mathbb{R}^{D+2}\}. \tag{3.25}$$

It is clear that this problem implies the original problem $(\mathcal{P}_\mathcal{K})$ (see (3.20)). In fact, a solution $\boldsymbol{\alpha}$ of $(\mathcal{P}'_\mathcal{K})$ necessarily satisfies the extremum relation

$$J'(\boldsymbol{\alpha}) = \langle \boldsymbol{m} \exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}) \rangle_\mathcal{K} - \boldsymbol{\rho} = 0. \tag{3.26}$$

Therefore, if we define $\mathcal{E}_k = \exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}_k)$ for all $k \in \mathcal{K}$, then $\mathcal{E}_\mathcal{K}$ realizes $\boldsymbol{\rho}$. Finally, this exponential form allows us to conclude that $\mathcal{E}_\mathcal{K}$ is the solution of $(\mathcal{P}_\mathcal{K})$. As in the continuous case (see [27]), by convexity of $t \log t$ and using (3.26), we have for all $g \in E_{\boldsymbol{\rho}}$

$$\begin{aligned} H_\mathcal{K}[\mathcal{E}_\mathcal{K}] &\leq H_\mathcal{K}[g] - \langle (1 + \boldsymbol{\alpha} \cdot \boldsymbol{m})(g - \mathcal{E}_\mathcal{K}) \rangle_\mathcal{K} \\ &= H_\mathcal{K}[g]. \end{aligned}$$

We want now to prove the existence and uniqueness of a solution to $(\mathcal{P}'_\mathcal{K})$. The advantage of this second problem is that conditions (3.21) and (3.23) appear naturally. First, we shall prove J to be strictly convex. To that effect, it is sufficient to note that the second derivative of J satisfies

$$\begin{aligned} J''(\boldsymbol{\beta})(\boldsymbol{\xi}, \boldsymbol{\xi}) &= \boldsymbol{\xi}^T \langle \boldsymbol{m} \otimes \boldsymbol{m} \exp(\boldsymbol{\beta} \cdot \boldsymbol{m}) \rangle_\mathcal{K} \boldsymbol{\xi} \\ &= \sum_{k \in \mathcal{K}} \exp(\boldsymbol{\beta} \cdot \boldsymbol{m}) (\sum_i \xi^{(i)} \boldsymbol{m}_k^{(i)})^2 \Delta v^D \end{aligned}$$

Therefore $J''(\boldsymbol{\beta})(\boldsymbol{\xi}, \boldsymbol{\xi}) \geq 0$ for all $\boldsymbol{\xi} \in \mathbb{R}^{D+2}$. In addition, assumption (3.21) reads

$$\boldsymbol{\xi} \cdot \boldsymbol{m}_k = 0 \quad \forall k \in \mathcal{K} \Leftrightarrow \boldsymbol{\xi} = 0, \tag{3.27}$$

which implies $J''(\boldsymbol{\beta})(\boldsymbol{\xi}, \boldsymbol{\xi}) > 0$ if $\boldsymbol{\xi} \neq 0$. This proves that $J''(\boldsymbol{\beta})$ is positive definite for all $\boldsymbol{\beta} \in \mathbb{R}^{D+2}$, and hence J is strictly convex. Now, we shall prove the coercivity of J. This is accomplished in two steps.

step 1: The first step will ensure that J is coercive "on each direction of $\mathbb{R}^{D+2}$", that is

$$\forall \boldsymbol{\omega} \in S^{D+1}, \quad J(\lambda \boldsymbol{\omega}) = \sum_{k \in \mathcal{K}} \exp(\lambda \boldsymbol{\omega} \cdot \boldsymbol{m}_k) \Delta v^D - \lambda \boldsymbol{\omega} \cdot \boldsymbol{\rho} \xrightarrow[\lambda \to +\infty]{} +\infty$$

locally uniformly around each $\boldsymbol{\omega}$. This problem is split into two cases. First, assuming that there exists k_0 such that $\boldsymbol{\omega} \cdot \boldsymbol{m}_{k_0} > 0$, we obtain $0 < c_1(\boldsymbol{\omega}) \leq \tilde{\boldsymbol{\omega}} \cdot \boldsymbol{m}_{k_0}$ for every $\tilde{\boldsymbol{\omega}}$ of a small enough neighborhood $B(\boldsymbol{\omega}, \varepsilon(\boldsymbol{\omega}))$ of $\boldsymbol{\omega}$. Consequently

$$J(\lambda\tilde{\boldsymbol{\omega}}) \geq \exp(\lambda c_1(\boldsymbol{\omega}))\Delta v^D - \lambda c_2(\boldsymbol{\omega}, \boldsymbol{\rho}), \tag{3.28}$$

which tends to $+\infty$ uniformly on $B(\boldsymbol{\omega}, \varepsilon(\boldsymbol{\omega}))$. In the second case $\boldsymbol{\omega} \cdot \boldsymbol{m}_k \leq 0$ for all k, and property (3.21) implies that there exists k_0 such that $\boldsymbol{\omega} \cdot \boldsymbol{m}_{k_0} < 0$. Therefore, $\boldsymbol{\omega} \cdot \boldsymbol{\rho} < 0$ and there exists a neighborhood $B(\boldsymbol{\omega}, \varepsilon(\boldsymbol{\omega}))$ of $\boldsymbol{\omega}$ in which $\tilde{\boldsymbol{\omega}} \cdot \boldsymbol{\rho} \leq c_3(\boldsymbol{\omega}) < 0$. Hence

$$J(\lambda\tilde{\boldsymbol{\omega}}) \geq -\lambda\tilde{\boldsymbol{\omega}} \cdot \boldsymbol{\rho} \geq -\lambda c_3(\boldsymbol{\omega}) \tag{3.29}$$

which tends to $+\infty$ uniformly on $B(\boldsymbol{\omega}, \varepsilon(\boldsymbol{\omega}))$.

step 2: We can now prove that J tends globally uniformly to $+\infty$ on each direction $\boldsymbol{\omega}$. Let us consider the neighborhoods $B(\boldsymbol{\omega}, \varepsilon(\boldsymbol{\omega}))$ defined above for all $\boldsymbol{\omega}$ of S^{D+1}. It is clear that the union of these open sets covers S^{D+1}. As we know S^{D+1} is compact, this union contains therefore a finite covering, *i.e.* , there exists $\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \ldots \boldsymbol{\omega}_L$ of S^{D+1} such that

$$S^{D+1} \subset \bigcup_{i=1}^{L} B(\boldsymbol{\omega}_i, \varepsilon(\boldsymbol{\omega}_i)). \tag{3.30}$$

From the previous step, we know that for all i and for all $M > 0$, there exists $R(\boldsymbol{\omega}_i, M) > 0$ such that

$$\lambda \geq R(\boldsymbol{\omega}_i, M) \Rightarrow J(\lambda\tilde{\boldsymbol{\omega}}) \geq M \qquad \forall \tilde{\boldsymbol{\omega}} \in B(\boldsymbol{\omega}_i, \varepsilon(\boldsymbol{\omega}_i)), \tag{3.31}$$

then we can define $R(M) = \max_{i=1..L} R(\boldsymbol{\omega}_i, M)$ which is finite. It is thus clear that from (3.31)

$$\forall \tilde{\boldsymbol{\omega}} \in S^{D+1} \qquad \lambda \geq R(M) \Rightarrow J(\lambda\tilde{\boldsymbol{\omega}}) \geq M,$$

i.e. $J(\lambda\tilde{\boldsymbol{\omega}})$ tends toward infinity uniformly in $\tilde{\boldsymbol{\omega}}$. The coercivity is then obvious.

It is now a classical result of optimization that problem $(\mathcal{P}'_{\mathcal{K}})$ has a unique solution $\boldsymbol{\alpha}$ in $\mathbb{R}^{D+2}$. As was noted at the beginning of the proof, this implies that $\mathcal{E}_{\mathcal{K}} = (\exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}_k))_{k\in\mathcal{K}}$ is the unique solution of $(\mathcal{P}_{\mathcal{K}})$. We have thus proved that under assumption (3.21), the solution of $(\mathcal{P}_{\mathcal{K}})$ is characterized if condition of strict realizability (3.23) is fulfilled. The converse of this statement is trivial, since $g = \mathcal{E}_{\mathcal{K}} = (\exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}_k))_{k\in\mathcal{K}}$ is necessarily strictly positive and realizes $\boldsymbol{\rho}$. □

The model is now completely defined except in the case where $E_{\boldsymbol{\rho}_{\mathcal{K}}}$ is empty. For that case, we can set $\mathcal{E}_{\mathcal{K}} = 0$. This has practically no interest because the model implicitly contains the fact that $\boldsymbol{\rho}_{\mathcal{K}}$ is realized by $f_{\mathcal{K}} \geq 0$. We can now state the model is well defined and has the expected properties.

Theorem 3.2 Let f_0 be a strictly positive vector of $\mathbb{R}^N$. Consider the initial value problem associated with model (3.18), where $\mathcal{E}_{\mathcal{K}}$ is defined either by $(\mathcal{P}_{\mathcal{K}})$, or by

$\mathcal{E}_\mathcal{K} = 0$ (if $E_{\boldsymbol{\rho}_\mathcal{K}} = \emptyset$). If this problem has a solution $f_\mathcal{K}$, then we have (at least formally)

$$f_k(t,x) > 0 \qquad \forall k, t, x, \tag{3.32}$$

$$\mathcal{E}_k = \exp(\boldsymbol{\alpha} \cdot \boldsymbol{m}_k) \qquad \forall k, \tag{3.33}$$

$$\partial_t \langle \boldsymbol{m} f_\mathcal{K} \rangle_\mathcal{K} + \nabla_x \langle \boldsymbol{m} v f_\mathcal{K} \rangle_\mathcal{K} = 0, \tag{3.34}$$

$$\partial_t \langle f_\mathcal{K} \log f_\mathcal{K} \rangle_\mathcal{K} + \nabla_x \langle v f_\mathcal{K} \log f_\mathcal{K} \rangle_\mathcal{K} \leq 0. \tag{3.35}$$

Proof. The proof of assertions (3.32,3.34,3.35) is very similar to that of the continuous case, and is left to the reader. We only need to show that the discrete equilibrium is always an exponential. Noting $f_\mathcal{K}$ is always strictly positive proves that the discrete moment $\boldsymbol{\rho}_\mathcal{K}$ is obviously strictly realizable, and theorem (3.1) then allows us to conclude. □

Note here that these properties permit us to obtain existence and uniqueness results for model (3.18), as well as convergence toward the continuous BGK (see [23]).

We feel that it is necessary to explain why condition (3.21) on $\mathcal{V}$ is not greatly restrictive. The following result demonstrates that it is sufficient to take a Cartesian grid with at least two points in each direction (including 0), and at least three points in a given direction.

Proposition 3.1 Let $\mathcal{V}$ be defined by (3.14) with

$$K^{(i)} \geq 1 \qquad \forall i = 1..D. \tag{3.36}$$

If there exists one direction i such that

$$K^{(i)} \geq 2, \tag{3.37}$$

then $\{\boldsymbol{m}_k, k \in \mathcal{K}\}$ satisfies (3.21).

Proof. Let us first assume that $a = 0$ (see (3.14)). Moreover, up to the renumbering of axes, we can assume there are at least three points in direction 1 (set $i = 1$ in (3.37)). With the assumptions of the proposition, we can then find a subset of $D+2$ multi-indexes $\{k_l\}_{l=1..D+2}$ of $\mathcal{K}$ such that

$$\begin{gathered} v_{k_1} = 0, \\ v_{k_l} = (v_{k_l}^{(j)} = \Delta v \text{ if } j < l,\ 0 \text{ otherwise}) \qquad \text{for } l = 2..D+1, \\ v_{k_{D+2}} = (2\Delta v, 0, \ldots, 0). \end{gathered}$$

We intent to prove $\{\boldsymbol{m}_{k_l}\}_{l=1..D+2}$ is a linearly independent set of $\mathbb{R}^{D+2}$, which will prove the proposition. Let us consider $\boldsymbol{\beta} \in \mathbb{R}^{D+2}$ such that $\boldsymbol{\beta} \cdot \boldsymbol{m}_{k_l} = 0$ for

$l = 1..D+2$, this reads $A\boldsymbol{\beta} = 0$ in a matrix form where

$$A = \left[\boldsymbol{m}_{k_1}, \ldots, \boldsymbol{m}_{k_{D+2}}\right] = \begin{pmatrix} 1 & 1 & . & . & . & . & 1 & 1 \\ 0 & \Delta v & . & . & . & . & \Delta v & 2\Delta v \\ 0 & 0 & . & & & & \Delta v & 0 \\ . & . & . & . & & & . & . \\ . & . & & . & . & & . & . \\ . & . & & & . & . & . & . \\ 0 & 0 & . & . & . & 0 & \Delta v & 0 \\ 0 & \frac{1}{2}\Delta v^2 & . & . & . & . & \frac{D}{2}|\Delta v|^2 & 2\Delta v^2 \end{pmatrix}.$$

It can easily be verified that the determinant of A is $\det A = \Delta v^{D+2} \neq 0$. This proves that $\boldsymbol{\beta} = 0$, and $\{\boldsymbol{m}_{k_l}\}_{l=1..D+2}$ is a linearly independent set, hence a basis of $\mathbb{R}^{D+2}$.

In the case $a \neq 0$, we set $\tilde{v_k} = v_k - a$ for all $k \in \mathcal{K}$. From the previous case it is clear that $\{\tilde{\boldsymbol{m}}_k = \boldsymbol{m}(\tilde{v_k})\}_{k=1..D+2}$ is a linearly independent set, which means that $\tilde{A} = \left[\tilde{\boldsymbol{m}}_{k_1}, \ldots, \tilde{\boldsymbol{m}}_{k_{D+2}}\right]$ is invertible. Setting $A = \left[\boldsymbol{m}_{k_1}, \ldots, \boldsymbol{m}_{k_{D+2}}\right]$, we have the following lemma:

Lemma 3.1 There exists an upper triangular matrix L such that $A^T = \tilde{A}^T L$ and $L_{ll} \neq 0$ for $l = 1..D+2$.

The simple proof of this lemma being left to the reader, we can immediately conclude that $\det A = \det \tilde{A} \times \det L \neq 0$, and the proof of the proposition is now complete. □

4. Other properties

4.1. *Temperature and velocity bounds in a general DVM*

In this section, given a discrete velocity set, we investigate the properties of macroscopic quantities, independently of the collision process (*i.e.* BGK or Boltzmann). The following result shows that once $\mathcal{V}$ is chosen, a discrete velocity model cannot describe any flow. Conversely, for a given flow, the discrete velocity set must be properly chosen to give a correct representation.

Proposition 4.2 Let $f \in \mathbb{R}^N$ be a nonnegative discrete distribution function. Then the macroscopic velocity and temperature associated with f on $\mathcal{V}$ by $u = \frac{1}{\rho}\langle vf\rangle_{\mathcal{K}}$ and $T = \frac{1}{DR\rho}\langle |v-u|^2 f\rangle_{\mathcal{K}}$ satisfy the bounds

$$\min_{\mathcal{K}} v_k^{(i)} \leq u^{(i)} \leq \max_{\mathcal{K}} v_k^{(i)} \qquad i = 1..D \tag{4.38}$$

$$\frac{1}{DR}\min_{\mathcal{K}} |v_k - u|^2 \leq T \leq \frac{1}{DR}\max_{\mathcal{K}} |v_k - u|^2. \tag{4.39}$$

The proof is an immediate consequence of definitions of u and T and is left to the reader.

These inequalities show that the macroscopic velocity and temperature are bounded above by velocity bounds. This implies that the discrete velocity set must be large enough to take into account high velocities and high temperatures which may appear in the flow.

Another, perhaps unexpected, feature of a discrete velocity flow is that temperature is bounded below. This bound is in fact strongly related to Δv, as it can be seen in the following example. Consider a 1D plane Couette flow (that is a plane flow between two plates of equal temperature T_W, one of which is motionless and the other moving at u_W). In such a flow, the macroscopic velocity increases continuously between 0 and u_W. This implies that there exists position x for which $\min_{\mathcal{K}} |u(x) - v_k| = \frac{1}{2}\Delta v$. Thus at this point $T(x)$ is bounded below by $\frac{\Delta v^2}{4DR}$, regardless of the wall temperature. As a consequence, a small T_W could lead to an overestimated temperature jump near the plates. This example suggests that Δv must be small enough to take into account low temperatures which may appear in the flow.

Remark 4.4 From Statistical Mechanics theory, it is preferable to define the temperature of a discrete velocity gas through the coefficient c such that $\exp(\boldsymbol{\alpha}\cdot\boldsymbol{m}_k) = a\exp(-\frac{1}{2c}|b-v_k|^2)$ (see [9]). The result of the previous proposition does not give any bounds on this temperature. However Rogier and Schneider [29] have proved that for a continuous Maxwellian in a bounded velocity domain, this coefficient may be either positive, if the domain is large enough, or negative if the domain is too small. Thus, it seems reasonable that this is also true for our discrete case. Note, however, that in the sequel, we shall only use the definition of the temperature given in proposition 4.2.

4.2. *A remark on plane flows*

This remark is motivated by the fact that in the continuous-velocity case (for instance with $D = 3$), if $u^{(3)} = 0$ (*i.e.* a plane flow), then (2.10) implies that $\alpha^{(3)} = 0$ as well. It is interesting to see if this would hold in a discrete velocity frame, since it would reduce the cost of computing discrete equilibrium. This would lead to a problem in $\mathbb{R}^{D+1}$ rather than in $\mathbb{R}^{D+2}$. Because explicit invertible relation between $\boldsymbol{\rho}$ and $\boldsymbol{\alpha}$ is not known, the situation is not as obvious as in the continuous case. However, we shall prove it to be true in the discrete case.

Proposition 4.3 Let $\mathcal{V}$ be a discrete velocity grid satisfying (3.21) and symmetric with respect to the j^{th} hyperplan $\{v^{(j)} = 0\}$. Let $\boldsymbol{\rho}$ be a strictly realizable vector of $\mathbb{R}^{D+2}$ on $\mathcal{V}$ (in the sense of (3.23)), such that $\rho^{(j)} = 0$ (*i.e.* $u^{(j)} = 0$). Then the Lagrange multiplier $\boldsymbol{\alpha}$ from theorem 3.1 satisfies equivalently $\alpha^{(j)} = 0$.

Proof. Note that the symmetry assumption on $\mathcal{V}$ means that for all k in $\mathcal{K}$, there exists k' in $\mathcal{K}$ such that $v_k^{(i)} = v_{k'}^{(i)}$ if $i \neq j$ and $v_k^{(j)} = -v_{k'}^{(j)}$. Consequently, defining

r_k such that $\boldsymbol{\alpha}\cdot\boldsymbol{m}_k = r_k + \alpha^{(j)}v_k^{(j)}$, yields

$$\rho^{(j)} = 2\sum_{\substack{k\ s.t.\\ v_k^{(j)}>0}} \exp(r_k)\sinh(\alpha^{(j)}v_k^{(j)})v_k^{(j)}\Delta v^D.$$

All the terms of this sum have the same sign, and are zero if and only if $\alpha^{(j)}=0$. □

5. Numerical schemes for the discrete velocity BGK model

5.1. *Explicit conservative entropic scheme*

For the sake of simplicity, our scheme is presented here in two spatial dimensions on a Cartesian grid, but all the properties stated here are valid for a general D-dimensional space and curvilinear meshes. The equation to be approximated is

$$\partial_t f_k + v_k^{(1)}\partial_x f_k + v_k^{(2)}\partial_y f_k = \frac{1}{\tau}(\mathcal{E}_k - f_k), \qquad k\in\mathcal{K}. \tag{5.40}$$

Consider a spatial Cartesian grid defined by nodes $(x_i,y_j)=(i\Delta x, j\Delta y)$ and cells $I=]x_{i-\frac12},x_{i+\frac12}[\times]y_{j-\frac12},y_{j+\frac12}[$. Consider also a time discretization with $t_n=n\Delta t$. If $f_{i,j}^n=(f_{k,i,j}^n)_{k\in\mathcal{K}}$ is an approximation of $f_{\mathcal{K}}(t_n,x_i,y_j)$, the moments of $f_{i,j}^n$ are naturally $\boldsymbol{\rho}_{i,j}^n=\langle\boldsymbol{m}f_{i,j}^n\rangle_{\mathcal{K}}$, and the corresponding discrete equilibrium is denoted by $(\mathcal{E}_{k,i,j}^n)_{k\in\mathcal{K}}$. If $\boldsymbol{\rho}_{i,j}^n$ is strictly realizable (in the sense of (3.23)), the discrete equilibrium is therefore $\mathcal{E}_{k,i,j}^n=\exp(\boldsymbol{\alpha}_{i,j}^n\cdot\boldsymbol{m}_k)$, where $\boldsymbol{\alpha}_{i,j}^n$ is the unique solution of the nonlinear set of equations (see remark 3.1)

$$\langle\boldsymbol{m}\exp(\boldsymbol{\alpha}_{i,j}^n\cdot\boldsymbol{m})\rangle_{\mathcal{K}}=\boldsymbol{\rho}_{i,j}^n. \tag{5.41}$$

The transport part is simply the linear convection equation, and can be approximated by a standard finite volume scheme. For the nonlinear relaxation term, a standard centered approximation technique is used. Our scheme thus reads

$$\begin{aligned} f_{k,i,j}^{n+1} = f_{k,i,j}^n &- \frac{\Delta t}{\Delta x}(\mathcal{F}_{k,i+\frac12,j}^n - \mathcal{F}_{k,i-\frac12,j}^n) - \frac{\Delta t}{\Delta y}(\mathcal{F}_{k,i,j+\frac12}^n - \mathcal{F}_{k,i,j-\frac12}^n)\\ &+ \frac{\Delta t}{\tau_{i,j}^n}(\mathcal{E}_{k,i,j}^n - f_{k,i,j}^n), \end{aligned} \tag{5.42}$$

where the numerical fluxes are defined by

$$\mathcal{F}_{k,i+\frac12,j}^n = \frac12\left(v_k^{(1)}f_{k,i+1,j}^n + v_k^{(1)}f_{k,i,j}^n - |v_k^{(1)}|(\Delta f_{k,i+\frac12,j}^n - \Phi_{k,i+\frac12,j}^n)\right) \tag{5.43}$$

with the notation $\Delta f_{k,i+\frac12,j}^n = f_{k,i+1,j}^n - f_{k,i,j}^n$, and the flux limiter function gives the order of the scheme. For instance

$$\Phi_{k,i+\frac12,j}^n = \begin{cases} \Delta f_{k,i+\frac12,j}^n & \text{for first order,}\\ \operatorname{minmod}(\Delta f_{k,i-\frac12,j}^n, \Delta f_{k,i+\frac12,j}^n, \Delta f_{k,i+\frac32,j}^n) & \text{for second order.}\end{cases} \tag{5.44}$$

With the appropriate definitions of our DVM, our scheme now possesses the expected properties mentioned in the first section. Theorem 3.2 can be expressed in its numerical form:

Proposition 5.4 Let $\{f^0_{k,i,j}\}_{k,i,j}$ be a strictly positive initial condition. Then the sequence $\{f^n\}_{n\geq 0}$ defined by scheme (5.42) satisfies

$$f^n_{k,i,j} > 0 \quad \text{and} \quad \mathcal{E}^n_{k,i,j} = \exp(\boldsymbol{\alpha}^n_{i,j} \cdot \boldsymbol{m}_k), \tag{5.45}$$

for all k, n, i, j. Also, for all n

$$\sum_{k,i,j} \boldsymbol{m}_k f^n_{k,i,j} \Delta v^D = \sum_{k,i,j} \boldsymbol{m}_k f^0_{k,i,j} \Delta v^D \tag{5.46}$$

$$\sum_{k,i,j} f^{n+1}_{k,i,j} \log f^{n+1}_{k,i,j} \Delta v^D \leq \sum_{k,i,j} f^n_{k,i,j} \log f^n_{k,i,j} \Delta v^D, \tag{5.47}$$

with the condition that for the first order scheme, the time step follows

$$\Delta t \left(\max_{i,j} (\frac{1}{\tau^n_{i,j}}) + \max_{\mathcal{K}} \left(\frac{|v^{(1)}_k|}{\Delta x} + \frac{|v^{(2)}_k|}{\Delta y} \right) \right) < 1. \tag{5.48}$$

Additionally, the second order scheme is subject to a similar condition.

Proof. We simply prove the entropy property (5.47) for the first order scheme. Relation (5.42) reads

$$\begin{aligned} f^{n+1}_{k,i,j} = & \left(1 - \frac{\Delta t}{\tau^n_{i,j}} - |v^{(1)}_k| \frac{\Delta t}{\Delta x} - |v^{(2)}_k| \frac{\Delta t}{\Delta y} \right) f^n_{k,i,j} - \frac{\Delta t}{\Delta x} v^{(1)-}_k f^n_{k,i+1,j} \\ & + \frac{\Delta t}{\Delta x} v^{(1)+}_k f^n_{k,i-1,j} - \frac{\Delta t}{\Delta y} v^{(2)-}_k f^n_{k,i,j+1} + \frac{\Delta t}{\Delta y} v^{(2)+}_k f^n_{k,i,j-1} \\ & + \frac{\Delta t}{\tau^n_{i,j}} \mathcal{E}^n_{k,i,j}, \end{aligned}$$

with the classical notations $a^+ = \frac{a+|a|}{2}$ and $a^- = \frac{a-|a|}{2}$. Using condition (5.48), this expression is a convex combination of $f^n_{k,i,j}$, $f^n_{k,i+1,j}$, $f^n_{k,i,j+1}$, $f^n_{k,i-1,j}$, $f^n_{k,i,j-1}$ and $\mathcal{E}^n_{k,i,j}$. Since the function $t \log t$ is also convex, we have

$$\begin{aligned} f^{n+1}_{k,i,j} \log f^{n+1}_{k,i,j} \leq & \left(1 - \frac{\Delta t}{\tau^n_{i,j}} - |v^{(1)}_k| \frac{\Delta t}{\Delta x} - |v^{(2)}_k| \frac{\Delta t}{\Delta y} \right) f^n_{k,i,j} \log f^n_{k,i,j} \\ & - \frac{\Delta t}{\Delta x} v^{(1)-}_k f^n_{k,i+1,j} \log f^n_{k,i+1,j} + \frac{\Delta t}{\Delta x} v^{(1)+}_k f^n_{k,i-1,j} \log f^n_{k,i-1,j} \\ & - \frac{\Delta t}{\Delta y} v^{(2)-}_k f^n_{k,i,j+1} \log f^n_{k,i,j+1} + \frac{\Delta t}{\Delta y} v^{(2)+}_k f^n_{k,i,j-1} \log f^n_{k,i,j-1} \\ & + \frac{\Delta t}{\tau^n_{i,j}} \mathcal{E}^n_{k,i,j} \log \mathcal{E}^n_{k,i,j}. \end{aligned}$$

By summation over i, j, entropy numerical fluxes vanish, and summing over k yields

$$\sum_{k,i,j} f^{n+1}_{k,i,j} \log f^{n+1}_{k,i,j} \leq \sum_{k,i,j} (1 - \frac{\Delta t}{\tau^n_{i,j}})\, f^n_{k,i,j} \log f^n_{k,i,j} + \sum_{k,i,j} \frac{\Delta t}{\tau^n_{i,j}}\, \mathcal{E}^n_{k,i,j} \log \mathcal{E}^n_{k,i,j}.$$

Since the discrete equilibrium minimizes entropy, relation (5.47) becomes obvious. □

5.2. *Linearized implicit scheme*

To efficiently compute steady flows without the restriction (5.48), we present a linearized implicit scheme.

Implicit schemes are usually derived from explicit schemes by evaluating at time t_{n+1} the terms leading to negative quantities for large Δt (*e.g.* numerical fluxes and source terms). If the nonlinear terms are differentiable, they are then linearized, otherwise they are kept at t_n (like the second order fluxes).

The scheme presented here is derived accordingly. For the source term, it is clear that $-f^n_{k,i,j}$ (the so-called loss term) must be written at t_{n+1}, since it produces undesirable negative distributions if Δt is large. The gain term, namely the discrete equilibrium $\mathcal{E}^n_{k,i,j}$, is positive, and therefore may be kept explicit (a strategy used in [34]). However, gain and loss terms are then evaluated at different times. This is observed to slow the convergence of the scheme considerably (see Sec.). Consequently, we decide to evaluate the gain term at t_{n+1} as well. However, defining an implicit relaxation time $\tau^{n+1}_{i,j}$ would yield a more difficult computation, and is not very useful. Since the discrete equilibrium is a nonlinear function of f, it may be linearized as follows

$$\mathcal{E}^{n+1}_{k,i,j} \approx \mathcal{E}^n_{k,i,j} + [D^n_{i,j}(f^{n+1}_{i,j} - f^n_{i,j})]_k, \tag{5.49}$$

where $D^n_{i,j}$ is the Jacobian of the mapping $g \in \mathbb{R}^N \mapsto \mathcal{E}[g]$ evaluated at $f^n_{i,j}$. We thus obtain the following linearized implicit scheme

$$\left(\frac{I}{\Delta t} + R^n + T\right)\delta f^n = RHS^n, \tag{5.50}$$

where $\delta f^n = f^{n+1} - f^n$, I is the unit matrix, T is a matrix such that $(Tf^n)_{k,i,j} = \frac{1}{\Delta x}(\mathcal{F}^n_{k,i+\frac{1}{2},j} - \mathcal{F}^n_{k,i-\frac{1}{2},j}) + \frac{1}{\Delta y}(\mathcal{F}^n_{k,i,j+\frac{1}{2}} - \mathcal{F}^n_{k,i,j-\frac{1}{2}})$ with only the first order fluxes, R^n is such that $(R^n f^n)_{i,j} = \frac{1}{\tau^n_{i,j}}(f^n_{i,j} - D^n_{i,j} f^n_{i,j})$, and

$$\begin{aligned} RHS^n_{i,j} = &- \frac{1}{\Delta x}(\mathcal{F}^n_{k,i+\frac{1}{2},j} - \mathcal{F}^n_{k,i-\frac{1}{2},j}) - \frac{1}{\Delta y}(\mathcal{F}^n_{k,i,j+\frac{1}{2}} - \mathcal{F}^n_{k,i,j-\frac{1}{2}}) \\ &+ \frac{1}{\tau^n_{i,j}}(\mathcal{E}^n_{i,j} - f^n_{i,j}) \end{aligned} \tag{5.51}$$

with first or second order fluxes. The Jacobian $D^n_{i,j}$ has the simple form

$$D^n_{i,j}[k,k'] = A^{-1}(\boldsymbol{\alpha}^n_{i,j}) : \boldsymbol{m}_k \otimes \boldsymbol{m}_{k'} \mathcal{E}^n_{k,i,j} \Delta v^D, \tag{5.52}$$

where $A(\boldsymbol{\alpha}_{i,j}^n) = \langle \boldsymbol{m} \otimes \boldsymbol{m} \exp(\boldsymbol{\alpha}_{i,j}^n \cdot \boldsymbol{m}) \rangle_{\mathcal{K}}$.

Due to the linearization of $\mathcal{E}_{\mathcal{K}}$ the scheme does not possess the conservation and entropy properties proved for the explicit scheme. However we are interested in the steady state solution only. More precisely, after convergence we have $f^{n+1} = f^n$ and the scheme gives the stationary conservation laws and entropy inequality.

The particular structure of matrices T and R^n may be noted. If quantities $f_{k,i,j}^n$ are stored as $f^n = (f_k^n)_{k\in\mathcal{K}}$ with $f_k^n = (f_{k,i,j}^n)$, then it can easily be seen that T is a $NM \times NM$ block diagonal matrix with $M \times M$ pentadiagonal blocks T_k (M is the number of cells), and that R^n is a full matrix of diagonal blocks $R_{k,k'}^n$. These structures are naturally due to the fact that relaxation process in BGK equation is local in space but global in velocity, whereas transport process is numerically global in space but local in velocity. One could observe that all the nonzero elements of a block T_k are equal, which would simplify the scheme. But in a practical use of this scheme, a curvilinear grid is used instead of a Cartesian one. This leads to geometric coefficients in the numerical fluxes which depend on (i, j). This is the reason why we do not employ all the simplifications awarded by the use of a Cartesian grid.

The linear system (5.50) to be solved at each iteration is quite large ($NM \times NM$), and an iterative method well adapted to different sparse structures of the matrices may be used. We present here an algorithm based on a coupling between Gauss-Seidel and Jacobi methods. First, R^n is separated into its block diagonal Δ^n and its block off-diagonal E^n, *i.e.* $R^n = \Delta^n - E^n$ (this is the Jacobi step). System (5.50) may then be solved by the following algorithm

1. $g^{(0)} = 0$
2. for $p = 0..P$ solve $\left(\frac{I}{\Delta t} + T + \Delta^n\right) g^{(p+1)} = RHS^n + E^n g^{(p)}$
3. set $\delta f^n = g^{(P+1)}$.

Since the matrix of these linear systems is block diagonal with pentadiagonal blocks $\frac{I}{\Delta t} + T_k + \Delta_k^n$, it is possible to use a line Gauss-Seidel method by setting $T_k = M_k - N_k$. This gives the following algorithm.

1. $g^{(0)} = 0$
2. $\forall k \in \mathcal{K}$ and for $p = 0..P$ solve

$$\left(\frac{I}{\Delta t} + M_k + \Delta_k^n\right) g_k^{(p+1)} = RHS_k^n + N_k g_k^{(p)} + [E^n g^{(p)}]_k \tag{5.53}$$

3. set $\delta f^n = g^{(P+1)}$.

The linear systems (5.53) may easily and exactly be solved by successive inversions of tridiagonal matrices. Note that calculating the product $E^n g$ is not very expensive

because the blocks of E^n are diagonal. In fact, we have

$$[E^n g]_{k,i,j} = \frac{1}{\tau^n_{i,j}} A^{-1}(\boldsymbol{\alpha}^n_{i,j}) \boldsymbol{m}_k \mathcal{E}^n_{k,i,j} \cdot \Big(\langle \boldsymbol{m} g_{i,j} \rangle_{\mathcal{K}} - \boldsymbol{m}_k g_{k,i,j} \Delta v^D \Big). \tag{5.54}$$

It is thus sufficient to compute $A^{-1}(\boldsymbol{\alpha}^n_{i,j}) \boldsymbol{m}_k \mathcal{E}^n_{k,i,j}$ at the beginning of the algorithm (a local computation in k and i,j), then to compute $\langle \boldsymbol{m} g_{i,j} \rangle_{\mathcal{K}}$ on each cell (which is local in i,j), and finally to form the dot product. The computation of $E^n g$ is thus local in i,j, and hence completely parallelizable; its cost is in $O(NM)$.

It is well known in CFD that since only a few iterations are needed to have the external process converge (the loop in n), it is not useful to carry on an algorithm like the previous one at convergence. The cost of our implicit scheme is then in $O(PNM)$ where $P = 2$ or 3.

5.3. *Computation of* $\boldsymbol{\alpha}^n_{i,j}$

The nonlinear set of equations (5.41) may be solved by the following Newton algorithm where J is defined in the proof of Theorem (3.1):

1. Let be given $\boldsymbol{\alpha}^{(0)}$
2. solve $J''(\boldsymbol{\alpha}^{(r)})\boldsymbol{\alpha}^{(r+1)} = \boldsymbol{\alpha}^{(r)} - J'(\boldsymbol{\alpha}^{(r)})$ until a stop criterion is satisfied
3. set $\boldsymbol{\alpha}^n_{i,j} = \boldsymbol{\alpha}^{(r)}$

Such an algorithm requires almost $O((D^2 + D)N)$ operations by iteration r and by cell (i,j) and is thus in $O(MN)$. However three problems may occur. First, although the inversion of $J''(\boldsymbol{\alpha}^{(r)}) = \langle \boldsymbol{m} \otimes \boldsymbol{m} \exp(\boldsymbol{\alpha}^{(r)} \cdot \boldsymbol{m}) \rangle_{\mathcal{K}}$ in step 2 is always theoretically possible (see the proof of Theorem 3.1), it is numerically difficult. Since the first element of $\boldsymbol{m}_k \otimes \boldsymbol{m}_k$ is 1 whereas the last one is $|v_k|^4$, the conditioning of this matrix may be very large. In fact, the conditioning is in fact as large as the velocity bounds, and it is often greater than 10^{20}. This leads to a lack of robustness which can be fixed by globalization technique such as the back-tracking linesearch (see [15]).

The second problem is the choice of initial data $\boldsymbol{\alpha}^{(0)}$. A natural choice is to take the parameter $\boldsymbol{\alpha}$ of the continuous equilibrium given in (2.10). But this parameter may be far from the solution, especially when N is small. Another choice is to take $\boldsymbol{\alpha}^{(0)} = \boldsymbol{\alpha}^{n-1}_{i,j}$ computed at the previous global iteration. The advantage of this method is that for small time steps (like in explicit scheme) and for stabilized flows, the initial data is close to the solution. For practical use of the implicit scheme, a good approach is to use the first method at the beginning of the computation, when the flow may have strong variations, and then to use the second one when the flow is almost stabilized.

This raises the third problem which could be a prohibitive large number of iterations of the Newton algorithm. In fact, with our choice $\boldsymbol{\alpha}^{(0)}$, the algorithm converges rapidly - only one iteration is needed for most cases tested.

Finally, using proposition 4.3 permits the reduction of the size of the system (5.41) to $D+1$ in the case of plane flows (with a well defined velocity set). This decreases the cost of the algorithm by a factor of almost 1.3.

6. Numerical results

Some numerical results are presented here to illustrate the ability of the method to describe steady flows. Our method is first validated on simple 1D flows which represent some typical kinetic non-equilibrium phenomena (shock and boundary layer). We also investigate the influence of the velocity discretization. Then, more complex 2D flow computations are presented, and comparisons with DSMC results are given.

Except in some cases, the linearized implicit scheme of second order is used in all the computations, with a CFL number of 10000 (*i.e.* Δt is 10000 times the explicit time step). The criterion used to determine whether the flow has reached steady state is the reduction of the quadratic global residual $\frac{1}{\Delta t}(\sum_{k,i,j} |RHS^n_{k,i,j}|^2)^{\frac{1}{2}}$ by a factor of 10^5. In all the results of this section, gas-surface interactions are Maxwellian reflections with total accommodation, *i.e.* incident molecules are supposed to be re-emitted by the wall with a Maxwellian distribution $\tilde{\rho} M[1, u_w, T_w]$ of mean velocity u_w and temperature T_w. The coefficient $\tilde{\rho}$ is determined to ensure a zero mass flux normal to the wall. In our discrete velocity model, the Maxwellian is naturally replaced by the discrete equilibrium function associated to u_w and T_w. Numerically, all the boundary conditions (gas-surface, symmetry axes, etc.) are treated by a classical ghost cell technique. Finally, the relation (2.8) where ω is given for each gas in [4], is used to compute the relaxation time of the model.

The first test case studied is the 1D stationary shock wave. The flow is initialized with two Maxwellian states related by classical Rankine-Hugoniot relations. In addition, the steady state shows the transition between upstream and downstream flows. In this instance, we have used $\rho_L = 6.63\,10^{-6} kg.m^{-3}$, $T_L = 293\,K$, $u_L = 2551\,m.s^{-1}$ for the upstream flow. These values yield a shock Mach number of 8. The gas considered is argon, and consequently $\omega = 0.81$. The width of the computational domain is $0.5\,m$, and a 1D grid of 200 cells is used. Fig.(1) shows the normalized profiles $\frac{q_L - q}{q_L - q_R}$ for $q = \rho, T, u$, obtained with $11 \times 9 \times 9$ discrete velocities in the (v_x, v_y, v_z) directions. The bounds are given by $[-3846, 5181] \times [-4513, 4513] \times [-4513, 4513]$ which ensures that left and right Maxwellian are correctly represented. This allows us to account almost for most flow information. These profiles are found to be somewhat less smooth than those by Bird [4]. The distribution functions in upstream flow, within the shock, and in downstream flow, are plotted in fig.(2). Despite the small number of discrete velocities, the results clearly show the kinetic nonequilibrium in the shock (the distribution is quite far from Maxwellian repartition). As expected, upstream and downstream distributions are approximately Maxwellian (upstream flow is less well represented because of its smaller temperature).

Fig.(3) shows a comparison of the shock-wave thicknesses given by our method

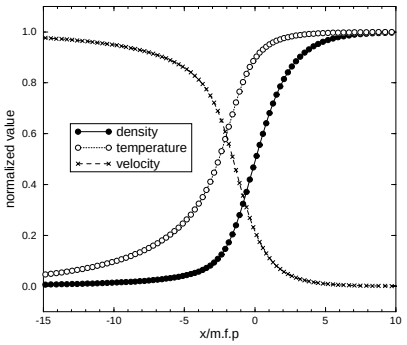

Fig. 1. Normalized profiles for a shock-wave at Mach 8.

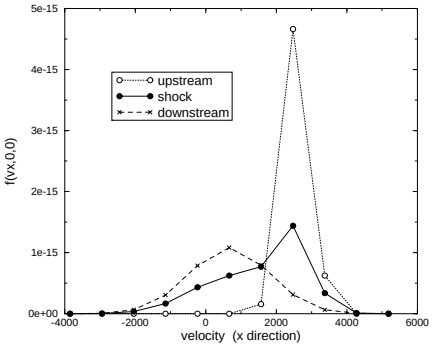

Fig. 2. Distribution functions in the shock-wave.

for different upstream Mach numbers with the experimental results taken from Alsmeyer's account [1] and with Navier-Stokes results. Our results are found to be in very good agreement with the experimental data.

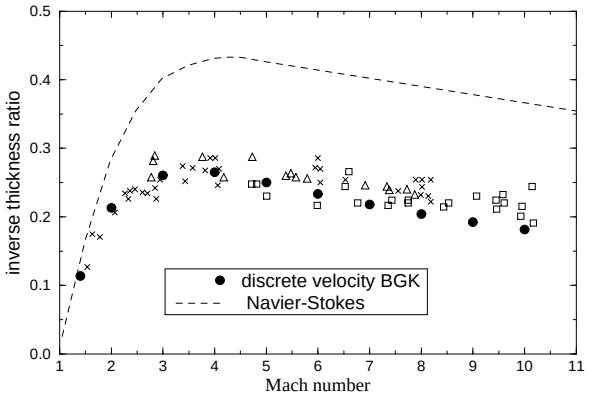

Fig. 3. Comparison of BGK, Navier-Stokes, and experiments for shock-wave thicknesses in argon.

In addition we have studied the influence of velocity bounds on both the macroscopic and microscopic quantities. Figure (4) represents the upstream, shock, and downstream distributions obtained for different bounds. The number of discrete velocities being adapted to each case in order to reach the same precision. The corresponding profiles for ρ, u, T are plotted in figure (5). As expected, employing very large bounds does not improve the results since all the essential information is already accounted with normal bounds (fig.(4)). However, when the bounds are too small, the profiles begin to loose their accuracy (they are not smooth enough).

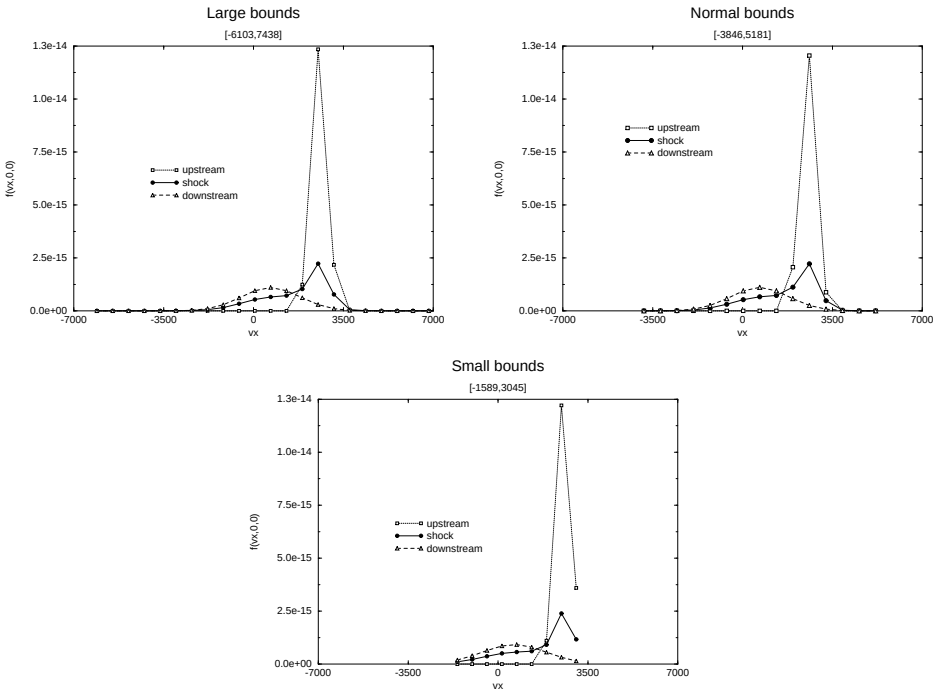

Fig. 4. Influence of velocity bounds on the discrete distribution functions.

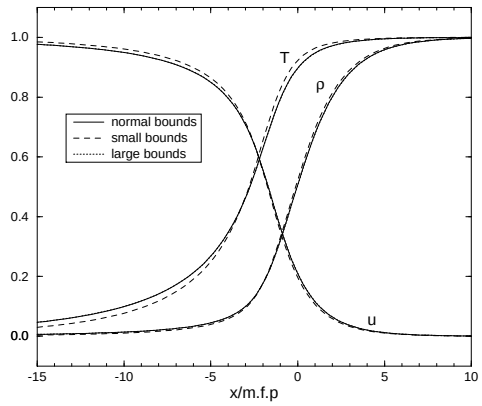

Fig. 5. Influence of velocity bounds on the macroscopic quantities: density ρ, temperature T, velocity u.

A lot of information is lost, especially in the population of high velocity molecules in the upstream state, since the shock and upstream distributions are truncated. Nevertheless, this truncation gives a good indication to check a posteriori if the discrete velocity set is large enough for the flow.

Finally, the improvement due to the explicit and implicit second order schemes (in space) is shown in figure (6) with a stationary shock wave of Mach number of 4. The upstream flow has the same density and temperature as in the previous case. The grid is still of 200 cells but we use now $9 \times 9 \times 9$ velocities and the bounds are $[-2013, 2771] \times [-2392, 2392] \times [-2392, 2392]$. This test is motivated by the fact that high order schemes that do not resolve the small relaxation time - as our linearized implicit scheme - may reduce to lower order. This is investigated by Jin in [20] for hyperbolic systems with stiff source terms. However, this drawback does not seem to affect our results. The first order scheme leads to a numerical dissipation which causes density, temperature, and velocity profiles too smooth. The dissipation is clearly reduced by the second order discretization in space, which yields stiffer profiles. Note that the results obtained with both explicit and implicit second order schemes are indistinguishable. This proves that the implicit scheme is really second order accurate. This improvement has also been observed with 2D computations.

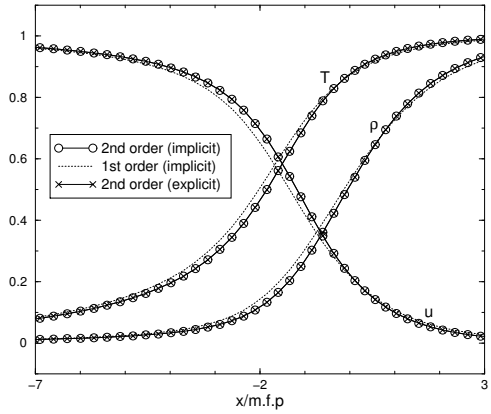

Fig. 6. Comparison between first and second order schemes (for a shock-wave at Mach 4).

The second test case is a 1D plane Couette flow. The data are those of Bird in [4]. The gas, argon, lies between two plates maintained at a temperature of $T_w = 273\,K$. One of which is stationary and the other is moving with a velocity of $u_w = 300\,m.s^{-1}$ in y direction. The gas is initially set to the same temperature as the plates, and its density is $9.28\,10^{-6}\,kg.m^{-3}$. The Knudsen number based on the distance between the plates ($1\,m$) is thus $9.25\,10^{-3}$. Using a grid of 200 cells in x direction,

$13 \times 17 \times 13$ discrete velocities with bounds $[-913, 913] \times [-1253, 1253] \times [-913, 913]$, we obtain the profiles plotted in figure (7). We find a good qualitative agreement with the DSMC results of Bird with, however, greater amplitudes for the density and temperature. This is probably due to the fact that the BGK model is known to lead to an unrealistic value of Prandtl number (1 instead of 2/3), *i.e.* heat conductivity is under-estimated. Note that the Knudsen layer can be observed in the pressure profile: this curve is almost constant except near the plates where there is a sharp decrease over a length-scale of a few mean free paths. This length-scale is known as the kinetic boundary layer.

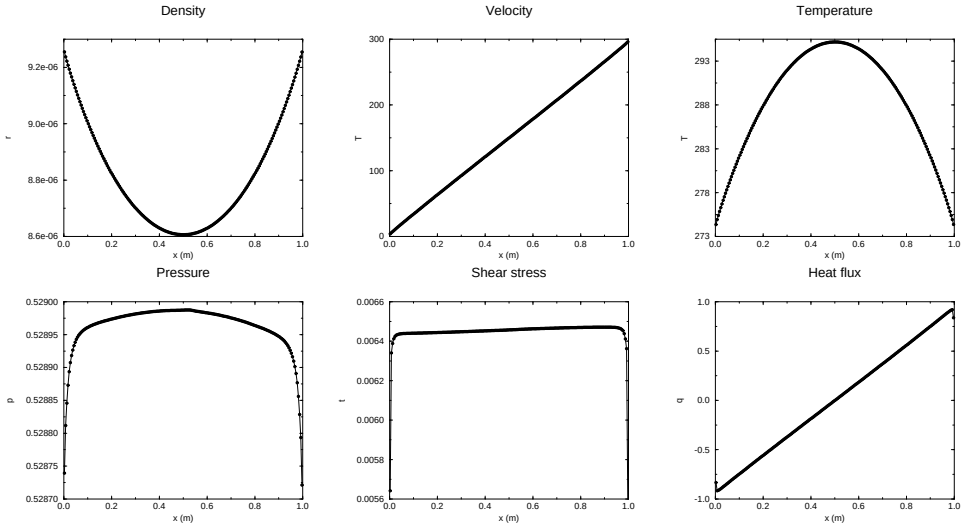

Fig. 7. Profiles in a 1D Couette flow.

The third test case is that of a supersonic flow past a cylinder. The gas considered here is argon. The upstream flow is such that $\rho_\infty = 0.317\,10^{-5}\,kg.m^{-3}$, $T_\infty = 249\,K$, $M_\infty = 4$, which gives a Knudsen number based on the radius of the cylinder ($1\,m$) of $3.56\,10^{-2}$. The temperature of the cylinder is $273\,K$. For symmetry reasons, only half of the flow is computed, and we have neglected the influence of the flow downstream from the cylinder by enforcing a supersonic out-going condition. We used a grid of 20×40 cells and $13 \times 13 \times 13$ discrete velocities. The velocity bounds are $[-2561, 2561] \times [-2461, 2461] \times [-2303, 2303]$. They are estimated by evaluating $\max(u + 4\sqrt{RT})$ with a Navier-Stokes code (they are essentially given by the temperature within the shock). The results for the density, velocity and temperature are shown in figure (8).

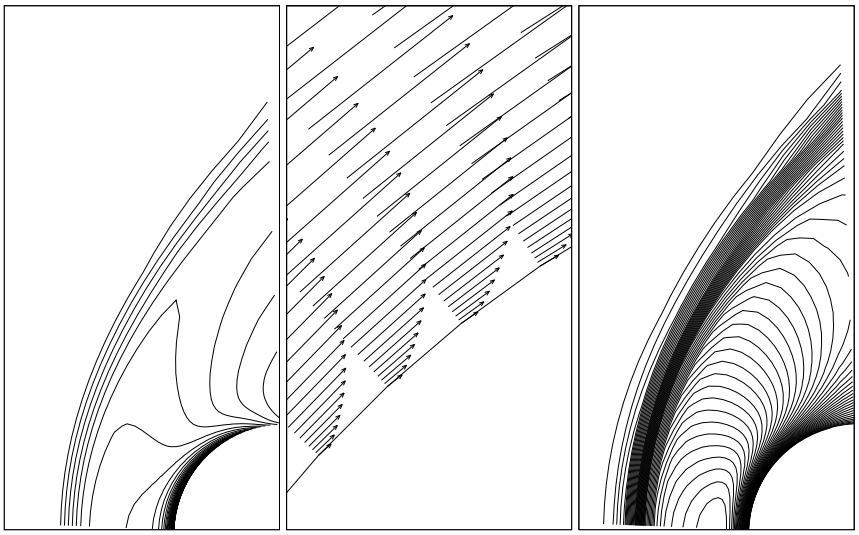

Fig. 8. Supersonic flow past a cylinder: density, velocity (zoom), temperature ($M_\infty = 4,\ Kn_\infty = 0.03$).

The last test is a comparison of our method with DSMC results (given by the code of J.-C. Lengrand [21]) for a supersonic flow past a flat plate of $5\,cm$ followed by a compression ramp of 10°. The parameters of the flow are $\rho_\infty = 1.288\,10^{-4}\,kg.m^{-3}, T_\infty = 72.2\,K, M_\infty = 3.67$, for the density, temperature and Mach number. The molecular mass is $4.815\,10^{-26}\,kg$ and the viscosity exponent is 0.77. This gives a Knudsen number of $6.7\,10^{-3}$ at infinity. The local Knudsen number (see [4]) at the leading edge is 0.13 which is beyond the validity range of Navier-Stokes equations. In fact, Bird notices in [4] that the error in Navier-Stokes results is significant in the regions of the flow where the local Knudsen number exceeds 0.1.

For both the DSMC and our method, we used a grid of 70×70 cells, and $13 \times 11 \times 11$ velocities are used for the DVM. The computation takes 260 iterations and 42 hours for the DVM on the single-processor IBM-SP2. For the DSMC, we used 2600 samples and an average of 20 particles per cell. It yields 46 hours of CPU time on the same computer. The contours of density and temperature are plotted in fig.(9) and (10) for both methods. The results are in good agreement, and this can be seen more clearly in fig.(11) where the distributions of density and temperature following three vertical lines are shown. The two methods are in very good agreement except within the shock where the results are slightly different. For the DSMC, note the noise induced by the stochasticity of the method. Also, it is apparent that the results of the DSMC are inaccurate in the small region in front of the downstream boundary. This is a direct consequence of a defect in the boundary conditions (see [7]). Although the CPU times of the two methods are provided for

this case, a fair comparison of computational speeds of the two methods is not easy because their criteria of convergence are very different. For instance one can obtain more accurate results for the DSMC or less accurate results for DVM by increasing or decreasing the CPU time by a factor 2. Yet, the comparison shown in figs.(9-11) remains almost unchanged.

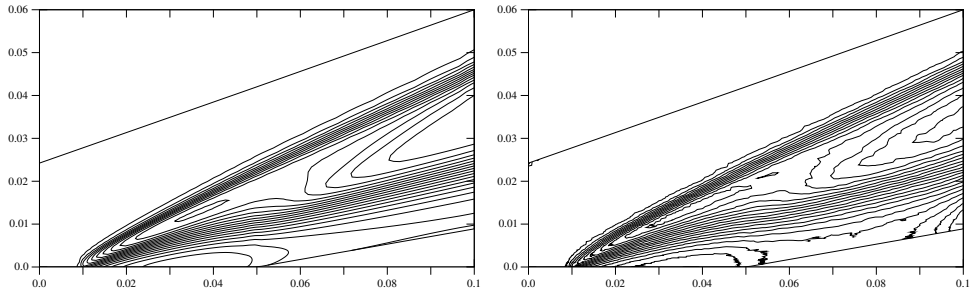

Fig. 9. Compression ramp: density (left BGK, right DSMC).

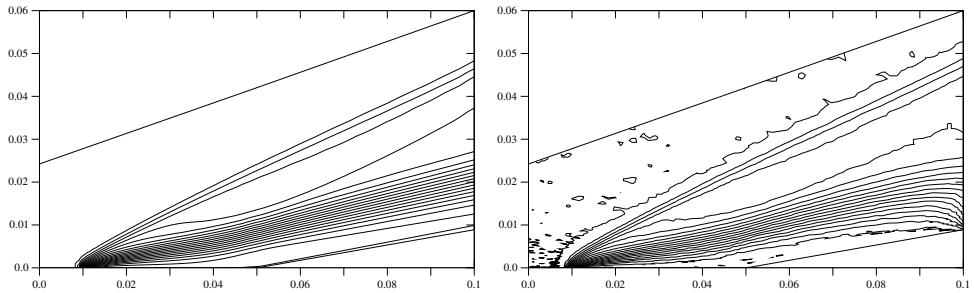

Fig. 10. Compression ramp: temperature (left BGK, right DSMC).

Finally we wish to emphasize the fact that our scheme is well suited for steady state computations. Table 1 shows the CPU time and the number of iterations obtained for both our explicit scheme and our linearized implicit scheme. The advantage of using the implicit scheme for stationary flows is undeniable. Although the explicit scheme does not require the solution of large linear systems (which divides the cost of one iteration by 2 compared to the implicit scheme), the number of iterations necessary for its convergence is so large that the implicit scheme is much more efficient. Figure (12) shows a comparison of the convergence history between implicit and semi-implicit scheme (where only the loss term is implicit, see Sec.). As mentioned earlier, the semi-implicit scheme is much slower than the implicit one. In fact the implicit scheme converges in almost 1000 iterations and 30 hours CPU time on the IBM-SP2, whereas approximately 4000 iterations and 158 hours are needed on the same computer in the semi-implicit case (note that with a

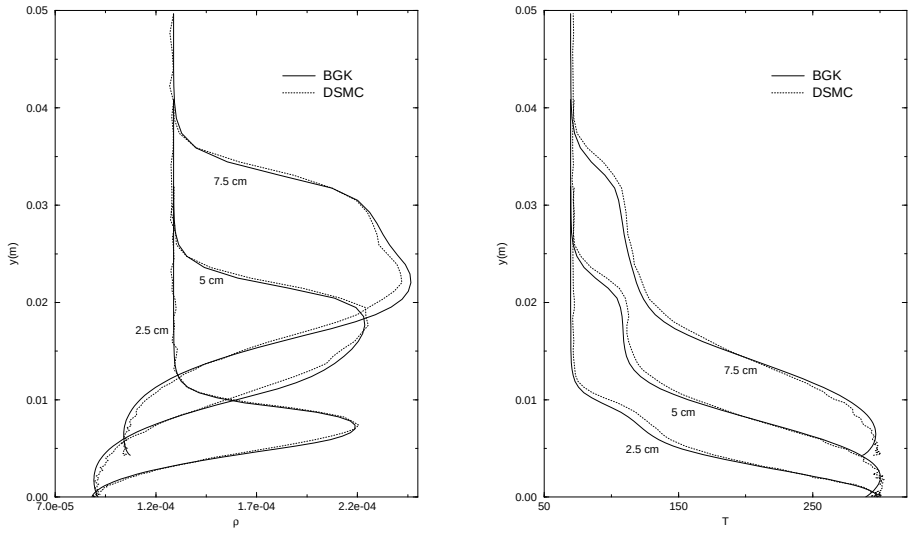

Fig. 11. Density and temperature distributions normal to the ramp along three vertical lines ($x = 2.5 - 5 - 7.5\ cm$).

Table 1. Comparison of explicit and implicit schemes for the shock-wave .

	iter.	CPU (s)
explicit	6000	16600
implicit	100	485

parallel computation, the implicit scheme would be even more enhanced since, as mentioned before, the linearization of $\mathcal{E}^n_{k,i,j}$ is completely parallelizable).

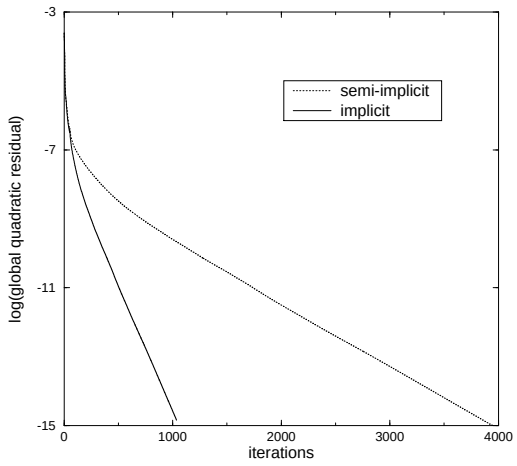

Fig. 12. Convergence history for implicit and semi-implicit scheme for the flow past the cylinder.

At last, we give a comparison of the mean CPU time per iteration for some of the previous test cases and some others. In table 2 are given for each case the mean CPU time per iteration, the number of velocities times the number of cells $N \times M$, and the ratio between these quantities. This ratio is almost constant. This confirms that, in accordance with Sec., the cost of one iteration for our algorithm is almost in $O(NM)$.

Table 2. CPU time per iteration for different test cases.

	CPU/iter.	$N \times M$	ratio
cylinder ($M = 20 \times 40$, $N = 13 \times 13 \times 13$)	100.11	1757600	17557
cylinder ($M = 20 \times 40$, $N = 11 \times 11 \times 11$)	59.81	1064800	17803
ramp ($M = 50 \times 40$, $N = 9 \times 9 \times 7$)	63.09	1134000	17974
Couette ($M = 200 \times 1$, $N = 13 \times 17 \times 13$)	32.96	574600	17433

7. Conclusion

We have presented a new discrete velocity model for the BGK equation such that conservation laws and entropy property are fully satisfied. This model is based on a rigorous discrete approximation of the Maxwellian equilibrium using a minimum entropy principle.

An explicit scheme for the approximation in space and time of the discrete velocity model has been proposed. It is also conservative and entropy decreasing. These properties allow an economic discretization in terms of the number of discrete velocities.

A linearized implicit scheme which satisfies these properties at convergence has been derived. It permits the model to reach steady state with large time steps, and is consequently a fast and robust method for stationary flows. Our results appear to compare very favorably with experiments and the DSMC method. For instance the results shown in figures (3,9,11) prove our method to be capable of computing transitional flows which ares in good agreement with that of the DSMC. In such regimes, Navier-Stokes is known to be insufficient to describe the flow, thus, with its low computational cost, the method presented here appears very attractive.

Finally, due to the linear complexity of our algorithm, the scheme may be extended to 3D flows and polyatomic gases without a prohibitive increase in the computational cost. It may also be easily extended to more realistic relaxation source terms in order to give better Prandtl number and thus reach better transport coefficients.

Acknowledgment

I thank B. Dubroca and P. Charrier for many helpful discussions. I am also very grateful to F. Veron for his help during the correction of the paper.

References

1. H. Alsmeyer, *Density Profiles in Argon and Nitrogen Shock Waves Measured by the Absorption of an Electron beam*, *J. Fluid Mech.* **74** (1976), no. 3, 497–513.
2. P. Andries, P. Le Tallec, J.-P. Perlat, and B. Perthame, *The Gaussian-BGK Model of Boltzmann Equation with Small Prandtl Number*, preprint.
3. K. Aoki, K. Kanba, and S. Takata, *Numerical Analysis of a Supersonic Rarefied Gas Flow Past a Flat Plate*, *Phys. Fluids* **9** (1997), no. 4.
4. G.A. Bird, *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford Science Publications, 1994.
5. A.V. Bobylev and J. Struckmeier, *Implicit and iterative methods for the Boltzmann equation*, *Transp. Th. Stat. Phys.* **25** (1996), no. 2, 175–195.
6. C. Buet, *Résolution déterministe de l'équation de boltzmann*, Tech. report, CEA, 1994.
7. C. Buet, *A Discrete-Velocity Scheme for the Boltzmann Operator of Rarefied Gas Dynamics*, *Transp. Th. Stat. Phys.* **25** (1996), no. 1, 33–60.
8. C. Cercignani, *The Boltzmann Equation and Its Applications*, vol. 68, Springer-Verlag, Lectures Series in Mathematics, 1988.
9. C. Cercignani, *Temperature, Entropy, and Kinetic Theory*, *J. Stat. Phys.* **87** (1997), no. 5-6.
10. S. Chapman and T.G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, Cambridge University Press, 1970.
11. P. Charrier, B. Dubroca, and J.-L. Feugeas, *Levermore's moment closure of discrete Boltzmann equations for non-equilibrium kinetic flows*, 21^{st} RGD Symposium, Book of Abstracts (Marseille), 1998.

12. P. Charrier, B. Dubroca, J.-L. Feugeas, and L. Mieussens, *Discrete-velocity models for kinetic nonequilibrium flows*, *C.R Acad.Sci.* **Serie I, 326** (1998), no. 11, 1347–1352, Paris.
13. P. Charrier, B. Dubroca, and L. Mieussens, *A numerical method for rarefied flow computation using a discrete velocity BGK model*, 21^{st} RGD Symposium, Book of Abstracts (Marseille), 1998.
14. F. Coron and B. Perthame, *Numerical Passage from Kinetic to Fluid Equations*, *SIAM J. Numer. Anal.* **28** (1991), no. 1, 26–42.
15. J.E. Dennis and R.B. Schnabel, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, Prentice-Hall Series in Computational Mathematics, 1983.
16. E. Gabetta, L. Pareschi, and G. Toscani, *Relaxation Schemes for Non Linear Kinetic Equations*, *SIAM J. Numer. Anal.* **34** (1997), no. 6, 2168–2194.
17. V. Garzó and A. Santos, *Comparison Between the Boltzmann and BGK Equations for Uniform Shear Flows*, *Physica A* **213** (1995), 426–434.
18. R. Gatignol, *Théorie Cinétique des Gaz à Répartition Discrète de Vitesses*, Lecture Notes in Physics, vol. 36, Springer-Verlag, 1975.
19. D. Issautier, *Convergence of a Weighted Particle Method for Solving The Boltzmann (B.G.K.) Equation*, *SIAM J. Numer. Anal.* **33** (1996), no. 6, 2099–2119.
20. S. Jin, *Runge-Kutta method for hyperbolic conservation laws with stiff relaxation terms*, *Journal of Comp. Phys.* **122** (1995), 51–67.
21. J.-C. Lengrand, *Mise en oeuvre de la méthode de Monte Carlo pour la simulation numérique d'un écoulement de gaz raréfié*, Tech. report, CNRS, 1986, RC 86-4.
22. C.D. Levermore, *Moment Closure Hierarchies for Kinetic Theories*, *J. Stat. Phys.* **83** (1996), 1021–1065.
23. L. Mieussens, *Convergence of a discrete-velocity model for the Boltzmann-BGK equation* , to appear in *C.R Acad. Sci. Paris*.
24. L. Mieussens, *An Implicit Discrete Velocity Scheme for the BGK Equation of Rarefied Gas Dynamics*, First European Symposium on Applied Kinetic Theory, Book of Abstracts (Toulouse), 1998.
25. L. Mieussens, *An Implicit Discrete Velocity Scheme for the BGK Equation of Rarefied Gas Dynamics*, 16^{th} ICNMFD, Book of Abstracts (Arcachon), 1998.
26. L. Mieussens, *Un schéma numérique pour un modèle discret en vitesse de l'équation BGK*, Rapport n^{o}98008, MAB, Université Bordeaux I, 1998.
27. B. Perthame, *Global existence to the BGK model of Boltzmann Equation*, *J. Diff. Eq.* **82** (1989), 191–205.
28. E.P. Gross P.L. Bathnagar and M. Krook, Phys. Rev. **94** (1954), 511.
29. F. Rogier and J. Schneider, *A Direct Method For Solving the Boltzmann Equation*, *Transp. Th. Stat. Phys.* **23** (1994), no. 1-3, 313–338.
30. J. Schneider, *Une méthode déterministe pour la résolution de l'équation de Boltzmann*, Thèse, Université Paris VI, 1993.
31. Ye. M. Shakhov, *The Two-dimensional Non-linear Problem of the Motion of a Rarefied Gas Between Two Parallel Plates.*, *Zh. Vychisl. Mat. Mat. Fiz.* **35** (1994), 83–94.
32. H. Struchtrup, *The BGK-Model with Velocity-Dependent Collision Frequency*, *Continuum Mech. Thermodyn.* **9** (1997), 23–31.
33. W. Waluś, Lecture Series in Mathematical Theory of the Boltzmann Equation, Series on Advances in Mathematics for Applied Sciences, ed. N. Bellomo (World Scientific), 1994.
34. J.Y. Yang and J.C. Huang, *Rarefied Flow Computations Using Nonlinear Model Boltzmann Equations*, *Journal of Computational Physics* **120** (1995), 323–339.
35. H. C. Yee, *A Class of high-resolution explicit and implicit shock-capturing methods*, von Karman Institute for Fluid Dynamics, Lectures Series, n^{o}4, 1989.