A Discrete-Velocity Scheme for the Boltzmann Operator of Rarefied Gas Dynamics

C. Buet

CEA-CEL-V
94195 Villeneuve Saint Georges Cedex, France

Abstract

We propose a conservative and entropic discrete-velocity method to compute the solutions of the Boltzmann equation in the case of monoatomic species. We begin by defining a discrete collision kernel on a velocity lattice which verifies all the properties of the continuous kernel. The continuous Boltzmann equation will be replaced by a Boltzmann equation for a discrete velocity gas, which is a hyperbolic system. This equation will be discretized by a finite volume scheme. For the evaluation of the collision term we use acceleration procedures of Monte Carlo type. The possibilities of our scheme will be illustrated by numerical tests in 1 and 2 space dimensions.

1 Introduction

In rarefied aerodynamics the kinetic model which is currently used is the Boltzmann equation, because we consider only binary collisions. Most of the numerical codes for rarefied aerodynamics are based on Monte-Carlo procedures, following the earlier work of gas dynamicists [1]. A description of these methods is given, for example, in [7, 13, 1]. However, in many situations, the numerical fluctuations which originate in the use of random sequences, lead to extremely noisy results and motivate the search for alternate, more accurate methods [11, 14, 15, 9, 10]. The aim of this paper is to present some attempts in this direction. The discrete velocity model that we use is close to the model presented in [11]. The most important point of our work is the technique to reduce the cost of the collision phase.

2 Conservation properties of the Boltzmann equation

We consider the Boltzmann equation for a monoatomic gas

$$\frac{\partial f}{\partial t} + v.\nabla_x f = Q(f,f), \quad f|_{t=0} = f_0(x,v) \tag{1}$$

$$Q(f,f) = \int_{\mathbb{R}^3} \left(\int_{S^2} q(v-v_*,\omega)(f'f'_* - ff_*) d\omega \right) dv_* \tag{2}$$

with the following notations

$$S^2 = \{\omega \in \mathbb{R}^3,\ |\omega|^2 = 1\}$$

$$f = f(x,v,t),\ f_* = f(x,v_*,t),\ f' = f(x,v',t),\ f'_* = f(x,v'_*,t)$$

$$v' = \frac{v+v_*}{2} + R_\omega(\frac{v-v_*}{2}), \quad v'_* = \frac{v+v_*}{2} - R_\omega(\frac{v-v_*}{2}), \tag{3}$$

where $R_\omega(\vec{u})$ is defined by

$$R_\omega(\vec{u}) = \cos\theta\vec{u} + |\vec{u}|\sin\theta(\cos\varphi\vec{i}_{\vec{u}} + \sin\varphi\vec{j}_{\vec{u}}),$$

with $\omega = (\cos\theta, \sin\theta\cos\varphi, \sin\theta\sin\varphi)$, $|\vec{j}_{\vec{u}}| = |\vec{j}_{\vec{u}}| = 1$ and $(\vec{u}, \vec{i}_{\vec{u}}, \vec{j}_{\vec{u}})$ form an orthogonal base of $\mathbb{R}^3$. The velocities v and v_* are pre-collisional velocities, while v' and v'_* are post-collisional velocities. They satisfy

$$v + v_* = v' + v'_* \quad \text{(conservation of momentum)} \tag{4}$$

$$|v|^2 + |v_*|^2 = |v'|^2 + |v'_*|^2 \quad \text{(conservation of energy)} \tag{5}$$

Finally, $q(v,\omega)$ is defined by

$$q(v,\omega) = |v|\sigma(v,\omega), \tag{6}$$

where $\sigma(v,\omega) = \sigma(|v|, cos\theta)$ is the differential scattering cross section. Well-known properties of the Boltzmann collision operator (2) are conservation of mass, momentum and energy, and decay of entropy

$$\int_{\mathbb{R}^3} Q(f,f) \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = 0 \tag{7}$$

$$\int_{\mathbb{R}^3} Q(f,f)\log(f)dv \leq 0. \tag{8}$$

More precisely, let $\psi(v)$ be any smooth test function. We have

$$\begin{aligned} &\int_{\mathbb{R}^3} Q(f,f)\psi dv = \\ &= -\frac{1}{4}\int_{\mathbb{R}^3}\int_{\mathbb{R}^3}\left(\int_{S^2} q(v-v_*,\omega)(f'f'_* - ff_*)(\psi' + \psi'_* - \psi - \psi_*)d\omega\right)dvdv_* \end{aligned} \tag{9}$$

It is well known that the only functions ψ such that $\int Q(f,f)\psi dv = 0$ are linear combinations of 1, v, and $|v|^2$.
Similarly, (8) follows from (9) (with $\psi = \log f$). Finally, from (9), it follows that any equilibrium distribution function, i.e., any f satisfying $Q(f,f) = 0$ is a Maxwellian

$$f(v) = \frac{\rho}{(2\pi RT)^{3/2}} \quad exp(-\frac{|v-u|^2}{2RT}), \tag{10}$$

where $\rho, T \in \mathbb{R}$, $\rho > 0, T > 0$, and $u \in \mathbb{R}^3$. (ρ, u, T) are the density, mean velocity and temperature of gas. In the homogeneous case, from (8) follows the H-theorem

$$\frac{d}{dt}\int\int_{\mathbb{R}^3} f(v,t)\log f(v,t)dxdv = \int_{\mathbb{R}^3} Q(f,f)(1+\log f)dv \leq 0, \tag{11}$$

showing that the entropy $H(f) = \int f\log f dv$ can only decrease during a time evolution. Furthermore, $H(f)$ can only be minimal if f is an equilibrium distribution, i. e. if f is a Maxwellian. The same result holds in the inhomogeneous case in the absence of boundaries

$$\frac{d}{dt}\int\int_{\mathbb{R}^3\times\mathbb{R}^3} f(x,v,t)\log f(x,v,t)dxdv = \int_{\mathbb{R}^3\times\mathbb{R}^3} Q(f,f)(1+\log f)dxdv \leq 0. \tag{12}$$

It is extremely important to preserve this latter property in any numerical discretization, because it expresses conditions that any dynamics must satisfy. Thus, a numerical scheme should satisfy discrete analogues of (7), (8), (9), should exhibit discrete invariants of collision, connected with discrete equilibrium distribution functions. The failure of deterministic particle methods to satisfy these requirements naturally leads to the following discrete-velocity model.

3 A discrete-velocity model

We first deal with the space homogeneous equation.

$$\frac{df}{dt} = Q(f,f), \quad f|_{t=0} = f_0(v), \tag{13}$$

where $Q(f,\ f)$ is given by (2). We introduce a regular discretization of $\mathbb{R}^3$: Let $\Delta v > 0$, $v_i = i\Delta v$, $i = (i_1, i_2, i_3) \in \mathbb{Z}^3$, and $f_i \simeq (\Delta v)^3 f(v_i)$. We derive an approximation of (2) by using a quadrature formula for the integral with respect to v_*, the quadrature points of which are the lattice points of $\Delta v.\mathbb{Z}^3$. We let

$$\begin{aligned} &[\int_{\mathbb{R}^3} (\int_{S^2} q(v-v_*,\omega)(f'f'_* - ff_*)d\omega)dv_*]_{v=v_i} \\ &\simeq \sum_{j\in\mathbb{R}^3} (\int_{S^2} q(v_i - v_j, \omega)(f(v_i^{'})f(v_j^{'}) - f(v_i)f(v_j))d\omega)\Delta v^3, \end{aligned} \tag{14}$$

where $v_i^{'}$, $v_j^{'}$ are defined by (3) and where v and v_* are replaced by v_i and v_j. At this point, the loss term already depends on the values $f(v_i)$, $f(v_j)$ of the distribution function f, while the gain term still depends on f, through an integral over $\omega \in S^2$. We have to find a quadrature formula for the integral at the right-hand side of (14), which involves values of f at the lattice points $\Delta v\mathbb{Z}^3$. Formula (3) shows that, when ω varies in S^2, $v^{'}$ varies on the sphere of largest diameter (v, v_*), and $v_*^{'}$ is diametrically opposed to $v^{'}$ (i. e. $(v^{'}, v_*^{'})$ is another largest diameter of the same sphere). Furthermore, quite generically, the sphere of largest diameter $(v_i,\ v_j)$ contains other lattice points v_k and such points appear in pairs of diametrically opposed points $(v_k,\ v_l)$. The set S_{ij} of such pairs can be defined by

$$S_{ij} = \{(k,l) \in \mathbb{Z}^3 \times \mathbb{Z}^3, \quad k+l = i+j, \quad |k|^2 + |l|^2 = |i|^2 + |j|^2\}.$$

Furthermore, for $(k,l) \in S_{ij}$, we can define a unique $\omega_{ij}^{kl} \in S^2$, such that formula (3) holds for $v = v_i$, $v_* = v_j$, $v' = v_k$, $v'_* = v_l$.

Now, for the integral with respect to ω which appears at the right hand-side of (14), we can use a quadrature formula where the ω_{ij}^{kl} are the quadrature points and the weights are all equal to $4\pi/\mathrm{Card}(S_{ij})$. This gives

$$\int_{S^2} q(v_i - v_j, \omega) f(v_i^{'}) f(v_j^{'}) d\omega \simeq \sum_{(k,l)\in S_{ij}} \frac{4\pi}{\mathrm{Card}(S_{ij})} q(v_i - v_j, \omega_{ij}^{kl}) f(v_k) f(v_l) \tag{15}$$

and in (15), the distribution function f has been replaced by its values $f(v_k)$, $f(v_l)$ at the lattice points. The overall approximate collision operator is now written by letting $\bar{f} = \{f_i, i \in$

$\mathbb{Z}^3\}$,

$$Q(f,f)(v_i) \simeq \tfrac{1}{(\Delta v)^3}\bar{Q}(\bar{f},\bar{f})_i$$

$$= \frac{1}{(\Delta v)^3}\sum_{j\in\mathbb{Z}^3}\sum_{(k,l)\in S_{ij}}\frac{4\pi}{\mathrm{Card}(S_{ij})}q(v_i-v_j,\omega_{ij}^{kl})(f_kf_l-f_if_j) \tag{16}$$

or

$$\bar{Q}(\bar{f},\bar{f})_i = \frac{1}{2}\sum_{(j,k,l)\in(\mathbb{Z}^3)^3}(A_{k,l}^{i,j}f_kf_l - A_{ij}^{kl}f_if_j), \tag{17}$$

witht

$$A_{ij}^{kl} = \begin{cases} \dfrac{4\pi q(v_i-v_j,w_{ij}^{kl})}{\mathrm{Card}(S_{ij})} & \text{if} \quad (k,l)\in S_{ij}. \\ \\ 0 \quad \text{otherwise} \end{cases} \tag{18}$$

Formula (17) shows that this approximation enters the class of discrete-velocity models. No error estimates for the quadrature formula (15) is available yet. One problem is that the number of quadrature points ω_{ij}^{kl} on S^2 is a function of $v_j - v_i$, which has an (almost) random behaviour. Another one is that, even if this number was known accurately, the location of the points ω_{ij}^{kl} on S^2 also varies (apparently) randomly. We have just an estimate of the number of quadrature points ω_{ij}^{kl} by adapting a very classical theorem of number theory, on how to split an integer into a sum of three squares of integers (see [8]). Indeed the center of the collisions spheres $S_{i,j}$ are in $\frac{1}{2}.\mathbb{Z}^3$, given a center of collisions sphere $\frac{\epsilon}{2}$ with $\epsilon = (\epsilon_1,\epsilon_2,\epsilon_3) \in \mathbb{Z}^3$, by translation, we can suppose that $\epsilon \in \mathbb{N}^3$ and $\max_{i=1}^{3}(\epsilon_i) \leq 2$. All the spheres having $\frac{\epsilon}{2}$ for center and which intersect $\mathbb{Z}^3$ have a radius of the form $|i-\frac{\epsilon}{2}|$ with $i\in\mathbb{Z}^3$ such that $\left|i-\frac{\epsilon}{2}\right|^2 \in \mathbb{N} + \left|\frac{\epsilon}{2}\right|^2$. If we let

$$r_\epsilon(n) = \mathrm{Card}\Big(\{i\in\mathbb{Z}^3 \;\; / \;\; \left|i-\frac{\epsilon}{2}\right|^2 = n + \left|\frac{\epsilon}{2}\right|^2\}\Big) \;,\; with\; n\in\mathbb{N}$$

be the number of points of $\mathbb{Z}^3$ on the sphere having the center $\frac{\epsilon}{2}$ and the radius $(n+\left|\frac{\epsilon}{2}\right|^2)^{\frac{1}{2}}$, we have the result:

Lemma 1

$$\sum_{k=0}^{n} r_\epsilon(k) = \frac{4}{3}\pi n^{\frac{3}{2}} \;\; + O(n)$$

and $r_\epsilon(n) = O(n^{\frac{1}{2}+\delta})$ *for all* $\delta > 0$ *or equivalently,* $r_\epsilon(n) = o(n^{\frac{1}{2}+\delta})$ *for all* $\delta > 0$.

For the proof see [2].

Since

$$\mathrm{Card}(S_{i,j}) = \frac{1}{2}r_\epsilon(\left|\frac{i-j}{2}\right|^2) \;\; with \;\; \epsilon \equiv i-j \pmod 2 \;,\;\; \max_{i=1}^{3}(\epsilon_i) \leq 2.$$

in the sense of the Cesaro mean value, $\mathrm{Card}(S_{i,j})$ behaves like $\frac{4}{3}\pi\left|\frac{i-j}{2}\right|$. In some sense, this tends to show that the approximation is "reasonably accurate". But a rigorous proof of

the accuracy of the approximation is missing up to now. We must now bound the velocity domain. The issue is to replace the Boltzmann equation in the whole velocity space domain, by a bounded space one, for which the algebraic properties displayed in section 3.1 still hold. We do that as in [14, 15]. Let V be a bounded velocity domain, and let $I(v, v_*, v', v'_*)$ be the following truncation function:

$$I(v, v_*, v', v'_*) = \begin{cases} 1 \quad \text{if} \quad (v, v_*, v', v'_*) \in V^4 \\ \\ 0 \quad \text{otherwise.} \end{cases} \tag{19}$$

Now, let us consider the Boltzmann operator:

$$Q(f, f)(v) = \int_V \int_{S^2} q(v - v_*, \omega)(f' f'_* - f f_*) I(v, v_*, v', v'_*) dv_* d\omega, \quad v \in V. \tag{20}$$

It is easily shown that properties (7) to (11) still hold, with the only difference that the coefficient of $|v|^2$ in (10) is no more positive. Indeed, its positivity for the full space case follows from integrability requirements on the Maxwellian, which can no longer be used because of the boundedness of the domain. The discrete velocity model is now restricted to approximations f_i of $(\Delta v)^3 f(v_i)$ for $v_i \in \Delta v \mathbb{Z}^3 \cap V$ and the discrete Boltzmann operator is of the form (17)

$$\tilde{S}_{ij} = \{(k, l) \in S_{ij} \quad \text{such that} \quad v_k, v_l \in V\}, \tag{21}$$

with i, j such that $v_i, v_j \in \Delta v \mathbb{Z}^3 \cap V$.

$$\tilde{A}_{ij}^{kl} = \begin{cases} \dfrac{4\pi q(v_i - v_j, \omega_{ij}^{kl})}{\text{Card}(\tilde{S}_{ij})} \quad \text{if} \quad v_i, v_j \in V \quad \text{and} \quad (k, l) \in \tilde{S}_{ij} \\ \\ 0 \quad \text{otherwise.} \end{cases} \tag{22}$$

In practice, all the discrete pre- and post-collisional velocities must be within the computational domain V, and the number of "allowed" post-collisional pairs must be considered in the quadrature formula for integrals on S^2.

4 Properties of the discrete-velocity model

First, since the cross section is supposed to depend only of $|v|$ and $cos(\theta)$, it is easy to see that the tensor A_{ij}^{kl} is positive and satisfies the following symmetry properties:

$$A_{ij}^{kl} = A_{ji}^{kl} = A_{ij}^{lk} \tag{23}$$

and

$$A_{ij}^{kl} = A_{kl}^{ij}. \tag{24}$$

Property (23) expresses that the two pre-collisional particles are undistinguishable. The same is also true for the two post-collisional particles. Property (24) is the microreversibility. We have the identity (9): let $\bar{\psi} = (\psi_i)_{i \in \mathbb{Z}^3}$ be a test sequence. Then

$$\sum_{i \in \mathbb{Z}^3} \bar{Q}(\bar{f}, \bar{f})_i \psi_i = -\frac{1}{8} \sum_{(i,j,k,l) \in (\mathbb{Z}^3)^4} (A_{kl}^{i,j} f_k f_l - A_{ij}^{kl} f_i f_j)(\psi_k + \psi_l - \psi_i - \psi_j). \tag{25}$$

Using the definition of the tensor A_{ij}^{kl} it is easy to see that we have the discrete analogue of conservation of mass, momentum and energy

$$\sum_{i\in\mathbb{Z}^3} \bar{Q}(\bar{f},\bar{f})_i \begin{pmatrix} 1 \\ v_i \\ |v_i|^2 \end{pmatrix} = 0, \tag{26}$$

and by the microreversibility property of the tensor, the decay of entropy hold (see [4]):

$$\sum_{i\in\mathbb{Z}^3} \bar{Q}(\bar{f},\bar{f})_i \log(f_i) \leq 0. \tag{27}$$

The equilibrium state $\bar{f}^\infty$ is characterized by one of the following properties (see [4]):

1. $\sum_{i\in\mathbb{Z}^3} \bar{Q}(\bar{f}^\infty,\bar{f}^\infty)_i \log(f_i^\infty) = 0$
2. $\bar{Q}(\bar{f}^\infty,\bar{f}^\infty)_i = 0$ for all i
3. $\overline{\log f^\infty}$ is an invariant of collision that is $\overline{\log f^\infty} \in \{\bar{\varphi}$ such that $\varphi_i + \varphi_j - \varphi_k - \varphi_l = 0$ if $A_{ij}^{kl} \neq 0\}$
4. $f_i^\infty f_j^\infty - f_k^\infty f_l^\infty = 0$ if $A_{ij}^{kl} \neq 0$

Now, for the specific model given by (18), it is noticeable that the reciprocal of (26) holds, like in the continuous case and so the equilibrium states are discrete Maxwellians:

Lemma 2 *With A_{ij}^{kl} defined by (18) the invariants of collisions are given by*

$$\psi_i = Av_i^2 + < B, v_i > + C$$

with A et $C \in \mathbb{R}$ and $B \in \mathbb{R}^3$.

Proof We say that $(i,j) \to (k,l)$ is a possible collision if $A_{i,j}^{k,l} \neq 0$, that implies by the symmetry of $\mathbb{Z}^3$ that the collision $(-i,-j) \to (-k,-l)$ is also possible. Let $e_1 = (1,0,0),\ e_2 = (0,1,0),\ e_3 = (0,0,1)$ be the canonical base of $\mathbb{Z}^3$. We search for ψ such that

$$\psi_i + \psi_j - \psi_k - \psi_l = 0 \quad for\ all\ A_{i,j}^{k,l} \neq 0.$$

We set

$$a_i = \psi_i + \psi_{-i}\ and\ b_i = \psi_i - \psi_{-i}.$$

By construction we have $a_{-i} = a_i$, $b_{-i} = -b_i$ and then $b_0 = 0$. We show recursively on $m = |i|_\infty = \max_{n=1}^3 |i_n|$ that

$$a_i - a_0 = |i|^2.(a_{e_1} - a_0) \quad b_i = i_1 b_{e_1} + i_2 b_{e_2} + i_3 b_{e_3}.$$

This is evidently true for a_i when $i = e_1$, e_3 or e_3 because $(e_k,-e_k) \to (e_l,-e_l)$ is a possible collision. For b_i in this case this is trivial. We suppose now that it is true until rank m. Let $i \in \mathbb{Z}^3$ such that $|i|_\infty = m+1$. If $(i,j) \to (k,l)$ is a possible collision then, by construction, $a_i + a_j = a_k + a_l$ and $b_i + b_j = b_k + b_l$. Since the following collisions are possible

$$(i,0) \to (i_1e_1 + i_2e_2, i_3e_3), \quad (i_1e_1 + i_2e_2, 0) \to (i_1e_1, i_2e_2),$$

we have

$$a_i + a_0 = a_{i_1e_1+i_2e_2} + a_{i_3e_3} = a_{i_1e_1} + a_{i_2e_2} + a_{i_3e_3} - a_0$$

and then

$$a_i - a_0 = (a_{i_1 e_1} - a_0) + (a_{i_2 e_2} - a_0) + (a_{i_3 e_3} - a_0)$$

and for b_i

$$b_i = b_{i_1 e_1} + b_{i_2 e_2} + b_{i_3 e_3}.$$

It suffices then to verify that:

$$a_{(m+1)e_k} - a_0 = (m+1)^2 (a_{e_1} - a_0) \quad , \quad b_{(m+1)e_k} = (m+1) b_{e_k}. \tag{28}$$

We define $u = (m-1)e_k$, $v = me_k + e_l$ and $w = me_k - e_l$. Since $|u|_\infty = m-1$ and $|v|_\infty = |w|_\infty = m$ the assertion holds for u, v, w. By the inductive hypothesis and since the following collision is possible

$$\left((m+1)e_k \ , \ u \right) \ \rightarrow \ \left(\ v \ , \ w \ \right)$$

(28) is true: we have for a_i

$$a_{(m+1)e_k} + a_u = a_v + a_w,$$

and then

$$\begin{aligned} a_{(m+1)e_k} - a_0 &= a_w - a_0 + a_v - a_0 - (a_u - a_0) \\ &= (|w|^2 + |v|^2 - |u|^2)(a_{e_1} - a_0) \\ &= (m^2 + 1 + m^2 + 1 - (m-1)^2)(a_{e_1} - a_0) \\ &= (m+1)^2 (a_{e_1} - a_0), \end{aligned}$$

and for b_i

$$b_{(m+1)e_k} = b_v + b_w - b_u = m b_{e_k} + b_{e_l} + m b_{e_k} - b_{e_l} - (m-1) b_{e_k} = (m+1) b_{e_k}.$$

Since $\psi_i = \dfrac{a_i + b_i}{2}$ we have the result for ψ_i with $A = \frac{a_{e_1} - a_0}{2\Delta v}$, $C = \frac{a_0}{2\Delta v}$, $B = (\dfrac{b_{e_1}}{2\Delta v}, \dfrac{b_{e_2}}{2\Delta v}, \dfrac{b_{e_3}}{2\Delta v})$.
□

Remark 1 *This proof shows also that, in the case of a bounded domain for v of the form $V = B(\vec{U}, R)$ or $V = \vec{U} + [-R, R]^3$ (which we use in practice) which define, after a translation of vector $\vec{U}$, a bounded domain for i of the form $\{i \in \mathbb{Z}^3 \ / \ i_1^2 + i_2^2 + i_3^2 \leq M\}$ or $\{i \in \mathbb{Z}^3 \ / \ sup_{k=1}^3 i_k \leq M\}$, the result for the form of the invariants of collisions remains valid.*

Since the only invariants of collision are $1, v, |v|^2$ the constants A, B, C only depend on the density, mean velocity, and temperature of $\bar{f}$(see [4]).

These properties show that the discrete models (17), (18), (17) or (22) satisfy the requirements that we have stated at the end of section (3.1).

5 Discretization in space and time

Now we define N as Card$((\Delta v \mathbb{Z}^3) \cap V)$ and we let $\mathcal{V}_N = \{v_i \in (\Delta v \mathbb{Z}^3) \cap V \ , \ i = 1, .., N\}$ be the finite set of velocities. We set now

$$S_{ij} = \{\{v_k, v_l\} \in (\Delta v \mathbb{Z}^3)^2, \quad v_k + v_l = v_i + v_j, \quad |v_k|^2 + |v_l|^2 = |v_i|^2 + |v_j|^2\}.$$

$$\tilde{S}_{i,j} = \{\{v_k, v_l\} \in S_{ij} \; and \; v_k, v_l \in V\}$$

The problem is now to solve

$$\frac{\partial f_i}{\partial t} + v_i.\nabla f_i = Q(\bar{f}, \bar{f})_i = \sum_{j=1}^{N} \sum_{\{v_k, v_l\} \in S_{ij}} (A_{kl}^{ij} f_k f_l - A_{ij}^{kl} f_i f_j) \tag{29}$$

where $f_i = f_i(x,t)$ is an approximation of $(\Delta v)^3 f(x,v,t)$ at the point v_i, A_{ij}^{kl} is defined by (22) and $\bar{f} = (f_1, .., f_N)$. As usual, we use a splitting method between the transport and the collision phase.

5.1 Numerical transport

The numerical treatment of the convective term can be done in a variety of ways, e.g., finite differences, finite volumes, the method of characteristics, or particle methods. We have developed the second one. The first one must be associated with directional splitting which leads to unpleasant directional effects. The third one implies a large dependency upon the data at the previous time step, and would involve heavy storage. The fourth one needs the storage of the positions of the particles on top of the storage of the value of f. This would also imply too heavy storage. We restrict the presentation to 2-D computations on quadrangular meshes and to a single equation of convection

$$\frac{\partial f(x,t)}{\partial t} + v.\nabla f(x,t) = 0.$$

We introduce a partitioning of the computation domain by a set $\mathcal{M}$ of cells M, and take a time increment $\Delta t > 0$. Given a cell M we suppose that we know an approximation f_M^n of

$$\frac{1}{|M|}.\int_M f(x, n\Delta t)dx.$$

Let A, B, C, D be the vertices of the cell M, (A,B), (B,C),.. the sides of the cell, n_{AB}, n_{BC},..the unit outward normals of each side, l_{AB}, l_{BC},.. the length of the sides, M_{AB}, M_{BC},.. the nearest neighbors of M and AB, BC,... the midpoint of each side. On each cell M we define a function g_M which verifies

$$f_M^n = \frac{1}{|M|}.\int_M g_M dx,$$

and we define g by $g = g_M$ on each cell M.

We let $g_{AB}^{in} = \lim_{x \to AB \ , \ x \in M} g(x)$ and $g_{AB}^{out} = \lim_{x \to AB, \ x \notin M} g(x)$ and similarly for the other sides.

The finite volume scheme is defined by

$$f_M^{n+1} = f_M^n - \frac{\Delta t}{|M|}\Bigg(flux(g_{AB}^{in}, g_{AB}^{out}).l_{AB} + flux(g_{BC}^{in}, g_{BC}^{out}).l_{BC} + flux(g_{CD}^{in}, g_{CD}^{out}).l_{CD} + flux(g_{DA}^{in}, g_{DA}^{out}).l_{DA} \Bigg),$$

where the function $flux(q^{in}, q^{out})$ is defined by

$$flux(q^{in}, q^{out}) = [\max(v.n, 0)]q^{in} + [\min(v.n, 0)]q^{out},$$

and n is the outward unit normal for the considered side.

The usual first order finite volume scheme, obtained by taking g constant on each cell M i.e., $g_M(x) = f_M^n$, is very dissipative and yields smoothed shock profiles, for which the Rayleigh line is no longer a straight line. We recall that the Rayleigh line is the set of $\mathbb{R}^2$ which contains the pairs $(\frac{1}{n}, p_{xx})$, where n is the density and p_{xx} is the longitudinal component of the pressure tensor, $p_{xx} = \int_{\mathbb{R}^3}(v_x - u_x)^2 f(x, v)dv$. When x ranges over the shock region, this set is a straight line. This is a good test for numerical methods as it is noted in [5], because how close the numerical Rayleigh line is to a straight line tells us how good the method is. The scheme also dissipates too much on a nonuniform grid.

We thus turn to Van Leer's method [16] to achieve second order accuracy in space and this leads to much better results. For this we take g linear on each cell:

$$g_M(x) = f_M^n + (\nabla f)_M^n.(x - x_M)$$

where x_M is the centroid of the cell and $(\nabla f)_M^n$ is an approximation of the gradient of f on the cell M limited such that:

$$g_{AB}^{in} \in [\min(f_M^n, f_{M_{AB}}^n), \max(f_M^n, f_{M_{AB}}^n)]$$

and similarly for the other sides.

The use of the Van-Leer method gives much better results, in particular in the case of shock profile, for which the Rayleigh line is now really close to a straight line.

A sufficient condition for stability for these two schemes is the conservation of positivity of f_M^{n+1}. For the first order scheme, one can verify that under the CFL condition

$$\max_{M\in\mathcal{M}} \frac{\Delta t}{|M|} (\max(v.n_{AB}, 0).l_{AB} + \max(v.n_{BC}, 0).l_{BC} + \max(v.n_{CD}, 0).l_{CD} + \max(v.n_{DA}, 0).l_{DA}) \leq 1,$$

the scheme is positive. In 1-D this reduces to

$$\max_{M\in\mathcal{M}} \frac{v\Delta t}{(\Delta x)_M} \leq 1.$$

Under this condition for the time step f_M^{n+1} is then a convex linear combination of the values $f_M^n, f_{M_{AB}}^n, f_{M_{BC}}^n, f_{M_{CD}}^n, f_{M_{DA}}^n$. Using the convexity of the function $x \to x \log x$ and in the absence of boundary we have then

$$\sum_{M\in\mathcal{M}} |M| f_M^{n+1} \log(f_M^{n+1}) \leq \sum_{M\in\mathcal{M}} |M| f_M^n \log(f_M^n),$$

which gives a discrete analogue of property (12) for the transport phase.

For the second order accuracy in space we have a much stronger CFL condition. By noting that by construction we have

$$g_{AB}^{in} + g_{BC}^{in} + g_{CD}^{in} + g_{DA}^{in} = 4f_M^n$$

and all of the values g_{AB}^{in},g_{AB}^{in}... are positives, then under

$$\max_{M\in\mathcal{M}} \frac{\Delta t}{|M|} sup\,(\max(v.n_{AB},0).l_{AB}, \max(v.n_{BC},0).l_{BC},$$
$$\max(v.n_{CD},0).l_{CD}, \max(v.n_{DA},0).l_{DA}) \leq \frac{1}{4}$$

the scheme is positive. In 1-D we recall that $\frac{1}{4}$ can be replaced by $\frac{1}{2}$.

5.2 Full collision phase

In each space cell, the problem is to compute an approximation $\bar{f}^{n+1}$ to the solution of the Cauchy problem

$$\left\{ \begin{array}{l} \dfrac{d\bar{f}(t)}{dt} = Q(\bar{f},\bar{f}), \\ \\ \bar{f}(0) = \bar{f}^n \end{array} \right. \tag{30}$$

where $Q(\bar{f},\bar{f}) = (Q(\bar{f},\bar{f})_1,..,Q(\bar{f},\bar{f})_N)$. The requirements of the time discretization scheme are that the solution at time Δt has the same first five moments that $\bar{f}^n$ and that the Maxwellians are steady state of the problem. An immediate solution is to use an Euler explicit scheme:

$$\bar{f}^{n+1} = \bar{f}^n + \Delta t Q(\bar{f}^n,\bar{f}^n).$$

The main questionable point about the method is its computational cost. The cost of the evaluation of the discrete collision operator in the right-hand side of (29) is of order $N^{2+(1/3)+\delta}$ for general differential cross section and if the differential cross section σ does not depend on ω, this cost can be reduced to order N^2 (see [2]). A computational complexity of order N^2 is much too large for a practical use of the algorithm. We propose different ways to reduce this cost. All these methods involve random choices (i.e., are of Monte Carlo type). The amount of "randomness" varies from one to the other.

5.3 Collision phase: first acceleration procedure using randomized sublattice.

The first acceleration technique (method A) which can be used is the following modification of the full method.

Let $b \in \mathbb{N}$, $b \geq 1$ and let $a \in \mathbb{N}^3$, such that

$$\max_{s=1,3} a_s \leq b-1$$

(where a_s denotes the s-th coordinate of a, $s = 1,2,3$). Then, $L_{a,b} = a\Delta v + b.\Delta v.\mathbb{Z}^3$ be a sublattice of $\Delta v.\mathbb{Z}^3$. We use it for the quadrature formula (14) instead of the original lattice $\Delta v.\mathbb{Z}^3 = L_{0,1}$. More precisely, we write

$$[\int_{\mathbb{R}^3}\int_{S^2} q(v-v_*,\omega)(f'f'_* - ff_*)d\omega dv_*]_{v=v_i}$$
$$\simeq \sum_{j/v_i+v_j\in L_{a,b}} (b\Delta v)^3 \int_{S^2} q(v_i-v_j,\omega)(f(v_i^{'})f(v_j^{'}) - f(v_i)f(v_j))d\omega.$$

We notice that the volume of the elementary cell of the coarser lattice $L_{a,b}$ is now $(b\Delta v)^3$. Then, the evaluation of the ω-integral is done exactly in the same way as previously, by taking the pairs $\{v_k, v_l\}$ of points of the finer lattice $\Delta v.\mathbb{Z}^3$, which belong to $\tilde{S}_{i,j}$. The resulting simplified operator of collisions can be written

$$Q(f,f)(v_i) \simeq \frac{1}{(\Delta v)^3} Q_{i,a,b}(\bar{f},\bar{f}) = \sum_{j/v_i+v_j \in L_{a,b}} \sum_{\{v_k,v_l\}\in \tilde{S}_{i,j}} b^3 \{A_{kl}^{ij} f_k f_l - A_{ij}^{kl} f_i f_j\}. \tag{31}$$

We can write

$$Q_{i,a,b}(\bar{f},\bar{f}) = G_{i,a,b}(\bar{f},\bar{f}) - p_{i,a,b}(\bar{f}).f_i$$

with

$$G_{i,a,b}(\bar{f},\bar{f}) = \sum_{j/v_i+v_j \in L_{a,b}} \sum_{\{v_k,v_l\}\in \tilde{S}_{i,j}} b^3 \{A_{kl}^{ij} f_k f_l\} \geq 0$$

which is the gain term for v_i and

$$p_{i,a,b}(\bar{f}) = \sum_{j/v_i+v_j \in L_{a,b}} \Big(\sum_{\{v_k,v_l\}\in \tilde{S}_{i,j}} A_{ij}^{kl} \Big) f_j \geq 0$$

is the collision frequency for v_i.

By using the facts that $\bigcup_{a\in A_b} L_{a,b} = \Delta v \mathbb{Z}^3$ and if $a_1 \neq a_2$ then $L_{a_1,b} \bigcap L_{a_2,b} = \emptyset$, where $A_b = \{0,1,...,b-1\}^3$, we remark that we have the relations

$$\frac{1}{b^3} \sum_{a\in A_b} p_{i,a,b}(\bar{f}) = p_i(\bar{f}) = p_{i,0,1}(\bar{f}) \tag{32}$$

and

$$\frac{1}{b^3} \sum_{a\in A_b} G_{i,a,b}(\bar{f},\bar{f}) = G_i(\bar{f},\bar{f}) = G_{i,0,1}(\bar{f},\bar{f}), \tag{33}$$

that implies the following decomposition of $Q(\bar{f},\bar{f})$:

$$Q(\bar{f},\bar{f}) = Q_{0,1}(\bar{f},\bar{f}) = \frac{1}{b^3} \sum_{a/\max_{s=1}^3 a_s < b-1}^{b^3} Q_{a,b}(\bar{f},\bar{f}).$$

We use the same time discretization as for the full method:

$$\bar{f}^{n+1} = \bar{f}^n + \Delta t Q_{a,b}(\bar{f}^n,\bar{f}^n).$$

Now, let us assume that b has been chosen ≥ 2. In order to preserve the accuracy of the finer lattice, (even though we use at some step of the discretization a coarser lattice), *the triple a is chosen randomly at each time step, in the set* A_b.

We can hope that the accuracy of the finer grid is reached for steady problems, if we look at a mean value result after some time steps, since the expectation of this random choice of $Q_{a,b}(\bar{f},\bar{f})$ is indeed $Q_{0,1}(\bar{f},\bar{f})$ as the relations (32) and (33) show it. In this sense, we can say that this scheme preserves the accuracy of the finer mesh.

It is important that the symmetry properties (23) and (24) are still satisfied. It is clear for (23). For (24), it suffices to notice that if $(v_k, v_l) \in \tilde{S}_{ij}$, then $v_k + v_l = v_i + v_j \in L_{a,b}$. Therefore,

Table 1 computational costs

	b=1	b=3	b=6
8^3	1	0.0 4	0.0 05
16^3	64	2. 5	0.3
32^3	4096	150	19

properties (25) to (27) still hold for the discrete collision operator (31). However, it is no longer obvious that the only invariants of collision, (which are sequences $\bar{\psi} = (\psi_1, .., \psi_N)$, such that $\psi_k + \psi_l - \psi_i - \psi_j = 0$ for $(v_i, v_j) \in (\Delta v \mathbb{Z}^3)^2$ s. t. $v_i + v_j \in L_{a,b}$ and $(v_k, v_l) \in \tilde{S}_{i,j}$) are linear combinations of 1, $v_i, |v_i|^2$ or that the system of equations is again coupled. With the choice of a at each time step, it is also clear, on homogeneous problems, that the only steady states are the same Maxwellians as those obtained with the full collision operator.

The same estimates of the computational efficiency can be made for this new method. The evaluation of the collision operator over one time step costs of the order of $\frac{(N^{2+(1/3)+\delta})}{b^3}$ operations in the general case and N^2/b^3 in the case of an ω-independent scattering cross section. For comparison, let us consider $8^3, 16^3$ and 32^3 points discretizations of the velocity space, for an ω-independent scattering cross section, without any sublatticing ($b = 1$) and with $b = 3$ and $b = 6$. Assume that computational cost is 1 for 8^3 and $b = 1$. Then, the costs which are obtained for these various situations are given in table 1. Any doubling of the number of discretization points in one direction multiplies the computational cost by a factor 4096. However, if the doubling is associated with sublatticing (with a doubling of the sublatticing while doubling the number of points), the computational cost increases milder. It has been verified numerically that important sublatticing (b of the order of 6) does not affect the results in any noticeable way, at least as far as moments of the distribution function are concerned, which are the most physically interesting quantities.

It is important for the problem that $\bar{f}^{n+1}$ remains positive. Since the scheme can be written

$$f_i^{n+1} = f_i^n(1 - \Delta t p_{i,a,b}(\bar{f}^n)) + \Delta t . G_{i,a,b}(\bar{f}^n, \bar{f}^n),$$

then we can see that under the condition

$$\sup_{i=1}^{N} p_{i,a,b}(\bar{f}^n)\Delta t \leq 1, \tag{34}$$

the scheme preserve the positivity of the solution.

5.4 Collision phase: second acceleration procedure using a 4-velocity model and splitting of operator.

When the support of the distribution function is too small, that is, when almost all of the mass is concentrated on a small number of points, one can verify that the sublattice method leads to collision frequencies $p_{i,a,b}(\bar{f}^n)$ which fluctuate too much. If we let $\Delta t_{a,b}$ be the maximum time step allowed by (34) for the sublattice method with parameters b and a, and if we suppose that $\bar{f}$ is like a δ function we have

$$\Delta t_{a,b} \sim \frac{\Delta t_{0,1}}{b^3},$$

and in this case we gain nothing compared to the full method. On the other hand, when all the f_i are equal to a constant we have

$$\Delta t_{a,b} \sim \Delta t_{0,1},$$

then in this case we have the maximum efficiency for the sublattice method. Moreover, it is not clear that if Δt satisfies (34) we have the decay of entropy, that is

$$\sum_{i=1}^{N} f_i^{n+1} \log(f_i^{n+1}) \leq \sum_{i=1}^{N} f_i^n \log(f_i^n). \tag{35}$$

It would be nice to have an acceleration method which is unconditionally stable in time, produces less fluctuating collision frequencies and such that the decay of entropy holds. We propose two other acceleration techniques satisfying the first and third points. They are based on a splitting method of the operator in order to reduce the scheme to a 4-velocity model, indeed in this case we have an exact solution. We hope that the second one satisfies the second wish.

For the sake of simplicity, we explain these methods in the case of an isotropic cross section i.e., σ does not depend on ω, as it is the case for the VHS model used in aerodynamics.

To begin with, we must do some algebraic manipulations on the collision operator. We call S_V the collection of the spheres $\tilde{S}_{i,j}$. We have the partition of the set of pairs $\{v_i, v_j\}$ where $v_i, v_j \in \mathcal{V}_N$:

$$\Big\{\{v_i, v_j\}, v_i \in \mathcal{V}_N, v_j \in \mathcal{V}_N\Big\} = \bigcup_{S \in S_V} S.$$

We can write $Q_{a,b}(\bar{f}, \bar{f}) = (Q_{1,a,b}(\bar{f}, \bar{f}), .., Q_{N,a,b}(\bar{f}, \bar{f}))$ defined by (31) as

$$Q_{a,b}(\bar{f}, \bar{f}) = \sum_{i=1}^{N} (Q_{i,a,b}) e_i = \sum_{S \in S_V} a_S Q_S(\bar{f}, \bar{f})$$

with $a_S = b^3$ if the center of S is in $L_{a,b}$ or $a_S = 0$ if it is not the case, the operator $Q_S(\bar{f}, \bar{f})$ is defined by

$$Q_S(\bar{f}, \bar{f}) = \sum_{\{v_i, v_j\} \in S} \frac{C_S}{\mathrm{Card}(S)} \left(\sum_{\{v_k, v_l\} \in S} \Big(f_k f_l - f_i f_j \Big) \right) (e_i + e_j),$$

$(e_1, ...e_N)$ is the canonical base of $\mathbb{R}^N$ and the constant C_S is defined by(see formula (22))

$$C_S = 4\pi q(\mathrm{diam}(S)).$$

If for two pairs $\{v_i, v_j\}$ and $\{v_l, v_k\}$ in S we define the operator $E_{i,j,k,l}(\bar{f}, \bar{f})$ by

$$E_{i,j,k,l}(\bar{f}, \bar{f}) = (f_k f_l - f_i f_j)(e_i + e_j - e_k - e_l),$$

using the symmetries properties

$$E_{i,j,k,l}(\bar{f}, \bar{f}) = E_{j,i,k,l}(\bar{f}, \bar{f}) = E_{i,j,l,k}(\bar{f}, \bar{f}) = E_{k,l,i,j}(\bar{f}, \bar{f}),$$

we can write

$$Q_S(\bar{f},\bar{f}) = \frac{C_S}{\mathrm{Card}(S)} \sum_{\{v_i,v_j\}\in S} \sum_{\{v_k,v_l\}\in S} \frac{1}{2} E_{i,j,k,l}(\bar{f},\bar{f}), \tag{36}$$

or, if we let $\mu = \{\mu_1, ..., \mu_N\}$so that $\mu_i > 0$ for all i,

$$Q_S(\bar{f},\bar{f}) = \frac{C_S}{\mathrm{Card}(S)} \sum_{\{v_i,v_j\}\in S} (\mu_i+\mu_j) \sum_{\{v_k,v_l\}\in S} \frac{1}{\mu_i+\mu_j+\mu_k+\mu_l} E_{i,j,k,l}(\bar{f},\bar{f}).$$

With this last decomposition of $Q_S(\bar{f},\bar{f})$ for each sphere S of S_V we can also write $Q(\bar{f},\bar{f}) = Q_{0,1}(\bar{f},\bar{f})$, since we shall use this form later, as

$$Q(\bar{f},\bar{f}) = \sum_{i=1}^{N}\sum_{j=1}^{N} \mu_j \sum_{\{v_k,v_l\}\in \tilde{S}_{ij}} \frac{C_{\tilde{S}_{ij}}}{\mathrm{Card}(\tilde{S}_{ij})(\mu_i+\mu_j+\mu_k+\mu_l)} E_{i,j,k,l}(\bar{f},\bar{f}). \tag{37}$$

Now we turn to the splitting method that we will use for time discretization. We suppose that we have defined an approximation $\tilde{Q}(\bar{f}^n,\bar{f}^n)$ of $Q(\bar{f}^n,\bar{f}^n)$ as a linear combination of terms $E_{i,j,k,l}(\bar{f}^n,\bar{f}^n)$, that is

$$\tilde{Q}(\bar{f}^n,\bar{f}^n) = \sum_{p=1}^{P} c_p E_{i_p,j_p,k_p,l_p}(\bar{f}^n,\bar{f}^n)$$

where P is some integer and the coefficients c_p are positive. With the definition of the operator $E_{i,j,k,l}$, it is clear that $\tilde{Q}(\bar{f},\bar{f})$ preserve the five first moments and the Maxwellians. In the Euler explicit scheme

$$\bar{f}^{n+1} = \bar{f}^n + \Delta t \tilde{Q}(\bar{f}^n,\bar{f}^n)$$

for solving problem (30), $\bar{f}^{n+1}$ can also be viewed as an approximation for the solution at time Δt to the differential equation

$$\frac{d\bar{f}(t)}{dt} = \tilde{Q}(\bar{f},\bar{f}),$$

with the initial condition

$$\bar{f}(0) = \bar{f}^n.$$

Another approximation of this solution can be obtained by the usual splitting technique for operator. Let $\pi(p)$ be a permutation of the indices p, and let $\bar{g}^p = (g_1^p, .., g_N^p)$ be the solution of the problem

$$\frac{d\bar{g}^p(t)}{dt} = c_{\pi(p)} E_{i_{\pi(p)},j_{\pi(p)},k_{\pi(p)},l_{\pi(p)}}(\bar{g}^p,\bar{g}^p),\ \bar{g}^p(0) = \bar{g}^{p-1}(\Delta t) \tag{38}$$

for $p = 1, .., P$ and by defining $\bar{g}^0$ as $\bar{f}^n$. Then $\bar{g}^P(\Delta t)$ is an approximation of $\bar{f}^{n+1}$.

The interest of this is that we can exhibit the solution for each step of this splitting technique. Solving $\frac{d\bar{g}}{dt} = c_{\pi(p)} E_{i_{\pi(p)},j_{\pi(p)},k_{\pi(p)},l_{\pi(p)}}(\bar{g},\bar{g})$ is equivalent to solving the homogeneous Boltzmann equation for a discrete velocity gas with only four velocities (four velocities Broadwell model). If we consider two pairs of velocities $\{v_1, v_2\}$ and $\{v_3, v_4\}$ which are two

diameters of a same sphere, the homogeneous Boltzmann equation for this discrete velocity gas is:

$$\frac{df_1(t)}{dt} = C\left(f_3(t)f_4(t) - f_1(t)f_2(t)\right)$$
$$\frac{df_2(t)}{dt} = C\left(f_3(t)f_4(t) - f_1(t)f_2(t)\right)$$
$$\frac{df_3(t)}{dt} = -C\left(f_3(t)f_4(t) - f_1(t)f_2(t)\right)$$
$$\frac{df_4(t)}{dt} = -C\left(f_3(t)f_4(t) - f_1(t)f_2(t)\right).$$

The Cauchy problem for this Boltzmann equation with the initial data

$$f_1(0) = f_1^0\ ,\ f_2(0) = f_2^0\ ,\ f_3(0) = f_3^0\ ,\ f_4(0) = f_4^0.$$

has for solution

$$f_1(t) = f_1^0 + A(t)\ ,\ f_2(t) = f_2^0 + A(t)\ ,\ f_3(t) = f_3^0 - A(t)\ ,\ f_4(t) = f_4^0 - A(t)$$

where

$$A(t) = \frac{(f_3^0 f_4^0 - f_1^0 f_2^0)}{\rho}(1 - e^{-\rho C t}) \quad and \quad \rho = \sum_{i=1}^{4} f_i^0.$$

Since we have an exact solution, the H-theorem holds

$$\frac{d}{dt}\sum_{i=1}^{4} f_i(t)\log(f_i(t)) \leq 0. \tag{39}$$

Since at each step of the splitting technique (38) we solve exactly the equation, the method is unconditionally stable in time and verifies (35), that is, we have the decay of entropy.

Remark 2 *As in computations the function* $B(t) = \frac{1}{\rho}(1 - e^{-\rho C t})$ *could be too expensive to evaluate, it would be possible to replace it by an approximation* $h(t)$ *which has the same behaviour:*

- $h(0) = 0$ *and* $0 < h(t) < \frac{1}{\rho}$
- $h'(0) = C$ *and* $h'(t) \geq 0$
- $lim_{t\to\infty} h(t) = \dfrac{1}{\rho}$

this leads to the approximation of the weights $f_1, .., f_4$

$$f_i(t) = f_i^0 \pm (f_3^0 f_4^0 - f_1^0 f_2^0)h(t).$$

We have then

$$\frac{d}{dt}\sum_{i=1}^{4} f_i(t)\log f_i(t) = \frac{h'(t)}{(1-\rho h(t))}(f_3(t)f_4(t) - f_1(t)f_2(t))\log(\frac{f_1(t)f_2(t)}{f_3(t)f_4(t)}) \tag{40}$$

and since the term $\dfrac{h'(t)}{(1-\rho h(t))}$ *remains positive the decay of the entropy still holds for* $t \in [0, \infty[$. *In the case of an explicit time discretization for solving the Boltzmann equation for the*

four velocities system, which correspond to the approximation of $B(t)$ by Ct, we are sure that the decay of entropy hold only for $t \in [0, \frac{1}{C\rho}[$. Another problem with this time discretization is that the relaxation to the Maxwellians is too fast. Another natural approximate is

$$h(t) = \frac{(1 - \frac{1}{1+\rho Ct})}{\rho}$$

which corresponds to the approximation of e^{-x} by $\dfrac{1}{1+x}$ but now the relaxation to the maxwellians is too slow. A solution could be to tabulate the function e^{-x} and to do linear interpolation between two consecutive values.

We now propose two methods using splitting for reduction to 4-velocities systems.

The first one (method B) is derived from the sublattice method. As we have seen before, in the sublattice method at each time step we take the following approximation of $Q(\bar{f}^n, \bar{f}^n)$

$$Q_{a,b}(\bar{f}^n, \bar{f}^n) = \sum_{S \in S_V} a_S Q_S(\bar{f}^n, \bar{f}^n)$$

where $Q_S(\bar{f}, \bar{f})$ is given by (36). We make then a new approximation for each $Q_S(\bar{f}^n, \bar{f}^n)$. Let $(c_1, ..., c_{\mathrm{Card}(S)})$ be the elements of S, which are pairs of velocities. Given a random permutation π of the indices of c we form the pair $(c_{\pi(2p)}, c_{\pi(2p+1)})$, we take the following approximation of $Q_S(\bar{f}^n, \bar{f}^n)$:

$$\tilde{Q}_S(\bar{f}^n, \bar{f}^n) = \Big(\frac{C_S \mathrm{Card}(S)}{\mathrm{Card}(S) + 1 - \epsilon}\Big) \sum_{p=1}^{2} E_{i(c_{\pi(2p)}), j(c_{\pi(2p)}), k(c_{\pi(2p+1)}), l(c_{\pi(2p+1)})}(\bar{f}^n, \bar{f}^n)$$

where $\epsilon = 0$ or 1, $\mathrm{Card}(S) \equiv \epsilon \pmod 2$. The expectation of this random approximation is indeed Q_S^n. Now we use the splitting method (38) in the following way: we pick the sphere S for which the center is in $L_{a,b}$ in a randomized fashion and for each S we solve all of the equations

$$\frac{d\bar{f}(t)}{dt} = a_S \Big(\frac{C_S \mathrm{Card}(S)}{\mathrm{Card}(S) + 1 - \epsilon}\Big) E_{i(c_{\pi(2p)}), j(c_{\pi(2p)}), k(c_{\pi(2p+1)}), l(c_{\pi(2p+1)})}(\bar{f}, \bar{f})$$

which is equivalent in fact to solve exactly the system

$$\frac{d\bar{f}(t)}{dt} = a_S \tilde{Q}_S(\bar{f}, \bar{f}),$$

because we have cut the set of velocities of S in systems of four velocities which are not coupled.

The second method (method C), using the splitting technique (38), starts by using formula (37) for the discrete collision kernel $Q(\bar{f}, \bar{f})$. We use a Monte Carlo quadrature formula with technique of importance sampling for obtaining an approximation of $Q(\bar{f}^n, \bar{f}^n)$. We suppose μ so that $\sum_{i=1}^N \mu_i = 1$. We choose M velocities v_{j_m} according to the probability law μ and we take

$$Q(\bar{f}^n, \bar{f}^n) \simeq \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{N} \sum_{\{v_k, v_l\} \in \tilde{S}_{ij_m}} \frac{C_{\tilde{S}_{ij_m}} E_{i,j_m,k,l}(\bar{f}^n, \bar{f}^n)}{\mathrm{Card}(\tilde{S}_{ij_m})(\mu_i + \mu_{j_m} + \mu_k + \mu_l)}$$

and for each sum over the sphere defined by v_i and v_{j_m} we choose randomly a pair (k_{i,j_m}, l_{i,j_m}). This gives the new approximation

$$Q(\bar{f}^n, \bar{f}^n) \simeq \tilde{Q}(\bar{f}^n, \bar{f}^n) = \frac{1}{M} \sum_{m=1}^{M} \sum_{i=1}^{N} \frac{C_{\tilde{S}_{ij_m}} E_{i,j_m,k_{i,j_m},l_{i,j_m}}(\bar{f}^n, \bar{f}^n)}{\mu_i + \mu_{j_m} + \mu_{k_{i,j_m}} + \mu_{l_{i,j_m}}}.$$

In order to have the best approximation possible, the probability law μ must be chosen to minimize the variance of the process. In practice we choose μ close as possible to $\frac{f_j^n}{\sum_{k=1}^N f_k^n}$. When we are far away from thermal equilibrium, we suggest to take $\mu = \frac{\bar{f}^n}{\sum_{k=1}^N f_k^n}$. Near equilibrium we suggest to take the Maxwellian which has the same five first moments that $\frac{f_j^n}{\sum_{k=1}^N f_k^n}$.

Now we apply the splitting method (38) with $\tilde{Q}(\bar{f}, \bar{f})$ defining by

$$\tilde{Q}(\bar{f}, \bar{f}) = \frac{1}{M} \sum_{i=1}^{N} \sum_{m=1}^{M} \frac{C_{\tilde{S}_{ijm}}}{\mu_i + \mu_{j_m} + \mu_{k_{i,j_m}} + \mu_{l_{i,j_m}}} E_{i,j_m,k_{i,j_m},l_{i,j_m}}(\bar{f}, \bar{f}).$$

For the permutation π in the process (38) we take a random permutation.

Remark 3 *In the absence of boundary it can be noticed that the methods described in this paragraph, since they verify (35) for all space cells, give, when they are combined with the first order finite volume method for the transport phase, a discrete analogue of (12).*

6 Numerical results

6.1 Shock wave

We compare the results obtained with our dicrete velocity method(DVM) with those obtained with a direct simulation Monte Carlo (DSMC) code in the case of a shock wave for a hard sphere gas. This DMSC code uses the Bird method without time counter for the collision phase.

The Mach number of the shock wave is approximatively 6.2. For the comparisons the calculations are made in an unsteady fashion by using a classical procedure to produce a shock. At the beginning the flow is uniform with a velocity $u_\infty = -8$ and a temperature such that $RT_\infty = 1.85$ and at x=0 we put a specular wall. The domain of the computation is $[0, 25\lambda_\infty]$, where λ_∞ is the mean free path at infinity. We look for the solution when the shock arrived at a distance $15\lambda_\infty$ of the wall which correspond at $t = 1.14$.

For the DSMC computation the time step is $\Delta t = .006 = 0.8\tau_\infty$, where τ_∞ is the mean free time at infinity and we take a uniform grid with $\Delta x = 0.5\lambda_\infty$. The number of particles at the initial time is 1200 in each cell of space. The DSMC computation takes 300 seconds on a Cray YMP.

The DVM computations are made with the same grid. The time step is $\Delta t = 0.8\tau_\infty$ and for a variant without splitting between collisions and transport $\Delta t = 0.533\tau_\infty$.

For the DVM we take $V = B(0, 12)$ and $\Delta v = 2.4$ and $\Delta v = 1.2$ which correspond to $N = 515$ and $N = 4169$ velocities. We use the sublattice acceleration (method A) and the values of the parameters b are:

$b = 3$ for $N = 515$, $b = 3$ and 6 for $N = 4169$.

On the Cray YMP the method with $N = 515$ and $b = 3$ takes 10 seconds. With 4169 velocities and $b = 6$ the computation takes 86 seconds.

The comparisons are made on the following profiles:

- Profiles of the density and the temperature.

- Rayleigh line: p_{xx} as a function of $\frac{1}{\rho}$. One can see that through the shock we have the relation $A\frac{1}{\rho} + p_{xx} = B$ where A and B are two constants (see [3]) and then the points $(\frac{1}{\rho}, p_{xx})$ are on a straight line called Rayleigh line [5].

As one can see on figures (1) and (2), the results are good for both choices of the number of velocities. For the Rayleigh line the best results are obtained with the second order scheme for the transport phase. We note that with a smaller Δv, $\Delta v = .8$ which correspond to approximatively 14000 velocities we do not improve the result corresponding to N=4169.

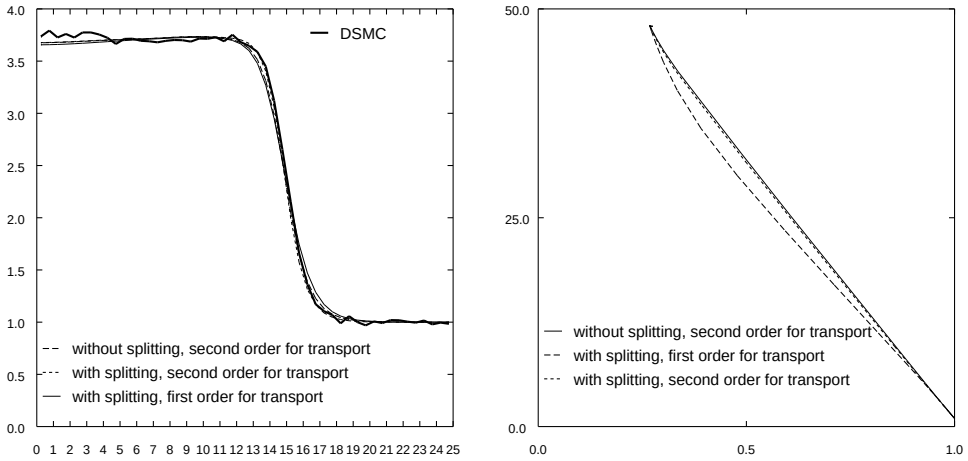

Figure 1: Density and Rayleigh line with $N = 515$ and $b = 3$.

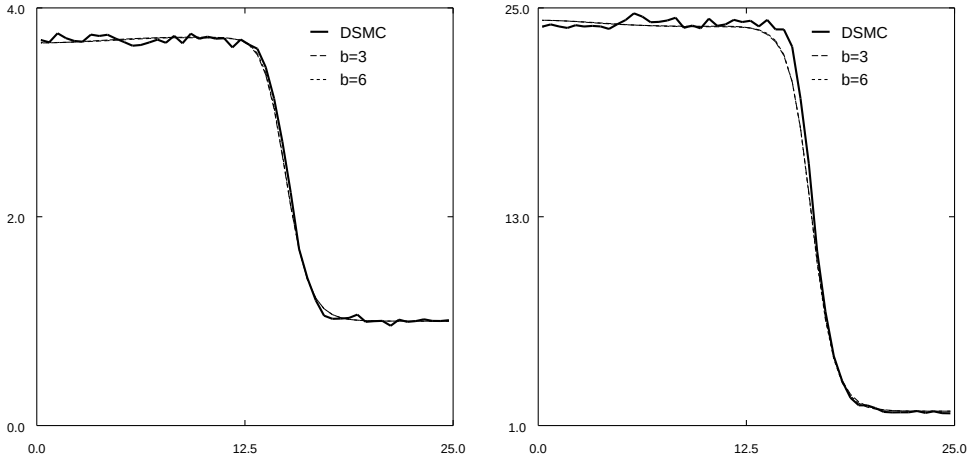

Figure 2: Variation of the parameter b. Density and RT with $N = 4169$ and second order for the transport step.

6.2 Two-dimensionnal results: compression ramp

We compare the results obtained with the same DSMC method used for monodimensionnal problems, with those obtained with our DVM for a compression ramp placed in a supersonic flow. The geometry is a flat plate of 5 cm followed by a ramp of 10^o. The characteristics for the flow at infinity are those for Mach 4 and Mach 20 as in [12]. Because we consider a monoatomic gas the Mach numbers are in fact 3.67 and 18.8.

At mach 3,67

-$v_\infty = 669,3\ m/s$

-$n_\infty = 2,769\ 10^{21}\ /m^3$

-$T_\infty = 69,76\ K$

-molecular mass: $4,815\ 10^{-26} kg$

-temperature of wall: $T_w = 336\ K$

-the mean free path at infinity: $\lambda_\infty = 2,348\ 10^{-4}\ m$

-the mean free path at the wall: $\lambda_\infty = 2,158\ 10^{-4}\ m$

-Knudsen number at infinity: $Kn_\infty = 0,0047$

At mach 18,8

-$v_\infty = 1503\ m/s$

-$n_\infty = 3,716\ 10^{20}\ /m^3$

-$T_\infty = 13,32\ K$

-molecular mass: $4,651\ 10^{-26}\ kg$

-temperature of wall: $T_w = 290\ K$

-the mean free path at infinity: $\lambda_\infty = 2,35\ 10^{-3}\ m$

-the mean free path at the wall: $\lambda_\infty = 1,03\ 10^{-3}\ m$

-Knudsen number at infinity: $Kn_\infty = 0,047$

The cross section in the two cases is of the VHS type. For the expression of the cross section, the mean free path and the values of parameters used in the VHS model see [12]. In the two cases we have a perfect accommodation at the wall. For DSMC and our scheme we used the same grid. For the flow at Mach 3,67 we used a nonuniform grid of 5250 quadrangulars elements. The size of the mesh in the direction perpendicular to v_∞ at the beginning of the flat plate and the corner are of the order of the mean free path near the wall. At Mach 18,8 we used a uniform grid with 3589 quadrangulars elements. The parameters for the DVM are the following:

-Mach 3.67: sublattice (method B) with $N = 515$, b=3, $\Delta t = 5.10^{-7} s$, and the numbers of iterations is 900.

-Mach 18.8: method C with $N = 3405$, $\Delta t = 3.89\ 10^{-6} s$, and the numbers of iterations is 58. Since the flow is far away from thermal equilibrium, we take $\mu = \dfrac{f_j^n}{\sum_{k=1}^N f_k^n}$.

For the DVM, we initialized the computations with Maxwellians such that they give the exact density mean velocity and temperature after projection on the velocity grid. The DSMC runs take 120 minutes on a CRAY YMP. At Mach 3,67 the number of samples is 1500 with an average of 20 particles in space cell. At Mach 18,8 the number of sample is 700 again with an average of 20 particles in space cell. At Mach 3,67 the DVM takes 120 minutes CPU (approximatively 80 per cent of the time for the transport phase) with the same computer as for DSMC. For Mach 18,8 the discrete velocity method takes 60 minutes (approximatively 80 per cent of the time for the collision phase).

On pictures (3) to (10) the density, temperature, velocity in the y direction and Mach number are plotted for the two methods. The DSMC calculations, as we have an incorrect account of the boundary condition at the downstream boundary (particles can leave the domain but none can enter through this boundary) the results are bad on a small region forward the boundary and near the wall but they are not affected in the rest of the domain. For the two Mach numbers, the results of the DVM are in good accordance with the results of the DSMC method (for the four quantities shown, the isolines have the same level) and are much less noisy than DSMC results.

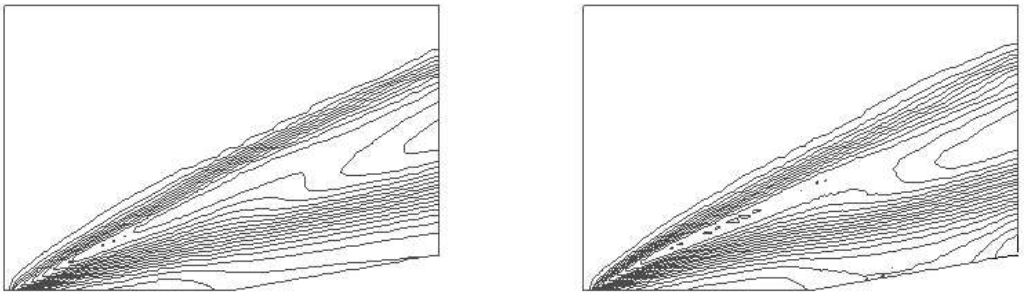

Figure 3: compression ramp at Mach 3.67, **density**, left DVM, right DSMC method.

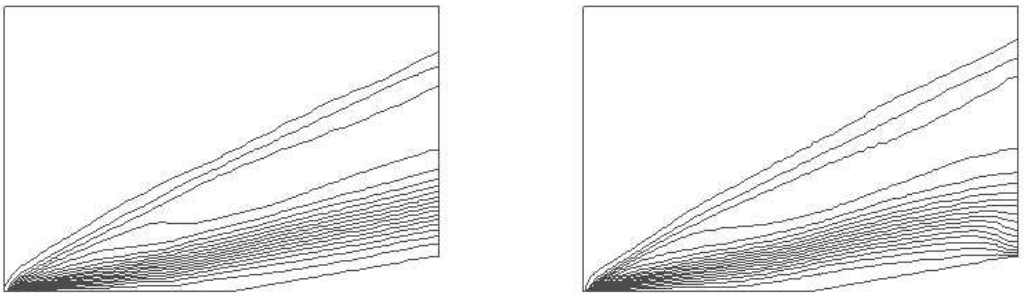

Figure 4: compression ramp at Mach 3.67, **temperature**, left DVM, right DSMC method.

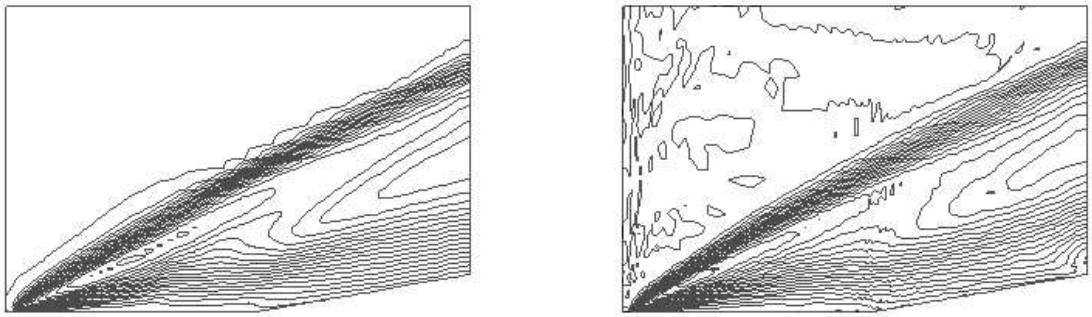

Figure 5: compression ramp at Mach 3.67, $\mathbf{v_y}$, left DVM, right DSMC method.

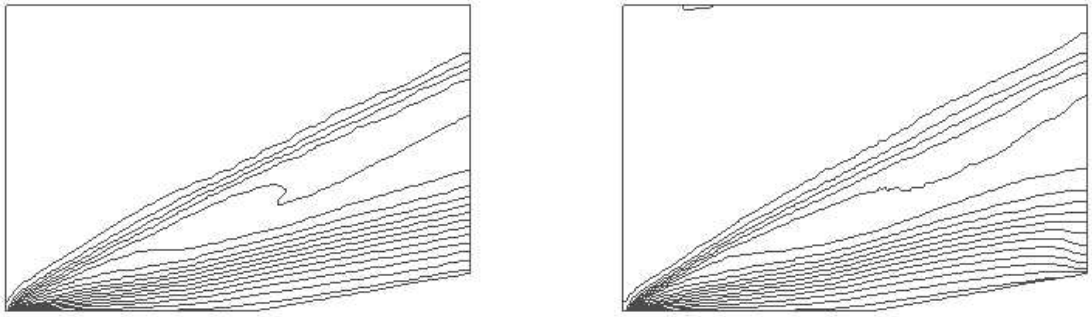

Figure 6: compression ramp at Mach 3.67, **Mach number**, left DVM, right DSMC method.

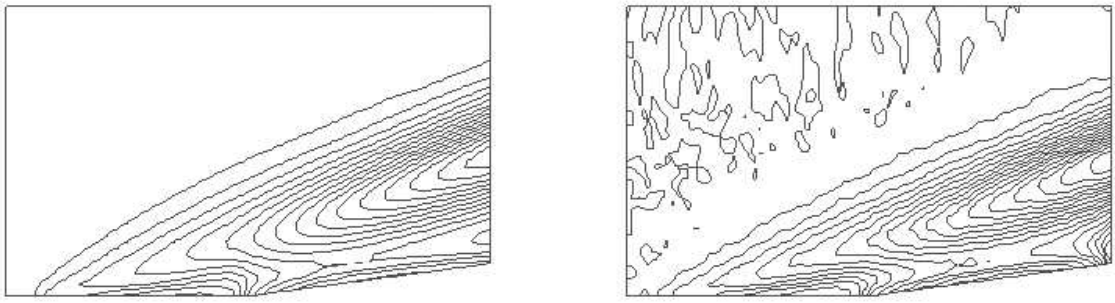

Figure 7: compression ramp at Mach 18.8, **density**, left DVM, right DSMC method.

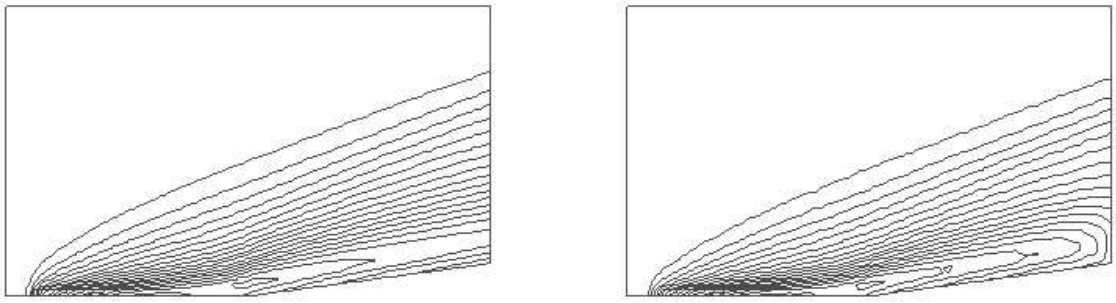

Figure 8: compression ramp at Mach 18.8, **temperature**, left DVM, right DSMC method.

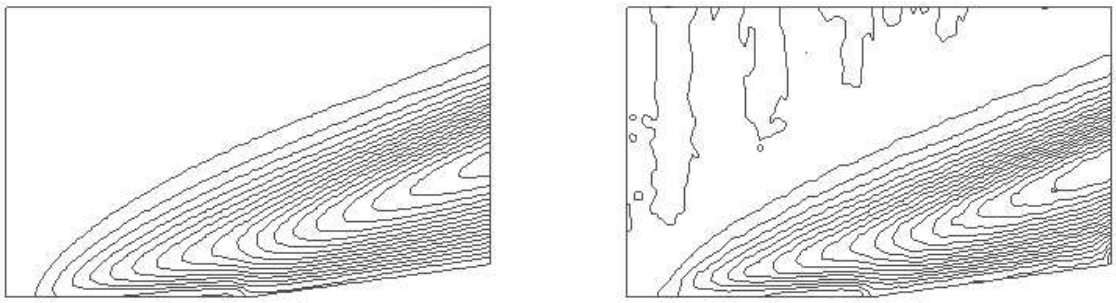

Figure 9: compression ramp at Mach 18.8, $\mathbf{v_y}$, left DVM, right DSMC method.

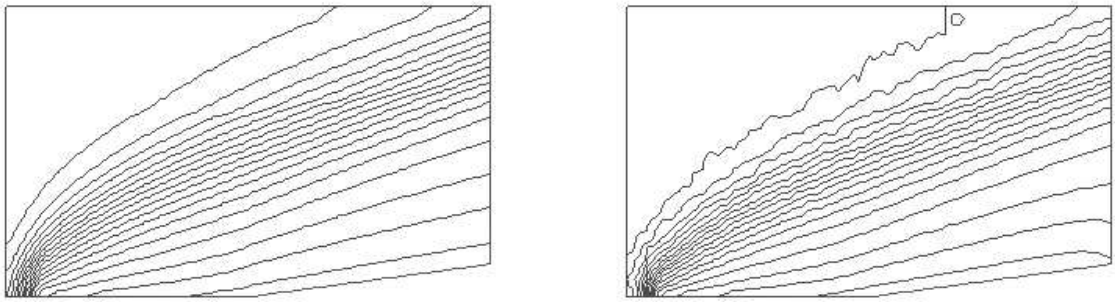

Figure 10: compression ramp at Mach 18.8, **Mach number**, left DVM, right DSMC method.

7 Conclusions

The Boltzmann equation for the discrete velocity model that we used seems to give good results in rarefied gas dynamics for monoatomic species as one can see with numerical results or in [11]. Acceleration procedures, like those we described in this paper, must be employed to give acceptable computationnal time. If employed, these acceleration techniques make the DVM an interesting alternative to the DSMC method in aerodynamics applications. Despite the fact that these acceleration procedures are of Monte Carlo type, the results remain good and seem to be almost free of noise. As one can see in [6] and by the use of our acceleration techniques, we think that we are able to extend this method to gas with internal degrees of freedom or to gas mixtures.

Acknowledgments

I am indebted to Pr P. DEGOND for many helpful discussions.

References

[1] G. A. BIRD, *"Molecular Gas Dynamics", Clarendon Press, Oxford*, (1976).

[2] C. BUET, *Résolution déterministe de l'équation de Boltzmann*, note interne CEA, (1994).

[3] C. CERCIGNANI, *The Boltzmann Equation and Its Applications*, Springer, New York, (1988).

[4] R. GATIGNOL, *Théorie cinétique des gaz à repartitions discrètes de vitesses*, Springer, New York, (1975).

[5] D. GOLDSTEIN, B. STURTEVANT and J. E. BROADWELL, *Investigations of the Motion of Discrete-Velocity Gases*, in "Rarefied Gas Dynamics: Theoretical and Computational Techniques", E. P. Muntz, D. P. Weaver and D. H. Campbell (eds), Progress in Astronautics and Aeronautics, Vol.118, AIAA, Washington DC, (1989).

[6] D. B. GOLDSTEIN, *Discrete-Velocity collision dynamics for polyatomic molecules*, Phys. Fluids A4 pp 1831-1839, (1992).

[7] F. GROPENGIESSER, H. NEUNZERT, J. STRUCKMEIER *Computational methods for the Boltzmann equation*. Venice 1989: The state of Art in Appl. and Industrial math., eds. R. Spigler, Kluwer Acad. Publ., (1990).

[8] G.H. HARDY and E.M. WRIGHT, *An introduction to the number theory*, Clarendon Press, Oxford, (1938).

[9] R. ILLNER and W. WAGNER, *A random dicrete velocity model and approximation of the Boltzmann equation* , Journal Stat. Phys. 70 (3/4) A2 pp 773-792, (1993).

[10] R. ILLNER and W. WAGNER, *random dicrete velocity models and approximation of the Boltzmann equation. Conservation of momentum and energy* , Transp. Th. Stat. Phys. 23 (1-3) A2 pp 27-38, (1994).

[11] T. INAMURO and B. STURTEVANT, *Numerical Study of Discrete-Velocity Gases*, Phys. Fluids A2 pp 2196-2203, (1990).

[12] J.C. LENGRAND, K.S. HEFFNER, A. CHPOUN, *RC 90-8 Etude 1 Rampe de compression en gaz rarefies, travaux dans le domaine de l'hypersonique GDR Hypersonique Rapport final de convention DRET(DGA) N°89.34.080.00.47075.01*, (1990).

[13] K. NANBU, *Direct simulation schemes derived from the Boltzmann equation*, J. Phys, Japan 49 p. 2042, (1980).

[14] F. ROGIER and J. SCHNEIDER, *A direct method for solving the Boltzmann Equation*, Transp. Th. Stat. Phys, (1994).

[15] J. SCHNEIDER, *Une méthode déterministe pour la résolution de l'équation de Boltzmann*, Ph.D thesis, University Paris 6, (1993).

[16] B. VAN LEER, *Towards the ultimate conservative difference scheme. V, A second order sequel of Godunov's method*, J. Comput. Phys.,Vol 32, (1979).