

Chapter 1

Conservative Discretization of the Boltzmann Equation and the Semicontinuous Model

L. Preziosi and L. Rondoni
Dip. Matematica - Politecnico di Torino - Italy
preziosi@polito.it rondoni@calvino.polito.it

1.1 Introduction

The Boltzmann equation provides an accurate description of the behavior of a rarefied gas. Unfortunately, it is very hard to numerically exploit all its capabilities and all the physical information it contains.

This is essentially related to the structure of the collision term, a five-fold integral which is very expensive to compute. Furthermore, this term needs to be evaluated with great precision, because all the mechanical properties characterizing particle collisions, e.g. conservation of mass, momentum, and energy, are included in it. A lack of conservation, even very small, may lead to serious consequences on the integration of the equation. For this reason the deduction of intrinsically conservative models results particularly important. On the other hand, as addressed in the chapter by Cercignani, in deducing discretized models one should avoid the appearance of spurious collision invariants, which is characteristic of many discrete velocity models and discretization procedures.

For this reason, in some numerical schemes proposed in the literature to integrate the Boltzmann equation the solution is suitably corrected at each time step in order to preserve mass, momentum and energy (see for instance [1, 2]). In the last decade some attention has been paid to the identification of discretization procedures that lead to models possessing most of the properties characteristic of the Boltzmann equation, namely, conservation of mass, momentum, and energy, validity of an H-theorem, and stationary states described by a Gaussian distribution. Most of the papers [3–10] deal with a discretization of the velocity space in a uniform cubic lattice, as described more in detail in other chapters of this volume. Another approach consists in uniformly discretizing the energy space, corresponding to a

discretization of the velocity space in an ensemble of spheres.

The idea of a polar discretization, already present in [11–13] has been developed in [14] where the definition of a suitable discretization of particles speeds and the introduction of a suitable interpolation procedure allows to deduce a model in which the collision operator is a sum of integrals over finite domains. These are the cartesian products of a unit circle and the portion of the unit sphere between two parallels symmetric with respect to the equatorial plane perpendicular to the velocity of the field particle. One can then exploit the fact that integration over spherical domains can be performed with good precision and small computational effort. The resulting kinetic model preserves all the mentioned physical properties.

The same characteristics are enjoyed by the model deduced in [15] where particle velocities are defined through their energy and velocity direction, the range of allowed energies is subdivided in intervals of equal width, and a stepwise interpolation is introduced. The main point addressed in this paper is the identification of the parameters involved in the discretization. This is based on the criterion that the spurious terms, which are present in the Euler equations related to the discretized models, are as small as requested by the application. The model has then been generalized by Koller and Schürerer [16–18] to handle gas mixtures, inelastic collisions, chemical reactions and photon excitation (see also the chapters by the Authors in this volume.)

In this chapter a particular attention is paid to the formulation in terms of energy and velocity direction, and to the consequences of each step of the discretization procedure. In particular the discretization procedure is divided in two main steps: 1) the definition of a limited subset of the velocity space containing all allowed velocities (Section 3); 2) its discretization (Section 4). In Section 5 it is shown that:

- The model preserves mass, momentum and energy;
- The modelling procedure is constructive, so that it is possible to obtain estimates on the “distance” between the discretized collision operator and the continuous Boltzmann operator, i.e. consistency of the quadrature rule.
- There exists an H-functional describing trend towards an equilibrium described by a Gaussian distribution.
- The fluid dynamic limit related to the discretized kinetic model tends rapidly towards the usual Euler equations with an isotropic pressure tensor, and vanishing heat flux when the number of allowed energies grows to infinity and the discretization interval for the particle energies becomes \mathbb{R}_+ .

For the sake of completeness, Section 6 gives the details of the collision dynamics in terms of energy and velocity direction. Finally, in Section 7 we draw our conclusions giving some hints for possible developments.

1.2 Splitting and Energy Formulation

In kinetic theory, the behaviour of a neutral, rarefied gas is usually described through the evolution of a distribution function, $f : [0, T] \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}_+$, which satisfies the Boltzmann equation

$$\frac{\partial f}{\partial t} = \mathcal{A}(f) + \mathcal{Q}(f, f) , \quad (1.1)$$

where

$$\mathcal{A}(f) = -\mathbf{v} \cdot \nabla_{\mathbf{x}} f , \quad (1.2)$$

is a linear differential operator and

$$\mathcal{Q}(f, f) = \int_{\mathbb{R}^3} \int_{S_2} \sigma(\gamma, g) g [f(t, \mathbf{x}, \mathbf{v}') f(t, \mathbf{x}, \mathbf{v}_*) - f(t, \mathbf{x}, \mathbf{v}) f(t, \mathbf{x}, \mathbf{v}_*)] d\hat{\mathbf{g}}' d\mathbf{v}_* , \quad (1.3)$$

is a nonlinear integral operator. In (1.3) $\sigma(\gamma, g)$ is the differential cross section, which depends on the interaction potential, $\gamma = \hat{\mathbf{g}} \cdot \hat{\mathbf{g}}'$ is related to the angle between the pre and the post collisional plane, and g is the modulus of the relative velocity \mathbf{g} of colliding particles

$$g \stackrel{def}{=} |\mathbf{v}_* - \mathbf{v}| = |\mathbf{v}'_* - \mathbf{v}'| . \quad (1.4)$$

The rule for collision between two particles is expressed by

$$\begin{cases} \mathbf{v}' = \mathbf{v} + (\mathbf{g} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \frac{\mathbf{R}}{2} + \frac{g}{2} \hat{\mathbf{g}}' , \\ \mathbf{v}'_* = \mathbf{v}_* - (\mathbf{g} \cdot \hat{\mathbf{n}}) \hat{\mathbf{n}} = \frac{\mathbf{R}}{2} - \frac{g}{2} \hat{\mathbf{g}}' , \end{cases} \quad (1.5)$$

where

$$\mathbf{R} \stackrel{def}{=} \mathbf{v} + \mathbf{v}_* = \mathbf{v}' + \mathbf{v}'_* , \quad (1.6)$$

and $\hat{\mathbf{n}}$ and $\hat{\mathbf{g}}'$ are unit vectors that span S_2 . In particular, $\hat{\mathbf{g}}'$ represents the direction of the relative velocity after collision. In this chapter hatted quantities refer to unit vectors and tilded ones to dimensionless quantities.

In the entire chapter, particles are considered as characterized by their energy per unit mass, $e = v^2/2$ (simply named energy for sake of brevity), and by their velocity direction $\hat{\mathbf{\Omega}}$. It is then useful to introduce the following quantities

$$E \stackrel{def}{=} e + e_* = e' + e'_* , \quad (1.7)$$

$$S \stackrel{def}{=} 2\sqrt{ee_*} \hat{\mathbf{\Omega}} \cdot \hat{\mathbf{\Omega}}_* = 2\sqrt{e'e'_*} \hat{\mathbf{\Omega}}' \cdot \hat{\mathbf{\Omega}}'_* , \quad (1.8)$$

$$g = \sqrt{2(E - S)}, \quad R \stackrel{def}{=} |\mathbf{R}| = \sqrt{2(E + S)}, \quad (1.9)$$

which, together with \mathbf{R} , are invariant during collision.

The description of the collision dynamics using the energy formulation is given for sake of completeness in Section 6.

The two operators in Eq.(1.1) present problems of different type to be addressed properly for a correct numerical evaluation. For this reason the so-called “splitting algorithm”, first applied by Temam [19] in the study of several evolution problems, like the Broadwell model, results very useful. This method, when applied to the Boltzmann equation, consists in decoupling the collision contributions from the free streaming ones, and looking for iterative solutions of the following problems

$$f_n^0(t=0) = \check{f}_n^0(t=0) = f_0, \quad (1.10)$$

$$\begin{cases} \frac{\partial f_n^k}{\partial t} = \mathcal{A}f_n^k, \\ f_n^k(t_k) = \check{f}_n^{k-1}(t_k), \end{cases} \quad \text{for } k > 0. \quad (1.11)$$

$$\begin{cases} \frac{\partial \check{f}_n^k}{\partial t} = \mathcal{Q}\check{f}_n^k, \\ \check{f}_n^k(t_k) = f_n^k(t_{k+1}), \end{cases} \quad \text{for } k > 0. \quad (1.12)$$

The convergence of the algorithm has been proven by Desvillettes and Mischler [20], under the same assumptions used by R.J. DiPerna and P.L. Lions [21] for their existence results for the Boltzmann equation.

Theorem 1.1 *Under assumptions 1 and 2, given below, the sequences f_n and \check{f}_n converge up to extraction of subsequences to the same nonnegative limit $f \in L^\infty([0, T]; L^1(\mathbb{R}^3 \times \mathbb{R}^3))$ weak-*, and this limit satisfies*

$$\frac{Q^\pm(f)}{1+f} \in L^1_{loc}([0, T] \times \mathbb{R}^3 \times \mathbb{R}^3), \quad (1.13)$$

and

$$\left(\frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla_x \right) \log(1+f) = \frac{Q(f)}{1+f}, \quad (1.14)$$

in the sense of distributions.

Assumption 1.1 The function $B = \sigma(\cdot, \cdot)g$ is in $L^1_{loc}(\mathbb{R}^3 \times \mathbb{S}_2)$, and depends only on g and γ . Moreover, the function

$$A(g) = \int_{\mathbb{S}_2} B(g, \gamma) d\hat{\mathbf{g}}' , \quad (1.15)$$

satisfies, for all $R > 0$,

$$\frac{1}{1+g^2} \int_{B_{\mathbf{v}}^R} A(|\mathbf{g} + \mathbf{v}|) d\mathbf{v} \rightarrow 0 , \quad \text{as } g \rightarrow \infty , \quad (1.16)$$

where $B_{\mathbf{v}}^R = \{\mathbf{v} \in \mathbb{R}^3 : |\mathbf{v}| < R\}$.

Assumption 1.2 The function f_0 is such that

$$\int_{\mathbb{R}^6} f_0(\mathbf{x}, \mathbf{v}) \{1 + |\mathbf{x}|^2 + |\mathbf{v}|^2 + |\log f_0(\mathbf{x}, \mathbf{v})|\} d\mathbf{x}d\mathbf{v} < \infty . \quad (1.17)$$

In the following we will focus on the collision step for the splitting algorithm, Eqs.(1.12), mainly for two reasons. On the one hand this step is rather heavy, from a computational viewpoint, because it involves the evaluation of a five-fold integral. On the other hand, the evaluation has to be done with great care and accuracy because the mechanical properties characterizing particle collisions, e.g. conservation of mass, momentum, and energy, are included in it. Our main concern will in fact be the link with the conservation equations, classically deduced from the moments of the Boltzmann equation

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0 , \\ \frac{\partial}{\partial t} (\rho \mathbf{U}) + \nabla \cdot (\mathbf{\Pi} + \rho \mathbf{U} \otimes \mathbf{U}) = \mathbf{0} , \\ \frac{\partial}{\partial t} \left(\mathcal{E} + \frac{1}{2} \rho U^2 \right) + \nabla \cdot \left[\mathbf{q} + \mathbf{\Pi} \mathbf{U} + \left(\mathcal{E} + \frac{1}{2} \rho U^2 \right) \mathbf{U} \right] = 0 , \end{cases} \quad (1.18)$$

where

$$\rho = m \int_{\mathbb{R}^3} f(\mathbf{v}) d\mathbf{v} = m \int_{\mathbb{R}_+} de \int_{\mathbb{S}_2} d\hat{\Omega} \sqrt{2e} f(e, \hat{\Omega}) , \quad (1.19)$$

is the mass density,

$$\mathbf{U} = \frac{m}{\rho} \int_{\mathbb{R}^3} \mathbf{v} f(\mathbf{v}) d\mathbf{v} = \frac{m}{\rho} \int_{\mathbb{R}_+} de \int_{\mathbb{S}_2} d\hat{\Omega} 2e \hat{\Omega} f(e, \hat{\Omega}) , \quad (1.20)$$

is the drift velocity, U is its modulus, and $\hat{\mathbf{U}} = \mathbf{U}/U$ its direction

$$\mathcal{E} = \frac{m}{2} \int_{\mathbb{R}^3} |\mathbf{v} - \mathbf{U}|^2 f(\mathbf{v}) d\mathbf{v} = \frac{m}{2} \int_{\mathbb{R}_+} de \int_{\mathbb{S}_2} d\hat{\Omega} |\sqrt{2e} \hat{\Omega} - \mathbf{U}|^2 \sqrt{2e} f(e, \hat{\Omega}) , \quad (1.21)$$

is the internal energy,

$$\begin{aligned}\mathbf{\Pi} &= m \int_{\mathbf{R}^3} (\mathbf{v} - \mathbf{U}) \otimes (\mathbf{v} - \mathbf{U}) f(\mathbf{v}) d\mathbf{v} \\ &= m \int_{\mathbf{R}_+} de \int_{S_2} d\hat{\Omega} \left(\sqrt{2e}\hat{\Omega} - \mathbf{U} \right) \otimes \left(\sqrt{2e}\hat{\Omega} - \mathbf{U} \right) \sqrt{2e} f(e, \hat{\Omega}),\end{aligned}\quad (1.22)$$

is the pressure tensor, and

$$\begin{aligned}\mathbf{q} &= \frac{m}{2} \int_{\mathbf{R}^3} |\mathbf{v} - \mathbf{U}|^2 (\mathbf{v} - \mathbf{U}) f(\mathbf{v}) d\mathbf{v} \\ &= \frac{m}{2} \int_{\mathbf{R}_+} de \int_{S_2} d\hat{\Omega} \left| \sqrt{2e}\hat{\Omega} - \mathbf{U} \right| \left(\sqrt{2e}\hat{\Omega} - \mathbf{U} \right) \sqrt{2e} f(e, \hat{\Omega}),\end{aligned}\quad (1.23)$$

is the heat flux.

The Maxwellian equilibrium, considered as a function of energy and velocity direction, can be written as

$$f(t, \mathbf{x}, e, \hat{\Omega}) = A \exp[\sqrt{B}e \hat{\mathbf{U}} \cdot \hat{\Omega} - Ce], \quad (1.24)$$

where the Maxwellian parameters A , B and C are related to the macroscopic observables through

$$A = \frac{\rho}{m} \left(\frac{3\rho}{4\pi\mathcal{E}} \right)^{3/2} \exp \left[-\frac{3\rho U^2}{4\mathcal{E}} \right], \quad B = \frac{1}{2} \left(\frac{3\rho U}{\mathcal{E}} \right)^2 \quad \text{and} \quad C = \frac{3\rho}{2\mathcal{E}}, \quad (1.25)$$

An important dimensionless parameter is the following:

$$r \stackrel{\text{def}}{=} \frac{B}{2C} = \frac{3\rho U^2}{2\mathcal{E}}, \quad (1.26)$$

which measures the ratio between kinetic and internal energy. It is also important to notice explicitly that if the dimensionless parameters

$$\tilde{B} = \sqrt{\frac{B}{2}} U \quad \text{and} \quad \tilde{C} = CU^2 \quad (1.27)$$

are introduced, one has

$$\tilde{B} = \tilde{C} = r \quad (1.28)$$

and the Maxwellian distribution can be written as

$$f(t, \mathbf{x}, \tilde{e}, \hat{\Omega}) = A \exp \left[r \left(\tilde{e} \hat{\mathbf{U}} \cdot \hat{\Omega} - \frac{\tilde{e}^2}{2} \right) \right] = \frac{\rho}{m} \left(\frac{3\rho}{4\pi\mathcal{E}} \right)^{3/2} \exp \left[-\frac{r}{2} \left(\tilde{e} \hat{\Omega} - \hat{\mathbf{U}} \right)^2 \right], \quad (1.29)$$

where

$$\tilde{e} = \frac{\sqrt{2e}}{U}. \quad (1.30)$$

The above definition of $\tilde{\epsilon}$ is preferred to others characterized by a proportionality between $\tilde{\epsilon}$ and ϵ (e.g. $\tilde{\epsilon} = 2\epsilon/U^2$) simply because of the simplifications in the following formula (in particular the presence of square roots is minimized).

Referring to [14] for further details, one can compute the macroscopic quantities (1.19–1.23) at Maxwellian equilibrium given as in (1.24) in terms of energy and velocity direction and, in particular, write the pressure tensor and the heat flux as

$$\mathbf{\Pi} = \Pi(\mathbf{I} + \eta_\pi \hat{\Omega} \otimes \hat{\Omega}) , \quad \mathbf{q} = \Pi(\varepsilon_q - \eta_\pi)\mathbf{U} , \quad (1.31)$$

where

$$\eta_\pi = \frac{\int_{\mathbf{R}_+^2} e^{-C(e+\epsilon')} [(Be+2)ss' - \sqrt{Be}cs' - B\sqrt{\epsilon\epsilon'}c\epsilon'] de d\epsilon'}{\int_{\mathbf{R}_+^2} e^{-C(e+\epsilon')} (\sqrt{Be}c-s)s' de d\epsilon'} , \quad (1.32)$$

$$\varepsilon_q = \frac{\int_{\mathbf{R}_+^2} \sqrt{Be}(e-\epsilon')e^{-C(e+\epsilon')}cs' de d\epsilon'}{2 \left[\int_{\mathbf{R}_+} e^{-Ce}(\sqrt{Be}c-s) de \right]^2} B - 1 , \quad (1.33)$$

and the following abbreviations have been used

$$s = \sinh \sqrt{Be} , \quad c = \cosh \sqrt{Be} , \\ s' = \sinh \sqrt{B\epsilon'} , \quad c' = \cosh \sqrt{B\epsilon'} .$$

The Euler equations can then be written in the following form

$$\begin{cases} \frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{U}) = 0 , \\ \frac{\partial}{\partial t} (\rho \mathbf{U}) + \nabla \cdot [(1 + \varepsilon_\pi) \rho \mathbf{U} \otimes \mathbf{U}] + \nabla \Pi = \mathbf{0} , \\ \frac{\partial}{\partial t} \left(\mathcal{E} + \frac{1}{2} \rho U^2 \right) + \nabla \cdot \left\{ \left[(1 + \varepsilon_q) \Pi + \mathcal{E} + \frac{1}{2} \rho U^2 \right] \mathbf{U} \right\} = 0 , \end{cases} \quad (1.34)$$

where

$$\varepsilon_\pi = \frac{\int_{\mathbf{R}_+^2} e^{-C(e+\epsilon')} [(Be+2)ss' - \sqrt{Be}cs' - B\sqrt{\epsilon\epsilon'}c\epsilon'] de d\epsilon'}{\left[\int_{\mathbf{R}_+} e^{-Ce}(\sqrt{Be}c-s) de \right]^2} . \quad (1.35)$$

The integrals in (1.32, 1.33, 1.35) can be computed analytically, and give $\eta_\pi = \varepsilon_\pi = \varepsilon_q = 0$, corresponding to an isotropic pressure tensor, a vanishing heat flux,

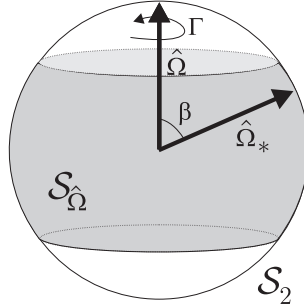


Fig. 1.1 The domain $S_{\hat{\Omega}}$.

and no extra terms in the Euler equations (1.34), as it must be. Obviously, that is not the case in numerical calculations, as explained in the next sections.

1.3 Working in a Finite Energy Interval

In this section we discuss the effect of working with a finite energy interval. The discretization procedure will be addressed in the following section.

The very first step of any discretization procedure consists in defining a limited subset of the velocity space containing the allowed velocities. As a consequence of this step, the two spurious terms in the momentum and energy Euler equations (1.34) related to the Boltzmann equation (one extra pressure tensor, in the form of a Reynolds stress term, and one extra heat flux term, in the form of a pressure) do not vanish any longer.

The presence of these terms may be directly related to the introduction of the “boundary” in the velocity space which corresponds to an approximation of the original equations, which breaks the Galilean invariance characteristic both of the continuous Boltzmann equation, and of the related Euler equations. To clarify this statement, one can consider the viewpoint of the particle which sees (or feels) this artificial boundary as a reference point. Therefore, one cannot expect that the Euler equations related to any discretization scheme have an isotropic pressure tensor and a vanishing heat flux.

In this section we give a criterion to identify an interval $\mathbb{E} = [e_m, e_M)$ such that the probability for a particle to have energy outside \mathbb{E} can be neglected.

The restriction of the energy range to a finite interval bears consequences for Eq.(1.8), because it leads to a restriction either on the outgoing energy e' , or on the angle between the incoming directions $\hat{\Omega} \cdot \hat{\Omega}_*$, if the outgoing energies are fixed. This is a well known fact:

- Given the incoming energies e, e_* and the angle between the incoming

directions $\hat{\Omega} \cdot \hat{\Omega}_*$, the outgoing energies are confined inside the interval

$$\mathcal{D}_{e'} = \left[\frac{E}{2} - \frac{gR}{4}, \frac{E}{2} + \frac{gR}{4} \right], \quad (1.36)$$

symmetrically with respect to its extrema.

For the construction of our model, it is useful to look at the same restriction from a different point of view.

- Given the incoming energies e and e_* , and an outgoing energy e' smaller than the largest energy the particle can possibly reach, $E = e + e_*$ (e'_* is then determined by the conservation of energy, $e'_* = E - e'$), there is a limited range for the angle between the pre-collisional velocity directions which yields the post-collisional energy e' . This can be expressed as:

$$(\hat{\Omega} \cdot \hat{\Omega}_*)^2 \leq \frac{e'e'_*}{ee_*}.$$

It is then useful to define the domain

$$S_{\hat{\Omega}} = \left\{ \hat{\Omega}_* : (\hat{\Omega} \cdot \hat{\Omega}_*)^2 \leq \frac{e'e'_*}{ee_*} \right\}, \quad (1.37)$$

which represents either the entire surface of the unit sphere S_2 , if $e'e'_* \geq ee_*$ or, taking $\hat{\Omega}$ as polar axis, the part between two parallels symmetric with respect to the equator, as shown in Fig. 1.

In order to understand the consequences of working in a finite energy interval, let us start with the basic steps of the derivation of the Boltzmann equation. The change during the time interval $[t, t + dt]$ of the number of molecules in the element $d\mathbf{x}d\mathbf{v}$ around (\mathbf{x}, \mathbf{v}) due to collision is

$$J d\mathbf{v} dt d\mathbf{x} = d\mathbf{v} dt d\mathbf{x} \times \int_{\mathbb{R}^3 \times S_2} \sigma(\gamma, g) g [f(t, \mathbf{x}, \mathbf{v}') f(t, \mathbf{x}, \mathbf{v}'_*) - f(t, \mathbf{x}, \mathbf{v}) f(t, \mathbf{x}, \mathbf{v}_*)] d\hat{\mathbf{g}}' d\mathbf{v}_* \quad (1.38)$$

Therefore computing the change due to collisions during the time interval $[t, t + dt]$ of the number of molecules in the volume element $d\mathbf{x}$ around \mathbf{x} , having velocity direction in the element $d\hat{\Omega}$ around $\hat{\Omega}$ and energy in $\mathcal{I} \subseteq \mathcal{E}$, one has

$$\begin{aligned} \int_{\mathcal{I}} J \sqrt{2e} de &= \int_{\mathcal{I}} \sqrt{2e} de \int_{\mathbb{R}_+} \sqrt{2e_*} de_* \int_{S_2} d\hat{\Omega}_* \times \\ &\times \int_{S_2} d\hat{\mathbf{g}}' \sigma(\gamma, g) g [f(\mathbf{v}') f(\mathbf{v}'_*) - f(\sqrt{2e}\hat{\Omega}) f(\sqrt{2e_*}\hat{\Omega}_*)] \end{aligned} \quad (1.39)$$

In our approach, it is useful to consider the output velocities not as functions of $\hat{\mathbf{g}}'$ as in (1.5), but as functions of an output energy e' and of the angle ϑ between

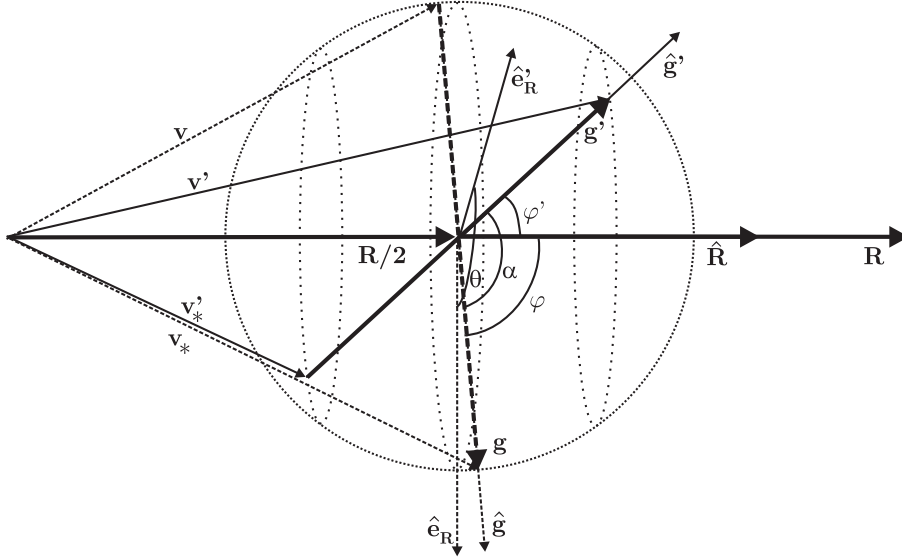


Fig. 1.2 Collision dynamics. Dashed arrows are in the plane of the figure which represents the pre-collisional plane. Full arrows refer to the post-collisional quantities which are in the plane which intersects the plane of the figure along \mathbf{R} and forms with it an angle ϑ .

the planes containing the pre- and post-collisional velocities, as described in (1.86) and (1.87) of Section 1.6, so that we can discretize the outgoing energy e' . This also means that in the scattering cross section, γ will be considered as a known function of e' and ϑ through Eq.(1.88), given in Section 1.6.

Referring to Fig. 2 (and using Eq.(1.85) of Section 1.6), one can change variable in the surface integral over $\hat{\mathbf{g}}'$ expressing $d\hat{\mathbf{g}}'$ as

$$d\hat{\mathbf{g}}' = \sin \varphi' d\varphi' d\vartheta = \frac{4}{gR} de' d\vartheta, \quad (1.40)$$

so that (1.39) can be rewritten as

$$\int_{\mathbf{I}} J\sqrt{2e} de = \int_{\mathbf{I}} de \int_{\mathbf{R}_+} de_* \int_{S_2} d\hat{\Omega}_* \times \int_{\mathbf{D}_{e'}} de' \int_0^{2\pi} d\vartheta \frac{8\sqrt{e e_*}}{R} \sigma(\gamma, g) [f(\mathbf{v}')f(\mathbf{v}_*) - f(\sqrt{2e}\hat{\Omega})f(\sqrt{2e_*}\hat{\Omega}_*)] \quad (1.41)$$

where $\mathbf{D}_{e'}$ is defined in (1.36) and depends on e, e_* and on $\hat{\Omega} \cdot \hat{\Omega}_*$.

In view of the velocity discretization, it is convenient to exchange the integrations over $\hat{\Omega}_*$ and over e' in order to group together the integrals over all the energy variables. In order to do that one has switch from the the post-collisional energies $\mathbf{D}_{e'}$, to the pre-collisional velocity directions $S_{\hat{\Omega}}$. This allows us to rewrite Eq.(1.41)

as

$$\int_{\mathbf{I}} J\sqrt{2e} de = \int_{\mathbf{I}} de \int_{\mathbf{R}_+} de_* \int_0^{e+e_*} \widehat{J}(e, e_*, e', \widehat{\Omega}) de' , \quad (1.42)$$

where

$$\begin{aligned} \widehat{J}(e, e_*, e', \widehat{\Omega}) \stackrel{def}{=} & \int_{\mathbb{S}_{\widehat{\Omega}}} d\widehat{\Omega}_* \int_0^{2\pi} d\vartheta \frac{8\sqrt{ee_*}}{R(e, e_*, \widehat{\Omega} \cdot \widehat{\Omega}_*)} \sigma(e, e_*, e', \widehat{\Omega} \cdot \widehat{\Omega}_*, \vartheta) \times \\ & \times [f(e', \widehat{\Omega}'(e, e_*, e', \widehat{\Omega}, \widehat{\Omega}_*, \vartheta)) f(e'_*, \widehat{\Omega}'_*(e, e_*, e', \widehat{\Omega}, \widehat{\Omega}_*, \vartheta)) - f(e, \widehat{\Omega}) f(e_*, \widehat{\Omega}_*)] , \end{aligned} \quad (1.43)$$

If one only considers collisions with energies in the interval $\mathcal{E} = [e_m, e_M)$, one can deduce a kinetic model of the same form as in (1.1) but with the collision operator replaced by

$$\mathcal{Q}_b(f, f) = \int_{\mathcal{E}} de_* \int_{\mathcal{E}'} \widehat{J}(e, e_*, e', \widehat{\Omega}) de' , \quad (1.44)$$

where $\mathcal{E}' = \mathcal{E} \cap [0, e + e_*]$. For this kinetic model, the conservation equations (1.18) still hold, and the Maxwellians can still be written as in (1.24), but the relations among the Maxwellian parameters and the macroscopic observables, computed over the finite domain $\mathcal{E} \times \mathbb{S}_2$, is not as simple as in (1.25).

The Euler equations can then be written in the form (1.34) with η_π , ε_π and ε_q given by formulae similar to (1.32), (1.33) and (1.35), in which \mathbb{R}_+ is replaced by \mathcal{E} . However, this replacement is not painless because these terms do not vanish any longer and give rise to spurious terms in the Euler equations.

In order to see that, take $e_m = 0$ first, and proceed in terms of the dimensionless variables introduced in (1.26) and (1.30). A lengthy calculation gives

$$\eta_\pi(0, \tilde{e}_M) = \frac{2[\mathcal{N}_\pi(\tilde{e}_M) - \tilde{\mathcal{E}}(\tilde{e}_M)\mathcal{M}_\pi(\tilde{e}_M)]}{\mathcal{D}_\pi(\tilde{e}_M)\mathcal{D}_q(\tilde{e}_M)} , \quad (1.45)$$

$$\varepsilon_\pi(0, \tilde{e}_M) = \frac{2[\mathcal{N}_\pi(\tilde{e}_M) - \tilde{\mathcal{E}}(\tilde{e}_M)\mathcal{M}_\pi(\tilde{e}_M)]}{\mathcal{D}_q^2(\tilde{e}_M)} , \quad (1.46)$$

$$\varepsilon_q(0, \tilde{e}_M) = \frac{2[\mathcal{N}_q(\tilde{e}_M) - \tilde{\mathcal{E}}(\tilde{e}_M)\mathcal{M}_q(\tilde{e}_M)]}{\mathcal{D}_q^2(\tilde{e}_M)} , \quad (1.47)$$

where $\tilde{\epsilon}_M = \sqrt{2\epsilon_M}/U$

$$\begin{aligned}
\mathcal{N}_\pi(\tilde{\epsilon}) &= 2e^{-r\tilde{\epsilon}^2} [(r+2) \sinh^2 r\tilde{\epsilon} - (r+1)r\tilde{\epsilon} \sinh r\tilde{\epsilon} \cosh r\tilde{\epsilon} - r^2\tilde{\epsilon}^2], \\
\mathcal{M}_\pi(\tilde{\epsilon}) &= e^{-r\tilde{\epsilon}^2/2} [(r+3+r^2\tilde{\epsilon}^2) \sinh r\tilde{\epsilon} - (r+3)r\tilde{\epsilon} \cosh r\tilde{\epsilon}], \\
\mathcal{N}_q(\tilde{\epsilon}) &= 2e^{-r\tilde{\epsilon}^2} \{ [r^2+4r-2 - (r+2)r^2\tilde{\epsilon}^2] \sinh^2 r\tilde{\epsilon} \\
&\quad - (r^2+3r-4)r\tilde{\epsilon} \sinh r\tilde{\epsilon} \cosh r\tilde{\epsilon} - (r+2)r^2\tilde{\epsilon}^2 \}, \\
\mathcal{M}_q(\tilde{\epsilon}) &= e^{-r\tilde{\epsilon}^2/2} [r(r+5) - r^2\tilde{\epsilon}^2] (\sinh r\tilde{\epsilon} - r\tilde{\epsilon} \cosh r\tilde{\epsilon}), \\
\mathcal{D}_\pi(\tilde{\epsilon}) &= \tilde{\mathcal{E}}(\tilde{\epsilon}) - 2e^{-r\tilde{\epsilon}^2/2} \sinh r\tilde{\epsilon}, \\
\mathcal{D}_q(\tilde{\epsilon}) &= r\tilde{\mathcal{E}}(\tilde{\epsilon}) - 2e^{-r\tilde{\epsilon}^2/2} [(r-1) \sinh r\tilde{\epsilon} + r\tilde{\epsilon} \cosh r\tilde{\epsilon}], \\
\tilde{\mathcal{E}}(\tilde{\epsilon}) &= \sqrt{\frac{\pi r}{2}} e^{r/2} \left[\operatorname{Erf} \left(\sqrt{\frac{r}{2}} (1+\tilde{\epsilon}) \right) - \operatorname{Erf} \left(\sqrt{\frac{r}{2}} (1-\tilde{\epsilon}) \right) \right] \\
&= \sqrt{2r} e^{r/2} \int_{\sqrt{\frac{r}{2}}(1-\tilde{\epsilon})}^{\sqrt{\frac{r}{2}}(1+\tilde{\epsilon})} e^{-w^2} dw,
\end{aligned}$$

and $\operatorname{Erf}(\cdot)$ is the error function. The zero in the argument of η_π , ε_π , and ε_q in (1.46–1.47) is due to our choice $\tilde{\epsilon}_m = \sqrt{2\epsilon_m}/U = 0$ as the lowest end of the dimensionless interval $[0, \tilde{\epsilon}_M]$.

The terms in (1.46–1.47) do not vanish, hence the pressure tensor is not isotropic, and the heat flux does not vanish either. However, the two spurious terms thus introduced in the Euler equation go exponentially to zero, as $\tilde{\epsilon}_M$ grows. The introduction of a lower bound, say $\tilde{\epsilon}_m > 0$, gives further contributions to the spurious terms which go to zero as $\tilde{\epsilon}_m^{3/2}$, when $\tilde{\epsilon}_m$ tends to zero. In this case η_π , ε_π and ε_q can be written as

$$\begin{aligned}
\eta_\pi(\tilde{\epsilon}_m, \tilde{\epsilon}_M) &= \tag{1.48} \\
&= \frac{2\{\mathcal{N}_\pi(\tilde{\epsilon}_M) + \mathcal{N}_\pi(\tilde{\epsilon}_m) - [\tilde{\mathcal{E}}(\tilde{\epsilon}_M) - \tilde{\mathcal{E}}(\tilde{\epsilon}_m)][\mathcal{M}_\pi(\tilde{\epsilon}_M) + \mathcal{M}_\pi(\tilde{\epsilon}_m)] + \mathcal{N}'_\pi(\tilde{\epsilon}_m, \tilde{\epsilon}_M)\}}{[\mathcal{D}_\pi(\tilde{\epsilon}_M) - \mathcal{D}_\pi(\tilde{\epsilon}_m)][\mathcal{D}_q(\tilde{\epsilon}_M) - \mathcal{D}_q(\tilde{\epsilon}_m)]}
\end{aligned}$$

$$\begin{aligned}
\varepsilon_\pi(\tilde{\epsilon}_m, \tilde{\epsilon}_M) &= \tag{1.49} \\
&= \frac{2\{\mathcal{N}_\pi(\tilde{\epsilon}_M) + \mathcal{N}_\pi(\tilde{\epsilon}_m) - [\tilde{\mathcal{E}}(\tilde{\epsilon}_M) - \tilde{\mathcal{E}}(\tilde{\epsilon}_m)][\mathcal{M}_\pi(\tilde{\epsilon}_M) + \mathcal{M}_\pi(\tilde{\epsilon}_m)] + \mathcal{N}'_\pi(\tilde{\epsilon}_m, \tilde{\epsilon}_M)\}}{[\mathcal{D}_q(\tilde{\epsilon}_M) - \mathcal{D}_q(\tilde{\epsilon}_m)]^2}
\end{aligned}$$

$$\begin{aligned}
\varepsilon_q(\tilde{\epsilon}_m, \tilde{\epsilon}_M) &= \tag{1.50} \\
&= \frac{2\{\mathcal{N}_q(\tilde{\epsilon}_M) + \mathcal{N}_q(\tilde{\epsilon}_m) - [\tilde{\mathcal{E}}(\tilde{\epsilon}_M) - \tilde{\mathcal{E}}(\tilde{\epsilon}_m)][\mathcal{M}_q(\tilde{\epsilon}_M) + \mathcal{M}_q(\tilde{\epsilon}_m)] + \mathcal{N}'_q(\tilde{\epsilon}_m, \tilde{\epsilon}_M)\}}{[\mathcal{D}_q(\tilde{\epsilon}_M) - \mathcal{D}_q(\tilde{\epsilon}_m)]^2}
\end{aligned}$$

where

$$\begin{aligned}
\mathcal{N}'_\pi(\tilde{\epsilon}_m, \tilde{\epsilon}_M) &= e^{-r(\tilde{\epsilon}_m^2 + \tilde{\epsilon}_M^2)/2} \{-4(r+2) \sinh r\tilde{\epsilon}_m \sinh r\tilde{\epsilon}_M \\
&\quad + (r+1)r[(\tilde{\epsilon}_m + \tilde{\epsilon}_M) \sinh r(\tilde{\epsilon}_m + \tilde{\epsilon}_M) - (\tilde{\epsilon}_M - \tilde{\epsilon}_m) \sinh r(\tilde{\epsilon}_M - \tilde{\epsilon}_m)] \\
&\quad + r^2[(\tilde{\epsilon}_m + \tilde{\epsilon}_M)^2 \cosh r(\tilde{\epsilon}_M - \tilde{\epsilon}_m) - (\tilde{\epsilon}_M - \tilde{\epsilon}_m)^2 \cosh r(\tilde{\epsilon}_M + \tilde{\epsilon}_m)]\},
\end{aligned}$$

and

$$\begin{aligned} \mathcal{N}'_q(\tilde{e}_m, \tilde{e}_M) = & e^{-r(\tilde{e}_m^2 + \tilde{e}_M^2)/2} \times \{-4(r^2 + 4r - 2) \sinh r\tilde{e}_m \sinh r\tilde{e}_M \\ & + 4(r + 2)r^2\tilde{e}_m\tilde{e}_M \cosh r\tilde{e}_m \cosh r\tilde{e}_M + (r^2 + 3r - 4)r \\ & \times [(\tilde{e}_m + \tilde{e}_M) \sinh r(\tilde{e}_m + \tilde{e}_M) - (\tilde{e}_M - \tilde{e}_m) \sinh r(\tilde{e}_M - \tilde{e}_m)] \\ & + r^3[(\tilde{e}_m + \tilde{e}_M)^2 \cosh r(\tilde{e}_M - \tilde{e}_m) - (\tilde{e}_M - \tilde{e}_m)^2 \cosh r(\tilde{e}_M + \tilde{e}_m)]\} . \end{aligned}$$

These quantities characterize the magnitude of the spurious terms introduced by the restriction of the velocity space to a sphere or a hollow sphere, and they provide us with a tool to control such terms. On these grounds, one may adopt the following procedure to select e_m and e_M :

1. Identify the range of internal energy and drift velocity characteristic of the flow to be described (and therefore the range of dimensionless parameter r);
2. Set a desired accuracy $\bar{\varepsilon}$ to limit the spurious terms ε_π (or η_π) and ε_q appearing in the Euler equations (1.34) or in the definition of heat flux and pressure tensor, and define \bar{e}_m and \bar{e}_M such that $\bar{\varepsilon}_b = \bar{\varepsilon}_M + \bar{e}_m < \bar{\varepsilon}$ (say, $\bar{e}_M = \bar{e}_m = \bar{\varepsilon}/3$);
3. Set \tilde{e}_M so that ε_π (or η_π) and ε_q in Eqs.(1.46–1.47) have absolute value smaller than $\bar{\varepsilon}_M$. For instance, for a given r Eq.(1.47) gives $\varepsilon_q = \varepsilon_q(0, \tilde{e}_M)$, which (locally) defines $\tilde{e}_M = \tilde{e}_M^q(0, \varepsilon_q)$. Therefore we can set

$$\tilde{e}_M = \bar{e}_M(\bar{\varepsilon}_M) \stackrel{def}{=} \max\{\tilde{e}_M^\pi(0, \bar{\varepsilon}_M), \tilde{e}_M^q(0, \bar{\varepsilon}_M)\} ; \quad (1.51)$$

4. Having set \tilde{e}_M according to (1.51), consider \tilde{e}_m so that ε_π (or η_π) and ε_q in Eqs.(1.48–1.50) have absolute value smaller than $\bar{\varepsilon}_b = \bar{\varepsilon}_m + \bar{\varepsilon}_M$. For instance, Eq.(1.50) determines the values of \tilde{e}_m which yield $\varepsilon_q(\tilde{e}_m, \bar{e}_M(\bar{\varepsilon}_M)) = \bar{\varepsilon}_m + \bar{\varepsilon}_M$, that is $\tilde{e}_m = \tilde{e}_m^q(\bar{\varepsilon}_m + \bar{\varepsilon}_M, \bar{\varepsilon}_M)$. Therefore, if computationally convenient, one can set

$$\tilde{e}_m = \bar{e}_m(\bar{\varepsilon}_m, \bar{\varepsilon}_M) \stackrel{def}{=} \min\{\tilde{e}_m^\pi(\bar{\varepsilon}_m + \bar{\varepsilon}_M, \bar{\varepsilon}_M), \tilde{e}_m^q(\bar{\varepsilon}_m + \bar{\varepsilon}_M, \bar{\varepsilon}_M)\} . \quad (1.52)$$

This procedure is used to plot Figure 3 and to give Tables 1 and 2. In particular, Figure 3 gives ε_π and ε_q as a function of \tilde{e}_M at different r , Table 1 gives $\bar{e}_M(\frac{\bar{\varepsilon}}{3})$ and Table 2 $\bar{e}_m(\frac{\bar{\varepsilon}}{3}, \frac{\bar{\varepsilon}}{3})$ at fixed values of r and overall accuracy $\bar{\varepsilon}$.

It should be noticed that for physical reasons e_m should be chosen smaller than $U^2/2$ and e_M larger than $U^2/2$, i.e. $\tilde{e}_m < 1$ and $\tilde{e}_M > 1$ (recall that $e_m = U^2\tilde{e}_m^2/2$ and $e_M = U^2\tilde{e}_M^2/2$). In Table 2 there are, however, some values of r and $\bar{\varepsilon}$ for which the contribution associated with any $\tilde{e}_m < 1$ is always less than $\bar{\varepsilon}/3$.

Furthermore, except for the very special cases where ε_q is close to changing sign (say, $|\varepsilon_q| < 10^{-6}$), ε_π is always an order of magnitude smaller than ε_q . Therefore, \bar{e}_m and \bar{e}_M are essentially determined by \tilde{e}_m^q and \tilde{e}_M^q . As a consequence, if the magnitude of the spurious term in the energy equation of (1.34) is not a concern,

	$\bar{\varepsilon} = 10^{-1}$ $\bar{\varepsilon}_M = \frac{1}{3}10^{-1}$	$\bar{\varepsilon} = 10^{-2}$ $\bar{\varepsilon}_M = \frac{1}{3}10^{-2}$	$\bar{\varepsilon} = 10^{-3}$ $\bar{\varepsilon}_M = \frac{1}{3}10^{-3}$	$\bar{\varepsilon} = 10^{-4}$ $\bar{\varepsilon}_M = \frac{1}{3}10^{-4}$
r=10	2.025	2.265	2.457	2.622
r=1	4.603	5.326	5.908	6.410
r=0.1	13.53	15.72	17.49	19.03

Table 1.1 Values of $\bar{\varepsilon}_M(r, \bar{\varepsilon}_M = \frac{\bar{\varepsilon}}{3})$ defined in (1.51) at different values of the expected macroscopic parameter $r = \frac{3\rho U^2}{2\bar{\varepsilon}}$, and overall accuracy $\bar{\varepsilon}$.

	$\bar{\varepsilon} = 10^{-1}$ $\bar{\varepsilon}_b = \frac{2}{3}10^{-1}$	$\bar{\varepsilon} = 10^{-2}$ $\bar{\varepsilon}_b = \frac{2}{3}10^{-2}$	$\bar{\varepsilon} = 10^{-3}$ $\bar{\varepsilon}_b = \frac{2}{3}10^{-3}$	$\bar{\varepsilon} = 10^{-4}$ $\bar{\varepsilon}_b = \frac{2}{3}10^{-4}$
r=10	0.406	0.208	0.103	0.0482
r=1	1.136	0.230	0.104	0.0466
r=0.1	2.606	0.684	0.303	0.136

Table 1.2 Values of $\bar{\varepsilon}_m(r, \bar{\varepsilon}_m = \frac{\bar{\varepsilon}}{3}, \bar{\varepsilon}_M = \frac{\bar{\varepsilon}}{3})$ defined in (1.52) at different values of the expected macroscopic parameter $r = \frac{3\rho U^2}{2\bar{\varepsilon}}$, and of the overall accuracy $\bar{\varepsilon}(\bar{\varepsilon}_b = \bar{\varepsilon}_m + \bar{\varepsilon}_M = \frac{2}{3}\bar{\varepsilon})$ and with $\bar{\varepsilon}_M$ given in Table 1. Values above 1 correspond to those cases in which the contribution $\bar{\varepsilon}_b$ deriving by introducing any non vanishing physically admissible $\bar{\varepsilon}_m < 1$ is always smaller than required.

and one only wants the magnitude of the spurious term in the momentum equation to be smaller than a certain $\alpha > 0$, even larger values of $\bar{\varepsilon}_m$ and smaller values of $\bar{\varepsilon}_M$ can be used. Roughly, one can use the values of $\bar{\varepsilon}_m$ and $\bar{\varepsilon}_M$ reported in Tables 1 and 2, which correspond to 10α instead of α . For instance, the choice $r = 10$, $\bar{\varepsilon}_M = 2.265$ gives rise to a spurious term in the momentum equation of order 10^{-3} instead of 10^{-2} .

1.4 Energy Discretization and Kinetic Model

Once the interval $\mathbb{E} = [e_m, e_M)$ has been identified, so that the probability for a particle to have energy outside \mathbb{E} is negligible, we adopt a discretization procedure which divides \mathbb{E} in $n + 1$ subintervals \mathbb{I}_i of equal width Δ :

$$\mathbb{I}_i = [e_m + i\Delta, e_m + (i + 1)\Delta) \quad , \quad i = 0, \dots, n \quad , \quad (1.53)$$

with

$$\Delta = \frac{e_M - e_m}{n + 1} \quad . \quad (1.54)$$

Denoting the mid-point energy of the interval \mathbb{I}

$$e_i = e_m + \left(i + \frac{1}{2}\right) \Delta \quad . \quad (1.55)$$

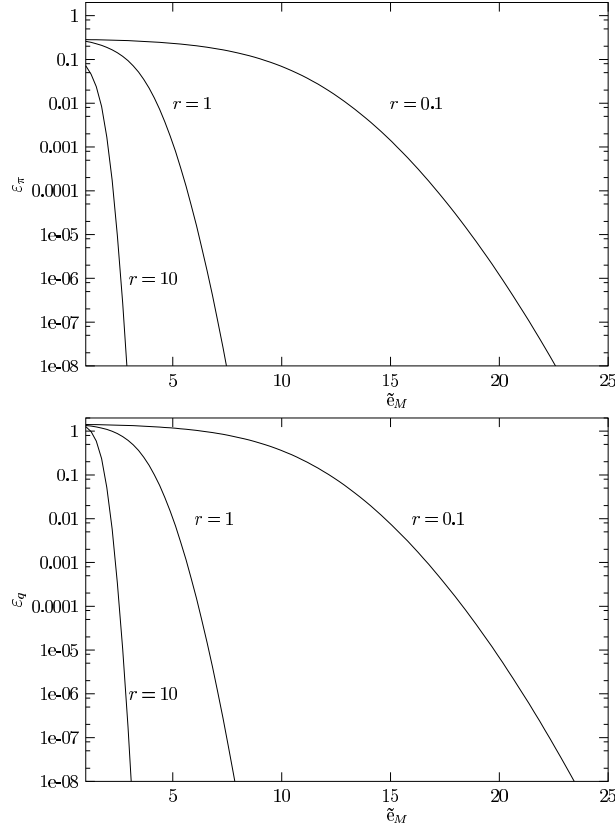


Fig. 1.3 Plots of ε_π and of ε_q as a function of $\tilde{\varepsilon}_M$ at different values of the expected macroscopic parameter $r = \frac{3\rho U^2}{2\varepsilon}$. The two plots give an indication of the value of the parameter $\tilde{\varepsilon}_M$ to be chosen, in the discretization, so that the induced magnitude of the spurious terms is small enough.

it is trivial to check that the following holds:

Proposition 1.1 *The discretization has the following properties*

- $\forall i, j, h$ s.t. $e_i + e_j - e_h \in \mathbb{E}$, $\exists k = i + j - h \in [0, n]$ s.t. $e_i + e_j - e_h = e_k$;
- If $e_m = \ell\Delta$ $\ell \in \mathbb{N}$, then $\forall i, j$ s.t. $e_i + e_j \leq e_m$, $e_i + e_j = \sup \mathbb{I}_{i+j+\ell}$.

It need to be mentioned that the condition $e_m = \ell\Delta$ does not represent a real restriction on the choice of \mathbb{E} , since \mathbb{E} can be easily adjusted to satisfy this condition.

Remark 1.1 The proposition assures that:

- a) For any couple of input energies belonging to the discretization, if one of the output energies also belongs to the discretization, then either the other output energy belongs to the discretization, or it is outside \mathbb{E} , but, as already stated, the occurrence of this last case has negligible probability.

b) If $e_m = \ell\Delta$, then either $[e_m, e_i + e_j)$ contains \mathbb{E} , or it is exactly equal to the union of $i + j + \ell + 1$ subintervals

$$[e_m, e_i + e_j) = \bigcup_{h=0}^{i+j+\ell} \mathbb{I}_h . \quad (1.56)$$

This is useful to properly handle the collision operator.

The following step consists in considering any function of velocity as a function of energy and velocity direction, and in approximating its energy dependence in \mathbb{E} by a stepwise interpolation defined on the partition \mathbb{I}_i . This means that we take

$$F(\mathbf{v}) = F(e, \hat{\Omega}) \approx \sum_{i=0}^n \chi_i(e) F_i(\hat{\Omega}) , \quad (1.57)$$

where $F_i(\hat{\Omega}) = F(e_i, \hat{\Omega})$ and $\chi_i(e)$ is the characteristic function of the interval \mathbb{I}_i . It is useful to note that

$$\int_{\mathbb{I}_i} \chi_j(e) de = \Delta \delta_{ij} , \quad (1.58)$$

where δ_{ij} is the Kronecker delta. Hence, one can approximate the integral of a function defined in \mathbb{E} as

$$\int_{\mathbb{E}} F(e) de \approx \int_{\mathbb{E}} \sum_{j=0}^n \chi_j(e) F_j de = \Delta \sum_{j=0}^n F_j , \quad (1.59)$$

with

$$\int_{\mathbb{I}_j} F(e) de = \Delta F_j . \quad (1.60)$$

The quadrature rule (1.59) is nothing else but the mid-point rule which gives an error that goes to zero as Δ^2 .

If we approximate the energy dependence of the Boltzmann one-particle distribution function using the stepwise interpolation, we can write (1.42) as

$$\int_{\mathbb{I}_i} J\sqrt{2e} de \approx \sum_{h,j=0}^n \int_{\mathbb{I}_i} \chi_h(e) de \int_{\mathbb{R}_+} \chi_j(e_*) de_* \int_0^{e_h+e_j} \hat{J}(e_h, e_j, e', \hat{\Omega}) de' , \quad (1.61)$$

where we have chosen $\mathbb{I} = \mathbb{I}_i$. Using (1.58, 1.59) we have

$$\begin{aligned} \int_{\mathbb{I}_i} J\sqrt{2e} de &\approx \sum_{h,j=0}^n \int_{\mathbb{I}_i} \chi_h(e) de \int_{\mathbb{R}_+} \chi_j(e_*) de_* \int_{e_m}^{e_h+e_j} \hat{J}(e_h, e_j, e', \hat{\Omega}) de' \\ &= \Delta^2 \sum_{j=0}^n \int_{e_m}^{e_+} \hat{J}(e_i, e_j, e', \hat{\Omega}) de' , \end{aligned} \quad (1.62)$$

where $e_+^h = \min\{e_h + e_j, e_M\}$ and $e_+ = \min\{e_i + e_j, e_M\}$.

Recalling Remark 1.1, and because of Proposition 1.1, the interval of integration over e' either results equal to \mathbb{E} , or is the union of $i + j + \ell + 1$ subintervals of the partition, that is

$$e_+ = \sup \mathbb{I}_{h_+} , \quad h_+ = \min \{i + j + \ell, n\} . \quad (1.63)$$

Discretizing also e' in the same fashion, gives

$$\begin{aligned} \int_{\mathbf{I}_i} J\sqrt{2e} \, de &\approx \Delta^2 \sum_{j=0}^n \int_{e_m}^{e_+} \sum_{h=0}^n \chi_h(e') \, de' \hat{J}(e_i, e_j, e_h, \hat{\Omega}) \\ &= \Delta^3 \sum_{j=0}^n \sum_{h=0}^{h_+} \hat{J}(e_i, e_j, e_h, \hat{\Omega}) . \end{aligned} \quad (1.64)$$

For consistency, one has to neglect the collisions with $e_k = e_i + e_j - e_h$ outside \mathbb{E} . This allows us to write the sum over the index h related to an output particle in a more symmetrical way

$$\int_{\mathbf{I}_i} J\sqrt{2e} \, de \approx \Delta^3 \sum_{j=0}^n \sum_{\substack{h,k=0 \\ h+k=i+j}}^n \hat{J}(e_i, e_j, e_h, \hat{\Omega}) . \quad (1.65)$$

Using the same procedure on the left hand side and equating gives

$$\begin{aligned} \frac{\partial f_i}{\partial t} &= 4\Delta^2 \sum_{j=0}^n \sqrt{2e_j} \\ &\times \sum_{\substack{h,k=0 \\ h+k=i+j}}^n \int_0^{2\pi} d\vartheta \int_{S_{\hat{\Omega}}(e_i, e_j, e_h)} d\hat{\Omega}_* \frac{\sigma_{ij}^{hk}(\hat{\Omega} \cdot \hat{\Omega}_*, \vartheta)}{R_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*)} (f'_h f'_{*k} - f_i f_{*j}) , \end{aligned} \quad (1.66)$$

where

$$\begin{aligned} f_i &= f_i(\hat{\Omega}) = f(\sqrt{2e_i}\hat{\Omega}) , \\ f_{*j} &= f_j(\hat{\Omega}_*) = f(\sqrt{2e_j}\hat{\Omega}_*) , \\ f'_h &= f_h(\hat{\Omega}'_{ij}{}^{hk}(\hat{\Omega}, \hat{\Omega}_*, \vartheta)) = f(\sqrt{2e_h}\hat{\Omega}'(e_i, e_j, e_h, \hat{\Omega}, \hat{\Omega}_*, \vartheta)) , \\ f'_{*k} &= f_k(\hat{\Omega}'_{*ij}{}^{hk}(\hat{\Omega}, \hat{\Omega}_*, \vartheta)) = f(\sqrt{2e_k}\hat{\Omega}'_*(e_i, e_j, e_h, \hat{\Omega}, \hat{\Omega}_*, \vartheta)) , \end{aligned}$$

and where σ_{ij}^{hk} , R_{ij} , $\hat{\Omega}'_{ij}{}^{hk}$ and $\hat{\Omega}'_{*ij}{}^{hk}$ are respectively the differential cross section (see also (1.88)), the collision momentum, and the post-collisional directions, which will be given in (1.87) for $e = e_i$, $e_* = e_j$, $e' = e_h$ and $e'_* = e_k$.

Remark 1.2 The integrand in (1.66) presents a denominator which vanishes in the case of head-on collisions, $e_i = e_j$, $\hat{\Omega} = -\hat{\Omega}_*$, i.e. for $R = 0$. Thus, we have to

check whether the integrand has a singularity, in that circumstance. Referring to Fig. 2 and choosing $\hat{\Omega}$ as polar axis in the integration over $S_{\hat{\Omega}}$, one has that

$$\frac{\sqrt{2e_i} d\hat{\Omega}_*}{R_{ii}(\hat{\Omega} \cdot \hat{\Omega}_*)} = \frac{\sin \beta d\beta d\Gamma}{\sqrt{2(1 + \hat{\Omega} \cdot \hat{\Omega}_*)}} = \frac{\sin \frac{\beta}{2} \cos \frac{\beta}{2}}{\sqrt{\frac{1 + \cos \beta}{2}}} d\beta d\Gamma = \sin \frac{\beta}{2} d\beta d\Gamma ,$$

where Γ is the longitude and β the colatitude ($\cos \beta = \hat{\Omega} \cdot \hat{\Omega}_*$), because $e_i = e_j$. Therefore there is no singularity as β goes to π .

1.5 Conservation and Euler Equations for the Discretized Model

Before deducing the conservation equations and the Euler equations for the discretized model (1.66) it is useful to introduce the following preliminary definitions

Definition 1.1 Given a regular function of energy and velocity direction $\phi(e, \hat{\Omega})$, its restriction to the sequence of energies e_i defines the vector function

$$\Phi(\hat{\Omega}) = (\Phi_0(\hat{\Omega}), \dots, \Phi_n(\hat{\Omega})) \quad \text{with} \quad \Phi_i(\hat{\Omega}) = \phi(e_i, \hat{\Omega}) , \quad i = 0, \dots, n .$$

We say that Φ is a collisional invariant if

$$\langle \Phi, \mathbf{J} \rangle \stackrel{def}{=} \sum_{i=0}^n \sqrt{2e_i} \int_{S_2} \Phi_i(\hat{\Omega}) J_i(\hat{\Omega}) d\hat{\Omega} = 0 , \quad \forall f_i \geq 0 , \quad (1.67)$$

where $\mathbf{J} = (J_0, \dots, J_n)$ is the collision operator defined by the right hand side of Eq.(1.66).

According to the discussion of Section 4, the following can be proved:

Theorem 1.2 *The following statements are equivalent:*

- a) Φ is a collisional invariant;
- b) $\Phi_i(\hat{\Omega}) = a_d + \sqrt{e_i} \mathbf{b}_d \cdot \hat{\Omega} - c_d e_i$, $i = 0, \dots, n$, $\hat{\Omega} \in S_2$ for all values of a_d , \mathbf{b}_d and c_d independent of i and $\hat{\Omega}$.

PROOF: The proof consists in proving the equivalence of both b) and a) to the following intermediate statement

$$\begin{aligned} \text{c) } \forall i, j, h, k = 0, \dots, n \quad & \text{and} \quad \forall \hat{\Omega}, \hat{\Omega}_*, \hat{\Omega}', \hat{\Omega}'_* \in S_2 \quad \text{such that} \\ e_i + e_j = e_h + e_k \quad & \text{and} \quad \sqrt{e_i} \hat{\Omega} + \sqrt{e_j} \hat{\Omega}_* = \sqrt{e_h} \hat{\Omega}' + \sqrt{e_k} \hat{\Omega}'_* \\ \implies \quad & \Phi_i(\hat{\Omega}) + \Phi_j(\hat{\Omega}_*) = \Phi_h(\hat{\Omega}') + \Phi_k(\hat{\Omega}'_*) \end{aligned}$$

a) \Leftrightarrow c) According to (1.66) Eq.(1.67) can be written as

$$\langle \Phi, \mathbf{J} \rangle = 8\Delta^2 \sum_{\substack{i,j,h,k=0 \\ h+k=i+j}}^n \int_0^{2\pi} d\vartheta \int_{S_2 \times S_{\hat{\Omega}}} d\hat{\Omega} d\hat{\Omega}_* \left[\sqrt{e_i e_j} \frac{\sigma_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*, \vartheta)}{R_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*)} (f'_h f'_{*k} - f_i f_{*j}) \right] \Phi_i(\hat{\Omega}),$$

where $S_{\hat{\Omega}} = S_{\hat{\Omega}}(e_i, e_j, e_h)$.

The integration domain $S_2 \times S_{\hat{\Omega}}$, the summation and the term in the square brackets are invariants with respect to the exchanges

$$\sqrt{e_i \hat{\Omega}} \longleftrightarrow \sqrt{e_j \hat{\Omega}_*}, \quad \sqrt{e_h \hat{\Omega}'} \longleftrightarrow \sqrt{e_k \hat{\Omega}'_*}.$$

Then, separating the gain from the loss contributions, i.e. writing

$$\langle \Phi, \mathbf{J} \rangle = \langle \Phi, \mathbf{G} \rangle - \langle \Phi, \mathbf{L} \rangle,$$

one obtains

$$\begin{aligned} \langle \Phi, \mathbf{L} \rangle = & 4\Delta^2 \sum_{\substack{i,j,h,k=0 \\ h+k=i+j}}^n \int_0^{2\pi} d\vartheta \int_{S_2 \times S_{\hat{\Omega}}} d\hat{\Omega} d\hat{\Omega}_* \\ & \times \left[\sqrt{e_i e_j} \frac{\sigma_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*, \vartheta)}{R_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*)} f_i(\hat{\Omega}) f_j(\hat{\Omega}_*) \right] [\Phi_i(\hat{\Omega}) + \Phi_j(\hat{\Omega}_*)], \end{aligned} \quad (1.68)$$

$$\begin{aligned} \langle \Phi, \mathbf{G} \rangle = & 4\Delta^2 \sum_{\substack{i,j,h,k=0 \\ h+k=i+j}}^n \int_0^{2\pi} d\vartheta \int_{S_2 \times S_{\hat{\Omega}}} d\hat{\Omega} d\hat{\Omega}_* \\ & \times \left[\sqrt{e_i e_j} \frac{\sigma_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*, \vartheta)}{R_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*)} f_h(\hat{\Omega}') f_k(\hat{\Omega}'_*) \right] [\Phi_i(\hat{\Omega}) + \Phi_j(\hat{\Omega}_*)], \end{aligned} \quad (1.69)$$

The second step is to make the substitutions

$$i \rightarrow h, \quad j \rightarrow k, \quad \hat{\Omega} \rightarrow \hat{\Omega}', \quad \text{and} \quad \hat{\Omega}_* \rightarrow \hat{\Omega}'_*,$$

in Eq.(1.69). Then, recalling that in [14] it was proved that for fixed incoming energies e_i and e_j , and outgoing energies e_h and e_k compatible with conservation of energy, and for a fixed angle ϑ between the pre- and post-collisional plane, the Jacobian of the transformation $(\hat{\Omega}', \hat{\Omega}'_*) \rightarrow (\hat{\Omega}, \hat{\Omega}_*)$ given by the conservation of momentum is

$$\frac{\partial (\hat{\Omega}', \hat{\Omega}'_*)}{\partial (\hat{\Omega}, \hat{\Omega}_*)} = \sqrt{\frac{e_i e_j}{e_h e_k}}.$$

This allows us to transform the integral over $(\hat{\Omega}', \hat{\Omega}'_*) \in S_2 \times S_{\hat{\Omega}}$ into an integral over $(\hat{\Omega}, \hat{\Omega}_*) \in S_2 \times S_{\hat{\Omega}}$, so that we can write

$$\begin{aligned} \langle \Phi, \mathbf{J} \rangle &= 4\Delta^2 \sum_{\substack{i,j,h,k=0 \\ h+k=i+j}}^n \int_0^{2\pi} d\vartheta \int_{S_2 \times S_{\hat{\Omega}}} d\hat{\Omega} d\hat{\Omega}_* \\ &\times \left[\sqrt{\epsilon_i \epsilon_j} \frac{\sigma_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*, \vartheta)}{R_{ij}(\hat{\Omega} \cdot \hat{\Omega}_*)} f_i(\hat{\Omega}) f_j(\hat{\Omega}_*) \right] \left[\Phi_h(\hat{\Omega}') + \Phi_k(\hat{\Omega}'_*) - \Phi_i(\hat{\Omega}) - \Phi_j(\hat{\Omega}_*) \right] \end{aligned}$$

Because the σ_{ij}^{hk} are positive in the domain of integration for all admissible collisions, we obtain:

$$\langle \Phi, \mathbf{J} \rangle = 0 \quad \iff \quad \Phi_h(\hat{\Omega}') + \Phi_k(\hat{\Omega}'_*) - \Phi_i(\hat{\Omega}) - \Phi_j(\hat{\Omega}_*) = 0 ,$$

which is the desired equivalence.

c) \iff b) We denote respectively by \mathcal{M}_b and \mathcal{M}_c the linear spaces of the vector functions $\tilde{\Phi}(\hat{\Omega}) = (\Phi_0(\hat{\Omega}), \dots, \Phi_n(\hat{\Omega}))$ verifying the conditions b) and c).

It is evident from the definitions of b) and c) that $\mathcal{M}_b \subseteq \mathcal{M}_c$, that is the model preserves mass, momentum and energy. We need to prove that $\mathcal{M}_c \subseteq \mathcal{M}_b$ meaning that there are no spurious invariants.

In order to do that we express \mathcal{M}_c (and consequently \mathcal{M}_b) as the direct sum $\mathcal{M}_c = \mathcal{M}_c^* \oplus \mathcal{M}'_c$, so that:

$$\tilde{\Phi}(\hat{\Omega}) \in \mathcal{M}_c \implies \Phi = \Phi^* + \Psi \quad \text{where } \Phi^* \in \mathcal{M}_c^*, \Psi \in \mathcal{M}'_c ,$$

with

$$\Phi_i^* = \frac{1}{4\pi} \int_{S_2} \Phi_i(\hat{\Omega}) d\hat{\Omega} , \quad (1.70)$$

$$\int_{S_2} \Psi_i(\hat{\Omega}) d\hat{\Omega} = 0 , \quad (1.71)$$

for all $i = 0, \dots, n$.

Consequently, one has $\mathcal{M}_b = \mathcal{M}_b^* \oplus \mathcal{M}'_b$ where

$$\begin{aligned} \Phi^* \in \mathcal{M}_b^* &\iff \Phi_i^* = a_d - c_d \epsilon_i \quad i = 0, \dots, n , \\ \Psi(\hat{\Omega}) \in \mathcal{M}'_b &\iff \Psi_i^* = \sqrt{\epsilon_i} \hat{\Omega} \cdot \mathbf{b}_d = \sqrt{\epsilon_i} \mathcal{L}(\hat{\Omega}) \quad i = 0, \dots, n . \end{aligned}$$

The proof of the implication $\mathcal{M}'_c \subseteq \mathcal{M}'_b$ is similar to the proof concerning the planar semicontinuous model given in [11] (pag.119-120) for two velocity moduli only, and extended in [12] (pag. 74-76) to any number of velocity moduli. Following the line of argument given there one can prove that for any $\Psi(\hat{\Omega})$ verifying (1.71) and the following condition

$$\begin{aligned}
 c') \quad \forall i, j, h, k &= 0, \dots, n \quad \text{and} \quad \forall \hat{\Omega}, \hat{\Omega}_*, \hat{\Omega}', \hat{\Omega}'_* \in S_2 \cap \Pi \quad \text{s.t.} \\
 e_i + e_j &= e_h + e_k \quad \text{and} \quad \sqrt{\epsilon_i} \hat{\Omega} + \sqrt{\epsilon_j} \hat{\Omega}_* = \sqrt{\epsilon_h} \hat{\Omega}' + \sqrt{\epsilon_k} \hat{\Omega}'_* \\
 \implies \Phi_i(\hat{\Omega}) + \Phi_j(\hat{\Omega}_*) &= \Phi_h(\hat{\Omega}') + \Phi_k(\hat{\Omega}'_*)
 \end{aligned}$$

where Π is any collisional plane one has

$$\Psi_i^* = \sqrt{\epsilon_i} \mathcal{L}(\hat{\Omega}) \quad i = 0, \dots, n . \quad (1.72)$$

This means that the linear space \mathcal{M}'_c defined by $c')$ is such that $\mathcal{M}'_c \subseteq \mathcal{M}'_b$. Keeping into account the arbitrariness of the collisional plane Π and that $c) \Rightarrow c')$ (that is, $\mathcal{M}'_c \subseteq \mathcal{M}'_b$) it necessarily follows that $\mathcal{M}'_c \subseteq \mathcal{M}'_b$.

It remains to verify that $\mathcal{M}_c^* \subseteq \mathcal{M}_b^*$. Keeping into account that the elements of \mathcal{M}_c^* are independent of $\hat{\Omega}$, from $c)$ it follows that

$$\Phi^* \in \mathcal{M}_c^* \implies \Phi_i^* + \Phi_j^* = \Phi_h^* + \Phi_k^* \quad \text{where} \quad i + j = h + k . \quad (1.73)$$

Equation (1.73) yields $n - 1$ linearly independent relations between the $n + 1$ components of Φ^* . There exist then $a_d, c_d \in \mathbb{R}$ such that

$$\forall i = 0, \dots, n \quad \Phi_i^* = a_d - c_d e_i ,$$

that is, $\mathcal{M}_c^* \subseteq \mathcal{M}_b^*$, which ends the proof of the equivalence of b) and c). ♣

Remark 1.3 This theorem not only guarantees the conservation of mass, momentum and energy for the discretized model (1.66), but also excludes the existence of spurious collisional invariants with no physical meaning.

The space \mathcal{M} of collisional invariants is then spanned by the basis

$$\left\{ \begin{array}{l} \Psi^{(0)}(\hat{\Omega}) = (1, \dots, 1) \\ \Psi^{(1)}(\hat{\Omega}) = (\sqrt{\epsilon_0}, \dots, \sqrt{\epsilon_n}) \hat{\Omega} \cdot \hat{e}_x \\ \Psi^{(2)}(\hat{\Omega}) = (\sqrt{\epsilon_0}, \dots, \sqrt{\epsilon_n}) \hat{\Omega} \cdot \hat{e}_y \\ \Psi^{(3)}(\hat{\Omega}) = (\sqrt{\epsilon_0}, \dots, \sqrt{\epsilon_n}) \hat{\Omega} \cdot \hat{e}_z \\ \Psi^{(4)}(\hat{\Omega}) = (\epsilon_0, \dots, \epsilon_n) \end{array} \right\} \begin{array}{l} \text{corresponding to conservation of mass} \\ \text{corresponding to conservation of momentum} \\ \text{corresponding to conservation of energy} \end{array}$$

The following theorem can then be proved following classical methods

Theorem 1.3 *The kinetic model (1.66) has the following properties:*

- It preserves mass, momentum and energy, defined by

$$\begin{aligned}\rho_d &= m\Delta \sum_{i=0}^n \sqrt{2e_i} \int_{S_2} f_i(\hat{\Omega}) d\hat{\Omega} , \\ \mathbf{U}_d &= \frac{2m\Delta}{\rho_d} \sum_{i=0}^n e_i \int_{S_2} \hat{\Omega} f_i(\hat{\Omega}) d\hat{\Omega} , \\ \mathcal{E}_d &= \frac{m\Delta}{2} \sum_{i=0}^n \sqrt{2e_i} \int_{S_2} |\sqrt{2e_i}\hat{\Omega} - \mathbf{U}|^2 f_i(\hat{\Omega}) d\hat{\Omega} ,\end{aligned}\tag{1.74}$$

- In the non-homogeneous case it yields the usual conservation equations

$$\left\{ \begin{aligned}\frac{\partial \rho_d}{\partial t} + \nabla \cdot (\rho_d \mathbf{U}_d) &= 0 , \\ \frac{\partial}{\partial t} (\rho_d \mathbf{U}_d) + \nabla \cdot (\mathbf{\Pi}_d + \rho_d \mathbf{U}_d \otimes \mathbf{U}_d) &= \mathbf{0} , \\ \frac{\partial}{\partial t} \left(\mathcal{E}_d + \frac{1}{2} \rho_d U_d^2 \right) + \nabla \cdot \left[\mathbf{q}_d + \mathbf{\Pi}_d \mathbf{U}_d + \left(\mathcal{E}_d + \frac{1}{2} \rho_d U_d^2 \right) \mathbf{U}_d \right] &= 0 ,\end{aligned}\right.\tag{1.75}$$

where

$$\begin{aligned}\mathbf{\Pi}_d &= m\Delta \sum_{i=0}^n \sqrt{2e_i} \int_{S_2} (\sqrt{2e_i}\hat{\Omega} - \mathbf{U}) \otimes (\sqrt{2e_i}\hat{\Omega} - \mathbf{U}) f_i(\hat{\Omega}) d\hat{\Omega} , \\ \mathbf{q}_d &= \frac{m\Delta}{2} \sum_{j=0}^n \sqrt{2e_j} \int_{S_2} |\sqrt{2e_j}\hat{\Omega} - \mathbf{U}|^2 (\sqrt{2e_j}\hat{\Omega} - \mathbf{U}) f_j(\hat{\Omega}) d\hat{\Omega} ;\end{aligned}\tag{1.76}$$

- The Maxwellian equilibrium state is defined by the following equivalent conditions

- $(\log f_0, \dots, \log f_n)$ belongs to the space of collisional invariants;
- $f_i = A_d \exp[\sqrt{B_d e_i} \hat{\mathbf{B}}_d \cdot \hat{\Omega} - C_d e_i]$, $A_d > 0$;
- $\mathbf{J}[\mathbf{f}, \mathbf{f}] = \mathbf{0}$ where $\mathbf{f} = (f_0, \dots, f_n)$;

- In the spatially homogeneous case the Boltzmann H-functional

$$H = \sum_{i=0}^n \sqrt{2e_i} \int_{S_2} f_i(\hat{\Omega}) \log f_i(\hat{\Omega}) d\hat{\Omega} ,\tag{1.77}$$

is such that $\frac{dH}{dt} \leq 0$ where the equality sign holds if and only if the system is in the Maxwellian equilibrium state.

It can be noticed that the macroscopic quantities defined in (1.74) and (1.76) related to the discretization are the approximations, through our approximation procedure, of the corresponding quantities obtained from the continuous Boltzmann equation.

The relation between the Maxwellian parameters and the definition of the macroscopic observables ρ_d , \mathbf{U}_d and \mathcal{E}_d can be obtained substituting in Eq.(1.74) the Maxwellian distribution densities. This leads to:

$$\begin{aligned}\rho_d &= 4\pi m A_d \Delta \sqrt{\frac{2}{B_d}} \sum_{i=0}^n e^{-C_d e_i} \sinh \sqrt{B_d e_i}, \\ \rho_d U_d &= \frac{8\pi m A_d \Delta}{B_d} \sum_{i=0}^n e^{-C_d e_i} \left(\sqrt{B_d e_i} \cosh \sqrt{B_d e_i} - \sinh \sqrt{B_d e_i} \right), \\ \hat{\mathbf{U}}_d &= \hat{\mathbf{B}}_d, \\ \mathcal{E}_d + \frac{1}{2} \rho_d U_d^2 &= 4\pi m A_d \Delta \sqrt{\frac{2}{B_d}} \sum_{i=0}^n e_i e^{-C_d e_i} \sinh \sqrt{B_d e_i},\end{aligned}\quad (1.78)$$

where the hats denote unit vectors. As for discrete models, the map

$$F : (\rho_d, U_d, \hat{\mathbf{U}}_d, \mathcal{E}_d) \longrightarrow (A_d, B_d, \hat{\mathbf{B}}_d, C_d)$$

represented by (1.78) is one-to-one, even though, differently from the continuous case, the inversion cannot be in general performed in terms of elementary functions. Here the subscript d stresses precisely the fact that the macroscopic observables are computed from the discretized model, and that the Maxwellian parameters are not those of Eq.(1.25) but are related to such a discretized kinetic model.

In particular, given B_d and C_d , the quantities U_d and \mathcal{E}_d/ρ_d are expressed by

$$\begin{aligned}U_d &= \sqrt{\frac{2}{B_d} \frac{\sum_{i=0}^n e^{-C_d e_i} (\sqrt{B_d e_i} \cosh \sqrt{B_d e_i} - \sinh \sqrt{B_d e_i})}{\sum_{i=0}^n e^{-C_d e_i} \sinh \sqrt{B_d e_i}}}, \\ \frac{2\mathcal{E}_d}{\rho_d} + U_d^2 &= \frac{2 \sum_{i=0}^n e_i e^{-C_d e_i} \sinh \sqrt{B_d e_i}}{\sum_{i=0}^n e^{-C_d e_i} \sinh \sqrt{B_d e_i}}.\end{aligned}\quad (1.79)$$

If the dimensionless values $\tilde{B}_d = \sqrt{\frac{B_d}{2}} U_d$, and $\tilde{C}_d = C_d U_d^2$ are introduced, for any given r they are determined by the system

$$\begin{cases} \sum_{i=0}^n e^{-\tilde{C}_d \tilde{e}_i/2} [\tilde{B}_d \tilde{e}_i \cosh \tilde{B}_d \tilde{e}_i - (\tilde{B}_d + 1) \sinh \tilde{B}_d \tilde{e}_i] = 0 \\ \sum_{i=0}^n e^{-\tilde{C}_d \tilde{e}_i/2} \left(\tilde{e}_i - 1 - \frac{3}{r} \right) \sinh \tilde{B}_d \tilde{e}_i = 0 \end{cases}\quad (1.80)$$

However, for not too small values of n , we have $\tilde{B}_d \approx \tilde{C}_d \approx r$, as is the case for continuous models. Moreover, using the values obtained from (1.80), or setting $\tilde{B}_d = \tilde{C}_d = r$, makes a very little difference -or no difference at all- in the identification of the discretization parameter n through the procedure described below.

If one writes the Euler equations related to the kinetic model (1.66), equations similar to (1.34) are found, where the macroscopic observables are replaced by their discretized version, and where ε_π and ε_q are expressed by

$$\varepsilon_\pi^d = \frac{\sum_{i,j=0}^n e^{-C_d(\varepsilon_i+\varepsilon_j)} [(B_d e_i + 2) s_i s_j - B_d \sqrt{\varepsilon_i \varepsilon_j} c_i c_j - \sqrt{B_d \varepsilon_i} c_i s_j]}{\left[\sum_{i=0}^n e^{-C_d \varepsilon_i} (\sqrt{B_d \varepsilon_i} c_i - s_i) \right]^2}, \quad (1.81)$$

$$\varepsilon_q^d = \frac{\sum_{i,j=0}^n e^{-C_d(\varepsilon_i+\varepsilon_j)} (\varepsilon_i - \varepsilon_j) \sqrt{\varepsilon_i} c_i s_j}{2 \left[\sum_{i=0}^n e^{-C_d \varepsilon_i} (\sqrt{B_d \varepsilon_i} c_i - s_i) \right]^2} B_d \sqrt{B_d} - 1, \quad (1.82)$$

where $c_i = \cosh \sqrt{B_d \varepsilon_i}$ and $s_i = \sinh \sqrt{B_d \varepsilon_i}$.

With respect to the spurious terms ε_π and ε_q encountered in Section 3, when limiting the range of allowed energies, ε_π^d and ε_q^d present a new contribution related to the discretization of the interval E in $n + 1$ subintervals and to the introduction of the mid-point rule which brings an error of order Δ^2 . Actually, the constant of proportionality could be explicitly written down, but it is useless, since what matters is the overall magnitude of the spurious terms, which is expressed by (1.81) and (1.82), rather than some information on the last discretization step.

Once $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_m$ have been chosen, together with the desired accuracy $\bar{\varepsilon}_n$ for this discretization, and the discretization interval has been identified by the values $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_m$ given by (1.51) and (1.52), respectively, one can set n (which will be of order $1/\sqrt{\bar{\varepsilon}_n}$) so that the overall discretization terms defined in Eqs.(1.81) and (1.82) are both smaller than $\bar{\varepsilon} = \bar{\varepsilon}_M + \bar{\varepsilon}_m + \bar{\varepsilon}_n$, that is

$$\varepsilon^d(\tilde{\varepsilon}_m, \tilde{\varepsilon}_M, n) \stackrel{def}{=} \max\{|\varepsilon_\pi^d|, |\varepsilon_q^d|\} < \bar{\varepsilon}.$$

The results of such a procedure are reported in Table 3 where the value of n determined for different values of r and accuracy $\bar{\varepsilon}$ with $\bar{\varepsilon}_M = \bar{\varepsilon}_m = \bar{\varepsilon}_n = \frac{\bar{\varepsilon}}{3}$ is reported.

The number of discrete energies needed to keep the spurious terms below a moderate magnitude is quite small, but rapidly grows if greater accuracy is required. In Table 4, the same thing is repeated with the values of $\tilde{\varepsilon}_M$ given in Table 1, and

	$\bar{\varepsilon} = 10^{-1}$	$\bar{\varepsilon} = 10^{-2}$	$\bar{\varepsilon} = 10^{-3}$	$\bar{\varepsilon} = 10^{-4}$
r=10	2	14	52	238
r=1	7	10	250	1451
r=0.1	5	9	263	1531

Table 1.3 Values of n at different values of r , such that the spurious terms defined in (1.81) and (1.82) are definitely smaller than the given accuracy $\bar{\varepsilon}$. The values of $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_m$ are those obtained in Tables 1 and 2, corresponding to $\bar{\varepsilon}_M = \bar{\varepsilon}_m = \bar{\varepsilon}_n = \frac{\bar{\varepsilon}}{3}$. In those cases in which any value of $\bar{\varepsilon}_m \leq 0.5$ gave a contribution smaller than $\bar{\varepsilon}/3$ the value $\bar{\varepsilon}_m = 0.5$ has been fixed.

	$\bar{\varepsilon} = 10^{-1}$	$\bar{\varepsilon} = 10^{-2}$	$\bar{\varepsilon} = 10^{-3}$	$\bar{\varepsilon} = 10^{-4}$
r=10	5	17	70	344
r=1	5	9	391	2208
r=0.1	5	9	411	2330

Table 1.4 Values of n at different values of r , such that the spurious terms defined in (1.81) and (1.82) are definitely smaller than the given accuracy $\bar{\varepsilon}$. The values of $\bar{\varepsilon}_M$ are those obtained in Table 1, while $\bar{\varepsilon}_m$ is always set to zero.

with $\tilde{\varepsilon}_m = 0$. In this way the effects of the presence of a non-vanishing $\tilde{\varepsilon}_m$ are evidenced. In our opinion, to take $\tilde{\varepsilon}_m \neq 0$ is worthwhile only for very small $\bar{\varepsilon}$, while the faster decay of the contribution related to the choice of $\tilde{\varepsilon}_M$, makes the possibility of setting $\bar{\varepsilon}_n > \bar{\varepsilon}_M$ might be of interest.

The values reported in Tables 3 and 4 are computed starting from large values of n , which, of course, yield an overall magnitude of the spurious terms $\varepsilon^d \approx \bar{\varepsilon}_m + \bar{\varepsilon}_M$, and reporting the value for which the value $\bar{\varepsilon}$ is achieved. However, it should be noticed that the terms due to the three steps of the discretization often do not have the same sign. Therefore, for given $\tilde{\varepsilon}_m$ and $\tilde{\varepsilon}_M$, it is possible to find much smaller values of n for which the overall magnitude ε^d practically vanishes. This behavior is evidenced in Fig. 4 where the cusps in the log scales locate the change of sign. Typically, ε^d is negative for very large values of n . Decreasing n , ε^d becomes positive first, and then negative again at very small values of n . In this last case, because of the poor resolution, the change of sign may generate in the log-scale plot an abrupt change of slope. An example of smaller values of n yielding a given overall magnitude of ε is reported in Table 5. The difference is particularly relevant for smaller values of the spurious terms.

We conclude that very few discrete energies are needed to keep the magnitude of the spurious terms moderately low ($\bar{\varepsilon} \approx 10^{-2}$). Furthermore, even fewer energies are needed to control the spurious term in the momentum equation, since its magnitude is one order smaller, e.g. $\tilde{\varepsilon}_m = 0$, $\tilde{\varepsilon}_M = 16$ and $n = 9$ would give $\bar{\varepsilon} \approx 10^{-3}$.

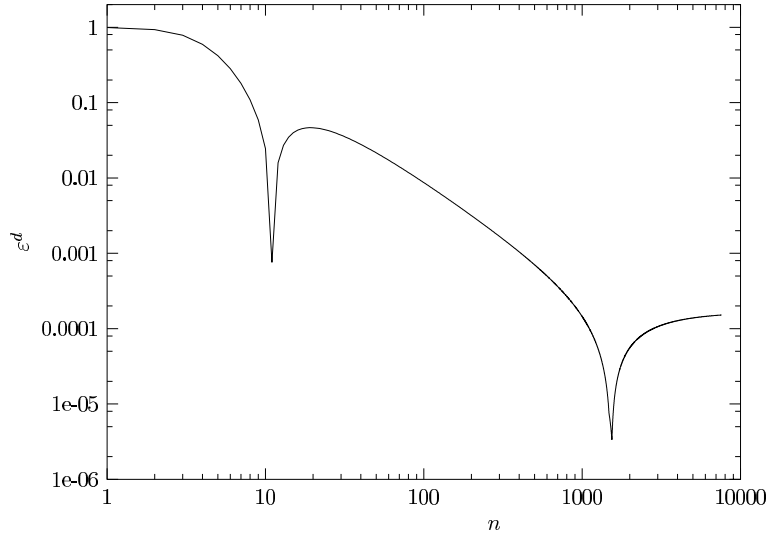


Fig. 1.4 Plot of ε^d as a function of n for $\tilde{\varepsilon}_m = 0$, $\tilde{\varepsilon}_M = 17.49$, and $r = 0.1$. The non regular minima correspond to change of sign in ε^d . Notice the asymptotic behaviour towards the value $\bar{\varepsilon}_b = \bar{\varepsilon}_M = \frac{1}{3}10^{-3}$ determined when \mathbf{R}_+ is substituted with $[0, 17.49]$.

	$\bar{\varepsilon} = 10^{-1}$	$\bar{\varepsilon} = 10^{-2}$	$\bar{\varepsilon} = 10^{-3}$	$\bar{\varepsilon} = 10^{-4}$
r=10	(0, 2.1, 5)	(0.39, 2.3, 7)	(0.3, 2.46, 15)	(0.21, 2.63, 33)
r=1	(0, 4.7, 5)	(0.1, 5.4, 9)	(0.35, 5.92, 22)	(0.27, 6.45, 72)
r=0.1	(0, 14, 5)	(0, 16, 9)	(0.71, 17.5, 13)	(0.5, 19.1, 208)

Table 1.5 Examples of values of $(\bar{\varepsilon}_m, \bar{\varepsilon}_M, n)$ with n smaller than the value reported in Table 4 and which yield the desired overall magnitude of the spurious terms.

From the application viewpoint, it seems that the method is more efficient for higher values of r , i.e. when the kinetic energy is much larger than the internal energy. Some flexibility can be obtained by adjusting the various contributions to the overall error $\bar{\varepsilon}$ (e.g. changing $\bar{\varepsilon}_M$). However, already from Table 1 it can be seen that the extrema of the energy interval weakly depend on slight changes of $\bar{\varepsilon}_M$ and $\bar{\varepsilon}_b$. In addition, changing slightly the energy interval influences the numerical complexity only through the possible increase in the number n of energy levels. The important thing is to keep this number low. For this reason, it seems better to allow a larger error contribution to the error directly related to the discretization than to the one related to the identification of the energy interval.

1.6 Energy Formulation of the Collision Dynamics

To conclude the description of our model, we must express the collision dynamics in terms of the particles' energy per unit mass e , and velocity direction. For the sake of simplicity, we drop the index referring to the energy discretization.

The pre- and post-collisional energies denoted respectively by (e, e_*) and (e', e'_*) , are related through energy conservation (1.7), while the pre- and post-collisional directions, denoted respectively by $(\hat{\Omega}, \hat{\Omega}_*)$ and $(\hat{\Omega}', \hat{\Omega}'_*)$, are related through the conservation of momentum, and are given, up to a scattering angle, by

$$\begin{aligned}\hat{\Omega}' &= \frac{1}{2\sqrt{e'}}(\sqrt{e}\hat{\Omega} + \sqrt{e_*}\hat{\Omega}_* + \sqrt{E - S}\hat{g}') , \\ \hat{\Omega}'_* &= \frac{1}{2\sqrt{e'_*}}(\sqrt{e}\hat{\Omega} + \sqrt{e_*}\hat{\Omega}_* - \sqrt{E - S}\hat{g}') ,\end{aligned}$$

where $\hat{g}' \in S_2$, and S is defined in (1.8).

To clarify the relation between the incoming and the outgoing velocity direction, refer to Figure 2 and decompose the direction of the pre- and post-collisional relative velocities \hat{g} and \hat{g}' as

$$\begin{aligned}\hat{g} &= \cos \varphi \hat{\mathbf{R}} + \sin \varphi \hat{\mathbf{e}}_R , \\ \hat{g}' &= \cos \varphi' \hat{\mathbf{R}} + \sin \varphi' \hat{\mathbf{e}}'_R ,\end{aligned}\tag{1.83}$$

where the unit vectors $\hat{\mathbf{e}}_R$ and $\hat{\mathbf{e}}'_R$ are orthogonal to $\hat{\mathbf{R}} = \mathbf{R}/R$ and belong respectively to the plane containing the pre- and post-collisional directions.

The expression of the angles $\varphi \in [0, \pi]$ and $\varphi' \in [0, \pi]$ in (1.83) formed by \hat{g} and $\hat{\mathbf{R}}$ and by \hat{g}' and $\hat{\mathbf{R}}$ respectively is given by

$$\cos \varphi = \hat{g} \cdot \hat{\mathbf{R}} = \frac{2(e_* - e)}{gR} = \frac{E - 2e}{\sqrt{E^2 - S^2}} \implies \sin \varphi = \sqrt{\frac{4ee_* - S^2}{E^2 - S^2}}\tag{1.84}$$

$$\cos \varphi' = \hat{g}' \cdot \hat{\mathbf{R}} = \frac{2(e' - e'_*)}{gR} = \frac{2e' - E}{\sqrt{E^2 - S^2}} \implies \sin \varphi' = \sqrt{\frac{4e'e'_* - S^2}{E^2 - S^2}}\tag{1.85}$$

while the unit vector $\hat{\mathbf{e}}'_R$ can, for instance, be defined by

$$\hat{\mathbf{e}}'_R = \begin{cases} \frac{(2e + S)\sqrt{2e_*}\hat{\Omega}_* - (2e_* + S)\sqrt{2e}\hat{\Omega}}{R\sqrt{4ee_* - S^2}} \cos \vartheta + \frac{\hat{\Omega} \times \hat{\Omega}_*}{|\hat{\Omega} \times \hat{\Omega}_*|} \sin \vartheta & \text{if } \hat{\Omega} \neq \pm \hat{\Omega}_* ; \\ \frac{\hat{\mathbf{e}}_x - (\hat{\Omega} \cdot \hat{\mathbf{e}}_x)\hat{\Omega}}{|\hat{\Omega} \times \hat{\mathbf{e}}_x|} \cos \vartheta + \frac{\hat{\Omega} \times \hat{\mathbf{e}}_x}{|\hat{\Omega} \times \hat{\mathbf{e}}_x|} \sin \vartheta & \text{if } \hat{\Omega} = \hat{\Omega}_* \neq \hat{\mathbf{e}}_x ; \\ \frac{\hat{\mathbf{e}}_y - (\hat{\Omega} \cdot \hat{\mathbf{e}}_y)\hat{\Omega}}{|\hat{\Omega} \times \hat{\mathbf{e}}_y|} \cos \vartheta + \frac{\hat{\Omega} \times \hat{\mathbf{e}}_y}{|\hat{\Omega} \times \hat{\mathbf{e}}_y|} \sin \vartheta & \text{if } \hat{\Omega} = \hat{\Omega}_* \neq \hat{\mathbf{e}}_y ; \end{cases}\tag{1.86}$$

The angle $\vartheta \in [0, 2\pi)$ appearing in (1.86) is the angle formed by the pre- and post-collisional plane and is not determined by the conservation laws. In conclusion, the outgoing velocity directions can be written as

$$\begin{aligned}\hat{\Omega}' &= \frac{\sqrt{4e'e'_* - S^2}}{\sqrt{2e'R}} \hat{e}'_R + \frac{2e' + S}{\sqrt{e'R^2}} (\sqrt{e'} \hat{\Omega} + \sqrt{e_*} \hat{\Omega}_*) , \\ \hat{\Omega}'_* &= -\frac{\sqrt{4e'e'_* - S^2}}{\sqrt{2e'_*R}} \hat{e}'_R + \frac{2e'_* + S}{\sqrt{e'_*R^2}} (\sqrt{e'} \hat{\Omega} + \sqrt{e_*} \hat{\Omega}_*) .\end{aligned}\tag{1.87}$$

From Eq.(1.5) and Figure 2, it is evident that, as $\hat{\mathbf{g}}'$ runs over S_2 , the post-collisional velocities \mathbf{v}' and \mathbf{v}'_* span the surface of the sphere centered in $\mathbf{R}/2$ with diameter equal to the modulus of the relative velocity g . Furthermore, for any $\hat{\mathbf{g}}'$ the post-collisional velocities point at two antipodal points. More in detail, if the polar axis of the unit sphere is set along \mathbf{R} , one of the post-collisional velocities runs, with varying ϑ , over a given northern parallel and the other over the relative southern parallel (at the antipodes). The other angle defining $\hat{\mathbf{g}}'$, the colatitude φ' , determines through (1.85) the energy of the outgoing particle.

As is well known, the scattering cross section is a function of the modulus of the relative velocity and of the deflection angle $\alpha \in [0, \pi]$ formed by the pre- and post-collisional relative velocities, which can be written as

$$\begin{aligned}\gamma = \cos \alpha = \hat{\mathbf{g}} \cdot \hat{\mathbf{g}}' &= \cos \varphi \cos \varphi' + \sin \varphi \sin \varphi' \cos \vartheta = \\ \frac{1}{E^2 - S^2} &\left\{ (e_* - e)(e' - e'_*) + 4ee_* \sqrt{[1 - (\hat{\Omega} \cdot \hat{\Omega}_*)^2] \left[\frac{e'e'_*}{ee_*} - (\hat{\Omega} \cdot \hat{\Omega}_*)^2 \right]} \cos \vartheta \right\} .\end{aligned}\tag{1.88}$$

Note that γ is an even function of $\hat{\Omega} \cdot \hat{\Omega}_*$ and of ϑ , and observe that

$$\begin{aligned}\gamma(e_*, e, e'_*, \hat{\Omega}_* \cdot \hat{\Omega}, \vartheta) &= \gamma(e, e_*, e', \hat{\Omega} \cdot \hat{\Omega}_*, \vartheta) , \\ \gamma(e', e'_*, e, \hat{\Omega}' \cdot \hat{\Omega}'_*, -\vartheta) &= \gamma(e, e_*, e', \hat{\Omega} \cdot \hat{\Omega}_*, \vartheta) ,\end{aligned}\tag{1.89}$$

which means that the deflection angle is symmetric under the exchange of the roles of the field and test particles. These properties were used in the proof of Theorem 1.2.

1.7 Concluding remarks

This chapter has presented a kinetic model based on a discretization procedure of the Boltzmann equation written in the energy formulation, in which particles take a discrete number of equidistributed energies, and which has the property to automatically preserve mass, momentum, and energy without giving rise to spurious collision invariants.

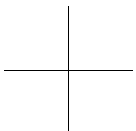
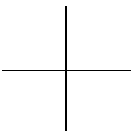
One of the advantages of the discretization of the velocity space in a particular sequence of spherical shells is that for given pre-collisional energies, conservation of energy is satisfied naturally. Then, conservation of momentum for fixed pre-collisional velocities and post-collisional energies restricts the post-collisional velocities to given (symmetric) parallels of the sphere with polar axis along the collision momentum \mathbf{R} , with center in $\mathbf{R}/2$, and with diameter equal to the modulus of the relative velocity.

In place of the five-fold collision integral over an infinite domain, the discretized model is characterized by a collision term which consists of a sum of integrals over finite domains, namely the cartesian product of a unit circle and the portion of the unit sphere between two parallels symmetric with respect to the equatorial plane perpendicular to the velocity of the field particle, it is known [22] that the quadrature rules for periodic functions or for functions over spheres have a faster convergence. Therefore the collision step of the splitting algorithm can be performed with good precision and small computational effort.

Deriving the model via a controllable approximation procedure makes it possible to obtain estimates on the “distance” between the discretized collision operator and the continuous Boltzmann equation, i.e. consistency of the quadrature rule. Finally, the fluid dynamic limit related to the discretized Boltzmann equation tends rapidly towards the usual Euler equations with an isotropic pressure tensor and vanishing heat flux, when the number of energies tends to infinity and the discretization interval tends to \mathbb{R}_+ . Actually, the evaluation of the magnitude of the spurious terms yields itself a procedure to set the discretized energies.

This procedure could be developed refining the interpolation step described in Section 4. In fact, in the model described in this chapter the distribution function is approximated by a piecewise constant function leading to a quadrature rule with a rate of convergence of the order of $1/n^2$. This is a rather rough approximation that can be improved using for instance, basic interpolants that span more than one energy interval.

A possibility is to use cubic splines $\chi_i(e)$ with support also in the neighbouring intervals, i.e. in $\mathbb{I}_{i-1} \cup \mathbb{I}_i \cup \mathbb{I}_{i+1}$. The structure of the model would still be similar to (1.66) with two extra summations related to the integration of $\chi_i(e)$ over the intervals $\mathbb{I}_{i\pm 1}$. The numerical complexity would then remain the same, but the fact that the distribution function would be approximated with a cubic spline (i.e. a piecewise C^2 function rather than a piecewise constant function) would lead to a faster rate of convergence and therefore to the possibility of using fewer nodes to achieve the same accuracy. However, the exact form of the discretized collision operator depends on how the interpolation at the extrema of the finite energy interval is performed. As stress throughout the chapter this step has to be done with care, in order to preserve the conservation laws and the other properties that characterize the Boltzmann equation.



Bibliography

- [1] Neunzert H., and Struckmeier J. (1995) "Particle methods for the Boltzmann equation", *Acta Numerica*, 417.
- [2] Aristov V. V., and Tcheremissine F. G. (1980) "The conservative splitting method for solving Boltzmann's equation", *USSR Comp. Math. Phys.* **21** 208.
- [3] Buet C. (1995) "A discrete-velocity scheme for the Boltzmann operator of rarefied gas-dynamics", in *Proc. of 19th Rarefied Gas Dynamics Symposium*, J. Harvey and G. Lord, eds., Oxford University Press, 878.
- [4] Buet C. (1996) "A discrete-velocity scheme for the Boltzmann operator of rarefied gas-dynamics", *Transp. Theory Statist. Phys.* **25**, 33.
- [5] Buet C. (1997) "Conservative and entropy schemes for the Boltzmann collision operator of polyatomic gases", *Math. Models Methods Appl. Sci.* **7**, 165.
- [6] Bobylev A.V., Palcewski A. and Schneider J. (1995) "Discretization of the Boltzmann equation and discrete velocity models", in *Proc. of 19th Rarefied Gas Dynamics Symposium*, J. Harvey and G. Lord, eds., Oxford University Press, 857.
- [7] Bobylev A.V., Palcewski A. and Schneider J. (1995) "On approximation of the Boltzmann equation by discrete velocity models", *Compt. Rend. Acad. Sci. serie I*, 639.
- [8] Palcewski A., Schneider J. and Bobylev A.V. (1997) "A consistency result for discrete-velocity schemes of the Boltzmann equation", *SIAM J. Num. Anal.* **34**, 1865.
- [9] Rogier F. and Schneider J. (1994) "A direct method for solving the Boltzmann equation", *Transp. Theory Statist. Phys.* **23**, 313.
- [10] Tcheremissine F.G. (1997) "Conservative discrete ordinates method for solving Boltzmann kinetic equation", *Comm. Appl. Math.*, ????
- [11] N. Bellomo and E. Longo (1991) "A new discretized model in nonlinear kinetic theory: The semicontinuous Boltzmann equation", *Math. Models Methods Appl. Sci.* **1**, 113.
- [12] E. Longo, L. Preziosi and N. Bellomo (1993) "The semicontinuous Boltzmann equation: Towards a model for fluid dynamic applications", *Math. Models Methods Appl. Sci.* **3**, 65.
- [13] L. Preziosi (1993) "The semicontinuous Boltzmann equation for gas mixtures", *Math. Models Methods Appl. Sci.* **3**, 665.
- [14] Preziosi L. and Longo E. (1997) "On a conservative polar discretization of the Boltzmann equation", *Japan J. Ind. Appl. Math.* **14**, 1.
- [15] Preziosi L., and Rondoni L. (1999) "Conservative energy discretization of Boltzmann operator", *Quarterly Appl. Math.* **57**(4), 699.
- [16] W. Koller, F. Hanser and F. Schürer (2000) "A semicontinuous extended kinetic

- model", *J. Phys. A* **33**, 3417.
- [17] W. Koller (2000) "A semicontinuous kinetic model for bimolecular chemical reactions", *J. Phys. A* **33**, 6081.
- [18] W. Koller and F. Schürer (2001), " P_n approximation of the nonlinear semicontinuous Boltzmann equation", *Trans. Th. Stat. Phys.* **30**, 471.
- [19] Temam R. (1968) "Sur la stabilité et la convergence de la method des pas fractionnaires", *Ann. Math. Pura Appl.* **79**, 191.
- [20] Desvillettes L. and Mischler S. (1996) "About the splitting algorithm for Boltzmann and BGK equations", *M³AS: Math. Models Methods Appl. Sci.* **6**, 1079.
- [21] Di Perna R.J. and Lions P.L. (1989) "On the Cauchy problem for Boltzmann equations: Global existence and weak stability", *Ann. Math.* **130**, 321.
- [22] P.J. Davis and P. Rabinowitz (1984) *Methods of Numerical Integration*, Academic Press.