Abstract

Discrete-velocity models provide a way for constructing conservative approximations of the collision operator in numerical methods for solving the Boltzmann equation. This paper discusses the convergence of a discrete-velocity collision operator to Boltzmann's collision integral for known conservative models and a new model which is based on Carleman's representation of the collision integral. For this model we prove a convergence result for continuous distribution functions and give error estimates depending on their regularity. In Carleman's variables, integration over the sphere in the collision term is replaced by integration over planes in $\mathbb{R}^3$. The resulting simplification in the structure of the quadrature formula which corresponds to the discrete-velocity model, allows us to prove the convergence using elementary properties of plane lattices. The proof is substantially simpler than for the earlier models which use approximate integration over the sphere.

Keywords: the Boltzmann equation, discrete-velocity models, Carleman's representation, convergence estimates

Acknowledgement

I wish to express my gratitude to my advisor Alexei Heintz, who suggested me this topic and supported me during my work, showing constant interest and answering all my questions. I would also like to thank Leif Arkeryd, Mohammad Asadzadeh and Bernt Wennberg for reading the first version of the manuscript and useful comments which helped to improve it significantly. I gratefully acknowledge the financial support given by the Swedish Institute. I am thankful to all my colleagues at Chalmers University of Technology and Gothenburg University who helped me in my work.

Introduction

The nonlinear Boltzmann equation is the basic equation of kinetic theory. It is used to describe the evolution of a gas which is considered as a collection of interacting particles. The model is physically relevant when the gas is rarefied and the binary collisions of particles prevail. The Boltzmann equation has a number of applications in various fields of natural science, such as high-speed aerodynamics, gas kinetics, and semiconductor physics. We refer to the books of Cercignani [17], Cercignani, Illner, and Pulvirenti [18], and Truesdell and Muncaster [52] for the physical background and the mathematical theory of this equation.

The kinetic theory approach uses a distribution function $f(\xi, x, t)$ which depends on the molecular velocity ξ, and space and time variables x, t. This function represents the mass density in (ξ,x)-space at time t. The physical parameters of a gas, such as mass density, mean velocity, and energy can be obtained easily if the distribution function is known. However, to determine these parameters from the kinetic description, the equation for the distribution function has to be solved, and this is a difficult problem. Firstly, the dimension of the problem is increased compared to the fluid-dynamic description, since the dependence on the molecular velocity is taken into account. Secondly, the collision terms of the kinetic equations are given by high-dimensional integrals, and their computation in a numerical method requires a huge amount of calculations.

Currently, the most efficient methods for solving the Boltzmann equation use stochastic techniques to overcome these difficulties. The most widely used methods of this group are known as the *particle simulation* methods. They use a stochastic simulation of the collision process with properties either derived from molecular dynamics or from the Boltzmann equation. The evolution of the particle system is described by a Markov process and its averages are considered as an approximation of the solution to the Boltzmann equation. The most successful realizations of particle methods are Bird's scheme [5]–[7] and its variants, which are usually referred to as *Direct Simulation Monte Carlo* schemes (DSMC), and also the Nanbu scheme [39] later modified by Babovsky [2]. The practical efficiency of these methods is due to the fact that the amount of computations for these schemes grows linearly with the number of particles, whereas for the majority of other schemes this growth is quadratic. Other variants of the DSMC method were introduced by Belotserkovskiy and Yanitskiy [4], Deshpande [24], and Ivanov and Rogazinsky [34].

The initial formulation of these methods was purely heuristic, and their consistency with the mathematical theory of the kinetic equation remained an open problem for a long time. The relation of these methods to the kinetic equation for a system of N particles was studied by Ivanov and Rogazinsky [34]. The first proof of convergence to the Boltzmann equation for the Nanbu method was given by Babovsky and Illner [3], and for Bird's method by Wagner [53] (see also the paper by Pulvirenti, Wagner, and Zavelani Rossi [42]).

The main drawback of the particle simulation methods is that they lead to large statistical fluctuations for the approximate solutions, which is especially noticeable if the number of particles in a computation cell is kept small. This is why the interest in alternative methods, and first of all in the ones that give less fluctuations, remains high.

We will concentrate here on the discussion of methods which are characterised as *deterministic*. The common feature of these methods is that they start from the Boltzmann equation (or some other kinetic equation) and combine the standard methods of approximating the equation (like finite difference, finite element or finite volume schemes) with a suitable discretization of the collision term. The development of these methods was initiated in the fifties by Nordsieck and Hicks [40], and further progress was made by Hicks and Yen [29], and Yen and Lee [58]. Their method combines a finite difference scheme for space and time variables with Monte-Carlo evaluation of the collision integral. A similar approach was developed by Cheremisin [19]–[21], who used the splitting method and conservative finite-difference schemes for approximation of the free-flow and relaxation parts. For the collision term, different types of Monte Carlo procedures, as well as regular quadrature methods, were used, and the error in conservation laws was corrected on each time step by multiplying the distribution function by a polynomial in ξ.

An approach combining the use of Hermite polynomial expansion for the distribution function, and a Gaussian quadrature for evaluating the collision integral is developed by Chorin [22]. Although an accurate solution to the model problems is obtained using this method, the computational costs are high. A faster method was suggested by Sod [49], but its limitations are connected with the use of Hilbert's expansion for the distribution function. More recently, Ohwada [41] applied an approach based on Laguerre polynomial expansion of the distribution function in the axially symmetric geometry, obtaining an efficient numerical solution to the shock-wave problem. This method is, however, geometry-specific.

Another type of methods is connected with the theory of the discrete-velocity models (DVM). These models approximate the distribution function by a function defined on a finite set of velocities and they consider a finite-dimensional system of quasilinear hyperbolic equations with quadratic interaction terms, instead of the Boltzmann equation. Such models were actively developed after the work of Broadwell [13] who considered a six-velocity model when analysing the shock wave problem. Several authors studied this and other models of the same type afterwards, and obtained numerical results for physically relevant problems, as well as some analytical solutions. The development of this theory is presented in the series of lecture notes by Gatignol [26], Cabannes [15], and Monaco and Preziosi [38]. The main feature of these models is that they have a structure similar to the Boltzmann equation and satisfy the discrete analogues of its basic macroscopic properties, expressed by the conservation laws and the entropy condition. Therefore, the advantage of these models is that no correction procedures are required to avoid a systematic error in numerical computation, as opposed to the methods which use nonconservative Monte Carlo or deterministic approximation for the collision term. The discrete-velocity models can also be easily adapted to more complicated physical situations involving the description of chemical reactions and gas mixtures.

The problem which was unsolved for a long time, was whether the discrete-velocity models reproduce the Boltzmann equation as the number of discrete points in the velocity space grows to infinity. The use of discrete models with a large number of velocities to approximate the Boltzmann equation started from the works of Aristov [1], Tan *et al.* [50], Bobylev and Dolgosheina [8], Inamuro and Sturtevant [33], and Goldstein and Sturtevant [27]. The latter used a discrete-velocity model in the context of particle simulation. Numerical experiments with these models, for different numbers of velocities, showed that if this number is sufficiently large, the results are close to each other and are in a good agreement with those obtained by other methods, at least on the level of macroscopic parameters. However, there was no theoretical proof of convergence, and even the question of constructing the DVM which satisfy the conservation laws, using large number of velocities, remained unanswered.

In order to overcome the difficulties of constructing convergent DVM, Illner, Rjasanow, and Wagner introduced the random discrete-velocity models [30], [32], [31]. In these models, the set of velocities and their weights in the approximation of the distribution function changes randomly in the time evolution. For such models the proof of consistency with the Boltzmann equation in the case of spatially uniform relaxation was given [54], but the original model failed to satisfy the conservation of momentum and energy and was inapplicable in the numerical computations. A modification of this model, for which all conservation laws were fulfilled, was given in [31]. The further investigations led to the development of generalized particle simulation methods [45].

More recently, progress was achieved by Rogier and Schneider [46], who constructed a fully conservative discrete-velocity approximation for the Boltzmann collision integral with two-dimensional velocity space, and gave a convergence proof for it. Generalization of this scheme to the case of three dimensions was given later by Michel and Schneider [36]. Buet [14] developed an $O(N)$-scheme based on the three-dimensional model, which was analogous to the one used by Sturtevant *et al.* [33], [27]; and he also gave a heuristic argument of why the convergence of this scheme should be expected. The rigorous convergence analysis of this scheme, presented by Bobylev, Palczewski and Schneider in [9] required the use of very recent and exact results from number theory concerning the distribution of integer points on spheres. Mischler [37] proved weak L_1-convergence for the solution to the Cauchy problem for DVM to the DiPerna-Lions solution of the Boltzmann equation.

In this paper we consider a discrete-velocity model with a similar structure to the one of Rogier, Schneider and Michel, and Bobylev, Palczewski and Schneider. This model is also conservative, satisfies an entropy condition and has only physical collision invariants. To obtain this model we follow the same procedure, as in [9], but use a different representation of the collision term, which is due to Carleman [16]. In this representation, integration over the sphere in the collision term is replaced by integration over planes in $\mathbb{R}^3$. This leads to a significant simplification in the structure of the quadrature formula and allows us to prove the convergence result, which is analogous to the one presented in [9], [10], but using quite elementary arguments. The Carleman transform adds a singularity in the region corresponding to small velocity changes in the collision. We study the effect of such a singularity on the accuracy of the quadrature formula and find the condition when it can be neglected. The error estimate for the quadrature formula was found to be between $Ch^{1/4}$ and $C_\varepsilon h^{1-\varepsilon}$ for Lipschitz continuous and Lipschitz differentiable distribution functions under certain assumption on the collision cross-section. This model was suggested to the author by A. Heintz. An analogous model with two-dimensional velocity space was previously considered by B. Wennberg and F. Golse [55].

1 The Boltzmann equation

The classical Boltzmann equation for a gas of identical particles with radially symmetric interaction has the form

$$\frac{\partial f}{\partial t} + \xi \cdot \nabla f = Q(f,f), \quad (x,t) \in \mathcal{D} \subseteq \mathbb{R}^3 \times \mathbb{R}, \quad \xi \in \mathbb{R}^3, \tag{1.1}$$

where the collision integral $Q(f,f)$ is given by the expression

$$Q(f,f) = Q^+(f,f) - Q^-(f,f). \tag{1.2}$$

Here

$$Q^+(f,g)(\xi) = \int_{\mathbb{R}^3}\int_{S^{(2)}} f(\xi')\, g(\eta')\, B(\xi-\eta,\omega)\, d\omega\, d\eta, \quad \xi \in \mathbb{R}^3, \tag{1.3}$$

$$\xi' = \xi - \omega(\omega, \xi-\eta), \quad \eta' = \eta + \omega(\omega,\xi-\eta), \tag{1.4}$$

and

$$Q^-(f,g)(\xi) = f(\xi) \int_{\mathbb{R}^3} g(\eta) \int_{S^{(2)}} B(\xi-\eta,\omega)\, d\omega\, d\eta, \quad \xi \in \mathbb{R}^3. \tag{1.5}$$

Expressions (1.3) and (1.5) are called the “gain” and the “loss” term respectively, and represent the effects of production of particles with the given velocity and their scattering in the collision process.

The function $B(u,\omega)$ is of the form

$$B(u,\omega) = B_0\left(|u|, \frac{|(u,\omega)|}{|u|}\right), u \in \mathbb{R}^3, \omega \in S^{(2)}. \tag{1.6}$$

It contains the information about the binary interactions of particles and reflects the physical properties of the model. The condition $B(u,\cdot) \in L_1(S^{(2)}), u \in \mathbb{R}^3$ is usually assumed to obtain the convergent integrals in (1.3), (1.5). If the particle interactions are modelled by inverse power forces with angular cut-off, then

$$B_0(r,x) = r^\gamma b(x), \tag{1.7}$$

where $\gamma \in (-3,1]$, and $b \in L_1([0,1])$. In the case of "hard sphere" molecules, $B_0(r,x) = rx$.

Important properties of the Boltzmann equation are conservation of mass, momentum and energy, and entropy condition

$$\int_{\mathbb{R}^3} Q(f,f)(\xi) \begin{pmatrix} 1 \\ \xi \\ \xi^2 \end{pmatrix} d\xi = 0 \tag{1.8}$$

$$\int_{\mathbb{R}^3} Q(f,f)(\xi) \log f(\xi)\, d\xi \geq 0. \tag{1.9}$$

They can all be deduced from the following identity expressing the symmetries of the collision integral:

$$\int_{\mathbb{R}^3} Q(f,f)(\xi)\psi(\xi)\, d\xi = -\tfrac{1}{4} \int_{\mathbb{R}^3\times\mathbb{R}^3\times S^{(2)}} (f'f_1' - ff_1)(\psi' + \psi_1' - \psi - \psi_1) B\, d\omega\, d\eta\, d\xi, \tag{1.10}$$

which holds for all functions ψ for which the left-hand side is defined. Here we used the notations $f' = f(\xi')$, $f_1' = f(\xi_1')$, etc.

It is well-known that all the functions ψ for which $\int Q(f,f)\psi\, d\xi = 0$ are given by linear combinations of 1, ξ, and ξ^2. All equilibrium distribution functions, that is those which satisfy $Q(f,f) = 0$, are Maxwellians:

$$f(\xi) = \frac{\rho}{(2\pi RT)^{\frac{3}{2}}} \exp\left(-\frac{|\xi - v|^2}{2RT}\right), \tag{1.11}$$

where ρ, $T > 0$, and $v \in \mathbb{R}^3$ are the density, temperature, and mean velocity of the gas, and R is an absolute constant.

In the homogeneous case the H-theorem follows from (1.10), (1.9):

$$\frac{d}{dt}\int_{\mathbb{R}^3} f(\xi,t) \log f(\xi,t)\, d\xi \leq 0, \tag{1.12}$$

showing that the entropy $S(f) = -\int f \log f \, d\xi$ can only increase in the time evolution. In the inhomogeneous case in the absence of boundaries the same result holds for the mean entropy:

$$\frac{d}{dt} \int\limits_{\mathbb{R}^3 \times \mathbb{R}^3} f(\xi, x, t) \log f(\xi, x, t) \, d\xi \, dx \le 0. \tag{1.13}$$

The properties (1.8), and (1.12)–(1.13) are important when constructing the numerical schemes for the Boltzmann equation, since otherwise a systematical error on the macroscopic level will be introduced.

2 The splitting method

The most common approach to the solution of the nonstationary problems for the Boltzmann equation (and stationary ones, if using the relaxation method) is to apply the splitting procedure, in which the solution process is decomposed into two phases: free-molecular flow and time relaxation. For the first phase the homogeneous transport equation is solved:

$$\begin{aligned} \frac{\partial f^{n+1/2}}{\partial t} + \xi \cdot \nabla f^{n+1/2} &= 0 \\ f^{n+1/2}(0) &= f^n, \end{aligned} \tag{2.1}$$

and for the second one the problem of time relaxation is solved using the initial data obtained in the previous step

$$\begin{aligned} \frac{\partial f^{n+1}}{\partial t} &= Q(f^{n+1}, f^{n+1}) \\ f^{n+1}(0) &= f^{n+1/2} \end{aligned} \tag{2.2}$$

Here we assume that the finite time interval $[0, T]$ is subdivided into the intervals $[t_n, t_{n+1}]$ of the same length Δt, and that f_n is the approximate solution at the time t_n.

Schemes of this type are widely used in situations when the right-hand side operator of the equation is the sum of two operators of a simpler structure, and there is an extensive literature considering these methods [56], [35]. Bogomolov [12] proved convergence of scheme (2.1)–(2.2) to the solution of the Boltzmann equation in the limit $\Delta t \to 0$ for the Cauchy problem in the case when there is a smooth classical solution. Recently the convergence of this scheme for the DiPerna-Lions solution of the Boltzmann equation was established by Desvillettes and Mischler [25].

The splitting procedure gives a base for using DSMC and other particle simulation schemes. In these methods, the standard approach involves the subdivision of the physical domain into cells and considers the relaxation phase in each cell separately. During the time relaxation the collision process is modelled for the particles which are in a given cell. In the free-flow phase the particles are moved into new locations corresponding to their current position and speed, thus modelling the particle exchange between the cells. Also in the free-flow stage the boundary conditions are taken into account.

In the deterministic setting some discretization of the distribution function in velocity and space variables is introduced, and numerical schemes for the equations (2.1) and (2.2) are constructed based on this discretization. The solution of the transport phase is the easier of the two problems. Here standard methods such as finite differences, finite volumes or finite elements can be applied. Initially, since only the simple geometry problems were considered, and the requirements on accuracy of the method were relatively low, a first-order, upwind, explicit, finite-difference scheme was mainly used, see [19]–[21]. A higher order explicit scheme was applied in the method proposed by Tan and Varghese [51]. In the numerical solution of problems with model kinetic equations such as the BGK model, the use of high-order non-oscillatory schemes of the type used in gas dynamics is common [57]. The finite-volume and finite-element discretizations applied to the solution of the Boltzmann equation [14], [46] have the advantage of easier adaptation to the geometry of the physical domain.

The most time-consuming step of the scheme is the relaxation phase. The two main goals in constructing a numerical scheme for this stage are thus reducing the computational cost and keeping the conservation properties of the scheme. The standard difference scheme used for the equation (2.2) is the first order forward Euler scheme

$$f^{n+1} = f^n + \Delta t\, Q(f^n, f^n), \tag{2.3}$$

which is both conservative and positive for sufficiently small Δt. However, there is still a problem with the choice of a conservative and numerically efficient calculation procedure for the collision integral. In the work of Buet [14], this problem is solved by applying a randomization procedure to a discrete-velocity model. By these means the number of multiplications in the computation of the collision integral for an N-velocities approximation is reduced to $O(N)$, while a direct calculation requires $O(N^2)$ multiplications.

Another way of improving the scheme performance at this stage would be the use of implicit schemes for the relaxation problem, which overcome the restrictions on the length of the time step Δt for (2.3). Applying such methods to the Boltzmann equation, however, presents certain difficulties, and the use of the implicit Euler scheme, for example, was introduced only recently [11].

3 Discrete-velocity schemes for the collision term

The main advantage of using the discrete-velocity models for calculation of the collision integral is that they satisfy exactly the conservation laws. However, constructing the models which retain these properties and approximate the Boltzmann equation, is difficult. The positive answer to the question of existence of such models was given recently in the works of Bobylev, Palczewski and Schneider [9], [10], Michel, Rogier and Schneider [46], [36]. The question of convergence for solutions to DVM to the Boltzmann equation was studied by Mischler [37]. We outline the classical approach of DVM and consider two models for which the convergence results were proved.

In discrete-velocity models it is assumed that the velocities of gas particles belong to a finite set $V \subseteq \mathbb{R}^3$. The distribution function $f(\xi)$, $\xi \in \mathbb{R}^3$, is replaced by a finite-dimensional approximation f_i, $\xi_i \in V$, and the following system of equations is considered for f_i:

$$\frac{\partial f_i}{\partial t} + \xi_i \cdot \nabla f_i = Q_i(f,f), \quad (x,t) \in \mathcal{D} \subseteq \mathbb{R}^3 \times \mathbb{R},\ \xi_i \in V, \tag{3.1}$$

$$Q_i(f,f) = \sum_{jkl} A_{ij}^{kl}\,(f_k f_l - f_i f_j), \tag{3.2}$$

where A_{ij}^{kl} are constant coefficients, and the summation is taken over all indices corresponding to the discrete velocities in V. If the coefficients A_{ij}^{kl} satisfy the conditions

$$A_{ij}^{kl} = A_{ji}^{lk}, \quad A_{ij}^{kl} = A_{kl}^{ij}, \tag{3.3}$$

and if

$$A_{ij}^{kl} \neq 0 \text{ only if } \xi_i + \xi_j = \xi_k + \xi_l \text{ and } \xi_i^2 + \xi_j^2 = \xi_k^2 + \xi_l^2, \tag{3.4}$$

then it is easy to see that the analogue of the relation (1.10) holds:

$$\sum_{i \in \mathbb{Z}^3} Q_i(f,f)\psi_i = -\frac{1}{4} \sum_{i,j,k,l} A_{ij}^{kl}\,(f_k f_l - f_i f_j)(\psi_k + \psi_l - \psi_i - \psi_j), \tag{3.5}$$

and thus, as in the case of the Boltzmann equation, the conservation laws and the entropy condition are satisfied:

$$\sum_i Q_i(f,f) \begin{pmatrix} 1 \\ \xi_i \\ \xi_i^2 \end{pmatrix} = 0; \qquad \sum_i Q_i(f,f)\,\log f_i \leq 0. \tag{3.6}$$

However, the question of the form of equilibrium solution is not so simple. It is known, see *e.g.* [26], that all solutions of the equation $Q_i(f,f) = 0$ are described by the condition that $\log f$ is a collision invariant, that is

$$\log f \in \Big\{ \psi \;\Big|\; \psi_i + \psi_j - \psi_k + \psi_l = 0 \text{ if } A_{ij}^{kl} \neq 0 \Big\}.$$

There are models for which collision invariants other than the classical ones exist, and the form (1.11) of the equilibrium state is not preserved. That is why the property of having only the proper collision invariants is important.

We describe here a model which has all the conservation properties and which approximates the Boltzmann equation. It was used in numerical computations in [27], [33], [14], and analyzed in [9], [10]. This model is based on a regular discretization of the velocity space: for some positive h the set

$$\mathcal{Z}_h = h\mathbb{Z}^3 = \left\{ h(i_1, i_2, i_3) \,\middle|\, i_{1..3} \in \mathbb{Z} \right\} \tag{3.7}$$

is considered, and a function f defined on $\mathbb{R}^3$ is approximated by a function on $\mathcal{Z}_h$: $f(\xi_i) \approx f_i$, $\xi_i \in \mathcal{Z}_h$. From now on, let us agree to use letters i, j, k, l to denote vectors in $\mathbb{Z}^3$. When writing $\xi_i \in \mathcal{Z}_h$ we mean $\xi_i = hi$, $i \in \mathbb{Z}^3$, etc. We use the same convention for functions on $\mathcal{Z}_h$.

Replacing the integration over $\mathbb{R}^3$ in (1.3), (1.5) by a rectangle quadrature approximation, we obtain

$$Q(f,f)(\xi_i) \approx h^3 \sum_{j \in \mathbb{Z}^3} \int_{S^{(2)}} \Big(f(\xi_i')\, f(\xi_j') - f(\xi_i)\, f(\xi_j) \Big)\, B(\xi_i - \xi_j, \omega)\, d\omega.$$

The idea of approximating the integrals over $S^{(2)}$ is to use those points of the velocity space which fall on the sphere Σ_{ij} with the diameter $|\xi_i - \xi_j|$ passing through the points ξ_i and ξ_j. It is based on the fact that for fixed ξ_i and ξ_j the velocities after the collision ξ_i' and ξ_j' run twice over all pairs of opposite points on this sphere, as ω runs over $S^{(2)}$. Thus, assuming that the distribution of the points of $\mathcal{Z}_h$ on such spheres is close to uniform when h is small, we can write

$$\begin{aligned}
&\int_{S^{(2)}} \Big(f(\xi_i')\, f(\xi_j') - f(\xi_i)\, f(\xi_j) \Big)\, B(\xi_i - \xi_j, \omega)\, d\omega \\
&= \frac{8}{|\xi_i - \xi_j|^2} \int_{\Sigma_{ij}} \Big(f(\xi_i')\, f(\xi_j') - f(\xi_i)\, f(\xi_j) \Big)\, \frac{B(\xi_i - \xi_j, \omega(\sigma))}{|\cos\theta(\sigma)|}\, d\sigma \\
&\approx \frac{8\pi}{|S_{ij}|} \sum_{k,l \in S_{ij}} \Big(f(\xi_k)\, f(\xi_l) - f(\xi_i)\, f(\xi_j) \Big)\, \frac{B(\xi_i - \xi_j, \omega_{ij}^{kl})}{|\cos\theta_{ij}^{kl}|},
\end{aligned} \tag{3.8}$$

where θ is the angle between $\xi_i - \xi_j$ and ω, and S_{ij} is the set of all indices k, l such that ξ_k, ξ_l are on the sphere Σ_{ij}. Thus, setting

$$A_{ij}^{kl} = \begin{cases} h^3 \dfrac{8\pi}{|S_{ij}|} \dfrac{B(\xi_i - \xi_j, \omega_{ij}^{kl})}{|\cos\theta_{ij}^{kl}|}, & \text{if } \begin{array}{c} k + l = i + j \\ k^2 + l^2 = i^2 + j^2 \end{array} \\ 0 \quad \text{otherwise,} \end{cases}$$

we obtain a model of the type (3.2) with the properties (3.3), and (3.4). It is known, see [14], that this model has no other collision invariants except mass, momentum, and energy, and thus, its equilibrium solutions have the form

$$f_i = \exp(A\xi_i^2 + (B, \xi_i) + C), \quad \xi_i \in \mathcal{Z}_h. \tag{3.9}$$

The convergence theorem for this model is due to Bobylev *et al.* [9], [10]:

Theorem 3.1 *Suppose f is a continuous function on $\mathbb{R}^3$ decaying faster than $|\xi|^{-3}$ as $|\xi| \to \infty$, and $q(u,\omega) = B(u,\omega)/\cos\theta$ is continuous and satisfies*

$$0 \le q(u,\omega) \le a + b|u|, \quad u \in \mathbb{R}^3, \quad \omega \in S^{(2)}.$$

Then

$$Q(f,f)(\xi_i) - Q_i(f,f) \underset{h\to 0}{\longrightarrow} 0$$

uniformly with respect to ξ_i on compact subsets of $\mathbb{R}^3$.
Suppose now that f and q are smooth ($C^{3,\alpha}$, $\alpha > 0$) functions, and f and all its derivatives up to the order 3 decay at infinity faster than $|\xi|^{-5}$. Then for sufficiently small h,

$$|Q(f,f)(\xi_i) - Q_i(f,f)| \le C_\varepsilon h^{1/175-\varepsilon}$$

with the same uniformity condition with respect to ξ_i.

The main difficulty of the proof consists in establishing the consistency of approximation (3.8) for the integrals over the spheres. The convergence of formula (3.8) for continuous integrands is connected with properties of distribution of integer points on the spheres of large integer radius, and the proof here is based on recent results from number theory (see references in [9], [10]). Notice also that the proof of Theorem 3.1 given in [9], [10] is not applicable to the analogous two-dimensional model used in computations by Inamuro and Sturtevant [33].

There is an alternative approach by Schneider, Michel and Rogier [46], [36], which works in the two-dimensional as well as in the three-dimensional situation. It uses the same velocity grid $\mathcal{Z}_h$, but changes the order of sphere and space integrations. First, a grid for calculating the integral over the sphere is defined as a central projection of the set

$$\mathcal{F}_N^d = \left\{(p_1,..,p_d) \in \mathbb{Z}^d \;\middle|\; |p_i| \le N, \text{ g.c.d.}(p_1,..,p_d) = 1\right\}$$

onto $S^{(2)}$. In the two-dimensional situation there is a natural interpretation of this set in terms of Farey series, see [46], which provides the weights for the quadrature formula. The generalization to the three-dimensional case, done in [36], involves introducing the subdivision of the sphere into the set of cells $R(p)$ centered around the points $p \in \mathcal{F}_N^d$, and defined in the following way:

$$R(p) = \left\{x \in S^{(d-1)} \;\middle|\; |p_m \frac{x_n}{x_m} - p_n| = \min_{q \in \mathcal{F}_N^d} |q_m \frac{x_n}{x_m} - q_n|, \; n = 1..d\right\},$$

where m is such that $|p_m| = \max(|p_1|, .., |p_d|)$. Now the quadrature formula

$$Q(f,f)(\xi_i) \approx \sum_{p \in \mathcal{F}_N^d} |R(p)| \int_{\mathbb{R}^d} \big(f(\xi')\, f(\eta') - f(\xi_i)\, f(\eta)\big)\, B(\xi_i - \eta, \frac{p}{|p|})\, d\eta \tag{3.10}$$

can be applied for the integration over $S^{(2)}$. A quadrature formula of the type (3.2) is obtained by using only those values of η in (3.10), for which ξ', η' belong to $\mathcal{Z}_h$. Using (1.4) we obtain that for fixed $\omega = \frac{p}{|p|}$, ξ' lies on the line $\xi_i + \mathbb{R}p = \{\xi_i + \alpha p \mid \alpha \in \mathbb{R}\}$, and η' lies on the plane $\xi_i + (\mathbb{R}p)^{\perp}$, and thus, it is possible to give the following quadrature formula:

$$\int_{\mathbb{R}^d} \big(f(\xi')\, f(\eta') - f(\xi_i)\, f(\eta)\big)\, B(\xi_i - \eta, \frac{p}{|p|})\, d\eta \approx h^3 |p|^2 \sum_{r \in \mathbb{Z}} \sum_{\substack{m \in \mathbb{Z}^3 \\ m \cdot p = 0}} B(h(rp+m), \frac{p}{|p|})$$
$$\times \Big(f(\xi_i + hrp)\, f(\xi_i + hm) - f(\xi_i)\, f(\xi_i + h(rp+m))\Big). \tag{3.11}$$

Here we used (1.4) to obtain $\eta = \xi' + \eta' - \xi_i$, and $h^3|p|^2$ stands for the volume of the fundamental cell of the integration lattice

$$\Big\{ h(rp+m) \,\Big|\, r \in \mathbb{Z},\ m \in \mathbb{Z}^3,\ p \cdot m = 0 \Big\}.$$

By combining the expressions (3.10) and (3.11) the equations for the discrete model of the type (3.2) are obtained, with the coefficients A_{ij}^{kl} given by

$$A_{ij}^{kl} = \begin{cases} h^3 |p_{ij}^{kl}|^2\, |R(p_{ij}^{kl})|\, B(\xi_i - \xi_j, \omega_{ij}^{kl}), & \text{if} \quad \begin{array}{l} (k-i) \perp (l-i) \\ j = k + l - i \end{array} \\ 0 \quad \text{otherwise,} \end{cases}$$

where p_{ij}^{kl} is a vector with relatively prime components in the direction of the collision parameter ω_{ij}^{kl}. Now it is easy to see that the symmetry conditions (3.3) are satisfied. Further, since

$$\left\{ \begin{array}{c} (k-i) \perp (l-i) \\ j = k + l - i \end{array} \right. \quad \text{if and only if} \quad \left\{ \begin{array}{c} k + l = i + j \\ k^2 + l^2 = i^2 + j^2, \end{array} \right. \tag{3.12}$$

the condition (3.4) is also true. In view of (3.5) this provides the conservation properties of the model. The proof that the equilibrium solutions have the form (3.9) is the same as in the case of the previous model, since it uses only symmetry properties of the velocity space and the conditions (3.3), (3.4). For the model considered, the convergence result of the type of Theorem 3.1 is obtained in [46], [36], and the convergence estimates $C_\varepsilon h^{1-\varepsilon}$ in the two-dimensional case, and $C_\varepsilon h^{6/11-\varepsilon}$ in the three-dimensional one are proved.

4 A discrete-velocity model using Carleman's variables

In this section we introduce a discrete-velocity model based on the Carleman representation of the Boltzmann collision integral. The motivations for considering this model are the difficulties in proving the consistency of the integral approximation (3.8) over the sphere and the weak error estimate for the model analyzed in [9], [10]. In Carleman's variables the integration over spheres in the collision term is replaced by integration over planes. Thus, after applying the rectangle formula for the integral over $\mathbb{R}^3$, the question of uniform distribution of integer points on the domain of integration becomes trivial. However, the quadrature accuracy for the plane integrals is not uniform with respect to the plane parameter, and so, some averaging is required in order to prove that the approximation of the collision term is consistent. In contrast to the proof in [9], [10], all the steps of the analysis of this model are quite elementary.

We begin by introducing the Carleman representation for the collision term [16]. It consists in the change of variables

$$(\eta,\ \omega)\ \mapsto\ \Big(p=\xi-\omega(\omega,\xi-\eta),\ q=\eta+\omega(\omega,\xi-\eta)\Big).$$

For fixed values of ξ and p, q runs twice over the plane $E_{\xi p}$, containing ξ and orthogonal to $\xi - p$, when ω runs over the unit sphere. The functional determinant of the inverse transform is $|\xi - p|^{-2}$.

It is convenient for our purposes to modify these new variables as follows:

$$u = p - \xi, \quad w = q - \xi.$$

Let us also use the notation E_u for the plane orthogonal to u:

$$E_u = \left\{ w \in \mathbb{R}^3 \,\middle|\, (u, w) = 0 \right\}.$$

In the new variables the "gain" and "loss" terms of the collision operator are transformed as follows:

$$Q^+(f,g)(\xi) = \int_{\mathbb{R}^3} f(\xi+u) \int_{E_u} g(\xi+w)\, \widetilde{B}(u,w)\, dw\, du, \quad \xi \in \mathbb{R}^3, \tag{4.1}$$

and

$$Q^-(f,g)(\xi) = f(\xi) \int_{\mathbb{R}^3} \int_{E_u} g(\xi+u+w)\, \widetilde{B}(u,w)\, dw\, du, \quad \xi \in \mathbb{R}^3, \tag{4.2}$$

where

$$\widetilde{B}(u,w) = 2\,|u|^{-2} B_0 \left(\sqrt{u^2+w^2}, \frac{|u|}{\sqrt{u^2+w^2}} \right). \tag{4.3}$$

We see that the function $\widetilde{B}(u,w)$ has a singularity at $u=0$, even if B_0 is smooth. This singularity produces an additional source of error when using a regular quadrature formula for approximating the integrals. We restrict ourselves by considering only the kernel functions $\widetilde{B}$ of the form

$$\widetilde{B}(u,w) = |u|^{-3+\beta}\bar{B}(u,w) \tag{4.4}$$

with some positive β such that $\bar{B}(u,w)$ is a globally continuous function. In the case of "hard spheres", this assumption is satisfied, since $\widetilde{B}(u,w) = 2|u|^{-1}$, and we can take $\beta = 2$ and $\bar{B}(u,w) = \text{const}$. However, the assumption (4.4) is rather restrictive when considering kernels of more general form. In the case of kernels (1.7), for example, this condition is true only if certain relations between γ and $b(x)$ are fulfilled. We refer to Section 8 for the further discussion of admitted collision kernels. Notice that for all γ and $b(x)$ in (1.7), for which the assumption (4.4) can be satisfied, the value of the parameter β can be chosen in the interval $(0,2]$.

We introduce the following notations for the integrands in the expressions (4.1), and (4.2):

$$F^{+}(\xi,u,w) = f(\xi+u)g(\xi+w)\widetilde{B}(u,w), \tag{4.5}$$

$$F^{-}(\xi,u,w) = f(\xi)g(\xi+u+w)\widetilde{B}(u,w), \tag{4.6}$$

$$G^{\pm}(\xi,u) = \int\limits_{E_u} F^{\pm}(\xi,u,w)\,dw. \tag{4.7}$$

Since the functions F^{+}, F^{-} and G^{+}, G^{-} have similar properties, we will use the notations F and G where either of these functions can be used. Notice that $G^{\pm}$ and $F^{\pm}$ are not defined for $u=0$. It is convenient, though, to assign certain value at this point to allow the use of 0 as a node of the quadrature rule. Thus, we set $G^{\pm}(\cdot,0)=0$, and $\widetilde{B}(0,\cdot)=0$, so that $F^{\pm}(\cdot,0,\cdot)=0$.

To construct the discrete-velocity model we use the same velocity space $\mathcal{Z}_h$ defined by (3.7) as in the two models considered above. Recall that the letters i, j, k, l always denote vectors in $\mathbb{Z}^3$. Using them as the indices of velocity variables, we always mean multiplication by h: $\xi_i = hi$, etc. When considering functions on $\mathcal{Z}_h$, the lower index is used to denote the value at the corresponding point: $f_i = f(\xi_i) = f(hi)$, $i \in \mathbb{Z}^3$.

To introduce a quadrature formula which uses the values of the distribution function only at the points of $\mathcal{Z}_h$ for approximating the integrals in (1.3) and (1.5), we follow the same procedure as in [14], [10]. First, the integrals over $\mathbb{R}^3$ are approximated by the three-dimensional rectangle formula:

$$\int\limits_{\mathbb{R}^3} G(\xi,u)\,du \approx h^3 \sum_{k\in\mathbb{Z}^3} G(\xi,u_k), \tag{4.8}$$

and then the values of the integrand at the points of $\mathcal{Z}_h$ lying on the planes E_{u_k} are used to calculate integrals over the planes. To give an expression for this last approximation, let us consider the set $\mathcal{L}_{u_k,h}$ which is the intersection of the discrete velocity space and the plane E_{u_k} for some fixed $k \in \mathbb{Z}^3$, $k \neq 0$:

$$\mathcal{L}_{u_k,h} \stackrel{\text{def}}{=} E_{u_k} \cap \mathcal{Z}_h = \left\{ hl \in \mathbb{R}^3 \;\middle|\; (u_k, l) = 0,\; l \in \mathbb{Z}^3 \right\} = hL_k, \tag{4.9}$$

where

$$L_k = \left\{ l \in \mathbb{Z}^3 \;\middle|\; (k, l) = 0 \right\}, \tag{4.10}$$

that is the set of the solutions of the linear Diophantine equation $(k, l) = 0$. This last set forms a lattice of rank 2 in $\mathbb{Z}^3$, i.e.,

$$L_k = \{ e_1 m + e_2 n \;\big|\; e_1, e_2 \in \mathbb{Z}^3;\; m, n \in \mathbb{Z} \},$$

where the vectors e_1 and e_2 are linearly independent over $\mathbb{R}$ as the vectors of $\mathbb{R}^3$. Then the standard lattice rule [48] can be used for calculation of the integrals over E_{u_k}:

$$\int\limits_{E_{u_k}} F(\xi, u_k, w)\, dw \;\approx\; h^2 \Delta_k \sum_{l \in L_k} F(\xi, u_k, w_l), \tag{4.11}$$

where Δ_k is the area of the fundamental cell of L_k, that is, of the parallelogram spanned by the base vectors of L_k: $\Delta_k = |e_1 \times e_2|$. Notice that though the basis of the lattice can be chosen in different ways, Δ_k does not depend on this choice [28]. Intuitively, one can interpret this quadrature formula as follows: each basis of the integration lattice $\mathcal{L}_{k,h}$ defines a splitting of the plane E_{u_k} into the set of equal parallelograms centered around the corresponding lattice points. They all are obtained by shifting the fundamental cell. Summing up the values of the integrand at the points of the lattice times the area of the cell gives the approximation (4.11).

Combining (4.8) and (4.11) we arrive at the following expressions for the discrete "gain" and "loss" terms:

$$Q_h^+(f, g)(\xi_i) = h^5 \sum_{k \in \mathbb{Z}^3} f(\xi_i + u_k) \Delta_k \sum_{l \in L_k} \widetilde{B}_{jk} g(\xi_i + w_l), \quad \xi_i \in \mathcal{Z}_h, \tag{4.12}$$

$$Q_h^-(f, g)(\xi_i) = h^5 f(\xi_i) \sum_{k \in \mathbb{Z}^3} \Delta_k \sum_{l \in L_k} \widetilde{B}_{jk} g(\xi_i + u_k + w_l), \quad \xi_i \in \mathcal{Z}_h, \tag{4.13}$$

where $\widetilde{B}_{jk} = \widetilde{B}(u_k, w_l)$.

Using (4.12), (4.13), and the relation (3.12)it is easy to obtain the expression for the discrete collision term in the form (3.2). The definition of the coefficients A_{ij}^{kl} is now the following:

$$A_{ij}^{kl} = \begin{cases} h^5 \Delta_{k-i} \widetilde{B}_{k-i,l-i}, & \begin{array}{l} k + l = i + j \\ k^2 + l^2 = i^2 + j^2, \end{array} \\ 0, \quad \text{otherwise}. \end{cases} \tag{4.14}$$

From this form of the discrete collision operator the symmetry conditions (3.3) can be observed easily. Clearly, (3.4) is also fulfilled. The proof of the statement that the equilibrium states have the form (3.9) can be easily adapted from the one for the two models considered in the previous section. Formally, the only difference that arises is that the values of A_{ij}^{kl} for $k = l$ and $i = j$ are put equal to zero in our model. However, this does not change the proof, since these coefficients correspond to the cases of "trivial" collisions which do not give any contribution to the collision term.

Thus, the obtained discrete-velocity model for the collision term satisfies all the conservation properties discussed above. We summarize these properties in the following theorem.

Theorem 4.1 *The discrete-velocity model defined by (3.1), (3.2) with the velocity space (3.7) and coefficients defined by (4.14) satisfies the discrete mass, momentum, and energy conservation laws as well as the entropy property, expressed by the conditions (3.6). All positive solutions to the equation*

$$Q_h(f, f) = 0$$

are given by the discrete Maxwellians

$$f^M(\xi_i) = \exp(A\xi_i^2 + (B, \xi_i) + C), \quad \xi_i \in \mathcal{Z}_h$$

where $A, C \in \mathbb{R}$, and $B \in \mathbb{R}^3$ do not depend on i. This means that there are no other collision invariants except the classical ones.

In the next three sections we give the convergence analysis for the discrete collision term of the model.

5 A convergence result for the discrete-velocity model

In this section two convergence theorems are formulated. We make a distinction between the cases of the continuous integrands, when no estimates are available, and the Lipschitz ones, for which convergence estimates are obtained. Denote by $\mathcal{P}$ the set where the kernel function $\widetilde{B}(u, w)$ is defined:

$$\mathcal{P} = \left\{(u, w) \in \mathbb{R}^3 \times \mathbb{R}^3 \;\middle|\; u \perp w\right\}.$$

We consider $\mathcal{P}$ as a submanifold of $\mathbb{R}^3 \times \mathbb{R}^3$ and denote by $C^{m-1,1}(\mathcal{P})$, $m \geq 1$, the space of C^{m-1}-functions on $\mathcal{P}$ having the locally Lipschitz m-th derivative. The notation $C_0^{m-1,1}(\mathcal{P})$ is used for compactly supported functions with the above properties.

The main theorems can now be formulated as follows:

Theorem 5.1 *Assume that f, $g \in C_0^{m-1,1}(\mathbb{R}^3)$ for some $m \geq 1$, and the kernel $\widetilde{B}$ defined by (4.3) satisfies (4.4) with $\bar{B} \in C^{m-1,1}(\mathcal{P})$ and $\beta > 0$. Then, for sufficiently small h*

$$|Q^{\pm}(f,g)(\xi_i) - Q_h^{\pm}(f,g)(\xi_i)| \leq Ch^r, \quad \xi_i \in \mathcal{Z}_h, \tag{5.1}$$

where

$$r = \min\left(\frac{m}{m+3}, \beta, m/\left(m+3+\tfrac{2-\beta-m}{1+\beta+m}\right)\right), \tag{5.2}$$

and the bound (5.1) is uniform with respect to ξ_i on compact subsets of $\mathbb{R}^3$.

Theorem 5.2 *Assume that $f, g \in C_0(\mathbb{R}^3)$, and the kernel $\widetilde{B}$ satisfies the assumption (4.4) with $\bar{B} \in C(\mathcal{P})$ and $\beta > 0$. Then*

$$Q^{\pm}(f,g)(\xi_i) - Q_h^{\pm}(f,g)(\xi_i) \to 0 \text{ as } h \to 0, \;\; \xi_i \in \mathcal{Z}_h \tag{5.3}$$

uniformly with respect to ξ_i on compact subsets of $\mathbb{R}^3$.

Remark. We notice that the third term in (5.2) can be omitted when $m \geq 3$.

It is convenient to reformulate the assumptions on f, g, and $\bar{B}$ in terms of the functions $F^{\pm}$ and $G^{\pm}$ from (4.5)–(4.7). Since these functions have a singularity of the form $|u|^{-3+\beta}$ at $u = 0$, we introduce the new functions

$$\bar{F}^{\pm}(\xi,u,w) = |u|^{3-\beta} F^{\pm}(\xi,u,w), \tag{5.4}$$
$$\bar{G}^{\pm}(\xi,u) = |u|^{3-\beta} G^{\pm}(\xi,u), \tag{5.5}$$

where this singularity is removed. If f, g are compactly supported continuous functions, then $\bar{F}^{\pm}$ and $F^{\pm}$ have compact support in $\mathcal{P}$, and the support bound is uniform, with respect to ξ on any compact set $K \subseteq \mathbb{R}^3$. Further, if f, g, and $\bar{B}$ are continuous or $C^{m-1,1}$-functions, then $\bar{F}^{\pm}$ are continuous or $C^{m-1,1}$-functions, respectively. For a fixed compact K let us denote by R and L the global bounds for the support and the Lipschitz constant of these functions:

$$R = \sup\left\{|u| + |w| \;\middle|\; F^{\pm}(\xi,u,w) \neq 0,\; (u,w) \in \mathcal{P},\; \xi \in K\right\} < +\infty, \tag{5.6}$$
$$L = \sup_{\xi \in K} \operatorname*{Lip}_{\mathcal{P}} \nabla^m \bar{F}^{\pm}(\xi,\cdot,\cdot) < +\infty. \tag{5.7}$$

Evidently, the constant R provides a bound for the support of $G(\xi,\cdot)$ for $\xi \in K$.

The first step in the proof of the convergence theorems is to consider the two possible error sources, namely, space and plane discretizations, separately. The difference in (5.3) is estimated as follows:

$$
\begin{aligned}
&\left|Q^{\pm}(f,g)(\xi_i) - Q_h^{\pm}(f,g)(\xi_i)\right| \\
&= \left| \int_{\mathbb{R}^3}\int_{E_u} F^{\pm}(\xi_i,u,w)\,du\,dw - h^5 \sum_{k\in\mathbb{Z}^3} \Delta_k \sum_{l\in L_k} F^{\pm}(\xi_i,u_k,w_l)\right| \\
&\leq \left| \int_{\mathbb{R}^3} G^{\pm}(\xi_i,u)\,du - h^3 \sum_{k\in\mathbb{Z}^3} G^{\pm}(\xi_i,u_k)\right| \\
&+ \left| \sum_{k\in\mathbb{Z}^3} h^3 \Big(\int_{E_{u_k}} F^{\pm}(\xi_i,u_k,w)\,dw - h^2\Delta_k \sum_{l\in L_k} F^{\pm}(\xi_i,u_k,w_l)\Big)\right| \\
&= \left| \sum_{|k|\leq R/h} \mathcal{S}_k(\xi_i,G^{\pm},h)\right| + \left| \sum_{|k|\leq R/h} h^3\,\mathcal{R}_k(\xi_i,F^{\pm},h)\right| = \mathcal{S}+\mathcal{R},
\end{aligned} \tag{5.8}
$$

where

$$
\begin{aligned}
\mathcal{R}_k(\xi,F,h) &= \int_{E^R_{u_k}} F(\xi,u_k,w)\,dw - h^2\Delta_k \sum_{l\in L^R_k} F(\xi,u_k,w_l), \\
\mathcal{S}_k(\xi,G,h) &= \int_{B_k} G(\xi,u)\,du - h^3 G(\xi,u_k).
\end{aligned}
$$

Here B_k is the cube of size h around u_k, and L^R_k and $E^R_{u_k}$ are the intersections of L_k and E_{u_k} with the ball or radius R centered at 0. Now, the term $\mathcal{R}$ represents the total error due to the discretization of the plane integrals, and $\mathcal{S}$ is the error of the three-dimensional rectangle formula on the function G. In the next two sections we give the bounds for each of these two terms.

6 Approximation of the integrals over planes

First, we focus our attention on the problem of approximating the integrals over the planes E_u using the lattice formula (4.11). For a fixed $k \in \mathbb{Z}^3$, and a basis for the integration lattice $\mathcal{L}_{k,h}$, the plane E_{u_k} is splitted into a set of equal parallelograms $\{D_{k,l}\}_{l\in L_k}$, centered at the lattice points. We estimate the local approximation error on each parallelogram cell, by using a standard approach based on the Bramble-Hilbert lemma (cf. [43]). Then, summation over all cells lying on the plane gives an estimate for the plane quadrature, which is subsequently used to estimate the total error given by the term $\mathcal{R}$ in (5.8).

The further analysis is based on the following lemma.

Lemma 6.1 *Let $\Omega_0 \subseteq \mathbb{R}^n$ be an open bounded set with Lipschitz boundary, such that $|\Omega_0| = 1$. Let A be a linear invertible mapping on $\mathbb{R}^n$, and $\Omega = A(\Omega_0)$. Let*

$g \in W_p^m(\Omega)$, $m = 1, 2$, $p \in (n/m, \infty]$, and

$$E(g) = \int\limits_{\Omega} g\,dx - |\Omega| g(0). \tag{6.1}$$

Then

$$|E(g)| \leq C|\Omega|^{1/q} \|A\|^m |g|_{m,p,\Omega}, \tag{6.2}$$

where C does not depend on Ω and g, $\|A\|$ denotes the matrix norm of the mapping A, q is the conjugate exponent to p, and

$$|g|_{m,p,\Omega} = \Big(\sum_{|\alpha|=m} \int_{\Omega} |\partial^\alpha g|^p \,dx \Big)^{1/p}, \; p < \infty, \quad |g|_{m,\infty,\Omega} = \max_{|\alpha|=m} \operatorname{ess.sup}_{\Omega} |\partial^\alpha g|. \tag{6.3}$$

Proof. We give only a sketch of the proof, referring to [43] for the details. Let us define the function $g_0(\xi) = g(A\xi)$, $\xi \in \Omega_0$, which evidently belongs to $W_p^m(\Omega_0)$. Then since $|\det A| = |\Omega|$, we find that

$$E(g) = \int\limits_{\Omega} g(x)\,dx - |\Omega|\, g(0) = |\Omega| \Big(\int\limits_{\Omega_0} g_0(\xi)\,d\xi - g(0) \Big) = |\Omega| E(g_0).$$

E can be considered as a linear bounded functional on W_p^m, which vanishes on all polynomials up to degree one. Thus, for the reference domain the inequality

$$|E(g_0)| \leq C|g_0|_{m,p,\Omega_0}$$

can be obtained as an application of the Bramble-Hilbert lemma. Now by using the inequality

$$|g_0|_{m,p,\Omega_0} \leq C|\Omega|^{-1/p} \|A\|^m |g|_{m,p,\Omega}. \tag{6.4}$$

(see [23]), we obtain the estimate (6.2). □

The application of this result to the case of the integration lattice $\mathcal{L}_{k,h}$ is straightforward: a square of the unit area D_0 can be considered as the reference domain Ω_0 and the fundamental cell $D_{k,l}$, $l = 0$ as Ω, with $A_{k,h}\colon D_0 \to D_{k,l}$ being a linear bijection transforming D_0 into $D_{k,l}$. Then using Lemma 6.1 the following bound for the error of plane integration can be obtained.

Lemma 6.2 *Let $K \subseteq \mathbb{R}^3$ be a compact set, and assume that $\bar{F}$ is such that $\bar{F}(\xi, \cdot, \cdot) \in C_0^{0,1}(\mathcal{P})$, and conditions (5.6) and (5.7) are satisfied. Let R and L be the constants defined by these conditions. Then*

$$|\mathcal{R}_k(\xi, \bar{F}, h)| \leq C(R, L) \|A_{k,h}\|, \tag{6.5}$$

and $C(R, L)$ does not depend on k and h and is uniform with respect to ξ on K.

Proof. Using the previous lemma we get the estimate

$$|\mathcal{R}_k(\xi,\bar F,h)| = \Big| \int\limits_{E^R_{u_k}} \bar F(\xi,u_k,w)\,dw - h^2\Delta_k \sum_{l\in L^R_k} \bar F(\xi,u_k,w_l)\Big|$$
$$\le \sum_{l\in L^R_k} \Big| \int\limits_{D_{k,l}} \bar F(\xi,u_k,w) - h^2\Delta_k \bar F(\xi,u_k,w_l)\Big|$$
$$\le C\|A_{k,h}\| \sum_{l\in L^R_k} |D_{k,l}||\bar F(\xi,u_k,\cdot)|_{1,\infty,D_{k,l}} \le CLR^2\|A_{k,h}\|,$$

which proves the lemma. □

Evidently, if $F(\xi,u,w) = |u|^{-3+\beta}\bar F(\xi,u,w)$, then

$$|\mathcal{R}_k(\xi,F,h)| \le C(F)|u_k|^{-3+\beta}\|A_{k,h}\|. \tag{6.6}$$

Remark. For functions of the class $W^m_p(\Omega)$ with $m\ge 3$, an estimate of the form (6.2) can be obtained for

$$E_m(g) = \int\limits_\Omega g\,dx - |\Omega|g(0) - \sum_{2\le|\alpha|\le m-1} d_\alpha\,|\Omega|^\alpha \int\limits_\Omega \partial^\alpha g(\xi)\,d\xi$$

with a suitable choice of the coefficients d_α. This can be easily proved by applying the inequality (6.4) in the original arguments given in [43]. So if the function $\bar F$ in the formulation of Lemma 6.2 is of the class $C^{m-1,1}_0(\mathcal{P})$, then the bound (6.5) can be changed to

$$|\mathcal{R}_k(\xi,\bar F,h)| \le C(R,L)\|A_{k,h}\|^m. \tag{6.7}$$

Lemma 6.2 expresses the fact that the quadrature formula (4.11) converges, whenever the linear size of the cell, given by the norm of $A_{k,h}$ tends to zero as $h\to 0$. In the case of the lattice (4.10), the value of the norm can be expressed in terms of the length of the maximal base vector. If $\{e_1,e_2\}$ is a basis of L_k, that is $\{he_1,he_2\}$ is the one of $\mathcal{L}_{k,h}$, we have (taking spherical norm for definiteness)

$$\|A_{k,h}\| = h\max(|e_1+e_2|,|e_1-e_2|) \le 2h\max(|e_1|,|e_2|). \tag{6.8}$$

A bound for the norms of the base vectors for the lattice defined by a linear Diophantine equation may be obtained using a rather elementary argument, which we consider in the next lemma.

Lemma 6.3 *Let $a\in\mathbb{Z}^3$, $(a,x)=0$ be a linear Diophantine equation. Then the basis of solutions $\{e_1,e_2\}$ of this equation can be chosen so that*

$$|e_1||e_2| \le \sqrt{\tfrac{4}{3}}\,|a|. \tag{6.9}$$

Proof. Let L_a be a lattice of solutions to $(a,x)=0$. Then $\det(L_a)=|a|/(a_1,a_2,a_3)$, where (a_1,a_2,a_3) is the greatest common divisor of a_1, a_2, a_3. This can be easily checked directly using the expression of the solution with parameters obtained by the Euclid algorithm, or it can be obtained as a corollary of a more general fact: [47], Lemma 4C, Ch.1. Now the inequality (6.9) follows from the Hermite's bound for the reduced basis of a lattice: [28], p.71. □

Remark. In the case of a two-dimensional lattice, which is considered in Lemma 6.3, Hermite's bound

$$|e_1||e_2| \le \sqrt{\tfrac{4}{3}}\,\det L_a. \tag{6.10}$$

can be proved using simple geometrical arguments. Namely, since $\det(L_a)$ is the area of a basis parallelogram, $\det(L_a)=|e_1||e_2|\sin(e_1,e_2)$, and the condition (6.10) is equivalent to $\sin(e_1,e_2)\ge\sqrt{3/4}=\sin(\pi/3)$. Then, if the angle between the vectors of the original basis is less than $\pi/3$, we can complement one of the vectors e_1, e_2 by either e_1-e_2 or e_1+e_2 to a new basis, for which this angle is greater than $\pi/3$, and thus the required condition is satisfied.

The inequality (6.9) implies that for the vectors of the reduced basis, the bound $|e_i|\le\sqrt{\tfrac{4}{3}}\,|a|$ is valid. In fact, the constant in this estimate can be improved to give

$$|e_i| \le |a|, \tag{6.11}$$

so that the norm of the linear mapping $A_{k,h}$ can be estimated as

$$\|A_{k,h}\| \le 2h|k| = 2|u_k|. \tag{6.12}$$

We see that this bound does not provide convergence of the norm to zero as $h\to 0$. It is easy to verify, that indeed there is no uniform convergence for all k. To show this, consider for example the integer vectors k of the form $(k_1,k_2,0)$, with k_1, k_2 relatively prime and $|k|>1/(2h)$. For such k, $\|A_{k,h}\|=2|u_k|>1$. Thus, the quadrature formulas over different planes do not converge uniformly. However, using the inequality (6.9) we can show that the fraction of points u_k, for which $\|A_{k,h}\|$ remains large, becomes small as h tends to zero. We use this in the next lemma to prove that the error of the integral plane quadrature tends to zero.

Lemma 6.4 *Let $F(\xi,u,w)=|u|^{-3+\beta}\bar F(\xi,u,w)$ with $\bar F$ being a $C_0^{0,1}(\mathcal{P})$-function, which satisfies the conditions (5.6) and (5.7) on every compact $K\subseteq\mathbb{R}^3$. Let R and L be the constants defined by these conditions. Then for sufficiently small h*

$$\left|\sum_{k\in\mathbb{Z}^3} h^3\mathcal{R}_k(\xi,F,h)\right| \le C(R,L)h^r, \tag{6.13}$$

with

$$r=\min\left(\frac{1}{4},1/\left(4+\tfrac{2-\beta}{1+\beta}\right)\right),$$

and this estimate is uniform with respect to ξ on compact subsets of $\mathbb{R}^3$.

Proof. Let us fix a compact $K \subseteq \mathbb{R}^3$, and put $N = N(h) = R/h$ for R defined by (5.6), and $\mathcal{A}_h = \left\{k \in \mathbb{Z}^3 \,\middle|\, |k| \le N\right\}$. Then the sum over $\mathbb{Z}^3$ in (6.13) can be replaced by the sum over $\mathcal{A}_h$.

By Lemma 6.3 there is a basis $\{e_1^{(k)}, e_2^{(k)}\}$ of integral solutions to $(k,x) = 0$, satisfying (6.9) and (6.11). Let $e_{\max}^{(k)}$ and $e_{\min}^{(k)}$ be the base vectors with maximal and minimal Euclidean norm, respectively.

Take some constant $\alpha \in (0,1)$ (to be fixed later).

If the inequality $|e_{\min}^{(k)}| \ge |k|^\alpha$ holds for some subset $\mathcal{B}_h = \mathcal{B}_h(\alpha)$ of $\mathcal{A}_h$, we get by (6.9) that $|e_{\max}^{(k)}| \le \sqrt{\frac{4}{3}}|k|^{1-\alpha}$. Hence, by (6.11)

$$\|A_{k,h}\| \le 2h|e_{\max}^{(k)}| = C|u_k|^{1-\alpha}h^\alpha. \tag{6.14}$$

Then,

$$\begin{aligned}\sum_{k\in\mathcal{B}_h} h^3|\mathcal{R}_k(\xi,F,h)| &\le C(R,L)h^\alpha \sum_{k\in\mathcal{B}_h} h^3|u_k|^{-2+\beta-\alpha}\\ &\le C(R,L)h^\alpha \int_{|u|\le R} |u|^{-2+\beta-\alpha}du = C(R,L)h^\alpha.\end{aligned}$$

For the remaining part of the indices k we have

$$\mathcal{A}_h \setminus \mathcal{B}_h = \left\{k \in \mathcal{A}_h \,\middle|\, |e_{\min}^{(k)}| < |k|^\alpha\right\} \subseteq \left\{k \in \mathcal{A}_h \,\middle|\, |e_{\min}^{(k)}| < N^\alpha\right\} \stackrel{\text{def}}{=} \mathcal{D}_h.$$

Now, it is easy to estimate the number of points in $\mathcal{D}_h$, and hence in $\mathcal{A}_h \setminus \mathcal{B}_h$. Simple geometrical arguments show that

$$\#\left\{j \in \mathbb{Z}^3 \,\middle|\, |j| \le n\right\} \le Cn^3, \tag{6.15}$$

$$\#E_{u_k} \cap \left\{j \in \mathbb{Z}^3 \,\middle|\, |j| \le n\right\} \le Cn^2, \tag{6.16}$$

where # denotes the number of elements in the set. Using (6.15) and (6.16) we conclude that the number of different vectors $e_{\min}^{(k)}$ satisfying the inequality $|e_{\min}^{(k)}| < N^\alpha$ is estimated by $CN^{3\alpha}$, and the number of vectors $k \in \mathcal{A}_h$ on each plane $\left\{k \in \mathbb{Z}^3 \,\middle|\, (e_{\min}^{(k)}, k) = 0\right\}$ is estimated by CN^2. This gives that

$$\#\mathcal{D}_h \le CN^{2+3\alpha},$$

and thus, if $\beta \ge 2$

$$\begin{aligned}\left|\sum_{k\in\mathcal{D}_h} h^3\mathcal{R}_k(\xi,F,h)\right| &\le C(R,L) \sum_{k\in\mathcal{D}_h} h^3|u_k|^{-2+\beta}\\ &\le C(R,L)h^3R^{-2+\beta}\left(\frac{R}{h}\right)^{2+3\alpha} = C(R,L)h^{1-3\alpha}.\end{aligned}$$

By setting $\alpha = \frac{1}{4}$ we get both parts of $\sum\limits_{j\in\mathcal{A}_h} h^3 \mathcal{R}_k(\xi, F, h)$ estimated by $C(F)h^{1/4}$.

If $\beta < 2$, then $|u_k|^{-2+\beta}$ is unbounded on $\mathcal{D}_h$, and the contributions of "large" and "small" k should be considered separately. For this purpose let us take some $\sigma \in (0,1)$ and split the above sum in the following way:

$$\sum_{k\in\mathcal{D}_h} h^3 |u_k|^{-2+\beta} \le \sum_{k\le N^{1-\sigma}} h^3 |u_k|^{-2+\beta} + \sum_{\substack{k\in\mathcal{D}_h \\ N^{1-\sigma}\le k\le N}} h^3 |u_k|^{-2+\beta}$$

$$\le C \int\limits_{|u|\le R^{1-\sigma}h^\sigma} |u|^{-2+\beta}\, du + h^3 \left(\frac{R}{h}\right)^{2+3\alpha} \max_{N^{1-\sigma}\le k\le N} |u_k|^{-2+\beta}$$

$$\le C(R)\Big(h^{\sigma(1+\beta)} + h^{1-3\alpha-\sigma(2-\beta)}\Big).$$

Thus, if $\sigma = \alpha/(1+\beta)$, and $\alpha = 1/\Big(4 + \frac{2-\beta}{1+\beta}\Big)$, we find

$$\Big|\sum_{k\in\mathcal{D}_h} h^3 \mathcal{R}_k(\xi, F, h)\Big| \le C(F)h^\alpha.$$

By combining the estimates for $\mathcal{B}_h$ and $\mathcal{D}_h$ we complete the proof of the lemma. □

Remark. For the functions from the class $C_0^{m-1,1}(\mathcal{P})$ the estimate of Lemma 6.4 can be improved. By applying inequality (6.7) instead of (6.5), we find that for $m = 2$ the value

$$r = \min\left(\frac{2}{5}, 2/\Big(5 + \frac{1-\beta}{2+\beta}\Big)\right),$$

can be obtained, and for $m \ge 3$,

$$r = \frac{m}{m+3},$$

since the source of unboundedness disappears in the sum over $\mathcal{D}_h$.

The proof of the convergence for continuous functions is obtained by analogy to the Lipschitz case.

Lemma 6.5 *Let $\bar{F}(\xi,\cdot,\cdot) \in C_0(\mathcal{P})$ satisfy the conditions (5.6) and (5.7) for every compact set $K \subseteq \mathbb{R}^3$, and $F(\xi, u, w) = |u|^{-3+\beta}\bar{F}(\xi, u, w)$. Then*

$$\sum_{k\in\mathbb{Z}^3} h^3 \mathcal{R}_k(\xi, F, h) \underset{h\to 0}{\longrightarrow} 0$$

uniformly with respect to ξ on compact subsets of $\mathbb{R}^3$.

Proof. Let us consider ξ in a fixed compact set $K \subseteq \mathbb{R}^3$, and define the sets $\mathcal{A}_h$, and for some fixed $\alpha \in (0, \frac{1}{3})$, $\mathcal{B}_h$ and $\mathcal{D}_h$ as in the proof of Lemma 6.4. Since the function $\bar{F}(\xi,\cdot,\cdot)$ is uniformly bounded on $\mathcal{P}$ with respect to $\xi \in K$, we clearly have

$$\sup_{\xi\in K} \Big|\mathcal{R}_k(\xi, \bar{F}, h)\Big| \le R^2 \sup_{\xi\in K} \sup_{\mathcal{P}} \bar{F}(\xi,\cdot,\cdot) \le C. \tag{6.17}$$

Thus, for any fixed positive ε we can choose a sufficiently small δ such that

$$\Big|\sum_{k\le\delta/h} h^3\mathcal{R}_k(\xi,F,h)\Big| \le C\sum_{k\le\delta/h} h^3|u_k|^{-3+\beta} \le C\int\limits_{|u|\le\delta} |u_k|^{-3+\beta}\,du < \varepsilon. \tag{6.18}$$

Denote by $\mathcal{A}_h^\delta$, $\mathcal{B}_h^\delta$, $\mathcal{D}_h^\delta$ the intersections of $\mathcal{A}_h$, $\mathcal{B}_h$, and $\mathcal{D}_h$ with the set $\{k\in\mathbb{Z}^3\,|\,|k|>\delta/h\}$. According to (6.14), the sizes of the cells of the integration lattices tend to zero uniformly for $k\in\mathcal{B}_h^\delta$. We show that on $\mathcal{B}_h^\delta$ the quadrature error for the function $\bar{F}$ tends to zero uniformly:

$$\sup_{\xi\in K}\sup_{k\in\mathcal{B}_h^\delta}\Big|\mathcal{R}_k(\xi,\bar{F},h)\Big| \underset{h\to0}{\longrightarrow} 0. \tag{6.19}$$

In order to prove this, fix some $u^0\in\mathbb{R}^3$ and consider $A_{u^0\to u}\in SO(\mathbb{R}^3)$, which is the rotation map with respect to the line orthogonal to both u^0 and u, and which transforms E_{u^0} into E_u. It is easy to see that if $|u|>\delta$, then for a fixed $w\in\mathbb{R}^3$, $u\mapsto A_{u^0\to u}w$ is a continuous function. Then,

$$\mathcal{R}_k(\xi,F,h) = \int\limits_{E_{u_k}} \bar{F}(\xi,u_k,w)\,dw - h^2\Delta_k\sum_{l\in L_k}\bar{F}(\xi,u_k,w_l) \tag{6.20}$$

$$= \int\limits_{E_{u^0}} \bar{F}(\xi,u_k,A_{u^0\to u_k}w^0)\,dw^0 - \det\Lambda_k\sum_{w_l\in\Lambda_k}\bar{F}(\xi,u_k,A_{u^0\to u_k}w_l),$$

where Λ_k is a lattice on E_{u^0}, and the size of its fundamental cell $|\Lambda|$ is the same as for hL_k. Consider now the function

$$\mathcal{F}(\xi,u,\Lambda) = \det\Lambda\sum_{w_l\in\Lambda}\bar{F}(\xi,u,A_{u^0\to u}w_l).$$

It converges to the value of the integral in (6.20) as $|\Lambda|\to0$ pointwise for all ξ, u. In order to prove (6.19), it suffices to show that $\mathcal{F}$ converges uniformly if $\xi\in K$, $u\in\{u\in\mathbb{R}^3\mid\delta\le|u|\le R\}$. This holds if $\mathcal{F}$ is uniformly equicontinuous in (ξ,u) with respect to $|\Lambda|\to0$: [44], p.29-30. This last property follows easily from the uniform continuity of $\bar{F}$. Now,

$$\sup_{\xi\in K}\Big|\sum_{k\in\mathcal{B}_h^\delta} h^3\mathcal{R}_k(\xi,F,h)\Big| \le \sup_{\xi\in K}\sup_{k\in\mathcal{B}_h^\delta}|\mathcal{R}_k(\xi,\bar{F},h)|\sum_{k\in\mathcal{B}_h^\delta}h^3|u_k|^{-3+\beta}$$

$$\le CR^3\sup_{\xi\in K}\sup_{k\in\mathcal{B}_h^\delta}|\mathcal{R}_k(\xi,\bar{F},h)| \underset{h\to0}{\longrightarrow} 0. \tag{6.21}$$

To show that the sum over the set $\mathcal{D}_h^\delta$ converges to zero, it suffices to use the bound (6.17) combined with a simple maximum estimate:

$$\sup_{\xi\in K}\Big|\sum_{k\in\mathcal{D}_h^\delta} h^3\mathcal{R}_k(\xi,F,h)\Big| \le |\mathcal{D}_h|h^3\sup_{\xi\in K}\Big|\mathcal{R}_k(\xi,\bar{F},h)\Big|$$

$$\le C\delta^{-3+\beta}h^3\left(\frac{R}{h}\right)^{2+3\alpha} \underset{h\to0}{\longrightarrow} 0. \tag{6.22}$$

Combination of (6.18), (6.21), and (6.22) proves the lemma. $\square$

7 Approximation of the integrals over $\mathbb{R}^3$

In this section we consider the approximation error of the quadrature formula over $\mathbb{R}^3$, which is given by the term $\mathcal{S}$ in (5.8). Let F be such that $F(\xi, u, w) = |u|^{-3+\beta}\bar{F}(\xi, u, w)$ with $\bar{F}(\xi, \cdot, \cdot) \in C_0^{0,1}(\mathcal{P})$, and let $\bar{F}$ satisfy the assumptions (5.6), (5.7) on every compact set $K \subseteq \mathbb{R}^3$. Consider

$$G(\xi, u) = \int_{E_u} F(\xi, u, w)\, dw \tag{7.1}$$

and

$$\bar{G}(\xi, u) = |u|^{3-\beta} G(\xi, u) = \int_{E_u} \bar{F}(\xi, u, w)\, dw. \tag{7.2}$$

Our aim here is to prove that the three-dimensional rectangle formula has sufficient accuracy when applied to the function $G(\xi, \cdot)$. The difficulty which occurs here, is that G has a singularity at $u = 0$, and the standard argument of the type of Lemma 6.2 can not be applied. Nevertheless, the bound of Lemma 6.1 applied to the local error on each cubic cell B_k can be used to show that the total quadrature error is small.

In order to obtain the estimate for the local quadrature error, we give a Lipschitz condition for the function G.

Lemma 7.1 *Let $K \subseteq \mathbb{R}^3$ be a fixed compact set, and let $G(\xi, u)$ be the function given by (7.1) with F satisfying the above assumptions. Let R and L be the constants defined by (5.6) and (5.7). Then for $u, u' \neq 0$, $\xi \in K$*

$$|G(\xi, u) - G(\xi, u')| \leq C(R, L)|u|^{-1}(|u|^{-3+\beta} + |u'|^{-3+\beta})|u - u'|.$$

Proof. First, we prove an analogous inequality for $\bar{G}$. Let us fix two nonzero vectors u and $u' \in \mathbb{R}^3$, and denote by A the rotation mapping $A_{u \to u'}$ introduced in the proof of Lemma 6.5. Then, using the compact support property and the Lipschitz condition for $\bar{F}$, we get

$$\begin{aligned} \left|\bar{G}(\xi, u) - \bar{G}(\xi, u')\right| &= \left| \int\limits_{E_u} \Big(\bar{F}(\xi, u, w)\, dw - \bar{F}(\xi, u', Aw)\Big)\, dw' \right| \\ &\leq L \int\limits_{E_u^R} \Big(|u - u'| + |w - Aw|\Big)\, dw', \end{aligned}$$

where $E_u^R = E_u \cap B_R(0)$. To estimate the difference $|w - Aw|$ in the integrand above, we notice that for all vectors x in the unit ball $B_1(0)$

$$|x - Ax| \leq \left| \frac{u}{|u|} - A\left(\frac{u}{|u|}\right) \right| = \left| \frac{u}{|u|} - \frac{u'}{|u'|} \right|,$$

and hence,

$$\|I-A\| \le \left|\frac{u'}{|u'|}-\frac{u}{|u|}\right| \le \left|\frac{u'}{|u|}-\frac{u}{|u|}\right| + \left|\frac{u'}{|u'|}-\frac{u'}{|u|}\right| = \frac{|u-u'|}{|u|}+\frac{||u|-|u'||}{|u|} \le 2\frac{|u-u'|}{|u|}.$$

Using this estimate of the norm we obtain

$$\left|\bar{G}(\xi,u)-\bar{G}(\xi,u')\right| \le L|u-u'|\int\limits_{E_u^R} dw + \frac{2L}{|u|}|u-u'|\int\limits_{E_u^R}|w|\,dw \le C(R,L)|u|^{-1}|u-u'|.$$

Now, it is easy to prove the desired inequality for G:

$$\begin{aligned}
&\left|G(\xi,u)-G(\xi,u')\right| = \left||u|^{-3+\beta}\bar{G}(\xi,u)-|u'|^{-3+\beta}\bar{G}(\xi,u')\right| \\
&\le \left||u|^{-3+\beta}\bar{G}(\xi,u)-|u'|^{-3+\beta}\bar{G}(\xi,u)\right| + \left||u'|^{-3+\beta}\bar{G}(\xi,u)-|u'|^{-3+\beta}\bar{G}(\xi,u')\right| \\
&= \left||u|^{-3+\beta}-|u'|^{-3+\beta}\right|\bar{G}(\xi,u) + |u'|^{-3+\beta}\left|\bar{G}(\xi,u)-\bar{G}(\xi,u')\right| \\
&\le C(R,L)\,|u|^{-1}(|u|^{-3+\beta}+|u'|^{-3+\beta})|u-u'|.
\end{aligned}$$

Here we used the bounds

$$|\bar{G}(\xi,u)| \le C, \quad |u| < R,$$

and

$$||u|^{-3+\beta}-|u'|^{-3+\beta}| \le C|u|^{-1}(|u|^{-3+\beta}+|u'|^{-3+\beta})|u-u'|,$$

where the latter one follows from the inequality

$$\frac{|x^\alpha-y^\alpha|}{|x-y|} = x^{\alpha-1}\frac{|1-\left(\frac{y}{x}\right)^\alpha|}{|1-\frac{y}{x}|} \le Cx^{\alpha-1}\left(1+\left(\frac{y}{x}\right)^\alpha\right) = Cx^{-1}(y^\alpha+x^\alpha)$$

holding for all x, $y>0$, $x\ne y$. □

Applying Lemma 6.1, we obtain

$$\left|\int\limits_{B_k} G(\xi,u)\,du - h^3G(\xi,u_k)\right| \le Ch^4|G(\xi,\cdot)|_{1,\infty,B_k} \le Ch^4 \operatorname*{Lip}_{B_k} G(\xi,\cdot), \tag{7.3}$$

where B_k is the cubic cell of size h around u_k, and $\mathrm{Lip}_{B_k}G(\xi,\cdot)$ denotes the Lipschitz constant over this cell. Thus, the local error $\mathcal{S}_k(\xi,G,h)$ on each cell B_k, $k\ne 0$ is estimated according to (7.3). Now, the Lipschitz constant of G can be estimated by Lemma 7.1 for $k\ne 0$, and summing up the bounds for $\mathcal{S}_k(\xi,G,h)$ over $k\in\mathbb{Z}^3$ will give the bound for the total spatial quadrature error.

Lemma 7.2 *Let $G(\xi,u)$ be as in Lemma 7.1. Let $K\subseteq\mathbb{R}^3$ be a compact set and let R and L be the constants given by (5.6) and (5.7). Then*

$$\left|\sum_{k\in\mathbb{Z}^3}\mathcal{S}_k(\xi,G,h)\right| \le C(R,L)h^r, \tag{7.4}$$

where $r=\min(1,\beta)$, and the bound is uniform with respect to ξ on K.

Proof. Using Lemma 7.1, the Lipschitz constant of G can be estimated by

$$\operatorname*{Lip}_{B_k} G(\xi, \cdot) \le C(R, L) \max_{B_k} |u|^{-4+\beta}. \tag{7.5}$$

Applying the estimates (7.3) and (7.5) and recalling that $G(\xi, 0) = 0$, we obtain

$$\begin{aligned}\Big|\sum_{k\in\mathbb{Z}^3} \mathcal{S}_k(\xi, G, h)\Big| &\le \sum_{|k|\le R/h} \Big|\int_{B_k} G(\xi, u)\, du - h^3 G(\xi, u_k)\Big| \\ &\le \int_{B_0} G(\xi, u)\, du + C(R, L) h \sum_{0<|k|\le R/h} h^3 \max_{B_k} |u|^{-4+\beta}.\end{aligned}$$

The integral over B_0 can be bounded as

$$\int_{B_0} G(\xi, u)\, du \le C \int_{B_0} |u|^{-3+\beta}\, du \le C \int_{|u|\le h} |u|^{-3+\beta}\, du = \frac{4\pi}{\beta} C h^\beta.$$

Since the maximum of $|u|^{-4+\beta}$ on each cubic cell B_k is attained either at one of the corners or at the centre of one of the sides, the last sum can be rewritten as follows (we skip writing $|k| \le R/h$, considering only such k):

$$\sum_{0<|k|\le R/h} h^3 \max_{B_k} |u|^{-4+\beta} = \Bigg(\sum_{k\in(\mathbb{Z}-\frac{1}{2})^3} + \sum_{\substack{k_1\in\mathbb{Z}-\frac{1}{2}\\ k_2=k_3=0}} + \sum_{\substack{k_2\in\mathbb{Z}-\frac{1}{2}\\ k_3=k_1=0}} + \sum_{\substack{k_3\in\mathbb{Z}-\frac{1}{2}\\ k_1=k_2=0}} \Bigg) h^3 |u_k|^{-4+\beta}$$

In the first sum, the contributions from the eight points closest to zero can be singled out, and the rest is estimated by an integral:

$$\sum_{k\in(\mathbb{Z}-\frac{1}{2})^3} h^3 |u_k|^{-4+\beta} \le 8h^3 \left(\frac{h}{2}\right)^{-4+\beta} + 2 \int_{\frac{h}{2}\le|u|\le R} |u|^{-4+\beta}\, du = C h^s$$

where $s = \min(0, -1+\beta)$. The factor 2 in front of the integral is due to duplication near the coordinate planes. In the same way,

$$\sum_{\substack{k_1\in\mathbb{Z}-\frac{1}{2}\\ k_2=k_3=0}} h^3 |u_k|^{-4+\beta} \le 2h^3 \left(\frac{h}{2}\right)^{-4+\beta} + \int_{\frac{h}{2}\le|u|\le R} |u|^{-4+\beta}\, du = C h^s,$$

and analogously for the remaining two sums. Combining the above estimates and noting that $r = s + 1$, we obtain the desired result. □

The proof of Theorem 5.1 is now achieved by combining the estimates of Lemma 6.4 and Lemma 7.2. The convergence for the error of the space integration in assumptions of continuity of integrands is proved in a much simpler way.

Lemma 7.3 *Assume that $G(\xi, u)$ is defined by (7.1), and that $\bar{F}(\xi, \cdot, \cdot) \in C_0(\mathcal{P})$ satisfies the conditions (5.6) and (5.7) on every compact set $K \subseteq \mathbb{R}^3$. Then*

$$\sum_{k \in \mathbb{Z}^3} \mathcal{S}_k(\xi, G, h) \underset{h \to 0}{\longrightarrow} 0$$

uniformly with respect to ξ on compact sets of $\mathbb{R}^3$.

Proof. Let us fix some positive δ and estimate the above sum as follows:

$$\Big| \sum_{k \in \mathbb{Z}^3} \mathcal{S}_k(\xi, G, h) \Big| \leq \sum_{|k| \leq R/h} \Big| \int_{B_k} G(\xi, u)\, du - h^3 G(\xi, u_k) \Big|$$

$$\leq \int_{|u| \leq \delta} G(\xi, u)\, du + \sum_{0 < |k| \leq \delta/h} h^3 \max_{B_k} G(\xi, u) + \sum_{\delta/h < |k| \leq R/h} \Big| \int_{B_k} G(\xi, u)\, du - h^3 G(\xi, u_k) \Big|.$$

Now, since the integral converges, the first term can be made as small as desired by choosing sufficiently small δ. For the second term we have:

$$\sum_{0 < |k| \leq \delta/h} h^3 \max_{B_k} G(\xi, \cdot) \leq \sup_{\mathcal{P}} \bar{G}(\xi, \cdot) \sum_{0 < |k| \leq \delta/h} h^3 \max_{B_k} |u|^{-3+\beta} \leq C \int_{|u| \leq \delta} |u|^{-3+\beta}\, du \underset{\delta \to 0}{\longrightarrow} 0.$$

The third term is a residual of the Riemann sum for the function $G(\xi, \cdot)$ outside a neighborhood of zero. We show that $\bar{G}(\xi, \cdot)$ is a continuous function outside of this neighborhood. In fact,

$$|\bar{G}(\xi, u) - \bar{G}(\xi, u')| \leq \int_{E_u} \Big| \bar{F}(\xi, u, w) - \bar{F}(\xi, u', A_{u \to u'} w) \Big|\, du,$$

and since $u' \to A_{u \to u'} u$ is a continuous function outside a neighborhood of zero, the continuity follows from the uniform continuity of $\bar{F}$. Now, $G(\xi, \cdot)$ is also continuous, and hence the third term converges to zero as a consequence of Riemann integrability of G.

All bounds can be made uniform with respect to ξ by taking supremums over a compact $K \subseteq \mathbb{R}^3$. $\square$

8 Conclusions

We showed that the discrete-velocity collision term constructed in Section 4 converges to the Boltzmann collision integral for compactly supported distribution functions. The case when these functions decay polynomially can also be treated in a similar way.

The expression for the exponent of convergence (5.2), shows that in the case of "hard sphere" molecules for which the value of the singularity exponent β can be chosen equal to 2, the convergence rate becomes $m/(m+3)$ and is determined only by the regularity of the distribution function. In the limit case when the distribution functions are from C_0^∞ (or the Schwartz class s) the quadrature error is estimated as $C_\varepsilon h^{1-\varepsilon}$, for all $\varepsilon > 0$. The same remark is true when $1 \leq \beta \leq 2$ and the function $\bar{B}$ in the kernel is of the class C^∞. However, if considering kernels of the form (1.7) with smooth function $b(x)$, this convergence rate is generally not attained. For example, for the so-called "variable hard sphere" cross-section, when $b(x) = x$ and $\gamma \in [0,1]$, the condition (4.4) with $\bar{B} \in C^{0,1}(\mathcal{P})$ gives that β must be taken $\leq \gamma$ and hence the value of the exponent of convergence given by Theorem 5.1 is equal to γ for γ close to zero, and vanishes for $\gamma = 0$. The estimate of Theorem 5.1 in this case can be improved by considering the kernel $\widetilde{B}$ in the form

$$\widetilde{B}(u,w) = |u|^{-\alpha}(u^2+w^2)^{-\beta/2}\,\bar{B}(u,w)$$

for some positive α and β and globally Lipschitz $\bar{B}$. For the "variable hard spheres" with $\gamma \in [0,1]$, for example, the convergence exponent $1/4$ can be obtained in this way.

It is also interesting to point out that the argument used in the proof of Lemma 6.4 cannot be applied for the analogous model with two-dimensional velocity space, and as in the case of the model considered in [9] the question of convergence for such a model is open. Thus, the only one of the three types of models considered in this paper, for which the convergence in two-dimensional case is proved, is the one by Rogier and Schneider [46]. Notice that the coefficients of the quadratic form $Q_h(f,f)$ are different for the three models discussed, and the convergence results proved for one model generally cannot be transferred to the other ones. The questions of the relation between these models and their possible generalizations could be the subject of future investigations.

References

[1] Aristov V. V., On solution of the Boltzmann equation for discrete velocities, *Dokl. Akad. Nauk SSSR*, 283 (1985), pp. 831–834 (in Russian)

[2] Babovsky H., On a simulation scheme for the Boltzmann equation, *Math. Methods Appl. Sci.*, 8 (1986), pp. 223–233

[3] Babovsky H., Illner R., A convergence proof for Nanbu's simulation method for the full Boltzmann equation, *SIAM J. Numer. Anal.*, 26 (1989), no. 1, pp. 45–65

[4] Belotserkovskiy O. M., Yanitskiy V. E., The statistical particles-in-cells method for solving rarefied gas dynamics problems, (Russian) *Zh. Vychisl. Mat. Mat. Fiz.*, 15 (1975), pp. 1195–1208 and 1553–1567; translation in *U.S.S.R Comp. Maths. Math. Phys.*, 15 (1975), no. 5, pp. 101-113 and no. 6, 184-197

[5] Bird G. A., Approach to translational equilibrium in a rigid sphere gas, *Phys. Fluids*, 6 (1963), pp. 1518–1519

[6] Bird G. A., Direct simulation and the Boltzmann equation, *Phys. Fluids*, 13 (1970), no. 11, pp. 2676–2681

[7] Bird G. A., *Molecular gas dynamics and the direct simulation of gas flows*, 2nd ed., Oxford Univ. Press (1994)

[8] Bobylev A. V., Dolgosheina E. B., Regular numerical solution for shock waves in two-dimensional Boltzmann gas, *17-th RGD Symposium, Aachen, July 1990, Book of Abstracts*, pp. 71–72

[9] Bobylev A. V., Palczewski A., Schneider J., On approximation of the Boltzmann equation by discrete velocity models, *C. R. Acad. Sci. Paris, Série I*, 320 (1995), pp. 639–644

[10] Bobylev A. V., Palczewski A., Schneider J., A consistency result for a discrete velocity model of the Boltzmann equation, preprint RW 95-05, IAMM, Warsaw University (1995); to appear in *SIAM J. Numer. Anal.*

[11] Bobylev A. V., Struckmeier, J. Implicit and iterative methods for the Boltzmann equation, *Transport Theory Statist. Phys.*, 25 (1996), no. 2, pp. 175–195

[12] Bogomolov S. V., Convergence of the total-approximation method for the Boltzmann equation, (Russian) *Zh. Vychisl. Mat. Mat. Fiz.*, 28 (1988), no. 1, pp. 119–126; translation in *U.S.S.R Comp. Maths. Math. Phys*, 28 (1988), no. 1, pp. 79–84

[13] Broadwell J. E., Shock structure in a simple discrete velocity gas, *Phys. Fluids*, 7 (1964), pp. 1243–1247

[14] Buet C., A discrete-velocity scheme for the Boltzmann operator of rarefied gas dynamics, *Transport. Theory Statist. Phys.*, 25 (1996), pp. 33–60

[15] Cabannes H., *The Discrete Boltzmann Equation: Theory and Applications*, lectures at univ. California, Berkeley (1980)

[16] Carleman T., *Problèmes Mathématiques dans la Théorie Cinétique des Gaz*, Uppsala (1957)

[17] Cercignani C., *Boltzmann Equation and its Applications*, Springer, NY (1988)

[18] Cercignani C., Illner R., Pulvirenti M. *Mathematical Theory of Dilute Gases* Springer, NY (1994)

[19] Cheremisin F. G., Numerical solution of the Boltzmann equation for one-dimensional stationary gas flows, (Russian) *Zh. Vychisl. Mat. Mat. Fiz.*, 10 (1970), no. 3, pp. 654–665; translation in *U.S.S.R Comp. Maths. Math. Phys*, 10 (1970), no.3, pp. 125–137

[20] Cheremisin F. G. Numerical methods for the direct solution of the kinetic Boltzmann equation. (Russian) *Zh. Vychisl. Mat. Mat. Fiz.*, 25 (1985), no. 12, 1840–1855; translation in *U.S.S.R Comp. Maths. Math. Phys*, 25 (1985), no. 12, pp. 156–166

[21] Aristov, V. V.; Cheremisin, F. G.; Shishkova, I. N. Unsteady reflection of a shock wave by a wedge for the Boltzmann equation. (Russian) *Zh. Vychisl. Mat. i Mat. Fiz.*, 36 (1996), no. 4, pp. 517–524; translation in *U.S.S.R. Comp. Math. Math. Phys.*, 36 (1996), no. 4, pp. 114–123

[22] Chorin A. J. Numerical solution of Boltzmann's equation, *Comm. Pure Appl. Math.*, 25 (1972), pp. 171–186

[23] P. Ciarlet, *Numerical Analysis of the Finite Element Method*, Les Presses de l'Université de Montréal (1976)

[24] Deshpande S. M., An unbiased and consistent Monte Carlo game simulating the Boltzmann equation, Report 78 FM 4, Indian Institute of Science (1978)

[25] Desvillettes L., Mischler S., About the splitting algorithm for Boltzmann and B.G.K. equations, *Math. Models Methods Appl. Sci.*, 6 (1996), no. 8, pp. 1079–1101.

[26] Gatignol R., *Theórie Cinétique des Gaz à Répartition Discrète de Vitesses*, Lecture Notes in Physics, 36, Springer (1975)

[27] Goldstein D. B., Sturtevant B., Broadwell J. E., Investigation of the motion of discrete-velocity gases, *Proc. 16-th Int. Symp on RGD, in: Ser. Progress in Astronautics and Aeronautics, vol. 118, ed. E.P. Muntz et al.* (1989), pp. 100–117

[28] Gruber P. M., Lekkerkerker C. G., *Geometry of Numbers*, North-Holland, Amsterdam (1987)

[29] Hicks B. L., Yen S. M., Collision integrals for rarefied gas problems, *Proc. 7th Int. Symp. on RGD*, vol. 2, p. 845–852, S. Nocilla, ed., Editrice Technico Scientica (1971)

[30] Illner R., Wagner W., A random discrete velocity model and approximation of the Boltzmann equation, *J. Statist. Phys.*, 70 (1993), no. 3-4, pp. 773–792.

[31] Illner R., Wagner W., Random discrete velocity models and approximation of the Boltzmann equation: conservation of momentum and energy, *Transp. Th. Stat. Phys.*, 23 (1994), no.1-3, pp. 27–38

[32] Illner R., Rjasanow S., Random discrete velocity models: possible bridges between the Boltzmann equation, discrete velocity models and particle simulation, in *Nonlinear Kinetic theory and Mathematical Aspects of Hyperbolic Systems*, World Scientific, Singapore (1992), pp. 152–158

[33] Inamuro T., Sturtevant B., Numerical study of discrete-velocity gases, *Phys. Fluids A*, 2 (1990), no.12, pp. 2196–2203

[34] Ivanov M. S., Rogazinsky S. V., Analysis of numerical techniques of the direct simulation Monte Carlo method in the rarefied gas dynamics, *Sov. Journ. Numer. Anal. Math. Modelling*, 3 (1988), no. 6, pp. 453–465

[35] Marchuk G. I., Splitting and alternating direction methods, in *Handbook of Numerical Analysis*, vol. I, P. G. Ciarlet and J. L. Lions eds., North-Holland, Amsterdam (1990), pp. 197–462

[36] Michel P., Schneider J., A quadrature formula for the Boltzmann equation, to appear in *SIAM J. Numer. Anal.*

[37] Mischler S., Convergence of discrete-velocity schemes for the Boltzmann equation, to appear in *Arch. Rat. Mech. Anal.*

[38] Monaco L., Preziosi L., *Fluid-dynamic applications of the discrete velocity models*, World Scientific, Singapore (1991)

[39] Nanbu K., Direct simulation scheme derived from the Boltzmann equation, I-III, *J. Phys. Soc. Japan*, 49 (1980), 5, pp. 2042–2058; IV, *J. Phys. Soc. Japan*, 50 (1981), 9, pp. 2829–2836

[40] Nordsieck A., Hicks B. L., Monte Carlo evaluation of the Boltzmann collision integral, *Proc 5th Int. Symp. on RGD*, C. L. Brundin ed., Acad. Press (1967) pp. 695–710

[41] Ohwada T., Structure of normal shock waves: Direct numerical analysis of the Boltzmann equation for hard-sphere molecules, *Phys. Fluids A*, 5 (1993), no.1, pp. 217–234

[42] Pulvirenti, M., Wagner, W., Zavelani Rossi, M. B.: Convergence of particle schemes for the Boltzmann equation. Preprint No.49, Inst. für Angew. Anal. und Stoch., Berlin (1994)

[43] P. A. Raviart, An analysis of particle methods, in *Numerical Methods in Fluid Dynamics*, Lecture Notes in Mathematics, 1127, Springer, (1984), pp. 243–324

[44] M. Reed, B. Simon, *Methods of Modern Mathematical Physics*, vol. I: *Functional Analysis*, Academic Press, NY (1972)

[45] Rjasanow S., Wagner W., A stochastic weighted particle method for the Boltzmann equation, *J. Comput. Phys.*, 124 (1996), no. 2, pp. 243–253.

[46] Rogier F., Schneider J., A direct method for solving the Boltzmann equation, *Transp. Th. Stat. Phys.*, 23 (1994), pp. 313–338

[47] Schmidt W. M., *Diophantine Approximation and Diophantine Equations*, Springer (1991)

[48] Sloan I. H., Joe S., *Lattice Methods for Multiple Integration*, Oxford University Press (1994)

[49] Sod G., A numerical solution of Boltzmann's equation, *Comm. Pure Appl. Math.*, 30 (1977), pp. 391–419

[50] Tan Z. Q., Chen Y.-K., Varghese P. L., New numerical strategy to evaluate the collision integral of the Boltzmann equation, *Proc. 16th Int. Symp on RGD, in: Ser. Progress in Astronautics and Aeronautics, vol. 118, ed. E.P. Muntz et al.* (1989) pp. 337–342

[51] Tan Z. Q., Varghese P. L., The $\Delta - \varepsilon$ method for the Boltzmann equation, *J. Comp. Phys.*, 110 (1994), pp. 327–340

[52] Truesdell C., Muncaster R. G., *Fundamentals of Maxwell's Kinetic Theory of a Simple Monatomic Gas : Treated as a Branch of Rational Mechanics* , Acad. Press, NY (1980)

[53] Wagner W., A convergence proof for Bird's Monte carlo method for the Boltzmann equation, *J. Stat. Phys.*, 66 (1992), no. 3-4, pp. 1011–1044

[54] Wagner W.: Approximation of the Boltzmann equation by discrete velocity models. Preprint No.96, Inst. für Angew. Anal. und Stoch., Berlin (1994)

[55] Wennberg B., private communication

[56] Yanenko N. N., *The Method of Fractional Steps*, Springer, N.Y. (1971)

[57] Yang J. Y., Huang J. C., Rarefied flow computations using nonlinear model Boltzmann equations, *J. Comp. Phys.*, 120 (1994), pp. 323–339

[58] Yen S. M., Lee K. D., Direct numerical solution of the Boltzmann equation for complex gas flow problems, *Proc. 16th Int. Symp on RGD, in: Ser. Progress in Astronautics and Aeronautics, vol. 118, ed. E.P. Muntz et al.* (1989) pp. 337–342