J. Indian Inst. Sci., May–June 2006, **86**, 169–192

Direct simulation Monte Carlo (DSMC): A numerical method for transition-regime flows–A review

P. S. PRASANTH* AND JOSE K. KAKKASSERY[†]
Department of Mechanical Engineering, National Institute of Technology Calicut, Kozhikode 673 601, Kerala, India.
*email: y3df07@nitc.ac.in; *Ph. No. 91-495-2286463.
[†]e-mail: josekkakkassery@nitc.ac.in; [†]Ph. No. 91-495-2286454.

Received on October 7, 2005; Revised on March 18, 2006.

Abstract

In fluid flow situations, the Navier–Stokes equations cannot adequately describe the transition and the free molecular regimes. In these regimes, the Boltzmann equation of kinetic theory is invoked to govern the flows. But this equation cannot be solved easily, neither by analytical techniques nor by numerical methods. Hence, in order to manoeuvre around this equation, Bird introduced a particle simulation technique called the 'Direct Simulation Monte Carlo' (DSMC) method. It is regarded as a numerical method for solving the Boltzmann equation. It has, over the years, become a very popular methodology for the simulation of fluids, where molecular effects are important. In this review, a brief description of the Boltzmann equation and its solution techniques are provided. Then the general DSMC procedure is explained. A brief summary of the advances, which have occurred on and due to DSMC, is given. A few limitations and constraints of the method are also briefly mentioned.

Keywords: Boltzmann equation, low-density flows, particle simulation, rarefied gas, transition regime.

1. Introduction

In fluid dynamics, when the average distance travelled by particles between collisions (the mean free-path, l) becomes comparable with or even larger than an important characteristic length (L) in the system, the continuum fluid approximation breaks down and the particle nature of matter must be taken explicitly into account. Such flows may be characterised by the Knudsen number (Kn). Often the local Knudsen number is used; L may be the length scale of a macroscopic gradient, $L = j / |\partial j / \partial x|$, where, j may be density, temperature, pressure or velocity. Use of local Knudsen number, instead of global quantity, avoids the confusion in the selection of L for large and complex systems. The effects of continuum breakdown are noticeable when $Kn \approx 0.1$ or even less [1].

$$Kn = \frac{l}{L}. \tag{1}$$

While there are no fixed boundaries, it is generally accepted that flows with Knudsen numbers less than 0.01 are considered to be continuum flows. It is also generally accepted that

---

*Author for correspondence.

free molecule flow may be assumed for Knudsen numbers larger than 10. The Knudsen number region between 0.01 and 10 is known as transitional regime, where neither the continuum nor a free molecule assumption would be appropriate. In this regime, the Navier–Stokes equations yield poor approximations to the physics of gas dynamics. The reason being that the constitutive equations that relate the shear stress and heat transfer to other variables break down. More specifically, the linear transport terms for mass, diffusion, viscosity, and thermal conductivity in the partial differential equations are no longer valid when the average distance moved by particles between collisions is comparable to typical length scales in the flow geometry. For such flows, the governing equation is the Boltzmann equation of kinetic theory.

A common example of flow regime, where the Boltzmann equation would be necessary, is the operation of flight vehicles in the upper atmosphere. This complicated problem involves catalytic reactions at the surface, internal degrees of freedom for molecules interacting with the surface, and other factors. While this equation may be appropriate for low-density flows, attempts to solve it numerically have met with much less success. To further complicate matters, these flows also tend to exhibit a large degree of chemical and thermal non-equilibriums. This can be attributed to the limited number of molecular interactions that occur within the local domain. Another prominent example would be the analysis of rarefied gas environments for orbiting platforms like satellites and space stations. In the aerodynamic and aerothermodynamic environments generated by the firing of control thrusters of orbiting platforms, although the flow in the local vicinity of the thrusters can typically be characterised as being continuum, its rapid expansion into vacuum of space negates the use of continuum-based solvers. Similar issues also exist for problems of interest in the semiconductor industry that take place in low-density environments, such as plasma etching and chemical vapour decomposition [2].

2. The Boltzmann equation

The Boltzmann equation describes the flow in all the regimes, continuum, continuum-transition and free molecular. It is an integrodifferential equation proposed by Boltzmann in 1872. Limiting cases of this equation yield the continuum description, commonly based on the Euler or Navier–Stokes equations, at the extreme of the small mean free paths and the collisionless (or free molecular) flow at the extreme of large free paths [3].

The Boltzmann equation for a simple dilute gas is [4]

$$\frac{\partial}{\partial t}(nf)+c.\frac{\partial}{\partial r}(nf)+F.\frac{\partial}{\partial c}(nf)=\int_{-\infty}^{\infty}\int_{0}^{4p} n^{2}(f^{*}f_{1}^{*}-ff_{1})c_{r}s\,\mathrm{d}\Omega\mathrm{d}c_{1} \qquad (2)$$

where (nf) is the product of number density (n) and the velocity distribution function (f). In this equation, c is the molecular velocity, c_r, the relative molecular speed, F, the external force per unit mass, the superscript (*) indicates post-collision values, f and f_1 represent the distribution functions of two different types of molecules of class c and c_1, respectively, s, is the collision cross-section, t represents time, r the physical space, and Ω represents solid angle.

In a gas mixture consisting of a total of *s* chemical species, a separate distribution function must be defined for each species. The Boltzmann equation then becomes a set of *s* simultaneous equations. If two particular species are represented by subscripts *p* and *q*, the Boltzmann equation for species *p* of the mixture may be written as,

$$\begin{aligned}&\frac{\partial}{\partial t}(n_p f_p)+c_p.\frac{\partial}{\partial r}(n_p f_p)+F.\frac{\partial}{\partial c}(n_p f_p)=\\&\sum_{q=1}^{s}\int_{-\infty}^{\infty}\int_{0}^{4p} n_p n_q(f_p^* f_{1q}^* - f_p f_{1q})c_{r_{pq}} s_{pq}\,\mathrm{d}\Omega\,\mathrm{d}c_{1q}.\end{aligned} \tag{3}$$

The right-hand side of the Boltzmann equation is called the collision term and is the source of problems in finding a solution.

2.1. *Solution techniques*

2.1.1. *Analytical techniques*

Problems, which involve large disturbances, are invariably nonlinear and there is then no analytical method that does not incorporate assumptions or approximations. The small perturbations approach may be based on the smallness of the Knudsen number or its reciprocal in order to obtain approximate solutions for near-continuum and near-collision-less regimes, respectively. Some methods are applicable only in situations where the flow disturbances are negligibly small. In particular, if the disturbance is sufficiently small that the velocity distribution function is perturbed only slightly from the equilibrium or Maxwellian form, the small-perturbation approach leads to the linearised Boltzmann equation. However, solutions are available only for a restricted set of problems with comparatively simple geometry and boundary conditions. These have limited applications to problems that are encountered in practice. However, for nonlinear problems that involve large perturbations and complex physical effects, analytical solutions are out of question and recourse must be taken to numerical methods.

A large number of the solution methods depend upon essentially arbitrary assumptions on the form of the distribution function and most of these can be classified as moment methods. The method of moments approach employs the moment equations that are obtained by first multiplying the Boltzmann equation by the molecular quantity Q (either the mass m, momentum mc or energy $\frac{1}{2}mc^2$ of a molecule) and subsequently integrating it over the velocity space. The moment equation for Q is of the form,

$$\frac{\partial}{\partial t}(n\bar{Q})+\nabla.(n\overline{cQ})-nF.\overline{\frac{\partial Q}{\partial c}}=\Delta[Q] \tag{4}$$

where the collision integral $\Delta[Q]$ is given as $\Delta[Q]=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{0}^{4p} n^2Q(f^* f_1^* - ff_1)c_r s\,\mathrm{d}\Omega\,\mathrm{d}c_1\mathrm{d}c$. Equation (4) is also called the transfer equation or the equation of change. The moment methods derive the higher-order fluid dynamics approximations beyond Navier–Stokes equations to account for departures from thermal equilibrium. The higher-order fluid dynamic models are known as the extended hydrodynamic equations (EHE) or the generalised hydrodynamic equations (GHE). The ‘Burnett equations’ represent a second-order departure from the equilibrium while the Navier–Stokes equations represent a first-order departure from thermal equilibrium. Higher-order approximations beyond Burnett equations are called the ‘super Burnett equations’. The higher-order extended or generalised hydrodynamic equations perform equally well in continuum and continuum-transition regimes. However, at present, the complexity of the highly nonlinear Burnett stress and heat-transfer terms itself is enormously challenging both computationally and in terms of understanding the physics.

The ‘model equations’ method is another area of major approximations, where the form of the Boltzmann equation itself is altered. The major mathematical difficulties posed by the Boltzmann equation are associated with the collision term and it is this term that is modified. The best-known model equation is called the BGK equation (5) after Bhatnagar, Gross, and Krook (1954).

$$\frac{\partial}{\partial t}(nf)+c.\frac{\partial}{\partial r}(nf)+F.\frac{\partial}{\partial c}(nf)=nn(f_0-f). \tag{5}$$

The parameter *n* is generally regarded as a collision frequency. Though it is assumed to be independent of velocity *c*, it has a restricted functional dependence, as it is proportional to density and may depend on temperature; f_0 is the Maxwellian distribution function. The BGK equation provides the correct collision-less or free molecule solution, since the form of the collision term is then immaterial. However, the approximation to the collision term would lead to an indeterminate error in the transition regime [4].

Numerical analysis is often required to obtain final results in the moment and model equation methods.

2.1.2. *Numerical simulation techniques*

While the analytical methods are almost invariably based directly on the Boltzmann equation with an accepted mathematical model of a gas flow at the molecular level, a large fraction of the numerical methods are based directly on the simulation of the physics of the flow and do not depend on the Boltzmann equation. The numerical methods, which are not based on either an assumed form of the distribution function or on a modification of the Boltzmann equation, may be divided into those methods, which take the Boltzmann equation as their starting point—like the Finite Pointset Method [5], Boltzmann CFD, etc.—and methods which are based directly on the molecular description provided by kinetic theory.

The most direct numerical approach to the Boltzmann equation would be to construct a conventional finite-difference formulation. The distribution function is the only dependent variable for a simple gas and this may be represented by its numerical value over a network of points in phase-space. But, a finite-difference approach would require a prohibitive amount of computer time and storage capacity for all but the simplest of problems. Despite these difficulties, satisfactory results were obtained for a number of problems; most notably the structure of normal shock waves. The Boltzmann CFD approach has also been extended to two-dimensional steady flows. The basic disadvantage of the Boltzmann CFD methods is the necessity to specify a bounded grid in velocity space. The Boltzmann equation must be accompanied by boundary conditions which describe the interaction of gas molecules with the solid walls. It is to this interaction of gas molecules with solid walls that one can trace the origin of the drag and the lift exerted by the gas on the body and the heat transfer between the gas and the solid boundary.

However, progress in computer hardware and improvements in numerical methods have made solution of the Boltzmann equation for rather complex gas dynamics problems real. Solution of the Boltzmann equation was initiated when the first powerful computers appeared. In these methods, the velocity-space was divided into cells of equal size and the collision integrals were evaluated either at the nodes of the grid or at the midpoints of the cells, employing Monte Carlo technique. However, a deficiency, namely, the nonfulfilment of conservation laws in the computed integrals, became apparent. Later, a splitting finite-difference scheme for the kinetic equation was proposed, and a special correction was developed to satisfy the conservation laws. The new method showed itself much more efficient, but the correction introduced some additional numerical viscosity, and required artificial assumptions in the case of gas mixtures [6]. The first successful method for the direct numerical solution of the Boltzmann equation was introduced in a series of papers by Nordsieck and Hicks [7] and Yen [8]. They dealt with one-dimensional steady flow problems and employed a conventional finite-difference technique for the ‘fluid-like’ terms on the left-hand side of the Boltzmann equation.

A discrete ordinate or discrete velocity method was introduced by Broadwell [9]. The discretization of velocity-space allows the Boltzmann equation to be replaced by a set of nonlinear hyperbolic differential equations. The number of equations in this set is equal to the number of discrete velocities, and the collision term involves a double summation over the velocity points. The solution of this set is formidable task and applications have been limited to simple geometries and simple molecular models.

The direct solution of the Boltzmann equation for the velocity distribution functions does not constitute the sole means of obtaining solutions to transition regime flows. Advantage may be taken of the particulate nature of the gas to produce a direct simulation model on a digital computer. An additional difficulty for traditional numerical methods is the fact that chemically reacting and thermally radiating flows (and even simpler flows of polyatomic gases) are hard to describe with theoretical models having the same degree of accuracy as the Boltzmann equation for monatomic nonreacting and nonradiating gases. These considerations paved the way to the development of simulation schemes. All methods that employ simulated particles or molecules may be classified as simulation methods. The key idea of particle simulation is rather simple, that is, to recall the underlying physical reality of a gas. The idea of particle simulations is really not just to return to the particle level, but to restrict the number of molecules to a manageable level. The molecular dynamics (MD) method [10], applicable only to dense gas flows, was the first physical simulation method.

In the simulation methods, the gas is represented by the positions and velocity components of a very large number of simulated molecules. While the distribution function may be established from the molecular information, the problem is formulated without reference to it. In the transition regime, it becomes necessary to compute typical intermolecular collisions in addition to the molecule–surface interactions. Each time a single molecule is considered, all other molecules must be examined as possible collision partners. This means that the computation task is proportional to square of the number of molecules. The computing requirement becomes manageable if probabilistic rather than deterministic procedures are adopted for the computation of collisions. The first probabilistic simulation method was introduced by Haviland and Lavin [11] and is best termed as the test particle Monte Carlo method. Further developments started with the work of Bird [12, 13] on the direct simulation Monte Carlo (DSMC) method, which is applied only to dilute gas flows. Hammersley and Handscomb [14] state that Monte Carlo methods comprise that branch of experimental mathematics which is concerned with experiments on random numbers. They also applied the term ‘direct simulation’ to the modelling of probabilistic problems, which is the simplest form of the Monte Carlo methods. The term ‘direct simulation Monte Carlo’ is therefore a generic term that covers the many variants of the method that have been introduced over the years.

Bird’s DSMC method, perhaps, can be considered as the most accurate and widely used technique for the computation of low-density flows. Given the physical basis of the DSMC method, the existence, uniqueness, and convergence issues that are important in the traditional mathematical analysis of equations are largely irrelevant. There is no stability criterion that is analogous to the Courant condition for finite-difference methods in continuum CFD. It has become a powerful tool for practical calculations. There appear to be very few limitations to the complexity of the flow fields that can be handled by this approach. Chemically reacting and ionised flows can be and have been analysed by this method. It may be regarded as a numerical method for solving the Boltzmann equation. It has been shown that its solutions converge to Boltzmann equation solutions in the limit of infinite number of particles, and vanishing cell size and time-step [15].

3. The DSMC method

The DSMC method is a technique for the computer modelling of a real gas flow by simulated molecules whose number has ranged from the order of hundreds in early calculations to millions in some contemporary simulations. The DSMC method employs simulated molecules of correct physical size and their number is reduced to a manageable level by regarding each simulated molecule as representing a fixed number of real molecules. This is typically a large number and advantage can sometimes be taken of its size to specify some properties through a distribution of values instead of a single value for each simulated molecule. The velocity components, position coordinates and internal states of simulated molecules are stored in the computer and are modified with time, as the molecules are concurrently followed through representative collisions and boundary interactions in simulated physical space. The velocity-space information is contained in the positions and velocities of simulated molecules. These positions and velocities are continuously distributed and the velocity-space is unbounded. The conservation of mass, momentum, and energy are imposed to within the roundoff error of the compiler. Bird [4] and Cercignani [3] dwell in detail on the procedures in applying the DSMC to steady and unsteady flows. The flowchart for a typical DSMC simulation is given in Fig. 1 [16].

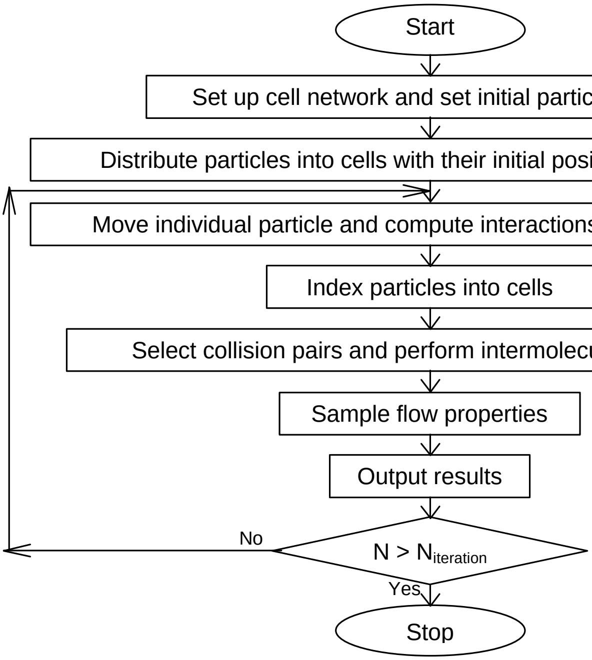

Fig. 1. DSMC flowchart.

The primary approximation of the DSMC is to decouple the molecular motions (modelled deterministically) and the intermolecular collisions (modelled probabilistically) over small time intervals. The time interval over which the solution is sought is subdivided into small subintervals over which the particle motions and collisions are decoupled. Though the computational time is directly proportional to the number of simulated molecules (N), the time parameter in the simulation may be identified with the physical time in the real flow. The main computational approximations associated with the DSMC method are the ratio of the number of real molecules to the number of simulated molecules (particle weight), the time-step and the finite cell and subcell sizes in the physical space. The selection of a time-step that is less than the mean collision time is of primary importance. If the time-step is too large, the molecules will be able to move too far without the opportunity to participate in a collision. This will cause a ‘smearing’ of properties of the flow, resulting in fictitious, nonphysical results. Execution of the method requires the use of a cell or spatial element network in physical space. The cell network provides geometric areas and volumes required to evaluate macroscopic flow properties. It is also used by the collision process model, in which only particles located within the same cell, at a given time, are allowed to interact. In the strictest sense, the size of the flow-field cells should relate to the mean free-path of the local molecules. This stands to reason, since the cells are used to group molecules into candidate collision pairs. If molecules, which are separated by distances much larger than the mean free-path, are allowed to collide, it would allow a nonphysical transfer of mass, momentum and energy. In a typical application of the direct simulation method, a cell has dimensions ($l/3$) and contains 30 molecules. This corresponds to the dilute gas assumption.

Excessive temporal and spatial discretization should also be avoided, as this will degrade the performance of the method.

A DSMC simulation proceeds from a set of prescribed initial conditions. The particle positions, velocities, collision cross-sections, and boundary conditions determine the subsequent evolution of the system. The initial states of the gas and the macroscopic boundary conditions such as vacuum and uniform equilibrium flow are prescribed. The flow develops from this initial state with time in a physically realistic manner rather than by iteration from an initial approximation to the flow. In this technique, molecules constantly enter into and leave the computational domain. The properties of entering molecules are specified by known boundary conditions.

The core of the DSMC algorithm consists of four primary processes: move the particles, index and cross-reference the particles, simulate collisions, and sample the flow-field. The molecular motion and the intermolecular collisions are uncoupled over the small time interval–the discrete time-step. In the first process, all the molecules are moved through distances appropriate to their velocity components and the discrete time-step. Appropriate action is taken if the molecule crosses boundaries representing solid surfaces, lines or surfaces of symmetry, or the outer boundary of the flow. As for the boundary conditions, the microscopic boundary conditions are specified by the behaviour of the individual molecules according to the conservation laws, rather than in terms of the distribution function. Collisions with surfaces can be treated as being either fully specular, fully diffuse, or some combinations of the two. Specular collisions involve a simple reversal of the molecular velocity component normal to the incident surface. Diffuse collisions cause a random reorientation of the reflected molecule, where the post-collision velocity is related to the temperature of the surface. There are also more complicated surface interaction models, where the orientation of the reflected molecule is also a function of the incident velocity. Maxwell’s model of diffuse reflection is adequate for many problems. There are many cases, however, in which it is far from adequate. The Cercignani–Lampis (CL) model [17] has been adapted and extended by Lord [18, 19] for application in DSMC studies. The resulting CLL model [18, 19] has been shown to provide realistic boundary conditions. More complicated models would be required to describe the chemical reactions that can occur at the surface at high impact energies. More complex surface interactions, such as surface catalysis, are easily included in the DSMC methodology.

The second DSMC process involves indexing and tracking the particles. A scheme for molecular referencing is the prerequisite for the next two steps: modelling the collisions and sampling the flow field.

The third step, simulating the collisions, is a probabilistic process that sets DSMC apart from deterministic simulation methods such as molecular dynamics. The probability of a collision between two molecules in a homogenous gas is proportional to the product of their relative speed and total collision cross-section. The mean value of the product is calculated for each cell, and the maximum value is also recorded. The collision pairs could then be chosen by the ‘acceptance-rejection’ method; the probability of a particular pair being chosen depends on the ratio of their product to the maximum product. However, this procedure would have a computation time directly proportional to the square of the total number of molecules in the cell. So the ‘no-time-counter’ technique was introduced [4] in order to obtain a computation time directly proportional to the number of molecules. When used in conjunction with the subcell technique, it ensures that collisions occur only between near neighbours and thereby improves the accuracy of the simulations. An appropriate number of representative collisions between randomly selected pairs of molecules within each cell is simulated.

The final process is sampling the macroscopic flow properties. The spatial coordinates and velocity components of molecules in a particular cell are used to calculate macroscopic quantities at the geometric centre of the cell.

The above-mentioned stages correspond to the solution of the Boltzmann equation (2) using the following splitting scheme within the short time-step [20].

$$\frac{\partial}{\partial t}(nf)+c.\frac{\partial}{\partial r}(nf)+F.\frac{\partial}{\partial c}(nf)=0 \tag{6}$$

for the collision-less movement of molecules in the first step, and

$$\frac{\partial}{\partial t}(nf)=\int_{-\infty}^{\infty}\int_{0}^{4p} n^2(f^* f_1^* - ff_1)c_r s\,\mathrm{d}\Omega\mathrm{d}c_1 \tag{7}$$

corresponds to the molecular collisions in the third step.

The DSMC procedure is explicit and time marching and it always produces a flow simulation that is unsteady. For an unsteady flow application, an ensemble of many computations may be assembled and averaged to obtain final results with an acceptable statistical accuracy. To simulate a steady problem, each independent computation proceeds until a steady flow is established at a sufficiently large time, and the desired steady result is time average of all values calculated after reaching the steady state. As a steady flow gets developed through a physically real unsteady flow, there are minimal problems with boundary conditions. There is no need for an initial estimate of the flow-field and there is no artificial iteration process. The DSMC, rather than solving an average description of macroscopic flow properties represented by partial differential equations, deals with gas flows at a molecular level. This allows the inclusion of chemical reactions and nonequilibrium effects into the methodology in a direct and physical manner. This has made the DSMC technique a very popular methodology for the simulation of fluids, where molecular effects are important.

Unlike other variants, where molecules are just conceptual elements, Bird’s method is a genuine particle method. It has been shown that this DSMC algorithm can be derived from the Leontovich mathematical model by introducing certain assumptions [21]. Conversely, it has been shown by Bird that the Boltzmann equation can be derived through the DSMC procedures.

4. Limitations of DSMC

The inherent constraints on the DSMC method are: 1) the cell size must be proportional to the local mean free-path, 2) the number of particles per cell must be roughly constant in order to preserve collision statistics, and 3) the simulation time-step must be chosen so that particles only traverse a fraction of the average cell length per time-step.

4.1. *Computational errors*

The DSMC is very attractive because it successfully ‘coarse grains’ (each simulated molecule represents several real molecules) the molecular description of the hydrodynamic regime, thus offering substantial computational efficiency over the ‘brute force’ (deterministic) molecular dynamics simulations. However, like most computational methods, in the limit of finite discretization both in space and time, numerical errors do contaminate the solution [22]. As a numerical and stochastic method, DSMC has numerical and statistical errors, which have to be quantified to verify the accuracy of the obtained solution.

The two types of errors arising in this method are: 1) a statistical error, which corresponds to the stochastic nature of the method, and 2) a deterministic error, which depends on the selection of numerical parameters such as the time-step, cell volume, and the total number of simulated particles. The source of the statistical errors is the process of taking a sample average in a finite volume and the use of a simulated molecule to represent a large number of actual molecules. The errors have an asymptotic behaviour and vanish when the sample size becomes infinitely large. Thus, these errors are inversely proportional to the sample size, and therefore they can be decreased by increasing the sample size, that is, the number of particles in a computational domain or number of time-steps used in the sample. There will be significant statistical ‘noise’ in an instantaneous sample. Successive sampling is done to increase the sample size and thus reduce the statistical ‘noise’. A weak disturbance can be completely lost in the scatter in an instantaneous sample or during the earlier stages of a simulation, but, it emerges correctly as the average builds up. The deterministic errors occur, as the numerical parameters cannot be taken as infinitesimally small. These errors are not statistical in nature, and therefore cannot be reduced by using a large sample size. The parameters, time-step and cell volume should be small enough compared to the characteristic time and length of the flow, which are physical characteristics of the specific flow. In fact, this type of error may have a greater impact on the results than the statistical error, especially when a large sample size is used, and hence the stochastic (statistical) error can be neglected [23].

Chen and Boyd [24] employed a root mean square (rms) error as an indicator of the statistical fluctuations. An equation is developed to model and analyse the rms errors. A range of the appropriate number of particles to be employed in the simulation is provided in order to achieve the smallest statistical error at a fixed computational cost. It is found that statistical errors depend on the models employed for intermolecular interaction. The errors are the smallest for hard sphere interaction, as this model produces the highest collision rate. This suggests that statistical errors may be minimised by selecting numerical parameters that allow the largest number of collisions to be simulated while preserving the physical validity of the computation. Computational efficiencies are also defined to measure the quality of a simulation with respect to computational cost and rate of reduction of statistical scatter. The larger the value of the computational efficiency, the more rapid will be the decrease in the rms error.

Truncation error due to discretization in space results from the need to select collision partners from cells of finite size. The resulting error is proportional to the square of the cell size; the constant of proportionality is such that the cell sizes of the order of one mean free-path results in errors of the order of 10%. Garcia and Wagner [25] show that the time-step truncation error in DSMC calculations is found to be proportional to the square of the time-step for a variety of simple flows, both transient and steady state. The time-step size exerts a more profound influence on simulation results, if the number of simulated molecules is small. Further, in his paper, Hadjiconstantinou [22] points out that the DSMC commits a 'discretization crime' by performing an instantaneous collision of molecules that would have normally collided within a time-step.

4.2. *Computational constraints*

The run-time and storage requirements are the two main computational limitations for DSMC simulations. Rieffel [26] presents a model, independent of the simulation architecture, girding technique, collision model and implementation technique, for predicting these requirements for the simulation of rarefied gas flows.

A problem that arises with probabilistic near-neighbour collisions is that the probability of the collision partner of a particular molecule being the same as one of those, which were involved, in an immediately preceding collision is high. This is physically impossible and would not occur in a deterministic simulation. The hypothesis of molecular chaos is valid only if the probability of the repeated collision of a colliding pair of molecules is negligibly small. Since the DSMC method employs a finite number of simulated molecules, the number of possible collision pairs in a cell is rather limited, and the probability of a repeated collision is fairly high. Therefore, there is a statistical dependence between the simulated molecules. A significant level of the statistical dependence means that the molecular chaos hypothesis used in the Boltzmann equation becomes invalid. The magnitude of statistical dependence can be estimated by calculating the fraction of repeated collisions. The smaller this parameter, the smaller is the statistical dependence between the simulated molecules. Computational results show that an increase in the time-step leads to an increase in the statistical dependence between the simulated molecules and the error associated with this dependence. The larger the particle weight (the ratio between the number of physical molecules and the computational particles), the greater is the growth of the statistical dependence between the simulated molecules with increasing time-step. The effect of repeated collisions on the computation results was examined [27] for classical problems of rarefied gas dynamics, which reveals the main features of the complicated rarefied gas flows. The studies showed that the fraction of repeated collisions is a reliable indicator of deviation of simulation results from the solution of the Boltzmann equation. The maximum limit of the fraction of repeated collisions, for which the contribution of statistical correlations to macro-parameter estimates is acceptable, has been found to be 20%. If the fraction of repeated collisions is higher than 50%, the simulation results have been found to be significantly different from the solution of the Boltzmann equation.

All computational methods are affected by the 'roundoff' error as a consequence of the representation of numbers in the computer by a finite number of binary bits and this limited word length places restrictions on the precision of computer arithmetic. The probabilistic simulation methods, such as the DSMC method, are also affected by the deviation of the computer-generated random numbers from the ideal distribution. Just as the finite length representation of numbers in a computer does not have all the properties of real numbers, the numbers from a random number generator on a computer do not conform to the properties of an ideal set of random numbers. Such a set is necessarily hypothetical; an ideal random number generator does not exist, and the computer-generated numbers are sometimes called pseudo-random numbers. A typical DSMC application involves the generation of hundreds of millions of random fractions so that those generators with very short periods are not desirable. However, the numbers are used in many procedures and the number of calls of the random number routine within each procedure is also subject to random variations. This means that most DSMC calculations are very tolerant of imperfections in the random number generator.

4.3. *Other constraints*

Subtle problems surface while simulating real gas effects like translational internal energy transfer and chemical reactions. The probabilities of internal energy transfer and chemical reactions are determined as the ratios of the corresponding cross-sections to the total collision cross-section. The cross-sections of vibration-translation energy transfer as well as chemical reactions increase significantly with the gas temperature. This may lead to the situation when the sum of all of these probabilities becomes larger than unity as the gas temperatures rise sufficiently high. Whereas the total inelastic probability exceeding unity may visibly distort the individual rates, the problem is in fact more complicated since both elastic and inelastic cross-sections are unknown at very high temperatures. Another aspect of the problem is related to the variable hard sphere (VHS) [28] and variable soft sphere (VSS) [29, 30] models typically used in the DSMC method, where after-collision molecular velocities are modelled according to the hard sphere law. This means that gliding collisions with small deflection angles are replaced by more efficient but fewer collisions with high deflection angles. This in turn increases the inelastic probabilities when the VHS model is used, as compared to the inverse power law interaction potential. Therefore, the problem with the total inelastic probability exceeding unity may only be avoided by using better reaction rates and employing a more realistic molecular potential with a small enough deflection angle cutoff.

A numerically optimal calculation is one in which the desired physics is resolved using a minimum number of particles. In a typical problem, the greatest amount of physics (collisions) and chemistry occurs in high-density regions. In order to model the flow correctly, it is necessary to resolve flow physics occurring in these regions. The particle weight for the calculation must be set to a low number so that there will be a reasonable number of particles in these high-density regions. As a result, the more rarefied portions of the flow-field will have a large number of particles and will be over-resolved. More computational time is spent for calculating the low-density regions than is needed, which results in less efficient calculations [31]. There are various ways in which the particle count distribution can be adjusted to approach the optimal case. In addition to particle weight scaling, variable time-step, cell sizes and shapes can be selected so as to improve the distribution. A combination of grid manipulation and variable time scales will often be the best means of controlling the distribution of particles.

The fluctuations inevitably occurring in a DSMC calculation can cause problems if the number of molecules is not large enough to representatively sample each area of the flow domain. The comparisons between the fluctuations in DSMC and the predictions of fluctuating hydrodynamic theory appear to be consistent with the fluctuations in a real gas. The most serious statistical problem occurs when a significant effect in the real gas is a consequence of the few molecules towards the extremes of the distribution. The number of simulated molecules may be so small in comparison with the real molecules, so that the important part of the distribution may not be adequately populated. This problem arises in connection with chemical reactions and thermal radiation.

5. Advances in DSMC

5.1. *Developments in algorithms and software*

The major advances in DSMC have been mainly in terms of innovative algorithms and in the direction of architecture-specific programming.

The Lagrangian DSMC [32] is an effort to expand the scope of DSMC to allow more grid adaptation including moving internal and external boundaries and involves combining DSMC with monotonic Lagrangian grid (MLG). The MLG is a method of constructing and maintaining data structures for large particle simulations, which are optimal for usage on parallel and vector computers. The combination of DSMC and MLG has several significant benefits, namely, an automatically adapting grid, improved prediction accuracy, reduced user effort and decreased computational requirements through efficient parallelisation. The MLG adds a time-varying grid that automatically adapts to the local number densities in the flow-field.

For the implementation of DSMC technique, it is advantageous to strictly separate physical modelling, geometrical calculations and organisational tasks, and to define communication interfaces between these modules. The ‘MONACO’ system [33] was developed with strong emphasis on using workstations as primary calculation platforms. A major design feature of MONACO is the separation of physical modelling, geometrical issues and managerial tasks. A kernel library handles the organizational tasks associated with a DSMC simulation, such as correct assignment of particles to their cell data structure; the geometric library handles such tasks as the movement of particles through the grid; and the physics library selects collision pairs and applies appropriate physical models to them.

The DSMC method is physically appropriate for rarefied gas flows in micro-electro-mechanical systems (MEMS). However, it is difficult for it to be statistically convergent, because the statistical fluctuations become insurmountably large at low Mach numbers that are characteristic of MEMS. For the purpose of variance reduction, a modification of DSMC called stochastic weighted particle method (SWPM) is proposed. In SWPM, a system of weighted particles is used, which allows one to resolve the low-density regions with a moderate number of simulation particles [34, 35]. A convergence proof for SWPM is proposed by Matheis and Wagner [36]. An information preservation (IP) technique [37–39] is also proposed to address this issue, which assigns two velocities to each simulated molecule. One is the molecular velocity used to compute the molecular motion following the same steps as the DSMC method and the other is called the information-velocity, which corresponds to the collective velocity of an enormous number of real molecules that the simulated molecule represents. The basic premise of this technique is to associate both microscopic and macroscopic information with each particle. The usual DSMC steps are applied to the microscopic properties and simple conservation laws are applied to the macroscopic properties. Meaningful results are obtained with a sample size, which is four orders of magnitude smaller than that required for ordinary DSMC. Therefore, there is tremendous reduction in CPU time.

Parallelisation of DSMC [40] has led to most significant reductions in computational time for large systems. To parallelise a code efficiently, a number of different aspects of the process including program control, data control, and programming tools must be considered. The program control or the program structure is either in data-parallel or task-parallel mode. The data-parallel mode, which is the mode usually used on massively parallel computers, is the one in which the same operations are performed simultaneously on different sets of data. In the task-parallel mode, individual processors run completely different operations on different data. Data control is the process by which data in one processor-memory unit is accessed by a second processor-memory unit. The process is usually carried out by using message-passing libraries, such as EXPRESS, PVM, and MPI. Programming tools include the programming language and various parallelising compilers. One major factor that hampers the parallelisation of DSMC code is that it requires a great deal of inter-processor communication, which is relatively slow. A second factor hampering parallelisation is load balancing among the processors. Load balancing keeps all the processors active and doing useful work for as much time as possible. For all parallel computers, a dynamic load-balancing scheme, which controls all working processors and minimises idle time, reduces interprocessor communication and so improves the performance.

SMILE [41] is a modern general DSMC program that has a number of users and a history of well-validated applications. The important features of SMILE are parallel capability, different collision and macroparameter grids with manual and automatic adaptations, and spatial weighting for axisymmetric flows.

The DSMC analysis code (DAC) software represents the National Aeronautics and Space Administration's (NASA) state-of-the-art implementation of DSMC method for simulating three-dimensional flows [42]. The DAC program employs 'virtual subcells' as an alternative to transient subcells. The distances from the first molecule to all the other molecules in the cell are determined and the nearest one is selected. This is a simpler procedure, but the computation time is proportional to the square of the number of molecules in the cell. In this, a number of significant processes of DSMC are automated. A prime example of this is the DAC preprocessor. The preprocessor is capable of automatically generating a flow-field discretization of the domain that satisfies the conditions necessary for an accurate simulation. The grid systems for the flow-field and the surface are uncoupled. The preprocessor takes special care of the Cartesian cells that intersect with the surface geometry.

The development of a general parallel three-dimensional direct simulation Monte Carlo code (PDSC) is reported by Wu *et al.* [43]. The PDSC is implemented on the memory-distributed parallel machine (PC-cluster system) using message passing interface with several important features, including tetrahedral unstructured mesh, solution-based adaptive mesh refinement, variable time-step method, dynamic load-balancing technique, extended conservative weighting scheme for trace species, very low-speed flows, iterative pressure boundary treatment method and hybrid NS-DSMC method. In addition, a graphic user interface using Visualisation Toolkit Library is developed to enhance the possibilities of including more features into the code. Also, a database of VSS collision model parameters for various species of interest in materials processing is included in PDSC.

A hybrid numerical technique is being developed that may employ either the CFD or DSMC techniques throughout the flow-field, based on local flow conditions. The primary issues are: 1) when to switch between the two methods, 2) how to pass information from one method to another, and 3) how to implement these processes in a single computer code that runs effectively on vector and massively parallel computer architecture. The study of continuum breakdown is the first step towards the development of hybrid numerical code. A parameter $(Kn)_{\text{GLL-D}}$ (gradient-length local Knudsen number), a Knudsen number determined by a mean-free-path, based on the local flow properties and a length scale based on the local gradient in density, was proposed by Boyd *et al*. [44] that successfully predicts failure of the continuum Navier–Stokes equations in hypersonic transitional flow. It was found that the continuum approach broke down whenever the value of $(Kn)_{\text{GLL-D}}$ exceeded 0.05. Wang and Boyd [45] proposed a new parameter, $Kn_{\max}$, that is modified from Boyd’s $(Kn)_{\text{GLL-D}}$ parameter and the continuum breakdown is best predicted when $Kn_{\max}$ exceeds 0.05. When the parameter $Kn_{\max}$ is utilised in the coupled DSMC-CFD computations, it is expected that the entire flow-field will approximately be separated into four layers. In the regions very close to the body surface, the DSMC technique will be employed and in between the near-body region and the shock region, the CFD technique will be used. The DSMC technique will be employed again throughout the shock. In the free-stream region, it is efficient to utilise the CFD technique.

New models of the DSMC method are being developed for a more detailed description of physical and chemical processes in gases. Their implementation into the existing DSMC codes, however, is rather difficult. The promising way of overcoming these difficulties is to develop object-oriented DSMC codes that can be readily modified to include the new models.

5.1.1. *The ideal program*

It is suggested by Bird [46] that the capabilities of the existing and any new general DSMC program should be evaluated against the following specifications for an ‘ideal program’.

The capabilities of the program should be such that it is easily applied to any problem, which falls within the computational capabilities of contemporary computers. The range of application of a program in terms of the lower Knudsen number limit will always be restricted by the performance of the computers that are available. Advances are likely to be increasingly dependent on parallel computation rather than on increase in CPU speed and hence the program should be capable of parallel operation. It should be capable of handling problems, which involve internal or external flows in one-, two- or three-dimensions, and either large-time steady flows, or time-accurate unsteady flows. The program should incorporate a range of realistic models for intermolecular collisions, internal degrees of freedom, gas-phase chemical reactions, and gas–surface interaction and reactions. It should be possible to simulate the effects of gravitational fields and similar body forces. The initial conditions should allow the specification of flow discontinuities and the boundaries should allow for periodic flows, constant pressure boundaries, and moving surfaces.

The programs should reproduce the solutions for all simple test cases that are amenable to exact analysis. The results from the program should be in agreement with any valid experimental measurement. Not only should the program automatically set the computational parameters to near optimal values, it should be able to internally check the quality of computation. The probable error in the results should be reported to the user. Also, should the quality of the calculation be outside the specified limits, the program should automatically abort the calculation.

The program should have a modern interactive graphical user interface (GUI). The data input should be through a fully specified file that can be generated by interactive menus within the program. There should be a user manual with representative demonstration cases. There should also be a help system within the menu structure that provides the necessary information on the data items. The computational details or the log of the calculation, the flow contours, and the distribution of the surface properties should be displayed continuously as the calculation proceeds. Calculations should be interactive such that the user should have control over the properties that are being displayed and should be able to continuously monitor the values of all properties at any point within the flow. Most importantly, apart from the error due to the improper positioning of the flow boundaries, the user should not obtain a bad result.

Finally, the program should be readily available for use on PCs; the cost should be within the established price range of specialist application programs. The program should be capable of being ported to other computers and to other operating systems. The current versions of the DS2V/3V codes by Bird come close to meeting the above specifications for the ideal general DSMC program.

5.2. *Variants of DSMC*

The DSMC is a generic term that covers the many variants of the method, which have been introduced over the years. Nanbu [47, 48] introduced a variation of the DSMC method that was derived directly from the Boltzmann equation. This method aims at the direct approximation of the Boltzmann equation. Here, molecules are just conceptual elements. The Finite Pointset Method is another variant of the DSMC method that is directly related to the Boltzmann equation. It employs discretization in both physical and phase space [49].

Macrossan [50, 51] describes a new approach to DSMC collision modelling called the 'viscosity DSMC' or the m-DSMC. In this the time-averaged temperature is used to set the characteristic collision cross-section in each cell such that the Chapman–Enskog viscosity is that given by any desired viscosity law $m = m\ (T)$. Realistic potentials, which produce realistic viscosity laws, are not widely used. This is because of the difficulty of using the Borgnakke–Larsen (BL) energy exchange scheme [52] when the collision probability does not match that of the VHS collision model. The combination of BL exchange model and the VHS model is the existing standard in DSMC. Further, the DSMC procedures for chemically reacting flows are built around the VHS model. In this new approach, an arbitrary viscosity law can be realised using simple (VHS) collision models which are compatible with the BL exchange scheme. The new method makes the VHS model display an arbitrary viscosity relationship by adjusting the number of collisions calculated in each cell, based on the cell temperature. In principle, the method can use directly experimental data on viscosity. There is virtually no extra computational overhead involved in obtaining the time-averaged temperature, which is available in DSMC codes as the evolution of the flow is tracked to steady state. The new method has been shown to be accurate for highly nonequilibrium flow in the interior of a shock and for the high-speed flow around a blunt cylinder, which involves an expansion around the sharp corner of the front face.

A variant of this new approach was proposed earlier by Macrossan [53]. This is called the 'collision-rate DSMC' or the *n*-DSMC. Here, the 'Maxwell VHS' model is used, which is a special case of the VHS model for which the collision probability is independent of the relative speed. The reference cross-section in each cell is adjusted according to the desired viscosity law. A large number of collisions are required in *n*-DSMC because, with the standard VHS model, collision pairs are weighted towards higher relative velocities so that an average VHS collision is more efficient in redistributing energy. The collision loops are particularly simple and increase the computational speed. This *n*-DSMC runs at twice the speed of the standard DSMC and produces results within 3% of those produced by DSMC for shock structure and within 5% for the Couette flow, for Knudsen numbers ranging between 0.1 and 0.005.

5.3. *Molecular models*

Several molecular models have been proposed by researchers to simulate various physical situations. No model is universally applicable and the choice of a particular model is a tradeoff between its capability to approximate a physical situation and the computational requirements/constraints, among other factors.

DSMC method is capable of simulating monatomic rarefied gas flows very well. Polyatomic rarefied gas flows, however, in which the gas molecules are able to transfer energy among translational, rotational and vibrational degrees of freedom cannot be accurately predicted using simple collision models. Tokumasu and Matsumoto [54] presented a dynamic molecular collision (DMC) model for nonpolar diatomic molecules. This model is applicable at moderate temperatures (up to a few hundred K), where most molecules are in the vibrational ground state and hence the vibrational degree of freedom can be neglected. The validity of the DMC model is verified by simulating the translational and rotational energy distributions at the equilibrium condition and the rotational relaxation of nitrogen through normal shock wave. The results are in good agreement with experimental and theoretical results. Consequently, the collision of diatomic molecules can be quite accurately predicted using the DMC model.

A physically realistic collision model is critical for DSMC to predict various gas flows reliably and accurately. In actual gases, the force between two molecules is repulsive at small distances and more weakly attractive at large distances. This behaviour is most simply described by the Lennard–Jones potential as,

$$f(r) = 4e\left[\left(\frac{s}{r}\right)^{12} - \left(\frac{s}{r}\right)^{6}\right]. \tag{8}$$

The parameters e and s, which have dimensions of energy and length, respectively, are characteristic constants of the chemical species of the colliding molecules and r is the intermolecular separation. For polar molecules like ammonia and water vapour, however, strong electrostatic interaction leads to anomalous behaviour with respect to the potential. This electrostatic contribution is proven to be proportional to the inverse third power of the molecular separation. A term is combined with the Lennard–Jones potential to describe the interaction between polar molecules as,

$$f(r) = 4e\left[\left(\frac{s}{r}\right)^{12} - \left(\frac{s}{r}\right)^{6} - d\left(\frac{s}{r}\right)^{3}\right] \tag{9}$$

where d is a dimensionless constant that measures the polarisation of a substance.

Fan [55] introduced a new collision model called the generalised soft-sphere model (GSS), which is a collision model suitable for polar and nonpolar molecules over a wide temperature range. Coefficients of viscosity, selfdiffusion and diffusion given by GSS model agree well with experimental data for a number of nonpolar and polar gases over a wide range of temperature. This GSS model may be regarded as an extension of the generalised hard-sphere (GHS) model [56] for attractive-repulsive potentials. At low temperatures, where the attractive intermolecular forces are important, the GHS model produces a more accurate variation of viscosity with temperature than the standard VHS model. It is conveniently implemented in an engineering simulation. However, the GHS model remains virtually unused because of its poor computational efficiency as compared to the VHS model. Macrossan and Lilley [57] introduced a slight modification to the GHS model that offers significant improvement in computational efficiency (not more than 5–15% more computationally expensive than the VHS model) with minimal effects on the viscosity behaviour. The computational efficiency of the GSS model, which has the same total cross-section as the GHS model, might be improved using the same modification.

The computing task of the simulation method varies with the molecular model. For models other than the Maxwell's molecular model (a special case of inverse power law model), it is proportional to N for Bird's method, while it is proportional to N^2 for Nanbu's method. Babovsky [58] and Babovsky and Illner [59] have introduced a procedure to reduce the computing task of Nanbu's method to make it proportional to N.

5.3.1. *Energy exchange schemes*

Rotational energy is associated only with di- and polyatomic molecules that consist of two or more atoms. In order to accommodate the properties of such gases, several models have been proposed and studied. The two centres of repulsion model, sphero-cylinder and the weighted sphere models, are a few of them. However, serious limitations of these models prompted a quick recourse to a phenomenological approach.

A phenomenological model is defined as a model that reproduces the overall effects in a gas flow of some physical feature of the real molecules without the explicit incorporation of that feature. The phenomenological models superimpose the energy exchange between the internal and translational energy upon an otherwise monatomic model. The models are therefore spherically symmetric and the phenomenological process allows the specification of the requisite number of rotational degrees of freedom that leads to equi-partition between these modes. There must also be exact equi-partition between rotational and translational modes, and, when the gas is in equilibrium, the distribution functions of each quantity must be the local Maxwellian.

The most widely used phenomenological model is that of Borgnakke and Larsen [52] in which a detailed balance is ensured by the selection of the postcollision values directly from the Maxwellian distributions. The relaxation rate, that is, the rate at which a gas tends to equilibrium is controlled by regarding a fraction of the collisions as elastic. It is a statistical model based on the relaxation concept, which is suitable for a Monte Carlo simulation scheme. Typically, the model formulation specifies the number of internal degrees of freedom, the collision probability and a postcollision probability, such as the angular distribution of the relative velocity. The model has a high degree of flexibility and requires little bit more computing time per collision than the hard-sphere gas of monatomic gas. It satisfies all the requirements for a successful phenomenological model. Moreover, it is to be noted that Borgnakke and Larsen in their model used hard spheres but with the collision probability varying as for the inverse-power potential, related to the viscosity power. Historically, the combination of hard sphere and a variation of the total collision cross-section corresponding to a general power viscosity law first appeared in their paper. Hence, it may be said that Borgnakke and Larsen have made important contributions towards the invention of the VHS model.

The BL model, though a successful phenomenological model, from the physical point of view, is quite unrealistic. In order to improve the physical realism of their model, Larsen and Borgnakke had introduced the ‘restricted energy exchange’ model. In this, all collisions are treated as inelastic, but the proportion of the energy of the collision pair that participates in the energy exchange is restricted. This is less efficient from the computational point of view and, more seriously, Pullin [60] showed that it does not satisfy detailed balancing. In Pullin’s model, every degree of freedom participates in the energy exchange for every collision. Though the model is more realistic, it is cumbersome. Nevertheless, it must be recognised that it is a possibility that may be exploited in the future.

In between, Bird [61] had proposed the ‘energy sink’ model, which forces equi-partition in a physically realistic manner. However, it fell out of use because it led to distortions of the equilibrium distribution function and did not satisfy detailed balancing.

5.4. *Gas–surface interaction studies*

A successful application of DSMC requires the development of accurate gas-surface interaction models, in addition to gas–gas molecular collision models. Gas–surface interaction study is one of the foremost directions of research related to rarefied gas dynamics [62]. The study of gas–surface interaction may be regarded as a bridge between the kinetic theory of gases and solid-state physics. Since the time when the first experiments with rarefied gases have been conducted by Knudsen and other classical physicists, the problem of gas–surface interaction has attracted considerable attention from both the theoreticians and experimental investigators. The difficulties of a theoretical investigation are due, mainly, to our lack of knowledge of the structure of surface layers of solid bodies and hence of the effective interaction potential of the gas molecules with the wall. Experimental research shows that the energy and momentum transfer between gas molecules and the solid surface is very sensitive to many factors. When a molecule impinges upon a surface, it is adsorbed and may form chemical bonds, dissociate, become ionised, or displace surface molecules. Its interaction with the solid surface depends on the nature of gas and that of the surface, the gas pressure, the temperature of both the gas and solids, chemical composition of the surface including induced phase due to adsorption process, and the surface roughness. During the preceding years, the experimental and theoretical studies have been more oriented towards ionised gases, polyatomic molecules, and catalytic reactions.

The theoretical study of the gas–surface interaction problem is carried out in several directions. The first step is concentrated on the development of the appropriate strategy of boundary conditions for the Boltzmann equation and on the use of a suitable scattering model for solving gas dynamics problems. This structural level of a theoretical approach can be related to the Boltzmann technique. Well-known models developed by Maxwell, Epstein, Cercignani and Lampis are used for solving the rarefied gas dynamics problems and they give quite satisfactory results. The main task in this approach is the development of a scattering kernel that transforms the distribution function of incident molecules into reflected ones from the surface. One effective approach for developing collision models is the use of the molecular dynamics (MD) method [63]. A multistage (MS) gas–surface interaction model for monatomic/diatomic gas molecule interacting with a solid surface is presented. The basic idea of the model is to separate the collisions into three stages. At stages 1 and 2, the energy and scattering direction are determined by the model equation. At stage 3, according to the translational energy, the molecule is determined to scatter, re-enter or to be trapped by the surface. Re-entering molecules return to stage 1.

5.5. *Analysis of real gas effects*

The analysis of real gas effects like translational–internal energy transfer, chemical reactions and also absorption and emission of radiation has been performed using various DSMC models. Bondar *et al.* [64] presented an applicability analysis of the DSMC models for real gas effects by comparing the rates obtained for chemical reactions and energy transfer processed in an adiabatic reservoir with the corresponding temperature-dependent experimental relaxation rates. The detailed comparison of the DSMC results, obtained using three models, with the Jeans (rotational relaxation), Landau–Teller (vibrational relaxation) and Arrhenius rates given for di- and polyatomic molecules.

The molecular model is ideally suited to the study of chemically reacting gas flows because the basic theory of gas-phase chemical reactions is primarily concerned with processes at the molecular level. In most DSMC simulations of reacting flows, the probability of reaction between reactants is calculated at each collision. Where the total collision energy exceeds some threshold criterion, the reactant species are changed to product species. The energy modes of the product particles are sampled from appropriate energy distributions. Bird [4] gives a detailed treatment of the collision theory for chemical reactions and the kinetic theory employed to extract microscopic data from the continuum rate constants. A model that treats dissociation as a consequence of vibrational excitation is also explained.

In general, the probability of a chemical reaction between colliding molecules is a function of the relative translational energy, internal states and relative orientation of the colliding molecules and the modelling of chemical reactions is an integral part of the collision routine. Lilley and Macrossan [65] presented a method different from the conventional DSMC chemistry methods, called the macroscopic (‘overlay’) chemistry methods for symmetrical diatomic dissociating gas. In the macroscopic chemistry method the simulation of collisions and reactions is decoupled; the reactions being computed independently after the collision routine. Nevertheless, the results obtained by this method, which requires lesser computing time than most DSMC chemistry methods, are in excellent agreement with those obtained through conventional DSMC chemistry methods [66]. This approach is phenomenological in nature as it does not attempt to model the detailed physics of the reaction process. The method is called a macroscopic method because the number of reaction events to perform in each cell is calculated using the macroscopic reaction rates, the macroscopic equilibrium constant and the local macroscopic flow conditions. One advantage of this method is that as macroscopic information about reaction rates is used directly in DSMC in the same way as in continuum codes, hybrid Navier–Stokes/DSMC codes can easily use the same chemical models in both rarefied and continuum flow regions.

6. Conclusions

Over the years, the capabilities of DSMC have grown enormously concomitant with the growth of computer capabilities. The need to model a variety of transition regime flows quantitatively has significantly advanced the DSMC methodology. If it were possible to use DSMC throughout its physically realistic range, the applicability and generality of the solution to transitional flow problems would be substantially greater. However, computational costs of DSMC are high, even in regimes, where it is being used now, and much of the potential range of the method is beyond current computational capabilities. Although, at present, the DSMC has no rivals for practical computations, some other methods may turn out to be of interest in the future as much more powerful computers become available.

The DSMC is not only a practical tool for engineers but also a good method for probing into uncovered areas of the theory of the Boltzmann equation. DSMC has already given quantitative answers to a broad range of questions that otherwise could not have been answered with sufficient accuracy. In addition to being used as a tool, the method can also be exploited towards carrying out numerical experimentation to address fundamental issues in gas dynamics, which are not readily susceptible to physical experimentation. The simplicity and flexibility of the implementation and the stability of the solution makes it an attractive tool for solving problems in physical sciences and engineering.

References

1. E. S. Oran, C. K. Oh, and B. Z. Cybyk, Direct simulation Monte Carlo: Recent advances and applications, *A. Rev. Fluid Mech.*, **30**, 403–441 (1998).
2. G. J. LeBeau, A parallel implementation of direct simulation Monte Carlo method, *Comput. Methods Appl. Mech. Engng*, **174**, 319–337 (1999).
3. C. Cercignani, *Rarefied gas dynamics. From basic concepts to actual calculations*, Cambridge University Press (2000).
4. G. A. Bird, *Molecular gas dynamics and the direct simulation of gas flows*, Clarendon Press (1994).
5. J. Struckmeier, and K. Steiner, A comparison of simulation methods for rarefied gas flows, *Phys. Fluids*, **7**, 2876–2885 (1995).
6. F. Tcheremissine, Direct numerical solution of the Boltzmann equation, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 677–685, American Institute of Physics (2005).
7. A. Nordseick, and B. L. Hicks, Monte Carlo evaluation of the Boltzmann collision integral, in *Proc. 5th Int. Rarefied Gas Dynamics Symp.* (C. L. Brundin, ed.), pp. 675–710, Academic Press (1967).
8. S. M. Yen, Monte Carlo solutions of nonlinear Boltzmann equation for problems of heat transfer in rarefied gases, *Int. J. Heat Mass Transfer*, **14**, 1865–1869 (1970).
9. J. E. Broadwell, Study of rarefied flow by the discrete velocity method, *J. Fluid Mech.*, **19**, 401–414 (1964).
10. B. J. Alder, and T. E. Waiwright, Studies in molecular dynamics, *J. Chem. Phys.*, **27**, 1208–1209 (1957).
11. J. K. Haviland, and M. L. Lavin, Application of the Monte Carlo method to heat transfer in a rarefied gas, *Phys. Fluids*, **5**, 1399–1405 (1962).
12. G. A. Bird, Approach to translational equilibrium in a rigid sphere gas, *Phys. Fluids*, **6**, 1518–1519 (1963).
13. G. A. Bird, Shock-wave structure in a rigid sphere gas, in *Proc. 4th Int. Rarefied Gas Dynamics Symp.* (J. H. de Leeuw, ed.), Vol. 2, pp. 216–222, Academic Press (1965).
14. J. M. Hammersley, and D. C. Handscomb, *Monte Carlo methods*, Methuen & Co. Ltd. (1964).
15. G. A. Bird, Direct simulation and the Boltzmann equation, *Phys. Fluids*, **13**, 2676–2681 (1969).
16. H. Xue, Q. Fan, and C. Shu, Prediction of micro-channel flows using direct simulation Monte Carlo, *Probabilistic Engng Mech.*, **15**, 213–219 (2000).
17. C. Cercignani, and M. Lampis, Kinetic models for gas–surface interactions, *Transport Theory Stat. Phys.*, **1**, 101–114 (1971).
18. R. G. Lord, Some extensions to the Cercignani–Lampis gas scattering kernel, *Phys. Fluids A*, **3**, 706–710 (1991).
19. R. G. Lord, Some further extensions of the Cercignani–Lampis gas–surface interaction model, *Phys. Fluids*, **7**, 1159–1161 (1995).
20. S. S. Sazhin, and V. V. Serikov, Rarefied gas flows: hydrodynamic versus Monte Carlo modelling, *Planetary Sp. Sci.*, **45**, 361–368 (1997).
21. W. Wagner, Monte Carlo methods and numerical solutions, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 459–466, American Institute of Physics (2005).
22. N. G. Hadjiconstantinou, Analysis of discretization in the direct simulation Monte Carlo, *Phys. Fluids*, **12**, 2634–2638 (2000).
23. D. A. Fedosov, S. V. Rogasinsky, M. I. Zeifman, M. S. Ivanov, A. A. Alexeenko, and D. A. Levin, Analysis of numerical errors in the DSMC method, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 589–594, American Institute of Physics (2005).
24. G. Chen, and I. D. Boyd, Statistical error analysis for the direct simulation Monte Carlo technique, *J. Comput. Phys.*, **126**, 434–448 (1996).
25. A. L. Garcia, and W. Wagner, Time step truncation error in direct simulation Monte Carlo, *Phys. Fluids*, **12**, 2621–2633 (2000).

26. M. A. Rieffel, A method for estimating the computational requirements of DSMC simulations, *J. Comput. Phys.*, **149**, 95–113 (1999).
27. A. A. Shevyrin, Ye. A. Bondar, and M. S. Ivanov, Analysis of repeated collisions in the DSMC method, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 565–570, American Institute of Physics (2005).
28. G. A. Bird, Monte Carlo simulation in an engineering context, in *Proc. 12th Int. Rarefied Gas Dynamics Symp.* (S. S. Fisher, ed.), Vol. 1, pp. 239–255, American Institute of Aeronautics and Astronautics (1981).
29. K. Koura, and H. Matsumoto, Variable soft sphere molecular model for inverse-power-law or Lennard–Jones potential, *Phys. Fluids A*, **3**, 2459–2465 (1991).
30. K. Koura, and H. Matsumoto, Variable soft sphere molecular model for air species, *Phys. Fluids A*, **4**, 1083–1085 (1992).
31. K. C. Kannenberg, and I. D. Boyd, Strategies for efficient particle resolution in the direct simulation Monte Carlo method, *J. Comput. Phys.*, **157**, 727–745 (2000).
32. B. Z. Cybyk, E. S. Oran, J. P. Boris, and J. D. Anderson Jr, Combining monotonic Lagrangian grid with a direct simulation Monte Carlo method, *J. Comput. Phys.*, **122**, 323–334 (1995).
33. S. Dietrich, and I. D. Boyd, Scalar and parallel optimized implementation of the direct simulation Monte Carlo method, *J. Comput. Phys.*, **126**, 328–342 (1996).
34. S. Rajasanow, and W. Wagner, Stochastic weighted particle method for the Boltzmann equation, *J. Comput. Phys.*, **124**, 243–253 (1996).
35. S. Rajasanow, and W. Wagner, Simulation of rare events by the stochastic weighted particle method for the Boltzmann equation, *Math. Comput. Modelling*, **33**, 907–926 (2001).
36. I. Matheis, and W. Wagner, Convergence of the stochastic weighted particle method for the Boltzmann equation, *SIAM J. Sci. Comput.*, **24**, 1589–1609 (2003).
37. Q. Sun, and I. D. Boyd, A direct simulation method for subsonic microscale gas flows, *J. Comput. Phys.*, **179**, 400–425 (2002).
38. J. Fan, and C. Shen, Statistical simulation of low-speed rarefied gas flows, *J. Comput. Phys.*, **167**, 393–412 (2001).
39. J. Fan, I. D. Boyd, C. P. Cai, K. Hennighausen, and G. V. Candler, Computation of rarefied gas flows around a NACA0012 airfoil, *AIAA J.*, **39**, 618–625 (2001).
40. Y. Matsumoto, and T. Tokumasu, Parallel computing of diatomic molecular rarefied gas flows, *Parallel Comput.*, **23**, 1249–1260 (1997).
41. M. S. Ivanov, G. N. Markelov, and S. F. Gimelshein, Statistical simulation of reactive rarefied flows: Numerical approach and applications, *AIAA paper 98-2669* (1998).
42. G. J. LeBeau, and F. E. Lumpkin III, Application highlights of DSMC analysis code (DAC) software for simulating rarefied flows, *Comput. Methods Appl. Mech. Engng*, **191**, 595–609 (2001).
43. J. S. Wu, K. C. Tseng, U. M. Lee, and Y. Y. Lian, Development of a general parallel three-dimensional direct simulation Monte Carlo code, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 559–564, American Institute of Physics (2005).
44. I. D. Boyd, G. Chen, and G. V. Candler, Predicting failure of continuum fluid equations in transitional hypersonic flows, *Phys. Fluids*, **7**, 210–219 (1995).
45. W. L. Wang, and I. D. Boyd, Predicting continuum breakdown in hypersonic viscous flows, *Phys. Fluids*, **15**, 91–100 (2003).
46. G. A. Bird, The DS2V/3V program suite for DSMC calculations, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 541–546, American Institute of Physics (2005).
47. K. Nanbu, Direct simulation scheme derived from the Boltzmann equation. I. Monocomponent gas, *J. Phys. Soc. Jap.*, **49,** 2042–2049 (1980).

48. K. Nanbu, Direct simulation scheme derived from the Boltzmann equation. II. Multicomponent gas, *J. Phys. Soc. Jap.*, **49**, 2050–2054 (1980).

49. F. Gropengiesser, H. Neunzert, J. Struckmeier, and B. Wiesen, Rarefied gas flow around a disc with different angles of attack, in *Proc. 17th Int. Rarefied Gas Dynamics Symp.* (A. E. Beylich, ed.), pp. 546–553, VCH (1991).

50. M. N. Macrossan, *m*-DSMC: A general viscosity method for rarefied flow, *J. Comput. Phys.*, **185**, 612–627 (2003).

51. M. N. Macrossan, A fast simulation method with arbitrary viscosity law, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 692–697, American Institute of Physics (2005).

52. C. Borgnakke, and P. S. Larsen, Statistical collision model for Monte Carlo simulation of polyatomic gas mixture, *J. Comput. Phys.*, **18**, 405–420 (1975).

53. M. N. Macrossan, -DSMC: A fast simulation method for rarefied flow, *J. Comput. Phys.*, **173**, 600–619 (2001).

54. T. Tokumasu, and Y. Matsumoto, Dynamic molecular collision (DMC) model for rarefied gas flow simulations by DSMC method, *Phys. Fluids*, **11**, 1907–1920 (1999).

55. J. Fan, A generalized soft-sphere model for Monte Carlo simulation, *Phys. Fluids*, **14**, 4399–4405 (2002).

56. H. A. Hassan, and D. B. Hash, A generalized hard-sphere model for Monte Carlo simulations, *Phys. Fluids A*, **5**, 738–744 (1993).

57. M. N. Macrossan, and C. R. Lilley, Modified generalized hard sphere collision model for direct simulation Monte Carlo calculations, *J. Thermophys. Heat Transfer*, **17**, 289–291 (2003).

58. H. Babovsky, A convergence proof for Nanbu’s Boltzmann simulation scheme, *Euro. J. Mech. B: Fluids*, **8**, 41–55 (1989).

59. H. Babovsky, and R. Illner, A convergence proof for Nanbu’s simulation method for the full Boltzmann equation, *SIAM J. Numerical Anal.*, **26**, 45–65 (1989).

60. D. I. Pullin, Kinetic models for polyatomic molecules with phenomenological energy exchange, *Phys. Fluids*, **21**, 209–216 (1978).

61. G. A. Bird, *Molecular gas dynamics*, Clarendon Press (1976).

62. S. F. Borisov, Progress in gas–surface interaction study, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 933–940, American Institute of Physics (2005).

63. N. Yamanishi, Y. Matsumoto, and K. Shobatake, Multistage gas–surface interaction model for the direct simulation Monte Carlo method, *Phys. Fluids*, **11**, 3540–3552 (1999).

64. Ye. A. Bondar, N. Gimelshein, S. Gimelshein, M. Ivanov, and I. Wysong, On the accuracy of DSMC modelling of rarefied flows with real gas effects, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 607–613, American Institute of Physics (2005).

65. C. R. Lilley, and M. N. Macrossan, A macroscopic chemistry method for the direct simulation of gas flows, *Phys. Fluids*, **16**, 2054–2066 (2004).

66. C. R. Lilley, and M. N. Macrossan, Modeling dissociation-vibration coupling with the macroscopic chemistry method, in *Proc. 24th Int. Rarefied Gas Dynamics Symp.* (M. Capitelli, ed.), pp. 1019–1024, American Institute of Physics (2005).