Implicit–Explicit Schemes for BGK Kinetic Equations

Sandra Pieraccini[1] and Gabriella Puppo[1,2]

In this work a new class of numerical methods for the BGK model of kinetic equations is presented. In principle, schemes of any order of accuracy in both space and time can be constructed with this technique. The methods proposed are based on an explicit–implicit time discretization. In particular the convective terms are treated explicitly, while the source terms are implicit. In this fashion even problems with infinite stiffness can be integrated with relatively large time steps. The conservation properties of the schemes are investigated. Numerical results are shown for schemes of order 1, 2 and 5 in space, and up to third-order accurate in time.

KEY WORDS: BGK model; kinetic equations; high-order schemes; implicit–explicit schemes.

1. INTRODUCTION

Navier Stokes equations describe the motion of a fluid when the continuum hypothesis underlying the model is valid: in particular, the ratio between the mean free path (λ) and the characteristic dimensions of the problem (L) is small: $Kn = \lambda/L \ll 1$, where Kn is the Knudsen number.

When $Kn = O(1)$, the kinetic theory of rarefied gas dynamics comes into play. Here the fundamental equation is Boltzmann equation. Traditionally, an important field of application of kinetic theory has been the motion of objects in the rarefied layers of the atmosphere, such as re-entry problems in aerospace engineering. In these cases in fact, the Knudsen number is large, because the mean free path is of orders of magnitude larger than the characteristic length of the space vehicle. Recently however a huge new field of applications of Boltzmann equations has

---

[1] Dipartimento di Matematica, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italia.
[2] To whom correspondence should be addressed. E-mail: gabriella.puppo@polito.it

1

begun to develop in the modeling of fluid flows in nanostructures: in this case, the Knudsen number is large because the scale of interest L is so small that the ratio λ/L is of order one and microscopic effects cannot be neglected.

The main tool to integrate Boltzmann equation numerically is the Direct Simulation Monte Carlo (DSMC) method (see for instance the classical book [8]); a more recent review can be found in [24]. The DSMC scheme however is very slow due to its low-convergence rate and to the complexity of the evaluation of the collision term. Moreover its results are polluted by stochastic noise and therefore lack smoothness. Interest has also focused on deterministic methods (see for instance [23]). A recent review on numerical methods for Boltzmann equation is [26], see also references therein.

From a numerical point of view, the BGK [7,10] model approximating Boltzmann equation for moderate Knudsen numbers is particularly attractive. It has a strong theoretical background (see for instance [27]). For the Boltzmann equation, the Chapman–Enskog expansion yields the system of compressible Euler equations for $Kn \to 0$ and Navier Stokes equations for moderate Kn [11]. An analogous procedure can be carried out for the BGK equations [33] and the BGK-ES model [2]. See also [30] for a recent work on the incompressible Navier Stokes limit of the BGK model and [34] for numerical applications, translating the BGK kinetic framework in the construction of reliable numerical schemes for Navier Stokes equations. Finally, relaxation schemes based on a BGK model approach have been extended to the case of degenerate parabolic equations [5].

Extensive numerical computations have tested the ability of the BGK model to approximate Boltzmann solutions for moderate Kn and Euler solutions for $Kn \ll 1$ (see [13,35]). More recently, extensive research on the elliptic BGK-ES model [16] has drawn attention to the improved approximation of the Navier Stokes regime for low-Knudsen numbers [2,19,35]. The BGK model has also been used to evaluate several flows of physical interest (see for instance [3,4], and for applications to reacting gas mixtures [1,20]). We also mention an application of BGK-like ideas to the development of models for the behavior of fluids in nanostructures [14].

The importance of the BGK model in applications has prompted a parallel development of numerical methods tailored to the particular structure of BGK equations. We start mentioning the first-order numerical scheme proposed in [13]. A second-order scheme is described in [4]. In this case, the scheme is constructed with linear second-order upwinding: the lack of limiters results in the possible onset of spurious oscillations. The third order in space scheme appearing in [35] exploits ideas derived from

high-order schemes for conservation laws: here the onset of spurious oscillations is prevented using Essentially Non-Oscillatory (ENO) reconstructions from [15]. The scheme relies on a first-order operator splitting, and thus its accuracy in time does not match the high accuracy of the space discretization.

The high-order schemes mentioned so far do not satisfy exact conservation of the macroscopic variables (mass, momentum, and energy) at the discrete level. This problem is solved in [18,19]. Exact conservation is obtained computing equilibrium at the discrete level. This construction requires the solution of a non-linear system of five equations for the BGK model and ten equations for the ES-BGK model at each grid point in space, even for explicit integration in time.

All schemes described so far are either explicit or fully implicit. Thus they require the solution of large linear systems of equations when the BGK model becomes stiff, i.e., for small relaxation times. Moreover the heavy non-linearity of the collision term usually requires further approximations in the evaluation of the Jacobian matrices needed to solve the non-linear system arising from the discretization of the BGK equations in the implicit case.

In this work, we propose a new high-order scheme in both space and time to solve the BGK equations. Our method is based on a high-order implicit–explicit discretization of the time derivative obtained with Runge–Kutta IMEX schemes, following the work of [6,17,25]. In our scheme, the discretization of the time derivative is carried out before the approximation of the space derivatives and the discretization of velocity space, as in [21]. In particular, the convective term is treated explicitly, unlike most numerical schemes for the BGK model in the stiff regime, while the source term is integrated implicitly. In the case of the BGK model, the implicit treatment of the collision term is very simple, because the local Maxwellian can be evaluated explicitly (see also [13]). This approach has the advantage of allowing to treat even problems with infinite stiffness easily (see [25,29] for applications to relaxation systems). In the implicit step, we first evaluate the macroscopic moments and next we update the distribution function: in this fashion, conservation is naturally enforced, and the solution of the implicit step is local, with no need to solve large systems of equations. We think that this approach yields a noticeable improvement in the efficiency of the scheme, and it allows to achieve very high-order accuracy. Space discretization is carried out with Weighted Essentially Non-Oscillatory (WENO) reconstructions (see [31]). An other application of WENO interpolation to kinetic problems can be found in [9]. When the grid in velocity space is introduced, exact conservation no longer holds at the discrete level. However, we find in our tests

that the error in the conserved variables is very small. Exact conservation can be enforced following the technique introduced in [18].

In this work, we show results obtained for schemes of order 1, 2 and 5 in space and 1–3 in time, respectively. Higher order schemes can be derived with the same techniques.

We start reviewing the BGK model in Sec. 2, and we describe the main characteristics of the new numerical scheme in Sec. 3. Next, we briefly describe the space and time discretizations in Sec. 4: this material is well established, but the application of these ideas in the context of kinetic equations is not standard and it allows some simplifications. We end with a few numerical tests in Sec. 5.

2. THE BGK MODEL

In this section, the BGK model is described. For simplicity, we only consider the classical BGK model introduced in [7], using the notation of [13, 18]. The scheme can be easily extended to more general BGK models.

We consider the initial value problem:

$$\frac{\partial f}{\partial t}(x, v, t) + v \cdot \nabla_x f(x, v, t) = \frac{1}{\tau}(f_{\mathrm{M}}(x, v, t) - f(x, v, t)) \qquad t \geq 0,$$
$$x \in \mathbb{R}^d, \quad v \in \mathbb{R}^N, \qquad (1)$$
$$f(x, v, 0) = f_0(x, v) \geqslant 0 \qquad \text{given initial data.}$$

Generally $d = 3$ and $N = 3$. In the 1D case, $d = 1$ and $N = 3$ and $\nabla_x = (\partial_{x_1}, 0, 0)$. In (1) f_{M} is the Maxwellian obtained from the moments of f, namely:

$$f_{\mathrm{M}}(x, v, t) = \frac{\rho(x, t)}{(2\pi RT(x, t))^{N/2}} \exp\left(-\frac{\|v - u(x, t)\|^2}{2RT(x, t)}\right).$$

The quantities ρ, u and T are, respectively, the macroscopic density, velocity, and temperature of the gas, and they are obtained from the moments of f, which are defined as follows. Given any function $g: \mathbb{R}^N \mapsto \mathbb{R}$, let us denote, as in [18], by $\langle g \rangle$ the quantity $\int_{\mathbb{R}^N} g(v)\, dv$; if $g: \mathbb{R}^N \mapsto \mathbb{R}^p$, $p > 1$, we still denote by $\langle g \rangle \in \mathbb{R}^p$ the vector whose components are given by $\langle g_i \rangle$. The moments of f are defined by

$$\begin{pmatrix} \rho \\ m \\ E \end{pmatrix} = \left\langle f \begin{pmatrix} 1 \\ v \\ \frac{1}{2}\|v\|^2 \end{pmatrix} \right\rangle \qquad (2)$$

(we dropped the dependence from x and t for simplicity). Here m is momentum, so that the macroscopic velocity is simply $u = m/\rho$, while E is the total energy, and the temperature is obtained from the internal energy e, through the relations: $\rho e = E - \frac{1}{2}\rho u^2$, $e = N R T/2$. In most applications, $N = 3$ which corresponds to a monoatomic gas with three translational degrees of freedom. In our tests, for simplicity, we will choose instead $N = 1$, as in [13]. This corresponds to a gas with a single degree of freedom, so that $e = RT/2$. The only difference with respect to a physical monoatomic gas appears in this rescaling of the temperature. With this choice, all velocity integrals will be evaluated in $\mathbb{R}$ instead of $\mathbb{R}^3$. Another approach to reduce the computational complexity of the velocity integrals, while maintaining the physical properties of the gas, has been introduced in [12] and used, among others, in [4,35].

In [4] the relaxation time τ is defined by

$$\tau^{-1} = A_c \rho, \tag{3}$$

where A_c is a constant, in such a way that $A_c \rho$ represents the collision frequency of the gas molecules. In [18] the BGK model is written with a relaxation time given by

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

where ω is the exponent of the viscosity law of the gas (e.g., for argon one has $\omega = 0.81$). In the adimensional case one has

$$\tau^{-1} = \frac{C}{Kn}. \tag{4}$$

The macroscopic moments of f are conserved, in the sense that:

$$\partial_t \langle f \rangle + \nabla_x \cdot \langle f v \rangle = 0, \tag{5a}$$

$$\partial_t \langle f v \rangle + \nabla_x \cdot \langle v \otimes v f \rangle = 0, \tag{5b}$$

$$\partial_t \left\langle \frac{1}{2}\|v\|^2 f \right\rangle + \nabla_x \cdot \left\langle \frac{1}{2}\|v\|^2 v f \right\rangle = 0. \tag{5c}$$

Moreover, an entropy principle holds, namely:

$$\partial_t \langle f \log f \rangle + \nabla_x \langle v f \log f \rangle \leqslant 0, \quad \forall f \geqslant 0 \tag{6}$$

the equality holding only for $f = f_{\mathrm{M}}$.

A numerical scheme for (1) should be able not only to yield an accurate solution to Eq. (1), but also to satisfy the conservation equations and the entropy principle in some discretized form.

3. THE NUMERICAL SCHEME

Following [21], we start discretizing the BGK problem (1) in time with a Runge–Kutta IMEX scheme (for the notation, see Sec. 4). Next the scheme will be discretized in space and, lastly, we will consider the discretization of velocity space and the evaluation of velocity integrals through quadrature. Finally, we will review the simplified algorithm that will be tested in the numerical problems.

3.1. Time Discretization

For the sake of simplicity we consider a one dimensional problem in space, and we introduce a uniform grid in time, with spacing Δt. Let $f^n(x, v) = f(x, v, n\Delta t)$. The updated density distribution is given by:

$$f^{n+1}(x, v) = f^n(x, v) - \Delta t \sum_{i=1}^{v} \tilde{w}_i v \, \partial_x f^{(i)}|_{(x,v)} + \Delta t \sum_{i=1}^{v} \frac{w_i}{\tau^{(i)}}$$
$$\times (f_{\mathrm{M}}^{(i)}(x, v) - f^{(i)}(x, v)), \tag{7}$$

where $\tau^{(i)}$ and $f_{\mathrm{M}}^{(i)}(x, v)$, $i = 1, ..., v$, are the relaxation time and the Maxwellian functions obtained from the moments of the intermediate stages $f^{(i)}(x, v)$, which are defined by the following relations (we drop the dependence on (x, v) for simplicity):

$$f^{(1)} = f^n + \Delta t \frac{a_{11}}{\tau^{(1)}} (f_{\mathrm{M}}^{(1)} - f^{(1)}), \tag{8a}$$

$$f^{(i)} = f^n - \Delta t \sum_{l=1}^{i-1} \tilde{a}_{il} v \, \partial_x f^{(l)} + \Delta t \sum_{l=1}^{i} \frac{a_{il}}{\tau^{(l)}} (f_{\mathrm{M}}^{(l)} - f^{(l)}). \tag{8b}$$

The coefficients a_{il}, $\tilde{a}_{il}$, w_i, and $\tilde{w}_i$ define the IMEX Runge–Kutta scheme and are specified in Sec. 4.

Let $\phi(v)$ denote the vector of the collision invariants of f, $\phi(v) = (1, v, \frac{1}{2}\|v\|^2)^T$. First we compute the moments of f. For the first stage value we have:

$$\left\langle f^{(1)}\phi \right\rangle = \langle f^n \phi \rangle + \Delta t \frac{a_{11}}{\tau^{(1)}} \left\langle \left(f_{\mathrm{M}}^{(1)} - f^{(1)} \right) \phi \right\rangle.$$

Since

$$\left\langle \left(f_{\mathrm{M}}^{(1)} - f^{(1)} \right) \phi \right\rangle = 0,$$

we immediately find the macroscopic variables $\rho^{(1)}$, $u^{(1)}$, and $T^{(1)}$ corresponding to $f^{(1)}$. With these quantities, we compute the corresponding Maxwellian:

$$f_{\mathrm{M}}^{(1)} = \frac{\rho^{(1)}}{(2\pi RT^{(1)})^{N/2}} \exp\left(-\frac{||v - u^{(1)}||^2}{2RT^{(1)}}\right) \tag{9}$$

and once the Maxwellian is known $\forall x$, the stage value $f^{(1)}$ can simply be found solving the linear equation (8a), for each x, v. For the following stage values we note that, iterating the construction, for all i we have $\left\langle \left(f_{\mathrm{M}}^{(i)} - f^{(i)} \right) \phi \right\rangle = 0$. Thus the moments at the ith stage are given by:

$$\left\langle f^{(i)}\phi \right\rangle = \left\langle f^n\phi \right\rangle - \Delta t \sum_{l=1}^{i-1} \tilde{a}_{il} \left\langle v\,\partial_x f^{(l)}\phi \right\rangle$$

and can be explicitly computed. Then the Maxwellian $f_{\mathrm{M}}^{(i)}$ can be computed as in (9). Now let:

$$B^{(i)} = f^n - \Delta t \sum_{l=1}^{i-1} \tilde{a}_{il} v\partial_x f^{(l)} + \Delta t \sum_{l=1}^{i-1} \frac{a_{il}}{\tau^{(l)}}(f_{\mathrm{M}}^{(l)} - f^{(l)}),$$

i.e., $B^{(i)}$ contains all information coming from the previously computed stages. Finally we obtain

$$f^{(i)} = \frac{\tau^{(i)} B^{(i)} + \Delta t\, a_{ii} f_{\mathrm{M}}^{(i)}}{\tau^{(i)} + \Delta t\, a_{ii}}.$$

Note that if $\tau^{(i)} \ll 1$, $f^{(i)}$ immediately relaxes on the local Maxwellian $f_{\mathrm{M}}^{(i)}$.

This completes the description of the time discretization of our scheme. We end this part formally stating the conservation property of the time discretization.

Proposition 1. The time discretization (7) and (8) preserves conservation of density, momentum, and total energy.

Proof. Computing the moments in (7) one has

$$\left\langle f^{n+1}\phi \right\rangle = \left\langle f^n\phi \right\rangle - \Delta t\,\partial_x \left\langle v \sum_{i=1}^{\nu} \tilde{w}_i f^{(i)}\phi \right\rangle.$$

Since this is an explicit Runge–Kutta discretization of (5), it is straightforward that density, momentum, and energy are conserved by the time discretization.   □

3.2. Space Discretization

We introduce a uniform grid in space, with grid points x_j, where the distribution function f will be evaluated: $f_j^n(v) = f(x_j, v, n\Delta t)$. The grid spacing will be denoted by $\Delta x = x_{j+1} - x_j$. The main task of the space discretization is to provide an accurate approximation of $\partial_x f$, satisfying the following requirements:

- the discretization of $\partial_x f$ should be non-oscillatory to prevent the onset of spurious oscillations, characteristic of high-order schemes in the presence of discontinuities or sharp gradients in the solution;
- the discretization should be conservative.

There are two main strategies to approximate the space derivative appearing in (7) in a conservative form. The most straightforward (finite volume formulation) is to integrate equation (7) on each cell $I_j = (x_j - \Delta x/2, x_j + \Delta x/2)$, and to update in time the cell averages of f:

$$\overline{f}_j^n(v) = \frac{1}{\Delta x} \int_{I_j} f^n(x, v)\, dx.$$

The advantage of this approach is that the cell average of the space derivative is simply:

$$\overline{v\partial_x f}_j^n = \frac{1}{\Delta x} v \left(f\left(x_j + \frac{\Delta x}{2}\right) - f\left(x_j - \frac{\Delta x}{2}\right) \right).$$

The disadvantage of this approach is that it requires to compute cell averages of the Maxwellian f_M. For first- and second-order accurate schemes, this is easily accomplished, because:

$$\overline{g}_j(v) = \frac{1}{\Delta x} \int_{I_j} g(x, v)\, dx = g(x_j, v) + O(\Delta x)^2$$

for a smooth function g. For schemes of higher order accuracy however, a more accurate quadrature rule in space is needed to compute cell averages, and then

$$\overline{\frac{1}{\tau}(f_M - f)} \neq \frac{1}{\tau}\left(\overline{f_M} - \overline{f}\right).$$

The evaluation of the cell averages of the source requires therefore to evaluate f_M, f and τ at quadrature nodes beyond the grid points x_j. This need couples the space cells together, and the scheme becomes much more complex.

The second approach is to evaluate (7) at the grid points, and approximate the derivative of f with a conservative finite difference formula, as in [32]. This is the approach we will follow in this work. Note that in this case, to preserve high accuracy the grid spacing Δx must be uniform or at least smoothly varying in x.

Suppose that the values of a flux function $F(f)$ are given at the grid points x_j. In the present case, clearly the flux function is linear, with $F(f) = v\,f$. The first step is to look for a function $\hat{F}$ that interpolates the data $F_j = F(f(.,x_j,.))$ in the sense of cell averages:

$$F_j = \frac{1}{\Delta x} \int_{I_j} \hat{F}\, dx,$$

thus the derivative of F will be written as:

$$\partial_x F|_{x_j} = \frac{1}{\Delta x} \left(\hat{F}(x_{j+1/2}) - \hat{F}(x_{j-1/2}) \right),$$

where $x_{j\pm 1/2} = x_j \pm \Delta x/2$. We outline the main steps of the construction: more details will be found in Sec. 4 and [31]. The approximation to $\hat{F}$ is typically a piecewise polynomial function, with jump discontinuities at the cell borders $x_{j\pm 1/2}$. To ensure stability, it is necessary to pick information coming from the correct direction. In other words, it is necessary to introduce upwinding. In this work, we will use flux splitting, which is particularly straightforward for the convective term of the BGK model.

Thus, we write the flux F as the sum of its positive and negative parts: $F = F^+ + F^-$ where F^+ and F^- have only non-negative (respectively: non-positive) eigenvalues. Clearly, in our case, the flux splitting will depend on the sign of v, namely:

$$F^+ = \begin{cases} v\,f, & \text{if } v > 0, \\ 0, & \text{otherwise,} \end{cases} \qquad F^- = \begin{cases} 0, & \text{if } v \leqslant 0, \\ v\,f, & \text{otherwise.} \end{cases}$$

Next, two reconstructions are computed, one for F^+ and one for F^-, which will be called, respectively, $\hat{F}^+$ and $\hat{F}^-$. Each of these has a jump discontinuity at $x_{j+1/2}$: to use upwinding, we pick the value from the left for the positive flux, $\hat{F}^+_{j+1/2} = \hat{F}^+(x^-_{j+1/2})$, and we pick the value from the right for the negative flux: $\hat{F}^-_{j+1/2} = \hat{F}^-(x^+_{j+1/2})$. Thus the numerical flux at each cell border will be given by:

$$\hat{F}_{j+1/2} = \hat{F}^+(x^-_{j+1/2}) + \hat{F}^-(x^+_{j+1/2}) \tag{10}$$

and the conservative approximation to the space derivative will be given by:

$$\partial_x(v\,f)|_j = \frac{1}{\Delta x}\left(\hat{F}_{j+1/2} - \hat{F}_{j-1/2}\right). \tag{11}$$

In the present case, the structure of F is particularly simple, and one has:

$$\hat{F}_{j+1/2}(v) = \max(v,0)\hat{f}(x_{j+1/2}^-) + \min(v,0)\hat{f}(x_{j+1/2}^+), \tag{12}$$

where $\hat{f}$ is a piecewise polynomial function such that $f_j = \frac{1}{\Delta x}\int_{I_j}\hat{f}\,dx$.

For example, for the first-order scheme $\hat{f}$ is piecewise constant, namely $\hat{f}(x)|_{I_j} \equiv f_j$. Thus $\hat{f}(x_{j+1/2}^-) = f_j$ while $\hat{f}(x_{j+1/2}^+) = f_{j+1}$ and the numerical flux in this case will be given by:

$$\hat{F}_{j+1/2}(v) = \max(v,0)f_j + \min(v,0)f_{j+1}. \tag{13}$$

Substituting the space discretization in (7), we find:

$$f_j^{n+1}(v) = f_j^n(v) - \frac{\Delta t}{\Delta x}\sum_{i=1}^{\nu}\tilde{w}_i\left(\hat{F}_{j+1/2}^{(i)}(v) - \hat{F}_{j-1/2}^{(i)}(v)\right)$$

$$+ \Delta t\sum_{i=1}^{\nu}\frac{w_i}{\tau_j^{(i)}}(f_{\mathrm{M},j}^{(i)}(v) - f_j^{(i)}(v)) \tag{14}$$

with a similar modification of (8) for the computation of the stage values $f^{(i)}$ at the grid points x_j:

$$f_j^{(1)}(v) = f_j^n(v) + \Delta t\frac{a_{11}}{\tau_j^{(1)}}(f_{\mathrm{M},j}^{(1)}(v) - f_j^{(1)}(v)), \tag{15a}$$

$$f_j^{(i)}(v) = f_j^n(v) - \frac{\Delta t}{\Delta x}\sum_{l=1}^{i-1}\tilde{a}_{il}\left(\hat{F}_{j+1/2}^{(l)}(v) - \hat{F}_{j-1/2}^{(l)}(v)\right)$$

$$+ \Delta t\sum_{l=1}^{i}\frac{a_{il}}{\tau_j^{(l)}}(f_{\mathrm{M},j}^{(l)}(v) - f_j^{(l)}(v)). \tag{15b}$$

We end stating the conservation properties of the semi-discrete scheme just described.

Proposition 2. The semi-discrete scheme (14) and (15) preserves conservation of density, momentum, and total energy.

Proof. Computing the moments of (14) and (15) in v space, we obtain a conservative and consistent discretization of the conservation equations (5) for ρ, m and E. In fact, the space discretization is carried out with an algorithm which is conservative by construction. Further, the numerical flux functions given in (13) and in Sec. 4 ensure consistency. By the Lax–Wendroff theorem, if the numerical solution converges, we conclude that the macroscopic quantities are conserved.   □

3.3. Velocity Discretization

The discretization of velocity arises several problems. The main difficulty is that the moments of f are computed approximately through a quadrature formula. As a consequence, it is no longer true that f and f_M have the same moments.

A second problem is due to the difficulty of selecting a suitable grid in velocity space. On one hand, a fine grid ensures a good accuracy of the velocity integrals which yield the macroscopic variables. On the other hand, f must be updated on each velocity grid point, which means that the scheme increases rapidly in computational cost as the velocity grid is refined. Moreover, the solution is quite sensitive to the choice of velocity grid points, which seems to be highly problem dependent, (see for instance [4], where the velocity grid changes at almost each test problem).

Choose a grid in velocity space, and let $\{v_k\}$, $k \in K$, be the set of the grid points, where k is a multi-index if $N > 1$. Given any function $g \colon \mathbb{R}^N \mapsto \mathbb{R}$, let:

$$\langle g \rangle_K = Q(g, \{v_k\}) \simeq \int_{\mathbb{R}^N} g(v)\, dv,$$

i.e. $\langle g \rangle_K$ denotes the approximation of $\langle g \rangle$ obtained by means of a suitable quadrature rule $Q(\cdot, \{v_k\})$ built on the nodes $\{v_k\}$; for a vector function $g \colon \mathbb{R}^N \mapsto \mathbb{R}^p$ we use a convention similar to the one already introduced for $\langle g \rangle$, i.e. $\langle g \rangle_K \in \mathbb{R}^p$ and the application of $\langle \cdot \rangle_K$ is meant componentwise.

The macroscopic variables now depend on the quadrature rule and on the grid used in velocity space. Let:

$$\begin{pmatrix} \rho_K \\ m_K \\ E_K \end{pmatrix} = \langle f\phi \rangle_K$$

be the moments computed from f with the quadrature $Q(\cdot, \{v_k\})$. Now construct an approximate Maxwellian with the formula:

$$M_K(f)(x, v, t) = \frac{\rho_K(x, t)}{(2\pi RT_K(x, t))^{N/2}} \exp\left(-\frac{||v - u_K(x, t)||^2}{2RT_K(x, t)}\right). \quad (16)$$

The problem is that there is no reason why f and $M_K(f)$ should have the same discrete moments, that is in general:

$$\langle (f - M_K(f))\phi \rangle_K \neq 0$$

as pointed out in [18]. However, again in [18], it is proven that it is possible to find a discrete Maxwellian $\mathcal{M}$ such that:

$$\langle (f - \mathcal{M}(f))\phi \rangle_K = 0, \qquad \mathcal{M}(f)(x, v, t) = \exp(\alpha(x, t) \cdot \phi(v)), \quad (17)$$

where $\alpha(x, t)$ is an unknown vector that depends on the macroscopic quantities. Note that α is computed precisely solving the non linear system defined by (17).

We can now introduce the discretization in velocity space in the semidiscrete system defined by (14) and (15). We describe two different schemes.

3.4. Conservative Scheme

Let $f_{jk}^n = f(x_j, v_k, t^n)$. The updated values of f will be given by:

$$f_{jk}^{n+1} = f_{jk}^n - \frac{\Delta t}{\Delta x} \sum_{i=1}^{\nu} \tilde{w}_i \left(\hat{F}_{j+1/2}^{(i)}(v_k) - \hat{F}_{j-1/2}^{(i)}(v_k)\right)$$

$$+ \Delta t \sum_{i=1}^{\nu} \frac{w_i}{\tau_{K,j}^{(i)}} (\mathcal{M}(f_j^{(i)})(v_k) - f_{jk}^{(i)}). \quad (18)$$

To compute the stage values, at each level i, for each space grid point x_j, first compute the discrete moments $(\rho_K, m_K, E_K)_j$. Then solve equation (17) to find the local components of the vector $\alpha_j^{(i)}$, and thus compute the discrete Maxwellian $\mathcal{M}(f_j^{(i)})$. Then the distribution function can

easily be evaluated at each space and velocity grid point:

$$f_{jk}^{(1)} = f_{jk}^n + \Delta t \frac{a_{11}}{\tau_{K,j}^{(1)}} (\mathcal{M}(f_j^{(1)})(v_k) - f_{jk}^{(1)}), \tag{19a}$$

$$f_{jk}^{(i)} = f_{jk}^n - \frac{\Delta t}{\Delta x} \sum_{l=1}^{i-1} \tilde{a}_{il} \left(\hat{F}_{j+1/2}^{(l)}(v_k) - \hat{F}_{j-1/2}^{(l)}(v_k) \right)$$

$$+ \Delta t \sum_{l=1}^{i} \frac{a_{il}}{\tau_{K,j}^{(l)}} (\mathcal{M}(f_j^{(l)})(v_k) - f_{jk}^{(l)}). \tag{19b}$$

Computing the discrete moments of (18), we find a consistent and conservative discretization of the conservation equations (5). Thus again density, momentum, and total energy are conserved.

It is worthwhile noting that the discrete Maxwellian $\mathcal{M}(f)$ is always computed explicitly, even if the method is implicit in the collision term. This is simpler than in the conservative scheme in [18], where the implicit version of the scheme requires an implicit evaluation of the discrete Maxwellian.

3.5. Simplified Scheme

Now we describe a simpler version of the scheme defined by (18) and (19), in which the velocity discretization is only approximately conservative.

We start from the semidiscrete equations (14) and (15). The updated values of f are now given by:

$$f_{jk}^{n+1} = f_{jk}^n - \frac{\Delta t}{\Delta x} \sum_{i=1}^{\nu} \tilde{w}_i \left(\hat{F}_{j+1/2}^{(i)}(v_k) - \hat{F}_{j-1/2}^{(i)}(v_k) \right)$$

$$+ \Delta t \sum_{i=1}^{\nu} \frac{w_i}{\tau_{K,j}^{(i)}} (f_{M,jk}^{(i)} - f_{jk}^{(i)}), \tag{20}$$

where $f_{M,jk}^{(i)}$ is the usual Maxwellian computed from the moments evaluated at the ith stage in the following way:

$$\left\langle f_j^{(i)} \phi \right\rangle_K = \left\langle f_j^n \phi \right\rangle_K - \frac{\Delta t}{\Delta x} \sum_{l=1}^{i-1} \tilde{a}_{il} \left\langle \left(\hat{F}_{j+1/2}^{(l)} - \hat{F}_{j-1/2}^{(l)} \right) \phi \right\rangle_K$$

$$+ \Delta t \sum_{l=1}^{i} \frac{a_{il}}{\tau_{K,j}^{(l)}} \left\langle \left(\mathcal{M}(f_j^{(l)}) - f_j^{(l)} \right) \phi \right\rangle_K. \tag{21}$$

Note that since in the equation above the discrete Maxwellian appears, the last term drops out. From the discrete moments, we compute the corresponding Maxwellian $f_{\mathrm{M},jk}^{(i)}$. Then the stage values are given by:

$$f_{jk}^{(i)} = f_{jk}^n - \frac{\Delta t}{\Delta x} \sum_{l=1}^{i-1} \tilde{a}_{il} \left(\hat{F}_{j+1/2}^{(l)}(v_k) - \hat{F}_{j-1/2}^{(l)}(v_k) \right)$$
$$+ \Delta t \sum_{l=1}^{i} \frac{a_{il}}{\tau_{K,j}^{(l)}} (f_{\mathrm{M},jk}^{(l)} - f_{jk}^{(l)}). \tag{22}$$

This scheme requires no solution of the non linear system (17), and is therefore much faster than the conservative scheme described above. In fact, the discrete Maxwellian $\mathcal{M}(f)$, although used, is never computed. We wish to stress the fact that this simplified scheme is not exactly conservative. However, in the numerical tests reported in Sec. 5 we will see that the errors on the conserved variables are indeed very small.

We end this section with a summary of the simplified scheme defined by (20) and (22). We use a ν stages IMEX scheme among those proposed in [22,25]. Let N_x be such that $N_x + 1$ is the number of points in the space grid and let N_v be the number of nodes in velocity space. The overall scheme reads as follows:

- Given $f^0 := f(x_j, v_k, 0) \in \mathbb{R}^{N_v \times N_x + 1}$.
- For $n = 0, 1, ...$
  - For $i = 1, ..., \nu$
    * If $i > 1$ construct the numerical fluxes $\hat{F}_{j+1/2}^{(i-1)}(v_k)$, $k = 1, ..., N_v$, with the reconstruction procedure described in Sec. 4, and according to eq. (10).
    * Compute the moments $m^{(i)}$ from (21) and the approximate Maxwellian $f_M^{(i)}$ corresponding to the point values (x_j, v_k), $j = 0, ..., N_x$, $k = 1, ..., N_v$, with $i = 1$ corresponding to the starting data at t^n.
    * Assemble the right-hand side of (22) and solve for $f^{(i)}$.
  - Compute f^{n+1} via (20).

4. TIME AND SPACE DISCRETIZATIONS

In this section, we give a few details on IMEX Runge–Kutta schemes and on high-order conservative finite difference formulas for the evaluation of the space derivatives appearing in (7) and (8). The algorithms we will overview in this section are drawn from [25] for the IMEX part and from [31] for the WENO differentiation formula.

4.1. IMEX Runge–Kutta Schemes

To set the notation for the Runge–Kutta IMEX schemes, we consider the autonomous ODE problem:

$$y'(t) = f(y) + \frac{1}{\tau} g(y), \quad y(t_0) = y_0. \tag{23}$$

We suppose that $0 < \tau \ll 1$, i.e., g/τ is stiff, so that we wish to integrate it implicitly, while f is non stiff, but highly non-linear, which means that for f an explicit scheme is more efficient.

Let $\tilde{A} = (\tilde{a}_{is})$ and $A = (a_{is})$ be two $\nu \times \nu$ matrices, with $\tilde{a}_{is} = 0$ for $s \geqslant i$ and $a_{is} = 0$ for $s > i$, and let $\tilde{w}$, w, $\tilde{c}$, c be coefficient vectors with ν elements. An IMEX Runge–Kutta scheme is represented by the following double Butcher's *tableau*:

$$\begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{w}^T \end{array} \qquad \begin{array}{c|c} c & A \\ \hline & w^T \end{array} \;,$$

where c and $\tilde{c}$ are included for completeness, although they are not needed for the autonomous system we are considering. The corresponding numerical scheme for (23) is:

$$y^{n+1} = y^n + \Delta t \sum_{i=1}^{\nu} \tilde{w}_i f(y^{(i)}) + \frac{\Delta t}{\tau} \sum_{i=1}^{\nu} w_i g(y^{(i)}), \tag{24}$$

where the stage values $y^{(i)}$ are given by:

$$y^{(1)} = y^n + \frac{\Delta t}{\tau} a_{11} g(y^{(1)}), \tag{25}$$

$$y^{(i)} = y^n + \Delta t \sum_{l=1}^{i-1} \tilde{a}_{il} f(y^{(l)}) + \frac{\Delta t}{\tau} \sum_{l=1}^{i} a_{il} g(y^{(l)}). \tag{26}$$

The coefficients of the Butcher's *tableaux* are computed in order to maximize accuracy. Further, the implicit scheme must be L-stable, to ensure that the numerical solution relaxes on the equilibrium solution, if τ is very small. Moreover, it is desirable that the IMEX scheme becomes a high-order explicit numerical scheme for the conserved variables, when $\tau \to 0$: for this reason, the matrix A must be invertible (see [22]).

In the numerical tests, we will apply IMEX schemes of orders 1–3. Here, we add the corresponding Butcher's *tableaux*.

IMEX1 First-order scheme [22]

$$
\begin{array}{c|c}
0 & 0 \\
\hline
  & 1
\end{array}
\qquad
\begin{array}{c|c}
1 & 1 \\
\hline
  & 1
\end{array}
$$

IMEX2 Second-order scheme [25]

$$
\begin{array}{c|ccc}
0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 \\
\hline
  & 0 & \frac{1}{2} & \frac{1}{2}
\end{array}
\qquad
\begin{array}{c|ccc}
\frac{1}{2} & \frac{1}{2} & 0 & 0 \\
0 & -\frac{1}{2} & \frac{1}{2} & 0 \\
1 & 0 & \frac{1}{2} & \frac{1}{2} \\
\hline
  & 0 & \frac{1}{2} & \frac{1}{2}
\end{array}
$$

IMEX3 Third-order scheme [25]

$$
\begin{array}{c|cccc}
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 \\
\frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\
\hline
  & 0 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{array}
\qquad
\begin{array}{c|cccc}
\alpha & \alpha & 0 & 0 & 0 \\
0 & -\alpha & \alpha & 0 & 0 \\
1 & 0 & 1-\alpha & \alpha & 0 \\
\frac{1}{2} & \beta & \eta & \frac{1}{2}-\beta-\eta-\alpha & \alpha \\
\hline
  & 0 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3}
\end{array}
$$

being $\alpha = 0.24169426078821$, $\beta = 0.06042356519705$, $\eta = 0.1291528696059$.

4.2. Conservative, High-Order Differentiation

The main formula for the computation of the numerical flux is (10). Here, we want to report the formulas that are needed to compute the numerical flux with first-, second- and fifth-order accuracy.

At first-order accuracy, the function $\hat{f}$ (see Sec. 3) is reconstructed as a piecewise constant function; for the second-order scheme, $\hat{f}$ is reconstructed as a piecewise linear function, while fifth-order accuracy is achieved reconstructing $\hat{f}$ as a piecewise parabolic function with WENO coefficients. We have already given the formula for the numerical flux of the first-order scheme in (13).

In the following, we drop the dependence on t and v.

4.2.1. Second order

For a second-order flux, the reconstruction for $\hat{f}$ is piecewise linear:

$$\hat{f}(x)|_{I_j} = f_j + \sigma_j(x - x_j),$$

where σ_j is an approximated slope, as for instance

$$\sigma_j = \text{MinMod}(f_{j+1} - f_j, f_j - f_{j-1})/\Delta x \quad \text{or} \quad \sigma_j = (f_{j+1} - f_{j-1})/(2\Delta x). \tag{27}$$

The first formula yields a second-order non-oscillatory reconstruction. The second formula is the classical centered difference formula we will use in the accuracy tests on smooth solutions, as in Sec. 5. Infact, the MinMod slope limiter deteriorates accuracy close to extrema, and would lead to an underestimation of the accuracy of the scheme.

Once the space reconstruction is given, the numerical flux can be written as:

$$\hat{F}_{j+1/2}(v) = \max(v, 0)(f_j + \tfrac{1}{2}\sigma_j) + \min(v, 0)(f_{j+1} - \tfrac{1}{2}\sigma_{j+1}). \tag{28}$$

4.2.2. WENO reconstruction

In the WENO reconstruction $\hat{f}(x)|_{I_j}$ is obtained through the superposition of the three parabolas $P_j^l(x)$, with $l = -1, 0, 1$ interpolating the data $f_{j+l-1}, f_{j+l}, f_{j+l+1}$ in the sense of cell averages.

We find three estimates of the left and right values of $\hat{f}$ at $x_{j+1/2}$, respectively:

$$\hat{f}^l(x_{j+1/2}^-) = P_j^l(x_{j+1/2}), \qquad \hat{f}^l(x_{j+1/2}^+) = P_{j+1}^l(x_{j+1/2}), \qquad l = -1, 0, 1.$$

The estimates are easily obtained as

$$\begin{aligned}
\hat{f}^{-1}(x_{j+1/2}^-) &= \tfrac{1}{3} f_{j-2} - \tfrac{7}{6} f_{j-1} + \tfrac{11}{6} f_j, \\
\hat{f}^0(x_{j+1/2}^-) &= -\tfrac{1}{6} f_{j-1} + \tfrac{5}{6} f_j + \tfrac{1}{3} f_{j+1}, \\
\hat{f}^1(x_{j+1/2}^-) &= \tfrac{1}{3} f_j + \tfrac{5}{6} f_{j+1} - \tfrac{1}{6} f_{j+2}
\end{aligned}$$

and

$$\begin{aligned}
\hat{f}^{-1}(x_{j-1/2}^+) &= -\tfrac{1}{6} f_{j-2} + \tfrac{5}{6} f_{j-1} + \tfrac{1}{3} f_j, \\
\hat{f}^0(x_{j-1/2}^+) &= \tfrac{1}{3} f_{j-1} + \tfrac{5}{6} f_j - \tfrac{1}{6} f_{j+1}, \\
\hat{f}^1(x_{j-1/2}^+) &= \tfrac{11}{6} f_j - \tfrac{7}{6} f_{j+1} + \tfrac{1}{3} f_{j+2}.
\end{aligned}$$

Next these data are combined with suitable weights to yield the desired values:

$$\hat{f}(x_{j+1/2}^-) = \sum_{l=-1}^{1} \omega_j^l \hat{f}^l(x_{j+1/2}^-), \qquad \hat{f}(x_{j+1/2}^+) = \sum_{l=-1}^{1} \tilde{\omega}_j^l \hat{f}^l(x_{j+1/2}^+). \tag{29}$$

The weights are defined by the relations:

$$\omega_j^l = \frac{\alpha_j^l}{\sum_{s=-1}^{1} \alpha_j^s}, \quad \alpha_j^l = \frac{d_l}{(\epsilon + \beta_j^l)^2}, \quad \tilde{\omega}_j^l = \frac{\tilde{\alpha}_j^l}{\sum_{s=-1}^{1} \tilde{\alpha}_j^s}, \quad \tilde{\alpha}_j^l = \frac{\tilde{d}_l}{(\epsilon + \beta_j^l)^2},$$

where

$$d_{-1} = \frac{1}{10}, \quad d_0 = \frac{3}{5}, \quad d_1 = \frac{3}{10}, \qquad \tilde{d}_{-1} = \frac{3}{10}, \quad \tilde{d}_0 = \frac{3}{5}, \quad \tilde{d}_1 = \frac{1}{10}$$

are the accuracy constants: they are computed in order to obtain a fifth-order scheme; ϵ prevents dividing by zero, and it is usually chosen as $\epsilon = 10^{-6}$. The quantities β_l's are the Smoothness Indicators: they prevent the inclusion in the combination of parabolas with a non-smooth stencil:

$$\beta_j^l = \sum_{k=1,2} \int_{I_j} \Delta x^{2k-1} \left(\frac{d^k P_j^l}{dx^k} \right)^2 dx.$$

In our case, the Smoothness Indicators are simply given by:

$$\begin{aligned}
\beta_j^{-1} &= \tfrac{13}{12}(f_{j-2} - 2f_{j-1} + f_j)^2 + \tfrac{1}{4}(f_{j-2} - 4f_{j-1} + 3f_j)^2, \\
\beta_j^0 &= \tfrac{13}{12}(f_{j-1} - 2f_j + f_{j+1})^2 + \tfrac{1}{4}(f_{j-1} - f_{j+1})^2, \\
\beta_j^1 &= \tfrac{13}{12}(f_j - 2f_{j+1} + f_{j+2})^2 + \tfrac{1}{4}(3f_j - 4f_{j+1} + f_{j+2})^2.
\end{aligned}$$

Once the values $\hat{f}(x_{j+1/2}^-)$ and $\hat{f}(x_{j+1/2}^+)$ are computed as in (29), we form the numerical flux as:

$$\hat{F}_{j+1/2}(v) = \max(v, 0)\hat{f}(x_{j+1/2}^-) + \min(v, 0)\hat{f}(x_{j+1/2}^+). \tag{30}$$

5. NUMERICAL EXPERIMENTS

In this section, we report some numerical results obtained with the simplified methods defined in (20) and (22). In particular, we will test the first order scheme, BGK1, obtained using the IMEX1 scheme defined in Sec. 4 and the first-order flux (13). The second-order scheme, BGK2, is defined by the second-order IMEX scheme IMEX2 of Sec. 4, and the second-order numerical flux (28). Finally the high-resolution scheme BGK3 is obtained coupling the third-order IMEX3 scheme with the fifth order accurate numerical flux (30).

Following [13], we consider a problem which is 1D in velocity space. We considered the following test problems.

Test 1. We first consider the model (4) with $C = 1$ (see [13]). We start with an initial distribution of the kind

$$f(x, v, 0) = \frac{\rho}{\sqrt{2\pi RT}} \cdot \exp\left(-\frac{(v - u_0(x))^2}{2RT}\right), \quad x \in [-1, 1]$$

with constant density $\rho = 1$ and temperature $T = 1$ and with

$$u_0 = \frac{1}{\sigma}\left(\exp\left(-(\sigma x - 1)^2\right) - 2\exp\left(-(\sigma x + 3)^2\right)\right)$$

with $\sigma = 10$. Thus initially the distribution function f is smooth, with a localized perturbation in velocity, in a gas with a uniform density and temperature.

Test 2. Consider again the model (4) with $C = 1$. We take as initial data a distribution which is discontinuous in space.

$$f(x, v, 0) = \begin{cases} \rho_L(2\pi RT_L)^{-1/2} \cdot \exp(-\frac{(u_L - v)^2}{2RT_L}), & 0 \leqslant x \leqslant 0.5, \\ \rho_R(2\pi RT_R)^{-1/2} \cdot \exp(-\frac{(u_R - v)^2}{2RT_R}), & 0.5 < x \leqslant 1 \end{cases} \quad (31)$$

with $(\rho_L, u_L, T_L) = (2.25, 0, 1.125)$ and $(\rho_R, u_R, T_R) = (3/7, 0, 1/6)$. This test is derived from [13].

Test 3. We consider the model (3) with $A_c = Kn^{-1}$. We take the same initial data used in Test 2.

We used free-flow boundary conditions in all the test problems. We considered several values of the Knudsen number Kn; further, we assumed $R = 1$. The velocity space was approximated by the finite interval $[-V, V]$ with $V = 10$; the velocity grid points are uniformly distributed around $v = 0$.

The time step is chosen following the CFL condition for the explicit, convective part, namely: $\Delta t = 0.9\Delta x / V$. The velocity integrals are approximated by trapezoidal or Simpson rule, or by the following rule

$$\int_{v_i}^{v_{i+1}} g(v)\, dv = \frac{\Delta v}{24}\left(-g(v_{i+2}) + 13g(v_{i+1}) + 13g(v_i) - g(v_{i-1})\right)$$
$$+ O(\Delta v^4). \quad (32)$$

5.1. Test 1

We integrated the first test problem up to time $t = 0.04$, when the solution is still smooth, in order to estimate the rate of convergence of the numerical solutions. In Table I, we report the errors computed, in the

**Table I.**   Test 1. Errors on ρ with $Kn = 10^{-5}$

	Scheme BGK1		Scheme BGK2		Scheme BGK3	
N_x	Error	Order	Error	Order	Error	Order
			L^1-errors			
20	8.984883e−03		6.645527e−03		5.892437e−03	
40	5.411987e−03	0.7313	2.078432e−03	1.6769	7.159745e−04	3.0409
80	3.103134e−03	0.8024	4.265673e−04	2.2846	5.544565e−05	3.6908
160	1.658142e−03	0.9042	8.301519e−05	2.3613	2.176432e−06	4.6710
320	8.565848e−04	0.9529	1.755677e−05	2.2413	4.944981e−08	5.4599
640	4.354495e−04	0.9761	4.040921e−06	2.1193	4.052476e−09	3.6091
			L^∞-errors			
20	2.542262e−02		1.362371e−02		1.643767e−02	
40	1.458513e−02	0.8016	4.776540e−03	1.5121	2.676986e−03	2.6183
80	9.001477e−03	0.6963	1.131382e−03	2.0779	2.119079e−04	3.6591
160	4.851294e−03	0.8918	2.879098e−04	1.9744	1.285543e−05	4.0430
320	2.551263e−03	0.9272	6.934110e−05	2.0538	3.401056e−07	5.2403
640	1.307679e−03	0.9642	1.689980e−05	2.0367	2.166468e−08	3.9726

case $Kn = 10^{-5}$, for different values of Δx and Δv with the three methods with respect to a reference solution obtained with scheme BGK3 on a finer grid, with $N_x = 1281$ and $N_v = 1281$. We computed the errors for the density, mean velocity and temperature. Here, we show only results corresponding to the density, since the errors in the other variables behave similarly. Specifically, we used the quadrature rule (32) to compute the integrals with $N_v = N_x$. For the accuracy test reported in these tables, in BGK2 we computed the slopes using the second choice in (27) in order to avoid the clipping phenomenon on local extrema caused by MinMod which deteriorates accuracy. Note that on all grids tested, except the coarsest one, the higher order BGK3 yields smaller errors. As the grid is fine enough, the convergence rate accelerates. We recall that scheme BGK3 is characterized by fifth-order accuracy in space and third-order in time. The higher order converge rate appearing in the table is justified by the fact that the space error dominates until very fine grids, where the accuracy deteriorates toward third-order. This motivates the use of higher order accuracy in space.

One of the density profiles used for the accuracy test in Table I appears in Fig. 1. Here we can also appreciate the difference in resolution of the various schemes tested. The data used are $Kn = 10^{-5}$, $N_x = 161$ and $N_v = 161$.
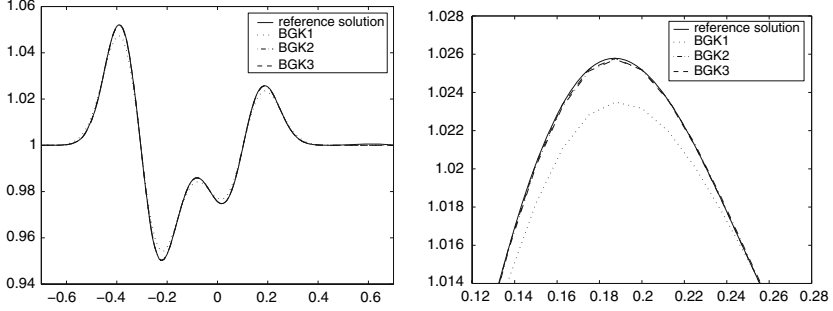
Fig. 1. Test 1, density. Left: scheme BGK1 (dotted line), scheme BGK2 (dash-dot line), scheme BGK3 (dashed line), reference solution (continuous line). Right: detail of the right peak. Results obtained with $N_x = 161$, $N_v = 161$, and $Kn = 10^{-5}$.
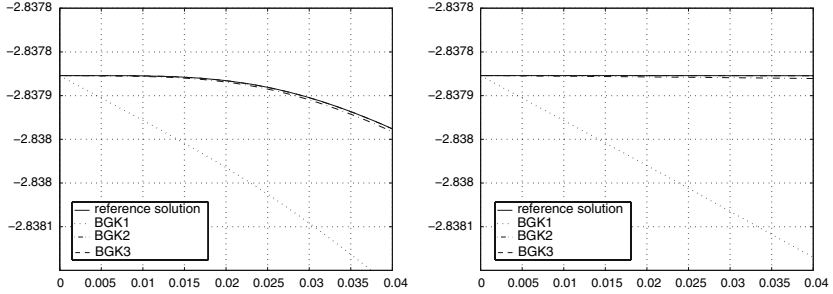
Fig. 2. Test 1, entropy; $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right). Scheme BGK1 (dotted line), scheme BGK2 (dash-dot line), scheme BGK3 (dashed line), reference solution (continuous line). In both figures the y-axis ranges from -2.8381 to -2.8378. Results obtained with $N_x = 161$, and $N_v = 161$.

Further, in Fig. 2 we plot the total entropy $H(f) = \int_{\mathbb{R}} \langle f \log(f) \rangle \, dx$ as a function of time with $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right), still with $N_x = 161$ and $N_v = 161$. Integrals in space are computed through the quadrature rule (32). We recall that for τ decreasing to 0, the BGK model approaches the hydrodynamic limit, for which the total entropy remains constant for smooth solutions. The left of Fig. 2 shows that for $Kn = 10^{-2}$ we are in the kinetic regime, where entropy decays due to kinetic collisional effects. The higher order schemes exhibit a decay close to the one of the reference solution, whereas for BGK1 entropy decays much faster. We suppose that the extra entropy dissipation is due to the numerical viscosity of the first order scheme, which can be interpreted as an artificial increase of the effective Knudsen number. This effect is present also in the higher order schemes although it is of course less pronounced.

Table II.   Test 1. Errors in conservation with $Kn = 10^{-5}$

	Scheme BGK1	Scheme BGK2	Scheme BGK3
		$N_x = N_v = 21$	
ρ	6.060198e−08	1.377644e−06	5.680021e−08
m	4.077876e−10	2.748174e−06	1.543149e−07
E	9.814734e−07	3.462126e−06	1.153762e−06
		$N_x = N_v = 41$	
ρ	9.561241e−13	5.940721e−10	2.193070e−10
m	1.365495e−12	1.091417e−09	3.901536e−10
E	1.459277e−12	1.105933e−09	3.240732e−10
		$N_x = N_v = 81$	
ρ	2.886580e−13	2.853273e−13	6.572520e−14
m	2.303713e−15	1.589354e−14	3.486794e−15
E	1.436629e−13	1.336709e−13	3.375078e−14
		$N_x = N_v = 161$	
ρ	5.906386e−13	1.938449e−13	1.278977e−13
m	3.712308e−15	1.571659e−15	2.359224e−16
E	2.953193e−13	9.992007e−14	6.550316e−14

On the right of Fig. 2 we report the entropy decay in the case $Kn = 10^{-5}$. Here we are very close to the hydrodynamic regime where total entropy remains constant with time. Again, BGK1 yields a pronounced entropy decay which is a purely numerical effect due again to numerical viscosity (as a matter of fact, the spurious amount of entropy production is almost the same in the two cases, $Kn = 10^{-2}$ and $Kn = 10^{-5}$). BGK2 and BGK3 yield an almost constant entropy production. For these reasons we believe that the entropy decay is a useful tool to assess the resolution of a numerical scheme when we are close to the hydrodynamic limit. For a link between entropy production and resolution see also [28].

Finally, in Table II we report the errors in conservation for total mass, momentum and energy in the case $Kn = 10^{-5}$ and for several values of N_x and N_v. The errors are obtained integrating in space the moments and computing the absolute value of the difference between the values at the final time ($t = 0.04$) and at the initial time. Integrals in space are approximated by formula (32). As expected, errors in conservation are very small. They decrease with the velocity grid refinement and become close to machine precision on relatively coarse grids, even for the simplified scheme (20)–(22) which does not require the evaluation of the discrete Maxwellian. The behavior on even finer grids (not shown) continues the trend established in the last rows of the table.
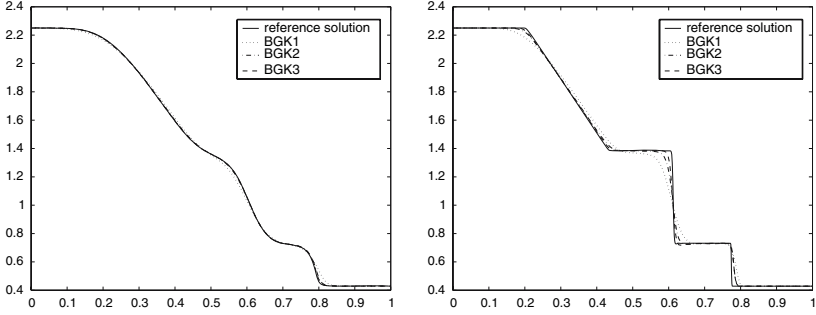
Fig. 3. Test 2, density: $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right). Scheme BGK1 (dotted line), scheme BGK2 (dash-dot line), scheme BGK3 (dashed line), reference solution (continuous line). Results obtained with $N_x = 200$ and $N_v = 81$.

5.2. Test 2

We present the results obtained for density and for the distribution function at $t = 0.16$, as in [13], for different values of Kn. We used $N_x = 200$ and $N_v = 81$ with the trapezoidal rule. The reference solution was obtained with $N_x = 1000$ and $N_v = 201$.

Figure 3 reports the results obtained for the density at time $t = 0.16$ with the three schemes BGK1, BGK2, and BGK3, for $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right). In Fig. 4, we plot the behavior of the total entropy $H(f)$ versus time, again for the two values of Kn.

We note that for $Kn = 10^{-2}$ there is a good agreement with the results shown in [13]. Since τ is relatively large, the profiles are still smooth. Note however that the higher order schemes yield somewhat sharper transitions. This is mirrored in the corresponding entropy decay plot, where again the first-order scheme overestimates kinetic effects.

In the case with $Kn = 10^{-5}$, the Knudsen number is so small that the solution has already converged to the solution of Euler equations for a monoatomic gas with a single degree of freedom. In the figures we can clearly detect the presence of a contact discontinuity and a shock wave. The results obtained with BGK3 are much sharper than the analogous results shown in [13] with a double number of grid points in space, and even the solution obtained with BGK1, which has an order of accuracy comparable with the scheme used in [13], has a better resolution and does not exhibit spurious oscillations. Again, the entropy decreases in time for all schemes tested. Here the entropy decay is non-zero even in the hydrodynamic limit. As a matter of fact, entropy is produced across a shock even in the compressible Euler equations. Again BGK1 produces an extra amount of entropy which may be accounted for by numerical
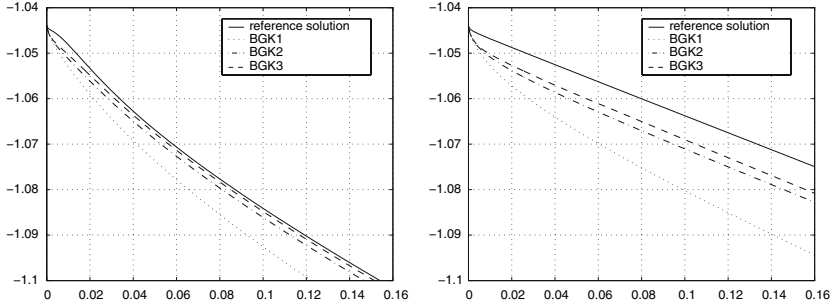
Fig. 4. Test 2, total entropy plotted versus time: $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right). Scheme BGK1 (dotted line), scheme BGK2 (dash-dot line), scheme BGK3 (dashed line), reference solution (continuous line). Results obtained with $N_x = 200$ and $N_v = 81$.

viscosity. Even the higher order schemes are characterized by a somewhat larger entropy production with respect to the reference solution. In this case, we conjecture that this behavior is due to the action of the limiters which are active on the discontinuous transitions and induce dissipation.

Figure 5 reports the distribution function f and the local Maxwellian f_M plotted versus v at $x = 0.32$, $x = 0.60$, $x = 0.78$. These locations are selected in order to lay at the center of the rarefaction, contact and shock waves, respectively, for $Kn = 10^{-5}$. These results are obtained with BGK3 at time $t = 0.16$ for $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right). Here we study the decay toward local equilibrium of f as a function of the Knudsen number. Note that for $Kn = 10^{-2}$ the distribution function f in the shock is away from equilibrium, while for $Kn = 10^{-5}$ mechanical equilibrium has already been reached also in the shock, since f coincides with the Maxwellian having the same moments. In the other waves the distribution function and the corresponding Maxwellian almost coincide and therefore in our further comparisons we will show only the behavior of the distribution function in the shock wave.

5.3. Test 3

We present in Fig. 6 the results obtained for density at $t = 0.16$ using different values of Kn for the two different collisional frequencies used for Tests 2 and 3. The reference solution is again computed with BGK3 on a very fine grid. As expected the two models are almost indistinguishable for small values of Kn (right of Fig. 6). For $Kn = 10^{-2}$ (left of Fig. 6), the model based on the relaxation time defined by (3) exhibits a more kinetic behaviour in those regions where the gas is more rarefied: thus the head
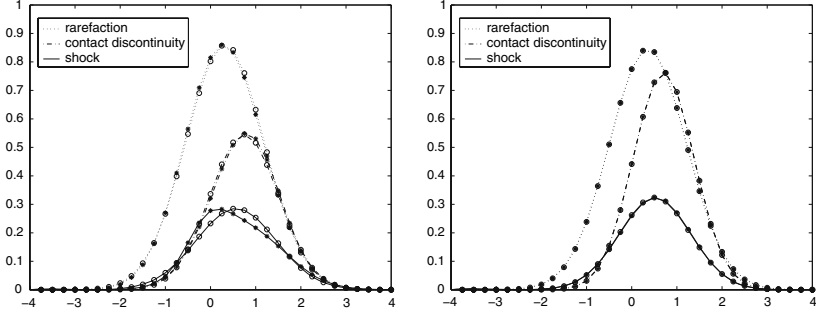
Fig. 5. Test 2, distribution function (stars) and local Maxwellian (circles) plotted versus v at selected points x ($x=0.32$, $x=0.60$, $x=0.78$). Results obtained with scheme BGK3, $N_x = 200$ and $N_v = 81$. Left: $Kn = 10^{-2}$, right: $Kn = 10^{-5}$.
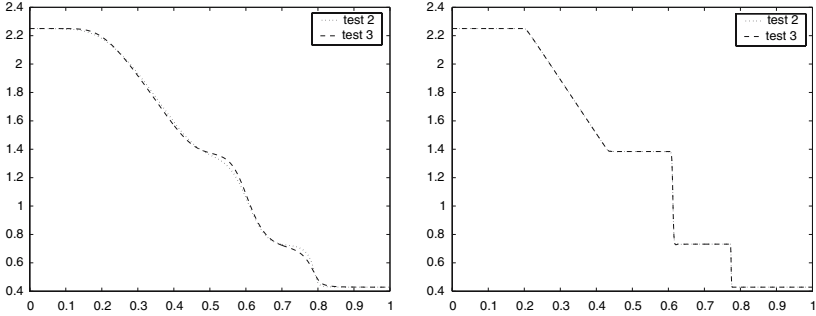

Fig. 6. Tests 2 and 3, density: $Kn = 10^{-2}$ (left) and $Kn = 10^{-5}$ (right). Tests 2 (dotted line) and 3 (dashed line). Results obtained with scheme BGK3, $N_x = 1000$ and $N_v = 201$.

and tail of the shock wave are more rounded than in the case of constant τ, see for a comparison the left part of Figure 3.

Finally Fig. 7 shows the behavior of the distribution function f and the local Maxwellian f_M plotted versus v within the shock wave for Tests 2 (top) and 3 (bottom). As in Test 2, $N_x = 200$, $N_v = 81$ and trapezoidal rule is emploied for quadrature in velocity space. Results are obtained with the three schemes at time $t = 0.16$ for $Kn = 10^{-2}$.

It is apparent that results obtained with Test 3 are further away from their local equilibrium with respect to the distribution functions computed with constant τ. Moreover, the lower order scheme underestimates the fluctuations from equilibrium.
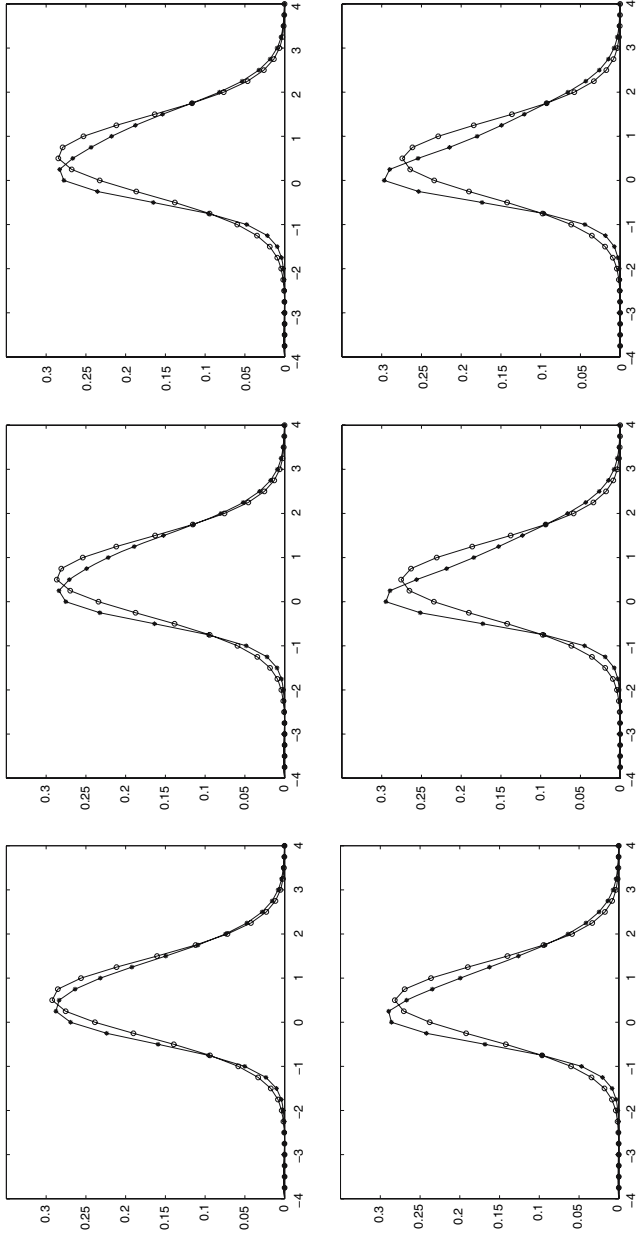
Fig. 7. Tests 2 (top) and 3 (bottom), distribution function (stars) and local Maxwellian (circles) plotted versus v at $x = 0.78$. Results obtained with scheme BGK1 (left), BGK2 (center), and BGK3 (right); $Kn = 10^{-2}$, $N_x = 200$, and $N_v = 81$.

REFERENCES

1. Andries, P., Aoki, K., and Perthame, B. (2002). A consistent BGK-type model for gas mixtures, J. Stat. Phys. **106**, 993–1018.
2. Andries, P., Bourgat, J. F., Le-Tallec, P., and Perthame, B. (2002). Numerical comparison between the Boltzmann and ES-BGK models for rarefied gases. *Comput. Methods Appl. Math. Eng.* **31**, 3369.
3. Aoki, K., Kanba, K., and Takata, S. (1997). Numerical analysis of a supersonic rarefied flow past a flat plate. *Phys. Fluids* **9**, 1144–1161.
4. Aoki, K., Sone, Y., and Yamada, T. (1990). Numerical analysis of gas flows condensing on its plane condensed phase on the basis of kinetic theory. *Phys. Fluids A* **2**, 1867–1878.
5. Aregba-Driollet, D., Natalini, R., and Tang, S. (2003). Explicit diffusive kinetic schemes for non linear degenerate parabolic systems. *Math. Comp.* **73**, 63–94.
6. Asher, U., Ruuth, S., and Spiteri, R. J. (1997). Implicit-explicit Runge-Kutta methods for time dependent partial differential equations. *Appl. Numer. Math.* **25**, 151–167.
7. Bhatnagar, P. L., Gross, E. P., and Krook, M. (1954). A model for collision processes in gases. Small amplitude processes in charged and neutral one-component systems. *Phys. Rev.* **94**, 511–525.
8. Bird, G. A. (1994). *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford University Press, Oxford.
9. Carrillo, J. A., Gamba, I. M., Majorana, A., and Shu, C.-W. (2003). A WENO-solver for the transients of Boltzmann-Poisson system for semiconductor devices: performance and comparisons with Monte Carlo methods. *J. Comput. Phys.* **184**, 498–525.
10. Cercignani, C. (2000). Rarefied Gas Dynamics, from Basic Concepts to Actual Calculations, Cambridge University Press, Cambridge.
11. Cercignani, C., Illner, R., and Pulvirenti, M. (1994). *The Mathematical Theory of Dilute Gases*. Applied Mathematical Sciences, vol. 106, Springer-Verlag, Berlin.
12. Chu, C. K. (1965). Kinetic-theoretic description of the formation of a shock wave. *Phys. Fluids* **4**, 12–22.
13. Coron, F., and Perthame, B. (1991). Numerical passage from kinetic to fluid equations. *SIAM J. Numer. Anal.* **28**, 26–42.
14. Guo, Z., Zhao, T. S., and Shi, Y. (2005). Simple kinetic model for fluid flows in the nanometer scale. *Phys. Rev. E* **71**, 035301-1, 035301-4.
15. Harten, A., Engquist, B., Osher, S., and Chakravarthy, S. (1987). Uniformly high order accurate essentially non-oscillatory schemes III. *J. Comput. Phys.* **71**, 231–303.
16. Holway, L. H. (1966). *Kinetic Theory of Shock Structure using an Ellipsoidal Distribution Function*, Academic Press, New York, pp. 193–215.
17. Kennedy, C. A., and Carpenter, M. H. (2003). Additive Runge–Kutta schemes for convection-diffusion-reaction equations. *Appl. Numer. Math.* **44**, 139–181.

18. Mieussens, L. (2000). Discrete velocity model and implicit scheme for the BGK equation of rarefied gas dynamics. *Math. Models Methods Appl. Sci.* **10**, 1121–1149.

19. Mieussens, L. (2000). Schemes for Boltzmann-BGK equation in plane and axisymmetric geometries. *J. Comput. Phys.* **162**, 429–466.

20. Monaco, R., Bianchi, M. P., and Soares, A. J. (2005). BGK-type models in strong reaction and kinetic chemical equilibrium regimes. *J. Phys. A Math. Gen.* **38**, 10413–10431.

21. Pareschi, L., Puppo, G., and Russo, G. (2005). Central Runge–Kutta schemes for conservation laws. *SIAM J. Sci. Comput.* **26**, 979–999.

22. Pareschi, L., and Russo, G. (2000). *Implicit-Explicit Runge–Kutta Schemes for Stiff Systems of Differential Equations, Recent Trends in Numerical Analysis*. In Brugnano, L., and Trigiante, D. (eds.), vol. 3, Nova Science, New York, pp. 269–289.

23. Pareschi, L., and Russo, G. (2000). Numerical solution of the Boltzmann equation I: spectrally accurate approximation of the collision operator. *SIAM J. Numer. Anal.* **37**, 1217–1245.

24. Pareschi, L., and Russo, G. (2001). An introduction to Monte Carlo methods for the Boltzmann equation. *ESAIM: Proceedings*, Vol. 10, Soc. Math. Appl. Indust., Paris, 1999, pp. 35–76.

25. Pareschi, L., and Russo, G. (2005). Implicit-explicit Runge–Kutta methods and applications to hyperbolic systems with relaxation. *J. Sci. Comput.* **25**, 129–155.

26. Pareschi, L., and Russo, G. (2005). An introduction to the numerical analysis of the Boltzmann equation. *Riv. Mater. Univ. Parma* **4****, 145–250.

27. Perthame, B. (1989). Global existence to the BGK model of Boltzmann equation. *J. Diff. Equat.* **82**, 191–205.

28. Puppo, G. (2003/04). Numerical entropy production for central schemes. *SIAM J. Sci. Comput.* **25**, 1382–1415.

29. Puppo, G., and Russo, G. (2005). Staggered finite difference schemes for balance laws. Proceedings of HYP 2004. Hyperbolic problems: Theory, Numerics, Applications, Vol II, 2006, Yokohama Publishers, pp. 243–250.

30. Saint-Raymond, L. (2003). From the BGK model to the Navier-Stokes equations. *Ann. Sci. Écn. Norm. Super. 4e série, t.* **36**, 271–317.

31. Shu, C.-W. (1998). *Essentially Non-Oscillatory and Weighted Essentially Non-Oscillatory Schemes for Hyperbolic Conservation Laws*, Quarteroni, A. (ed.), Lecture notes in Mathematics, Vol. 1697, Springer, Berlin, pp. 325–432.

32. Shu, C.-W. and Osher, S. (1988). Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.* **77**, 438–471.

33. Vincenti, W. G., and Kruger, C. H. (1986). *Introduction to Physical Gas Dynamics*, Krieger FL, USA.

34. Xu, K. (2001). A gas-kinetic BGK scheme for the Navier–Stokes equations and its connection with artificial dissipation and Godunov method. *J. Comput. Phys.* **171**, 289–335.

35. Yang, J. Y., and Huang, J. C. (1995). Rarefied flow computations using nonlinear model Boltzmann equations. *J. Comput. Phys.* **120**, 323–339.