Implicit schemes for the Fokker-Planck-Landau equation

Mohammed Lemou and Luc Mieussens

11th February 2005

MIP (UMR CNRS 5640), UFR MIG, Université Paul Sabatier,
118 Route de Narbonne, 31062 Toulouse Cedex, France
emails : lemou@mip.ups-tlse.fr, mieussens@mip.ups-tlse.fr

Abstract

We propose time implicit schemes to solve the homogeneous Fokker-Planck-Landau equation in both the isotropic and 3D geometries. These schemes have properties of conservation and entropy. Moreover, they allow for large time steps, making them faster than the usual explicit schemes. To solve the involved linear systems, we prove that the use of Krylov-like solvers preserves the conservation properties. We show in particular that the Conjugate Gradient method can be used. Numerical tests are performed for the isotropic case and demonstrate an important gain in terms of CPU time, with the same accuracy as explicit schemes. This work is a first step to the development of fast implicit schemes to solve a class of inhomogeneous kinetic equations.

Keywords. Kinetic equation, Fokker-Planck-Landau equation, implicit schemes, conservative schemes, Krylov methods

AMS subject classifications. 82C40, 82D10, 82C80, 65M06, 65Y20, 65F10

1 Introduction

The Fokker-Planck-Landau (FPL) is a kinetic collisional model used to describe a system of particles in plasma physics (see [17] for instance). The particles are described through a distribution function $f(t,x,v)$ depending on time t, particle position $x \in \mathbb{R}^d$, and their velocity $v \in \mathbb{R}^d$ ($d = 2,3$). In this paper we are concerned with the homogeneous case where $f(t,x,v)$ does not depend on x. The model writes in the so-called Landau form

$$\partial_t f(t,v) = Q(f)(v) = \nabla \cdot \int_{\mathbb{R}^d} \Phi(v - v_*) \left(f(v_*) \nabla f(v) - f(v) \nabla f(v_*) \right) \, dv_*, \tag{1}$$

where $\Phi(w)$ is the following $d \times d$ matrix

$$\Phi(w) = C|w|^{\gamma+2} S(w) = C|w|^{\gamma+2} \left(I_d - \frac{w \otimes w}{|w|^2} \right).$$

In this expression, C is a positive constant and γ is a parameter leading to the standard classification in hard potentials ($\gamma > 0$), Maxwellian potential ($\gamma = 0$) and soft potentials ($\gamma < 0$). This last case includes the most physically interesting case, the Coulombian case ($\gamma = -3$). The $d \times d$ matrix $S(w)$ is simply the orthogonal projection onto the plane orthogonal to w. For all $w \neq 0$, $\Phi(w)$ is a positive matrix whose null-space is

$$\operatorname{Ker} \Phi(w) = \mathbb{R}w.$$

Throughout this paper, when no confusion is possible, the values of any function f under the integral signs will be denoted by f for $f(v)$ and by f_* for $f(v_*)$. Beside, it is also useful to write the FPL collision operator in the so called "Log form":

$$Q(f)(v) = \nabla \cdot \int_{\mathbb{R}^d} \Phi(v - v_*) f f_* \left(\nabla \log f - \nabla \log f_*\right) \, dv_*. \tag{2}$$

This collision operator $Q(f)$ satisfies the following weak formulation

$$\int_{\mathbb{R}^d} Q(f)\phi dv = -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \Phi(v - v_*) f f_* \left(\nabla \log f - \nabla \log f_*\right) \cdot \left(\nabla\phi - \nabla\phi_*\right) dv dv_*, \tag{3}$$

for all distribution function f and all test function ϕ. From this formulation, one can immediately derive the following conservation and entropy properties

(i) the only functions $\phi(v)$ such that

$$\int_{\mathbb{R}^d} Q(f)\phi \, dv = 0, \qquad \text{for all } f,$$

are linear combinations of $1, v, |v|^2$. In particular, total mass, momentum, and energy are conserved.

(ii) the entropy dissipation inequality

$$\int_{\mathbb{R}^d} Q(f)\log(f)\, dv \leq 0, \qquad \text{for all } f > 0.$$

This also gives the well known H-theorem, saying that the functional $H(f) = \int_{\mathbb{R}^d} f \log(f)\, dv$ is a time non-increasing function. Furthermore, this inequality becomes an equality if and only if f is a Maxwellian

$$f_{eq}(v) = \frac{\rho}{(2\pi T)^{\frac{d}{2}}} \exp(-\frac{|v-u|^2}{2T}), \tag{4}$$

where ρ, u and T are velocity independent parameters. This is formally equivalent to say that f is an equilibrium function, that is $Q(f) = 0$.

In a recent past, numerous works have been concerned with constructing discretizations of the collision operator that obey the above physical properties of conservation and entropy. The first scheme in this direction was established in [8]. Unfortunately, it turned out to be very expensive in terms of CPU time. Later, various fast algorithms have been constructed to reduce this cost (see [6, 15, 4, 1]). Numerical experiments usually show that the exact preservation of properties (i) and (ii) is crucial to design efficient numerical schemes. In particular, a small error on conservation properties (i) may result in a substantial error on the equilibrium state and may generate some instabilities. On the other hand, non conservative schemes have to be sufficiently accurate to reduce this error. This task can be achieved by refining the velocity grid (which is expensive), or by using spectral schemes as in [18, 19].

In all these works, the time discretization problem has not been completely solved. Indeed the used schemes are explicit in time, as for instance the usual Euler explicit scheme

$$f^{n+1} = f^n + \Delta t Q(f^n) \tag{5}$$

or higher order versions (explicit Runge-Kutta methods). It is known that such schemes induce a strong parabolic CFL condition of the form $\Delta t \leq C\Delta v^2$, Δt and Δv being the time and velocity steps. This condition is due to the diffusive nature of the FPL operator and has been rigorously established in [4] for the isotropic case (that is where the distribution function only depends on the modulus of the velocity). Therefore, to reach a given simulation time, a large number of iterations n_{it} is required. For instance, in the isotropic case, we have $n_{it} = N^2$, N being the dimension of the approximation space. In that case, even with fast evaluations of the collision operator in $O(N)$ or $O(N \log N)$ (as proposed in the previous works), these explicit schemes require a total simulation cost of the order of $n_{it}O(N) = O(N^3)$ or $O(N^3 \log N)$. However, according to some recent works, it is possible to use explicit schemes with slightly larger time steps. For instance in [10] a high order explicit scheme with a large stability interval has been used. But the gain obtained with this method remains relatively small: the time step can be taken 5 at 10 times as large as that of the Euler method (5), and the total CPU time is divided by a factor 4 only.

Consequently, it is attractive to use time implicit schemes, since it is known that they can be free of restrictive time step conditions. This consists in replacing $Q(f^n)$ in (5) by an approximation that also depends on f^{n+1}. Then the problem is how to construct such implicit schemes in order to satisfy the following two requirements

- the properties (i) and (ii) must be preserved;
- the total computational cost must be smaller than that of the explicit scheme with almost the same accuracy.

Many works in plasma physics area use implicit schemes to solve the FPL equation (see [12, 13] for instance and the references therein). However, the problem of exact preservation of (i) and (ii) is generally not addressed. Moreover, the total complexity of these algorithms is not optimal. We note that, beyond the fact that implicit schemes usually induce an additional computational cost (non-sparse matrix inversion), they may affect also the properties of conservation and entropy. In [9] for instance, Epperlein has proposed an implicit scheme which is conservative, whereas its total numerical complexity is comparable to the usual explicit resolution. This has been clearly shown in [5]. A recent work by Chacón, Barnes, Knoll and Miley [7] uses a fast linear solver to reduce the cost of their implicit scheme. However, their approach does not exactly respect the conservation and entropy properties. In fact, whereas they claim that their scheme preserves the energy, they also point out that the solvers and the velocity boundary conditions that they used in practice affect the conservation properties.

In this paper we develop a strategy leading to exactly conservative implicit schemes with a reduced computational cost. One of our scheme also satisfies the entropy property. Iterative Krylov solvers are used to efficiently solve the linear systems generated by the implicit schemes. This strategy is also proved to be conservative, even if the linear systems are only solved approximately, and an important gain in terms of computational cost is obtained. All the schemes developed in this work concern both the 2 and 3-dimensional cases as well as the isotropic FPL equation.

We point out that our method is based on the Landau form (1) or the "Log" form (2) of the FPL equation. Hence it can be used with various potentials, and easily be extended to quantum and relatvistic cases. This does not seem to be possible if one uses the so-called Rosenbluth form of the classical FPL equation, as done in [7]. Note also that our schemes can easily be used in the resolution of inhomogeneous problems via standard splitting techniques.

The outline of the paper is as follows. In the next section we focus on the time discretization only, and present the different implicit schemes. In section 3, the velocity variable is also discretized, leading to completely discrete implicit schemes. All these schemes require the resolution of large and non-sparse linear systems. This is addressed in section 4 where fast linear solvers are proposed. In section 5, we discuss the numerical complexity of our algorithms, and we present various numerical tests in the isotropic case. The validation of the present strategy for the 3-dimensional case needs further investigations to solve the linear systems. Therefore the numerical tests for the 3-dimensional geometries are defered to a forthcoming paper.

We finally mention that a part of this work has been summarized in a previous Note [16].

2 Implicit schemes

In this section, we focus on the time discretization only. Suitable discretizations in the velocity variable will be developed in sec. 3. Below, we present different strategies to construct linear time implicit schemes that have properties of conservation and entropy.

2.1 Contracted implicit scheme

We first note that the FPL operator (1) can be rewritten in the following diffusive form

$$Q(f) = \nabla \cdot (D(f)\nabla f + F(f)f)\,, \tag{6}$$

where $D(f) = \int_{\mathbb{R}^d} \Phi(v-v_*)f(v_*)\,dv_*$ and $F(f) = \int_{\mathbb{R}^d} \Phi(v-v_*)\nabla f(v_*)\,dv_*$. This shows in particular that the CFL condition resulting from the use of time explicit schemes is due to the diffusive term $\nabla \cdot (D(f)\nabla f)$ in (6). Therefore the first idea is to make ∇f implicit in this last expression. On the other hand, the conservation and entropy properties are a consequence of the symmetry property (between v and v_*) of the collision operator. To preserve this symmetry, the friction coefficient $F(f)$ has to be implicit too. This leads to the following contracted implicit scheme

$$\frac{f^{n+1} - f^n}{\Delta t} = q^c(f^n, f^{n+1}), \qquad \text{with} \tag{7}$$

$$q^c(f,g) = \nabla \cdot \int_{\mathbb{R}^d} \Phi(v - v_*)\left(f_* \nabla g - f \nabla g_*\right)\,dv_*. \tag{8}$$

Note that this operator q^c is not the linearization of Q around f. The linearized operator around f is in fact

$$\begin{aligned} q(f,g) &= \nabla \cdot \int_{\mathbb{R}^d} \Phi(v - v_*)\left(f_* \nabla g - f \nabla g_* + g_* \nabla f - g \nabla f_*\right)\,dv_* \\ &= q^c(f,g) + q^c(g,f). \end{aligned} \tag{9}$$

Therefore, the contracted scheme (7) is not the usual linearized implicit scheme as used by Epperlein in [9]. The operator q^c is contracted in the sense that we only keep the terms that are necessary to ensure the symmetry between v and v_* in the linearized operator q.

Proposition 2.1. *(i) The operator q^c satisfies the following weak formulation:*

$$\int_{\mathbb{R}^d} q^c(f,g)\phi dv = -\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \Phi(v - v_*)\left(f_* \nabla g - f \nabla g_*\right) \cdot \left(\nabla \phi - \nabla \phi_*\right) dv dv_*, \tag{10}$$

for any test function ϕ.

(ii) The contracted scheme given by (7-8) is conservative:

$$\forall n, \quad \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^n\,dv = \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^0\,dv.$$

(iii) The contracted scheme is first order in time.

Proof. The weak form is classically obtained using integration by parts and the permutation of the variables v and v_*. The conservation property is a direct consequence of the weak form: $\nabla\phi - \nabla\phi_*$ is zero for $\phi = 1$ and v, and is in $\operatorname{Ker}\Phi(v - v_*)$ for $\phi = \frac{1}{2}|v|^2$. Finally, a simple Taylor expansion and the consistency relation $q^c(f,f) = Q(f)$ give the time order of the contracted scheme. □

2.2 A θ-scheme

The present approach first consists in a time integration of equation (1) using the standard θ-scheme:

$$\frac{f^{n+1} - f^n}{\Delta t} = (1-\theta)Q(f^n) + \theta Q(f^{n+1}). \tag{11}$$

Then we linearize $Q(f^{n+1})$ around f^n:

$$Q(f^{n+1}) = Q(f^n) + DQ(f^n)(f^{n+1} - f^n), \tag{12}$$

where $DQ(f)(g) = q(f,g)$ is given by formula (9). With this linearization, the θ-scheme (11) turns to

$$\frac{f^{n+1}-f^n}{\Delta t} = \theta q(f^n, f^{n+1}) + (\frac{1}{2} - \theta) q(f^n, f^n), \tag{13}$$

for any $\theta \in \mathbb{R}$.

If we introduce the classical symmetrized bilinear operator $q_s(f,g) = \frac{1}{2}q(f,g)$, then the previous scheme can be written as a convex combination:

$$\frac{f^{n+1}-f^n}{\Delta t} = 2\theta q_s(f^n, f^{n+1}) + (1-2\theta) q_s(f^n, f^n).$$

Note that the contracted scheme (7) is not a θ-scheme and that the linearized implicit scheme used by Epperlein in [9] is obtained for $\theta = 1$. Now we give the main properties of this scheme.

Proposition 2.2. *(i) The θ-scheme given by (13) is conservative:*

$$\forall n, \quad \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^n \, dv = \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^0 \, dv.$$

(ii) The θ-scheme is second order in time if $\theta = \frac{1}{2}$, else it is first order.

Proof. The same arguments as for proposition 2.1 give the results. □

However, we are not able to prove any entropy property for both the contracted and θ-scheme, except in the isotropic case for the contracted scheme (see [16]). Therefore, we propose another strategy that uses the "Log form" (2) of the FPL collision operator. This leads to conservative and entropic schemes that are detailed in the next section.

2.3 "Log" implicit schemes

The first step is to make implicit only the log terms in the "Log form" (2) of the collision operator. This gives the following non-linear implicit scheme:

$$\frac{f^{n+1}-f^n}{\Delta t} = q^{log}(f^n, f^{n+1}) = \nabla \cdot \int_{\mathbb{R}^d} \Phi(v - v_*) f^n f^n_* \left(\nabla \log f^{n+1} - \nabla \log f^{n+1}_* \right) dv_*. \tag{14}$$

Proposition 2.3. *(i) The operator q^{log} satisfies the following weak formulation:*

$$\int_{\mathbb{R}^d} q^{log}(f,g) \phi dv = -\frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Phi(v - v_*) f f_* \left(\nabla \log g - \nabla \log g_* \right) \cdot \left(\nabla \phi - \nabla \phi_* \right) dv dv_*, \tag{15}$$

for any test function ϕ.

(ii) Scheme (14) is conservative:

$$\forall n, \quad \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^n \, dv = \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^0 \, dv.$$

(iii) The collision operator q^{log} dissipates the entropy in the following sense

$$\int_{\mathbb{R}^d} q^{log}(f,g) \log g \, dv \leq 0.$$

(iv) Discrete H-theorem: the entropy sequence $H_n = \int_{\mathbb{R}^d} f^n \log f^n \, dv$ is non increasing.

Proof. Property (i) is again obtained by classical arguments as in the proof of proposition 2.1. Properties (ii) and (iii) are direct consequences of (i). To prove (iv), we first compute the entropy variation:

$$H(f^{n+1}) - H(f^n) = \int_{\mathbb{R}^d} (f^{n+1} - f^n) \log f^{n+1} \, dv + \int_{\mathbb{R}^d} f^n \log \frac{f^{n+1}}{f^n} \, dv.$$

Using (14) and (iii) shows that the first integral is non positive. Moreover, it is classical to use a convexity inequality to prove that the second integral is also non positive. This completes the proof of (iv). □

However, $q^{log}(f^n, f^{n+1})$ is non-linear with respect to f^{n+1} which makes it difficult to use. In the second step, we propose the following linearization around f^n. We write

$$\log f^{n+1} \approx \log f^n + \frac{f^{n+1} - f^n}{f^n},$$

which is inserted in (14) to obtain the following "log"-linear implicit scheme

$$\frac{f^{n+1} - f^n}{\Delta t} = Q(f^n) + q^l(f^n, f^{n+1}), \tag{16}$$

with

$$q^l(f, g) = \nabla \cdot \int_{\mathbb{R}^d} \Phi(v - v_*) f f_* \left(\nabla \left(\frac{g}{f} \right) - \nabla \left(\frac{g}{f} \right)_* \right) dv_*. \tag{17}$$

Note that $q^l(f, g) = q^l(f, g - f)$ which clearly shows that the scheme is consistent.

Proposition 2.4. *(i) Weak formulation for q^l:*

$$\int_{\mathbb{R}^d} q^l(f, g)\phi dv = -\frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \Phi(v - v_*) f f_* \left(\nabla \left(\frac{g}{f} \right) - \nabla \left(\frac{g}{f} \right)_* \right) \cdot (\nabla\phi - \nabla\phi_*) \, dv dv_*. \tag{18}$$

(ii) Scheme (16-17) is conservative:

$$\forall n, \quad \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^n \, dv = \int_{\mathbb{R}^d} (1, v, \tfrac{1}{2}|v|^2)^T f^0 \, dv.$$

(iii) Collisional part of scheme (16-17) dissipates the entropy in the following sense:

$$\int_{\mathbb{R}^d} (Q(f) + q^l(f, g))(\log f + \frac{g}{f}) \, dv \leq 0.$$

(iv) Discrete H-theorem: the entropy sequence $H_n = \int_{\mathbb{R}^d} f^n \log f^n \, dv$ is non increasing if

$$\inf_{n \in \mathbb{N}, v \in \mathbb{R}^d} \left(\frac{f^{n+1}}{f^n} \right) \geq \frac{1}{2}. \tag{19}$$

(v) For all positive f, the linear operator $g \mapsto q^l(f, g)$ is a non-positive self-adjoint operator in the following sense:

$$\langle q^l(f, g), h \rangle_{\frac{1}{f}} := \int_{\mathbb{R}^d} q^l(f, g) h \frac{dv}{f} = \langle q^l(f, h), g \rangle_{\frac{1}{f}}, \quad \textit{and} \quad \langle q^l(f, g), g \rangle_{\frac{1}{f}} \leq 0,$$

for all g and h.

Proof. Again, we use an integration by parts and the symmetry between v and v_* in (17) to obtain (i). Using (3) and (i), we obtain the following weak formulation

$$\int_{\mathbb{R}^d}(Q(f)+q^l(f,g))\phi\,dv = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Phi(v-v_*)ff_*\left[\nabla\left(\log f+\frac{g}{f}\right)-\nabla\left(\log f+\frac{g}{f}\right)_*\right][\nabla\phi-\nabla\phi_*]\,dvdv_*.$$

This gives (ii) with $\phi(v)=1, v, \frac{1}{2}|v|^2$ and (iii) with $\phi(v)=\log f+\frac{g}{f}$. Now we prove (iv): first a simple calculation shows that the entropy variation is

$$H(f^{n+1})-H(f^n)=\int_{\mathbb{R}^d}(f^{n+1}-f^n)\left(\log f^n+\frac{f^{n+1}}{f^n}\right)dv+\int_{\mathbb{R}^d}B(f^n,f^{n+1})\,dv, \qquad (20)$$

where

$$B(f^n,f^{n+1})=f^n\left(\frac{f^{n+1}}{f^n}\log\frac{f^{n+1}}{f^n}-\left(\frac{f^{n+1}}{f^n}\right)^2+\frac{f^{n+1}}{f^n}\right).$$

From relation (16), the first integral of (20) is

$$\int_{\mathbb{R}^d}(f^{n+1}-f^n)\left(\log f^n+\frac{f^{n+1}}{f^n}\right)dv=\Delta t\int_{\mathbb{R}^d}(Q(f^n)+q^l(f^n,f^{n+1}))\left(\log f^n+\frac{f^{n+1}}{f^n}\right)dv,$$

which is non positive thanks to (iii).

To study the second integral of (20), we set $\theta(X)=X\log X-X^2+X$ for $X\in\mathbb{R}^+$, which is concave on $[\frac{1}{2},+\infty)$. Therefore, if $\frac{f^{n+1}}{f^n}\geq\frac{1}{2}$ we can use the Jensen inequality to get

$$\int_{\mathbb{R}^d}B(f^n,f^{n+1})\,dv=\int_{\mathbb{R}^d}f^n\theta\left(\frac{f^{n+1}}{f^n}\right)dv\leq\theta\left(\frac{\int_{\mathbb{R}^d}f^{n+1}\,dv}{\int_{\mathbb{R}^d}f^n\,dv}\right)\int_{\mathbb{R}^d}f^n\,dv=\theta(1)\int_{\mathbb{R}^d}f^n\,dv=0.$$

Consequently, the two integrals of the RHS of (20) are non-positive under assumption (19) of the proposition, which completes the proof of (iv).

For property (v), we use weak formulation (18) with $\phi=\frac{h}{f}$:

$$\langle q^l(f,g),h\rangle_{\frac{1}{f}}=-\frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\Phi(v-v_*)ff_*\left(\nabla\left(\frac{g}{f}\right)-\nabla\left(\frac{g}{f}\right)_*\right)\cdot\left(\nabla\left(\frac{h}{f}\right)-\nabla\left(\frac{h}{f}\right)_*\right)dvdv_*,$$

which is clearly a symmetric formula with respect to g and h. Moreover, taking $h=g$ in the previous relation gives the last inequality of (v). □

Remark 1. *The condition (19) of the proposition is reasonable, since in practice the ratio $\frac{f^{n+1}}{f^n}$ should be close to 1. However, to get a weaker condition, one could consider the following modified scheme*

$$\frac{f^{n+1}-f^n}{\Delta t}=Q(f^n)+Aq^l(f^n,f^{n+1}),$$

where A is a free positive parameter. This scheme is still consistent and condition (19) turns to

$$\inf_{n\in\mathbb{N},v\in\mathbb{R}^d}\left(\frac{f^{n+1}}{f^n}\right)\geq\frac{1}{2A}.$$

The last property of the proposition is of practical importance since in that case, the Conjugate Gradient (CG) method could be used in the inversion process of the linear system (see details in section 4). However, the weight $\frac{1}{f^n}$ which is used to construct an inner product in the CG algorithm should be positive at any time. Unfortunately, we cannot prove that this property is satisfied. We point out that the two first schemes (contracted and θ-schemes) do not have the self-adjointness property. Therefore, we propose modified versions of the θ and log-linear schemes in which the linear operators are self-adjoint at any time (in particular, the positivity of the weight is guaranteed at any time).

2.4 Equilibrium linearized implicit schemes

We write implicit terms $q(f^n, f^{n+1})$ and $q^l(f^n, f^{n+1})$ of schemes (13) and (16-17) under the following perturbation form: $q(f^n, f^{n+1} - f^n) + q(f^n, f^n)$ and $q^l(f^n, f^{n+1} - f^n)$. The idea is to replace the first argument f^n of the perturbative terms $q(f^n, f^{n+1} - f^n)$ and $q^l(f^n, f^{n+1} - f^n)$ by its associate Maxwellian equilibrium f_{eq} of the form (4) that has the same mass, momentum and energy as f^n:

$$q(f^n, f^{n+1} - f^n) \approx q(f_{eq}, f^{n+1} - f^n) \quad \text{and} \quad q^l(f^n, f^{n+1} - f^n) \approx q^l(f_{eq}, f^{n+1} - f^n).$$

This leads to the following schemes:

- Equilibrium θ-scheme:

$$\frac{f^{n+1} - f^n}{\Delta t} = Q(f^n) + \theta q(f_{eq}, f^{n+1} - f^n). \tag{21}$$

  Because f_{eq} is a Maxwellian, it is well known that the linear operator $g \mapsto q(f_{eq}, g)$ is non-positive self-adjoint for the weight $\frac{1}{f_{eq}} > 0$. Moreover, this scheme is conservative.

- Equilibrium "log" linear scheme:

$$\frac{f^{n+1} - f^n}{\Delta t} = Q(f^n) + q^l(f_{eq}, f^{n+1} - f^n). \tag{22}$$

  Because f_{eq} is positive, from property (v) of proposition 2.4, the linear operator $g \mapsto q^l(f_{eq}, g)$ is non-positive self-adjoint for the weight $\frac{1}{f_{eq}}$. Moreover, this scheme is conservative.

We mention that the use of equilibrium approximation has already been exploited in other contexts. We refer for instance to [11] where equilibrium is used to truncate the so-called Wild sums, and also to the standard BGK model of the Boltzmann equation [2].

Remark 2. *Schemes (21),(22) are not obtained by a linearization of $Q(f^{n+1})$ near the equilibrium f_{eq}.*

2.5 Summary

We summarize in table 1 the properties of the previous implicit linear schemes: time order, conservation of mass and energy, the entropy dissipation property of the approximated collision operator, the decreasing of the entropy functional (H-theorem), and the properties of the linear operators $g \mapsto q(f,g), q^c(f,g), q^l(f,g), q(f_{eq}, g)$ and $q^l(f_{eq}, g)$.

Table 1: Properties of the different implicit schemes

	time order	conservation	entropy	H-theorem	linear operator
contracted scheme (7-8)	1	yes	?	?	?
θ-scheme (11)	2 if $\theta = \frac{1}{2}$ 1 if not	yes	?	?	?
log-linear scheme (16-17)	1	yes	yes	yes	non-positive self-adjoint
equilibrium θ-scheme (21)	1	yes	?	?	non-positive self-adjoint
equilibrium log-linear scheme (22)	1	yes	?	?	non-positive self-adjoint

2.6 A remark about an unconditionally positive implicit scheme

In this section, we introduce an alternative idea to make linear the non-linear implicit scheme (14). To that purpose, we modify the approximation of the time derivative in the following way

$$\partial_t f(n\Delta t) = (f\partial_t \log f)(n\Delta t) \approx f^n \frac{\log f^{n+1} - \log f^n}{\Delta t}.$$

According to this approximation, the non-linear log scheme (14) is modified into

$$f^n \frac{\log f^{n+1} - \log f^n}{\Delta t} = q^{log}(f^n, f^{n+1}). \tag{23}$$

This is still non-linear in f^{n+1}, but it is linear in $F^{n+1} = \log f^{n+1}$ and can be easily solved. Then $f^{n+1} = \exp(F^{n+1})$ is unconditionally positive.

However, it is shown in the following proposition that this scheme cannot preserve the mass $\int f^n \, dv$.

Proposition 2.5. *(i) The scheme (23) preserves the positivity: if $f_0 > 0$ then $f_n > 0$ for all n.*

(ii) If f^n is not a Maxwellian then

$$\int f^{n+1} \, dv > \int f^n \, dv.$$

In other words, either there exists n_0 such that f^{n_0} is a Maxwellian, and then the scheme is stationary for $n \geq n_0$, or the mass $\int f^n \, dv$ is a strictly increasing sequence.

Proof. By construction, the proof of (i) is straightforward. Now we prove (ii): first, integrating (23) with respec to v gives the following relation

$$\int f^n \log\left(\frac{f^{n+1}}{f^n}\right) dv = 0. \tag{24}$$

Now using Jensen inequality yields

$$\int f^n \log\left(\frac{f^{n+1}}{f^n}\right) dv \leq \log\left(\frac{\int f^{n+1} \, dv}{\int f^n \, dv}\right) \int f^n \, dv.$$

Since f^n is positive, the previous inequality gives $\log\left(\frac{\int f^{n+1} \, dv}{\int f^n \, dv}\right) \geq 0$ which is equivalent to $\int f^{n+1} \, dv \geq \int f^n \, dv$.

To prove that this last inequality is strict, we assume the contrary, that is $\int f^{n+1} \, dv = \int f^n \, dv$, and use the Czisar-Kullback inequality to obtain

$$\int f^n \log\left(\frac{f^{n+1}}{f^n}\right) dv \leq -\frac{1}{4(\int f^n \, dv)^2} \|f^{n+1} - f^n\|^2_{L^1}.$$

Using (24) then gives $f^{n+1} = f^n$. Then we replace f^{n+1} by f^n in scheme (23) and this gives $q^{log}(f^n, f^n) = Q(f^n) = 0$. Thus f^n is a Maxwellian, which is impossible due to the assumption of the proposition. Consequently, $\int f^{n+1} \, dv > \int f^n \, dv$. □

Despite its simplicity and its positivity property, this scheme does not provide attractive numerical results. This is probably due to the fact that it does not satisfy the conservation properties.

3 Implicit schemes and velocity discretization of FPL equation: the isotropic case

The previous schemes can naturally be discretized in the velocity variable using standard conservative and entropic discretizations [8, 6, 15, 4, 1]. In this section, we illustrate this assertion with a simple velocity discretization of the FPL equation in the isotropic case.

The isotropic FPL model is equation (1) in which the distribution function f only depends on time t and on the particle kinetic energy $\varepsilon = |v|^2$. In this case, the FPL equation writes

$$\partial_t f(t,\varepsilon) = Q(f) = \frac{1}{\sqrt{\varepsilon}}\frac{\partial}{\partial \varepsilon}\int_0^{+\infty} K(\varepsilon,\varepsilon_*)\left(f(\varepsilon_*)\partial_\varepsilon f(\varepsilon) - f(\varepsilon)\partial_\varepsilon f(\varepsilon_*)\right)\, d\varepsilon_*, \tag{25}$$

with $K(\varepsilon,\varepsilon_*) = \frac{16}{3}\pi \inf(\varepsilon^{3/2},\varepsilon_*^{3/2})$ for Coulombian interactions, and $K(\varepsilon,\varepsilon_*) = \frac{16}{3}\pi\varepsilon^{3/2}\varepsilon_*^{3/2}$ for Maxwellian interactions.

For any distribution function f, the collision operator $Q(f)$ satisfies the following weak formulation

$$\int_0^{+\infty} Q(f)\phi(\varepsilon)\sqrt{\varepsilon}\, d\varepsilon = -\frac{1}{2}\int_0^{+\infty}\int_0^{+\infty} K(\varepsilon,\varepsilon_*)\left(f_*\partial_\varepsilon f - f\partial_\varepsilon f_*\right)\left(\partial_\varepsilon\phi - \partial_\varepsilon\phi_*\right)\, d\varepsilon d\varepsilon_*, \tag{26}$$

ϕ being any test function. From this formulation, one can immediately derive the following conservation and entropy properties

$$\int_0^{+\infty}(1,\varepsilon)^T Q(f)\sqrt{\varepsilon}\, d\varepsilon = 0, \qquad \int_0^{+\infty} Q(f)\log(f)\sqrt{\varepsilon}\, d\varepsilon \le 0.$$

3.1 A conservative and entropic discretization of the isotropic FPL equation

We briefly recall here the discretization used by Berezin, Khudic and Pekker [3] (also studied by Buet and Cordier [5]). The energy domain is replaced by a regular energy grid of step $\Delta\varepsilon$ and of nodes $\varepsilon_i = (i-1)\Delta\varepsilon$, with $i = 1$ to N. The case of an irregular discretization is also considered at the end of this section. Here, the length of the grid is $e = (N-1)\Delta\varepsilon$. Any function f of ε is approximated on the grid by values $(f_i)_{i=1}^N$ supposed to be approximations of $(f(\varepsilon_i))_{i=1}^N$.

Integrals on $\mathbb{R}^+$ with respect to the measure $\sqrt{\varepsilon}d\varepsilon$ are approximated by the following weighted trapezoidal quadrature formula

$$\int_0^{+\infty}\phi(\varepsilon)\sqrt{\varepsilon}\, d\varepsilon \approx \sum_{i=1}^N \phi(\varepsilon_i)\omega_i,$$

where $\omega_1 = \frac{1}{2}\int_{\varepsilon_1}^{\varepsilon_2}\sqrt{\varepsilon}\, d\varepsilon$, $\omega_i = \int_{\varepsilon_{i-1}}^{\varepsilon_{i+1}}\sqrt{\varepsilon}\, d\varepsilon$ for $i = 2$ to $N-1$, and $\omega_{N-1} = \frac{1}{2}\int_{\varepsilon_{N-1}}^{\varepsilon_N}\sqrt{\varepsilon}\, d\varepsilon$.

A simple discretization of (25) is

$$\partial_t f = Q(f),$$

where $Q(f)$ now is a N-vector of components

$$Q_i(f) = -\frac{1}{\omega_i}(D^*\mathcal{F})_i, \quad i = 1:N. \tag{27}$$

The operator D^* is defined by $(D^*f)_i = f_{i-1} - f_i$, and the vector $\mathcal{F}$ is defined by

$$\mathcal{F}_i = \sum_{j=1}^{N-1} K_{ij}\Big(f_j(Df)_i - f_i(Df)_j\Big), \quad i = 1:N-1, \tag{28}$$

and $\mathcal{F}_0 = \mathcal{F}_N = 0$. The finite difference operator D is the formal adjoint of D^* defined by $(Df)_i = f_{i+1} - f_i$.

This discretization is constructed so that the discrete collision operator satisfies the following weak formulation

$$\sum_{i=1}^{N} Q_i(f)\phi_i\omega_i = -\frac{1}{2}\sum_{i=1}^{N-1}\sum_{j=1}^{N-1} K_{ij}\Big(f_j(Df)_i - f_i(Df)_j\Big)\Big((D\phi)_i - (D\phi)_j\Big).$$

This implies that conservation and entropy properties are preserved (see the proof in [5]), namely we have

$$\sum_{i=1}^{N} Q_i(f)(1,\sqrt{\varepsilon_i})^T\omega_i = 0, \quad \sum_{i=1}^{N} Q_i(f)(\log f_i)\,\omega_i \leq 0.$$

Another discretization is deduced from the "log" form of (25) similarly to (2) (see [3]). In that case the discrete collision operator has the same form as (27) but the vector $\mathcal{F}$ now is

$$\mathcal{F}_i = \sum_{j=1}^{N-1} K_{ij} f_i f_j \Big((D\log f)_i - (D\log f)_j\Big), \quad i = 1:N-1,$$

and $\mathcal{F}_0 = \mathcal{F}_N = 0$. This scheme has the same properties as the "non-log" scheme.

The main drawback of a regular discretization is the fact that the resolution is not accurate near $\varepsilon = 0$, while there are too many points for large ε. Therefore, it is interesting to consider an irregular discretization with the energy step $\Delta\varepsilon_i = \varepsilon_{i+1} - \varepsilon_i$ as in [4]. Following the same procedure as for the regular case, we obtain this discrete collision operator

$$Q_i(f) = \frac{1}{\omega_i}(\mathcal{F}_i - \mathcal{F}_{i-1}) \qquad i = 1:N, \tag{29}$$

with

$$\mathcal{F}_i = \sum_{j=1}^{N-1} K_{ij}\Big(f_j(Df)_i - f_i(Df)_j\Big)\Delta\varepsilon_j, \quad i = 1:N-1,$$

and $\mathcal{F}_0 = \mathcal{F}_N = 0$. Here D is the following finite difference operator $(Df)_i = \frac{f_{i+1}-f_i}{\Delta\varepsilon_i}$, for $i = 1:N-1$. This approximation is conservative, but the associated equilibrium states are only approximations of the Maxwellians, *i.e.*

$$(f_{eq})_i = \beta\prod_{j=1}^{i}(1+\alpha\Delta\varepsilon_j),$$

β and α being some constants. Maxwellian equilibrium states can be obtained by using the conservative and entropic approximation of [4], which is based on the "log" form.

Consequently, we have a discrete collision operator that possesses all desired properties (both in "log" and "non-log" forms). Our implicit schemes can now be derived exactly as in the continuous case, as it is illustrated in the next section.

3.2 Implicit schemes

In this section, we give completely discretized implicit schemes (in both time and velocity). For the sake of simplicity we only present the case of a regular discretization. The case of an irregular grid can be treated in the same way.

The different implicit schemes are given by equations (7, 13, 16, 21, 22), where the discrete collision operators are the N-vectors defined as follows:

- contracted scheme: $q_i^c(f,g) = -\frac{1}{\omega_i}(D^*\mathcal{F}^c(f,g))_i$, for $i = 1:N$, where the flux $\mathcal{F}^c$ is

$$\mathcal{F}_i^c(f,g) = \sum_{j=1}^{N-1} K_{ij}\Big(f_j(Dg)_i - f_i(Dg)_j\Big), \quad i = 1:N-1.$$

- θ-scheme: $q_i(f,g) = q_i^c(f,g) + q_i^c(g,f)$, $i = 1 : N$.
- log-linear scheme: $q_i^l(f,g) = -\frac{1}{\omega_i}(D^* \mathcal{F}^l(f,g))_i$, for $i = 1 : N$, where the flux $\mathcal{F}^l$ is

$$\mathcal{F}_i^l(f,g) = \sum_{j=1}^{N-1} K_{ij} f_i f_j \Big[\Big(D\Big(\frac{g}{f}\Big)\Big)_i - \Big(D\Big(\frac{g}{f}\Big)\Big)_j \Big], \quad i = 1 : N-1.$$

For all these definitions, numerical fluxes are zero for $i = 0$ and $i = N$, and D and D^* are the finite difference operators defined in section 3.1. Following the same strategy as in the continuous case, we can write the discrete versions of equilibrium schemes (21) and (22). In that case, the discrete equilibrium is the discrete Maxwellian $f_{eq,i} = \exp(\alpha + \beta\varepsilon_i)$ whose coefficients α and β are computed so as f_{eq} has the same discrete mass and energy as f.

It is now an easy matter to prove that these discrete collision operators satisfy weak formulations similar to those satisfied by the continuous model. Hence the discrete analogous properties of propositions 2.1, 2.2 and 2.4 are satisfied by all these discrete schemes.

4 Linear solvers

In this section, we present a strategy to solve the linear implicit schemes (7, 13, 16, 21, and 22), the unknown being f^{n+1}. The present method obey the conservation properties as soon as the approximation of the collision operator Q is conservative. Here, we deal with both the FPL equation (1) with $d = 2$ or 3 and the isotropic case (25). In this section, these models only differ by the dimension of the integration domain: $d = 2$ or 3 for (1) and $d = 1$ for (25).

We first assume that the collision operator is discretized with N velocity points and write the schemes (7, 13, 16) in the following matrix-forms:

$$\begin{array}{ll} \text{contracted scheme (7)} & L^c(f^n)f^{n+1} = f^n, \\ \theta\text{-scheme (13)} & L_\theta(f^n)f^{n+1} = f^n + \Delta t(1-2\theta)Q(f^n), \\ \text{log-linear scheme (16)} & L^l(f^n)f^{n+1} = f^n + \Delta t Q(f^n), \end{array}$$

where $L^c(f)$, $L_\theta(f)$ and $L^l(f)$ are the $N \times N$-matrices defined by

$$L^c(f)g = g - \Delta t\, q^c(f,g), \tag{30}$$

$$L_\theta(f)g = g - \Delta t\, \theta q(f,g), \tag{31}$$

$$L^l(f)g = g - \Delta t\, q^l(f,g), \tag{32}$$

for every vector g in $\mathbb{R}^N$.

We can also consider the following equivalent (and more convenient) δ-form

$$L^c(f^n)\delta f^n = \Delta t Q(f^n), \tag{33}$$

$$L_\theta(f^n)\delta f^n = \Delta t Q(f^n), \tag{34}$$

$$L^l(f^n)\delta f^n = \Delta t Q(f^n), \tag{35}$$

where $\delta f^n = f^{n+1} - f^n$. Equilibrium implicit schemes (21) and (22) are defined in δ-form by

$$L_\theta(f_{eq})\delta f^n = \Delta t Q(f^n), \tag{36}$$

$$L^l(f_{eq})\delta f^n = \Delta t Q(f^n). \tag{37}$$

Before going to the resolution of these linear systems, we give some of their important algebraic properties.

Proposition 4.1.
- *For all $\Delta t > 0$ and all $f > 0$, $L^l(f)$ is a positive definite self-adjoint matrix for the inner product with weight $\frac{1}{f}$, and, in particular, it is invertible.*

- *For all $\Delta t > 0$, $\theta \in [0,1]$, and all discrete Maxwellian f_{eq}, $L_\theta(f_{eq})$ is a positive definite self-adjoint matrix for the inner product with weight $\frac{1}{f_{eq}}$, and, in particular, it is invertible.*
- *For all $f > 0$, $L^c(f)$ and $L_\theta(f)$ are invertible if Δt is small enough.*

Proof. We do not give the proof of this proposition, since it is a direct consequence of the discrete versions of the properties of q^c, q, and q^l given in propositions 2.1, 2.2, and 2.4. □

These linear systems are non sparse, with generally large dimension. For a numerical resolution, one can mainly investigate three classes of methods: direct (as LU), iterative non-stationary (as Krylov methods), iterative stationary (as relaxation methods).

Direct methods have been used by Epperlein [9] for the isotropic FPL equation with LU factorization. However, it is clear that such solvers are not usable in multidimensional cases. The complexity of the algorithms is $O(N^3)$, and the memory storage for the matrices is $O(N^2)$. Even in the one dimensional case as isotropic equation, the total complexity for a given simulation time is asymptotically the same as the simple explicit scheme (see section 5.1).

Consequently, it is clear that iterative methods must be considered. At this stage, we want to point out a crucial fact: since iterative methods generally construct an approximate solution to the linear system, *it should be investigated whether conservation properties of the implicit schemes are preserved or not.* Indeed, it is questionable to construct perfectly conservative schemes if that property is destroyed by the linear solvers.

To make this point more precise, we define the $(d+2) \times N$ matrix M that associates to every vector $f \in \mathbb{R}^N$ its corresponding $(d+2)$-vector of moments Mf (*i.e* an approximation of $\int_{\mathbb{R}^d}(1, v, \frac{1}{2}|v|^2)^T f\, dv$). For instance, we can set $Mf = \sum_{i=1}^N (1, v_i, \frac{1}{2}|v_i|^2)^T f_i\, \Delta v^d$ for a regular discretization and $d = 2$ or 3. For the isotropic case, a possible definition is $Mf = \sum_{i=1}^N (1, \varepsilon_i)^T f_i \sqrt{\varepsilon_i} \Delta\varepsilon$. Since we assume the discrete collision operator Q to be conservative, the following relations are satisfied for every f and g in $\mathbb{R}^N$:

$$Mq^c(f,g) = 0, \qquad Mq(f,g) = 0 \quad \text{and} \quad Mq^l(f,g) = 0.$$

This implies the following relations for the matrices (30,31,32)

$$ML^c(f) = M, \qquad ML_\theta(f) = M \quad \text{and} \quad ML^l(f) = M.$$

Multiplying by M the δ-forms (33-37) of our implicit schemes is another way to check the conservation property $Mf^{n+1} = Mf^n$.

Consequently, the implicit schemes written under the δ-forms (33-37) are of the same type as the following general linear system in $\mathbb{R}^N$

$$Ax = b,$$

where the $N \times N$-matrix A and N-vector b satisfy

$$MA = M \text{ and } Mb = 0. \tag{38}$$

Then we consider the following problem: find an iterative solver such that if we start with an initial vector $x^{(0)}$ satisfying $Mx^{(0)} = 0$, then any iterate $x^{(k)}$ also satisfies $Mx^{(k)} = 0$. Such solvers will be called *conservative iterative linear solvers.* They guarantee that the implicit scheme is still conservative even if convergence to the exact solution of the linear system is not achieved.

In the sequel, we prove that Krylov subspace methods are conservative iterative linear solvers. In our study, these methods can be set in the following general frame (see [20] for an introduction to Krylov solvers)

Algorithm 4.1. *1. give $x^{(0)}$ such that $Mx^{(0)} = 0$ and set $r^{(0)} = b - Ax^{(0)}$;*

2. for $k = 1$ to K, find $x^{(k)}$ in the affine subspace $x^{(0)} + \mathcal{K}_k$, where

$$\mathcal{K}_k = \{r^{(0)}, Ar^{(0)}, \dots, A^{k-1}r^{(0)}\}.$$

The different versions of Krylov methods arise from different choices of $x^{(k)}$ in $\mathcal{K}_k$. We now prove the following

Proposition 4.2. *All iterative methods that can be set under the form of algorithm 4.1 are conservative. This means that we have $Mx^{(k)} = 0$ for every k.*

Proof. Using conservation property on the initial vector $x^{(0)}$ and property (38) yields

$$Mr^{(0)} = M(b - Ax^{(0)}) = 0 - Mx^{(0)} = 0,$$

and therefore $MA^p r^{(0)} = MAA^{p-1}r^{(0)} = MA^{p-1}r^{(0)} = \ldots = Mr^{(0)} = 0$ for every $p \geq 1$. Consequently, we have $M\mathcal{K}_k = \{0\}$, and necessarily $x^{(k)} \in x^{(0)} + \mathcal{K}_k$ implies $Mx^{(k)} = 0$. □

It is remarkable that this conservation property holds even if the linear system is not exactly solved. Another advantage of these methods is well known: they are "matrix-free", i.e. the matrix A only appears in matrix-vector products Ay in the solver. Thus we do not need to form and store the matrix A. Moreover in our case, if the quantities $q^c(f,g)$, $q(f,g)$ and $q^l(f,g)$ can be computed in $O(N)$ operations, then this is also possible for products Ay in Krylov solvers, since $A = L^c(f^n), L_\theta(f^n)$, or $L^l(f^n)$ which are given by (30,31,32).

Finally, we would emphasize the fact that all iterative solvers are not necessarily conservative. For instance, consider the general class of stationary methods, such as Jacobi, Gauss-Seidel, or relaxations methods. These methods consist in the splitting matrix A into $A = E - F$, and can be written as

```
ALGORITHM 4.2.  1. give $x^{(0)}$ such that $Mx^{(0)} = 0$ and set $r^{(0)} = b - Ax^{(0)}$;
    2. for $k = 0$ to $K$ solve $Ex^{(k+1)} = b + Fx^{(k)}$.
```

In this case, we only have $MEx^{(k+1)} = 0 + MFx^{(k)} = M(E-A)x^{(k)} = MEx^{(k)} - Mx^{(k)}$. Even if we assume $Mx^{(k)} = 0$, there is no reason to have $Mx^{(k+1)} = 0$. For instance with Jacobi method, matrix E is the diagonal part of A, and we do not have any information about the value of ME.

Among the above Krylov methods, we mainly use the GMRES and the CG methods. The CG method is used to solve linear systems whose matrices are positive definite self-adjoint. As stated in proposition 4.1, this is the case for the log-linear and equilibrium log-linear schemes (16, 22) and the equilibrium θ-scheme (21). This is a real advantage, since the CG method is one of the most efficient Krylov solver (in terms of CPU cost and memory storage). For the other schemes, we simply use the GMRES method.

5 Complexity and numerical tests

5.1 Complexity of the algorithms

In this section, we assume that for an arbitrary number of dimensions $d \geq 1$, the number of operations for computing $q^c(f,g), q(f,g)$, and $q^l(f,g)$ is $O(N)$. This is true for $d = 1$ in the case of isotropic equation with Coulomb or Maxwell potential (see [5]). For $d \geq 2$, cost reductions to $O(N \log N)$ can be obtained through rapid algorithms as Multipole methods [15] and spectral methods [18]. The cost $O(N)$ could be reached by an extension of isotropic wavelet methods [1] to the multi-dimensional case. We also assume that the velocity domain is discretized with a step Δv or $\Delta\varepsilon = \frac{1}{n}$ in each direction and a total number of points of $N = n^d$. For a problem with a time scale τ, we want to compare the complexity of the usual Euler explicit scheme (5) and our linear implicit schemes for a given simulation time T.

For the explicit scheme, one iteration requires only one evaluation of $Q(f)$, which costs $O(N)$. Assume that the CFL condition imposes a time step $\Delta t = O(\frac{1}{n^2})$ (this seems to be true from numerical experiments and has been proved in the isotropic case [4]). Then the number of iterations is $\frac{T}{\Delta t} = O(n^2)$. Consequently, the total complexity is $O(N^{1+2/d})$.

Table 2: Complexity of explicit and linear implicit schemes for $d = 1, 2, 3$.

	$d = 1$	$d = 2$	$d = 3$
explicit	$O(N^3)$	$O(N^2)$	$O(N^{5/3})$
implicit	$KO(N)$	$KO(N)$	$KO(N)$

For linear implicit schemes with Krylov solvers, the cost of one Krylov iteration is $O(N)$ (since only evaluations of $q^c(f, g), q(f, g)$, and $q^l(f, g)$ are needed). We assume that K iterations of the Krylov solver are necessary to get a correct approximation of the time iterate. Thus the cost of one time iteration is $KO(N)$. Finally, if the time step can be set to τ, the number of time iterations is $\frac{T}{\tau}$, and therefore, the total complexity is $KO(N)$.

In table 5.1, we give a comparison of these complexities for $d = 1, 2, 3$. We can see that for $d = 1$, there is an important gain even if $K = N$. For $d = 2$, there is still a gain if $K << N$. If $K = O(N)$, the explicit and implicit schemes have the same asymptotic cost $O(N^2)$. For $d = 3$, we have a significant gain only if $K << N^{2/3}$. Therefore, the efficiency of the implicit schemes decreases as the dimension increases. Consequently, the reduction of the number K of Krylov iterations becomes necessary in multi-dimensional cases. This could be done by using adapted preconditioning techniques. As noted by [7], such techniques can render K virtually independent of N. In that case, our implicit schemes would always be advantageous as compared to the explicit scheme. This is the subject of a future work.

5.2 Numerical tests

In order to check the properties of the implicit schemes that we introduced in this work, we present various numerical tests with both the Maxwellian and Coulombian potentials.

Maxwellian potential: comparison with an exact solution.

First, in the case of Maxwellian potential, it is proved in [14] that

$$f(t, \varepsilon) = \frac{\rho}{(2\pi T)^{3/2}} \exp(-\frac{\varepsilon}{2T}) \left(1 + \frac{11}{120} \left(\left(\frac{\varepsilon}{T}\right)^2 - 10\frac{\varepsilon}{T} + 15\right) \exp(-8\rho t)\right)$$

is an exact solution of (25). Then, in figure 1 and 2, we compare the numerical solution obtained with the second-order θ-scheme (13) to this exact solution. The energy domain is $[0, 2]$ discretized with 500 points. The parameters of the exact solution are $\rho = 2$ and $T = 0.01$. The time step used with the implicit scheme is about 300 times the time step required by an explicit computation. In figure 1, we plot f as a function of $v = \sqrt{\varepsilon}$ at different time steps. We observe that the numerical solution is very close to the exact one, and the equilibrium is reached in only ten iterations. This corresponds to a final physical time equal to $t_{max} = 10$. This shows that the dynamics described by the exact FPL equation can efficiently be simulated with a much larger time step than those of usual explicit schemes. This is also clear on figure 2 where the kinetic entropy is plotted. The slight difference that can be observed in the stationary regime is due to the velocity discretization itself. In fact the discrete Maxwellian is different from the exact one.

Coulombian potential: regular energy discretization.

The second test case uses Coulombian potential with the so called Rosenbluth initial data:

$$f^0(\varepsilon) = 0.01 \exp(-10((\sqrt{\varepsilon} - 0.3)/0.3)^2).$$

On figure 3, we plot the kinetic entropy obtained by the explicit scheme (5) with time step Δt_{exp}, which is the largest step ensuring the stability of the scheme. This entropy is compared with that obtained by the contracted implicit scheme (7). For this last scheme, we take a time step $\Delta t_{imp} = 50\Delta t_{exp}$. We use a regular energy discretization of $[0, 2]$ with $N = 100$ points. Tolerance for CG algorithm is 10^{-6}. The dynamics is well described by the implicit scheme even if the time step is much larger. However the gain in terms of CPU time is not obtained unless the number of energy points is sufficiently large. Indeed, the implicit scheme is really advantageous for $N \geq 500$ only. This is clearly shown by figure 4 where the CPU time versus N is plotted for explicit and implicit schemes. According to what we explained in section 5.1, the CPU time of the explicit scheme behaves as $O(N^3)$. A contrario, the implicit scheme only requires $O(KN)$ operations, where K is the number of iterations in the Krylov procedure. It is known that $K \leq N$, and we observe that on this test case, K is much smaller than N. The test on figure 4 confirms that the numerical complexity of the implicit scheme behaves like $O(N^2)$. Indeed, for $N \geq 500$, the slope of the curve is 1.95 for the implicit scheme whereas it is 2.8 for the explicit scheme.

Coulombian potential: irregular energy discretization.

The use of irregular discretizations could be necessary when one is only interested in the dynamics of a localized region in the velocity space for instance, or in general when using adaptive meshes. In such cases, one can expect a more important gain when using implicit schemes, since the time step of the explicit scheme is constrained by the smallest step size of the irregular mesh.

To illustrate this fact, we perform the previous Rosenbluth test case using an irregular energy discretization where the discrete energy points are $\varepsilon_i = 2(\frac{i}{N})^3$ for $i = 1$ to N. Here we use the equilibrium θ-scheme (11) with $\theta = 1$, and plot on figure 5 the CPU time versus N for explicit and implicit schemes. We observe that for $N = 50$ points, the implicit scheme is 1.5 times faster than the explicit scheme. The CPU time is much more reduced for larger N: it is divided by a factor 6 for $N = 100$ points and by a factor 33 for $N = 200$.

To go further in this direction, we mention that in most cases, the evaluation of the collision operator is very expensive (the present isotropic case is very particular). Example are 3D geometries with classical or relativistic particles. In these cases, the arguments of the kernel K cannot be separated in a simple way, and the use of fast evaluation algorithms is not as efficient as in the isotropic case. Consequently, in such situations it is attractive to reduce the number of time iterations by using implicit schemes. To illustrate this fact, we consider the isotropic case in which we do not use any fast algorithm: we directly compute the collision operator with formula like (27-28), which costs $O(N^2)$ operations. Of course, this is just used as a prototype of the previously mentioned realistic situations. In the case of the previous irregular grid with the same initial data, we obtain that for $N = 50$ points, the same implicit scheme is 2 times faster than the explicit scheme. The CPU time is much more reduced for $N = 100$ since it is divided by a factor 10.

Comparison of the different implicit schemes.

For the same initial data, we compare the different implicit schemes studied in this work to the explicit one: contracted (7), θ-scheme (13), log-linear implicit scheme (16), equilibrium θ-scheme (21), equilibrium log-linear scheme (22). On figure 6 we plot the fourth order moment $\int f(\varepsilon)\varepsilon^2\sqrt{\varepsilon}\,d\varepsilon$ as a function of time. We only compare first order (in time) schemes with $\Delta t_{imp} = 50\Delta t_{exp}$ and $N = 100$ on a regular grid. We observe that the equilibrium θ-scheme with $\theta = 1$ is the most accurate in this case. Despite its good mathematical properties, the log-linear scheme is not sufficiently accurate. This is probably due to the linearization of the log function.

On figure 7, we compare our second order (in time) θ-scheme (with $\theta = 0.5$) to a second order explicit scheme (Runge-Kutta scheme). We use the same parameters as on figure 6. We can see that the two results are in a very good agreement.

Conservation properties.

Finally, we point out that all the schemes proposed in this paper are perfectly conservative, up to the machine error. For instance, on the second test case, the variation amplitude of density and energy during the time evolution is about 10^{-16}, as for the explicit scheme. This confirms that the approximate resolution of the linear systems does not affect the conservation properties.

6 Conclusion

We have constructed various linear implicit schemes to solve the homogeneous FPL equation. It has been shown that these schemes possess important properties of conservation and entropy. Moreover, numerical tests in the isotropic case have shown a significant gain in terms of CPU time with the same accuracy, when compared with usual explicit schemes.

As we pointed out in the introduction, our strategy applies to multidimensional collision operators, but still some details must be investigated in that case. Indeed, the involved linear systems are much larger, and suitable preconditioners are required. We recall that in a Krylov method with a good preconditioner, the number of iterations is independent of N. Thus the complexity of our implicit schemes could be reduced to $O(N)$ only. Moreover, the rapid matrix-vector product that has been used in the isotropic case cannot be applied to the multidimensional cases. Therefore, to get a fast implicit solver in several dimensions, other acceleration techniques are required. We believe that this can be done using multipole or wavelet techniques [15, 1].

Note that for the inhomogeneous (space dependent) case, our implicit schemes should be more efficient. Indeed, in many inhomogeneous situations the transition phase is rapid and does not need to be accurately described. In such cases, our implicit schemes allow to directly reach the hydrodynamic behavior with a few time steps.

Finally, the extension of our strategy to other collision operators (including the relativistic and quantum effects for instance) is the subject of a future work.

Acknowledgments: Support by the European network HYKE (EC contract HPRN-CT-2002-00282) is acknowledged.

References

[1] X. Antoine and M. Lemou. Wavelet approximation of a collision operator in kinetic theory. *C.R Acad. Sci.*, Ser. I(337), 2003.

[2] P.L. Bathnagar, E.P. Gross, and M. Krook. A model for collision processes in gases. I. small amplitude processes in charged and neutral one-component systems. *Phys. Rev.*, 94:511–525, 1954.

[3] Yu. A. Berezin, V. N. Khudick, and M. S. Pekker. Conservative finite-difference schemes for the Fokker-Planck equation not violating the law of an increasing entropy. *J. Comput. Phys.*, 69(1):163–174, 1987.

[4] C. Buet and S. Cordier. Conservative and entropy decaying numerical scheme for the isotropic Fokker-Planck-Landau equation. *J. Comput. Phys.*, 145(1):228–245, 1998.

[5] C. Buet and S. Cordier. Numerical analysis of the isotropic Fokker-Planck-Landau equation. *J. Comput. Phys.*, 179(1):43–67, 2002.

[6] C. Buet, S. Cordier, P. Degond, and M. Lemou. Fast algorithms for numerical, conservative, and entropy approximations of the Fokker-Planck-Landau equation. *J. Comput. Phys.*, 133(2):310–322, 1997.

[7] L. Chacón, D. C. Barnes, D. A. Knoll, and G. H. Miley. An implicit energy-conservative 2D Fokker-Planck algorithm. II. Jacobian-free Newton-Krylov solver. *J. Comput. Phys.*, 157(2):654–682, 2000.

[8] P. Degond and B. Lucquin-Desreux. An entropy scheme for the Fokker-Planck collision operator of plasma kinetic theory. *Numer. Math.*, 68:239–262, 1994.

[9] E.M. Epperlein. Implicit and conservative difference scheme for the Fokker-Planck equation. *J. Comput. Phys.*, 112(2):291–297, 1994.

[10] F. Filbet and L. Pareschi. A numerical method for the accurate solution of the Fokker-Planck-Landau equation in the nonhomogeneous case. *J. Comput. Phys.*, 179(1):1–26, 2002.

[11] E. Gabetta, L. Pareschi, and G. Toscani. Relaxation Schemes for Non Linear Kinetic Equations. *SIAM J. Numer. Anal.*, 34(6):2168–2194, 1997.

[12] A. Girard, C. Lécot, and K. Serebrennikov. Numerical simulation of the plasma of an electron cyclotron resonance ion source. *J. Comput. Phys.*, 191(1):228–248, 2003.

[13] R. J. Kingham and A. R. Bell. An implicit Vlasov-Fokker-Planck code to model non-local electron transport in 2-D with magnetic fields. *J. Comput. Phys.*, 194(1):1–34, 2004.

[14] M. Lemou. Exact solutions of the Fokker-Planck equation. *C. R. Acad. Sci Série 1*, 319:579–583, 1994.

[15] M. Lemou. Multipole expansions for the Fokker-Planck-Landau operator. *Numer. Math.*, 78(4):597–618, 1998.

[16] M. Lemou and L. Mieussens. Fast implicit schemes for the Fokker-Planck-Landau equation. *C. R. Acad. Sci. Paris, Ser. I*, 338:809–814, 2004.

[17] E. M. Lifchitz and L. P. Petaevski. *Kinetic Theory.* MIR, Moscow, 1979.

[18] L. Pareschi and B. Perthame. A Fourier spectral method for homogeneous Boltzmann equations. In *Proceedings of the Second International Workshop on Nonlinear Kinetic Theories and Mathematical Aspects of Hyperbolic Systems (Sanremo, 1994)*, volume 25, pages 369–382, 1996.

[19] L. Pareschi, G. Russo, and G. Toscani. Fast spectral methods for the Fokker-Planck-Landau collision operator. *J. Comput. Phys.*, 165(1):216–236, 2000.

[20] Y. Saad. *Iterative methods for sparse linear systems.* available online at `http://www-users.cs.umn.edu/~saad`, 2000.

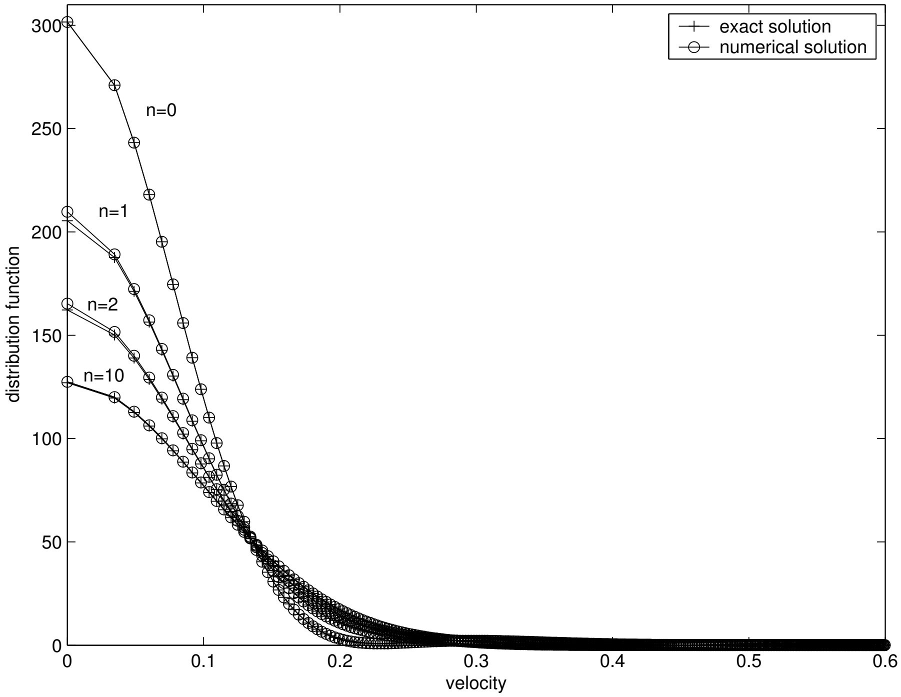

Figure 1: Exact solution and implicit scheme for Maxwellian potential: distribution function at different time steps.

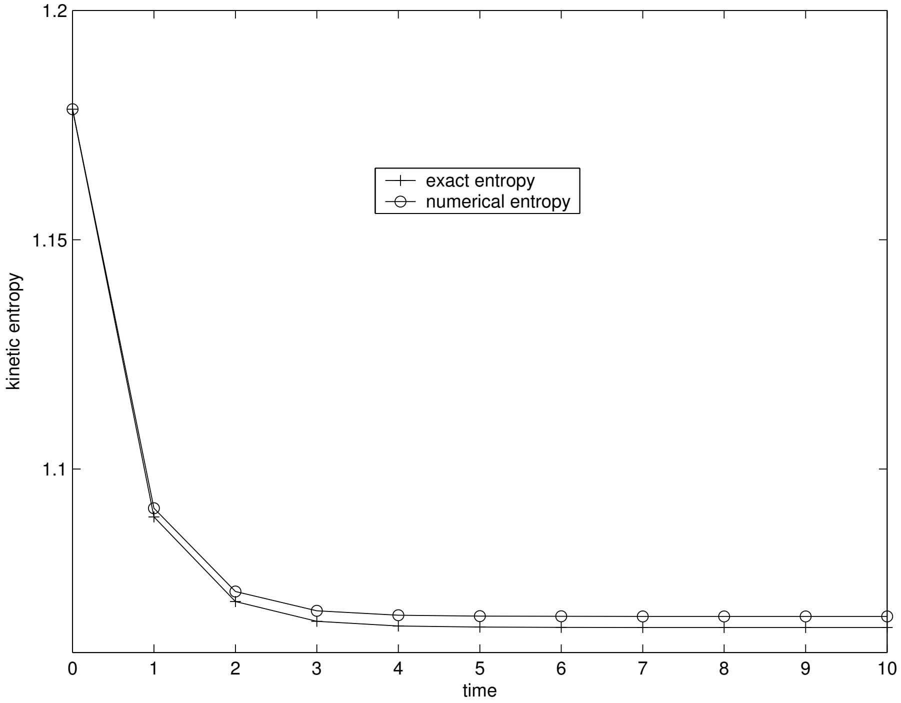

Figure 2: Exact solution and implicit scheme for Maxwellian potential: entropy.

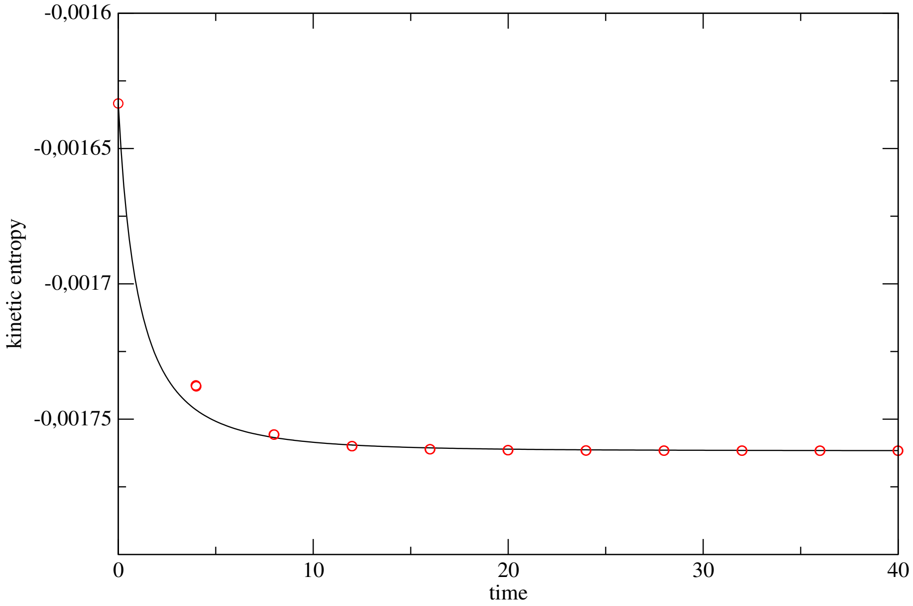

Figure 3: Kinetic entropy for explicit (-) and contracted implicit scheme (7) (o) for Coulombian potential. Case of a regular grid.

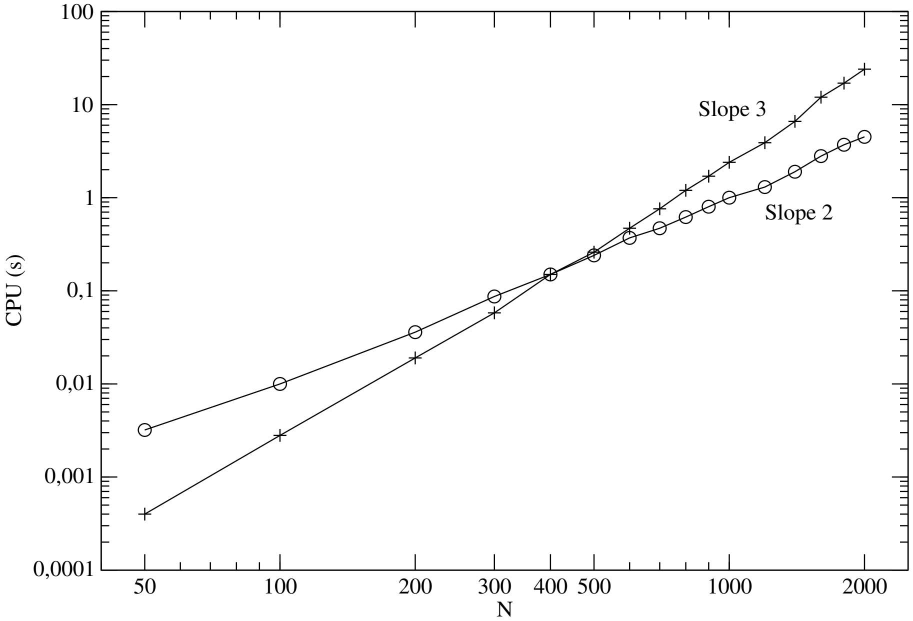

Figure 4: CPU cost of explicit (+) and contracted implicit scheme (7) (o) versus the number N of energy points in a logarithmic scale (Coulombian potential). Tolerance for CG is 10^{-6}, time step in implicit scheme is 3. Regular energy grid is used.

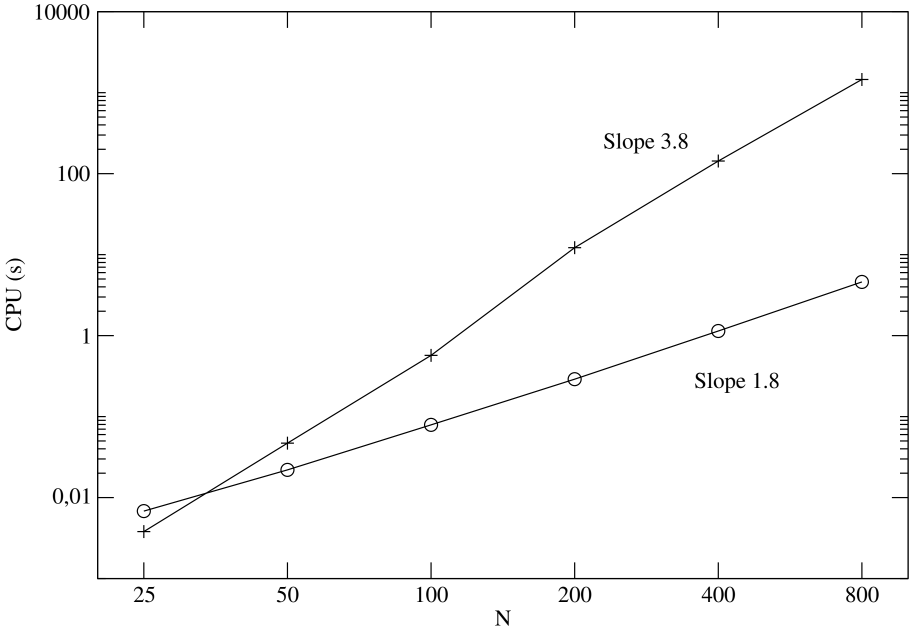

Figure 5: CPU cost of explicit (+) and equilibrium θ-scheme (11) with $\theta = 1$ (o) versus the number N of energy points in a logarithmic scale (Coulombian potential). Tolerance for CG is 10^{-6}, time step in implicit scheme is 3. Irregular energy grid is used.

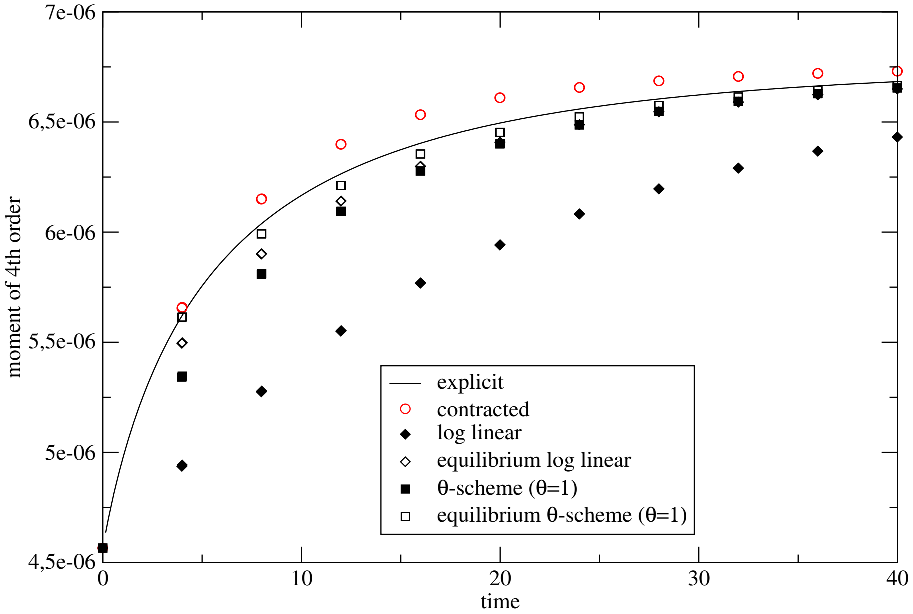

Figure 6: Fourth order moment for different implicit schemes.

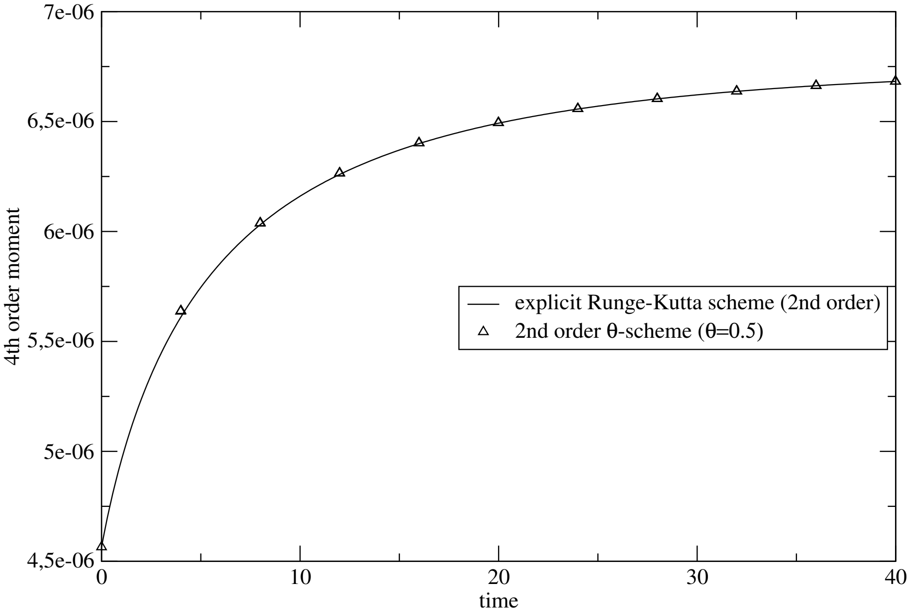

Figure 7: Fourth order moment for second order (in time) schemes.