QUARTERLY OF APPLIED MATHEMATICS
VOLUME LVII, NUMBER 3
SEPTEMBER 1999, PAGES 521–541

ENTROPY PRODUCTION AND THE RATE OF CONVERGENCE TO EQUILIBRIUM FOR THE FOKKER-PLANCK EQUATION

By

G. TOSCANI

Dept. of Mathematics, University of Pavia, via Abbiategrasso 209, 27100 Pavia, Italy

Abstract. We reckon the rate of exponential convergence to equilibrium both in relative entropy and in relative Fisher information, for the solution to the spatially homogeneous Fokker-Planck equation. The result follows by lower bounds of the entropy production which are explicitly computable. Second, we show that the Gross's logarithmic Sobolev inequality is a direct consequence of the lower bound for the entropy production relative to Fisher information. The entropy production arguments are finally applied to reckon the rate of convergence of the solution to the heat equation towards the fundamental one in various norms.

1. Introduction. Rarefied gas dynamics is usually described by the Boltzmann equation

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f = J(f) \tag{1.1}$$

where $f(x, v, t)$ is the density of particles which at time $t \in \mathbf{R}^+$ and point $x \in \mathbf{R}^n$, move with velocity $v \in \mathbf{R}^n$, $n \geq 1$. J is a quadratic operator, described in [TM80] and [Ce88].

The mathematical difficulties of solving the Boltzmann equation have made it worthwhile to look for alternatives or modifications that convey essentially an equivalent amount of physical information. The most widely known of these alternative collision models is usually called the Bhatnagar Gross and Krook (BGK) model [BGK54]

$$J_1(f) = \nu(M^f - f). \tag{1.2}$$

In (1.2) $M^f(x, v, t)$ denotes the locally Maxwellian distribution function,

$$M^f(x, v, t) = \rho(x, t)[2\pi\theta(x, t)]^{-n/2} \exp\left\{\frac{-|v - u(x, t)|^2}{2\theta(x, t)}\right\}, \tag{1.3}$$

where

$$\rho(x, t) = \int_{\mathbf{R}^n} f(x, v, t)\, dv \tag{1.4}$$

Received April 14, 1997.
1991 *Mathematics Subject Classification.* Primary 82C40.

is the mass density,

$$u(x,t) = \frac{1}{\rho}\int_{\mathbf{R}^n} v f(x,v,t)\,dv \tag{1.5}$$

is the mean velocity, and $T = \theta/R$ is the local temperature defined by

$$T(x,t) = \frac{1}{n\rho R}\int_{\mathbf{R}^n} |v-u|^2 f(x,v,t)\,dv. \tag{1.6}$$

R is the Boltzmann constant, the presence of which is due to the fact that the energy and temperature are, by tradition, commonly expressed in different units. For the purpose of this paper this is inconvenient, and we will assume henceforth that the temperature T is expressed in the same units as θ, i.e., in Eq. (1.6) with $R = 1$. The parameter ν, the collision frequency, is taken to be a function of ρ, u, T, and plays the role of the reciprocal of the relaxation time in this theory.

The BGK model satisfies the usual conservation of mass, mean velocity, and temperature, and

$$\int_{\mathbf{R}^n} \log f J_1(f)\,dv \le 0. \tag{1.7}$$

Another model is offered by a Fokker-Planck collision term,

$$J_2(f) = \gamma \sum_{k=1}^{n} \left\{ \frac{\partial^2 f}{\partial v_k^2} + \frac{1}{\theta}\frac{\partial}{\partial v_k}[(v_k - u_k) f] \right\}. \tag{1.8}$$

The one-particle friction constant γ is usually assumed to be a function of ρ, u, T.

The Fokker-Planck equation appears in many different contexts. It was originally derived for the distribution function of a Brownian particle in a fluid [Ch43], and is applicable in a more general form to a plasma [ChC58]. A detailed investigation of this model has been performed by Frisch, Helfand, and Lebowitz [FHL60] in connection with the kinetic theory of liquids. As shown by Cercignani [Ce88], it provides also a good description of the grazing collisions in a gas. The Fokker-Planck collision operator has the usual conservation properties of mass, mean velocity, and temperature, and

$$\int_{\mathbf{R}^n} \log f J_2(f)\,dv \le 0. \tag{1.9}$$

It is interesting to remark that, if the friction γ is taken to be proportional to the pressure $p = \rho\theta$, then $J_2(f)$ has the same kind of nonlinearity (quadratic) as the true Boltzmann equation.

One of the most significant problems connected with the Boltzmann equation is the reckoning of the rate at which, in the spatially homogeneous situation, the solution to the initial value problem converges to equilibrium. This is important for the following reason. The unit time scale relevant for Eq. (1.1) is the mean time between collisions. This time scale is much shorter than the time scale governing macroscopic transport phenomena, so that it is commonly believed that the spatially homogeneous equation governs the rate of approach to local equilibrium even in nonhomogeneous settings.

The relaxation toward equilibrium is usually assumed to be a consequence of the monotonicity in time of the Boltzmann H-functional,

$$H(f) = \int_{\mathbf{R}^n} f \log f \, dv. \tag{1.10}$$

Given a collisional operator $J_i(f)$, the entropy production at f is defined by

$$-\int_{\mathbf{R}^n} \log f J_i(f) \, dv, \tag{1.11}$$

since at $f(t)$ this equals $\frac{d}{dt} H[f(t)]$, when $f(t)$ is a solution to

$$\frac{\partial f(t)}{\partial t} = J_i[f(t)] \tag{1.12}$$

and the integrand in (1.11) is integrable at $f(t)$.

A more precise statement of the Boltzmann H-theorem is that

$$-\int_{\mathbf{R}^n} \log f(t) J_i[f(t)] \, dv \geq 0, \tag{1.13}$$

with equality exactly when $f = M^f$.

According to inequalities (1.7) and (1.9), both the BGK and Fokker-Planck models satisfy the Boltzmann H-theorem.

To see how the Boltzmann H-theorem can be strengthened to yield information on the rate at which convergence to equilibrium occurs, introduce the relative entropy $D(f)$ of f, always with respect to M^f, by

$$D(f) = \int_{\mathbf{R}^n} \left[\frac{f}{M^f}\right] \log \left[\frac{f}{M^f}\right] M^f \, dv. \tag{1.14}$$

Because $\log M^f$ is quadratic,

$$D(f) = H(f) - H(M^f).$$

It follows from Jensen's inequality that $D(f) \geq 0$, with equality exactly when $f = M^f$; this is known as Gibb's lemma. More is true; when $D(f)$ is close to zero, then f is close to M^f in the L^1-norm. This is expressed by the Csiszar-Kullback inequality [Cs62] and [Ku67],

$$\|f - M^f\|^2_{L^1(\mathbf{R}^n)} \leq 2D(f). \tag{1.15}$$

Hence, we can use entropy production to control the strong L^1-convergence to equilibrium. It is clear that a precise statement about the relaxation to equilibrium in relative entropy depends on the possibility of obtaining a lower bound of the entropy production of the form

$$-\int_{\mathbf{R}^n} \log f J_i(f) \, dv \geq \phi[D(f)], \tag{1.16}$$

where the function $\phi(r)$ is bounded away from zero whenever r is strictly positive.

The choice of the model operator J_1 provides a useful example. The spatially homogeneous BGK model is linear in consequence of the conservation of moments. If f denotes

the initial density, then

$$\frac{\partial f(t)}{\partial t} = \nu[M^f - f(t)]. \tag{1.17}$$

Since $\log M^f$ is quadratic and the solution conserves mass, mean velocity, and temperature,

$$\begin{aligned}\nu \int_{\mathbf{R}^n} \log f(t)[M^f - f(t)]\, dv &= \nu \int_{\mathbf{R}^n} [\log f(t) - \log M^f][M^f - f(t)]\, dv \\ &= -\nu \int_{\mathbf{R}^n} f(t) \log\left[\frac{f(t)}{M^f}\right] dv - \nu \int_{\mathbf{R}^n} M^f \log\left[\frac{M^f}{f(t)}\right] dv.\end{aligned}$$

Because the second term on the right is nonpositive, we obtain

$$-\int_{\mathbf{R}^n} \log f(t) J_1[f(t)]\, dv \geq \nu D[f(t)], \tag{1.18}$$

which implies exponential convergence in relative entropy at a rate ν. It is a fact that, due to the simplicity of the BGK model in the spatially homogeneous case, the exponential L^1-convergence to equilibrium can be obtained directly. Indeed

$$f(t) - M^f = (f - M^f)e^{-\nu t}. \tag{1.19}$$

The rate of exponential convergence to equilibrium for the spatially homogeneous Fokker-Planck or Boltzmann equation for initial data far from equilibrium, and for general kernels, is not yet known.

Lower bounds for the entropy production for the Kac model, and for the Boltzmann and Fokker-Planck-Landau equations and their connection with the speed of convergence to equilibrium, have first been investigated by Desvillettes [De89]. Subsequently, Carlen and Carvalho recovered entropy production lower bounds in the form (1.16), first for a model Boltzmann equation [CC92] and second for the hard sphere model [CC94]. Using these lower bounds, they show that the rate at which strong L^1-convergence to equilibrium occurs is uniform in wide classes of initial data.

Few papers are concerned with the exponential rate of convergence. For intermolecular forces harder than Maxwellian ones, and in the presence of a cut-off, Arkeryd [Ar88] obtained stability results in L^1. These results were extended to pseudomaxwellian molecules by Wennberg [We93]. Here the method of proof is based on the spectral theory of the linearized collision operator, and gives exponential convergence to equilibrium if the initial data belong to an appropriately small neighborhood of the equilibrium itself.

Such an approach, while fully successful in establishing exponential convergence, gives no information as to what the exponential rate might be.

An explicitly computable rate of exponential convergence has been obtained by Gabetta, Toscani, and Wennberg [GTW95] for the Kac model and for Maxwellian molecules in a metric equivalent to the weak-* convergence of measures. Very recently, the optimal rate of exponential convergence for the aforementioned models in the strong L^1-norm has been derived by Carlen, Gabetta, and Toscani [CGT96]. This result is the first showing that the spectral gap in the linearized collision operator does govern the rate of approach to equilibrium for the full nonlinear Boltzmann equation, even for initial data that is far from equilibrium.

The structure of the paper is as follows. In the next section we introduce Fisher information and the logarithmic Sobolev inequality by Gross [Gr75]. In Sec. 3 we derive lower bounds for the entropy production. These bounds are a consequence of the logarithmic Sobolev inequality, and permit us to obtain the sharp rate of exponential convergence to equilibrium. In Sec. 4 we extend our analysis to the relative Fisher information. We show that Fisher information is nonincreasing in time, and we find a lower bound for the entropy production that enables us to find the sharp rate of exponential convergence to equilibrium. In Sec. 5 we briefly comment on some possible extension of the method to other functionals. In Sec. 6 we give a new proof of the logarithmic Sobolev inequality that follows as a direct consequence of the results of Sec. 4. Applications of the same arguments to the heat equation are discussed in Sec. 7.

2. Fisher information and the logarithmic Sobolev inequality. Let $f(v)$ be a probability density on $\mathbf{R}^n$. For $\sqrt{f} \in H^1$, we define the Fisher information of f, $L(f)$ by

$$L(f) = 4\int_{\mathbf{R}^n} |\nabla f^{1/2}(v)|^2\, dv = \int_{\mathbf{R}^n} |\nabla \log f(v)|^2 f(v)\, dv. \tag{2.1}$$

This quantity was introduced by Fisher [Fi25] in his theory of sufficient statistics. In the kinetic theory of rarefied gases, after the paper by McKean on the Kac equation [Mk66], $L(f)$ is often named Linnik's functional [BT92], [LT95]. Suppose f has the second moment bounded. The relative Fisher information $J(f)$ of f with respect to M^f is defined by

$$\begin{aligned} J(f) &= 4\int_{\mathbf{R}^n} \left|\nabla\left[\frac{f(v)}{M^f(v)}\right]^{1/2}\right|^2 M^f(v)\, dv \\ &= \int_{\mathbf{R}^n} |\nabla \log f(v) - \nabla \log M^f(v)|^2 f(v)\, dv. \end{aligned} \tag{2.2}$$

Easy computations reveal that

$$J(f) = L(f) - L(M^f) \tag{2.3}$$

and that $L(M^f) = n/\theta$.

The Fisher information functional is a convex functional closely related to the Boltzmann H-functional. A consequence of (2.3) is that $L(f) \geq L(M^f)$ with equality exactly when $f = M^f$; this gives the analog of the Gibb's lemma for H. Two further connections are important here. Let M_α denote the Maxwellian function in $\mathbf{R}^n$ of unit mass, zero mean velocity, and temperature α. First, Fisher information comes out in a natural way in evaluating the difference between $H(f \star M_\alpha)$ and $H(f)$, because of the relationship [St59], [Bl65]

$$L(f \star M_\alpha) = -2\frac{dH(f \star M_\alpha)}{d\alpha}. \tag{2.4}$$

The second connection is a consequence of the Gross logarithmic Sobolev inequality [Gr75], which asserts that

$$\int_{\mathbf{R}^n} |f|^2 \log |f|^2 M_\alpha\, dv \leq 2\alpha \int_{\mathbf{R}^n} |\nabla f|^2 M_\alpha\, dv \tag{2.5}$$

for all f such that $\int_{\mathbf{R}^n} |f|^2 M_\alpha\, dv = 1$.

Because $M_\alpha\, dv$ is a probability measure, the left-hand side is always well defined, infinity is admitted as a possible value, and is finite with (2.5) holding whenever the right-hand side is finite.

Inequality (2.5) can be rewritten with respect to the Lebesgue measure. In this case, for all functions $g \in H^1$,

$$\begin{aligned}\int_{\mathbf{R}^n} |g|^2 \log(|g|^2/\|g\|^2_{L^2(\mathbf{R}^n)})\, dv + \Big(n + \frac{n}{2}\log 2\pi a\Big) \int_{\mathbf{R}^n} |g|^2\, dv \\ \leq 2a \int_{\mathbf{R}^n} |\nabla g|^2\, dv\end{aligned} \tag{2.6}$$

for all $a \geq 0$.

The statement concerning cases of equality in (2.6) was established by Carlen [Ca91]. There is equality in (2.6) if and only if g^2 is a multiple and translate of M_a.

Inequality (2.6) has been recently used by the author to investigate the asymptotic behaviour of the solution to the heat equation [To96].

Given a probability density $h(v)$, with zero mean velocity and temperature θ, choose $f^2 = h/M_\theta$ in (2.5). Then, the left-hand side is nothing than the relative entropy $D(h)$, whereas the right-hand side is a multiple of the relative Fisher information. Thus, the logarithmic Sobolev inequality can be rewritten as

$$D(h) \leq \frac{\theta}{2} J(h). \tag{2.7}$$

By a simple change of variable, we conclude that (2.7) holds even if the bulk velocity is different from zero. Hence, (2.7) provides an upper bound for the relative entropy in terms of both the temperature and the Fisher information. A similar inequality can be derived by (2.6), inserting $g^2 = h$. We obtain for all $a > 0$

$$H(h) + \Big(n + \frac{n}{2}\log 2\pi a\Big) \leq \frac{a}{2} L(h). \tag{2.8}$$

Moreover, there is equality in (2.8) if and only if h is a multiple and translate of M_a.

3. Entropy production for the Fokker-Planck equation. This section contains results on the sharp rate of convergence to equilibrium in relative entropy of the solution to the Fokker-Planck equation. Most of the preliminary results are known, and can be found, for example, in a recent paper by Carlen and Soffer [CS91] devoted to the entropy production in the central limit theorem.

Here and in the next section we will be mainly interested in studying the time evolution of different Lyapunov functionals of the solution (the H-functional here, Fisher information and others there). Because our goal is different from that of [CS91], many of their results need to be rewritten.

Let us consider the initial value problem for the (spatially homogeneous) Fokker-Planck equation

$$\frac{\partial f(t)}{\partial t} = \gamma \sum_{k=1}^{n} \left\{ \frac{\partial^2 f}{\partial v_k^2} + \frac{1}{\theta}\frac{\partial}{\partial v_k}[(v_k - u_k) f] \right\} \tag{3.1}$$

with the initial condition

$$f(v, t = 0) = \varphi(v). \tag{3.2}$$

We remark that Eq. (3.1) is usually known as the Fokker-Planck equation linked to the Ornstein-Uhlenbeck process [Ri84]. In its physical meaning, the distribution function should be nonnegative and have finite moments up to the second order. Suppose the initial value φ has mean velocity u and temperature θ. Easy computations show that the mass, mean velocity, and temperature of the solution to (3.1) do not change with time. It is convenient to normalize f to be a probability density instead of a mass density, and change problem (3.1–3.2) to a dimensionless form. To do this, we will introduce the dimensionless variables $\overline{v}, \overline{t}$, and the dimensionless functions $\overline{\varphi}, \overline{f}$ defined by the formulas

$$\begin{gathered} \overline{v} = \frac{v-u}{\sqrt{\theta}}, \qquad \overline{t} = \frac{\gamma}{\theta} t, \\ \varphi(v) = \rho\theta^{-n/2}\overline{\varphi}(\overline{v}), \quad f(v,t) = \rho\theta^{-n/2}\overline{f}(\overline{v}, \overline{t}). \end{gathered} \tag{3.3}$$

Substituting (3.3) into (3.1), (3.2), carrying out elementary calculations, and then omitting the bars, we obtain that the function $f(v,t)$ will now satisfy the equation

$$\frac{\partial f(t)}{\partial t} = \sum_{k=1}^{n} \left\{ \frac{\partial^2 f}{\partial v_k^2} + \frac{\partial}{\partial v_k}(v_k f) \right\}, \tag{3.4}$$

with the initial condition $\varphi(v)$ and consequently $f(v,t)$ satisfying the following simple normalization conditions:

$$\rho = 1, \quad u = 0, \quad \theta = 1. \tag{3.5}$$

The normalization (3.5) corresponds to the equilibrium Maxwell distribution

$$M(v) = M_1(v) = (2\pi)^{-n/2} \exp\left\{ -\frac{v^2}{2} \right\}. \tag{3.6}$$

Let φ be any probability density on $\mathbf{R}^n$ satisfying conditions (3.5). Let X be any random variable with this density, and let W be any independent Gaussian random variable with density M given by (3.6). For every $t > 0$ define

$$Z_t = e^{-t}X + (1 - e^{-2t})^{1/2}W. \tag{3.7}$$

Then the random variable Z_t has a density $f(v,t)$ at each t, and it is well known that $f(t)$ is evolved from φ under the action of the adjoint Ornstein-Uhlenbeck semigroup. Therefore $f(v,t)$ satisfies Eq. (3.4), which can of course be checked directly from the definition.

For any $\alpha > 0$, we set $f_\alpha(v) = \alpha^{-n/2} f(\alpha^{-1/2} v)$. Then, $f(v,t)$ is expressed by the convolution formula

$$f(v,t) = \varphi_{\alpha(t)} \star M_{\beta(t)} \tag{3.8}$$

where $\alpha(t) = e^{-2t}$, $\beta(t) = 1 - e^{-2t}$.

Let us set

$$\nu(t) = \frac{\beta(t)}{\alpha(t)} = e^{2t} - 1. \tag{3.9}$$

Then, $f(v,t)$ can be expressed in the equivalent form

$$f(v,t) = \alpha(t)^{-n/2} \int_{\mathbf{R}^n} \varphi(w) M_{\nu(t)} \left(\frac{v}{\alpha^{1/2}(t)} - w \right) dw. \tag{3.10}$$

Using the formula (3.10) it is a straightforward matter to show that

$$H(f(t)) = H(\varphi \star M_{\nu(t)}) + nt \tag{3.11}$$

and

$$L(f(t)) = \alpha^{-1}(t) L(\varphi \star M_{\nu(t)}). \tag{3.12}$$

(3.11) and (3.12) establish that $H(f(t))$ and $L(f(t))$ exist and are finite for all $t > 0$.

Let us differentiate with respect to the time equation (3.11). Thanks to (2.4) we obtain

$$\begin{aligned} \frac{dH(f(t))}{dt} &= \frac{d}{dt} H(\varphi \star M_{\nu(t)}) + n \\ &= -\frac{1}{2} L(\varphi \star M_{\nu(t)}) \frac{d\nu(t)}{dt} + n = -L(f(t)) + n. \end{aligned} \tag{3.13}$$

Since $L(M) = n$, and $H(M)$ is constant, (3.13) is

$$\frac{dD(f(t))}{dt} = -J(f(t)). \tag{3.14}$$

The right-hand side of (3.14) can be bounded from above thanks to the logarithmic Sobolev inequality (2.7), obtaining

$$\frac{dD(f(t))}{dt} \leq -2D(f(t)). \tag{3.15}$$

We proved

THEOREM 3.1. Let $f(v,t)$ be the solution to the initial value problem for Eq. (3.4), where the initial value $\varphi(v)$ satisfies (3.5) and has finite entropy. Then $f(v,t)$ converges in relative entropy to $M(v)$, and

$$D(f(t)) \leq e^{-2t} D(\varphi). \tag{3.16}$$

REMARK 3.1. Reverting to the old variables, we obtain the rate of convergence to equilibrium for the solution to (3.1) with initial data of finite mass, temperature, and entropy. The rate of exponential convergence depends in this case on γ and θ, and it is equal to $2\gamma/\theta$. Convergence in relative entropy and formula (3.14) are proven also in Lemma 2.2 of [CS91]. There the proof leading to (3.14) is different, and the rate of convergence to equilibrium is not investigated.

REMARK 3.2. The application of the Csiszar-Kullback inequality (1.15) to (3.16) shows that the solution to (3.1) decays exponentially in the L^1-norm towards the Maxwellian at a rate γ/θ.

4. Time decay of the relative Fisher information. In this section we will investigate the time decay of the relative Fisher information. For $t > 0$, let us differentiate (3.12) with respect to time. We obtain

$$\frac{dL(f(t))}{dt} = 2\alpha^{-1}(t)L(\varphi \star M_{\nu(t)}) + 2\alpha^{-2}(t)\frac{dL(\varphi \star M_{\nu(t)})}{d\nu(t)}. \tag{4.1}$$

Hence, recalling (2.4), we need to evaluate the second derivative of $H(\varphi \star M_\nu)$ with respect to ν. In the one-dimensional case, the subsequent derivatives of $H(\varphi \star M_\nu)$ have first been considered by McKean [Mk66]. In particular, a detailed study of $dL(\varphi \star M_\nu)/d\nu$ is due to Gabetta [Ga95], in connection with the trend towards equilibrium of the solution to the Kac equation. Here, we extend the analysis of [Ga95] to higher dimensions.

The estimates of the following lemma are simple variants of heat semigroup estimates that can be found in [CS91].

LEMMA 4.1. Suppose $\varphi(v)$ is a probability density on $\mathbf{R}^n$, and let us set $g(v,t) = \varphi \star M_t$. Then, given $p > 1$, there exist finite constants c depending only on the dimension n, p, and t such that, for all $k \geq 1$

$$\sum_{i_1, i_2, \ldots, i_k} \left| \frac{\partial^k g(v,t)}{\partial v_{i_1}, \partial v_{i_2} \cdots \partial v_{i_k}} \right| \leq c(p,k,t) g^{1/p}(v,t). \tag{4.2}$$

Proof. For any $t > s > 0$, $g(v,t)$ is continuously differentiable in v, and

$$\frac{\partial g(v,t)}{\partial v_i} = -\int_{\mathbf{R}^n} \varphi(w) \frac{v_i - w_i}{t} M_t(v-w)\, dw. \tag{4.3}$$

Hence, by the Hölder inequality, for $p > 1$ and $1/p + 1/q = 1$,

$$\begin{aligned} \left| \frac{\partial g(v,t)}{\partial v_i} \right| &\leq \int_{\mathbf{R}^n} (\varphi(w) M_t(v-w))^{1/p} \\ &\quad \cdot \left[\varphi(w) \left(\frac{|v_i - w_i|}{t} \right)^q M_t(v-w) \right]^{1/q} dw \\ &\leq g^{1/p}(v,t) \left[\int_{\mathbf{R}^n} \varphi(w) \left(\frac{|v_i - w_i|}{t} \right)^q M_t(v-w)\, dw \right]^{1/q}. \end{aligned} \tag{4.4}$$

The result for $k = 1$ follows since $|v_i|^q M_t(v)$ is clearly bounded above by a constant depending only on t and $q = p/(p-1)$.

The inequalities for $k > 1$ follow in a similar fashion from the fact that, for any $t > 0$,

$$\frac{\partial^k M_t(v)}{\partial v_i^k} = (-1)^k (\sqrt{2t})^{-k} H_k\left(\frac{v_i}{\sqrt{2t}}\right) M_t(v),$$

where H_k is the mth Hermite polynomial.

REMARK 4.1. Let $f(v,t)$ be the solution to the initial value problem for (3.4). Then, by formula (3.8), the result of Lemma 4.1 can be easily extended to $f(v,t)$.

LEMMA 4.2. Suppose f is a probability density on $\mathbf{R}^n$ such that $L(f)$ is finite. Then, for all $\alpha \geq 0$

$$\frac{dL(f \star M_\alpha)}{d\alpha} = -Q(f \star M_\alpha) \tag{4.5}$$

where, for $\sqrt{f} \in H^2(\mathbf{R}^n)$, $Q(f)$ is the operator

$$Q(f) = \sum_{i,j=1}^{n} \int_{\mathbf{R}^n} \left[\frac{1}{f(v)} \frac{\partial^2 f(v)}{\partial v_i \partial v_j} - \frac{1}{f^2(v)} \frac{\partial f(v)}{\partial v_i} \frac{\partial f(v)}{\partial v_j} \right]^2 f(v)\, dv. \tag{4.6}$$

Proof. Let $\alpha > 0$. Differentiating $f \star M_\alpha$, we get the diffusion equation

$$\frac{\partial (f \star M_\alpha)}{\partial \alpha} = \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 (f \star M_\alpha)}{\partial v_i^2}. \tag{4.7}$$

Thus, exchanging the integral with the derivative, and denoting $g = f \star M_\alpha$ for the sake of simplicity, we obtain

$$\begin{aligned} \frac{dL(g)}{d\alpha} &= \sum_{j=1}^{n} \frac{d}{d\alpha} \int_{\mathbf{R}^n} \frac{1}{g} \left(\frac{\partial g}{\partial v_j} \right)^2 dv \\ &= \sum_{j=i}^{n} \int_{\mathbf{R}^n} \left[-\frac{1}{g^2} \frac{\partial g}{\partial \alpha} \left(\frac{\partial g}{\partial v_j} \right)^2 + \frac{2}{g} \frac{\partial g}{\partial v_j} \frac{\partial}{\partial \alpha} \left(\frac{\partial g}{\partial v_j} \right) \right] dv \\ &= \sum_{i,j=1}^{n} \int_{\mathbf{R}^n} \left[-\frac{1}{2g^2} \frac{\partial^2 g}{\partial v_i^2} \left(\frac{\partial g}{\partial v_j} \right)^2 + \frac{1}{g} \frac{\partial g}{\partial v_j} \frac{\partial^3 g}{\partial v_i^2 \partial v_j} \right] dv. \end{aligned} \tag{4.8}$$

The differentiability of $L(g)$ under the sign of integration will be justified by showing that, for every fixed i and j, the integrals on the right side of (4.8) converge uniformly for all values of $\alpha \geq \alpha_0 > 0$, and are continuous functions of the parameter α. The first property can be easily verified. For each i and j, let us set

$$A_i(g) = \frac{1}{g^{1/2}} \left| \frac{\partial^2 g}{\partial v_i^2} \right|, \qquad B_j(g) = \frac{1}{g^{3/2}} \left(\frac{\partial g}{\partial v_j} \right)^2. \tag{4.9}$$

Then, by the Cauchy-Schwarz inequality

$$\begin{aligned} \int_{\mathbf{R}^n} \frac{1}{2g^2} \left| \frac{\partial^2 g}{\partial v_i^2} \right| \left(\frac{\partial g}{\partial v_j} \right)^2 dv &= \int_{\mathbf{R}^n} A_i(g) B_j(g)\, dv \\ &\leq \left[\int_{\mathbf{R}^n} A_i^2(g)\, dv \right]^{1/2} \left[\int_{\mathbf{R}^n} B_j^2(g)\, dv \right]^{1/2} \\ &\leq \left[\int_{\mathbf{R}^n} A_i^2(M_\alpha)\, dv \right]^{1/2} \left[\int_{\mathbf{R}^n} B_j^2(M_\alpha)\, dv \right]^{1/2} = c_1(\alpha). \end{aligned} \tag{4.10}$$

The last inequality follows by Lemma 2.1 of Lions and Toscani [LT95]. The same lemma shows that

$$\lim_{\beta \to 0^+} \int_{\mathbf{R}^n} A_i^2(g \star M_\beta)\, dv = \int_{\mathbf{R}^n} A_i^2(g)\, dv \tag{4.11}$$

and

$$\lim_{\beta \to 0^+} \int_{\mathbf{R}^n} B_j^2(g \star M_\beta)\, dv = \int_{\mathbf{R}^n} B_j^2(g)\, dv. \tag{4.12}$$

Since $A_i(g \star M_\beta)$ and $B_j(g \star M_\beta)$ converge a.e. to $A_i(g)$ and $B_j(g)$ respectively, (4.11) and (4.12) imply that $A_i(g \star M_\beta)$ and $B_j(g \star M_\beta)$ converge to $A_i(g)$ and $B_j(g)$ in L^2. Consequently

$$\lim_{\beta \to 0^+} \int_{\mathbf{R}^n} |A_i(g \star M_\beta) B_j(g \star M_\beta) - A_i(g) B_j(g)| \, dv = 0. \tag{4.13}$$

The same procedure can be applied to the integral

$$\int_{\mathbf{R}^n} \frac{1}{g} \frac{\partial g}{\partial v_j} \frac{\partial^3 g}{\partial v_i^2 \partial v_j} \, dv.$$

This justifies formula (4.8).

Let us fix i and j. Integrating by parts with respect to v_i we obtain

$$\begin{aligned} &\int_{\mathbf{R}^n} -\frac{1}{2g^2} \frac{\partial^2 g}{\partial v_i^2} \left(\frac{\partial g}{\partial v_j} \right)^2 dv \\ &\quad = \int_{\mathbf{R}^n} \left[\frac{1}{g^2} \frac{\partial g}{\partial v_i} \frac{\partial g}{\partial v_j} \frac{\partial^2 g}{\partial v_i \partial v_j} - \frac{1}{g^3} \left(\frac{\partial g}{\partial v_i} \right)^2 \left(\frac{\partial g}{\partial v_j} \right)^2 \right] dv. \end{aligned} \tag{4.14}$$

In fact, by inequality (4.2),

$$\left| \frac{1}{2g^2} \frac{\partial g}{\partial v_i} \left(\frac{\partial g}{\partial v_j} \right)^2 \right| \le c(p, 1, \alpha)^3 g^{(3-2p)/p},$$

and, choosing $p < 3/2$,

$$\lim_{v_i \to \pm\infty} \frac{1}{g^2} \frac{\partial g}{\partial v_i} \left(\frac{\partial g}{\partial v_j} \right)^2 = 0. \tag{4.15}$$

Likewise, integrating by parts with respect to v_i,

$$\begin{aligned} &\int_{\mathbf{R}^n} \frac{1}{g} \frac{\partial^3 g}{\partial v_i^2 \partial v_j} \frac{\partial g}{\partial v_j} \, dv \\ &\quad = \int_{\mathbf{R}^n} \left[-\frac{1}{g} \left(\frac{\partial^2 g}{\partial v_i \partial v_j} \right)^2 + \frac{1}{g^2} \frac{\partial g}{\partial v_i} \frac{\partial g}{\partial v_j} \frac{\partial^2 g}{\partial v_i \partial v_j} \right] dv. \end{aligned} \tag{4.16}$$

Also in this case, by applying (4.2) with $p < 2$

$$\lim_{v_i \to \pm\infty} \frac{1}{g} \frac{\partial g}{\partial v_i} \frac{\partial^2 g}{\partial v_i \partial v_j} = 0. \tag{4.17}$$

The following lemma provides the analog of the logarithmic Sobolev inequality in the form (2.7), and furnishes an upper bound for the relative Fisher information $J(f)$ in terms of $Q(f)$ and $L(f)$.

LEMMA 4.3. Suppose f is a probability density on $\mathbf{R}^n$ that satisfies conditions (3.5), and such that $Q(f)$ is finite. Then

$$J(f) \le Q(f) - L(f). \tag{4.18}$$

In addition

$$Q(f) \ge Q(M) \tag{4.19}$$

and there is equality in (4.19) if and only if $f = M$.

Proof. Let $\delta_{i,j}$ denote as usual the Kronecker delta. Then

$$\begin{aligned} 0 &\leq \sum_{i,j=1}^{n} \int_{\mathbf{R}^n} \left[\frac{1}{f(v)} \frac{\partial^2 f(v)}{\partial v_i \partial v_j} - \frac{1}{f^2(v)} \frac{\partial f(v)}{\partial v_i} \frac{\partial f(v)}{\partial v_j} + \delta_{i,j} \right]^2 f(v)\, dv \\ &= Q(f) + n - 2L(f). \end{aligned} \tag{4.20}$$

Because $L(M) = n$, (4.20) implies (4.18). Moreover, $Q(M) = n$ and $L(f) \geq L(M)$ imply (4.19). Finally, the statement concerning equality in (4.19) follows by the convexity of L and (2.3).

We will now apply the above results to get a bound on the right-hand side of (4.1). Combining (4.5) with the identity $Q(f(t)) = \alpha(t)^{-2} Q\ (\varphi \star M_{\nu(t)})$, Eq. (4.5) takes the form

$$\frac{dL(f(t))}{dt} = -2[Q(f(t)) - L(f(t))]. \tag{4.21}$$

Next, by (4.18) we obtain

$$\frac{dL(f(t))}{dt} = \frac{dJ(f(t))}{dt} \leq -2J(f(t)). \tag{4.22}$$

We proved

THEOREM 4.1. Let $f(v,t)$ be the solution to the initial value problem for Eq. (3.4), where the initial value $\varphi(v)$ satisfies (3.5) and has finite Fisher information. Then $f(v,t)$ converges in relative Fisher information to $M(v)$, and

$$J(f(t)) \leq e^{-2t} J(\varphi). \tag{4.23}$$

REMARK 4.2. As in Sec. 3, reverting to the old variables, we obtain the rate of convergence to equilibrium for the solution to (3.1) with initial data of finite mass, temperature, and Fisher information. We mention that inequality (4.19) has been first derived by McKean [Mk66] in the one-dimensional case. As we shall see in Sec. 6, inequality (4.18) is the key for a new proof of the logarithmic Sobolev inequality.

5. Other functionals. In the previous section, we showed that the derivative of the Fisher information of the solution to (3.4) satisfies the differential equation

$$\frac{dL(f(t))}{dt} = -2[Q(f(t)) - L(f(t))]. \tag{5.1}$$

The same procedure could be applied to evaluate the time derivative of the operator $Q(f(t))$. Since $Q(f(t)) = \alpha(t)^{-2} Q\ (\varphi \star M_{\nu(t)})$,

$$\begin{aligned} \frac{dQ(f(t))}{dt} &= 4\alpha^{-2}(t) Q(\varphi \star M_{\nu(t)}) + 2\alpha^{-3}(t) \frac{dQ(\varphi \star M_{\nu(t)})}{d\nu(t)} \\ &= -2[Q_1(f(t)) - 2Q(f(t))]. \end{aligned} \tag{5.2}$$

In (5.2) the operator $Q_1(f \star M_\nu)$ equals the derivative of $Q(f \star M_\nu)$ with respect to ν. Because of the logarithmic Sobolev inequality, we proved that

$$-\frac{dH(f(t))}{dt} = L(f(t)) - L(M) \geq 2[H(f(t)) - H(M)]. \tag{5.3}$$

Moreover, thanks to Lemma 4.3,

$$-\frac{dL(f(t))}{dt} = 2[Q(f(t)) - L(f(t))] \geq 2[L(f(t)) - L(M)]. \tag{5.4}$$

This suggests that, if f satisfies (3.5),

$$Q_1(f) - 2Q(f) \geq Q(f) - Q(M), \tag{5.5}$$

and, if for $i \geq 1$ we define

$$Q_i(f \star M_\nu) = \frac{d^{i+1}L(f \star M_\nu)}{d\nu^{i+1}}, \tag{5.6}$$

it is conjectured that

$$Q_{i+1}(f) - (i+2)Q_i(f) \geq Q_i(f) - Q_i(M), \tag{5.7}$$

but we could not prove it.

6. A new proof of the logarithmic Sobolev inequality. A non-secondary consequence of the analysis of Sec. 4 is a new proof of inequality (2.8), and so a new proof of the logarithmic Sobolev inequality with respect to the Lebesgue measure, including the cases of equality. To this purpose, we need to add two further results. The first one is a simple variant of a lemma that can be found in McKean [Mk66], or, in a more general setting, in Lions and Toscani [LT95]. For this reason we skip the proof.

LEMMA 6.1. Suppose φ is a probability density on $\mathbf{R}^n$. Then, $H(\varphi \star M_s)$ and $L(\varphi \star M_s)$ are decreasing with respect to $s \geq 0$, and, whenever $H(\varphi)$ and $L(\varphi)$ are finite

$$\lim_{s \to 0^+} H(\varphi \star M_s) = H(\varphi), \qquad \lim_{s \to 0^+} L(\varphi \star M_s) = L(\varphi). \tag{6.1}$$

The next result is concerned with the cases of equality in (4.18). We have

LEMMA 6.2. Suppose φ is a density on $\mathbf{R}^n$ that satisfies conditions (3.5), and let $f(v,t)$ be the solution to the initial value problem for (3.4) with φ as initial datum. Then, unless $\varphi = M$,

$$J(f(t)) < Q(f(t)) - L(f(t)). \tag{6.2}$$

Proof. We remark that for $t > 0$, the solution to (3.4) can be bounded from below in terms of a time-dependent Maxwellian. Indeed (cf. Lemma 1.1 of [CS91])

$$f(v,t) \geq A_t M_{1-e^{-t}}, \tag{6.3}$$

where A_t is an explicitly computable function bounded away from zero for t finite.

The sum of the integrals in (4.20) is equal to zero if and only if, for all i, j, the integrand is equal to zero on every compact set. Since on $\{|v| \leq R\}$, $R > 0$, $f(v,t)$ satisfies (6.3), the aforementioned condition is equivalent to requiring that $f(v,t)$ be a solution to the partial differential equations

$$\frac{\partial^2 f}{\partial v_i \partial v_j} = \frac{1}{f}\frac{\partial f}{\partial v_i}\frac{\partial f}{\partial v_j} + \delta_{i,j} f, \qquad i,j = 1,\ldots,n \tag{6.4}$$

on the set $\{|v| \leq R\}$. Elementary computations show that $M(v)$ satisfies (6.4). Since the solution to (6.4) satisfying conditions (3.5) is unique, the result of the lemma follows.

We can now pass to proving inequality (2.8). Let φ be the initial value for (3.4). By (3.13), for any $t > 0$

$$-\frac{dH(f(t))}{dt} = L(f(t)) - L(M). \tag{6.5}$$

Adding the positive quantity $Q(f(t)) + n - 2L(f(t))$ on both sides of (6.5) we obtain

$$-\frac{dH(f(t))}{dt} + Q(f(t)) + n - 2L(f(t)) = Q(f(t)) - L(f(t)) = -\frac{1}{2}\frac{dL(f(t))}{dt}. \tag{6.6}$$

By Lemma 6.1, with a simple exchange of the roles of φ and M, we conclude that $\lim_{t\to+\infty} H(f(t)) = H(M)$, and $\lim_{t\to+\infty} L(f(t)) = L(M)$. Hence, we can integrate (6.6) on the time interval $[s, +\infty)$, $s > 0$, obtaining

$$H(f(s)) - H(M) + R_s(\varphi) = \tfrac{1}{2}[L(f(s)) - L(M)], \tag{6.7}$$

where

$$R_s(\varphi) = \int_s^{+\infty} [Q(f(t)) + n - 2L(f(t))]\, dt \tag{6.8}$$

is nonnegative by (4.20). By Lemma 6.1, $\lim_{s\to 0^+} H(f(s))$ and $\lim_{s\to 0^+} L(f(s))$ are always well defined, and infinity is admitted as a possible value. Thus, letting $s \to 0$ in (6.8) we obtain (2.8) whenever $L(\varphi)$ is finite. We proved

THEOREM 6.1. Suppose φ is a density on $\mathbf{R}^n$ that satisfies conditions (3.5), and such that $L(\varphi)$ is finite. Then, $H(\varphi)$ is also finite, and

$$H(\varphi) - H(M) + R(\varphi) = \tfrac{1}{2}[L(\varphi) - L(M)]. \tag{6.9}$$

The "remainder term" $R(\varphi)$, given by

$$R(\varphi) = \int_0^{+\infty} [Q(\varphi_{\alpha(t)} \star M_{\beta(t)}) + n - 2L(\varphi_{\alpha(t)} \star M_{\beta(t)})]\, dt \tag{6.10}$$

is positive, and is equal to zero if and only if $\varphi = M$.

REMARK 6.1. In a recent paper, Carlen [Ca91] showed that the strict form of the Gross inequality is a direct consequence of an inequality due to Stam [St59] and Blachman [Bl65], and that this in turn is a direct consequence of strict superadditivity of the Fisher information. The Stam-Blachman inequality says that for any pair of independent random variables X, Y with densities that satisfy (3.5), and for all positive constants a, b such that $a^2 + b^2 = 1$,

$$H(aX + bY) \le a^2 S(X) + b^2 S(Y), \tag{6.11}$$

with equality if and only if X and Y have Gaussian densities with

$$\mathrm{cov}(X) = \mathrm{cov}(Y). \tag{6.12}$$

In his paper, Carlen determines the cases of equality and the Gross logarithmic Sobolev inequality with a "remainder term".

The relation between entropy and Fisher information obtained through the dynamics of the Fokker-Planck equation was used before by Bakry and Emery [BE85] in their proof of the logarithmic Sobolev inequality. This proof makes no use of the functional $Q(f)$

and does not take advantage of the identification of the cases of equality that we obtained in Lemma 6.2.

REMARK 6.2. The analysis of Secs. 3 and 4 only requires that φ has finite second moment, possibly different from the second moment of M. In this last case, one has only to take care of substituting $M^{f(t)}$ to M in the various inequalities. Hence, inequality (2.6) can be easily justified by the same arguments.

The general case of a function with unbounded second moment follows easily. Indeed, if h has unbounded energy, for any $N > 0$ we consider the function $h_N(v)$ that is equal to $f(v)$ on the set $\{|v| \leq N\}$, and 0 elsewhere. Then, h_N has bounded energy and satisfies (2.8). Furthermore, $L(h_N) \leq L(h)$, and letting $N \to \infty$, (2.8) for h follows.

7. Some applications to the heat equation. In this last section, we investigate the asymptotic behaviour of the solution to the heat equation

$$\frac{\partial u}{\partial t} = \frac{1}{2}\Delta u \tag{7.1}$$

by means of the arguments we developed in the first part of the paper. The decay in relative entropy was used in [To96] to prove that the fundamental solution to (7.1) gives the asymptotic representation of the solution of the Cauchy problem for the same equation, with an explicit rate of decay in the L^1-norm. Here we will find the decay in relative Fisher information and its consequences on the rate of decay (in one dimension) in $C(\mathbf{R})$. We will consider initial values $\varphi(v)$ that are probability densities on $\mathbf{R}^n$. Moreover

$$\int_{\mathbf{R}^n} v\varphi(v)\,dv = 0, \qquad \int_{\mathbf{R}^n} |v|^2\varphi\,dv = nE < \infty. \tag{7.2}$$

The main result of [To96] is

THEOREM 7.1. Let φ be a probability density on $\mathbf{R}^n$ that satisfies (7.2) and such that $H(\varphi)$ is finite. Then the solution $u(t)$ to the Cauchy problem for (7.1) converges in relative entropy to M_{E+t}, and

$$D(u(t)) \leq D(\varphi)\frac{E}{E+t}. \tag{7.3}$$

The proof of Theorem 7.1 is based on the logarithmic Sobolev inequality (2.8). Indeed, the solution to (7.1) is represented by the convolution

$$u(v,t) = (\varphi \star M_t)(v), \tag{7.4}$$

and, by (2.4)

$$\frac{dH(u(t))}{dt} = -\frac{1}{2}L(u(t)). \tag{7.5}$$

Since $M_{E+t} = M_E \star M_t$,

$$\frac{dH(M_{E+t})}{dt} = -\frac{1}{2}L(M_{E+t}). \tag{7.6}$$

Hence, subtracting (7.6) from (7.5) we obtain

$$\frac{dD(u(t))}{dt} = -\frac{1}{2}J(u(t)). \tag{7.7}$$

Since the second moment of $u(t)$ is equal to $n(E+t)$, inequality (2.8) implies

$$D(u(t)) \leq \frac{E+t}{2}J(u(t)), \tag{7.8}$$

and (7.3) follows. In particular, by the Csiszar-Kullback inequality it follows that $u(t)$ decays towards the fundamental solution at a rate $\sqrt{E+t}$.

The argument leading to Theorem 7.1 can be applied to evaluate the decay of Fisher information. By (4.5) we have

$$\frac{dL(u(t))}{dt} = -Q(u(t)) \tag{7.9}$$

and

$$\frac{dL(M_{E+t})}{dt} = -Q(M_{E+t}). \tag{7.10}$$

Hence, subtracting (7.10) from (7.9) we obtain

$$\frac{dJ(u(t))}{dt} = -[Q(u(t)) - Q(M_{E+t})]. \tag{7.11}$$

It is easy to see that (4.20), written for $u(t)$ is

$$Q(u(t)) - \frac{2}{E+t}L(u(t)) + \frac{n}{(E+t)^2} \geq 0, \tag{7.12}$$

and this inequality implies

$$Q(u(t)) - Q(M_{E+t}) \geq \frac{2}{E+t}[L(u(t)) - L(M_{E+t})]. \tag{7.13}$$

Indeed, $L(M_{E+t}) = \frac{n}{E+t}$, $Q(M_{E+t}) = \frac{n}{(E+t)^2}$. Finally, inserting (7.13) into (7.11) we establish that

$$\frac{dJ(u(t))}{dt} \leq -\frac{2}{E+t}J(u(t)). \tag{7.14}$$

Therefore we proved

THEOREM 7.2. Let φ be a probability density on $\mathbf{R}^n$ that satisfies (7.2) and such that $L(\varphi)$ is finite. Then the solution $u(t)$ to the Cauchy problem for (7.1) converges in relative Fisher information to M_{E+t} and

$$J(u(t)) \leq J(\varphi)\frac{E^2}{(E+t)^2}. \tag{7.15}$$

REMARK 7.1. Unlike the behaviour of entropy and Fisher information in the Fokker-Planck equation, the relative Fisher information of the solution to the heat equation decays at a rate $(E+t)^2$, namely at the square of the rate of decay of the relative entropy. This result has interesting consequences, at least in one space dimension, in that it permits us to find the optimal rate of decay of the solution to (7.1) towards the fundamental solution in $C(\mathbf{R})$.

To this end, we need to improve a recent result by Gabetta and Toscani [GT94], which is in one sense the corresponding result to the Csiszar-Kullback inequality for Fisher information.

LEMMA 7.1. Let φ be a probability density on $\mathbf{R}$ that satisfies

$$\int_{\mathbf{R}} v\varphi(v)\,dv = 0, \quad \int_{\mathbf{R}} v^2\varphi(v)\,dv = \sigma, \quad L(\varphi) < \infty. \tag{7.16}$$

Then

$$\|\varphi - M_\sigma\|_\infty \le C(\varphi)[L(\varphi) - L(M_\sigma)]^{1/2}, \tag{7.17}$$

where

$$C(\varphi) = 1 + \frac{\sqrt{5}}{2}\{1 + \sigma^{1/4}[L(\varphi) - L(M_\sigma)]^{1/4}\} + \frac{\sqrt{5}}{4\sqrt{\pi}}\sigma^{-1/4}\int_{\mathbf{R}} \sqrt{\varphi(v)}\,dv. \tag{7.18}$$

Proof. Given $s > 0$, we shall prove inequality (7.17) first for the function $\varphi \star M_s$. The result for general f satisfying (7.16) will follow by Lemma 6.1 letting $s \to 0$. Moreover, to avoid inessential heavy notation, we will set $f = \varphi \star M_s$ and $\nu = \sigma + s$. Then $f \in C^1(\mathbf{R})$, and it satisfies the equation

$$\frac{d}{dx}\{e^{\frac{v^2}{2\nu}}[f(v) - M_\nu(v)]\} = e^{\frac{v^2}{2\nu}}\left[f'(v) + \frac{v}{\nu}f(v)\right]. \tag{7.19}$$

Integration of (7.19) on $[0, v]$, $v > 0$, leads to the inequality

$$\begin{aligned} |f(v) - M_\nu(v)| &\le |f(0) - M_\nu(0)|e^{-\frac{v^2}{2\nu}} \\ &\quad + \int_0^v \left|f'(w) + \frac{w}{\nu}f(w)\right| e^{-\frac{v^2}{2\nu} + \frac{w^2}{2\nu}}\,dw. \end{aligned} \tag{7.20}$$

By application of the Cauchy-Schwarz inequality and (2.2) we get

$$\begin{aligned} \int_0^v \left|f'(w) + \frac{w}{\nu}f(w)\right| &\le \left\{\int_0^v \left[\frac{f'(w)}{f(w)} + \frac{w}{\nu}\right]^2 f(w)\,dw\right\}^{1/2} \left\{\int_0^v f(w)\,dw\right\}^{1/2} \\ &\le [L(f) - L(M_\nu)]^{1/2}. \end{aligned} \tag{7.21}$$

Hence, (7.20) implies that

$$|f(v) - M_\nu(v)| \le |f(0) - M_\nu(0)|e^{-\frac{v^2}{2\nu}} + [L(f) - L(M_\nu)]^{1/2}. \tag{7.22}$$

Since f and M_ν are probability densities,

$$\begin{aligned} |f(0) - M_\nu(0)| &= \left|\int_{\mathbf{R}} [f(0)M_\nu(v) - M_\nu(0)f(v)]\,dv\right| \\ &\le M_\nu(0)\int_{\mathbf{R}} |f(v) - f(0)e^{-\frac{v^2}{2\nu}}|\,dv. \end{aligned} \tag{7.23}$$

Let us integrate the identity

$$\frac{d}{dx}\{e^{\frac{v^2}{4\nu}}[\sqrt{f(v)} - \sqrt{M_\nu(v)}]\} = e^{\frac{v^2}{4\nu}}\left[\frac{f'(v)}{2\sqrt{f(v)}} + \frac{v}{2\nu}\sqrt{f(v)}\right] \tag{7.24}$$

on the interval $[0, v]$, $v > 0$. We obtain

$$\sqrt{f(v)} - \sqrt{f(0)}e^{-\frac{v^2}{4\nu}} = \int_0^v \left[\frac{f'(w)}{2\sqrt{f(w)}} + \frac{w}{2\nu}\sqrt{f(w)}\right] e^{-\frac{v^2}{4\nu}+\frac{w^2}{4\nu}}\, dw. \tag{7.25}$$

Hence

$$\begin{aligned} |f(v) - f(0)e^{-\frac{v^2}{2\nu}}| &\le \tfrac{1}{2}[\sqrt{f(v)} + \sqrt{f(0)}e^{-\frac{v^2}{4\nu}}] \\ &\cdot \int_0^v \left|\frac{f'(w)}{\sqrt{f(w)}} + \frac{w}{\nu}\sqrt{f(w)}\right| e^{-\frac{v^2}{4\nu}+\frac{w^2}{4\nu}}\, dw. \end{aligned} \tag{7.26}$$

By the Cauchy-Schwarz inequality and (2.2) we get

$$\begin{aligned} &\int_0^v \left|\frac{f'(w)}{\sqrt{f(w)}} + \frac{w}{\nu}\sqrt{f(w)}\right| e^{-\frac{v^2}{4\nu}+\frac{w^2}{4\nu}}\, dw \\ &\le [L(f) - L(M_\nu)]^{1/2} \left\{\int_0^v e^{-\frac{v^2}{2\nu}+\frac{w^2}{2\nu}}\, dw\right\}^{1/2}. \end{aligned} \tag{7.27}$$

Easy computations reveal that for all $v > 0$

$$\int_0^v e^{-\frac{v^2}{2\nu}+\frac{w^2}{2\nu}}\, dw \le \frac{5}{2}\sqrt{\nu}. \tag{7.28}$$

Since the integrand is less than or equal to one, if $v \le 2\sqrt{\nu}$, then (7.27) holds. Thus, suppose $v > 2\sqrt{\nu}$, and at the same time $w \ge 2\sqrt{\nu}$. Since $v \ge w$ we have

$$\frac{v^2}{2\nu} - \frac{w^2}{2\nu} = \frac{1}{2\nu}(v-w)(v+w) \ge \frac{2}{\sqrt{\nu}}(v-w)$$

and

$$\int_{2\sqrt{\nu}}^v e^{-\frac{v^2}{2\nu}+\frac{w^2}{2\nu}}\, dw \le \int_{2\sqrt{\nu}}^v e^{-\frac{2}{\sqrt{\nu}}(v-w)}\, dw \le \frac{\sqrt{\nu}}{2}.$$

So (7.26) implies the inequality

$$\begin{aligned} |f(v) - f(0)e^{-\frac{v^2}{2\nu}}| &\le \frac{\sqrt{5}}{4\sqrt{\pi}\sqrt[4]{\nu}}[\sqrt{f(v)} + \sqrt{f(0)}e^{-\frac{v^2}{4\nu}}] \\ &\cdot [L(f) - L(M_\nu)]^{1/2}. \end{aligned} \tag{7.29}$$

To conclude the proof of the lemma, we have only to observe that

$$\begin{aligned} |f(v)| &= \left|\int_{-\infty}^v f'(w)\, dw\right| \\ &\le \int_{-\infty}^v \left|\frac{f'(w)}{\sqrt{f(w)}}\right| \sqrt{f(w)}\, dw \le L(f)^{1/2}. \end{aligned} \tag{7.30}$$

Hence,

$$\sqrt{f(0)} \le L(f)^{1/4} \le [L(f) - L(M_\nu)]^{1/4} + L(M_\nu)^{1/4} \tag{7.31}$$

and

$$\int_{\mathbf{R}} \sqrt{f(0)}e^{-\frac{v^2}{4\nu}}\, dv \le 2\sqrt{\pi\nu}\{\nu^{-1/4} + [L(f) - L(M_\nu)]^{1/4}\}. \tag{7.32}$$

LEMMA 7.2. Let φ be a probability density on $\mathbf{R}$ that satisfies

$$\int_{\mathbf{R}} v\varphi(v)\,dv = 0, \quad \int_{\mathbf{R}} v^2\varphi(v)\,dv = \sigma, \quad H(\varphi) < \infty. \tag{7.33}$$

Then

$$\int_{\mathbf{R}} |\varphi(v) - M_\sigma(v)|^{1/2}\,dv \le 2^{3/4}\sigma^{1/2}D(\varphi)^{1/4} + 2^{3/2}. \tag{7.34}$$

Proof. By the Cauchy-Schwarz inequality,

$$\int_{|v|\le\sigma} |\varphi(v) - M_\sigma(v)|^{1/2}\,dv \le (2\sigma)^{1/2}\left[\int_{|v|\le\sigma} |\varphi(v) - M_\sigma(v)|\,dv\right]^{1/2}, \tag{7.35}$$

so that, by (1.15)

$$\int_{|v|\le\sigma} |\varphi(v) - M_\sigma(v)|^{1/2}\,dv \le 2^{3/4}\sigma^{1/2}D(\varphi)^{1/4}. \tag{7.36}$$

To conclude the proof, consider that

$$\begin{aligned}
&\int_{|v|>\sigma} |\varphi(v) - M_\sigma(v)|^{1/2}\,dv \\
&\quad\le \left[\int_{|v|>\sigma} v^{-3/2}\,dv\right]^{1/2}\left[\int_{|v|>\sigma} v^{3/2}|\varphi(v) - M_\sigma(v)|\,dv\right]^{1/2} \\
&\quad\le 2\sigma^{-1/4}\left[\sigma^{-1/2}\int_{|v|>\sigma} v^2|\varphi(v) - M_\sigma(v)|\,dv\right]^{1/2} \le 2^{3/2}.
\end{aligned} \tag{7.37}$$

Finally we have

THEOREM 7.3. Let φ be a probability density on $\mathbf{R}$ that satisfies (7.2) and such that $L(\varphi)$ is finite. Then the solution $u(t)$ to the Cauchy problem for (7.1) converges to M_{E+t} in $C(\mathbf{R})$, and there exists a constant $C(E, L(\varphi))$ such that, for all $t \ge 0$

$$\|u(t) - M_{E+t}\|_\infty \le C(E, L(\varphi))\frac{E}{E+t}. \tag{7.38}$$

Proof. Since $L(\varphi)$ is finite, by the logarithmic Sobolev inequality (2.7), $H(\varphi)$ is finite too, and if $u(t)$ is the solution to the initial value problem for (7.1), for all $t \ge 0$

$$D(u(t)) \le \frac{E+t}{2}J(u(t)) \le J(\varphi)\frac{E^2}{2(E+t)}. \tag{7.39}$$

In (7.39) we used (7.15). By Lemma 7.2,

$$\begin{aligned}
\int_{\mathbf{R}} \sqrt{u(v,t)}\,dv &\le \int_{\mathbf{R}} |u(v,t) - M_{E+t}(v)|^{1/2}\,dv + \int_{\mathbf{R}} \sqrt{M_{E+t}(v)}\,dv \\
&\le 2^{3/4}E^{1/2}(E+t)^{1/4}J(\varphi)^{1/4} + 2^{3/2} + 2^{3/4}\sqrt{\pi}(E+t)^{1/4}.
\end{aligned} \tag{7.40}$$

Le us consider inequality (7.18) for $u(t)$. If we bound the last integral by inequality (7.39), the result follows. Note that the constant $C(E, L(\varphi))$ is computable by (7.18) and (7.34).

8. Concluding remarks. In this paper we investigated the asymptotic behaviour of the solution of the Fokker-Planck equation by means of the study of the decay, both of the relative entropy and the relative Fisher information. Thus, in most theorems it is required that the initial density has finite entropy or finite Fisher information. This is clearly unnecessary, since, in consequence of Lemma 6.1, at any time $t \geq s > 0$, both entropy and Fisher information of the solution are finite. In particular, an upper bound of these two quantities in terms of t can be obtained considering that for all $s > 0$ [Mk66], [LT95]

$$H(\varphi \star M_s) \leq H(M_s), \qquad L(\varphi \star M_s) \leq L(M_s). \tag{8.1}$$

Second, as far as the heat equation is concerned, Theorem 7.3 shows that the solution converges towards the fundamental solution in $C(\mathbf{R})$ at a rate $(E+t)^{-1}$. Since the solution itself converges to zero in $C(\mathbf{R})$ at a rate $t^{-1/2}$, we proved that taking the C-norm of the difference between $u(t)$ and M_{E+t} we gain a decay of $t^{-1/2}$. The same gain of decay follows by Theorem 7.1 for the L^1-norm of the difference. Indeed, the L^1-norm of the solution is conserved in time, whereas $u(t)$ converges in $L^1(\mathbf{R})$ towards M_{E+t} at a rate $t^{-1/2}$. Since Theorem 7.1 holds in any dimension of the space, one could conjecture that the same gain of decay is achieved in more than one dimension.

Unfortunately, the discussion of the decay of the n-dimensional heat equation encounters several obstacles. A first step in proving this conjecture could be the derivation of a bound on the L^2-norm of the difference between $u(t)$ and M_{E+t} in terms of the "correct" power of $L(u(t)) - L(M_{E+t})$, namely the corresponding bound of Lemma 7.1 in $L^2(\mathbf{R}^2)$. Indeed, in this case Sobolev imbeddings suggest that one could have the control. On the other hand, the optimal decay of the solution to the heat equation in L^2 can be obtained by the Nash inequality [Na58].

Acknowledgments. This work has been written within the activities of the National Group of Mathematical Physics (GNFM) of the National council for Researches (CNR), Project "Applications of Mathematics for Technology and Society". Partial support of the Institute of Numerical Analysis of the CNR is kindly acknowledged.

References

[Ar88] L. Arkeryd, *Stability in L^1 for the spatially homogeneous Boltzmann equation*, Arch. Rat. Mech. Anal. **103**, 151–167 (1988)

[BE85] D. Bakry and M. Emery, *Diffusions hypercontractives*, Seminaire de Probabilités XIX, Lecture Notes in Math., Vol. 1123, Springer-Verlag, Berlin, 1985, pp. 179–206

[Bl65] N. M. Blachman, *The convolution inequality for entropy powers*, IEEE Trans. Info. Theory **2**, 267–271 (1965)

[BGK54] P. L. Bhatnagar, 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)

[BT92] A. V. Bobylev and G. Toscani, *On the generalization of the Boltzmann H-theorem for a spatially homogeneous Maxwell gas*, J. Math. Phys. **33**, 2578–2586 (1992)

[Ca91] E. A. Carlen, *Superadditivity of Fisher's information and logarithmic Sobolev inequalities*, J. Funct. Anal. **101**, 194–211 (1991)

[CC92] E. A. Carlen and M. C. Carvalho, *Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation*, J. Stat. Phys. **67**, 575–608 (1992)

[CC94] E. A. Carlen and M. C. Carvalho, *Entropy production estimates for Boltzmann equations with physically realistic collision kernels*, J. Stat. Phys. **74**, 743–782 (1994)

[CGT96] E. A. Carlen, E. Gabetta, and G. Toscani, *Propagation of smoothness and the rate of exponential convergence to equilibrium for a spatially homogeneous Maxwellian gas*, preprint (1996)

[CS91] E. A. Carlen and A. Soffer, *Entropy production by block variable summation and central limit theorem*, Commun. Math. Phys. **140**, 339–371 (1991)

[Ce88] C. Cercignani, *Theory and applications of the Boltzmann equation*, Springer-Verlag, New York, 1988

[Ch43] S. Chandrasekhar, *Stochastic problems in physics and astronomy*, Rev. Modern Phys. **15**, 1–110 (1943)

[ChC58] S. Chapman and T. G. Cowling, *The Mathematical Theory of Non-Uniform Gases*, Cambridge University Press, Cambridge, England, 1958

[Cs62] I. Csiszar, *Informationstheoretische Konvergenzbegriffe im Raum der Wahrscheinlichkeitsverteilungen*, Publications of the Mathematical Institute, Hungarian Academy of Sciences, VII, Series A, 137–158 (1962)

[De89] L. Desvillettes, *Entropy dissipation rate and convergence in kinetic equations*, Commun. Math. Phys. **123**, 687–702 (1989)

[Fi25] R. A. Fisher, *Theory of statistical estimation*, Proc. Cambridge Philos. Soc. **22**, 700–725 (1925)

[FHL60] H. L. Frisch, E. Helfand, and J. L. Lebowitz, *Nonequilibrium distribution functions in a fluid*, Phys. of Fluids **3**, 325–338 (1960)

[Ga95] E. Gabetta, *On a conjecture of McKean with application to Kac's model*, Special issue devoted to the Proceedings of the 13th International Conference on Transport Theory (Riccione, 1993), Transport Theory Statist. Phys. **24**, 305–317 (1995)

[GT94] E. Gabetta and G. Toscani, *On convergence to equilibrium for Kac's caricature of a Maxwell gas*, J. Math. Physics **35**, 190–208 (1994)

[GTW95] E. Gabetta, G. Toscani, and B. Wennberg, *Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation*, J. Statist. Phys. **81**, 901–934 (1995)

[Gr75] L. Gross, *Logarithmic Sobolev inequalities*, Amer. J. Math. **97**, 1061–1083 (1975)

[Ku67] S. Kullback, *A lower bound for discrimination information in terms of variation*, IEEE Trans. Information Theory **4**, 126–127 (1967)

[LT95] P. L. Lions and G. Toscani, *A strengthened central limit theorem for smooth densities*, J. Funct. Anal. **128**, 148–167 (1995)

[Mk66] H. P. McKean, Jr., *Speed of approach to equilibrium for Kac's caricature of a Maxwellian gas*, Arch. Rational Mech. Anal. **21**, 343–367 (1966)

[Na58] J. Nash, *Continuity of solutions of parabolic and elliptic equations*, Amer. J. Math. **80**, 931–954 (1958)

[Ri84] H. Risken, *The Fokker-Planck Equation, Methods of Solution and Applications*, Springer-Verlag, Berlin, 1984

[St59] A. J. Stam, *Some inequalities satisfied by the quantities of information of Fisher and Shannon*, Information and Control **2**, 101–112 (1959)

[To96] G. Toscani, *Kinetic approach to the asymptotic behaviour of the solution to diffusion equations*, Rendic. di Matem., Serie VII **16**, 329–346 (1996)

[TM80] C. Truesdell and R. G. Muncaster, *Fundamentals of Maxwell's kinetic theory of a simple monoatomic gas*, Academic Press, New York, 1980

[We93] B. Wennberg, *Stability and exponential convergence for the Boltzmann equation*, Thesis, Chalmers University of Technology, 1993