ENTROPY PRODUCTION
AND CONVERGENCE TO EQUILIBRIUM

C. VILLANI, ENS LYON
cvillani@umpa.ens-lyon.fr

This set of notes was used to complement my short course on the convergence to equilibrium for the Boltzmann equation, given at Institut Henri Poincaré in November-December 2001, as part of the *Hydrodynamic limits* program organized by Stefano Olla and François Golse. The informal style is in accordance with the fact that this is neither a reference book, nor a research paper. The reader can use my review paper, *A review of mathematical topics in collisional kinetic theory*, as a reference source to dissipate any ambiguity with respect to notation for instance. Apart from minor corrections here and there, the main changes with respect to the original version of the notes were the addition of a final section to present some more recent developments and open directions, and the change of the sign convention for the entropy, to agree with physical tradition. Irene Mazzella is warmly thanked for kindly typesetting a preliminary version of this manuscript.

CONTENTS

1. The entropy production problem for the Boltzmann equation.

I shall start with Boltzmann's brilliant discovery that the H functional (or negative of the entropy) associated with a dilute gas is nonincreasing with time. To explain the meaning of this statement, let me first recall the model used by Boltzmann.

1.1. The Boltzmann equation: notation and preliminaries

<u>Unknown</u>: $f(t, x, v) = f_t(x, v) \geq 0$ is a time-dependent probability distribution on the phase space $\Omega_x \times \mathbb{R}_v^N$, where $\Omega_x \subset \mathbb{R}^N$ ($N = 2$ or 3) is the spatial domain where particles evolve and $\mathbb{R}_v^N$ is the space of velocities (to be thought of as a tangent space).

<u>Evolution equation</u>:

$$
\text{(BE)} \quad
\begin{cases}
\dfrac{\partial f}{\partial t} + v.\nabla_x f \;=\; Q(f, f) \\[2mm]
\qquad\qquad\quad := \displaystyle\int_{S^{N-1}} \int_{\mathbb{R}^N} (f' f'_* - f f_*) B(v - v_*, \sigma)\, dv_*\, d\sigma
\end{cases}
$$

+ boundary conditions.

<u>Notation</u>:

- $f = f(t, x, v)$, $f' = f(t, x, v')$, $f_* = f(t, x, v_*)$, $f'_* = f(t, x, v'_*)$,

- $v' = \dfrac{v + v_*}{2} + \dfrac{|v - v_*)}{2}\sigma, \quad v'_* = \dfrac{v + v_*}{2} - \dfrac{|v - v_*|}{2}\sigma \qquad (\sigma \in S^{N-1})$

  Think of (v', v'_*) as possible pre-collisional velocities in a process of elastic collision between two particles, leading to post-collisional velocities $(v, v_*) \in \mathbb{R}^N \times \mathbb{R}^N$.

<u>Physical quantity</u>: $B = B(v - v_*, \sigma) \geq 0$, the collision kernel (= cross-section times relative velocity) keeps track of the microscopic interaction. It is assumed to depend only on $|v - v_*|$ and $\cos\theta$, where

$$
\cos\theta = \left\langle \frac{v - v_*}{|v - v_*|}, \sigma \right\rangle.
$$

(Brackets stand for scalar product.) By abuse of notation I may sometimes write $B(v - v_*, \sigma) = B(|v - v_*|, \cos\theta)$.

The picture of collisions is as follows (in $\mathbb{R}^N_v$):
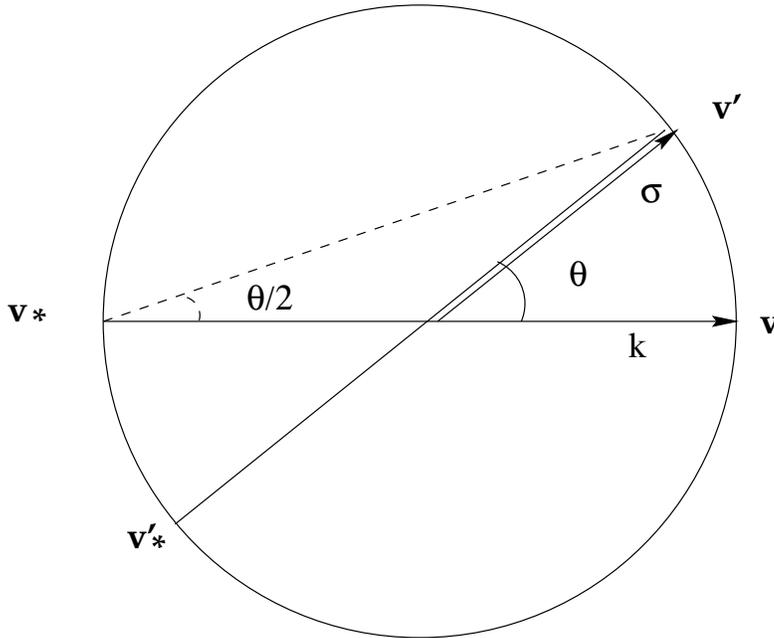
Boundary conditions: I shall consider three simple cases:

1) periodic condition: $\Omega = \mathbb{T}^N$ (not really a subset of $\mathbb{R}^N$!), no boundaries.

2) bounce-back condition: Ω smooth bounded,

$$f_t(x, v) = f_t(x, -v) \quad \text{for} \quad x \in \partial\Omega$$

3) specular reflection: Ω smooth bounded,

$$f_t(x, R_x v) = f_t(x, v) \quad \text{for} \quad x \in \partial\Omega$$

$$\text{where} \begin{cases} R_x v = v - 2\langle n(x), v\rangle n(x) \\ n(x) = \text{normal to } \partial\Omega \quad \text{at } x \end{cases}$$

3

<u>Local hydrodynamic fields</u>:

The following definitions constitute the bridge between the kinetic theory of Maxwell and Boltzmann on one hand, and the classical hydrodynamics on the other hand. Whenever $f(x, v)$ is a kinetic distribution, define the

- local density $\rho(x) = \displaystyle\int_{\mathbb{R}_v^N} f(x, v) \, dv$

- local velocity (mean) $u(x) = \dfrac{\displaystyle\int f(x, v) \, v \, dv}{\rho(x)}$

- local temperature $T(x) = \dfrac{\displaystyle\int f(x, v)|v - u(x)|^2 dv}{N\rho(x)}$

A simple symmetry argument shows that $\displaystyle\int_{\mathbb{R}^N} Q(f, f)\varphi \, dv = 0$ for $\varphi = \varphi(v)$ in Vect $(1, v_i, |v|^2)_{1 \leq i \leq N}$, as soon as $f = f(v)$ is integrable enough at large velocities. Those φ's are called **collision invariants**.

<u>Global conservation laws</u>:

Let $(f_t)_{t \geq 0}$ be a well-behaved solution of the BE. Then

$$\begin{cases} \dfrac{d}{dt} \displaystyle\int f_t(x, v) \, dv \, dx = 0 \quad \text{(conservation of mass)} \\[4mm] \dfrac{d}{dt} \displaystyle\int f_t(x, v)\dfrac{|v|^2}{2} \, dv \, dx = 0 \quad \text{(conservation of kinetic energy)} \end{cases}$$

- Also $\dfrac{d}{dt} \displaystyle\int f_t(x, v)v \, dv \, dx = 0$ in the case of periodic boundary conditions (conservation of momentum).

- When Ω has an axis of symmetry $\vec{k}$ and specular reflection is enforced, then there is an additional conservation law:

$$\dfrac{d}{dt} \int f_t(x, v) \, v_0 \, (k \wedge n) \, dv \, dx = 0 \quad \text{(conservation of angular momentum)}$$

($|k| = 1$, and $n = n(x)$ is still the normal).

4

<u>Normalizations</u>: Without loss of generality I shall assume

- $\displaystyle\int f_t(x,v)dvdx = 1 \qquad \int f_t(x,v)\frac{|v|^2}{2}\,dv\,dx = \frac{N}{2}$

- and $\displaystyle\int f_t(x,v)\,v\,dv\,dx = 0$ in the periodic case

- $|\Omega| = 1$ ($|\Omega| = N$-dimensional Lebesgue measure of Ω)

Moreover, in this course I will <u>not</u> consider the case when Ω has an axis of symmetry and specular boundary condition is imposed. A discussion would have to take into account angular momentum, and consider separately the particular case when Ω is a ball.

1.2. H functional and H Theorem

Let us now introduce <u>Boltzmann's H functional</u>: when f is a probability distribution on $\Omega \times \mathbb{R}^N$, define

$$H(f) = \int f \log f \; dv\,dx.$$

This quantity is well-defined in $\mathbb{R} \cup \{+\infty\}$ provided that $\displaystyle\int f(x,v)|v|^2\,dv\,dx$ is finite, and will be identified with the negative of the entropy associated with f.

The following theorem, essentially due to Boltzmann, will be our starting point.

<u>Theorem 1</u>: *Let $(f_t)_{t\geq 0}$ be a well-behaved (smooth) solution of the BE (in particular with finite entropy), with one of the boundary conditions discussed above. Then*

i) $\dfrac{d}{dt}H(f_t) \leq 0$. *Moreover, one can define a functional D on $L^1(\mathbb{R}_v^N)$, called "entropy production functional", or "dissipation of H functional", such that*

$$\frac{d}{dt}H(f_t) = -\int_{\Omega_x} D(f_t(x,.))\,dx.$$

ii) Assume that the collision kernel $B(v - v_*, \sigma)$ is > 0 for almost all $(v, v_*, \sigma) \in \mathbb{R}^{2N} \times S^{N-1}$. Let $f(x, v)$ be a probability distribution distribution on $\Omega \times \mathbb{R}^N$, with $\int f(x, v)|v|^2 \, dv \, dx < +\infty$. Then

$$\int_\Omega D(f(x, .)) \, dx = 0 \iff f \text{ is in local equilibrium, i.e. there exist}$$

functions $\rho(x) \geq 0$, $u(x) \in \mathbb{R}^N$, $T(x) > 0$,

such that $f(x, v) = \rho(x) \dfrac{e^{-\frac{|v - u(x)|^2}{2T(x)}}}{[2\pi T(x)]^{N/2}}$.

iii) Assume that the boundary condition is either periodic, or bounce-back, or specular, and in the latter case assume that the dimension is either 2 or 3 and that Ω has no axis of symmetry (is not a disk or a cylinder or an annulus or a ball or a shell). Without loss of generality, assume that f satisfies the normalizations discussed above. Then

$$(f_t)_{t \geq 0} \text{ is stationary} \iff \forall t \geq 0 \int_\Omega D(f_t(x, .)) dx = 0$$

$$\iff f_t(x, v) = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{N/2}} \equiv M(v) \quad \forall t \geq 0$$

The proof of this theorem is well-known (actually there are several proofs for point (ii), even though not so many), but is is useful to sketch it in order to help understanding refinements to come.

<u>Proof of theorem 1</u> (sketch):

i)

$$\frac{d}{dt} \int f_t \log f_t = \int Q(f_t, f_t)(\log f_t + 1) - \int (v.\nabla_x f_t)(\log f_t + 1)$$

$$= \int Q(f_t, f_t) \log f_t - \int \nabla_x \cdot (v f_t \log f_x)$$

$$= \int_{\Omega \times \mathbb{R}^N} Q(f_t, f_t) \log f_t - \int_{\partial\Omega \times \mathbb{R}^N} [v \cdot n(x)] f_t \, \log f_t$$

6

Under any one of the boundary conditions that we use, the second integral is 0. As for the first one, it can be rewritten as

$$\int_\Omega \int_{\mathbb{R}^{2N}} \int_{S^{N-1}} (f'f'_* - ff_*) \log f \ B(v - v_*, \sigma) \ dv \, dv_* \, d\sigma \, dx$$

By a simple symmetry trick, this is also

$$-\frac{1}{4} \int_\Omega \int_{\mathbb{R}^{2N}} \int_{S^{N-1}} (f'f'_* - ff_*) \log \frac{f'f'_*}{ff_*} \ B \, d\sigma \, dv_* \, dv \, dx$$

which takes the form $-\int_\Omega D(f) \, dx$ if one defines the **entropy production functional**:

$$D(f) = \frac{1}{4} \int_{\mathbb{R}^{2N} \times S^{N-1}} (f'f'_* - ff_*) \ \log \frac{f'f'_*}{ff_*} B(v - v_*, \sigma) \, d\sigma \, dv_* \, dv.$$

Clearly $D(f) \geq 0$ because $B \geq 0$ and $(X - Y) \log \frac{X}{Y} \geq 0$ as a consequence of log being increasing.

ii) Since $B > 0$ almost everywhere, the equality means that for (almost) all $x \in \Omega$ the L^1 function $f = f(x, \cdot)$ satisfies the functional equation of Maxwell–Boltzmann:

(MB) $\qquad f(v')f(v'_*) = f(v)f(v_*) \quad$ for almost all v, v_*, σ

(and also $\int f(v)(1 + |v|^2) \, dv < +\infty$, up to deletion of a negligible set of x's).

Integrate equation (MB) with respect to $\sigma \in S^{N-1}$, to find that

$$\begin{aligned}
f(v)f(v_*) &= \frac{1}{|S^{N-1}|} \int_{S^{N-1}} f(v')f(v'_*) \, d\sigma \\
&= \frac{1}{|S(v, v_*)|} \int_{S(v,v_*)} f(\alpha)f(\tilde{\alpha}) \, d\alpha
\end{aligned}$$

where

- $S(v, v_*)$ is the collision sphere, centered at $\dfrac{v + v_*}{2}$, with radius $\dfrac{|v - v_*|}{2}$

- $\widetilde{\alpha}$ is the symmetric of α with respect to $\dfrac{v + v_*}{2}$.

The important point about this average over $S(v, v_*)$ is that it only depends upon $S(v, v_*)$, whence only upon $\dfrac{v + v_*}{2}$ and $\dfrac{|v - v_*|}{2}$, or (which is equivalent) upon the physically meaningful variables

$$
\begin{cases}
m = & v + v_* \quad \text{(total momentum)} \\[2em]
e = & \dfrac{|v|^2 + |v_*|^2}{2} \quad \text{(total kinetic energy)}
\end{cases}
$$

Thus $f(v)f(v_*) = G(m, e)$.

Note: In this argument, due to Boltzmann, the Maxwell distribution arises from this <u>conflict of symmetries</u> between the tensor product structure of ff_* and the dependence of g upon a reduced set of variables: m and e.

Let us continue with the proof of (ii). We first assume f to be smooth (C^1, positive). Taking logarithms, we find

$$
\begin{aligned}
\log \; f(v) + \log \; f(v_*) \;\; &= \;\; \log \; G(m, e) \\
\frac{\partial}{\partial v} \implies \quad \nabla \log \; f(v) + 0 \;\; &= \;\; \nabla_v \Big[\log \; G(m, e) \Big] \\
&= \;\; \nabla_m \Big[\log \; G(m, e) \Big] + \frac{\partial}{\partial e} \Big[\log \; G(m, e) \Big] v
\end{aligned}
$$

Similarly, $\nabla \log \; f(v_*) = \nabla_m \Big[\log \; G(m, e) \Big] + \dfrac{\partial}{\partial e} \Big[\log \; G(m, e) \Big] v_*$.

So $(\nabla \log \; f)(v) - (\nabla \log \; f)(v_*) \; /\!/ \; v - v_* \quad \forall v, v_* \in \mathbb{R}^N \times \mathbb{R}^N$.

As a purely algebraic consequence of this (here $N \geq 2$ is crucial), there exists $\lambda \in \mathbb{R}$ and $\mu \in \mathbb{R}^N$ such that

$$
\forall v \in \mathbb{R}^N \quad \nabla \log \; f(v) = \lambda v + \mu.
$$

This in turn implies that f is a Maxwellian distribution.

What happens if f is not smooth? As remarked by Desvillettes, the Maxwell–Boltzmann equation (MB), written with a suitable parametrization of the collisions, is invariant under convolution with a Maxwellian. In fact the following property will be enough for us: if

$$M_\delta(v, v_*) = M_\delta(v) M_\delta(v_*) = \frac{e^{-\frac{|v|^2}{2\delta}}}{(2\pi\delta)^{N/2}} \frac{e^{-\frac{|v_*|^2}{2\delta}}}{(2\pi\delta)^{N/2}}$$

then from $f(v)f(v_*) = G(m, e)$ we maydeduce

$$M_\delta \underset{v,v_*}{*} \; ff_* \;\; \equiv \;\; M_\delta \underset{v,v_*}{*} \;\; G(m, e)$$

which is $(f * M_\delta)(f * M_\delta)_* = M_\delta * G(m, e)$. But the convolution by M_δ preserves the class of functions of the form $G(v + v_*, |v - v_*|)$.
Hence, for all $\delta > 0$, $\quad (f * M_\delta)(f * M_\delta)_* = G^\delta(m, e)$, so $f * M_\delta$ is a Maxwellian distribution, and we conclude by letting $\delta \to 0$ that f also is.

<u>Remark</u>: Lions (1994) has a beautiful direct proof that $(\text{MB}) \Rightarrow [f \text{ is } C^\infty]$.

Now let us go on with the proof of Theorem 1 (iii). It is clear from (ii) that $Q(M, M) = 0$. Also $v \cdot \nabla_x M = 0$, so M is a stationary solution. Conversely, let $(f_t)_{t \geq 0}$ be a solution which does not produce entropy. From (ii) we know that

$$f_t(x, v) \;=\; \rho(t, x) \, \frac{e^{-\frac{|v - u(t,x)|^2}{2T(t,x)}}}{[2\pi T(t,x)]^{N/2}}.$$

So we can plug this into the BE, which reduces to $\dfrac{\partial f}{\partial t} + v \cdot \nabla_x f = 0$ since $Q(f_t, f_t) = 0$. Assuming f smooth > 0, we write

$$\frac{1}{f}\left(\frac{\partial f}{\partial t} + v \cdot \nabla_x f\right) = 0$$

and see that it is in fact a system of polynomial equations in v:

$$\frac{1}{\rho}\frac{\partial \rho}{\partial t} + \frac{|v - u|^2}{2T^2}\frac{\partial T}{\partial t} + \langle \frac{v - u}{T}, \frac{\partial u}{\partial t} \rangle - \frac{N}{2T}\frac{\partial T}{\partial t}$$
$$+\frac{v \cdot \nabla_x \rho}{\rho} + \frac{|v - u|^2}{2T^2}v \cdot \nabla_x T - \frac{N}{2T}v \cdot \nabla_x \rho + \frac{\nabla_x u}{T} : [v \otimes (v - u)] = 0$$

Let us use the shorthand $X' = \partial X /\partial t$, and identify powers of v:

9

(0) $\dfrac{\rho'}{\rho} + \dfrac{|u|^2}{2T^2}T' - \dfrac{N}{2T}T' - \dfrac{\langle u, u'\rangle}{T} = 0$

(1) $-\dfrac{T'}{T^2}u + \dfrac{u'}{T} + \dfrac{\nabla\rho}{\rho} + \dfrac{|u|^2}{2T^2}\nabla T - \dfrac{N}{2T}\nabla T - \dfrac{1}{T}\nabla\left(\dfrac{|u|^2}{2}\right) = 0$

(2) $\dfrac{T'}{2T^2}\delta_{ij} - \dfrac{u_j\partial_i T + u_i\partial_j T}{2T^2} + \dfrac{\partial_i u_j + \partial_j u_i}{2T} = 0$

(3) $\dfrac{\nabla T}{2T^2} = 0$

For a detailed version of the reasoning which follows, see Desvillettes (1990).

- From (3) we have $\nabla_x T = 0$, so $T = T(t)$.

- Plugging this into (2) we find

$$\dfrac{\partial_i u_j + \partial_j u_i}{2} = -\dfrac{T'}{2T}\delta_{ij} \quad \text{(in matrix notation, } \nabla_{sym}u = -\dfrac{T'}{2T}\,Id)$$

and this expression is independent of x. From this one easily sees that $\partial_{ij}u_i = 0$, then $\partial_{ii}u_j = 0$, then u is an affine map which can be written as $u(x) = \lambda x + \Lambda \wedge x + u_0$. ($\lambda \in \mathbb{R}$, $\Lambda \in \mathbb{R}^3$)

From the boundary conditions, $\displaystyle\int_\Omega (\nabla \cdot u)\,dx = \int_{\partial\Omega} u \cdot n = 0$ (periodic $\Rightarrow \partial\Omega = \emptyset$, bounce-back $\Rightarrow u = 0$ on $\partial\Omega$, specular $\Rightarrow u \cdot n = 0$). So $\lambda = 0$ and $u(x) = \Lambda \wedge x + u_0$. This is possible for $\Lambda \neq 0$ only if Ω has an axis of symmetry $k \,/\!/\, \lambda$ and we are considering specular reflection, a case which we have excluded. So $u(x) = u_0(t)$, but u_0 can be $\neq 0$ only in the periodic case. To sum up, at this point we know that $u \equiv 0$; or $u \equiv u(t)$ in the periodic case.

- Going back to (1): if we know $u \equiv 0$, then

$$\dfrac{\nabla\rho}{\rho} = 0, \quad \text{so} \quad \rho = \rho(t);$$

and if we know $u = u(t)$, then $\dfrac{\nabla\rho}{\rho}$ is independent of x, so $\rho = \exp\Big(A(t)\cdot x + B(t)\Big) = \rho_0(t)e^{A_0(t)\cdot x}$.

10

• Then, from (0) ρ' only depends on t, and so does ρ.

Conclusion: In all the cases, we know that
- ρ, u, T do not depend on x (case of the torus)
- or ρ, T do not depend on x and $u \equiv 0$ (other cases).

But the global conservation laws imply $\int \rho = 1$, $\int \rho \frac{|u|^2}{2} + \frac{N}{2} \int \rho T = \frac{N}{2}$ (conservation of mass and energy). Also in the periodic case $\int \rho u = 0$ (conservation of momentum). This implies $\rho \equiv 1$, $u \equiv 0$, $T \equiv 1$. The proof is complete. $\square$

Theorem 1 is very important! Point (ii) is at the basis of the problem of hydrodynamic limit, while point (iii) is crucial in the problem of trend to equilibrium. We should note that in the particular situation when $f_t(x, v) = f_t(v)$ [spatial homogeneity, $x \in \mathbb{T}^N$] then (iii) is superfluous.

1.3. What this course is about: convergence to equilibrium

<u>Problem of trend to equilibrium</u>: Let be given a solution $(f_t)_{t \geq 0}$ of the BE, starting from some initial datum f_0 which is out of equilibrium (i.e., with our normalizations, $f_0 \neq M$). Is it true that

$$f_t \xrightarrow[t \to \infty]{} M \quad ?$$

<u>Remark 1</u>: A major difference between the problem of hydrodynamic limit and the problem of trend to equilibrium is that in the latter case, one is interested in proving that the solution approaches a <u>global</u> equilibrium, while in the former one only expects to come close to a <u>local</u> equilibrium. Accordingly, as a general rule the problem of trend to equilibrium is much more sensitive to boundary conditions, than the hydrodynamic limit problem.

<u>Remark 2</u>: Another difference is that one expects shocks occurring asymptotically as the Knudsen number goes to O, at least in the compressible hydrodynamic limit, while for fixed Knudsen number there is no known reason for appearance of shocks — at least in a strictly convex domain.

<u>Remark 3</u>: The H Theorem is <u>not</u> the only way to attack the problem of trend to equilibrium. But it seems to be by far the most robust.

<u>Remark 4</u>: A lot of other kinetic (or not) models have similar features. The choice of the BE is interesting because of historical reasons, because it is one of the very, very few models which have been derived from Newton's laws of motion "alone" (in some particular regimes), and also because it concentrates a lot of serious difficulties.

As soon as one is able to get a little bit of information about solution of the BE (which is very difficult!), one can turn Theorem 1 into some statement of trend to equilibrium. For instance we have the following result by some easy compactness argument:

<u>Soft Theorem</u>: Let $(f_t)_{t\geq 0}$ be a solution (meaning not made precise here) of the BE, with $H(f_0) < +\infty$ and

$$\sup_{t\geq 0} \int f_t(x,v)|v|^{2+\delta}\,dv\,dx < +\infty \quad \text{for some} \quad \delta > 0$$

then

$$f_t \quad \xrightarrow[t\to\infty]{} \quad M \quad \text{in weak} - L^1 \text{ sense}$$

Let me insist that nobody knows how to construct such solutions under natural assumptions on f_0 (say, f_0 belonging to Schwartz class of C^∞ rapidly decaying functions)!! Only under some restrictions on f_0 are such a priori estimates available. But this is a different issue, it concerns the study of the Cauchy problem, not so much the mechanism of trend to equilibrium.

<u>Goal of the course</u>: Present a program on entropy production, indirectly inspired by the pioneering works of Kac (in the fifties) and McKean (in the sixties), really started by Carlen and Carvalho at the beginning of the nineties, currently very active. Its goal is to **provide explicit bounds on the rate of convergence by a detailed analysis of the entropy production in situations where the Cauchy problem is well understood.**

When Kac and McKean began to think about convergence to equilibrium, the models they were able to "attack" were extremely simplified compared to those that experts in the Cauchy problem were able to treat. This remained so for a long time, but now it is the reverse situation: the study of the Cauchy problem, while very active, is still very far from completion, while the

entropy production program, likely, will soon cover all interesting situations under assumptions of suitable a priori estimates (exactly which estimates is an interesting point because this may be taken as a guide for the Cauchy problem).

The recent story of entropy production has been very much influenced by the field of logarithmic Sobolev inequalities (but as we shall see, there are a lot of distinctive, very interesting features) and by some theoretical portions of information theory.

The basic idea behind the entropy production program is quite natural and simple: given that (f_t) is far from equilibrium at some time t, try to estimate from below the amount of entropy which has to be produced at later times.

Here is an example of a recent result (based on work by C. Mouhot and myself) obtained by this program:

Let $f_0(v)$ satisfy $\displaystyle\int f_0(v)\,dv = 1$, $\displaystyle\int f_0(v)\,v\,dv = 0$, $\displaystyle\int f_0(v)\frac{|v|^2}{2}\,dv = \frac{N}{2}$, $\displaystyle\int f_0(v)\,|v|^3\,dv < +\infty$, $f_0 \in L^p$ for some $p > 1$. Let $(f_t)_{t \geq 0}$ be the unique solution of the spatially homogeneous Boltzmann equation (SHBE) starting from f_0 with the collision kernel $B(v - v_*, \sigma) = |v - v_*|$ (in dimension $N = 3$ this is the usual hard spheres collision kernel). Then

$$\left\| f_t - M \right\|_{L^1} = O(t^{-\infty}).$$

More precisely, for any $\epsilon > 0$ there exists an explicit constant $C_\epsilon(f_0)$ such that $\forall t > 0 \quad \|f_t - M\|_{L^1} \leq C_\epsilon t^{-1/\epsilon}$.

Nota: Since the first construction of solutions of the SHBE with hard sphere cross-section by Carleman (1932), this is the very first result asserting better decay than just $O(t^{-K})$ for some $K > 0$ (Toscani & Villani, 1999b) with explicit constants.

I shall conclude this introduction with two remarks:

Remark 5: As I said above, the program was intended to cover situations where the Cauchy problem is well understood. Does it provide information

in more general situations? I don't know! The reader can try to extract from the arguments some principles which would help in such settings.

Remark 6: In principle the entropy production program can also be applied to hydrodynamic limits. This was in fact the main motivation of Carlen & Carvalho (1992, 1994). Carlen et al. (1998) have demonstrated on an interesting baby model (for which the hydrodynamic model is a set of two nonlinear ODE's) that this goal is sometimes feasible. But there is in fact not much in common between these two problems, besides the fact that tools which are used to estimate entropy production in a certain context, are likely to be useful to estimate entropy production in another context.

2. Tentative panomara

The entropy production program is closely linked to several other topics which possess their own interest:

- the central limit theorem for Maxwellian molecules: this is an alternative approach of trend to equilibrium, which has been first suggested by McKean (1966), developing on an idea by Kac. Just as the central limit theorem of i.i.d. random variables, the central limit theorem for Maxwellian molecules was implemented quantitatively by several methods, in particular via:

  - Fourier transform (Toscani and co-authors)

  - Monge–Kantorovich distances (Tanaka)

  - information-theoretical techniques (Carlen and Carvalho)

  (McKean's method was a clever variant of Trotter's elementary proof of the central limit theorem). This program has recently led to very sharp results for the study of trend to equilibrium, but it seems to concern only the case where there is no x (position) variable and in addition $\overline{B(|v - v_*|, \cos\theta)} = b(\cos\theta)$ [spatially homogeneous Maxwellian case].

- Kac's problem of trend to equilibrium for an N-particle system: this is an attempt of understanding trend to equilibrium on the basis of the asymptotic (as $N \to \infty$ and $t \to \infty$ together) properties of the master equation describing the evolution of a system of N colliding particles. After decades of sleep, this program has been revived by recent works (Janvresse; Carlen, Carvalho and Loss; Maslen), but it is still very far from satisfactory understanding.

- the behavior of Fisher's information along solutions of the Boltzmann equation: McKean (1966) made the striking discovery that the Fisher information $I(f) = \int \dfrac{|\nabla f|^2}{f}$ is a Lyapunov functional for a simple one-dimensional caricature of the Boltzmann equation. He believed this was a particular feature of dimension 1, but it is not, and this result has been extended to multi-dimensional Boltzmann equations (Toscani and subsequent work), however only in the case of "Maxwellian molecules".

There is a tight connection between this problem and the entropy production issue, as we shall see.

I shall briefly review these three topics in section 6. Some information can also be found in my 2003 paper about "Cercignani's conjecture".

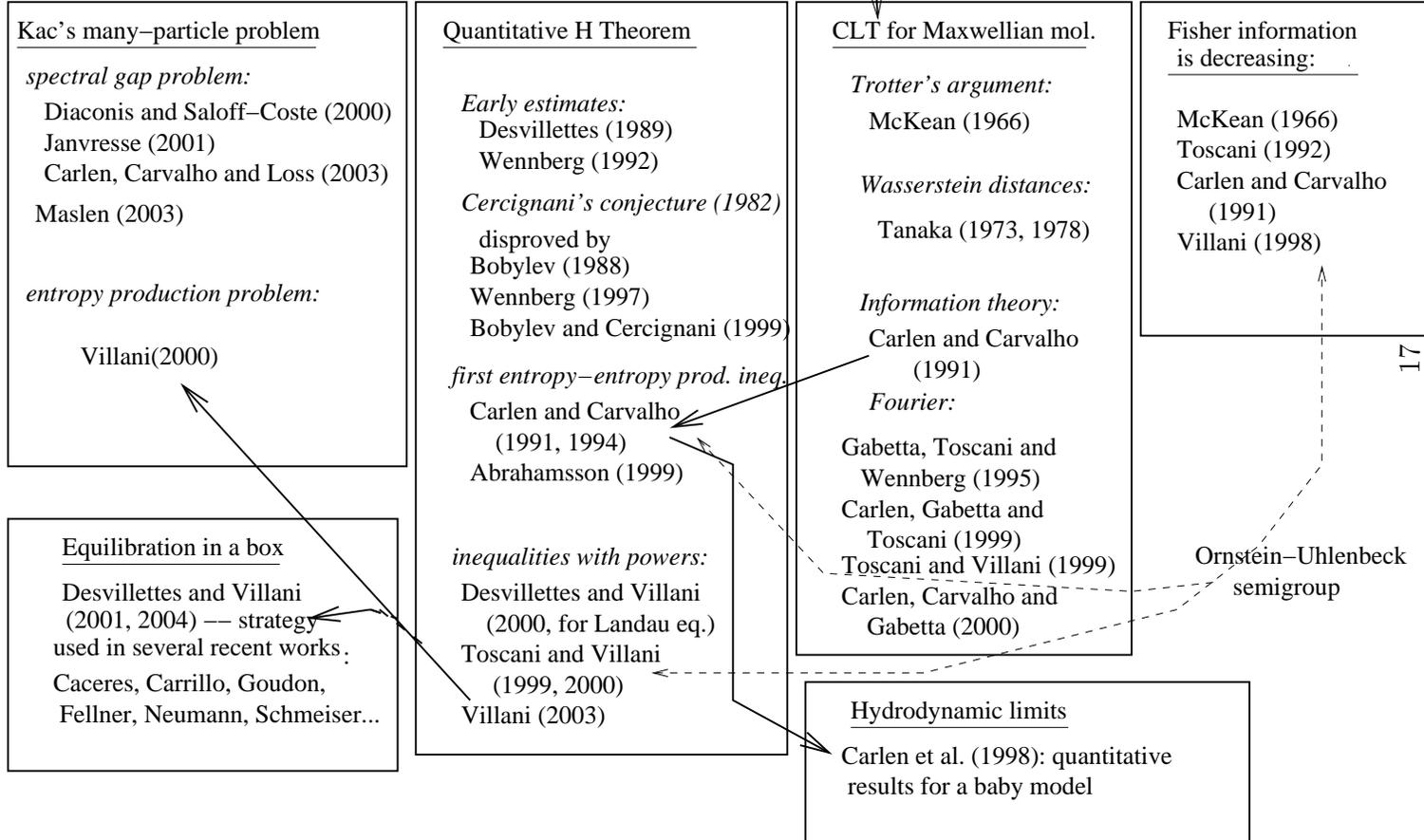Also, there are alternative programs based on entropy production:

- compactness arguments (possibly for weak solutions), non quantitative.

    $\longrightarrow$ In the context of hydrodynamic limits, compactness is crucial for the Bardos–Golse–Levermore program. The text of my 2001 Bourbaki seminar (published in 2002) tries to describe the essentials of that program, which was recently led to completion by the joint efforts of Bardos, Golse, Levermore, Lions, Masmoudi, Saint-Raymond.

    $\longrightarrow$ In the problem of trend to equilibrium, compactness arguments are rather easy and general but quite disappointing: at present, they can essentially be applied only in situations where quantitative arguments lead to infinitely stronger results (and yet, not all of them!). This is because the bound on $\int f_t(x,v)|v|^{2+\delta}\, dv\, dx$ is very difficult to establish.

- linearization in a close-to-equilibrium setting. In this process $H(f)$ turns into the $(L^2\text{norm})^2$ of the perturbation $f - M$ in some weighted L^2 space: $\int \dfrac{(f-M)^2}{M}\, dv\, dx$ (extremely strong norm!). After a change of unknown, trend to equilibrium can hopefully be studied by spectral gap methods. This approach is unbeatable in a (very-)close-to-equilibrium setting, but <u>cannot</u> predict anything quantitative when f_0 is not very close to M.

On the next page is an attempt to summarize the whole picture.

Kac's question:
How fast is the approach to equilibrium?

many–particle system

one–particle marginal

Kac's many–particle problem

spectral gap problem:
  Diaconis and Saloff–Coste (2000)
  Janvresse (2001)
  Carlen, Carvalho and Loss (2003)
 Maslen (2003)

entropy production problem:

  Villani(2000)

Quantitative H Theorem

 Early estimates:
  Desvillettes (1989)
  Wennberg (1992)

 Cercignani's conjecture (1982)
  disproved by
  Bobylev (1988)
  Wennberg (1997)
  Bobylev and Cercignani (1999)

 first entropy–entropy prod. ineq.
  Carlen and Carvalho
   (1991, 1994)
  Abrahamsson (1999)

 inequalities with powers:
  Desvillettes and Villani
   (2000, for Landau eq.)
  Toscani and Villani
   (1999, 2000)
  Villani (2003)

CLT for Maxwellian mol.

 Trotter's argument:
  McKean (1966)

 Wasserstein distances:
  Tanaka (1973, 1978)

 Information theory:
  Carlen and Carvalho
   (1991)
 Fourier:

  Gabetta, Toscani and
   Wennberg (1995)
  Carlen, Gabetta and
   Toscani (1999)
  Toscani and Villani (1999)
  Carlen, Carvalho and
   Gabetta (2000)

Fisher information is decreasing:

  McKean (1966)
  Toscani (1992)
  Carlen and Carvalho
   (1991)
  Villani (1998)

Ornstein–Uhlenbeck semigroup

Equilibration in a box

 Desvillettes and Villani
 (2001, 2004) –– strategy
 used in several recent works .
 Caceres, Carrillo, Goudon,
  Fellner, Neumann, Schmeiser...

Hydrodynamic limits
 Carlen et al. (1998): quantitative
  results for a baby model

17

3. Reminders from information theory

In this section I shall review some necessary preliminaries before turning to the study of the Boltzmann equation.

3.1. Background and definitions

Information theory was founded by Shannon (1948) for applications to signal transmission, error-correcting, etc. An excellent reference about modern information theory is Cover and Thomas (1991), Elements of information theory. For the particular topics which will be discussed here, another user-friendly reference is Chapter 10 of the recent book Sur les inégalités de Sobolev logarithmiques (Panoramas et Synthèses, n^{0}10, SMF, 2000, by Ané et al., in french). Part of information theory deals with the study of functionals which aim at quantifying the "information" contained in a random variable; the two most popular of these functionals were introduced respectively by Fisher (1925) and Shannon (1948):

$$\bullet \text{Shannon's entropy} \quad S(f) \;\; = \;\; - \int_{\mathbb{R}^N} f \, \log \, f$$

$$= \;\; \text{Boltzmann's } H \text{ function, up to change of sign}$$

$$\bullet \text{Fisher's information } I(f) \;\; = \;\; \int_{\mathbb{R}^N} \frac{|\nabla f|^2}{f} = 4 \int_{\mathbb{R}^N} |\nabla \sqrt{f}|^2 = - \int_{\mathbb{R}^N} f \Delta(\log f)$$

Whenever X is a random variable whose law has Lebesgue density f, one defines $S(X) = S(f)$, $I(X) = I(f)$. In the sequel I shall work with $H = -S$ instead of S.

Historical motivations: In Shannon's theory of communication, the entropy measures the (asymptotically) optimal rate of compression that one can apply to a signal without (asymptotically) losing any of the information which it carries. Here "asymptotically" means: in the limit when one cuts the signal into very long samples.

In Fisher's theory of "efficient statistics", Fisher's information measures how difficult it is to infer the value of an unknown parameter (appearing in the law of a random variable) from the empirical observation. Typically: assume that you observe a signal $X_\theta = X + \theta$, where the law of X is known and equal to f, and $\theta \in \mathbb{R}$ (say) is deterministic but unknown to the observer. Given a large sample of independent observations of X_θ, one can construct

a statistic $\hat{\theta}$ (observed values) which is hoped to give a good approximation of θ, with high probability. But the Cramér-Rao inequality tells you that, no matter how cleverly you design $\hat{\theta}$, if it is unbiased (i.e. $E\hat{\theta} = \theta$, where E stands for expectation), then its variance cannot be smaller than $1 / [nI(f)]$ (here n = number of observations). Hence, the lower $I(f)$ is, the most difficult it is to construct efficient statistics. Morever, one can show that, asymptotically as $n \to \infty$, the maximum likehood estimate achieves this lower bound $[nI(f)]^{-1}$. In higher dimensions ($N \geq 1$), the Fisher information *matrix* is a basic object in statistics:

$$I_{ij}(f) = \int_{\mathbb{R}^N} \frac{\partial_i f \partial_j f}{f},$$

of which $I(f)$ is just the trace.

<u>Remark</u>: Fisher's information is sometimes encountered in kinetic theory under the name "Linnik functional", for the following reason: McKean (1966) borrowed from Linnik (1959) the idea to use Fisher's information in the context of the central limit theorem, and in his paper omitted to quote Fisher. This terminology should be banned to avoid further confusion.

A first hint that H and I can be used to represent some kind of "information" is their superadditive properties: an information about the joint law of (X, Y) is usually much more valuable than an information about the marginal laws of X and Y. And, as a matter of fact, it is the case that

$$\begin{aligned}
H((X,Y)) &\geq H(X) + H(Y) \\
I((X,Y)) &\geq I(X) + I(Y),
\end{aligned}$$

and in both cases, equality occurs if and only if X, Y are independent.

The first inequality (superadditivity of H) is well-known (see any textbook on information theory). The second one can be proven by a simple and short computation, which the reader can try to find out as an exercise; it is also a particular case of more general and more delicate information-theoretical inequalities proven by Carlen (1991).

19

<u>Conventions</u>

In a more intrinsic definition, H and I should be defined relatively to some <u>reference measure</u> ν (not necessarily a probability):

$$H_\nu(f) = \int f \log f \, d\nu \quad I_\nu(f) = \int \frac{|\nabla f|^2}{f} \, d\nu$$

I shall use the notation

$$H(\mu|\nu) = H_\nu\left(\frac{d\mu}{d\nu}\right) \qquad I(\mu|\nu) = I_\nu\left(\frac{d\mu}{d\nu}\right)$$

and call these functionals the relative entropy of μ with respect to ν, relative Fisher information of μ with respect to ν. In information theory $H(\mu|\nu)$ is called the Kullback entropy of μ with respect to ν, or informational divergence, or Kullback-Leibler distance, and probably has many other names.

<u>Abuse of notation</u>: When $d\mu = f(x)\,dx$, $d\nu = g(x)\,dx$, I shall write $H(f|g)$, $I(f|g)$ for $H(\mu|\nu)$, $I(\mu|\nu)$.

Finally, an object which is very popular in information theory, ever since the time of Shannon, is the <u>power entropy</u>:

$$\mathcal{N}(f) = \frac{e^{-\frac{2}{N}H(f)}}{2\pi e} = \mathcal{N}(X) \quad \text{if law } (X) = f(x)\,dx$$

3.2. Inequalities

We can now state three famous inequalities about H and I.

1. <u>the Blachman–Stam inequality</u> (three <u>equivalent</u> formulations)

$$
\begin{aligned}
I(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) &\leq \lambda I(X) + (1-\lambda)I(Y) \\
I(X+Y) &\leq \lambda^2 I(X) + (1-\lambda)^2 I(Y) \\
\frac{1}{I(X+Y)} &\geq \frac{1}{I(X)} + \frac{1}{I(Y)}
\end{aligned}
$$

whenever $\lambda \in [0,1]$ and X, Y are independent random variables.

20

2. the Shannon–Stam inequality (three equivalent formulations)

$$
\begin{aligned}
H(\sqrt{\lambda}X + \sqrt{1-\lambda}Y) &\leq \lambda H(X) + (1-\lambda)H(Y) \\
H(X+Y) &\leq \lambda H(X) + (1-\lambda)H(Y) + \frac{N}{2}[\lambda \log \lambda + (1-\lambda)\log(1-\lambda)] \\
\mathcal{N}(X+Y) &\geq \mathcal{N}(X) + \mathcal{N}(Y)
\end{aligned}
$$

3. the Stam-Gross logarithmic Sobolev inequality (three equivalent formulations)

$$
\begin{cases}
H(f \mid M) &\leq \quad \dfrac{1}{2}I(f \mid M) \\[2mm]
\mathcal{N}(f)I(f) &\geq \quad N
\end{cases}
$$

whenever f is a probability density on $\mathbb{R}^N$

$$
\int h^2 \log h^2 M \leq 2 \int |\nabla h|^2 M + \left(\int h^2 M \right) \log \left(\int h^2 M \right)
$$

Paternity issues: Shannon (1948) conjectured inequality 2 (in the $\mathcal{N}$ formulation) and found it useful for certain issues in signal transmission. Then Stam (1959) proved 1 and 2, using a suggestion by de Bruijn. In the same work he noticed the validity of 3 (in the 2nd formulation) for $N = 1$ (actually, the general case can be reduced to this particular case). Later Blachman (1965) gave a implified proof for 1. In 1975 Gross, unaware of Stam's work, proved 3 (in the 3rd formulation) and, most importantly, showed that this inequality is equivalent to the hypercontractivity of the Ornstein–Uhlenbeck semigroup. (Although this is not widely acknowledged, the concept of hypercontractivity goes back at last to Bonami (1971).) A few years later, Lieb (1978), unaware of Shannon and Stam, re-proved 2 as a consequence of optimal Young inequalities. It is only at the beginning of the nineties that the whole picture was clarified, with the papers of Carlen (1991), providing synthetical and simple proofs of all this, together with multiple links.

I myself wrote a very short paper about the entropy power functional (Villani, 2000c).

De Bruijn's identity: Let $(P_t)_{t\geq 0}$ be the heat semigroup (generated by the PDE $\partial_t f = \Delta f$), then, regardless of regularity issues (which do not really

bother here),

$$\frac{d}{dt}H(P_t f) = -I(P_t f).$$

It is easy to check this fact, which as we shall see is very convenient.

Geometrical interpretation: There is an appealing interpretation of all this in geometrical language. From Boltzmann's original definition of entropy ("$S = k \ln W$") we know that $S = -H$ can be thought of as an infinite-dimensional analogue of $A \mapsto \dfrac{1}{n} \log |A|_n$ ($A \subset \mathbb{R}^n$, $|A|_n = $ Lebesgue measure of A), or equivalently that $\mathcal{N}$ plays the role of $A \mapsto |A|_n^{2/n}$. Under this analogy, one can develop the following correspondence:

$$
\begin{array}{rcl}
\text{probability measures } f & \to & \text{convex sets } A \text{ in } \mathbb{R}^n \ (n \to \infty) \\
\text{entropy power } \mathcal{N}(f) & \to & |A|_n^{1/N} \\
\text{Maxwellian distribution } M & \to & \text{ball of radius } \sqrt{n} \text{ in } \mathbb{R}^n, B_n(\sqrt{n}) \\
\text{Shannon} - \text{Stam ineq.} & \to & \text{Brunn} - \text{Minkowski inequality} \\
& & |A + B|_n^{1/n} \geq |A|_n^{1/n} + |B|_n^{1/n} \\
& & (\text{or rather the restricted version :} \\
& & |A + B|_n^{2/n} \geq |A|_n^{2/n} + |B|_n^{2/n}) \\
\text{Logarithmic Sobolev ineq.} & \to & \text{isoperimetric inequality} \\
& & \dfrac{|\partial A|_{n-1}}{n|A|_n}|A|_n^{1/n} \geq |B_n(1)|_n^{1/n} \\
\text{Fisher information } I(f) & \to & \dfrac{\text{surface }(A)}{n|A|_n} = \dfrac{|\partial A|_{n-1}}{n|A|_n}.
\end{array}
$$

In this array there is just one missing inequality for the analogy to be complete: this is the Blachman–Stam inequality. The following finite-dimensional analogue of it was suggested by Dembo, Cover and Thomas (1991):

$$\frac{|(A + B)|_n}{|\partial(A + B)|_{n-1}} \geq \frac{|A|_n}{|\partial A|_{n-1}}$$

Here $|\partial A|_{n-1}$ is the surface of A, defined as $\lim\limits_{\epsilon \to 0} \dfrac{|A + \epsilon B_n(1)|_n - (A)}{\epsilon}$.

This inequality is apparently not known. It cannot hold time for general compact sets A, B, but it has a chance to be true for convex bodies, and by

means of some Alexandrov–Fenchel inequalities for mixed volumes, one can check that it <u>is</u> true if A is convex and B is a ball.

Before going back to Boltzmann's H theorem, let us record three particular formulations of the above-mentioned inequalities; these are the ones which one should mostly keep in mind for applications in kinetic theory. Let

$$f_\lambda = \text{law of } \sqrt{\lambda}X \quad \text{if } X \text{ is random with law } f(x)\,dx.$$

Then

$$\begin{cases} I(f_{\frac{1}{2}} * g_{\frac{1}{2}}) \leq \ \dfrac{1}{2}\Big[I(f) + I(g)\Big] \\[2ex] H(f_{\frac{1}{2}} * g_{\frac{1}{2}}) \leq \ \dfrac{1}{2}\Big[H(f) + H(g)\Big] \\[2ex] H(f \mid M) \leq \ \dfrac{1}{2}I(f \mid M). \end{cases}$$

4. Quantitative H theorem

4.1. Entropy-entropy production inequalities

Recall part (ii) of Theorem 1: whenever $B > 0$ a.e. and $f \in L^1_2(\mathbb{R}^N_v)$ $(f \geq 0)$ satisfies $D(f) = 0$ [B enters in the expression of D, so an assumption on B is mandatory], then f is a Maxwellian distribution:

$$
(1) \qquad f(v) = \rho \, \frac{e^{-\frac{|v-u|^2}{2T}}}{(2\pi T)^{N/2}}.
$$

We now wish to get a quantitative replacement for that, stating that if B is positive enough and $D(f)$ is small, then f is close to a Maxwellian. The first such results were established by Desvillettes (1989) in the form (essentially)

$$
(2) \qquad D(f) \geq K_{f,R} \, \inf_{m \in \mathcal{M}} \int_{|v|<R} |\log f - \log m| \, dv
$$

where $\mathcal{M}$ is the space of all Maxwellian distributions, and $K_{f,R}$ depends on R and f via certain bounds, in particular on the strict positivity of f. Desvillettes' theorem was established under the assumption that B is bounded from below, but later Wennberg (1992) relaxed this restriction.

Essentially the best that one can hope for is an entropy-entropy production inequality, which would relate $D(f)$ on one hand, $H(f \mid M^f)$ on the other hand, where M^f is the Maxwellian distribution having the same hydrodynamic parameters (mass, momentum, energy) as f. In the present context, an EEP inequality is a functional inequality of the form

$$
(3) \qquad D(f) \geq \Theta\Big(H(f \mid M^f)\Big)
$$

[Note that $H(f \mid M^f) = H(f) - H(M^f)$], where $\Theta : \mathbb{R}_+ \to \mathbb{R}_+$ is some increasing function, with $\Theta(\alpha) > 0$ if $\alpha > 0$. Such an inequality implies at once convergence to equilibrium for the spatially homogeneous BE (only collisions are taken into account, there is no dependence on the position). Moreover, one can then estimate the speed of convergence to equilibrium in term of the behavior of $\Theta(\alpha)$ close to $\alpha = 0$; the most favorable situation being that when $\Theta(\alpha) = K \cdot \alpha$ for α small enough (exponential decay).

We already encountered an EEP inequality for another model: consider the (linear) Fokker–Planck equation

(4)
$$\frac{\partial f}{\partial t} = \Delta f + \nabla \cdot (fv) \quad t \geq 0, \ v \in \mathbb{R}^N$$

For this equation the role of the (negative of the) entropy is played by the Kullback information $H(f \mid M)$, and its dissipation along the Fokker–Planck semigroup is

$$D_{\text{FP}}(f) = I(f \mid M)$$

Thus the logarithmic Sobolev inequality of Chapter 3 states that, for any probability distribution f,

(5)
$$D_{\text{FP}}(f) \geq 2H(f \mid M)$$

which incidentally implies by Gronwall's lemma $H(S_t f \mid M) \leq e^{-2t} H(f \mid M)$ S_t =semigroup associated to the FP eq. (4). The assumption that f is a probability distribution is of course not the point: an adequate change of parameters allows to treat any nonnegative distribution.

Inequality (5) is a very strong EEP inequality: first because the function Θ is linear ($\Theta(\alpha) = 2\alpha$), secondly because it holds true without restriction upon f. For more complicated models, certainly we will not be so lucky and Θ may be vanishing to higher order, or the inequality may hold only under some restrictions (smoothness, decay, positivity...).

4.2. EEP inequalities for the Boltzmann equation

The goal of this chapter is to discuss EEP inequalities for the Boltzmann collision operator, under various assumptions on the kernel B and on f.

Normalization: Assume

(6)
$$\int_{\mathbb{R}^N} f \, dv = 1, \quad \int_{\mathbb{R}^N} fv \, dv = 0, \quad \int_{\mathbb{R}^N} f|v|^2 \, dv = N.$$

Then

$$M^f = M = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{N/2}}.$$

Notation: This class of f's will be denoted by $\mathcal{C}(1, 0, 1)$.

It is easy to extend the results which will be established under this normalization, to the general case. Indeed, if f has mass ρ, momentum u, temperature T, then

$$f(v) = \frac{\rho}{T^{N/2}}\left(\frac{v-u}{\sqrt{T}}\right) \quad \text{where} \quad \int \widetilde{f}(v)\left(\begin{array}{c} 1 \\ v \\ |v|^2 \end{array}\right)dv = \left(\begin{array}{c} 1 \\ 0 \\ N \end{array}\right)$$

and one has

$$\left\{ \begin{array}{ll} H(f\mid M^f) & = \rho\, H(\widetilde{f}\mid M) \\ D(f) & = \rho^2\, \widetilde{D}(\widetilde{f}) \end{array}\right.$$

where $\widetilde{D}$ is the entropy production functional associated with the modified kernel $\widetilde{B}(v-v_*,\sigma) = B(\sqrt{T}(v-v_*),\sigma)$.

It was once conjectured (Cercignani, 1982) that under suitable assumptions on B, there would exist a constant $K > 0$ such that for all f satisfying (6), $D(f) \geq K H(f\mid M)$. This conjecture was then disproved under more and more general assumptions (Bobylev, 1988; Wennberg, 1997; Bobylev and Cercignani, 1999). In particular, the following important "counter-theorem" by Bobylev and Cercignani (1999) shows that there is no hope of such an inequality holding true, even if we restrict to probability densities f which are very smooth (in the sense of Sobolev norms), decay quite fast at infinity (in the sense of having a lot of finite moments), and are strictly positive (in the strong sense of being bounded below by a Maxwellian distribution). Before I state the theorem, some further remarks and notation:

Remark 1: I consider these estimates (Sobolev norms, moments, lower bound like $Ke^{-c|v|^q}$) as the natural estimates associated with the nonlinear Boltzmann equation. This will be discussed in more detail in the next chapter. I stress that, in a linearized context, it would be natural to look at stronger estimates like $L^p(M^{-1}dv)$ norms.

Functional spaces:

$$\|f\|_{L^p_s} = \left[\int_{\mathbb{R}^N} f(v)^p\left(1+|v|^p\right)^s dv\right]^{1/p};$$

in particular $\|f\|_{L^1_s} = \int f(v)(1+|v|)^s dv$;

$$\|f\|_{H^k} = \Big[\sum_{j \leq k} \int_{\mathbb{R}^N} |D^j f(v)|^2 dv \Big]^{1/2}.$$

In the sequel I shall use L^1_s and H^k norms as a natural way to quantify the decay at infinity and the smoothness, respectively. It would also be possible to combine both informations in a single two-parameter family of norms:

$$\|f\|_{H^k_s} = \Big[\int_{\mathbb{R}^N} |D^k f(v)|^2 (1+|v|^2)^s dv \Big]^{1/2},$$

but let's rather consider separately decay and smoothness. At last, here comes the above-mentioned theorem (this is no exactly the formulation given in the reference quoted, but it is straightforward to make the transformation):

<u>Theorem 2</u> (Bobylev & Cercignani, 1999). *Let B satisfy*

$$\int_{S^{N-1}} B(v-v_*, \sigma) d\sigma \leq C_B (1+|v-v_*|^\gamma)$$

for some $\gamma \in [0,2]$. Then there exist sequences $(M_s)_{s \geq 0}$, $(S_k)_{k \geq 0}$ [to be understood as <u>moment</u> and <u>smoothness</u> bounds] and ρ_0, $A_0 > 0$, such that

$$\inf \left\{ \begin{array}{l} \dfrac{D(f)}{H(f \mid M)}; \quad f \in \mathcal{C}(1,0,1); \quad \forall k \in \mathbb{N} \ \|f\|_{H^k} \leq S_k; \\[2mm] \forall s \in \mathbb{N} \ \|f\|_{L^1_s} \leq M_s; \quad \forall v \in \mathbb{R}^N \ f(v) \geq \rho_0 e^{-A_0 |v|^2} \end{array} \right\} = 0.$$

In particular, it is impossible to have an inequality like $D(f) \geq K H(f \mid M)$ if we only assume bounds on $\|f\|_{L^1_s}$, $\|f\|_{H^k}$, and a lower bound.

<u>Remark 2</u>: One can moreover impose that any finite number of the bounds M_s, S_k, ρ_0, A_0, be arbitrarily close to their equilibrium values. More precisely, for any s_*, k_*, and any $\epsilon > 0$, one can impose

$$\left\{ \begin{array}{l} M_s \leq \|M\|_{L^1_s} + \epsilon \quad \forall s \leq s_* \\[1mm] S_k \leq \|M\|_{H^k} + \epsilon \quad \forall k \leq k_* \\[1mm] \rho_0 \geq 1 - \epsilon \\[1mm] A_0 \leq \dfrac{1}{2} + \epsilon. \end{array} \right.$$

Principle of the proof: it consists in constructing an explicit sequence of functions f_n) with $\dfrac{D(f_n)}{H(f_n \mid M)} \xrightarrow[n \to \infty]{} 0$. The sequence is simply defined as a perturbation of M, with "fat tails" (in a very weak sense):

$$f_n = M + \epsilon_n \; \varphi\left(\frac{\cdot}{\theta_n}\right),$$

where θ_n is large enough, ϵ_n is small enough, and φ is an $\underline{\text{arbitrary}}$ smooth, positive function bounded below by a Maxwellian.
Then f_n does not exactly belong to $\mathcal{C}_{1,0,1}$, but a slight perturbation of it (multiply by a constant and rescale slightly) does. The important point is that θ_n is very large, so that $\varphi\left(\dfrac{v}{\theta_n}\right)$ decays quite "slowly" at infinity, but ϵ_n is very, very small, so that f_n really is a perturbation of the equilibrium.

Next come some positive results. The first EEP inequality for the Boltzmann operator was established by Carlen and Carvalho (1992, 1994); This a very clever, complicated construction, and the function Θ appearing in the EEP inequality was very, very flat close to 0. A huge improvement was then provided by Toscani and myself (1999b), and recently I improved again on these results (Villani, 2003). In the sequel I shall describe a subset of what is presently known.

Theorem 3 (Toscani & Villani, 1999). *Let B satisfy*

$$B(v - v_*, \sigma) \geq K_B (1 + |v - v_*|)^{-\beta} \quad \text{for some} \quad \beta \geq 0.$$

Then, for all $\underline{\epsilon > 0}$,

$$\forall p > 1, \quad \forall I, \quad \forall \rho_0, \forall A_0, \forall q_0, \; \exists s = s(\epsilon, p, q_0 B); \quad \forall M$$

$$\exists K_\epsilon > 0; \quad \left[\|f\|_{L^p} \geq I, \quad \|f\|_{L^1_s} \leq M, \quad f \geq \rho_0 e^{-A_0 |v|^{q_0}}\right]$$

$$\implies D(f) \geq K_\epsilon H(f \mid M)^{1+\epsilon}.$$

In a more intelligible statement: $D(f) \geq K_\epsilon(f) H(f \mid M)^{1+\epsilon}$, where $K_\epsilon(f)$ only depends on an upper bound on $\|f\|_{L^p}$ ($p > 1$) arbitrary), an upper bound on high enough order moments of f, and a lower bound on f. (In short $K_\epsilon(f)$ depends on the $\underline{\text{integrability}}$ - $\underline{\text{decay}}$ - $\underline{\text{strict positivity}}$ of f.)

Shorthand: Let's just write "$D(f) \geq K(f)H(f \mid M)^{1+0}$ if f has some integrability, decays very fast, is strictly positive".

Theorem 4 (Villani, 2003). *Let B satisfy*

$$B(v - v_*, \sigma) \geq K_B \inf\left(|v - v_*|^\gamma, |v - v_*|^{-\beta}\right) \quad \text{for some} \quad \beta, \gamma \geq 0.$$

Then, for all $\epsilon > 0$

$$\forall \rho_0 \quad \forall A_0 \quad \forall q_0 \quad \exists k = k(\epsilon, q_0, B) \quad \exists s = s(\epsilon, q_0, B) \; / \; \forall S \; \forall M$$

$$\exists K_\epsilon > 0; \; \left[\|f\|_{H^k} \leq Z, \; \|f\|_{L^1_s} \leq M, \; f \geq \rho_0 e^{-A_0|v|^{q_0}}\right]$$

$$\implies D(f) \geq K_\epsilon H(f \mid M)^{1+\epsilon}$$

In words: This theorem extends the previous result to the case when B is allowed to vanish (ex: hard spheres, $B = |v - v_*|$), at the price of requiring high-order <u>regularity</u> (Sobolev norms at any order). This result can be rewritten as

$$D(f) \geq K_\epsilon(f) H(f \mid M)^{1+\epsilon},$$

where $K_\epsilon(f)$ only depends on an "upper bound" on the Sobolev-regularity for some large enough exponent, an upper bound on moments of f of order high enough and a lower bound on f. (In short K_ϵ depends on the <u>smoothness</u> - <u>decay</u> - <u>strict positivity</u> of f.)

<u>Comment 1</u>: Putting together Theorems 2, 3 and 4 we see that EEP inequalities for the Boltzmann equation are very tricky: then is no linear EEP inequality (with $\Theta(\alpha) = K\alpha$), but there are EEP inequalities as soon as one allows the power to be > 1 ($\Theta(\alpha) = K\alpha^{1+\epsilon}$). Probably there are some refinements like $K\alpha \log \dfrac{1}{\alpha}$, but this is not so immediate. Further see Section 4.3 below.

<u>Comment 2</u>: At the level of the spatially homogeneous Boltzmann equation,

$$\left\{ \begin{array}{l} \text{functional inequality} \\ D(f) \geq K H(f \mid M)^{1+\epsilon} \end{array} \right\} \quad \overset{\text{a priori}}{\underset{\text{estimates}}{\Longrightarrow}} \quad \left\{ \begin{array}{l} \text{estimate of rate of convergence} \\ H(f_t \mid M) = O(t^{-\infty}) \end{array} \right\}$$

Here $O(t^{-\infty})$ means $O(t^{-K})$, for any $K > 0$. I shall come back to this in the next chapter.

<u>Comment 3</u>: As explained in the next chapter, too, the stronger assumptions in Theorem 4 (Sobolev regularity) will not necessarily result in stronger assumptions at the level of the trend of equilibrium problem.

<u>Comment 4</u>: All constants K_ϵ appearing in this theorem are <u>explicit</u>— but a bit complicated! See Villani (2003) for more details.

In all the sequel of this chapter, I will try to describe the proofs of Theorems 3 and 4 in an intelligible way.

<u>Principle of the proof of Theorem 3</u>: It is an adaptation to this Boltzmann context of the celebrated (among those who know it!) proof of the Shannon–Stam inequality by Stam himself, together with some of the ingredients used in Boltzmann's proof of Theorem 1 (ii).

In short, the main idea is to introduce an <u>auxiliary diffusion semigroup</u> which is compatible with the important symmetries of the problem, and try to show that the <u>dissipation</u> of our functional $D(f)$, along this semigroup, is bounded below in terms of how far f is from equilibrium. Once we have this information, we can integrate along the semigroup, to recover some information on the value of $D(f)$ itself.

To illustrate this strategy, I shall first recall how the argument goes for the Shannon–Stam inequality. Stam (building on an idea by de Bruijn) used the <u>heat semigroup</u> $(P_t)_{t \geq 0}$ to link the Shannon–Stam inequality with the Blachman–Stam inequality thanks to

1) the identity $-\dfrac{d}{dt} H(P_t f) = I(P_t f)$

2) the nice properties of the heat semigroup with respect to the rescaled covolution, i.e. the operation which to f and g (laws of X and Y respectively) associate $f_\lambda * g_{1-\lambda}$, the law of $\sqrt{\lambda}\, X + \sqrt{1-\lambda}\, Y$:

$$\left[f_\lambda(x) = \frac{1}{\lambda^{N/2}}\, f\left(\frac{x}{\sqrt{\lambda}}\right) \right].$$

The advantage of this strategy is that I is much simpler than H in several respects, in particular because it involves quadratic quantities. This makes it rather simple to prove the Blachman–Stam inequality by some Cauchy–Schwarz inequalities, then use this semigroup argment to deduce from this the Shannon–Stam inequality.

Remark: This principle is also underlying the Γ_2 calculus of Bakry and Émery, or many semigroup arguments arising in the theory of logarithmic Sobolev inequalities. Semigroup arguments were once very popular in certain circles dealing with probability theory, geometry and convex bodies, then were superseded by other methods, and currently seem to regain vitality again.

Here is how the Shannon–Stam inequality can be proven elegantly, thanks to a variation of this principle (Carlen & Soffer, 1991). First note that it is equivalent to

$$H(f_\lambda * g_{1-\lambda} \mid M) \leq \lambda H(f \mid M) + (1 - \lambda)H(g \mid M)$$

Next introduce the <u>Fokker–Planck</u> (or adjoint Ornstein-Uhlenbeck) semigroup $(S_t)_{t\geq 0}$, defined by the solution of $\partial_t f = \Delta f + \nabla \cdot (fv)$.(Here v is the variable in $\mathbb{R}^N$.) This semigroup has the advantage over (P_t), to possess a steady state. Then one shows that

$$-\frac{d}{dt}H(S_t f \mid M) = I(S_t f \mid M)$$

$$-\frac{d}{dt}H\Big((S_t f)_\lambda * (S_t g_{1-\lambda}) \mid M\Big) = I\Big((S_t f)_\lambda * (S_t g)_{1-\lambda} \mid M\Big),$$

and therefore one has the representation formula

$$H\Big((S_t f)_g * (S_t g)_{1-\lambda} \mid M\Big) \qquad = \qquad \int_0^{+\infty} I\Big((S_t f)_\lambda * (S_t g)_{1-\lambda} \mid M\Big)dt$$

$$\uparrow$$

$$\text{because} \quad H(S_t h \mid M) \xrightarrow[t\to 0]{} \quad 0$$

By Blachman–Stam inequality, this is

$$\leq \quad \lambda \int_0^{+\infty} I(S_t f \mid M)dt + (1 - \lambda) \int_0^{+\infty} I(S_t g \mid M)dt$$
$$= \quad \lambda H(f \mid M) + (1 - \lambda)H(g \mid M).$$

In the case of the Boltzmann entropy production problem, this strategy will turn out to work also (which is somewhat unexpected), but will lead to much more complications (as we can expect).

The auxiliary semigroup will still be the Fokker–Planck semigroup; in fact it is natural to use a Gaussian semigroup in view of the symmetries of the Boltzmann operator. So essentially we shall try to estimate the dissipation, by the Fokker–Planck semigroup, of the entropy production (which itself is computed along the Boltzmann equation). This is not very natural from the physical point of view — all the more that we shall not apply the Fokker–Planck semigroup directly to the entropy production functional, but rather to a suitable lower bound. And the incredible thing is that, essentially, the object which will appear rather naturally in this process is another interesting object in kinetic theory: the Landau entropy production functional,

$$D_L(f) = \int_{\mathbb{R}^{2N}} \Psi(|v - v_*|) f f_* \left| \Pi_{(v-v_*)^\perp} \left[\nabla(\log f) - \nabla_*(\log \ f)_* \right] \right|^2 dv \, dv_*.$$

Thus, in some sense the Landau entropy production functional appears as a kind of infinitesimal variation of the Boltzmann entropy production functional, along the Fokker–Planck semigroup.

<u>Remark</u>: The Landau equation reads

$$\begin{aligned}
\frac{\partial f}{\partial t} + v \cdot \nabla_x f &= Q_L(f, f) \\
&= \nabla_v \cdot \left(\int_{\mathbb{R}^N} a(v - v_*)[f_* \nabla f - f(\nabla f)_*] \, dv_* \right)
\end{aligned}$$

where $a(v - v_*) = \Psi(|v - v_*|) \, \Pi_{(v-v_*)^\perp}$, and $\Pi_{z^\perp}$ is the orthogonal projection (in $\mathbb{R}^N$) onto the orthogonal hyperplane to z). This equation appears in the physics of charged (plasma) particles, only in the particular case $\Psi(|z|) = K/|z|$ ($K > 0$). From the mathematical point of view, it can be studied for various choices of Ψ, and for our purpose it is $\Psi(z) = |z|^2$ which shows up. This equation is related to the Boltzmann equation by the asymptotics of grazing collisions: for references and rigorous results see, e.g., Villani (2002b), Alexandre and Villani (2004).

<u>Sketch of proof of Theorem 3</u>: The goal is to estimate from <u>below</u>

$$D(f) \qquad = \qquad \frac{1}{4} \int \left(f' f'_* - f f_* \right) \log \frac{f' f'_*}{f f_*} B(v - v_*, \sigma) \, d\sigma \, dv \, dv_*$$

$$\geq \frac{K_B}{4} \int \left(f'f'_* - ff_* \right) \log \frac{f'f'_*}{ff_*} (1 + |v - v_*|)^{-\beta} \, d\sigma \, dv \, dv_*$$

by assumption

where $\beta \geq 0$ ($\beta = 0$ would be enough here, but would not simplify things very much). I shall divide the argument into eight lemmas, from which the whole proof can be reconstructed fairly easily. Some of them are very simple and some other are quite tricky.

Lemma 1 *(Boltzmann's angular integration trick)*:

$$D(f) \geq \frac{K_B |S^{N-1}|}{4} \overline{D}(f), \quad \text{where}$$

$$\overline{D}(f) \equiv \int (ff_* - \fint f'f'_* d\sigma) \log \frac{ff_*}{\fint f'f'_* d\sigma} (1 + |v - v_*|)^{-\beta} dv dv_*$$

and I uses the notation $\fint \cdot \, d\sigma = \dfrac{1}{|S|^{N-1}} \int_{S^{N-1}} \cdot \, d\sigma$ *(average).*

This lemma is a direct consequence of Jensen's inequality and the joint convexity of the function $(X - Y) \log \dfrac{X}{Y}$ on $\mathbb{R}_+ \times \mathbb{R}_+$.

Let me recall that $\fint f'f'_* \, d\sigma$ is a function of $\left\{ \begin{array}{c} m = v + v_* \\ e = \dfrac{|v|^2 + |v_*|^2}{2} \end{array} \right\}$ only.

Let us write $\left\{ \begin{array}{ll} F(v, v_*) & = f(v)f(v_*) \\ G(v, v_*) & = \fint f'f'_* d\sigma = G(m, e). \end{array} \right.$

Lemma 2 *(The semigroup $(S_t)_{t \geq 0}$ is compatible with the basic symmetries)*

$$\forall t \geq 0 \quad \left\{ \begin{array}{l} S_t F(v, v_*) = S_t(ff_*) = (S_t f)(S_t f)_* \\ S_t G(v, v_*) \quad \text{is still a function of } m \text{ and } e \text{ only.} \end{array} \right.$$

Also $\displaystyle\int S_t f \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}.$

Notation: I have used the same notation (S_t) for the FP semigroup acting on $L^1(\mathbb{R}^N)$ and acting on $L^1(\mathbb{R}^{2N}) = L^1(\mathbb{R}_v^N \times \mathbb{R}_{v_*}^N)$. This lemma is a consequence of the fact that the FP semigroup acts as the rescaled convolution with M.

Remark: In fact there is a stronger identity:

$$S_t\Big(\fint_{S^{N-1}} f' f'_* \, d\sigma\Big) = \fint_{S^{N-1}} (S_t f)'(S_t f)'_* \, d\sigma$$

Notation: In the sequel, $X = (v, v_*) \in \mathbb{R}^{2N}$.

Lemma 3: *(nonlinear commutator formula). Let $(S_t)_{t \geq 0}$ be any semigroup with generator $L = \Delta + a(X) \cdot \nabla + b(X)$, acting on $L^1(\mathbb{R}^d_X)$. Then*

$$\frac{d}{dt}\Big|_{t=0}\Big[S_t\Big((F-G)\log\frac{F}{G}\Big) - (S_t F - S_t G)\log\frac{S_t F}{S_t G}\Big] = \Big|\frac{\nabla F}{F} - \frac{\nabla G}{G}\Big|^2 (F+G).$$

The proof is by direct computation: for instance the object to compute can be written as $L\Big[(F-G)\log\frac{F}{G}\Big] - (LF - LG)\log\frac{F}{G} - (F-G)\big(\frac{LF}{F} - \frac{LG}{G}\big)$. The terms in $a \cdot \nabla$ disappear because they are differentiation operators, the terms in b disappear because $(F-G)\log\frac{F}{G}$ is homogeneous of degree 1.

Remark: This is a kind of "tensorial analogue" of the usual formula

$$\Delta\Phi(f) - \Phi'(f)\Delta f = \Phi''(f)|\nabla f|^2,$$

a "Γ_1 formula" for the Laplace operator.

The following lemma is crucial. Note that t below has nothing to do with the time in the Boltzmann equation!

Lemma 4: *(representation formula for $\overline{D}(f)$): Let ψ be any function of X, F and G be any probability distributions on $\mathbb{R}^d_X$, then*

$$\int \psi(X)(F-G)\log\frac{F}{G}\, dX = \int_0^{+\infty}\Big[\int \psi(X)\Big|\frac{\nabla S_t F}{S_t F} - \frac{\nabla S_t G}{S_t G}\Big|^2 (S_t F + S_t G)\, dX\Big]dt$$

"main" contribution

$$- \int_0^{+\infty}\Big[\int L^*\psi(X)(S_t F - S_t G)\log\frac{S_t F}{S_t G}\, dX\Big]dt$$

"error"

Here $(S_t)_{t \geq 0}$ is again the Fokker–Planck semigroup on $\mathbb{R}^d_X$. This formula holds true under the assumption that F, G are integrable enough (enough

moments), and that ψ does not increase faster than polynomially in X. Formally, this is an immediate consequence of Lemma 3, integration against ψ, and the identity

$$\int \psi(X)(F-G)\log\frac{F}{G}\,dX = \int_0^{+\infty}\int\left[\frac{d}{dt}(S_tF - S_tG)\log\frac{S_tF}{S_tG}\right]\psi(X)\,dX\,dt$$

which itself comes from $(S_tF - S_tG)\log\dfrac{S_tF}{S_tG} \xrightarrow[t\to\infty]{} 0$ since $S_tF \xrightarrow[t\to\infty]{} M$. The technical point in establishing the formula rigorously consist in controlling the convergence as $t \to \infty$, and the continuity as $t \to 0$, of the map $t \mapsto \int (S_tF - S_tG)\log\dfrac{S_tF}{S_tG}\psi\,dX$.

Warning: ∇ here is with respect to $X = (v, v_*)$ in our application.

Lemma 5 (symmetries): *Let P be the linear operator on $\mathbb{R}^{2N}$, depending on X, defined by*

$$P : [A, B] \mapsto \Pi_{(v-v_*)^\perp}\,[A - B]$$

(in fact $P \in L^\infty(\mathbb{R}_v^N \times \mathbb{R}_{v^*}^N; \mathcal{L}(\mathbb{R}^{2N}, \mathbb{R}^N)))$ ($\mathcal{L}$ = the set of linear mappings). Then, as soon as G only depends on $m = v + v_*$, $e = \dfrac{|v|^2 + |v_*|^2}{2}$, one has

$$\left|\frac{\nabla F}{F} - \frac{\nabla G}{G}\right|^2 \geq \frac{1}{\|P\|^2}\left|P\left(\frac{\nabla F}{F}\right)\right|^2$$

$$= \frac{1}{4}\left|\frac{P\nabla F}{F}\right|^2.$$

Corollary:

$$\left|\frac{\nabla S_tF}{S_tF} - \frac{\nabla S_tG}{S_tG}\right|^2 S_tF \geq \frac{1}{4}\left|\Pi_{(v-v_*)^\perp}\left[\frac{\nabla S_tf}{f} - \left(\frac{\nabla S_tf}{f}\right)_*\right]\right|$$

The proof is elementary linear algebra, plus the key property that

$$\frac{\nabla_X F}{F} = \left[\frac{\nabla f}{f}, \left(\frac{\nabla f}{f}\right)_*\right] \quad \text{if} \quad F = ff_*$$

$$\frac{\nabla_X G}{G} = \left[\frac{\nabla_m G}{G} + v\frac{\partial_e G}{\partial e}, \ \frac{\nabla_m G}{G} + v_*\frac{\partial G}{\partial e}\right]$$

36

together with the definition of P.

Remark: It is worthwile noting that, at this particular point, we are able to implement the symmetry part of the proof of Theorem 1 (ii), without having had to differentiate f — because the semigroup has done it for us! So what we have introduced here is the bounded operator P, instead of an unbounded operator of differentiation.

Lemma 6 *(A strong form of EEP inequality for the Landau equation):*

$$\text{Let } f(v) \geq 0 \text{ satisfy } \int_{\mathbb{R}^N} f(v) \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}.$$

$$\text{Let } T_f = \inf_{e \in S^{N-1}} \int f(v)(v \cdot e)^2 dv \ (\text{``minimum directional temperature''})$$

$$\text{and } D_L(f) = \frac{1}{2} \int_{\mathbb{R}^{2N}} \Psi(|v - v_*|) f f_* \left| \Pi_{(v-v_*)^\perp} \left[\nabla(\log \ f) - \nabla_*(\log f)_* \right] \right|^2 dv \, dv_*.$$

$$\text{Then } \Psi(|v - v_*|) \leq K_\Psi |v - v_*|^2 \implies D_L(f) \geq K_\Psi \left(\frac{N-1}{2} \right) T_f I(f \mid M).$$

Remark 1: The assumption on Ψ is rather unrealistic in the case of the Landau equation (but used in numerical simulations and some theoretical studies). As we already mentioned, the physical Ψ is $1/|z|$.

Remark 2: If we apply the logarithmic Sobolev inequality we conclude that $D_L(f) \geq \left[K_\Psi(N-1)T_f \right] H(f \mid M)$ which is an EEP for the Landau equation in this situation — notice that this one is "linear"!

This lemma is one of the key estimates in Desvillettes & Villani (2000). To prove it, one first reduces to the case when $\int f(v) v_i v_j = T_i \delta_{ij}$ by a change of orthonormal basis, then expands everything, taking advantage of the quadratic nature of things. One then applies the two inequalities

$$\sum_{ij} \int \left(|v|^2 \delta_{ij} - v_i v_j \right) \frac{\partial_i f \partial_j f}{f} \ dv \ \geq \ 0$$

$$\sum_i \alpha_i \int \frac{(\partial_i f)^2}{f} dv \ \geq \ \sum_i \frac{\alpha_i}{T_i} \quad (\text{a variant of Heisenberg's inequality})$$

Remark 3: The estimate $T_f > 0$ means that all directions are represented

37

int he support of f, which is crucial in the proof of

$$\frac{\nabla f}{f} - \left(\frac{\nabla f}{f}\right)_* \,/\!/\, v - v_* \Longrightarrow f \text{ is Gaussian.}$$

<u>Lemma 7</u> (Going from I back to H by the FP semigroup):

$$\int_0^{+\infty} I(S_t f \mid M)\, dt \;=\; H(f \mid M).$$

This lemma is well-known.

<u>Lemma 8</u> $((S_t)_{t\geq 0}$ is well adapted to the above-mentioned inequalities): *If* $f \in \mathcal{C}(1,0,1)$ *(normalization of mass, momentum, energy), and satisfies assumptions of boundedness of all moments, some L^p estimates and some lower bound estimates, then*

(1) $$T_{S_t f} \geq T_f$$

(2)
$$\|S_t F\|_{L_s^1} \;\leq\; C_s(\|F\|_{L_s^1} + 1)$$
$$\|S_t F\|_{L_s^2 \log L} \;\leq\; C_s(\|F\|_{L_s^1 \log L} + \|F\|_{L_{s+1}^1} + 1)$$
$$\qquad\qquad\qquad \text{[obtained by ODE techniques and use of LSI]}$$
$$H(S_t F \mid M) \;\xrightarrow[t\to\infty]{}\; 0 \text{ exponentially fast(LSI)}$$

(3)
$$\forall R \geq 1, \ \forall s$$

$$\int_{|X|\geq R} |X|^2 (F-G) \log \frac{F}{G}\, dX \leq \frac{C_s(f)}{R^s}$$

$$\int_0^{+\infty} \left[\int_{|X|\geq R} |X|^2 (S_t F - S_t G) \log \frac{S_t F}{S_t G}\, dX\right] dt \leq \frac{C_s(f)}{R^s}$$

This step is just technical: a lot of bounds by use of Cauchy–Schwarz, Chebyshev, elementary inequalities, etc.

Finally, one ties up lemmas 1–8 to get the desired estimate (Theorem 3). A conflict arises in the implementation, between what suggests Lemma 4

weight with inverse polynomial decay in |z| (cuts large relative velocities in the Boltzmann EP)

apply the Landau entropy inequality with this Ψ

square–like function

small constant

R (large)

|z| variable

square–like function
|z|²
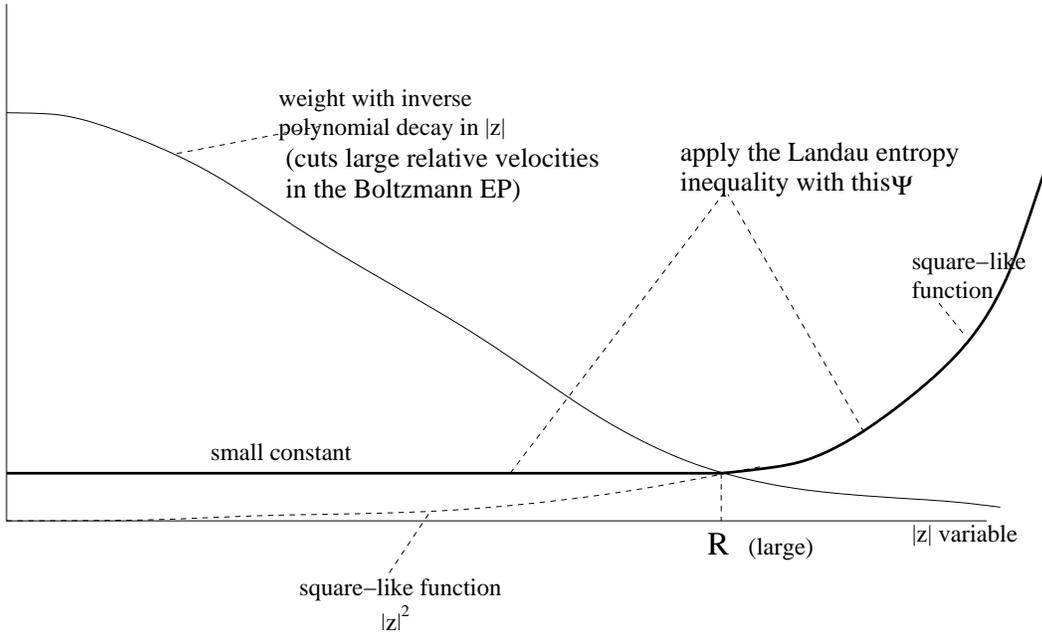
(choose $\psi(X)$ close to a constant, so that $L^*\psi = \Delta\psi - X \cdot \nabla\psi$ is small, and the "error" is indeed small), and what suggests Lemma 6 (choose $\psi(X)$ close to $\geq |v - v_*|^2$ to apply the estimate). This conflict is resolved schematically as in the picture.

In the argument, this leads to an error term for velocities $\geq R$, where ψ is not dominated by $(1 + |v - v_*|)^{-\beta}$. This error term is however very small, as a function of R, in view of Lemma 8. This small error term is the reason why, in the end, we do not recover a bound below by $H(F \mid M)$, but only by $[H(f \mid M)$-small error] (small term which can be chosen to be like $H(f \mid M)^{1+\epsilon}$ for arbitrarily small ϵ), due to the fact that s in (3) of Lemma 8 is as large as one whishes — so error terms are $O(R^{-\infty})$.

All of this is implemented carefully in Toscani & Villani (1999b). A technically simpler proof has been later devised in Villani (2003): in that variant, the error terms do not involve any integration with respect to the semigroup time variable, which makes things less intricate.

Once Theorem 3 is proven, Theorem 4 follows by a very tricky "interpo-

lation" estimate: $\forall \epsilon > 0, \ \delta \leq 1,$

$$\int_{|v-v_*|\leq\delta} (f'f'_* - ff_*) \log \frac{f'f'_*}{ff_*} \, dv \, dv_* \, d\sigma \leq C_\epsilon(f) H(f \mid M)^{1-\epsilon} \delta^{1/4};$$

C_ϵ depends on moments, Sobolev norms and lower bound on f. The difficulty here is to show that the left-hand side vanishes as $\delta \to 0$ (this is easy), but at the same time as $f \to M$, and with an order very close to 2 (thing of $H(f \mid M)$ as a square norm). Il would not be sufficient to have just $O(H(f \mid M)^\alpha \delta^{1/4})$ for some $\alpha < 1$.

The bounds on the Sobolev norms of f enter the proof somewhere via the interpolation inequality

$$\|u\|_{L^2} \leq C_N \|u\|_{L^1}^{1-\eta} \|u\|_{H^{\frac{N+1}{\eta}}}^\eta.$$

I won't say more about the argument because it is "purely technical": see Villani (2003).

4.3. Proof of Cercignani's conjecture in a special case

In fact, with the tools presented before, it is possible to prove the following theorem which bridges a gap (or rather fills a hole) between Theorems 2 and 3:

Theorem 3-ϵ: *Assume* $B(v - v_*, \sigma) \geq K_B(1 + |v - v_*|^2)$, *and let* f *be a probability density on* $\mathbb{R}^N$ *with* $\int f \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}$. *Then,*

$$D(f) \geq \left(K_B \frac{|S^{N-1}|}{8} \frac{N-1}{1+2N} \right) T_f H(f \mid M)$$

where $T_f = \inf_{e \in S^{N-1}} \int_{\mathbb{R}^N} f(v)(v \cdot e)^2 dv.$

Remark 1: This is the first case in which Cercignani's conjecture holds true! (It was discovered during the process of preparing this course, in an attempt to answer some questions by T. Bodineau.)

Remark 2: Besides its own interest, this result is the starting point for the simplified proof of Theorem 2 which was mentioned above and which is written in Villani (2003).

Sketch of proof: As before, let us introduce the Fokker–Planck regularization semigroup and start with

$$D(f) \geq \frac{|S^{N-1}|}{4}\overline{D}(f) \geq \frac{K_B|S^{N-1}|}{4}\overline{D}(f) \quad (\overline{D} \text{ with kernel } (1 + |v - v_*|^2))$$

$$-\frac{d}{dt}\overline{D}(S_t f) = \int_{\mathbb{R}^{2N}} \psi(X) \left| \frac{\nabla_x S_t f}{S_t F} - \frac{\nabla_x S_t G}{S_t G} \right|^2 (S_t F + S_t G) \, dX$$
$$- \int_{\mathbb{R}^{2N}} (L^* \psi)(S_t F - S_t G) \log \frac{S_t F}{S_t G} dx$$

$$F = f f_*, \quad G = \textstyle\fint f' f'_* d\sigma$$

$$X = (v, v_*), \quad \psi(X) = 1 + |v - v_*|^2, \quad L^* = \Delta_X - X \cdot \nabla_X.$$

$$L^* \psi = 4N - 2|v - v_*|^2 \leq 4N\psi.$$

Recall that

$$\int |v - v_*|^2 \left| \frac{\nabla_X S_t F}{S_t F} - \frac{\nabla_X S_t G}{S_t G} \right|^2 S_t F \, dX \geq \left(\frac{N-1}{2}\right) T_{S_t f} I(S_t f \mid M)$$
$$\geq \frac{N-1}{2} T_f I(S_t f \mid M)$$

$$-\frac{d}{dt}\overline{D}(S_t f) \geq \frac{N-1}{2} T_f I(S_t f \mid M) - 4N\overline{D}(S_t f)$$

From the Blachman–Stam inequality [using $S_t f = M_{1-e^{2t}} * f_{e^{2t}}$ with the notation $g_\lambda = \frac{1}{\lambda^{N/2}} g(\frac{\cdot}{\sqrt{\lambda}})$] or elementary Γ_2 theory,

$$I(S_\tau g \mid M) \leq e^{2\tau} I(g \mid M).$$

In particular, $I(S_{t+2Nt} f \mid M) \leq e^{-4Nt} I(S_t f \mid M)$. Thus

$$-\frac{d}{dt}\left[e^{-4Nt}\overline{D}(S_t f) \right] \geq \frac{N-1}{2} T_f \, I(S_{(1+2N)t} f \mid M)$$

41

Now take $\int_0^{+\infty} dt$, to find

$$\overline{D}(f) = \overline{D}(S_0 f) \;\geq\; \frac{N-1}{2} T_f \int_0^{+\infty} I(S_{(1+2N)t} f \mid M) dt$$

$$= \frac{N-1}{2} T_f \int_0^{+\infty} I(S_t f \mid M) \frac{dt}{1+2N} \quad \text{[change of var.'s]}$$

$$= \frac{N-1}{2(1+2N)} T_f \, H(f \mid M). \; \square$$

5. The state of the art for the spatially homogeneous Boltzmann equation.

In this chapter I shall describe how the results of the last chapter can be used to study the trend to equilibrium for the BE, under the very restrictive assumption that f_t does not depend on x. To this end, I shall appeal to many results from the theory of the Cauchy problem for the Boltzmann equation. Explaining precisely how one obtains these results would be a complete course on its own, so I shall not go into this.

5.1. Preparations

The theory of the Cauchy problem for the spatially homogeneous BE is complicated and many situations have to be studied separately. More precisely, the theory naturally splits <u>twice</u> into two branches:

- depending on the <u>kinetic</u> behavior of the collision kernel $B(v - v_*, \sigma)$ as $|v - v_*| \to 0$ [kinetic here means: as a function of $|v - v_*|$]: one says that the context is of

  - hard potential type if $B \sim |v - v_*|^\gamma \quad \gamma > 0$

  - soft potential type if $B \sim |v - v_*|^\gamma \quad \gamma < 0$

  - Maxwellian type if B does not depend on $|v - v_*|$, but only on
  $$\cos\theta = < \frac{v - v_*}{|v - v_*|}, \sigma > .$$

  The behavior of B as $|v - v_*| \to \infty$ has a decisive impact on the <u>tail behavior</u> of solutions: for hard potentials the solutions have a tendency to decrease very fast at infinity, while for soft potential this is not necessarily the case.

- depending on the <u>angular</u> behavior of the collision kernel as $\theta \to 0$ [again, $\theta = $ <u>deviation angle</u>]: one speaks of

  - cutoff kernel if $\displaystyle\int B(v - v_*, \sigma) d\sigma < +\infty$

  - non-cutoff kernel if $\displaystyle\int B(v - v_*, \sigma) d\sigma = +\infty$ (for almost all $|v - v_*|$).

Non-cutoff kernels naturally arise in the modelling of <u>long-range interactions</u> (even interactions which would decay extremely fast at large distances). As a

general rule, from the stochastic point of view, the cutoff Boltzmann equation represents a jump process and is associated with propagation of regularity and of singularity; while the non-cutoff one is intermediate between a jump and a diffusion process (I refrain from using the words "fractional diffusion", because these terms are commonly used in a completely different meaning), and is associated with regularizing effects. To quantify these statements, more precise assumptions have to be imposed on the kernel.

There may be a third dichotomy in the theory, depending on the behavior of the kernel at small relative velocities (singular or not). Nobody is sure...

The theory of cutoff Boltzmann equation is by now rather well-understood; the theory of non-cutoff BE is still under construction (Alexandre, Desvillettes, Fournier, Méléard, Villani...). In these lectures, for simplicity I only consider the cutoff case, in the model situations

$$\begin{aligned} B(v - v_*, \sigma) &= |v - v_*|^\gamma \quad (\gamma > 0) \\ B(v - v_*, \sigma) &= (1 + |v - v_*|)^{-\beta} \quad (\beta \geq 0) \end{aligned}$$

The most important case of application is certainly that of hard spheres: in dimension 3, $B(v - v_*, \sigma) = |v - v_*|$ (up to a constant).

5.2. Current state of regularity theory

By gathering many results from different authors, one obtains

Theorem 5: *Consider the kernel $B = |v - v_*|^\gamma$, $0 < \gamma \leq 1$, and let f be a probability density on $\mathbb{R}^N$, $\displaystyle\int f(v) \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}$. Then*

(i) *[Mischler & Wennberg (1999)] There exists a unique solution of the SHBE $\dfrac{\partial f}{\partial t} = Q(f, f)$ (in the class of solutions whose kinetic energy is nonincreasing — this restriction is important!).*

Moreover $\forall t$ $\displaystyle\int f(v) \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}$.

Remark: *Carleman (1932) obtained existence and uniqueness under*

the assumptions f_0 continuous, radially symmetric, $f_0(v) = O\left(\dfrac{1}{1 + |v|^6}\right)$, $\gamma = 1$.

(ii) [Desvillettes (1993); Wennberg (1994a, 1996); Mischler & Wennberg (1999)]

$$\forall s \geq 2 \quad \forall t_0 > 0 \quad \sup_{t \geq t_0} \left\| f_t \right\|_{L^1_s} < +\infty.$$

$\longrightarrow$ all moments are bounded, uniformly in $t \geq t_0$.

(iii) [Pulvirenti & Wennberg (1997)] $\forall t_0 > 0 \quad \exists \rho_0 > 0 \quad \exists A_0 > 0$;

$$t \geq t_0 \Longrightarrow \forall v \in \mathbb{R}^N \quad f_t(v) \geq \rho_0 e^{-A_0 |v|^2}.$$

$\longrightarrow$ the solution satisfies a Maxwellian lower bound, uniform in $t \geq t_0$.

Remark: Carleman (1932) obtained $\rho_0 e^{-A_0 |v|^{2+\epsilon}}$.

(iv) [Gustafsson (1986, 1988)] $f_0 \in L^1_{s_1} \cap L^p_s$ (s given, s_1 large enough)

$\Longrightarrow \sup_{t \geq 0} \left\| f_t \right\|_{L^p_s} < +\infty$ — not explicit.

[Mouhot & Villani (2004)] $f_0 \in L^1_3 \cap L^p$ ($1 < p < \infty$)

$\Longrightarrow \forall t_0 > 0 \quad \sup_{t \geq t_0} \left\| f_t \right\|_{L^p} < +\infty$ — explicit constants

[Arkeryd (1983)] same as Gustafsson with $p = \infty$ — explicit constants.

(v) [Mouhot & Villani (2004)] $f_0 \in L^1_3 \cap H^k \quad k \geq 0$

$\Longrightarrow \quad \forall t_0 > 0 \quad \sup_{t \geq t_0} \left\| f_t \right\|_{H^k} < +\infty.$

$\longrightarrow$ Propagation of H^k smoothness. Previous partial results had been obtained by Wennberg (1994b): H^1 smoothness in some cases, nonexplicit constants, $N = 3$; and Carlen, Gabetta and Toscani (1999): for Maxwellian collision kernels, with f_0 having all moments finite.

Remark: There is no regularization, so $f_0 \notin H^k_{loc} \Longrightarrow f_t \notin H^k_{loc}$.

However, one can prove the following:

(vi) [Mouhot & Villani (2004)] $f_0 \in L^1_3 \cap L^p$ $1 < p < \infty$ *then* $\forall k_0 \geq 0$
$\forall s_0 > 0$, $\exists \rho_0, A_0, \lambda_0 > 0$ s.t.

$$t \geq t_0 \Longrightarrow f(t,.) = \quad f^S(t,.) + \quad f^R(t,.)$$
$$\text{``smooth''} \quad \text{``remainder''}$$

with

$$\begin{cases} \sup_{t \geq t_0} \left\| f^S_t \right\|_{H^{k_0}} < +\infty, \quad \sup_{t \geq t_0} \left\| f^S_t \right\|_{L^1_{s_0}} < +\infty, \quad f^S_t(v) \geq \rho_0 e^{-A_0 |v|^2} \\ \\ \left\| f^R_t \right\|_{L^p} = O(e^{-\lambda_0 t}) \end{cases}$$

The proofs use:
- moment inequalities (Povzner inequalities...);
- mixing properties of Q^+ operator, similar to Lions' theorem: under some (very stringent) assumptions on B,

$$\left\| Q^+(g,f) \right\|_{H^{\frac{N-1}{2}}} \leq C \left\| g \right\|_{L^1} \left\| f \right\|_{L^2};$$

- Duhamel's formula (and its iterated versions),

$$f_t(v) = f_0(v) e^{\int_0^t L f_s(v) ds} + \int_0^t Q^+(f_s, f_s)(v) e^{-\int_s^t L f_\tau(v) d\tau} ds$$

$$L f = \left(\int B \, d\sigma \right) * f$$

- a lot of cutting into pieces;
- a lot of interpolation in weighted Lebesgue and Sobolev spaces.

5.3. Convergence to equilibrium

Here is an application of the previous regularity results, when combined with the entropy production estimates of the last chapter:

__Theorem 6__ (based on Mouhot, Villani) *Let* $f_0 \in L^1_3 \cap L^p$ $(1 < p < \infty)$,

$$\int f_0(v) \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}, \, B(v - v_*, \sigma) = |v - v_*|^\gamma, \, then$$

$$\left\| f_t - M \right\|_{L^1} = O(t^{-\infty}) \quad \text{with explicit constants}$$

46

(the statement means: $\forall \epsilon > 0, \exists C_\epsilon$ explicit s.t. $\left\| f_t - M \right\|_{L^1} \leq C_\epsilon t^{-1/\epsilon}$).

Sketch of proof: Use the decomposition $f_t = f_t^S + f_t^R$ (k_0 to be chosen later on, $t_0 = 1$).

Let t_1 be an indermediate time, to be chosen later; introduce

$$\begin{cases} \widetilde{f}_t & = \text{sol. of BE starting from } f_{t_1}^S \text{ at } t = t_1 \\ \widetilde{M} & = \text{Maxwellian distribution associated with } \widetilde{f}_{t_1} = f_{t_1}^S. \end{cases}$$

Since $\widetilde{f}_{t_1}$ is very smooth, so is $\widetilde{f}_t$, and in fact $\widetilde{f}_t$ is uniformly bounded in $L_{s_0}^1 \cap H^{k_0}$ for $t \geq t_1 + 1$, with a Maxwellian lower bound. This implies (Theorem 4)

$$D(\widetilde{f}_t) \geq K_\epsilon H(\widetilde{f}_t \mid \widetilde{M})^{1+\epsilon}$$

with ϵ as small as one wishes if k_0 is chosen large enough, depending on ϵ. As a consequence (Gronwall), $H(\widetilde{f}_t \mid \widetilde{M}) = O((t - t_1)^{-1/\epsilon})$ with explicit constants which do not depend on t_1 (they only depend on t_0 used in the decomposition $f^S + f^R$).

On the other hand, from the statement in Theorem 5 we have

$$\left\| f_t^R \right\|_{L^1} = O(e^{-\lambda_0 t}) \quad (t \geq 1)$$

$$\left\| M - \widetilde{M} \right\|_{L^1} = O(e^{-\lambda t_1}) \quad \text{for some } \lambda > 0$$

Next, by a stability result about solutions of the SHBE,

$$\left\| f_t - \widetilde{f}_t \right\|_{L^1} \leq C \, e^{A(t-t_1)} \left\| f_{t_s} - \widetilde{f}_{t_1} \right\|_{L_s^q}$$

[at most exponential divergence]

After a little bit of interpolation, one gets the bound

$$\begin{aligned} \left\| f_t - M \right\|_{L^1} &\leq \left\| f_t - \widetilde{f}_t \right\|_{L^1} + \left\| \widetilde{f}_t - \widetilde{M} \right\|_{L^1} + \left\| \widetilde{M} - M \right\|_{L^1} \\ &\leq C\left(e^{A(t-t_1)} e^{-\lambda t_1} + (t - t_1)^{-\frac{1}{\epsilon}} + e^{-\lambda t_1} \right) \end{aligned}$$

Now, if one chooses $t_1 \gg t - t_1$, the exponential divergence can be compensated for. More precisely, if $A(t - t_1) = \frac{\lambda}{2} t_1$, i.e. $t_1 = \left(\frac{A}{\frac{\lambda}{2} + A} \right) t$, then one gets $\left\| f_t - M \right\| \leq C(t_1) t^{-1/\epsilon}$, as announced. $\square$

Let me conclude this chapter with a result about "mollified soft potentials": $B(v - v_*, \sigma) = (1 + |v - v_*|^2)^{-\beta/2}$ ($\beta \geq 0$), or $B = (1 + |v - v_*|)^{-\beta}$, which gives the same bounds.

Theorem 7 (Toscani & Villani, 1999). *Assume* $B(v - v_*, \sigma) = (1 + |v - v_*|)^{-\beta}$, $0 \leq \beta < 2$. *Let* f_0 *be a probability density satisfying* $\int f_0 \begin{pmatrix} 1 \\ v \\ |v|^2 \end{pmatrix} dv = \begin{pmatrix} 1 \\ 0 \\ N \end{pmatrix}$, *with* $f_0 \in \cap_{s>0} L^1_s$, $f_0 \in L^p$ *for some* $p > 1$, $f_0 \geq \rho_0 e^{-A_0 |v|^2}$ *for some* $\rho_0 > 0$, $A_0 > 0$. *Then* $H(f_t \mid M) = O(t^{-\infty})$.

Remark 1: This theorem is based on Theorem 3 together with the propagation of bounds. These bounds are <u>not</u> uniform in time: we only know $\left\| f_t \right\|_{L^1_s} = O((1+t)^\epsilon)$, $\left\| f_t \right\|_{L^p} = O((1+t)^{1+\epsilon})$, and $f_t(v) \geq \rho_t e^{-A_0 |v|^2}$ for $\rho_t = e^{Ct} \rho_0$. It turns out that this deterioration of the bounds is <u>slow enough</u> to prove the result. This of course is possible only because we have a precise control of the bounds in the entropy production inequalities.

Remark 2: Up to very recently, this was the only known result for trend to equilibrium for collision kernels with that kind of decay — with the exception of Caflisch (1980) in a close-to-equilibrium setting. In particular, compactness arguments spectacularly fail in this context because of the non-uniformity in the moment bounds. Very recently, Guo was able to construct close-to-equilibrium solutions for which the moment bounds are uniform, and then both the compactness and the constructive methods can be applied.

6. A review of some closely related topics

6.1. Particles systems and Kac's problem

Assume you have an n-particle system governed by microscopic equations, which is well-described in some sense by a macroscopic system as $n \to \infty$. Typically, the macroscopic equations would be those satisfied by the one-particle marginal as $n \to \infty$. If one is interested in the large-time behavior of the particle system, it is natural to a) carefully investigate the large-time behavior of the macroscopic system, b) carefully investigate how precise the approximation is in the limit $n \to \infty$ (say, for instance, how close the microscopic system is to a chaotic system governed by the macroscopic equation).

However, in some situations it is possible to directly study the asymptotic behavior of the particle system as $t \to \infty$, uniformly in n. When this is the case, then the large-time behavior of the macroscopic equation can be deduced from that of the particle system, and one expects that the limit $n \to \infty$ and $t \to \infty$ are exchangeable.

When Kac first began to think about the rate of convergence for the BE, he thought that a direct approach of the large-time behavior of a particle system might succeed — and would be more satisfactory than the strategy described in the first paragraph. Since the problem for the full Boltzmann equation was much too complicated, he devised simpler models, in which only velocity was taken into account. They are often called "master equations". Below are two such models. They take the form of evolution equations for the density $f^{(n)}$ of n particles (joint density).

- the Boltzmann master equation: the phase space is

$$\{(v_1, \cdots, v_n) \in (\mathbb{R}^N)^n;\ \Sigma |v_i|^2 = 2nE;\ \Sigma v_i = nV\}$$
$$(E = \text{ mean kinetic energy};\ V = \text{ mean velocity})$$

and the equation is

$$\frac{\partial f^{(n)}}{\partial t} = n \sum_{ij}' \int_{S^{n-1}} d\sigma\, B(v_i - v_j, \sigma) \left[A_\sigma^{ij} f^{(n)} - f^{(n)} \right]$$

$$\left[\mathcal{A}_\sigma^{ij} f^{(n)} \right] (v_1, \cdots, v_n) = f^{(n)}(v_1, \cdots, v_i', \cdots, v_j', \cdots, v_n)$$

49

$$\begin{cases} v_i' = \dfrac{v_i + v_j}{2} + \dfrac{|v_i - v_j|}{2}\sigma \\ v_j' = \dfrac{v_i + v_j}{2} - \dfrac{|v_i - v_j|}{2}\sigma \end{cases}$$

where $\fint_{S^{N-1}} d\sigma = \dfrac{1}{|S^{N-1}|} \displaystyle\int d\sigma$, $\displaystyle\sum_{ij}{}' = \dfrac{1}{\binom{n}{2}} \displaystyle\sum_{i<j}$

[If the initial datum $f_0^{(n)}$ is chaotic.]

In the limit $n \to \infty$, it is expected that the 1-particle marginal $f^{(1)}$ of $f^{(n)}$ converges towards a solution of the Boltzmann equation

$$\frac{\partial f}{\partial t} = \fint_{S^{N-1}} d\sigma \int dv_* B(v - v_*, \sigma)(f'f_*' - ff_*).$$

Remark: In this model v_i', v_j' should be thought of as the velocities <u>after</u> collision, instead of <u>before</u> for the Boltzmann equation.

The limit $n \to \infty$ has been studied by various authors, including Kac himself, Sznitman, Méléard and collaborators.

- the Kac master equation: this is a simplified model of the above. The phase space is $\{(v_1, \cdots, v_n) \in \mathbb{R}^N; \Sigma v_i^2 = 2nE\} = \sqrt{2nE}\, S^{n-1}$, and the equation is

$$(K) \qquad \frac{\partial f^{(n)}}{\partial t} = n \sum_{ij}{}' \fint_0^{2\pi} \left[f_0^{(n)} \circ R_\theta^{ij} - f^{(n)} \right] d\theta$$

where $R_\theta^{ij}(v_1, \cdots, v_n) = (v_1, \cdots, v_i', \cdots, v_j', \cdots, v_m)$ and $(v_i', v_j') = R_\theta(v_i, v_j)$, where R_θ is the direct rotation with angle θ in the plan $\mathbb{R}^2$.

For notational simplicity let us set $E = \dfrac{1}{2}$, so the phase space is $\sqrt{n}S^{n-1}$.

Let us focus on Kac's master equation; the discussion would be similar for the Boltzmann master equation. If one rewrites (K) in the form

$$\frac{\partial f^{(n)}}{\partial t} = L_n f^{(n)},$$

50

then the operator L_n is a beautiful linear operator, symmetric in $L^2(d\sigma^n_{\sqrt{n}})$ where $\sigma^n_{\sqrt{n}}$ is the uniform probability measure on $\sqrt{n}S^{n-1}$. Therefore Kac thought is was possible to attack the problem of trend to equilibrium by the apparatus of linear operator theory. It is easy to see that for each n, L_n is a nonpositive operator, admitting 0 as a simple eigenvalue, associated to the eigenfunction 1 (conservation of mass). Moreover, 0 is an isolated eigenvalue, so one can define

$$\lambda_n = \text{ the gap between 0 and the largest nonzero eigenvalue.}$$

This naturally suggests

Kac's spectral gap problem: *Find a uniform lower bound on λ_n.*

Here uniform means of course uniform as $n \to \infty$. This problem stood open from 1956 to 1999, the date when it was solved by Janvresse, by the use of Yau's martingale method (see Janvresse (2001)). Very shortly after, the result was spectacularly improved by Carlen, Carvalho and Loss (2001):

Theorem 8 (Carlen, Carvalho, Loss): $\Lambda_n = \dfrac{n+2}{2(n-1)}$. *Moreover the corresponding eigenfunction is* $\displaystyle\sum_{j=1}^{n} v_j^4 - \dfrac{3}{n(n+2)}$.

A more precise description of the spectrum was actually obtained before, by a very different method, in a preprint by Maslen which remained unpublished for a long time. Recently, Carlen, Carvalho and Loss obtained a result rather similar to Theorem 8 for the Boltzmann master equation, even though less explicit (in the case where $B(v-v_*,\sigma)$ only depends on $< \dfrac{v-v_*}{|v-v_*|}, \sigma >$). Their method is very interesting and certainly deserves further study; it was applied recently by Caputo (2003) to solve a tricky spectral gap problem for spin systems.

However, once the result is proven, it is not clear at all what can be deduced from it for the nonlinear Boltzmann equation, or for its variant for

Kac's master equation, which is the Kac model

$$
\text{(Km)} \qquad
\begin{cases}
\dfrac{\partial f}{\partial t} = \displaystyle\int_{\mathbb{R}} \int_0^{2\pi} (f' f'_* - f f_*)\, d\theta\, dv_* \\[4mm]
(v', v'_*) = R_\theta(v, v_*).
\end{cases}
$$

The problem is that, for instance, one has

$$
\left\| f^{(n)} - \mathbb{1} \right\|_{L^2(d\sigma^n_{\sqrt{n}})} \;\leq\; e^{-\lambda t}
$$

for some $\lambda > 0$ independent of n. But in the limit going to the BE, the L^2 norm blows up very quickly, something like C^n for some constant C. So one cannot a priori conclude that some quantity, converging as $n \to \infty$ towards something, will decrease exponentially fast. Certainly the problem lies in the limit which is not adapted to L^2 norm — note that $L^2(d\sigma^n_{\sqrt{n}})$ transform into $L^2(M^{-1}dv)$ at the level of $f^{(1)}$. I personally think that the right thing to do here is to go to the <u>linearized</u> Boltzmann equation by a different scaling.

Remark: As a corollary of his study, Kac was led to the following strange conjecture, which I have never seen mentioned anywhere except in his text. Whenever f solves (Km), or maybe the BE with Maxwell molecules, and has zero mean velocity and unit temperature, then the following functional $K(f)$ would be nonincreasing with time:

$$
\begin{cases}
K(f) = \displaystyle\int e^{z_0(f^2)(1-v^2)} f^2(v)\, dv \\[4mm]
z_0(f^2) \text{ sol. of } \displaystyle\int (1 - v^2) e^{-z_0(f^2)v^2} f^2(v)\, dv = 0
\end{cases}
$$

To investigate rates of convergence towards the BE, a better idea is to look for the entropy production versions of Kac's problem.

<u>Entropy variant of Kac's spectral gap problem</u>: *Let μ_n be the best constant in the inequality*

$$
D(f^{(n)}) \geq \mu_n \Big[H(f^{(n)}) - H(1) \Big] = \mu_n H(f^{(n)})
$$

52

Is there a uniform lower bound on μ_n?

Here $D(f^{(n)}) = \sum_{ij}' \dfrac{n}{2} \displaystyle\int_{\sqrt{n}} d\theta \displaystyle\int_{S^{n-1}} \left(f^{(n)} \circ R_\theta^{ij} - f^{(n)} \right) \log \dfrac{f^{(n)} \circ R_\theta^{ij}}{f^{(n)}} d\sigma_{\sqrt{n}}^n(v);$ of course this is the entropy production functional associated to Kac's master equation.

If there is a uniform lower bound $\mu > 0$ for μ_n, then

$$H\left(f_t^{(n)} \right) \le H\left(f_0^{(n)} \right) e^{-\mu t}$$

If there is a strong enough chaos property, then $f_t^{(n)} d\sigma^n$ looks like $[f_t(v)dv]^{\otimes n}$, where f_t is a solution of the Boltzmann equation. A well-known lemma states that $d\sigma_{\sqrt{n}}^n$ roughly speaking, looks like $M^{\otimes n}$, where M is the standard Gaussian distribution on $\mathbb{R}$. Then

$$H\left(f_t^{(n)} \right) = \int f_t^{(n)} \log f_t^{(n)} d\sigma_{\sqrt{n}}^{(n)} \simeq \int \dfrac{f_t^{\otimes n}}{M^{\otimes n}} \log \dfrac{f_t^{\otimes n}}{M^{\otimes n}} M^{\otimes n} dv?$$

$$\simeq nH\left(f_t \mid M \right)$$

(Here I just encountered notational problems because $f_t^{(n)}$ is a density with respect to $d\sigma^n$ which is morally a Gaussian measure, while f_t is a density with respect to Lebesgue measure...)

Then, providing everything passes to the limit we would have

$$H\left(f_t \mid M \right) \le H\left(f_0 \mid M \right) e^{-\mu t},$$

as the limit of

$$\dfrac{H\left(f_t^{(n)} \right)}{n} \le \dfrac{H\left(f_0^{(n)} \right)}{n} e^{-\mu t}.$$

However, this would probably be asking too much: passing to the limit in the entropy-entropy production inequality for fixed n, we would recover an entropy-entropy production inequality of linear type for Kac's model, which would contradict the intuition furnished by the Bobylev–Cercignani counterexamples. Actually, Carlen, Carvalho and Loss recently checked that the optimal μ_n goes to 0 as $n \to \infty$. However, the picture is still not so clear because the Bobylev–Cercignani counterexamples exploit quite well the structure of the Boltzmann operator, and cannot be directly adapted to the

53

context of the Kac equation. In fact, after working on that question with Carlen, Carvalho and Loss I have come to suspect that linear entropy-entropy production inequalities may after all exist for the Kac equation under some moment conditions; there is something to understand here.

Anyway, here is now some positive result: it was shown in my Habilitation (2000), that $\mu_n^{-1} = O(n)$.

So one can ask:

- what is the exact order of μ_n?

- whether there are some initial data $f_0^{(n)}$ which provide a better rate than μ_n.

- whether there are related models with μ_n bounded below.

Here is an answer to the third question. The next theorem was proven for the sake of this course, then slightly generalized in Villani (2003).

<u>Theorem 9</u> *Consider the following modified Kac master equation:*

$$\frac{\partial f^{(n)}}{\partial t} = n \sum_{ij}' \int_0^{2\pi} (1 + v_i^2 + v_j^2)[f^{(n)} \circ R_\theta^{ij} - f^{(n)}]$$

[NOTE THAT v_i^2, v_j^2 are typically $O(1)$ if $v \in \sqrt{n}S^{n-1}$, so this equation has the same physical scaling as the previous one.]

Define $H(f^{(n)}) = \int_{\sqrt{n}S^{n-1}} f^{(n)} \log f^{(n)} d\sigma_{\sqrt{n}}^n \quad \left(\int f^{(n)} d\sigma^n = 1 \right)$

$$D(f^{(n)}) = \frac{n}{2} \int_{\sqrt{n}s^{n-1}} d\sigma_{\sqrt{n}}^n \sum_{ij}' \int_0^{2\pi} d\theta (1+v_i^2+v_j^2)[f^{(n)} \circ R_\theta^{ij} - f^{(n)}] \log \frac{f^{(n)} \circ R_\theta^{ij}}{f^{(n)}}$$

Then $D(f^{(n)}) \geq \frac{n}{5n - 1} H(f^{(n)}) \geq \frac{1}{4} H(f^{(n)})$.

Proof: To simplify notations let us rescale everything to work on the unit sphere S^{n-1}. Let σ^n be the uniform probability measure on S^{n-1}. The problem becomes: bounding below

$$D(f^{(n)}) = \frac{n}{2\binom{n}{2}} \sum_{i<j} \frac{1}{2\pi} \int_0^{2\pi} d\theta \int_{S^{n-1}} (1+nv_i^2+nv_j^2)[f^{(n)} \circ R_\theta^{ij} - f^{(n)}] \log \frac{f^{(n)} \circ R_\theta^{ij}}{f^{(n)}}$$

54

in terms of $H(f^{(n)}) = \int_{S^{n-1}} f^{(n)} \log f^{(n)} \, d\sigma^n$, whenever $\int f^{(n)} d\sigma^n = 1$.

The proof follows the same lines as Thm 3-ϵ. First, by Jensen,

$$D(f^{(n)}) \geq \overline{D}(f^{(n)}) \equiv \frac{n}{2\binom{n}{2}} \sum_{i<j} \int_{S^{n-1}} (1 + nv_i^2 + nv_j^2)(f - f^{ij}) \log \frac{f}{f^{ij}} \, d\sigma^n$$

where $f^{ij}(v) = \frac{1}{2\pi} \int_0^{2\pi} f(R_\theta^{ij} v) d\theta$.

Let Δ_S be the Laplace-Beltrami operator on the sphere S^{n-1}; by definition (this is one of the possible definitions),

$$\Delta_S f = \sum_{i<j} (D^{ij})^2 f,$$

where $D^{ij} f = v_j \partial_i f - v_i \partial_j f = \frac{d}{d\theta}\Big|_{\theta=0} f \circ R_\theta^{ij}$.

An easy computation shows that

$$\int_{S^{n-1}} f \Delta_S f = - \int_{S^{n-1}} \sum_{i<j} (D^{ij} f)^2$$

$$= - \int_{S^{n-1}} |v|^2 |\nabla f|^2 + \int_{S^{n-1}} (v \cdot \nabla f)^2 = - \int_{S^{n-1}} |\nabla f|^2$$

(tangential gradient or gradient on S^{n-1}).

Let $(S_t)_{t \geq 0}$ be a semigroup associated with the heat equation

$$\partial_t f = \Delta_S f.$$

This is the equivalent of the Ornstein–Uhlenbeck semigroup $\partial_t h = \Delta h - v \cdot \nabla h$. A computation similar to that in the proof of Theorem 3-ϵ yields

$$-\frac{d}{dt} \overline{D}(S_t f^{(n)}) =$$

$$\frac{n}{2\binom{n}{2}} \sum_{i<j} \int_{S^{n-1}} (1 + nv_i^2 + nv_j^2)(S_t f + S_t f^{ij}) |\nabla \log S_t f - \nabla \log S_t f^{ij}|^2 d\sigma^n$$

$$-\frac{n}{2\binom{n}{2}} \sum_{i<j} \int_{S^{n-1}} \Delta_S (1 + nv_i^2 + nv_j^2)(S_t f - S_t f^{ij}) \log \frac{S_t f}{S_t f^{ij}}$$

55

Here I used $(S_t f)^{ij} = S_t(f^{ij})$.

Next, $\Delta_S(1 + nv_i^2 + nv_j^2) = n[4 - 2n(v_i^2 + v_j^2)] \le 4n$

And Also, if P_{ij} denotes the orthogonal projection on the (i,j) plane,

$$(S_t f + S_t f^{ij}) \left| \nabla \log(S_t f) - \nabla \log(S_t f)^{ij} \right|^2$$

$$\ge S_t f \left| P_{ij} \nabla \log S_t f - P_{ij} \nabla \log(S_t f)^{ij} \right|^2$$

But $P_{ij} \nabla \log S_t f^{ij}$ is always colinear to $(v_i, v_j) \in$ plane (i,j).

If T is the operator with unit norm $(\mathbb{R}^2 \to \mathbb{R})$ $[x_i, x_j] \xrightarrow{T} \dfrac{x_i v_j - x_j v_i}{v_i^2 + v_j^2}$

then $TP_{ij} \nabla \log(S_t f)^{ij} = 0$ and we have

$$
\begin{aligned}
|P_{ij} \nabla \log S_t f - P_{ij} \nabla \log(S_t f)^{ij}|^2 &\ge |TP_{ij} \nabla \log S_t f|^2 \\
&= \frac{(D^{ij} \log S_t f)^2}{v_i^2 + v_j^2}.
\end{aligned}
$$

All in all,

$$
\begin{aligned}
-\frac{d}{dt}\overline{D}\left(S_t f^{(n)}\right) &\ge \frac{n}{2\binom{n}{2}} \sum_{i<j} \int_{S^{n-1}} n S_t f \left(D^{ij} \log S_t f\right)^2 d\sigma^n \\
&\quad - \frac{n}{2\binom{n}{2}} \sum_{i<j} \int_{S^{n-1}} 4n(S_t f - S_t f^{ij}) \log \frac{S_t f}{S_t f^{ij}} d\sigma \\
&= \frac{n^2}{2\binom{n}{2}} \int_{S^{n-1}} S_t f |\nabla \log S_t f|^2 d\sigma^n - 4n\overline{D}\left(S_t f^{(n)}\right)
\end{aligned}
$$

which is

$$
\begin{aligned}
-\frac{d}{dt}\overline{D}\left(S_t f^{(n)}\right) + 4n\overline{D}\left(S_t f^{(n)}\right) &\ge \frac{n}{n-1} \int_{S^{n-1}} S_t f |\nabla \log S_t f|^2 d\sigma^n \\
-\frac{d}{dt}\left[e^{-4nt}\overline{D}(S_t f^{(n)})\right] &\ge \frac{n}{n-1} e^{-4nt} \int_{S^{n-1}} S_t f |\nabla \log S_t f|^2 d\sigma^n \\
&= \frac{n}{n-1} e^{-4nt} I\left(S_t f^{(n)}\right)
\end{aligned}
$$

Here $I(f) = \displaystyle\int \frac{|\nabla f|^2}{f}\, d\sigma^n$. From the theory of logarithmic Sobolev inequalities, $I(S_t f) \le e^{-\rho t} I(f)$ with $\rho = \dfrac{n-1}{(\text{radius})^2} = n-1$. So

$$I\left(S_{t+\frac{4n}{n-1}t}\, f\right) \le e^{-4nt} I(S_t f),$$

and we have

$$-\frac{d}{dt}\left[e^{-4nt}\overline{D}(S_t f^{(n)})\right] \ge \frac{n}{n-1} I\left(S_{(1+\frac{4n}{n-1})t}\, f^{(n)}\right)$$

If one $\displaystyle\int_0^{+\infty} dt$, one finds

$$
\begin{aligned}
\overline{D}\left(f^{(n)}\right) \ge \frac{n}{n-1}\int_0^{+\infty} I\left(S_{(1+\frac{4n}{n-1})t}\, f^{(n)}\right) dt &= \left(\frac{n}{n-1}\right)\frac{n-1}{5n-1}\int_0^{+\infty} I(S_t f) \\
&= \frac{n}{5n-1} H\left(f^{(n)}\right). \;\square
\end{aligned}
$$

<u>Physical interpretation</u>: This model enhances the relaxation of <u>fast</u> particles, and this is precisely what the Boltzmann equation needs for the fast decrease of entropy.

6.2. The central limit theorem for Maxwellian molecules

In the particular case when $B(v - v_*, \sigma) = b(\cos\theta)$ [$\theta =$ deviation angle as usual], then there is a strong analogy between the Q^+ operator

$$Q^+(f, f) = \int dv_*\, d\sigma\; b(\cos\theta) f' f'_*$$

and the rescaled convolution operator

$$f_{1/2} * f_{1/2} = \frac{1}{2^{N/2}} f\left(\frac{\cdot}{\sqrt{2}}\right) * \frac{1}{2^{N/2}} f\left(\frac{\cdot}{\sqrt{2}}\right).$$

Thus, many tools which work (more or less classically) for the study of rate of convergence in the central limit theorem, also have a chance to work in the study of the Boltzmann equation.

McKean (1966) worked out the brilliant idea, apparently first suggested by Kac, to explicitly formulate the asymptotic behavior for the BE (spatially homogeneous with Maxwellian molecules!!) in a form similar to the CLT. The idea was that f_t could be written as a sum of terms taking into account all possible collisional scenarii. One term would take into account particles who have never collided, another one particles which have collided once with a particle having never collided before, etc. It is easy to have such representation by <u>iteration of Duhamel's formula</u>: starting from

$$\frac{\partial f}{\partial t} = Q^+(f, f) - f$$

(f is the simple form taken by the loss term, due to the particular collision kernel and the normalization),
one gets

$$f_t(v) = f_0(v)e^{-t} + \int_0^t e^{-(t-s)}Q^+(f_s, f_s)\, ds$$

then one can replace f_s in the integral by $f_0 e^{-s} + \int_0^s e^{-(s-s_2)}Q^+(f_{s_2}, f_{s_2})\, ds_2$. If one does this an infinite number of times, one arrives at the

<u>Wild sum representation</u>:

$$f_t = \sum_{n=1}^{\infty} e^{-t}(1 - e^{-t})^{n-1} \sum_{\gamma \in \Gamma(n)} \alpha(\gamma)\, Q_\gamma^+(f_0)$$

- n is the number of particles involved

- $e^{-t}(1 - e^{-t})^{n-1}$ are coefficients which go to 0 as $t \to \infty$, for given n

- γ is a tree with n nodes, describing the collision history for the n particles

- α is a combinatorial coefficient

- everything is expressed in terms of the initial datum f_0

Here $\Gamma(n)$ is the set of all binary trees with n leaves, $Q_\gamma^+(f_0)$ is defined by induction as follows: let $\gamma \in \Gamma(n)$, then γ defines two sub-trees (take the root out), γ_1 and γ_2. Then

$$Q_\gamma^+(f_0) = Q^+(Q_{\gamma_1}^+(f_0), Q_{\gamma_2}^+(f_0)),$$

with the definition $Q^+(g, f) = \int b(\cos\theta) f' g'_* \, dv_* \, d\sigma$. If $n = 0$ one agrees $Q_\gamma^+(f_0) = f_0$. The term f_0 corresponds to particles having undergone no collision (so they are still in their initial state). The term $Q^+(f_0, f_0)$ to particles having undergone one collision; etc. For instance,

$$Q_\gamma^+(f_0) = Q^+(Q^+(f_0, Q^+(f_0, f_0)), \; Q^+(f_0, f_0))$$

stands for particles having undergone, first a collision with another particle having undergone no collision, second a collision with a particle who had previously twice collided a particle having undergone no collision before (Exercise: draw the corresponding tree).

Finally, $\alpha(\gamma)$ is a combinatorial coefficient, recursively defined by

$$\alpha(\gamma) = \frac{\alpha(\gamma_1)\alpha(\gamma_2)}{n-1}; \qquad \alpha(\gamma) = 1 \text{ if } n = 1.$$

(γ_1, γ_2 are the subtrees of γ).

Intuitively, each time a particle undergoes a collision, this will take it closer to equilibrium. This is not really true: the collision partner should not be too far from equilibrium itself. The worst example is a tree in which each node has a terminal leaf; then $Q_\gamma^+(f_0) = Q^+(f_0, Q^+(f_0, Q^+(f_0, ...)))$, which certainly is not close to equilibrium... The "good" scenario is one in which all particles undergo many collisions, so that all branches in the tree are quite long. So the game suggested by McKean is to

1) show that if a tree is well-balanced (all particle undergo many collisions), then $Q_\gamma^+(f_0)$ is close to $M=$ equilibrium;

2) show that in the Wild sum representation, well-balanced trees will play the most important role as $t \to \infty$.

This looks quite delicate. Nevertheless, McKean (1966) was able to complete this program in the case of the Kac (N = 1) model (getting exponential convergence). The key estimate was

$$\sum \left\{ \alpha(\gamma); \sup_{a<b} \left| \int_a^b Q_\gamma^+(f_0) - \int_a^b M \right| > \delta \right\} \le c(\delta, f_0) n^{-\delta}, \qquad \delta = 1 - (8/3\pi).$$

In higher dimension, McKean failed to get similar estimates. The problem was recently solved in a lot of generality by Carlen, Carvalho and Gabetta (2000).

How can one quantify the fact that collisions have a tendency to drive the system to equilibrium? The key point is to find some inequality expressing the fact that $Q^+(f,g)$ is closer to equilibrium than f, or g. The previous authors prove that if

$$\Phi(f) = \left\| (f - P_f) - M \right\|_\alpha + K \sqrt{\sum_{ij} p_{ij}^2}$$

$$P_f = \mathcal{F}^{-1}\left(-\frac{1}{2} \sum_{ij} p_{ij} \xi_i \xi_j \psi(|\xi|) \right)$$

$$M \quad \text{Maxwellian}$$

with $p_{ij} = \int_{\mathbb{R}^N} \left[v_i v_j - \frac{1}{N} d_{ij} |v|^2 \right] f(v) dv$

and $\mathcal{F}^{-1}$ = inverse Fourier transform

and ψ is a smooth, monotone decreasing function, $\psi \equiv 1$ close to 0, $\psi \equiv 0$ at infinity

and K is large enough

and $\|F\|_\alpha = \sup \dfrac{|\mathcal{F}(F)(\xi)|}{|\xi|^{2+\alpha}}$ ($\mathcal{F}$ =Fourier transform), then

$$\exists c \in (0,1); \ \ \Phi(Q^+(f,g)) \leq \frac{c}{2}[\Phi(f) + \Phi(g)]$$

This is closely related to several interesting formulas in the same spirit:

<u>Tanaka's theorem</u>: $W_2(Q^+(f,f), Q^+(g,g)) \leq W(f,g)$;
in particular $W_2(Q^+(f,f), M) \leq W(f,M)$. Here W_2 is the (improperly called) Wasserstein distance with exponent 2,

$$W_2(f,g) \ = \ \sqrt{\inf \int_{\mathbb{R}^N \times \mathbb{R}^N} |v - w|^2 d\pi(v,w)}$$

where the inf runs over all π's with marginal densities f and g.

Nonexpansivity of Q^+ in Toscani's distance:

$$d_2(Q^+(f,f), Q^+(g,g)) \leq d_2(f,g) \text{ where } d_2(f,g) = \sup_{\xi \in \mathbb{R}^N} \frac{|\hat{f}(\xi) - \hat{g}(\xi)|}{|\xi|^2}$$

(hat stands for Fourier transform).

This can be improved, as shown by Gabetta, Toscani and Wennberg (1995) and by Carlen, Gabetta and Toscani (1999) by looking at

$$d_s(f,g) = \sup_{\xi \in \mathbb{R}^N} \frac{|\hat{f}(\xi) - \hat{g}(\xi)|}{|\xi|^s}$$

for $s > 2$. Since this expression only makes sense if the Taylor polynomials of $\hat{f}$ and $\hat{g}$ at $\xi = 0$ coincide up to order $[s] + 1$, it is necessary to substract a well-chosen polynomial to f before applying d_s to the solution of the BE.

Spectacular results have been obtained by these Fourier-based metrics: in particular, Carlen, Gabetta and Toscani (1999) show that, if f_0 has all its moments bounded, as well as all its Sobolev norms, then

$$\left\| f_t - M \right\|_{L^1} = O\left(e^{-(\lambda - 0)t} \right)$$

(λ = spectral gap of the linearized Boltzmann operator) which is almost optimal. The strong assumptions on the smoothness and moments are used to interpolate between the distance d_s and smoothness/moments, to recover convergence in L^1 norm. I think that the assumptions on the initial datum may be considerably relaxed by a study similar to that in Mouhot and Villani (2004).

Stam-Boltzmann inequalities: These are the analogues of the Stam inequalities:

$$H\left(Q^+(f,g) \right) \leq \frac{1}{2}\Big[H(f) + H(g) \Big]$$
$$I\left(Q^+(f,g) \right) \leq \frac{1}{2}\Big[I(f) + I(g) \Big]$$

They were proven for general Maxwell kernels in Villani (1998). Earlier works by Carlen and Carvalho (1992), and Toscani (1992a,b) considered particular cases. In particular, in the case of a particular kernel, Carlen and Carvalho gave an interesting <u>lower bound</u> on $-\Big[H(Q^+(f,f)) - H(f) \Big]$. This was the

starting point of their theory for trend to equilibrium, based on the important remark that if $B(v - v_*, \sigma) = b(\cos\theta)$, $\int b(\cos\theta)d\sigma = 1$, then

$$D(f) \geq H(f) - H\Big(Q^+(f, f)\Big).$$

(D = entropy production).

As we saw in previous chapters, now there are very good bounds for $D(f)$ — but what for $H(f) - H(Q^+(f, f))$? Recent results by Ball, Barthe and Naor (2003) suggest that a bound like

$$H(f) - H(Q^+(f, f)) \geq \text{const. } [H(f) - H(M)]$$

can be hoped for under *very* strong conditions on f, namely that it satisfy a Poincaré inequality:

$$\int fh = 0 \Longrightarrow \int f|\nabla h|^2 \geq \text{const. } \int fh^2.$$

The story here is certainly not finished.

6.3. The role of high energy tails

In the particular case of Maxwellian molecules, the role of high energy tails to hinder the convergence to equilibrium has been illustrated by rather spectacular works of Carlen and Lu (2003), Carlen, Gabetta and Regazzini (2007). Carlen and Lu showed how to construct solutions converging arbitrarily slowly to equilibrium, say $\|f(t) - M\|_{L^1} \geq K/\log\log\log\log\log t$ as $t \to \infty$; this is achieved by means of very heavy distributional tails at high velocities (still with finite energy). Then Carlen, Gabetta and Regazzini showed that if the initial datum has infinite energy, there is no convergence to equilibrium at all; instead, all the energy goes to high velocities, in a way which can be somewhat quantified.

There exist related works by Bobylev and Cercignani (2002a,b,c), constructing "eternal" self-similar solutions (with infinite energy) of the spatially homogeneous Boltzmann equation.

In all these contributions, Maxwellian interactions play a very particular role and some of these results (maybe all of them) do not hold for hard potentials.

6.4. Behavior of the Fisher information

As a consequence of the Stam-Boltzmann inequalities, one can show that

$$\frac{d}{dt}I(f_t) \leq 0$$

for Maxwell collision kernels. This was first proven by McKean (1965) for the Kac model, then Toscani (1992a,b) and Carlen and Carvalho (1992) for other special cases; finally I treated the general case (1998). A few years ago, this could be considered as a precious regularity estimate; but now much stronger regularity estimates have appeared in the spatially homogeneous case.

There is an interesting connection between this problem and the proof of Theorem 3-ϵ or 3. In the case of Maxwell molecules, the Boltzmann semigroup $(B_t)_{t \geq 0}$ commutes with the Fokker–Planck semigroup $(S_t)_{t \geq 0}$ (this observation goes back to Bobylev). On the other hand,

$$I(S_t f \mid M) = -\frac{d}{dt}H(S_t f \mid M).$$

Therefore

$$
\begin{aligned}
\frac{d}{ds}\Big|_{s=0} I(B_s f \mid M) &= -\frac{d}{ds}\Big|_{s=0}\frac{d}{dt}\Big|_{t=0} H(B_s S_t f \mid M) \\
&= \frac{d}{dt}\Big|_{t=0}\frac{d}{ds}\Big|_{s=0} H(S_t B_s f \mid M) \\
&= \frac{d}{dt}\Big|_{t=0} D(S_t f).
\end{aligned}
$$

This surprising relation was noticed by McKean (1965), later rediscovered by Toscani and Villani (1999b). It implies that $\frac{d}{dt}D(S_t f) \leq 0$ for Maxwell kernel, and confirms the fact that the Fokker–Planck semigroup is a good thing to let act on $D(f)$!

6.5. Variations on a theme

The H Theorem states that the H functional goes down in time along solutions of Boltzmann's equation; and we just saw that, at least in a particular case, the Fisher information I also goes down in time along these

solutions. It is natural to ask if there is more to uncover here. There are in fact several possible generalizations, of various interest.

More general equations: The decrease of the Fisher information can be easily extended to the Landau equation with Maxwellian molecules, either directly (Villani, 2000b), or by passing to the limit of grazing collisions, which relates the Boltzmann and Landau equation. But one can ask whether the restriction to Maxwellian molecules is mandatory. My personal guess is that the decreasing property is not true for general kernels, although I have been somewhat surprised to see numerical simulations suggesting a decreasing behavior for the Landau equation with Coulomb interaction.

At the level of the linear Fokker–Planck equation, it is remarkable that the (relative) Fisher information is indeed decreasing: this is part of the Γ_2 theory of Bakry and Émery, and was underlying McKean's work in dimension 1. Actually, Toscani (1999) showed that a simple adaptation of a trick by McKean yields the result in any dimension. This is somehow the method which I adapted in (Villani, 2000b) to a mildly nonlinear setting.

For the linear Fokker–Planck equation, the decrease of the Fisher information means the convexity (in time) of the Kullback information. McKean (1966) suggested that for Boltzmann's equation, or at least Kac's caricature, the H functional would be not only decreasing with time, but also a convex and even a completely monotone function of time (all derivatives would have alternate signs). This "super-H-Theorem" immediately attracted the attention of the physicists' community, and some people tried to prove it for various models. There was a kind of thrill when it was discovered that certain particular solutions constructed by Bobylev and others did satisfy such a monotonicity property for derivatives up to order about 100; nevertheless the conjecture was in the end disproved. References on this episode can be easily traced from Lieb (1982) and Olaussen (1982). Then physicists withdrew from this problem; to my knowledge, even convexity in time has not been proved or disproved. Numerical simulations often show a curve which vaguely seems convex, in the spatially homogeneous case. On the contrary, there is rather good numerical evidence that such a convexity property is downright false for spatially inhomogeneous equations, and one can also construct explicit examples in the linear case (see Chapter 7 for some intuition).

A final comment about this issue: It so happens that a decade ago, a theoretical physicist named Frieden predicted that Fisher's information would

be decreasing for a large variety of physical systems, and in fact developed a very peculiar vision of the world, with Fisher's information as a unifying principle for most of quantum and classical physics, plus other sciences as well. After publishing his book *Physics from Fisher information: a Unification* (Cambridge University Press, 1998), he recently released the second edition as *Science from Fisher information: a Unification* (2004)[1]. There is actually a small community of researchers working actively on the subject (A. Plastino, A.R. Plastino, B.H. Soffer...). I don't have any opinion on the relevance of this line of research, but just conclude that the decreasing property of the Fisher information along certain physical systems was expected by some, independently of McKean's contribution.

More general functionals: Could the H and I functionals be just the first two members of a large family of functionals which are all decreasing in time along solutions of, say, the spatially homogeneous Boltzmann equation with Maxwellian kernel? This was more or less what McKean (1966) suggested for the Kac caricature, considering the family of functionals obtained by successively differentiating the H functional along the heat equation: define $I_n(f)$ as $(-d/dt)^n H(e^{t\Delta} f)$, hope that I_n is nonincreasing with time. This apparently crazy suggestion becomes less surprising if one recalls McKean's interest in completely integrable partial differential equations. Nobody has proven or disproven this conjecture, but there is reason to believe it is false after Ledoux's formal study (1995) on the linear Fokker–Planck equation.

Functionals involving two solutions: A very striking remark was made by Plastino & Plastino (1996): if f_t and g_t are two solutions of the linear Fokker–Planck equation, then not only are $H(f_t|M)$ and $H(g_t|M)$ nonincreasing functions of t, but so is also $H(f_t|g_t)$. I am not aware of any important application, but this property is "too striking to be useless". It would be interesting to know whether this "nonexpansivity" property, which is somewhat reminiscent of Tanaka's work, also holds true for, say, the Boltzmann equation with Maxwellian molecules.

---

[1]Some more detailed information on these topics can be found on Frieden's Web page at `http://www.optics.arizona.edu/faculty/resumes/frieden.htm`. Not all readers did appreciate the book, as can be seen from the severe review at `http://cscs.umich.edu/~crshalizi/reviews/physics-from-fisher-info/`, but many did find it valuable, and some consider it as a milestone in theoretical physics.

7. Wiping out spatial inhomogeneities

If one is interested in estimating rates of convergence in the hydrodynamic limit, one should only retain from the preceding discussion that inequalities like

$$D(f) \geq K_f H(f \mid M^f)^\alpha$$

are a quantitative way to express statements like: small entropy production $\Longrightarrow f \simeq M^f$.

This is a <u>local</u> statement, "local" meaning "for each x". In the hydrodynamic limit, one only wants to control the local behavior of f and establish the asymptotic equation satisfied by M^f, or equivalently by ρ, u, T.

On the other hand, if one is interested in the trend to equilibrium problem, in a spatially-dependent context, then one wants to establish that <u>globally</u> f looks like a Maxwellian as $t \to \infty$; i.e., there exists a Maxwellian distribution $M(v)$, independent on x, such that $f_t \xrightarrow[t \to \infty]{} M$.

Entropy production estimates can only give an information on how close f is from local equilibrium, i.e. $f \simeq M^f = M(\rho, u, T))$, for some $\rho(x), u(x), T(x)$. In fact, if f is any $M(\rho, u, T)$, then $D(f) = 0$. So if one wants to show $f_t \to M$, then one has to introduce additional input.

The theory of the Cauchy problem for the full Boltzmann equation is <u>incredibly</u> difficult, and situations where one can prove existence of smooth, well-behaved solutions are extremely rare. However, it seems reasonable (to me) to conjecture the following:

<u>**"Conjecture"**</u>: *Let $B(v - v_*, \sigma) = |v - v_*|^\gamma b(\cos\theta)$ (say) with $\gamma > 0$, b integrable, and let f_0 be a smooth (C^∞, rapidly decaying) nonnegative density on $\Omega_x \times \mathbb{R}_v^N$, where $\Omega_x = \mathbb{T}^N$ (torus) or Ω_x is a smooth open subset of $\mathbb{R}^N$, strictly convex [nobody really knows what should happen if Ω_x is not convex], then there exists a unique smooth solution $(f_t)_{t \geq 0}$ to the BE with specular reflection or periodic or bounce-back boundary condition, and initial datum f_0, satisfying*

$$\sup_{t \geq t_0} \left\| f_t \right\|_{L_s^1} < +\infty \quad \forall s$$

$$\sup_{t \geq t_0} \left\| f_t \right\|_{H^k} < +\infty \quad \forall k$$

$$t \geq t_0 \Longrightarrow f_t(x, v) \geq \rho_0 e^{-A_0 |v|_0^q}$$

66

Here

$$\|f\|_{L_s^1} = \int f(x,v)(1+|v|^s)\, dx\, dv,$$

$$\|f\|_{H^k} = \|f\|_{H^k(\Omega_x \times \mathbb{R}_v^N)}.$$

 At present such a theorem can be deduced only in a close-to-equilibrium context from the recent works by Guo.

The program that I started with Desvillettes some time ago is the following: <u>assume</u> that (f_t) satisfies all those conjectured bounds, and get <u>explicit</u> rates of convergence for f_t towards global equilibrium. This should be <u>seen as</u> the last step of a complete argument. Of course a compactness proof solves the problem of convergence in a few lines, but gives no information about explicit rates.

The first, naive approach is the following. Start from the H Theorem,

$$\frac{d}{dt} H(f \mid M) = -\int_{\Omega_x} D(ft)dx, \quad (M = \text{ global eq.})$$

and from this deduce that

$$\int_0^{+\infty} D(f_t)\, dx\, dt < +\infty;$$

as a consequence

$$\int_0^{+\infty} H\left(f_t \mid M^{f_t}\right)^{1+\epsilon} dt < +\infty$$

and in particular

$$\int_0^{+\infty} \left\|f_t - M^{f_t}\right\|_{L^1}^{2(1+\epsilon)} dt < +\infty.$$

Thus it seems that f_t approaches local equilibrium, at a rate no more than $O\left(\dfrac{1}{t^{\frac{1}{2}-0}}\right)$ or something like that.

But this is downright false: A convergent integral may indeed diverge arbitrarily slowly, no matter how smooth the function is: just think of the case where it admits a lot of "plateaux"... (see figure)

Such a behavior may seem unlikely, but there is some reason to suspect that it can happen; in particular, it is possible to cook up simple linear models, completely integrable, for which the entropy production is oscillating
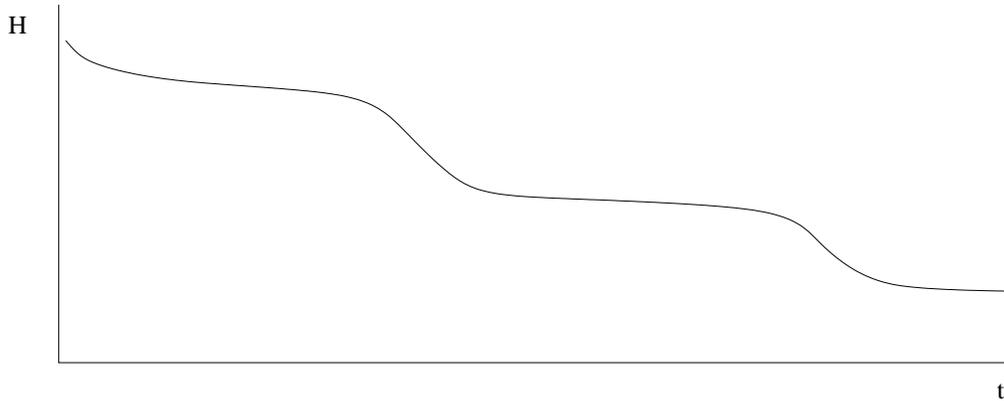
at regular pace. On my request, F. Filbet has performed some very precise numerical simulations on some instances of the spatially inhomogeneous Boltzmann equation in a one-dimensional periodic box (velocity being two-dimensional), and observed beautiful oscillations in the entropy production... As a conclusion, not only is it difficult to rule out the behavior sketched in the above figure, but I personally suspect that it is often the rule! A more precise discussion can be found in Desvillettes & Villani (2005).

From the physical point of view, it is also a heresy to proceed along the lines sketched above. Indeed, contrarily to the hydrodynamic limit, here there is no separation of scales: we are looking at a regime where the mean free path is of the same order as the typical length scale (Knudsen number = $O(1)$). So the right thing to do is to work at the same time on the deviation of f from M^f, and on the deviation from M^f to M (or equivalently of ρ, u, T from their equilibrium values). In short,

- like in the hydrodynamic limit problem we have to study the approximation of M^f by f, and the behavior of ρ, u, T (hydrodynamic fields);

- but on the contrary to the hydrodynamical limit problem, here we should study at the same time the convergence of f towards M^f, and of ρ, u, T towards the equilibrium value. Let's say that the trend to equilibrium problem is at the same time local and global, while the hydrodynamical limit problem is essentially local. I won't discuss this in more detail since both problems are still far to be fully understood.

Theorem 10 (Desvillettes & Villani, 2005): *Let $B(v - v_*, \sigma)$ satisfy*

$$K_B |v - v_*|^\gamma \wedge (1 + |v - v_*|)^{-\beta} \leq B(v - v_*, \sigma) \leq C_B (1 + |v + v_*|)^{k_B} \frac{\theta^{-(1+1/\nu_B)}}{\sin^{N-2}\theta}$$

with $\beta, \gamma \geq 0$, $K_B < +\infty$, $\nu_B < 2$.
Let Ω_x be $\mathbb{T}^N$ or a smooth open subset of $\mathbb{R}^N$, connected, and let $(f_t)_{t \geq 0}$ be a very smooth solution of the BE

$$\frac{\partial f}{\partial f} + v \cdot \nabla_x f = Q(f, f),$$

with periodic boundary conditions or bounce-back reflection or specular reflection. In the case of specular reflection, further assume that Ω has no axis of symmetry and $N = 2$ or 3. Without loss of generality, assume that

$$|\Omega_x| = 1 \quad \int f_0 \, dv \, dx = 1 \quad \int f_0(x, v)|v|^2 \, dv \, dx = N,$$

and if $\Omega_x = \mathbb{T}^N$ assume $\int f_0(x, v) v \, dv \, dx = 0$. Further assume that

$$\forall s \quad \sup_{t \geq 0} \left\| f_t \right\|_{L_s^1} < +\infty$$

$$\forall k \quad \sup_{k \geq 0} \left\| f_t \right\|_{H^k} < +\infty$$

$\exists \rho_0, A_0, q_0; \ \forall t \geq 0 \quad \forall v \ \forall x \quad f_t(x, v) \geq \rho_0 e^{-A_0 |v|^{q_0}}$.
Then, if M stands for the global equilibrium

$$M(x, v) = M(v) = \frac{e^{-\frac{|v|^2}{2}}}{(2\pi)^{N/2}},$$

one has $H(f_t \mid M) = O(t^{-\infty})$ in the sense that for all $\epsilon > 0$ there exists a constant C_ϵ, only depending on B, Ω_x, on the boundary conditions, on ρ_0, A_0, q_0, on $\sup\limits_{t \geq 0} \|f_t\|_{L_s^1}$ and $\sup\limits_{t \geq 0} \|f_t\|_{H^k}$ for some $s = s(\epsilon, q_0, B)$ and $k = k(\epsilon, q_0, B)$, such that $H(f_t \mid M) \leq C_\epsilon t^{-1/\epsilon}$. In particular

$$\left\| f_t - M \right\|_{L^1(\Omega_x \times \mathbb{R}^N)} = O(t^{-\infty}).$$

69

I will not reproduce the proof, which includes dozens of pages of "pure computation", but only try to explain the main (hopefully interesting) ideas and the overall plan.

First of all, the general strategy is, as in the spatially homogeneous case, to reduce to a system of differential inequalities for which the rate of convergence can be estimated by elementary means.

As long as $H(f \mid M^f)$ [here H is with respect to $dx\,dv$] is significantly large with respect to $H(f \mid M)$, one can use the quantitative H Theorem and get

$$-\frac{d}{dt}H(f \mid M) = \int_{\Omega_x} D(f_t(x, \cdot))dx \geq K_{f_t}H(f_t \mid M)^{1+\epsilon}.$$

K constant depending on Ω_x and on bounds on f_t.

This differential inequality, were it true for all times, would ensure $O(t^{-\infty})$ convergence by Growvall's lemma. BUT it may very well happen that $H(f \mid M^f)$ becomes very small at some time, and actually the total entropy production $\displaystyle\int_{\Omega_x} D(f_t(x, \cdot))dx$ would become very small compared to $H(f \mid M)$.

A general plan of attack against this degeneracy was proposed by Desvillettes & Villani (2001). In the present context it would consist in estimating from below the second time-derivative of a well-chosen functional measuring the discrepancy between between f_t and M^{f_t}. In our context the most natural choice would be to look at $\dfrac{d^2}{dt^2}H(f_t \mid M^{f_t})$; but this leads to very tricky problems to control errors everywhere, and it is much better to look at $\dfrac{d^2}{dt^2}\|f_t - M^{f_t}\|_{L^2}^2$. In fact, this is one of the necessary ingredients to avoid using too strong a priori estimates like $f_t \in L^\infty(M^{-1})$ or something like that.

Behind this calculation is the following idea: if, at some time t_*, $f_{t_*} = M^{f_{t_*}}$, then f_t has to strictly depart from M^{f_t} just after t_*, unless $M^{f_{t_*}} = M$ [because M is the only local Maxwellian solving the BE]. Thus the second-order derivative $\dfrac{d^2}{dt^2}\|f_t - M^{f_{t_*}}\|_{L^2}^2$ should be strictly positive whenever f_t is very close to M^{f_t} as ρ_t, u_t, T_t are not at their equilibrium values $(1, 0, 1)$. Then one would combine this differential inequality on $\left\|f_t - M^{f_t}\right\|_{L^2}^2$ with

the entropy production estimate (H Theorem) to get a system of differential inequalities involving both the distance between F_t to M^{f_t}, and the distance between M^{f_t} and M.

This strategy was successfully applied to the model case of the linear Fokker–Planck equation (spatially dependent). However, the Boltzmann equation is nastier, and the above guess about $\dfrac{d^2}{dt^2}\left\| f_t - M^{f_t} \right\|^2_{L^2}$ is <u>false</u>. There are a lot of hydrodynamic fields ρ, u, T (x-dependent!), which are not in equilibrium, and such that this second derivative vanishes. Desvillettes and I called these states <u>quasi-equilibria</u>. They correspond to a considerable slowing down of the entropy production: whenever t passes through a quasi-equilibrium at some time t_*, then not only does the total entropy production $\displaystyle\int_{\Omega_x} D(f_{t_*}(x, \cdot))\, dx$ vanish, but also its time derivative at time t_*; so formally, $H(f_t | M^{f_t}) = O((t - t_*)^4)$ near t_*.

These quasi-equilibria depend on the global geometry of the problem, i.e. both the shape of Ω_x and the boundary conditions. They are determined as the solutions of the equations

$$\left[\begin{array}{l} T = \text{constant} \\ \{\nabla u\} = 0 \end{array} \right.$$

where $\{\nabla u\}$ is the deviator of u, i.e. the traceless part of the symmetric gradient of u

$$\left\{ \nabla u \right\}_{ij} = \frac{1}{2}\left[\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right] - \frac{(\nabla . u)}{N}\delta_{ij}.$$

Obviously, for any Ω_x one can construct quasi-equilibria by imposing $T = \text{constant}$, $u = 0$, ρ arbitrary (x-dependent). The existence of solutions with non-trivial velocity field, on the other hand, depends on Ω_x:

- if $\Omega_x = \mathbb{T}^N$, then $\{\nabla u\} = 0 \implies \nabla u = 0$ (so $u = 0$, with our normalization)

- if $\Omega_x \subset \mathbb{R}^N$ with bounce-back condition, then $\{\nabla u\} = 0 \implies u = 0$

- If $\Omega_x \subset \mathbb{R}^2$ with specular reflection, then there always exists a 3-dimensional vector space of solution to $\{\nabla u\} = 0$

- If $\Omega_x \subset \mathbb{R}^3$ with specular reflection, then there is no quasi-equilibrium with non-trivial velocity field unless Ω_x has an axis of symmetry (if Ω_x is a spherical ball, then there is, in addition to the obvious ones, a strange quasi-equilibrium, as shown to me by É. Ghys).

The remedy which we suggest is to estimate, not only the distance between f and $M(\rho, u, T) = M^f$, but also between f and $M(\rho, u, 1)$, and between f and $M(\rho, 0, 1)$; which is an attempt to separate as much as possible the roles of the three hydrodynamic fields.

Now we have all the main ideas for establishing a system of differential inequalities.

First step: Write a quantitative H Theorem

(E.1) $\qquad -\dfrac{d}{dt} H(f_t \mid M) = \displaystyle\int_{\Omega_x} D(f_t(x,.)) dx \geq K_1 H(f_t \mid M^{f_t})^{1+\epsilon}$

Here K_1 depends on f, on B and on ϵ, which will be fixed until the end but arbitrarily small.

Second step: Explicity compute $\dfrac{d^2}{dt^2}\left\| f - M(\rho, u, T)\right\|_{L^2}^2$, $\dfrac{d^2}{dt^2}\left\| f - M(\rho, u, 1)\right\|$ $\dfrac{d^2}{dt^2}\left\| f - M(\rho, 0, 1)\right\|_{L^2}^2$, where

$$\rho = \rho_t = \int f_t dv, \quad \rho u = \rho_t u_t = \int f_t v dv,$$

$$\rho|u|^2 = \rho_t|u_t|^2 = \frac{1}{N}\int f_t |v - v_t|^2 dv$$

This is a monster computation; for the reader's amusement, I reproduced only the term $\dfrac{1}{M(\rho, u, T)}\dfrac{\partial^2}{\partial t^2} M(\rho, u, T)$ at the very end of the notes. This expression may look frightening, but from the formula given one sees that it is formally $O(\|f - M\|)$ and this is all one has to retain from this intermediate computation. After one has separated the important contributions from the rest, one arrives at

$$\frac{d^2}{dt^2}\left\| f - M(\rho, u, T)\right\|_{L^2}^2 \geq K_2\left[\int_{\Omega_x}\left\{\nabla u\right\}^2 + \int_{\Omega_x} |\nabla T|^2\right]$$

72

(E.2)

$$-\frac{C_2}{\delta_2^{1-\epsilon}}\left\|f - M(\rho, u, T)\right\|_{L^2}^{2(1-\epsilon)} - \delta_2 H(f \mid M)$$

Here δ_2 is arbitrarily small, and the constants depend on the a priori estimates on f (and thus on $M(\rho, u, T)$ as well). The appearence of the ϵ in the second term of the right-hand side is due to an interpolation procedure, which is mandatory if one wants to measure the distance between f and $M(\rho, u, T)$ with the L^2 norm. Indeed, very strong norms (with derivatives with respect to v and x) of $f - M(\rho, u, T)$ appear in the process of estimating $\frac{d^2}{dt^2}\|f - M(\rho, u, T)\|_{L^2}^2$; then to "transform" these strong norms in L^2, one interpolates with the smoothness bounds on f and $M(\rho, u, T)$, and an ϵ exponent is lost in the battle. Note that the lower bound assumption on f is important here to ensure the smoothness of ρ, u and T (and the strict positivity of ρ and T).

Also note the asymmetry between the way we quantify the distance from f to $M(\rho, u, T)$, and the distance from f to M. Using $H(f \mid M(\rho, u, T))$ would require much stronger assumptions on the tail behavior of f.

Equation (E.2) can interpreted as follows: if the distance between f and $M^f = M(\rho, u, T)$, ever becomes extremely close to 0, and if $\{\nabla u\} \not\equiv 0$ or $\nabla T \not\equiv 0$ [so that $M(\rho, u, T)$ is not a quasi-equilibrium] then, up to a small error $(\delta_2 H(f \mid M))$ the rate at which f_t will depart from M^{f_t} at later times is estimated from below.

One establishes in a similar way

$$\frac{d^2}{dt^2}\left\|f - M(\rho, u, 1)\right\|_{L^2}^2 \geq K_3 \int_{\Omega_x} |\nabla^{sym} u|^2 \ - \ \frac{C_3}{\delta_3^{1-\epsilon}}\left\|f - M(\rho, u, 1)\right\|_{L^2}^{2(1-\epsilon)}$$
$$- \ \delta_3 H(f \mid M)$$

$$\frac{d^2}{dt^2}\left\|f - M(\rho, 0, 1)\right\|_{L^2}^2 \geq K_4\left(\int_{\Omega_x} |\nabla \rho|^2 + \int_{\Omega_x} (\nabla_x . u)^2\right) \ - \ \frac{C_4}{\delta_4^{1-\epsilon}}\left\|f - M(\rho, u, 1)\right\|_{L^2}^{2(1-\epsilon)}$$
$$- \ \delta_4 H(f \mid M)$$

Here $\nabla^{sym}u$ is the symmetrized gradient of u:

$$\left[\nabla^{sym}u\right]_{ij} = \frac{\partial_i u_j + \partial_j u_i}{2}$$

Note that on the whole, the quantities $\int |\nabla \rho|^2$, $\int |\nabla^{sym}u|^2$, $\int |\nabla T|^2$ have appeared, and that the combination of all of them controls how far ρ, u, T are from equilibrium. This actually is the content of

Step 3: One establishes <u>functional inequalities</u> to quantify how the quadratic ∇_x quantities seen above control the distance between the hydrodynamic fields and their equilibrium values. At this point one uses

- the lower bound on ρ

- the boundary conditions

- the shape of Ω_x

- the <u>global</u> conservation laws.

<u>Remark</u>: the fact that these quadratic gradient quantities appear is a consequence of our strategy which looks at second derivatives along the differential operator $v.\nabla_x$, in such a way that the square of the partial derivative $\dfrac{\partial}{\partial t}$ contributes to the main part.

- <u>Inequality for ρ</u> : Under the constraint $\int \rho = 1$, $|\Omega_x| = 1$, one has

    (E.5) $\displaystyle\int_{\Omega_x} |\nabla \rho|^2 \geq K_5 \int \rho \log \rho$ (Poincaré inequality in fact)

- <u>Inequality for u</u>: If $\Omega_x = \mathbb{T}^N$ or $\Omega_x \subset \mathbb{R}^N$ with bounce-back condition ($u = 0$ on $\partial\Omega_x$), or $\Omega_x \subset \mathbb{R}^N$ with specular reflection ($u \cdot n = 0$ on $\partial\Omega_x$) with no axis of symmetry, then

    (E.6) $\displaystyle\int_{\Omega_x} |\nabla^{sym}u|^2 \geq K_6 \int_{\Omega_x} |\nabla u|^2$

74

This is a variant of Korn's inequality, well-known in elasticity and in hydrodynamics. One original thing here is that it appears with the tangency constraint $u \cdot n = 0$, while in both abovementioned fields it usually comes with no slip boundary conditions: $u = 0$ on some portion of $\partial \Omega_x$. The constant K_6 can be explicitly estimated in terms of how much the domain Ω_x departs from being axisymmetric.

Next, by Poincaré's inequality with reference measure $\rho(x)dx$, [which is uniformly smooth and bounded from below, and has unit integral!]

$$(\text{E.6'}) \qquad \int_{\Omega_x} |\nabla u|^2 \geq \frac{1}{||\rho||_{L^\infty}} \int \rho |\nabla u|^2 \geq \frac{K_6'}{||\rho||_{L^\infty}} \int \rho |u|^2.$$

(here we use the boundary condition on u again; and $\int \rho u = 0$ if $\Omega_x = \mathbb{T}^N$.)

- Inequality for T: Using the fact that ρ is bounded from below and above,

$$\int \rho T + \frac{1}{N} \int \rho |u|^2 = 1,$$

and a Poincaré inequality, one shows

$$(\text{E.7}) \int_{\Omega_x} |\nabla T|^2 \geq K_7 \times \frac{N}{2} \int \rho (T - \log T - 1) \, dx - C_7 \int \rho \frac{|u|^2}{2} \, dx$$

Our strange choices for estimating the distance of ρ, u, T to $1, 0, 1$ will become clear after the next step.

Step 4 : It consists in quantifying the idea that if f is close to M^f, and M^f is close to M, then f is close to M. This is provided by a well-known formula of additivity of the entropies:

$$H(f \mid M) =$$

(E.8)

$$H(f \mid M(\rho, u, T)) + \left[\int \rho \log \rho + \int \rho \frac{|u|^2}{2} + \frac{N}{2} \int \rho (T - \log T - 1) \right]$$

75

Let us complement it with the similar formulas

(E.9) $H(f \mid M) = H(f \mid M(\rho, u, 1)) + \left[\int \rho \log \rho + \int \rho \frac{|u|^2}{2} \right]$

(E.10) $H(f \mid M) = H(f \mid M(\rho, 0, 1)) + \int \rho \log \rho$

Equations (E.1)-(E.10) constitute a system of differential inequalities involving the quantities $H(f \mid M)$, $H(f \mid M(\rho, u, T))$, $H(f \mid M(\rho, u, 1))$, $H(f \mid M(\rho, 0, 1))$, $\|f - M(\rho, u, T)\|_{L^2}^2$, $\|f - M(\rho, u, 1)\|_{L^2}^2$, $\|f - M(\rho, 0, 1)\|_{L^2}^2$. To reduce the number of quantities involved, we can use interpolation again:

(E.11) $\left\| f - M(\rho, u, T) \right\|_{L^2}^2 \leq C_{11} H(f \mid M(\rho, u, T))^{1-\epsilon}$

(E.12) $\left\| f - M(\rho, u, 1) \right\|_{L^2}^2 \leq C_{12} H(f \mid M(\rho, u, 1))^{1-\epsilon}$

(E.13) $\left\| f - M(\rho, 0, 1) \right\|_{L^2}^2 \leq C_{13} H(f \mid M(\rho, 0, 1))^{1-\epsilon}.$

Now we have a closed system of differential inequalities on the quantities $H(f \mid M)$, $\|f - M(\rho, u, T)\|_{L^2}^2$, $\|f - M(\rho, u, 1)\|_{L^2}^2$, $\|f - M(\rho, 0, 1)\|_{L^2}^2$. To study its time behavior, we need to quantify the fact that a solution of

$$y''(t) \geq K - Cy(t)^{1-\epsilon}$$

cannot be always small if $K > 0$. For this we make systematic use of the following lemma from Desvillettes and Villani (2001):

Lemma L.1: Let $y(t) \geq 0$ be a C^2 function solving the differential inequality

$$\frac{d^2}{dt^2} y(t) + \frac{CA}{\delta} y^{1-\epsilon} \geq K\alpha \quad \text{for} \quad T_1 \leq t \leq T_2$$

where C, A, δ, K, α are positive constants, A bounded below, ϵ small enough. Then,

- either $[T_1, T_2]$ is short:

$$T_2 - T_1 \leq \text{const.} \left(\frac{A}{\delta} \right)^{-\frac{1}{2(1-\epsilon)}} \alpha^{\frac{\epsilon}{2(1-\epsilon)}}$$

76

- or $y(t)$ is bounded below <u>on the average</u>:

$$\frac{1}{T_2 - T_1} \int_{T_1}^{T_2} y(t)\, dt \geq \text{const} \left(A^{-1}\delta\right)^{\frac{3}{2}+10\epsilon} \alpha^{\frac{1}{1-\epsilon}}.$$

with constants only dependng on C, K and ϵ (not on A, δ, α).

In short: either $T_2 - T_1$ is bounded in $O(\alpha^0)$ or the average value of y is bounded below in $O(\alpha^{1+0})$.

The dichotomy has to do with the fact that we are dealing with <u>second order</u> differential inequalities, for which there is no such thing as Gronwall's lemma, and there may be some <u>delay</u> before the positivity of the right-hand side is "transfered" onto $y(t)$. One can easily see that anything can happen if $[T_1, T_2]$ is too short.

The lemma is established "by hand" and relies on a careful cutting of $[T_1, T_2]$ into many sub-pieces depending on the values take by $y(t)$. the main idea is that if y becomes very small, then, due to the inequality, it has to become strictly convex as a function of t, then either it will very soon leave this regime where it is very small, or it will stay very small for a long time, but then, due to uniform convexity, finally get out of the smallness regime with a very fast rate of increase.

If we want to apply lemma L.1 to the differential inequalities established in previous steps, we run into a serious difficulty: we would like to use equation (E.2) only when $\int |\nabla T|^2$ (or $\int \rho(T - \log T - 1)$) is large and $H(f \mid M(\rho, u, T))$ is very small; equation (E.3) only when $\int \rho|u|^2$ is large and $H(f \mid M(\rho, u, 1))$ is very small; equation (E.4) only when $\int \rho \log \rho$ is large and $H(f \mid M(\rho, 0, 1))$ is very small. In each case, we need lower bounds on the lengths of the time intervals on which each event occurs. What if f_t switches very quickly from one of these "bad" states to the other, and keeps doing that? Then we cannot apply lemma L.1 in an interesting way. So the next step consists in ruling out this possibility, by establishing that the hydrodynamic quantities vary "slowly" in term:

Step 5: By using the Boltzmann equation again, together with interpolation

and lower bounds, one establishes that

$$\left| \frac{d}{dt} \int \rho_t \log \rho_t, \ \frac{d}{dt} \int \rho_t |u_t|^2, \ \frac{d}{dt} \int \rho_t (T_t - \log T_t - 1) \right| \le C_{14} H(f \mid M)^{1-\epsilon}$$
(E.14)
(when the system approaches global equilibrium the hydrodynamic quantities vary slowly.)

Note that we do <u>not</u> know how to establish a similar bound for $\left| \dfrac{d}{dt} H(f_t) \right|$ [kinetic entropy!] without stronger assumptions on the tail behavior of f_t (pointwise exponential bounds or so...).

All the functional job has now been done and we are left with a problem on differential inequalities (which does not mean that the rest is easy.....). The last step will yield the conclusion:

Step 6: Using the fact that the quantities

$$H(f_t \mid M), \quad \|f_t - M(\rho, u, T)\|_{L^2}^2, \quad \|f_t - M(\rho, u, 1)\|_{L^2}^2, \quad \|f_t - M(\rho, 0, 1)\|_{L^2}$$

and the quantities $\displaystyle\int \rho_t \log \rho_t, \ \int \rho_t |u_t|^2, \ \frac{N}{2} \int \rho_t (T_t - \log T_t - 1)$ satisfy the system of differential inequalities given by (E.1)-(E.14), and are bounded quantities, prove that
$$H(f_t \mid M) = O(t^{-\frac{1}{250\epsilon}}).$$

This step is again done by cutting into sub- pieces a time-interval on which $\alpha \ge H(f_t \mid M) \ge \lambda\alpha$, where $\lambda < 1$ is fixed ($\lambda = 4/5$), such that on the subpieces the entropy production is either high or small. Then re-cut into pieces the intervals on which the entropy production is small, according to whether $\displaystyle\int |\nabla T|^2$ is large or small; then repeat the process for $\displaystyle\int \rho|u|^2$, then again for $\displaystyle\int \rho \log \rho \ ...$. Then use all the differential inequalities, together with lemma L.1, to show a lower bound <u>on the average</u> on the entropy production, and in the end to bound the length of the time interval into consideration by a function of α.

78

8. Towards exponential convergence?

For various reasons, our method of proof cannot yield exponential decay. In certain situations it is not expected that exponential decay occurs: in particular, for soft potentials, the linearized Boltzmann operator has no spectral gap and no exponential convergence is expected (Caflisch (1980) has some results in $O(e^{-t^{\alpha}})$ for some cases). However, there are other cases, such as hard spheres, in which exponential decay is expected. There is no hope in trying to fix our proof to get such a result, it is much better to turn to a *linearized* study. One of our goals when working on this problem was to find a method which does not use linearization, mainly because linearization is not so general and does not yield any control on the time it takes until one enters the linearized regime. But now that we have some quantitative estimates, we can hope to complement them with a linearized study.

A major difficulty arising here is that our setting is not the natural one for the linearization: to get a self-adjoint linear Boltzmann collision operator, the natural space is $L^2(M^{-1})$, which is much, much narrower than the spaces we have been working in. It looks extremely doubtful that we can establish strong enough estimates to reduce to that space, so the best thing to do is accept the idea that we are going to work in a non-self-adjoint setting. Such a study was already performed by Arkeryd (1988), with later refinements due to Wennberg (1993a,b). Still, their work was limited to the spatially homogeneous case, and also involved a compactness argument to reduce from the self-adjoint to the non-self-adjoint case. Moreover, spectral gap in the self-adjoint case itself was based on a standard compactness argument (a consequence of Weyl's perturbation theorem for linear operators), thus preventing any hope of getting explicit estimates whatsoever.

So here is the program:

1. Find a constructive method for bounding below the spectral gap in the self-adjoint case, i.e. in the natural space $L^2(M^{-1})$, say for hard spheres;

2. Find a constructive argument to go from spectral gap in $L^2(M^{-1})$ to spectral gap in L^1, with all the subtleties associated with spectral theory of non-self-adjoint operators in infinite dimension...

3. Find a constructive argument to overcome the degeneracy in the space variable, to get an exponential decay for the linear semigroup associated with the linearized spatially inhomogeneous Boltzmann equation; something

similar to hypo-ellipticity techniques.

4. Combine the whole thing with a perturbative and linearization analysis to get the exponential decay for the nonlinear equation, very close to equilibrium.

This of course is a long-run program, on which my student and collaborator Mouhot has been able to make substantial progress, in collaboration with Baranger on one hand, Gallay on the other hand. As of now, Step 1 has been solved in Baranger & Mouhot (2005) with a clever method based on intermediate collisions. The reader can consult their paper for a very clear exposition. In a much simpler context, this method of intermediate collisions was used by Carrillo, McCann and Villani (2003); and in a different study, I had tried once to use intermediate collisions to improve some tricky regularity estimates for the Boltzmann equation with grazing collisions (Villani, 1999); moreover, I recently learnt from Diaconis that the main idea is very close to a useful trick in the study of Markov kernels. In spite of all these precursors, the proof by Baranger and Mouhot is extremely original and, in my opninion, the physical idea which underlies it is quite appealing.

In relation with this, Mouhot and Strain (2006) recently clarified the role of the grazing collisions (non-cutoff assumption) in the spectral theory of the linearized Boltzmann equation; a topic which had been studied by Pao and Klaus several decades ago.

Now for the rest of the program: Steps 2 and 4 have been worked out by Mouhot (2006) in the case of spatially homogeneous Boltzmann equation with hard potentials; one of the basic ingredients comes from an earlier work by Gallay and Wayne (2002), but required significant effort to adapt to the Boltzmann case. Up to refinements and improvements, it seems fair to state that the spatially homogeneous situation is now understood.

All the final effort should now bear on Step 3. I have worked on that recently, developing estimates for various linear operators, in a way which is somewhat reminiscent of hypo-ellipticity but has a distinct flavor — a phenomenon which I call "hypo-coercivity", following a suggestion by Gallay. Many other people have been recently working on that problem with various tools and for various inhomogeneous linear Fokker–Planck equations: for instance Talay (2002), Rey-Bellet & Thomas (2000), Mattingly & Stuart [& Higham] (2002), using ergodic Markov chain theory; Hérau & Nier (2004), Eckmann & Hairer (2003), using simple pseudo-differential theory. The work

by Hérau and Nier was actually the first one in which an explicit exponential rate was derived for an inhomogeneous Fokker–Planck equation; a recent volume of Lecture Notes by Helffer and Nier (2005) gives a simplified and more general variant of their approach, together with an overview of the field. The typical example of such an equation is the following:

$$\frac{\partial f}{\partial t} + v \cdot \nabla_x f - \nabla V(x) \cdot \nabla_v f = \nabla_v \cdot (\nabla_v f + fv).$$

As the reader can see, the left-hand side is transport and confinement by a potential, while the right-hand side is a standard linear Fokker–Planck operator, which contains drift and diffusion, BUT only in the velocity variable.

In this problem (at least in kinetic theory) an analytic approach has in the end many advantages over a stochastic approach, one of the reasons being that most kinetic equations have a very complicated stochastic interpretation; they cannot be written just as the law of diffusion process as is the case for the inhomogeneous Fokker–Planck equation. Note that the nonlinear method described in the previous chapter has been applied to the very same linear inhomogeneous Fokker–Planck equation by Desvillettes & Villani (2001), only to get an $O(t^{-\infty})$ rate of convergence.

All my recent progress on this question is gathered in a preprint (*Hypocoercivity*, 2006), which will appear as a Memoir of the American Mathematical Society.

The story is not over and clearly it will soon get into new developments. For instance I would bet that the exponential convergence for the inhomogeneous Boltzmann equation, conditional to strong regularity estimates, can be obtained in the next few years.

References

[Alexandre and Villani, 2004] Alexandre, R. and Villani, C. (2004). On the Landau approximation in plasma physics. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 21(1):61–95.

[Ané et al., 2000] Ané, C., Blachère, S., Chafaï, D., Fougères, P., Gentil, I., Malrieu, F., Roberto, C., and Scheffer, G. (2000). *Sur les inégalités de Sobolev logarithmiques*, volume 10 of *Panoramas et Synthèses [Panoramas and Syntheses]*. Société Mathématique de France, Paris. With a preface by Dominique Bakry and Michel Ledoux.

[Arkeryd, 1983] Arkeryd, L. (1983). L^∞ estimates for the space-homogeneous Boltzmann equation. *J. Statist. Phys.*, 31(2):347–361.

[Arkeryd, 1988] Arkeryd, L. (1988). Stability in L^1 for the spatially homogeneous Boltzmann equation. *Arch. Rational Mech. Anal.*, 103(2):151–167.

[Ball et al., 2003] Ball, K., Barthe, F., and Naor, A. (2003). Entropy jumps in the presence of a spectral gap. *Duke Math. J.*, 119(1):41–63.

[Baranger and Mouhot, 2005] Baranger, C. and Mouhot, C. (2005). Explicit spectral gap estimates for the linearized Boltzmann and Landau operators with hard potentials. *Rev. Mat. Iberoamericana*, 21(3):819–841.

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

[Bobylev, 1988] Bobylev, A. V. (1988). The theory of the nonlinear, spatially uniform Boltzmann equation for Maxwellian molecules. *Sov. Sci. Rev. C. Math. Phys.*, 7:111–233.

[Bobylev and Cercignani, 1999] Bobylev, A. V. and Cercignani, C. (1999). On the rate of entropy production for the Boltzmann equation. *J. Statist. Phys.*, 94(3-4):603–618.

[Bobylev and Cercignani, 2002a] Bobylev, A. V. and Cercignani, C. (2002a). Exact eternal solutions of the Boltzmann equation. *J. Statist. Phys.*, 106(5-6):1019–1038.

[Bobylev and Cercignani, 2002b] Bobylev, A. V. and Cercignani, C. (2002b). The inverse Laplace transform of some analytic functions with an application to the eternal solutions of the Boltzmann equation. *Appl. Math. Lett.*, 15(7):807–813.

[Bobylev and Cercignani, 2002c] Bobylev, A. V. and Cercignani, C. (2002c). Self-similar solutions of the Boltzmann equation for non-Maxwell molecules. *J. Statist. Phys.*, 108(3-4):713–717.

[Bonami, 1970] Bonami, A. (1970). Étude des coefficients de Fourier des fonctions de $L^p(G)$. *Ann. Inst. Fourier (Grenoble)*, 20(fasc. 2):335–402 (1971).

[Caflisch, 1980] Caflisch, R. (1980). The Boltzmann equation with a soft potential. *Comm. Math. Phys.*, 74:71–109.

[Caputo, 2003] Caputo, P. (2003). Uniform Poincaré inequalities for unbounded conservative spin systems: the non-interacting case. *Stochastic Process. Appl.*, 106(2): 223–244.

[Carleman, 1932] Carleman, T. (1932). Sur la théorie de l'equation intégrodifférentielle de Boltzmann. *Acta Math.*, 60:369–424.

[Carlen, 1991a] Carlen, E. A. (1991a). Some integral identities and inequalities for entire functions and their application to the coherent state transform. *J. Funct. Anal.*, 97(1):231–249.

[Carlen, 1991b] Carlen, E. A. (1991b). Superadditivity of Fisher's information and logarithmic Sobolev inequalities. *J. Funct. Anal.*, 101(1):194–211.

[Carlen and Carvalho, 1992] Carlen, E. A. and Carvalho, M. C. (1992). Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation. *J. Statist. Phys.*, 67(3-4):575–608.

[Carlen and Carvalho, 1994] Carlen, E. A. and Carvalho, M. C. (1994). Entropy production estimates for Boltzmann equations with physically realistic collision kernels. 74(3-4):743–782.

[Carlen et al., 2000] Carlen, E. A., Carvalho, M. C., and Gabetta, E. (2000). Central limit theorem for Maxwellian molecules and truncation of the Wild expansion. *Comm. Pure Appl. Math.*, 53(3):370–397.

[Carlen et al., 2001] Carlen, E. A., Carvalho, M. C., and Loss, M. (2001). Many-body aspects of approach to equilibrium. In *Séminaire: Équations aux Dérivées Partielles, 2000–2001*, Sémin. Équ. Dériv. Partielles, pages Exp. No. XIX, 12. École Polytech., Palaiseau.

[Carlen et al., 2003] Carlen, E. A., Carvalho, M. C., and Loss, M. (2003). Determination of the spectral gap for Kac's master equation and related stochastic evolution. *Acta Math.*, 191(1):1–54.

[Carlen et al., 1997] Carlen, E. A., Carvalho, M. C., and Wennberg, B. (1997). Entropic convergence for solutions of the Boltzmann equation with general physical initial data. *Transport Theory Statist. Phys.*, 26(3):373–378.

[Carlen et al., 1998] Carlen, E. A., Esposito, R., Lebowitz, J. L., Marra, R., and Rokhlenko, A. (1998). Kinetics of a model weakly ionized plasma in the presence of multiple equilibria. *Arch. Rational Mech. Anal.*, 142(3):193–218.

[Carlen et al., 2007] Carlen, E. A., Ester, G., and Regazzini, E. (2007). On the rate of explosion for infinite energy solutions of the spatially homogeneous Boltzmann equation. Preprint.

[Carlen et al., 1999] Carlen, E. A., Gabetta, E., and Toscani, G. (1999). Propagation of smoothness and the rate of exponential convergence to equilibrium for a spatially homogeneous Maxwellian gas. *Comm. Math. Phys.*, 199(3):521–546.

[Carlen and Lu, 2003] Carlen, E. A. and Lu, X. (2003). Fast and slow convergence to equilibrium for Maxwellian molecules via Wild sums. *J. Statist. Phys.*, 112(1-2):59–134.

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

[Carrillo et al., 2003] Carrillo, J. A., McCann, R. J., and Villani, C. (2003). Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Rev. Mat. Iberoamericana*, 19(3):971–1018.

[Cercignani, 1982] Cercignani, C. (1982). *H*-theorem and trend to equilibrium in the kinetic theory of gases. *Arch. Mech.*, 34:231–241.

[Cover and Thomas, 1991] Cover, T. M. and Thomas, J. A. (1991). *Elements of information theory*. Wiley Series in Telecommunications. John Wiley & Sons Inc., New York. A Wiley-Interscience Publication.

[Dembo, 1989] Dembo, A. (1989). Simple proof of the concavity of the entropy power with respect to added Gaussian noise. *IEEE Trans. Inform. Theory*, 35(4):887–888.

[Dembo et al., 1991] Dembo, A., Cover, T., and Thomas, J. (1991). Information theoretic inequalities. *IEEE Trans. Inform. Theory*, 37(6):1501–1518.

[Desvillettes, 1989] Desvillettes, L. (1989). Entropy dissipation rate and convergence in kinetic equations. *Comm. Math. Phys.*, 123(4):687–702.

[Desvillettes, 1990] Desvillettes, L. (1990). Convergence to equilibrium in large time for Boltzmann and BGK equations. *Arch. Rational Mech. Anal.*, 110(1):73–91.

[Desvillettes, 1993] Desvillettes, L. (1993). Some applications of the method of moments for the homogeneous Boltzmann equation. *Arch. Rational Mech. Anal.*, 123(4):387–395.

[Desvillettes, 2000] Desvillettes, L. (2000). Convergence to the thermodynamical equilibrium. In G. Iooss, O. G. and Nouri, A., editors, *Trends in Applications of Mathematics to Mechanics, Monographs and Surveys in Pure and Applied Mathematics*, pages 115–126. Chapman and Hall, Boca Raton.

[Desvillettes and Villani, 2000] Desvillettes, L. and Villani, C. (2000). On the spatially homogeneous Landau equation for hard potentials. II. *H*-theorem and applications. *Comm. Partial Differential Equations*, 25(1-2):261–298.

[Desvillettes and Villani, 2001] Desvillettes, L. and Villani, C. (2001). On the trend to global equilibrium in spatially inhomogeneous entropy-dissipating systems: the linear Fokker-Planck equation. *Comm. Pure Appl. Math.*, 54(1):1–42.

[Desvillettes and Villani, 2005] esvillettes, L. and Villani, C. (2005). On the trend to global equilibrium for spatially inhomogeneous kinetic systems: the Boltzmann equation. *Invent. Math.*, 159(2): 245–316.

[Eckmann and Hairer, 2003] Eckmann, J.-P. and Hairer, M. (2003). Spectral properties of hypoelliptic operators. *Comm. Math. Phys.*, 235(2):233–253.

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

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

[Gallay and Wayne, 2002] Gallay, T. and Wayne, C. E. (2002). Invariant manifolds and the long-time asymptotics of the Navier-Stokes and vorticity equations on $\mathbf{R}^2$. *Arch. Ration. Mech. Anal.*, 163(3):209–258.

[Gustafsson, 1986] Gustafsson, T. (1986). L^p-estimates for the nonlinear spatially homogeneous Boltzmann equation. *Arch. Rational Mech. Anal.*, 92(1):23–57.

[Gustafsson, 1988] Gustafsson, T. (1988). Global L^p-properties for the spatially homogeneous Boltzmann equation. *Arch. Rational Mech. Anal.*, 103:1–38.

[Helffer and Nier, 2005] Helffer, B. and Nier, F. (2005). *Hypoelliptic estimates and spectral theory for Fokker-Planck operators and Witten Laplacians*, volume 1862 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin.

[Hérau and Nier, 2004] Hérau, F. and Nier, F. (2004). Isotropic hypoellipticity and trend to equilibrium for the Fokker-Planck equation with a high-degree potential. *Arch. Ration. Mech. Anal.*, 171(2):151–218.

[Janvresse, 2001] Janvresse, E. (2001). Spectral gap for Kac's model of Boltzmann equation. *Ann. Probab.*, 29(1):288–304.

[Kac, 1956] Kac, M. (1956). Foundations of kinetic theory. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III*, pages 171–197, Berkeley and Los Angeles. University of California Press.

[Ledoux, 1995] Ledoux, M. (1995). L'algèbre de Lie des gradients itérés d'un générateur markovien – développements de moyennes et d'entropies. *Ann. Scient. Ec. norm. sup.*, 28:435–460.

[Lieb, 1982] Lieb, E. (1982). Comment on : "Approach to Equilibrium of a Boltzmann-Equation Solution". *Phys. Rev. Lett.*, 48(15):1057.

[Lieb, 1978] Lieb, E. H. (1978). Proof of an entropy conjecture of Wehrl. *Comm. Math. Phys.*, 62(1):35–41.

[Linnik, 1959] Linnik, Y. V. (1959). An information-theoretic proof of the central limit theorem with the Lindenberg condition. *Theory Probab. Appl.*, 4:288–299.

[Lions, 1994] Lions, P.-L. (1994). Compactness in Boltzmann's equation via Fourier integral operators and applications, I. *J. Math. Kyoto Univ.*, 34(2):391–427.

[Maslen, 2003] Maslen, D. K. (2003). The eigenvalues of Kac's master equation. *Math. Z.*, 243(2):291–331.

[Mattingly and Stuart, 2002] Mattingly, J. C. and Stuart, A. M. (2002). Geometric ergodicity of some hypo-elliptic diffusions for particle motions. *Markov Process. Related Fields*, 8(2):199–214. Inhomogeneous random systems (Cergy-Pontoise, 2001).

[Mattingly et al., 2002] Mattingly, J. C., Stuart, A. M., and Higham, D. J. (2002). Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232.

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

[Mischler and Wennberg, 1999] Mischler, S. and Wennberg, B. (1999). On the spatially homogeneous Boltzmann equation. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 16(4):467–501.

[Mouhot, 2006] Mouhot, C. (2006). Rate of convergence to equilibrium for the spatially homogeneous Boltzmann equation with hard potentials. *Comm. Math. Phys.*, 261(3):629–672.

[Mouhot and Strain, 2006] Mouhot, C. and Strain, R. (2006). Spectral gap and coercivity estimates for the linearized boltzmann collision operator without angular cutoff. Preprint.

[Mouhot and Villani, 2004] Mouhot, C. and Villani, C. (2004). Regularity theory for the spatially homogeneous Boltzmann equation with cut-off. *Arch. Ration. Mech. Anal.*, 173(2):169–212.

[Olaussen, 1982] Olaussen, K. (1982). Extension of the Boltzmann H theorem. *Phys. Rev. A*, 25(6):3393–3395.

[Plastino and Plastino, 1996] Plastino, A. R. and Plastino, A. (1996). Symmetries of the fokker-planck equation and the fisher-frieden arrow time. *Phys. Rev. E*, 54:4423.

[Pulvirenti and Wennberg, 1997] Pulvirenti, A. and Wennberg, B. (1997). A Maxwellian lower bound for solutions to the Boltzmann equation. *Comm. Math. Phys.*, 183:145–160.

[Rey-Bellet and Thomas, 2000] Rey-Bellet, L. and Thomas, L. E. (2000). Asymptotic behavior of thermal nonequilibrium steady states for a driven chain of anharmonic oscillators. *Comm. Math. Phys.*, 215(1):1–24.

[Shannon, 1948] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Tech. J.*, 27:379–423, 623–656.

[Shannon and Weaver, 1949] Shannon, C. E. and Weaver, W. (1949). *The Mathematical Theory of Communication*. The University of Illinois Press, Urbana, Ill.

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

[Talay, 2002] Talay, D. (2002). Stochastic Hamiltonian systems: exponential convergence to the invariant measure, and discretization by the implicit Euler scheme. *Markov Process. Related Fields*, 8(2):163–198. Inhomogeneous random systems (Cergy-Pontoise, 2001).

[Tanaka, 1973] Tanaka, H. (1973). An inequality for a functional of probability distributions and its application to Kac's one-dimensional model of a Maxwellian gas. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete*, 27:47–52.

[Tanaka, 1978] Tanaka, H. (1978). Probabilistic treatment of the Boltzmann equation of Maxwellian molecules. *Z. Wahrsch. Verw. Gebiete*, 46(1):67–105.

[Toscani, 1992a] Toscani, G. (1992a). Lyapunov functionals for a Maxwell gas. *Arch. Rational Mech. Anal.*, 119:301–307. *NOTA: The argument used to apply the main result in this paper to the Fisher information (a.k.a. Linnik functional) is wrong in dimension 3. See Villani (1998) for a (hopefully) correct proof.*

[Toscani, 1992b] Toscani, G. (1992b). New a priori estimates for the spatially homogeneous Boltzmann equation. *Cont. Mech. Thermodyn.*, 4:81–93.

[Toscani, 1999] Toscani, G. (1999). Entropy production and the rate of convergence to equilibrium for the Fokker-Planck equation. *Quart. Appl. Math.*, 57(3):521–541.

[Toscani and Villani, 1999a] Toscani, G. and Villani, C. (1999a). Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas. *J. Statist. Phys.*, 94(3–4):619–637.

[Toscani and Villani, 1999b] Toscani, G. and Villani, C. (1999b). Sharp entropy dissipation bounds and explicit rate of trend to equilibrium for the spatially homogeneous Boltzmann equation. *Comm. Math. Phys.*, 203(3):667–706.

[Villani, 1998] Villani, C. (1998). Fisher information bounds for Boltzmann's collision operator. *J. Math. Pures Appl.*, 77:821–837.

[Villani, 1999] Villani, C. (1999). Regularity estimates via the entropy dissipation for the spatially homogeneous Boltzmann equation without cut-off. *Rev. Mat. Iberoamericana*, 15(2):335–352.

[Villani, 2000a] Villani, C. (2000a). Contribution à l'étude mathématique des collisions en théorie cinétique. Master's thesis, Univ. Paris-Dauphine, France.

[Villani, 2000b] Villani, C. (2000b). Decrease of the Fisher information for solutions of the spatially homogeneous Landau equation with Maxwellian molecules. *Math. Models Methods Appl. Sci.*, 10(2):153–161.

[Villani, 2000c] Villani, C. (2000c). A short proof of the "concavity of entropy power". *IEEE Trans. Inform. Theory*, 46(4):1695–1696.

[Villani, 2002a] Villani, C. (2002a). Limites hydrodynamiques de l'équation de Boltzmann (d'après C. Bardos, F. Golse, C. D. Levermore, P.-L. Lions, N. Masmoudi, L. Saint-Raymond). *Astérisque*, (282):Exp. No. 893, ix, 365–405. Séminaire Bourbaki, Vol. 2000/2001.

[Villani, 2002b] Villani, C. (2002b). A review of mathematical topics in collisional kinetic theory. In *Handbook of mathematical fluid dynamics, Vol. I*, pages 71–305. North-Holland, Amsterdam.

[Villani, 2003] Villani, C. (2003). Cercignani's conjecture is sometimes true and always almost true. *Comm. Math. Phys.*, 234(3):455–490.

[Villani, 2006] Villani, C. (2006). Hypocoercivity. Preprint, available online via `www.umpa.ens-lyon.fr/~cvillani`.

[Wennberg, 1992] Wennberg, B. (1992). On an entropy dissipation inequality for the Boltzmann equation. *C.R. Acad. Sci. Paris, Série I,*, 315:1441–1446.

[Wennberg, 1993a] Wennberg, B. (1993a). Stability and exponential convergence in L^p for the spatially homogeneous Boltzmann equation. *Nonlinear Anal.*, 20(8):935–964.

[Wennberg, 1993b] Wennberg, B. (1993b). Stability and exponential convergence in L^p for the spatially homogeneous Boltzmann equation. *Nonlinear Anal.*, 20(8):935–964.

[Wennberg, 1994a] Wennberg, B. (1994a). On moments and uniqueness for solutions to the space homogeneous Boltzmann equation. *Transport Theory Statist. Phys.*, 24(4):533–539.

[Wennberg, 1994b] Wennberg, B. (1994b). Regularity in the Boltzmann equation and the Radon transform. *Comm. Partial Differential Equations*, 19(11-12):2057–2074.

[Wennberg, 1995] Wennberg, B. (1995). Stability and exponential convergence for the Boltzmann equation. *Arch. Rational Mech. Anal.*, 130(2):103–144.

[Wennberg, 1996] Wennberg, B. (1996). The Povzner inequality and moments in the Boltzmann equation. In *Proceedings of the VIII International Conference on Waves and Stability in Continuous Media, Part II (Palermo, 1995)*, number 45, part II, pages 673–681.

[Wennberg, 1997] Wennberg, B. (1997). Entropy dissipation and moment production for the Boltzmann equation. *J. Statist. Phys.*, 86(5–6):1053–1066.

APPENDIX

$$\frac{(M^f)''}{M^f} = \frac{1}{\rho} \operatorname{div}_x \operatorname{div}_x \int (f-M)v \otimes v \, dv + \frac{v-u}{\rho T} \cdot \operatorname{div}_x \operatorname{div}_x \int (f-M)v \otimes v \otimes v \, dv$$

$$-\frac{(v-u)}{\rho T} \cdot u \operatorname{div}_x \operatorname{div}_x \int (f-M)v \otimes v \, dv$$

$$-2\frac{v-u}{T} \cdot \frac{\operatorname{div}_x(\rho u)}{\rho^2} \operatorname{div}_x \left(\int (f-M)v \otimes v \, dv \right)$$

$$+2\frac{v-u}{T} \frac{u}{\rho^2} \operatorname{div}_x(\rho u) \operatorname{div}_x \left(\int (f-M) \, v \, dv \right)$$

$$+\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{1}{N\rho T} \operatorname{div}_x \operatorname{div}_x \left(\int (f-M)|v|^2 v \otimes v \, dv \right)$$

$$-\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{1}{N\rho T} (NT + |u|^2) \operatorname{div}_x \operatorname{div}_x \int (f-M)v \otimes v \, dv$$

$$-\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{4u}{N\rho^2 T} \operatorname{div}_x(\rho u) \left[\operatorname{div}_x \int (f-M)v \otimes v \, dv - u \operatorname{div}_x \int (f-M) \, dv \right]$$

$$-\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{2}{N\rho^2 T} \left(u \operatorname{div}_x(\rho u) - \int f v \otimes v \, dv \right)$$

$$\times \left[u \operatorname{div}_x \int (f-M) \, dv - \operatorname{div}_x \left(\int (f-M)v \otimes v \, dv \right) \right]$$

$$-\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{2u}{N\rho T} \operatorname{div}_x \operatorname{div}_x \int (f-M)v \otimes v \otimes v \, dv$$

$$+\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{u^2}{N\rho^2 T} \operatorname{div}_x \operatorname{div}_x \int (f-M)v \otimes v \, dv$$

$$+4\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{u}{N\rho^2 T} \operatorname{div}_x(\rho u) \operatorname{div}_x \int (f-M)v \otimes v \, dv$$

$$-4\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{u}{N\rho^2 T} \operatorname{div}_x(\rho u) \nabla_x \int (f-M) \, dv$$

$$-2\left(\frac{|v-u|^2}{2T} - \frac{N}{2}\right) \frac{\operatorname{div}_x(\rho u)}{N\rho^2 T} \operatorname{div}_x \int (f-M)|v|^2 v \, dv$$

$$+2\Big(\frac{|v-u|^2}{2T}-\frac{N}{2}\Big)\frac{|u|^2+NT}{N\rho^2 T}\operatorname{div}_x(\rho u)\operatorname{div}_x\int(f-M)v\,dv$$

$$-\Big(\frac{|v-u|^2}{2T}-\frac{N}{2}\Big)\frac{2u}{N^2\rho^2 T}\cdot\operatorname{div}_x\int(f-M)v\otimes v\,dv$$

$$+\Big(\frac{|v-u|^2}{2T}-\frac{N}{2}\Big)\frac{2|u|^2}{N^2\rho^2 T}\operatorname{div}_x\int(f-M)v\,dv$$

$$-\Big[-\frac{1}{\rho^2 T}\operatorname{div}_x\int fv\otimes v\,dv+\frac{u}{\rho^2 T}\operatorname{div}_x(\rho u)\Big]\Big(u\operatorname{div}_x\int(f-M)\,dv-\operatorname{div}_x\int(f-M)v\otimes v\,dv\Big)$$

$$-\frac{|v-u|^2}{2T^3}\Big(-\frac{1}{\rho N}\operatorname{div}_x\int f|v|^2v\,dv+\frac{(|u|^2+NT)}{\rho N}\operatorname{div}_x(\rho u)+\frac{2u}{\rho N}\cdot\operatorname{div}_x\int fv\otimes v\,dv-2\frac{|u|^2}{\rho N}\nabla_x\rho\Big)$$

$$\times\Big(-\frac{1}{\rho N}\operatorname{div}_x\int(f-M)|v|^2v\,dv+\frac{|u|^2+NT}{\rho N}\operatorname{div}_x\int(f-M)v\,dv$$

$$+\frac{2u}{\rho N}\cdot\operatorname{div}_x\int(f-M)v\otimes v\,dv-\frac{2|u|^2}{\rho N}\operatorname{div}_x\int(f-M)\,dv\Big)$$

$$+\Big(\frac{|v-u|^2}{2T}-\frac{N+4}{2}\Big)\Big(-\frac{1}{\rho NT}\operatorname{div}_x\int f|v|^2v\,dv+\Big(\frac{|u|^2+NT}{\rho NT}\Big)\operatorname{div}_x(\rho u)$$

$$+\frac{2u}{\rho NT}\cdot\operatorname{div}_x\int fv\otimes v\,dv-2\frac{|u|^2}{\rho NT}\nabla_x\rho\Big)$$

$$\times\frac{v-u}{\rho T}\Big[u\operatorname{div}_x\int(f-M)\,dv-\operatorname{div}_x\int(f-M)v\otimes v\,dv\Big]$$

$$+\frac{v-u}{\rho T}\cdot\Big[u\operatorname{div}_x(\rho u)-\operatorname{div}_x\int fv\otimes v\,dv\Big]$$

$$\times\frac{v-u}{\rho T}\Big[u\operatorname{div}_x\int(f-M)\,dv-\operatorname{div}_x\int(f-M)v\otimes v\,dv\Big]$$

$$-2\frac{v-u}{\rho^2 T}\cdot\Big[u\operatorname{div}_x(\rho u)-\operatorname{div}_x\int fv\otimes v\,dv\Big]\operatorname{div}_x\int(f-M)v\,dv$$

$$-\frac{2}{\rho T}\Big(\frac{|v-u|^2}{2T}-\frac{N}{2}\Big)\Big[-\frac{1}{\rho N}\operatorname{div}_x\int f|v|^2v\,dv+\frac{|u|^2+NT}{\rho N}\operatorname{div}_x(\rho u)$$

$$+\frac{2u}{\rho N}\cdot\operatorname{div}_x\int fv\otimes v\,dv-\frac{2|u|^2}{\rho N}\operatorname{div}_x(\rho u)\Big]\operatorname{div}_x\int(f-M)v\,dv.$$