Riv. Mat. Univ. Parma (7) **2*** (2003), 101-149

E. A. CARLEN and M. C. CARVALHO (*)

Probabilistic methods in kinetic theory (**)

Contents

1 - Introduction

Statistical mechanics is an inherently probabilistic subject, as the name suggests. Yet, in many areas of non equilibrium statistical mechanics, and kinetic theory in particular, probabilistic reasoning does not usually play much of a role in the analysis of bulk behavior.

(*) E. A. Carlen: School of Mathematics, Georgia Institute of Technology, Atlanta, GA, 30332 U.S.A., carlen@math.gatech.edu; M. C. Carvalho: School of Mathematics, Georgia Institute of Technology, Atlanta, GA, 30332 U.S.A., on leave from Department of Mathematics, University of Lisbon, Av. Prof. Gama Pinto 2, 1649-003 Lisbon, Portugal, mcarvalh@cii.fc.ul.pt

(**) Received October 1st, 2003. AMS Classification 82 C 40.

Work partially supported by U.S. N.S.F. grant DMS 00-70589, by PRAXIS XXI and by TMR ERB-FMRX CT97 0157.

This is not altogether surprising. A great deal of the power of the methodology developed by the founders of statistical mechanics lies in the exploitation of the law of large numbers to eliminate a complicated stochastic description in favor of a description in terms of simple local averages that obey deterministic evolution equations. Thus, while randomness of some sort enters on the microscopic scale, on larger scales it has been eliminated.

In the kinetic theory of gases, the evolution is described on the mesoscopic scales by the Boltzmann equation, while on the macroscopic it is described by the Euler or Navier-Stokes equations.

Although these are deterministic equations, one might hope that their probabilistic origins make them amenable to analysis by probabilistic means. Indeed, consider the case of the Schrödinger equation. There is no underlying stochastic model here, yet the mere presence of the Laplacian opens the way to a very effective analysis in terms of Wiener integrals. Might it also be possible to use probabilistic methods to analyze the Boltzmann equation, which does, after all, describe a physical phenomenon that is probabilistic at the microscopic level?

This question is not new, and much work has been done to validate the answer «yes», particularly by Kac and McKean. In what follows, we will describe some of their ideas and results, along with more recent results, and some open problems.

The problems examined here in detail concern the spatially homogeneous Boltzmann equation. Probabilistic methods are particularly effective for studying the effects of collisions on the the distribution of velocities. There is a large literature in which probabilistic methods are applied to study the spatially inhomogeneous Boltzmann equation, particularly by Méléard and her collaborators, who have used a Malliavin calculus for jump processes to prove mild regularity results for quite singular (non cutoff – see below) Boltzmann equations, and Rezakahnlou, who has obtained results on propagation of chaos (see below) in the spatially inhomogeneous setting. The methods and goals of these authors are rather different than the ones discussed here, where our focus is on quantitative questions pertaining to the rate of relaxation to equilibrium.

2 - The central limit theorem and the Wild sum

2.1 - *Maxwellian molecules and the Wild sum*

The spatially homogeneous Boltzmann equation governs the time evolution of the probability density function $f(v, t)$ for the velocities of the molecules in a dilu-

te gas. Thus $f(\cdot, v)$ is a time dependent probability measure on $\boldsymbol{R}^3$, and the equation takes the form

$$\frac{\partial}{\partial t} f(v, t) = Q(f(\cdot, v), f(\cdot, v)), \tag{2.1}$$

where Q is given by

$$Q(f, g)(v) = \iint_{\boldsymbol{R}^3 \times \boldsymbol{S}^2} B\left(\left\langle \frac{v - v_*}{|v - v_*|}, \sigma \right\rangle, |v - v_*|\right) [f(v') g(v_*') - f(v) g(v_*)] \, d\sigma \, dv_*.$$

and $\langle \cdot, \cdot \rangle$ denotes the inner product in $\boldsymbol{R}^3$. This describes the effect of a binary collision between two identical molecules with *pre-collisional velocities* v and v_*, and *post-collisional velocities* v' and v_*'. It is assumed that the collisions conserve momentum and kinetic energy, so that

$$v' + v_*' = v + v_*, \quad |v'|^2 + |v_*'|^2 = |v|^2 + |v_*|^2. \tag{2.2}$$

All possible pairs of post collisional velocities – that is, pairs satisfying these constraints – may be parameterized by a vector in $\boldsymbol{S}^2$.

One such parametrization is

$$v' = \frac{v + v_*}{2} + \frac{|v - v_*|}{2} \sigma, \quad v_*' = \frac{v + v_*}{2} - \frac{|v - v_*|}{2} \sigma, \quad \sigma \in \boldsymbol{S}^2, \tag{2.3}$$

and another is

$$v' = v + (\langle v_* - v, \sigma \rangle) \, \sigma \quad v_*' = v_* - (\langle v_* - v, \sigma \rangle) \, \sigma \quad \sigma \in \boldsymbol{S}^2. \tag{2.4}$$

The function B specifies the rate at which collisions with parameter σ occur when the incoming velocities are v and v_*, and as indicated, it depends only on the magnitude of $v - v_*$ and the angle between $v - v_*$ and σ. It is called the *collision kernel*, and it is a non negative Borel function of its arguments. The specific form depends on which of the two parameterizations are being employed.

In what follows we shall impose two conditions on B. First, we specialize to the case of so-called *Maxwellian molecules*; i.e., we assume that the kernel B does not depend on $|v - v_*|$:

$$B\left(\left\langle \frac{v - v_*}{|v - v_*|}, \sigma \right\rangle, |v - v_*|\right) = B\left(\left\langle \frac{v - v_*}{|v - v_*|}, \sigma \right\rangle\right). \tag{2.5}$$

The condition (2.5) holds if there is a $1/r^5$ force law between the molecules. Maxwell guessed that nature would opt for a special simple force law as it often does, and posited such an interaction between the molecules. In retrospect, this may

look overly bold. On the other hand, there is a lot of wisdom to the old maxim saying that if one drops ones keys at night, one should look for them first under the street lamp. It is easier to see things through to the end with Maxwellian molecules, and it is still useful, even today, to look first at the Maxwellian case when investigating any new problem. For more background, see [**18**], [**35**].

Second, we impose the *angular cutoff condition*, which is

$$\int_0^\pi B(\cos(\theta))\sin(\theta)\,d\theta < \infty\ . \tag{2.6}$$

Under the angular cutoff condition, the collision integral operator $Q(f, g)$ can be split into its so-called *gain* and *loss* terms:

$$Q(f, g)(v) = Q^+(f, g)(v) - Q^-(f, g)(v) \tag{2.7}$$

where

$$Q^+(f, g)(v) = \iint_{\boldsymbol{R}^3\times\boldsymbol{S}^2} B\left(\left\langle \frac{v - v_*}{|v - v_*|}\ ,\ \sigma\right\rangle\right) f(v')\, g(v_*')\, d\sigma\, dv_*,$$

$$Q^-(f, g)(v) = f(v)\int_{\boldsymbol{R}^3}\left[\int_{\boldsymbol{S}^2} B\left(\left\langle \frac{v - v_*}{|v - v_*|}\ ,\ \sigma\right\rangle\right) d\sigma\right] g(v_*)\, dv_*.$$

Without the cutoff condition, the cancellations in $Q(f, g)$ are crucial, and solutions to (2.1) have only been shown to exist and studied in a certain weak form [**4**], [**22**], [**40**], [**3**] and [**2**]. With the cutoff condition, it is possible to produce strong solutions that can be studied in greater detail.

Having both the cut-off and Maxwellian molecules, one can actually give a construction of the solutions that permits one to analyze them in detail. This construction has an interesting probabilistic interpretation, discovered by McKean, and we shall now explain what it is, and how it may be used.

The story begins with a paper of Wild [**39**]. Let the *Wild convolution* $f\circ g$ of two integrable functions f and g on $\boldsymbol{R}^3$ be defined by

$$f\circ g = Q^+(f, g)\,. \tag{2.8}$$

From what has been said above, under the condition that B is even, the Wild convolution is commutative. However, it is not associative under any natural condition on B. *Therefore, we make the blanket assumption, unless otherwise stated, that the kernel* $B(\cdot)$ *is a nonnegative, even function in* $L^1(\boldsymbol{R})$ *with* $\operatorname{supp} B \subset [-1, 1]$.

We make a further blanket assumption – normalizing out units of time – that

$$\int_{S^2} B(\langle \cdot, \sigma \rangle)\, d\sigma = 2\pi \int_0^{\pi} B(\cos(\theta)) \sin(\theta)\, d\theta = 2\pi \|B\|_{L^1[-1,1]} = 1 . \tag{2.9}$$

In this case,

$$Q^-(f, g) = \left(\int_{\boldsymbol{R}^3} f(w)\, \mathrm{d}w \right) g(v) . \tag{2.10}$$

Using (2.7), (2.8) and (2.10), we can rewrite (2.1) in the form

$$\frac{\partial}{\partial t} f(\cdot, v)(v) = (f(\cdot, v) \circ f(\cdot, v))(v) - f(\cdot, v)(v) . \tag{2.11}$$

Let F denote the initial data so that

$$f_t \,|_{t=0} = F . \tag{2.12}$$

Then the Cauchy problem specified by (2.11) and (2.12) is equivalent to

$$f(v, t) = e^{-t} F(v) + \int_0^t e^{-(t-s)} f \circ f(v, s)\, \mathrm{d}s .$$

We therefore fix the initial data F and define a function Φ from $\mathcal{C}(\boldsymbol{R}, L^1(\boldsymbol{R}^3))$ to itself by

$$\Phi(f) = e^{-t} F(\cdot) + \int_0^t e^{-(t-s)} f \circ f(\cdot, s)\, \mathrm{d}s .$$

Then $f \in \mathcal{C}(\boldsymbol{R}, L^1(\boldsymbol{R}^3))$ is a solution of our Cauchy problem if and only if

$$\Phi(f) = f .$$

As Wild pointed out, this fixed point problem may be solved by iteration: put $f_{(0)} = 0$, and then define

$$f_{(j+1)} = \Phi(f_{(j)}) \quad \text{for all} \quad j \geqslant 1 . \tag{2.13}$$

This yields:

$$
\begin{aligned}
f_{(1)}(t) &= e^{-t}F \\
f_{(2)}(t) &= e^{-t}F + e^{-t}(1-e^{-t})\,F\circ F \\
f_{(3)}(t) &= e^{-t}F + e^{-t}(1-e^{-t})\,F\circ F + \\
&\qquad e^{-t}(1-e^{-t})^2((1/2)\,F\circ(F\circ F) + (1/2)(F\circ F)\circ F)
\end{aligned}
\tag{2.14}
$$

and so on.

As one sees in these examples, $f_{(n)}(t) - f_{(n-1)}(t)$ is positive, so that $f_{(n)}(t) \geq f_{(n-1)}(t)$. To see this more generally, introduce the sequence of probability densities $Q_1^+(F)$, $n \geq 1$, defined recursively by

$$Q_1^+(F) = F$$

and

$$Q_n^+(F) = \frac{1}{n-1}\sum_{k=1}^{n-1} Q_k^+(F)\circ Q_{n-k}^+(F), \qquad n \geq 2\,. \tag{2.15}$$

Then one easily checks that

$$Q_2^+(F) = F\circ F$$

$$Q_3^+(F) = \frac{1}{2}\,(F\circ(F\circ F) + (F\circ F)\circ F)\,.$$

We can then rewrite (2.14) as

$$
\begin{aligned}
f_{(1)}(t) &= e^{-t}Q_1^+(F) \\
f_{(2)}(t) &= e^{-t}Q_1^+(F) + e^{-t}(1-e^{-t})\,Q_2^+(F) \\
f_{(3)}(t) &= e^{-t}Q_1^+(F) + e^{-t}(1-e^{-t})\,Q_2^+(F) + e^{-t}(1-e^{-t})^2 Q_3^+(F)\,.
\end{aligned}
\tag{2.16}
$$

From here you guess the pattern:

$$f_{(n)}(t) = \sum_{k=1}^{n} e^{-t}(1-e^{-t})^{k-1} Q_k^+(F)\,. \tag{2.17}$$

Notice that

$$\sum_{k=1}^{\infty} e^{-t}(1-e^{-t})^{k-1} = 1\,. \tag{2.18}$$

Therefore, an easy application of the monotone convergence theorem leads to

Wild's result [**39**] that

$$\lim_{n\to\infty} f_{(n)}(t) = f(t) \tag{2.19}$$

exists and is a solution of (1.8). This gives us the explicit formula

$$f(\cdot, v) = \sum_{n=1}^{\infty} e^{-t}(1 - e^{-t})^{n-1} Q_n^+(F) \tag{2.20}$$

for the unique solution of our Cauchy problem.

This formula reduces the study of the time evolution $t \mapsto f(\cdot, t)$ to the study of the sequence $n \mapsto Q_n^+(F)$. As we shall show, one can obtain fairly precise information on how the large n behavior of $n \mapsto Q_n^+(F)$ depends on F, and this is a very effective route to information on how the large t behavior of $t \mapsto f(\cdot, t)$ depends on f. Moreover, *probabilistic tools* are natural for analyzing the large n behavior of $n \mapsto Q_n^+(F)$.

For example, let M_F denote the Maxwellian density with the same mean and variance as F. We recall that the Maxwellian densities are those Gaussian densities that have isotropic covariance. We henceforth assume, without loss of generality, that F has zero mean and unit variance, so that

$$M_F(v) = \left(\frac{1}{2\pi}\right)^{3/2} e^{-|v|^2/2}.$$

Where convenient, we simply write M in place of M_F.

2.2 - *Maxwellian molecules and the Central Limit Theorem*

McKean conjectured [**28**] an analog of the Central Limit Theorem for Maxwellian molecules. More specifically, he conjectured that

$$\lim_{n\to\infty} Q_n^+(F) = M_F. \tag{2.21}$$

To clearly see the connection with the Central Limit Theorem, let $\{V_j\}$ be a sequence of independent identically distributed random variables with values on $\boldsymbol{R}^3$. Suppose that they are zero mean with unit variance, and suppose that for each j and each Borel set $A \subset \boldsymbol{R}^3$,

$$\Pr(V_j) = \int_A F(v)\, dv\,.$$

Define the *normalized convolution* $F \star G$ of two probability densities F and G on

$\boldsymbol{R}^3$ by

$$F \star G(v) = \int_{\boldsymbol{R}^3} F\left(\frac{v+w}{\sqrt{2}}\right) G\left(\frac{v-w}{\sqrt{2}}\right) \mathrm{d}w\,.$$

This is simply the normal convolution of F and G after rescaling to preserve the variance. In particular $F \star F$ is the probability density of $\dfrac{V_1 + V_2}{\sqrt{2}}$, and $(F \star F) \star (F \star F)$ is the probability density of

$$\frac{V_1 + V_2 + V_3 + V_4}{\sqrt{4}}$$

and so on. If one defines a sequence of probability densities $F_{(n)}$ recursively by $F_{(1)} = F$ and

$$F_{(n+1)} = F_{(n)} \star F_{(n)}, \tag{2.22}$$

then $F_{(n+1)}$ is the probability density of

$$\frac{\sum_{k=1}^{2^n} V_k}{2^{n/2}}\,.$$

Therefore, the Central Limit Theorem tells us that

$$\lim_{n \to \infty} F_{(n)} = M_F.$$

The Wild convolution $F \circ F$ is somewhat more complicated than the normalized convolution $F \star F$. Nonetheless, there is a strong resemblance. Also, the recursion (2.15) is more complicated than the recursion (2.22). Still, it is tempting to conjecture that (2.21) holds, for then we would have from (2.20) that

$$\lim_{t \to \infty} f(\cdot, t) = M_F(\cdot)\,,$$

and moreover, we can relate the rates at which $Q_n^+(F)$ tends to M_F and at which $f(\cdot, t)$ tends to M_F.

The following demonstration of this is taken from [**13**], [**14**]. We have from (2.18) that

$$M_F = \sum_{k=1}^{\infty} e^{-t}(1 - e^{-t})^{k-1} M_F$$

so that in any norm $\|\cdot\|$, and any positive integer N,

$$\begin{aligned}\|f(\cdot, t) - M_F\| &\leq \sum_{k=1}^{\infty} e^{-t}(1-e^{-t})^{k-1}\|Q_k^+(F) - M_F\| \\ &\leq \sum_{k=1}^{N-1} e^{-t}(1-e^{-t})^{k-1}\|Q_k^+(F) - M_F\| \\ &\quad + \sum_{k=N}^{\infty} e^{-t}(1-e^{-t})^{k-1}\|Q_k^+(F) - M_F\|.\end{aligned}$$

Suppose that $\lim_{n\to\infty}\|Q_k^+(F) - M_F\| = 0$, and that in fact,

$$\|Q_k^+(F) - M_F\| \leq Ak^{-p} \tag{2.23}$$

for some $1 > p > 0$ and $A < \infty$.

Then

$$\begin{aligned}\sum_{k=1}^{N-1} e^{-t}(1-e^{-t})^{k-1}\|Q_k^+(F) - M_F\| &\leq e^{-t}A\sum_{k=1}^{N-1} k^{-p} \\ &\leq e^{-t}\frac{A}{1-p}N^{1-p}.\end{aligned}$$

Also,

$$\begin{aligned}\sum_{k=N}^{\infty} e^{-t}(1-e^{-t})^{k-1}\|Q_k^+(F) - M_F\| &\leq e^{-t}\sum_{k=N}^{\infty} e^{-t}Ak^{-p} \\ &\leq AN^{-p}.\end{aligned}$$

We therefore have that for all N,

$$\|f(\cdot, t) - M_F\| \leq A((1-p)^{-1}e^{-t}N^{1-p} + N^{-p}).$$

Choosing $N = e^t$ we see that for all $t > 0$,

$$\|f(\cdot, t) - M_F\| \leq A\frac{2-p}{1-p}e^{-pt}. \tag{2.24}$$

We record this as a theorem:

Theorem 2.1 ([13]). *Let $\|\cdot\|$ be any norm such that for all initial data F,*

$$\|Q_k^+(F) - M_F\| \leq Ak^{-p}$$

for some $1 > p > 0$ and $A < \infty$. Then for the same A and p,

$$\|f(\cdot, t) - M_F\| \leq A\frac{2-p}{1-p}e^{-pt}.$$

That is, algebraic decay of $\|Q_k^+(F) - M_F\|$ at rate k^{-p} implies exponential decay of $\|f(\cdot, t) - M_F\|$ at rate e^{-pt} for the same p. Since exponential decay in t of $\|f(\cdot, t) - M_F\|$ is the best that we can hope for, algebraic decay in n of $\|Q_k^+(F) - M_F\|$ is the best we can hope for. This is reasonable, as in general the distribution of

$$\frac{\sum_{k=1}^{n} V_k}{\sqrt{N}}$$

only approaches the normal distribution at an algebraic rate; see [**20**].

The question now is: how can one prove an estimate of the type (2.23) – that is, a quantitative version of the Central Limit Theorem for Maxwellian molecules?

The proof given in [**13**] uses explicit representation for $Q_n^+(F)$ in terms of a random walk on certain tree graphs that was invented by McKean [**29**] that we now explain.

2.3 - *McKean's walk on tree graphs*

Observe that $Q_n^+(F)$ will be a weighted sum of n-fold Wild convolutions of F. Since the Wild convolution is not associative [**28**], and so there are many different n-fold Wild convolutions in general. To see this, consider two such convolutions for $n = 8$:

$$(((F \circ F) \circ (F \circ F)) \circ ((F \circ F) \circ (F \circ F))) \tag{2.25}$$

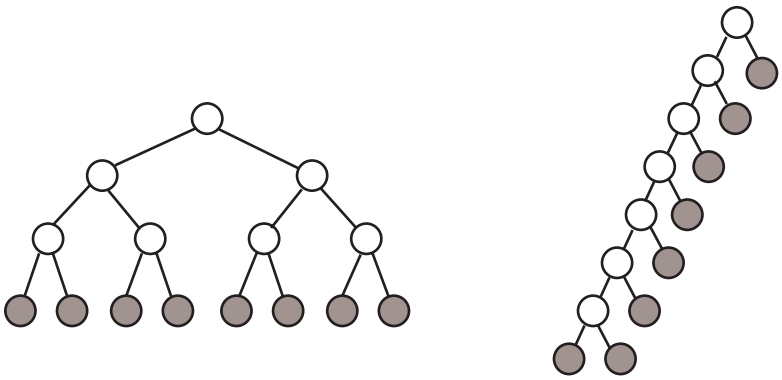
and

$$(((((((F \circ F) \circ F) \circ F) \circ F) \circ F) \circ F) \circ F). \tag{2.26}$$

The arrangement of the parentheses matters because the Wild convolution is not associative.

We can enumerate the different n-fold Wild convolutions easily by identifying them with certain graphs. Let $\Gamma(n)$ denote the set of all binary tree graphs with n leaves and $n-1$ nodes in which each node has either two «children» or no «children». *Single «children» are not allowed.* We refer to $\Gamma = \bigcup_{n=2}^{\infty} \Gamma(n)$ as the set of *McKean graphs*. Drawing the graph so that the two «children» of any node can be identified, respectively, as the «left child» and the «right child» provides a natural left to right order on the nodes and leaves in any given «generation». The

particular McKean graphs corresponding to the two convolutions in (2.25) and (2.26) are

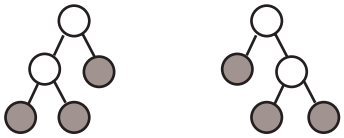


where the graph on the left corresponds to (2.25). Both graphs have 8 leaves and 7 nodes. The key difference lies in the number of generations separating the leaves from the node at the top – the root; i.e. in the depth of the leaves.

The correspondence between graphs γ and convolutions of F is the following: Pick any $\gamma \in \Gamma(n)$, and write «F» in each of the leaves. Now find the left-most pair of leaves in the deepest row. Erase this pair of leaves, which makes the former parent node a leaf, and write down $(F \circ F)$ in the new leaf, producing a graph in $\Gamma(n-1)$. Then repeat the procedure, working our way back to the root, but with the difference that at the later steps, we will be erasing pairs of leaves that have various iterated convolutions of F written in them. *If the left leaf we are erasing has G_ℓ written in it, and the right leaf has G_r written in it, put $(G_\ell \circ G_r)$ in the new leaf created upon erasure.* Once one has done this until only the root is left, it has written in it some n-fold iterated convolution of F, and that is the convolution $C_\gamma(F)$ corresponding to the graph $\gamma \in \Gamma(n)$ that we started with. It is easy to check the two graphs in the diagram above do in fact lead to (2.25) and (2.26) as claimed. Henceforth, we shall write $C_\gamma(F)$ to denote the convolution corresponding to γ.

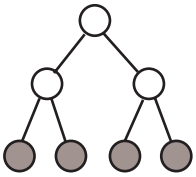
We now introduce the *McKean walk* which is a random walk on Γ that passes through $\Gamma(n)$ at the n-th step. The McKean walk starts at the unique element γ_0 in $\Gamma(2)$.

Pick a leaf at random – uniformly – and attach another copy of γ_0. This produces three leaves, and we get one of the two graphs in $\Gamma(3)$.



Clearly, each one is equally likely.

Next, pick one of the three leaves at random, and again attach a copy of γ_0. There are 5 elements in $\Gamma(4)$, and this time they are not all equally likely: the graph



is produced with probability 1/3, while each of the others is produced with probability 1/6. This process is repeated until n leaves are produced. We call this random walk on McKean graphs the *McKean process*.

Let γ_n denote the state of the McKean walk at time n, $n \geq 2$. For any $\gamma \in \Gamma(n)$, let

$$p_n(\gamma) = \Pr(\gamma_n = \gamma).$$

McKean's expression for $Q_n^+(F)$ is

$$Q_n^+(F) = e^{-t}(1 - e^{-t})^{n-1}\Big[\sum_{\gamma \in \Gamma(n)} p_n(\gamma)\, C_\gamma(F)\Big] \tag{2.27}$$

where the $p_n(\gamma)$ are probabilities

$$\sum_{\gamma \in \Gamma(n)} p_n(\gamma) = 1 \tag{2.28}$$

expressing the relative frequency with which any particular convolution $C_\gamma(F)$ contributes to $Q_n^+(F)$. McKean proved the following:

McKean's Lemma [**29**]. *Let γ be any member of $\Gamma(n)$. Then $p_n(\gamma)$ in* (2.27) *is the probability that the McKean process passes through γ at time $n-2$.*

We refer to the number of generations separating a leaf from the root as the *depth* of the leaf. In the graph corresponding to (2.25), all of the leaves have the same depth, namely 3. In the graph corresponding to (2.26) there are two leaves of depth 7, and one each of every depth from 1 to 6. The depths of the leaves will be crucial in our analysis.

The point is that the depth of a leaf represents the number of collisions that the particle corresponding to the leaf experiences before it contributes to the probability distribution for the velocity of the observed particle. If all of the contributing leaves have undergone many collisions, then they are all almost at equilibrium, and we will expect $C_\gamma(F) \approx M_F$.

On the other hand if there is a leaf that contributes directly to $C_\gamma(F)$, as is the case in (2.26), then there is no reason to expect $C_\gamma(F) \approx M_F$. That is, if we try to prove that $Q_h^+(F) \approx M_F$, there will be «good» and «bad» graphs in the sum (2.27). The good graphs in $\Gamma(n)$ will be those in which all of the leaves have a depth that is not too far from $\log_2(n)$, and the bad ones will be those in which there are leaves of low depth. (Notice that if $n = 2^k$, and each leaf has the same depth, as in (2.26), then the depth is k, which is where the $\log_2(n)$ comes from).

We now introduce a quantitative measure of how good (or bad) a particular graph $\gamma \in \Gamma(n)$ is: the following function is small when all of the leaves are deep, and is large when any one is shallow:

Definition. *Fix any constant $0 < c < 1$. Then for all n and all $\gamma \in \Gamma(n)$ define*

$$W(\gamma) = \sum_{j=1}^{n} \left(\frac{c}{2} \right)^{d(j)}$$

where $d(j)$ denotes the depth of the j-th leaf of γ from the left and define

$$T(n) = \sum_{\gamma \in \Gamma(n)} p_n(\gamma)\, W(\gamma)\,.$$

If $n = 2^k$, and γ_{best} is the «best possible» graph, each of the leaves of γ has

depth k, then

$$W(\gamma_{\text{best}}) = 2^k \left(\frac{c}{2} \right)^k = c^k$$

which is small for k large since $c < 1$.

On the other hand, let γ_{worst} has even a single leaf of depth 2 for example, then $W(\gamma_{\text{worst}}) > (c/2)^2$ no matter how big k is.

The quantity $T(n)$ is the expected value of W:

$$T(n) = \mathrm{E} W(\gamma(n)) .$$

The closer $T(n)$ is to c^k, the more typical good graphs are, and the closer $T(n)$ $(c/2)^2$, the more typical bad graphs are. The following theorem says that good graphs are fairly typical.

Theorem 2.2 ([13]). *There is a finite constant A such that*

$$T(n) \leqslant A b^{\ln n} \quad \text{for all } n \geqslant 2$$

where b is any number satisfying $b > e^{c-1}$. Furthermore, the constant A depends only on the choice of b, and is explicitly computable.

The identification of $p_n(\gamma)$ as the probability that a random walk on McKean graphs hits the particular graph $\gamma \in \Gamma(n)$ – McKean's Lemma – allows us to use (2.27) and Theorem 2.2 to estimate $\| Q_+^n(F) - M_F \|$ in appropriate norms, as we now explain.

First, note that it follows from (2.27) that for any convex function ϕ on $L^1(\boldsymbol{R}^3)$,

$$\begin{aligned} \phi(Q_+^n(F)) &= \phi\left(\sum_{\gamma \in \Gamma(n)} p_n(\gamma)(C_\gamma(F)) \right) \\ &= \sum_{\gamma \in \Gamma(n)} p_n \phi(C_\gamma(F)) . \end{aligned} \tag{2.29}$$

Now suppose that we can find a convex function ϕ on $L^1(\boldsymbol{R}^3)$ such that for some constant $c < 1$

$$\phi(F \circ G) \leqslant \frac{c}{2} (\phi(F) + \phi(G)) . \tag{2.30}$$

Then taking apart any McKean graph from the «top down»; i.e., starting at the root, one obtains:

Theorem 2.3 ([13]). *Let ϕ be any convex function on $L^1(\boldsymbol{R}^3)$ such that for some constant $c<1$,*

$$\phi(F\circ G)\leqslant \frac{c}{2}\,(\phi(F)+\phi(G)) \tag{2.31}$$

for all probability densities F and G with mean zero and unit variance. Then

$$\phi(C_\gamma(F))\leqslant W(\gamma)\,\phi(F)\,.$$

Proof of the Lemma. Consider the example $\phi(F\circ(F\circ F))$. Using (2.30) twice we have

$$\phi(F\circ(F\circ F))\leqslant$$

$$\frac{c}{2}\,\phi(F)+\frac{c}{2}\,\phi(F\circ F)\leqslant$$

$$\frac{c}{2}\,\phi(F)+\frac{c}{2}\left(\frac{c}{2}\,\phi(F)+\frac{c}{2}\,\phi(F)\right).$$

The general case follows by a simple induction argument based on this pattern. ■

Combining the lemma, (2.29) and Theorem 2.2, we have that

$$\phi(Q_+^n(F))\leqslant Ab^{\ln(n)}\phi(F)\,.$$

Therefore, if we can a norm $\|\cdot\|$ and a convex function ϕ such that (2.31) holds, and also

$$\|F-M_F\|\leqslant\phi(F)\,,$$

then we have

$$\|Q_+^n(F)-M_F\|\leqslant Ab^{\ln(n)}\phi(F)\,.$$

2.4 - *A convolution-convex functional ϕ.*

Definition. We say that a functional ϕ on $L^1(\boldsymbol{R}^3)$ is *strictly convolution-convex* in case (2.30) holds with some $c<1$ for all probability densities F and G of the same mean and variance.

We now see how to find such functionals. It would have been nice if we could use $\phi(F) = \|F - M\|_{L^1}$ but unfortunately, while this is convex in the usual sense, which we also require, it is not convolution-convex.

To devise an appropriate functional, we recall some norms recently introduced by Gabetta, Toscani and Wennberg [**21**].

Definition. *Let $\alpha > 0$, and then for any integrable function H on $\boldsymbol{R}^3$ define*

$$|||H|||_\alpha = \sup_{\xi \in \boldsymbol{R}^3} \frac{|\widehat{H}(\xi)|}{|\xi|^{2+\alpha}}$$

where $\widehat{H}(\xi)$ denotes the Fourier transform of $H(v)$.

If F and G are two probability densities with the same first and second moments, then by a simple Taylor expansion,

$$\frac{|\widehat{F}(k) - \widehat{G}(k)|}{|k|^2}$$

is bounded. If furthermore, for some $\delta > 0$, both

$$\int_{\boldsymbol{R}^3} |v|^{2+\delta} F(v)\, dv < \infty \quad \text{and} \quad \int_{\boldsymbol{R}^3} |v|^{2+\delta} F(v)\, dv < \infty .$$

Then estimating the remainder in the Taylor expansion, we have that with $\alpha = \delta/(1+\delta)$,

$$\frac{|\widehat{F}(k) - \widehat{G}(k)|}{|k|^{2+\alpha}}$$

is bounded.

Therefore, if F is an isotropic probability density such that $\int_{\boldsymbol{R}^3} |v|^{2+\delta} F(v)\, dv < \infty$, then $|||F - M_F|||_\alpha < \infty$.

In this case, we define the functional

$$\phi(F) = |||F - M_F|||_\alpha .$$

It is clear that ϕ is convex. We now show that is also convolution-convex. In fact, the following lemma shows something more:

Lemma ([**13**]). *Consider any four integrable functions* F_1, F_2, G_1 *and* G_2 *such that their Fourier transforms* $\widehat{F}_1$ $\widehat{F}_2$, $\widehat{G}_1$ *and* $\widehat{G}_2$ *satisfy*

$$\|\widehat{F}_i\|_\infty, \quad \|\widehat{G}_i\|_\infty \leqslant 1+\varepsilon \qquad \text{for } i=1,2$$

where $(1+\varepsilon)\, c_\alpha = c < 1$ *and*

$$c_\alpha = \int |\cos(\theta/2)|^{2+\alpha} B(\theta)\cos(\theta)\,\mathrm{d}\theta\,.$$

Then

$$|||F_1\circ G_1 - F_2\circ G_2|||_\alpha \leqslant \frac{c_\alpha}{2}\,(|||F_1-F_2|||_\alpha + |||G_1-G_2|||_\alpha). \tag{2.32}$$

Before giving the proof, we note that if we choose $F_1 = F$, $G_1 = G$ and $F_2 = G_2 = M$, then since $M\circ M = M$ for any Maxwellian, the inequality (2.32) reduces to

$$|||F\circ G - M|||_\alpha \leqslant \frac{c_\alpha}{2}\,(|||F-M|||_\alpha + |||G-M|||_\alpha) \tag{2.33}$$

and this is the desired convolution convexity.

Proof. We start with the formula for the Fourier transform of $F\circ G$, first worked out by Bobylev [**7**]:

$$\widehat{F\circ G}(\xi) = \int_{S^2} \widehat{F}(\xi_+)\,\widehat{G}(\xi_-)\,B(\boldsymbol{n}\cdot\xi/|\xi|)\,\mathrm{d}\boldsymbol{n} \tag{2.34}$$

where

$$\xi_\pm = \frac{\xi \pm |\xi|\boldsymbol{n}}{2}\,.$$

Defining the angle θ by $\cos(\theta) = \boldsymbol{n}\cdot\xi/|\xi|$, we have from e.g., [**16**], that

$$|\xi_+| = \cos(\theta/2)\,|\xi| \quad \text{and} \quad |\xi_-| = \sin(\theta/2)\,|\xi|\,.$$

Next,

$$F_1\circ G_1 - F_2\circ G_2 = (F_1-F_2)\circ G_1 + F_2\circ(G_1-G_2)\,. \tag{2.35}$$

Fourier transforming $(F_1 - F_2) \circ G_1$,

$$\int_{S^2} (\widehat{F}_1(\xi_+) - \widehat{F}_2(\xi_+))\, \widehat{G}_1(\xi_-)\, B(\boldsymbol{n}\cdot\xi/|\xi|)\, \mathrm{d}\boldsymbol{n}\,.$$

Next,

$$\frac{|(\widehat{F}_1(\xi_+) - \widehat{F}_2(\xi_+))\, \widehat{G}_1(\xi_-)|}{|\xi|^{2+\alpha}} =$$

$$\frac{|(\widehat{F}_1(\xi_+) - \widehat{F}_2(\xi_+))\, \widehat{G}_1(\xi_-)|}{|\xi_+|^{2+\alpha} \cos(\theta/2)^{-2-\alpha}} \leqslant$$

$$|||F_1 - F_2|||_\alpha \cos(\theta/2)^{2+\alpha} \|G_1\|_\infty \leqslant$$

$$|||F_1 - F_2|||_\alpha \cos(\theta/2)^{2+\alpha}(1+\varepsilon)\,.$$

Therefore, the contribution of $(F_1 - F_2) \circ G_1$ to

$$|||F_1 \circ G_1 - F_2 \circ G_2|||_\alpha$$

is bounded by

$$|||F_1 - F_2|||_\alpha (1+\varepsilon) \int_{S^2} \cos(\theta/2)^{2+\alpha} B(\boldsymbol{n}\cdot\xi/|\xi|)\, \mathrm{d}\boldsymbol{n}\,.$$

Likewise, the contribution of $F_2 \circ (G_1 - G_2)$ to

$$|||F_1 \circ G_1 - F_2 \circ G_2|||_\alpha$$

is bounded by

$$|||F_1 - F_2|||_\alpha (1+\varepsilon) \int_{S^2} \sin(\theta/2)^{2+\alpha} B(\boldsymbol{n}\cdot\xi/|\xi|)\, \mathrm{d}\boldsymbol{n}\,.$$

Therefore, we have the inequality with c_α given by

$$c_\alpha = \int_{S^2} (\sin(\theta/2)^{2+\alpha} + \cos(\theta/2)^{2+\alpha})\, B(\boldsymbol{n}\cdot\xi/|\xi|)\, \mathrm{d}\boldsymbol{n}\,.$$

This is strictly less than 1 since

$$\int_{S^2} B(\boldsymbol{n}\cdot\xi/|\xi|)\, \mathrm{d}\boldsymbol{n} = 1\,. \qquad \blacksquare$$

This provides us with the desired convex and convolution-convex functional for isotropic densities F. When the densities are not isotropic, a modification must be made, for then $|||F - M_F|||_\alpha$ is not finite for $\alpha > 0$, and $c_\alpha < 1$ only if $\alpha > 0$.

However, the non-isotropic component of the covariance of the velocities decays very rapidly to zero, and using this, one may modify the functional ϕ so that it works without restriction on the covariance. See [**13**] for details.

We also note that these methods may be used to estimate the rate of relaxation to equilibrium in stronger, more physically meaningful norms, such as the $L^1(\boldsymbol{R}^3)$ norm. The following lemma is proved in [**13**]. In this lemma H^k denotes the Sobolev space of functions on $\boldsymbol{R}^3$ with k-fold square integrable derivatives.

Lemma ([**13**]). *Let f be any integrable function on $\boldsymbol{R}^3$ such that*

$$\int_{\boldsymbol{R}^3} |f(v)|(1 + |v|^{2+\delta})\, dv \leq B < \infty$$

and

$$\|f\|_{H^{2+\delta}} \leq C < \infty$$

for some $\delta > 0$. Then there is a constant $D < \infty$ depending only on B, C and δ so that

$$\|f\|_{L^1}^q \leq D\, |||f|||_{\delta/(1+\delta)}.$$

Moreover, for any $k > 0$, if F is a probability density of zero mean and unit variance such that

$$\|F\|_{H^k} \leq B < \infty$$

then there is a constant $\tilde{B} < \infty$ depending only on B so that

$$\|C_\gamma(F)\|_{H^k} \leq \tilde{B} < \infty$$

for all $\gamma \in \Gamma(n)$ and all n.

The second part of the Lemma is what makes the first part uniformly applicable to all of the terms in the Wild sum at least for initial data F belonging to $H^{2+\delta}$.

There are two directions to go from here: one can maintain smoothness and moment assumptions on the initial data, and then try to obtain the best possible

results on the rate of algebraic decay in the Central Limit Theorem for Maxwellian molecules.

The second is that one can try to remove the extra smoothness and moment conditions, and obtain more qualitative estimates on the rate of relaxation.

Work in both directions has been carried out, by Carlen, Carvalho and Gabetta in the first case, and by Carlen and Lu in the second case.

Carlen, Carvalho and Gabetta [**14**] have shown that for nice initial data,

$$\|Q_+^n(F) - M_F\|_{L^1(\boldsymbol{R}^3)} \leq C_\varepsilon n^{-(1-\varepsilon)\lambda}$$

where λ is the least negative eigenvalue of the linearized collision operator. By Theorem 2.1, this is the best possible result one could obtain, because it implies a bound

$$\|f(\cdot, t) - M_F\|_{L^1(\boldsymbol{R}^3)} \leq C_\varepsilon e^{-t(1-\varepsilon)\lambda},$$

and apart from the arbitrarily small ε, this is the best that one can have. This shows that nothing essential is given up in the probabilistic methods used to prove Theorem 2.1.

Carlen and Lu [**17**] have obtain precise qualitative information on the rates of relaxation for general initial data. This work shows that the longer the tails, the slower the decay, and gives the first examples of solutions that decay more slowly than exponentially.

3 - Kinetic theory from a many body perspective

3.1 - *Introduction*

Kinetic theory, as it was originally developed by Maxwell and Boltzmann describes the behavior of a system of many colliding molecules on a *mesoscopic scale* which permits one, *in principle*, to calculate the transport of bulk mass, energy and momentum in the system without directly taking into account all of the detailed behavior of the individual molecules. However, this is very much a matter of «in principle»; the Boltzmann equation is notoriously difficult to analyze.

Over forty years ago, Mark Kac made a radical proposal: instead of working with the Boltzmann equation, go back to a system of many colliding molecules. *However*, do not go back to the complicated underlying model of physical dynamics, but instead go back to a simpler stochastic model for the collisions. In-

deed, the idea is to take the *simplest* many body system that has the Boltzmann equation as a scaling limit, and to directly analyze that.

When using a microscopic model to study mesoscopic or macroscopic physics, it is not necessary that the microscopic dynamics exactly model the actual microscopic physics. Instead, we require only that it correctly capture the physics on the mesocopic scale.

In other words, for our purposes, all microscopic dynamics having the Boltzmann equation as their mesoscopic limit are equally valid. While they are all equally valid, they are not equally *good*. The simpler they are the better. Kac proposed such a microscopic system, and he did show that it had the Boltzmann equation as its mesoscopic limit. He proposed that a quantitative study of his stochastic model would provide new insight into the Boltzmann equation. In Kac's model, the collision mechanism is random. Making the collision mechanism random – by fiat – provides a much simpler connection between the microscopic dynamics and the irreversible evolution described by the Boltzmann equation, and therefore a more tractable setting in which to analyze the Boltzmann equation.

For a system of N particles with positions x_j and velocities v_j, $j = 1, 2, \dots, N$, let $(\vec{x}, \vec{v}) \in (\boldsymbol{R}^3)^{2N}$ the full microscopic state of the system - the totality of all of the positions and velocities. Let $F(\vec{x}, \vec{v}, t)$ denote the probability density for finding the system in the state $(\vec{x}, \vec{v})$ at time t.

The initial results obtained by Kac concerned the *spatially homogeneous* case in which F does not depend on the positions $\vec{x}$. He considered a system with one dimensional velocities, and pair collisions that conserved energy but not momentum. (With only two degree of freedom, two conservation laws would trivialize matters). In this case, if the total energy is E, the state space for the system is the sphere $S^{N-1}(\sqrt{E})$ given by

$$\sum_{j=1}^{N} v_j^2 = E \,. \tag{3.1}$$

Let $\vec{v} = (v_1, v_2, \dots, v_N)$ denote a point in the state space.

Kac proposed to let $\vec{v}$ evolve by undergoing a certain random walk on the state space in which the steps model pair collisions. Consider a collision of particles i and j with initial velocities v_i and v_j, and final velocities v_i^* and v_j^*. Then the conservation of the total energy implies that $v_i^2 + v_j^2 = (v_i^*)^2 + (v_j^*)^2$ and hence that for some angle θ,

$$v_i^* = \cos(\theta)\, v_i - \sin(\theta)\, v_j \qquad \text{and} \qquad v_j^* = \sin(\theta)\, v_i + \cos(\theta)\, v_j .$$

For $i<j$ define the $N\times N$ rotation matrix $R(\theta, i, j)$ by

$$R(\theta, i, j)\ \vec{v} = (v_1, \dots, v_i^* \dots, v_j^*, \dots, v_N).$$

Now let pick a pair of particles $i<j$ uniformly at random, and pick an angle θ with $-\pi \leqslant \theta \leqslant \pi$ according to some probability density ϱ on $[-\pi, \pi]$. At each step, with probability $\dfrac{2}{N(N-1)}\varrho(\theta)\,\mathrm{d}\theta$, the «collision»

$$\vec{v} \ \to\ R(\theta, i, j)\ \vec{v}$$

occurs. From this description, one easily sees that the Markov transition operator Q for this process is given by

$$Q\phi(\vec{v}) = \frac{2}{N(N-1)} \int_{-\pi}^{\pi} \phi(R(\theta, i, j)\ \vec{v})\ \varrho(\theta)\,\mathrm{d}\theta\,. \tag{3.2}$$

We require that ϱ is symmetric about $\theta = 0$. This makes Q self adjoint on $L^2(S^{N-1}(\sqrt{E}), \mathrm{d}\mu)$ where $d\mu$ denotes the uniform probability measure on $S^{N-1}(\sqrt{E})$. We also require that $\varrho(0) > 0$, and that ϱ be continuous near the origin. This ensures that 1 is a simple eigenvalue of Q. The eigenvector is of course the constant function 1.

As one sees from (3.2), Q is an average over isometries on $L^2(S^{N-1}(\sqrt{E}), \mathrm{d}\mu)$, and so it is a contraction. Moreover, the spectrum is discrete. In fact, if $Y(\vec{v})$ is any h-th degree spherical harmonic, so is each $Y(R(\theta, i, j)\ \vec{v}))$, and hence so is $QY(\vec{v})$. Therefore, for each n, the space of n-th degree spherical harmonics is an invariant subspace for Q. As Kac observed, this means that the eigenvectors of Q are spherical harmonics, and the spectrum is discrete.

Kac proposed to run the random walk just described in continuous time with the collisions arriving in a Poisson stream with rate N. This means that the probability transition operator at time t is

$$e^{-Nt} \sum_{j=0}^{\infty} \frac{(Nt)^j}{j!} Q^j = e^{tN(Q-I)}.$$

Let $\mathcal{L}_N$ denote the self adjoint linear operator $N(Q-I)$:

$$\mathcal{L}_N F(\vec{v}) = N\left(\frac{2}{N(N-1)} \sum_{i<j} \frac{1}{2\pi} \int_0^{2\pi} F(R(\theta, i, j)\ \vec{v})\,\mathrm{d}\theta - F(\vec{v}) \right).$$

Now let $F(\vec{v}, t)$ be the probability density (with respect to the uniform probability measure μ) for finding the system in state $\vec{v}$ at time t. Since Q is self adjoint, this density satisfies the equation

$$\frac{\partial}{\partial t} F(\vec{v}, t) = \mathcal{L} F(\vec{v}, t). \tag{3.3}$$

This linear equation is called the *Kac Master equation.*

Kac then went on to prove that the mesoscopic limit of this stochastic evolution is governed by a model Boltzmann equation for particles with one dimensional velocities. To study the limit as N tends to infinity, let $F^{(N)}$ be a sequence of probability densities on $S^{N-1}(\sqrt{N})$, that are to be used as initial data for the Kac master equation (3.3). (We take the total energy to be N so that independent of N, each particle has on average one unit of energy). Suppose that each is a symmetric function of the velocities $\{v_1, v_2, \ldots, v_n\}$.

Let $f^{(N)}$ be the probability density on $\boldsymbol{R}$ defined by

$$\int_a^b f^{(N)}(v)\, \mathrm{d}v = \int_{S^{N-1}(\sqrt{N})} 1_{[a,b]}(v_1)\, F^{(N)}(\vec{v})\, \mathrm{d}\mu\,.$$

Suppose that for each fixed k, and any test function ϕ on $\boldsymbol{R}^k$,

$$\lim_{N\to\infty} \int_{S^{N-1}} \phi(v_1, v_2, \ldots, v_k)\, F^{(N)}(\vec{v})\, \mathrm{d}\mu = \int_{\boldsymbol{R}^k} \phi(v_1, v_2, \ldots, v_k) \prod_{j=1}^{k} f^{(N)}(v_j)\, \mathrm{d}^k v \tag{3.4}$$

and for any test function ψ on R,

$$\lim_{N\to\infty} \int \psi(v) f^{(N)}(v)\, \mathrm{d}v = \int_R \psi(v) f(v)\, \mathrm{d}v \tag{3.5}$$

where f is a probability density on R. Then the family of densities $\{F^{(N)}\}$ is said to be a *consistent, chaotic* family of initial data for the Kac master equation.

Kac then proved the following: if $F^{(N)}(\vec{v}, t)$ denotes the corresponding solution; i.e.,

$$F^{(N)}(\vec{v}, t) = e^{tN(Q-I)} F^{(N)}(\vec{v})$$

then for any test function ψ,

$$\lim_{N\to\infty} \int_{S^{N-1}(\sqrt{N})} \psi(v)\, F^{(N)}(\vec{v}, t)\, \mathrm{d}\mu = \int_R \psi(v)\, f(v, t)\, \mathrm{d}v$$

exists and defines $f(v, t)$, which satisfies a model Boltzmann equation, called the *Kac equation.*

$$\frac{\partial}{\partial t} f(v, t) = \mathcal{Q}(f, f)(v, t) \tag{3.6}$$

where

$$Q(f, f)(v) = f \circ f - f \tag{3.7}$$

and now $f \circ f$ denotes the modified Wild convolution

$$f \circ f(v) = \int_{-\pi}^{\pi} \varrho(\theta) \left[\int_R (f(\cos(\theta)\, v + \sin(\theta)\, w)\, f(-\sin(\theta)\, v + \cos(\theta)\, w))\, \mathrm{d}w \right] \mathrm{d}\varrho(\theta).$$

There is just one probability density that is an equilibrium solution of this equation, and which has unit energy; namely

$$M(v) = \frac{1}{\sqrt{2\pi}}\, e^{-v^2/2}.$$

Kac proposed that one could determine the rate of approach to equilibrium for solutions $f(v, t)$ of (3.7); that is, the rate at which $f(v, t)$ tends to $M(v)$ by determining bounds on the rate at which solutions $F(v, t)$ of (3.3) tend toward 1, the uniform density, which is the unique equilibrium density for the Master equation.

The equation (3.3) is linear, while (3.7) is non linear. However, rather precise control on how the rate of approach to equilibrium for the linear equation depends on N is required to draw any conclusions relevant to the non linear equation. Thus, it is not a trivial matter to exploit the linearity of the master equation. Indeed, Kac was unable to carry out his program, and it was many years before his ideas were vindicated. However, at the time he wrote, there was very little mathematical work on the Boltzmann equation. The paper [**10**] was the most significant work up to that time, but it provided no means to address the problem of determining the rate of approach to equilibrium.

The way in which Kac's ideas were vindicated, many years later, appears to set the stage for further progress, as we now explain. Let $-\Delta_N$ denote the negative eigenvalue closest to zero, so that Δ_N is the *spectral gap* for $\mathcal{L}_N$. That is,

$$\Delta_N = \sup\{\langle f, N(I-Q)f\rangle_{L^2} : \langle f, 1\rangle_{L^2} = 0 \text{ and } \|f\|_2 = 1\}. \tag{3.8}$$

Kac made the following specific conjecture about $\mathcal{L}_N$:

$$\liminf_{N\to\infty} \Delta_N = C > 0. \tag{3.9}$$

This conjecture would have the following consequence: consider any consistent, chaotic sequence $\{F^{(N)}\}$ of initial data for the Kac master equation. Let $\{F^{(N)}(\vec{v}, t)\}$ denote the corresponding sequence of solutions. Then for any $c < C$, we have that

$$\|F^{(N)}(\cdot, t) - 1\|_2 \leq \|F^{(N)}(\cdot) - 1\|_2 e^{-ct}. \tag{3.10}$$

That is, there would be exponentially fast relaxation to equilibrium uniformly in N -at least as far as the exponential rate is concerned. Kac then hoped to use this to draw conclusions about the rate at which solutions of the Kac equation (3.6) approached equilibrium.

Kac's conjecture was finally recently proved by Janvresse [**25**] using Yau's martingale method [**43**], [**44**]. Her method gives no information on what C might be, and seems difficult to generalize to physically realistic momentum conserving collisions. However, shortly afterwards, Carlen, Carvalho and Loss [**15**] found a very simple relation between Δ_N and Δ_{N-1}. To explain this, we first introduce a simple measure of the correlations between the velocities on the sphere.

Let π_j be the j-th coordinate projection on the unit sphere so that $\pi_j(\vec{v}) = v_j$. Define a quantity κ_N by

$$\kappa_N = \sup\left\{ \frac{\int_{S^{N-1}} g(\pi_1)\, g(\pi_2)\, d\mu}{\int_{S^{N-1}} |g(\pi_1)|^2 d\mu} : g \in \mathcal{C}(\boldsymbol{R}), \int_{S^{N-1}} g(\pi_1)\, d\mu = 0 \right\}. \tag{3.11}$$

Notice that the marginal distribution of (v_1, v_2) induced by μ_N is

$$\frac{|S^{N-3}|}{N|S^{N-1}|}\left(1 - \frac{v_1^2 + v_2^2}{N}\right)^{(N-4)/2} dv_1\, dv_2.$$

As N tends to infinity, this tends to

$$\frac{1}{2\pi} e^{-(v_1^2+v_2^2)/2}\,\mathrm{d}v_1\,\mathrm{d}v_2,$$

and under this limiting measure, the two coordinate functions v_1 and v_2 are independent. Hence for any admissible trial function g in (3.11),

$$\begin{aligned}
&\lim_{N\to\infty}\int_{S^{N-1}} g(\pi_1)\,g(\pi_2)\,\mathrm{d}\mu = \\
&\frac{1}{2\pi}\int_{\mathbb{R}^2} g(v_1)\,g(v_2)\,e^{-(v_1^2+v_2^2)/2}\,\mathrm{d}v_1\,\mathrm{d}v_2 = \\
&\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(v_1)\,e^{-v_1^2/2}\,\mathrm{d}v_1\;\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} g(v_2)\,e^{-v_1^2/2}\,\mathrm{d}v_2 = \\
&\lim_{N\to\infty}\langle g\circ\pi_1,\,1\rangle\langle g\circ\pi_2,\,1\rangle = 0
\end{aligned} \tag{3.12}$$

which implies that $\lim_{N\to\infty}\kappa_N = 0$, without, however, showing how fast.

We can now explain the relation between Δ_N and Δ_{N-1}:

Theorem 3.1 ([**15**]). *For all* $N \geq 3$.

$$\Delta_N \geq (1-\kappa_N)\,\Delta_{N-1}.$$

The variational problems (3.11) and (3.8) are very different. While (3.8) involves the operator $\mathcal{L}_N$ and all of the details of the dynamics, (3.11) does not: κ_N *is completely independent of the details inherent in the master equation; it is a purely «kinematical» measure of the lack of independence of the coordinate function* v_1 *and* v_2 *under* μ.

In fact, it is not hard to compute κ_N:

Theorem 3.2 ([**15**]). *For all* $N \geq 3$,

$$\kappa_N = \frac{3}{N^2-1}\,.$$

The last two theorems lead to a proof of Kac's conjecture: we see that

$$\begin{aligned}\Delta_N &= \prod_{j=3}^{N}\left(1-\frac{3}{N^2-1}\right)\Delta_2\\ &= \prod_{j=3}^{N}\left(\frac{N^2-4}{N^2-1}\right)\Delta_2\\ &= \prod_{j=3}^{N}\left(\frac{(N-2)(N+2)}{(N-1)(N+1)}\right)\Delta_2\\ &= \prod_{j=3}^{N}\left(\frac{(N+2)}{(N-1)}\right)\left(\frac{(N+1)}{(N-2)}\right)^{-1}\Delta_2\\ &= \frac{(N+2)}{(N-1)}\,\frac{1}{4}\Delta_2.\end{aligned}$$

Moreover, the computation of Δ_2 is very simple. For two particles, Q is an operator on functions on S^1:

$$\langle f, Qf\rangle_{L^2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(\psi)\, f(\psi-\theta)\, \varrho(\theta)\, \mathrm{d}\theta\, \mathrm{d}\psi\,, \tag{3.13}$$

and writing this in terms of Fourier series leads to the following formula for λ_2, the second largest eigenvalue of Q for $N=2$:

$$\lambda_2 = \sup_{k\neq 0}\left\{\int_{-\pi}^{\pi}\varrho(\theta)\cos(k\theta)\,\mathrm{d}\theta\right\}. \tag{3.14}$$

By the Riemann-Lebesgue lemma, $\lambda_2 < 1$, and hence $\Delta_2 = 2(1-\lambda_2) > 0$. This leads to the following result the first part of the following result:

Theorem 3.3 ([**15**]). *For all* $N \geqslant 2$,

$$\Delta_N \geqslant \frac{1-\lambda_2}{2}\,\frac{N+2}{N-1}\,. \tag{3.15}$$

Moreover, this bound is actually an exact value for the constant density ϱ, *in which case* $\lambda_2 = 0$, *and more generally whenever*

$$\int_{-\pi}^{\pi}\varrho(\theta)\cos(k\theta)\,\mathrm{d}\theta \leqslant \int_{-\pi}^{\pi}\varrho(\theta)\cos(4\theta)\,\mathrm{d}\theta \tag{3.16}$$

for all $k \neq 0$. *In all of these cases*

$$\sum_{k=1}^{N} (v_k^4 - \langle v_k^4, 1\rangle) \tag{3.17}$$

is an eigenfunction of $\mathcal{L}_N$ *with*

$$\mathcal{L}_N f_N = -\Delta_N f_N.$$

Proof. Apart from the theorems whose proof has been postponed, it remains only to check that the bound is sharp, and that the multiplicity of the eigenvalue is one. For

$$f_N(\vec{v}) = \sum_{j=1}^{N} (v_j^4 - \langle 1, v_j^4\rangle), \tag{3.18}$$

$$Qf_N = \left(1 - \frac{2\gamma(N+2)}{N(N-1)}\right) f_N \tag{3.19}$$

where

$$\gamma = \frac{1}{4}\left(1 - \int_{-\pi}^{\pi} \cos(4\theta)\, \varrho(\theta)\, \mathrm{d}\theta\right). \tag{3.20}$$

This explicitly computed eigenvalue leads to the upper bound

$$\Delta_N \leqslant 2\gamma \frac{N+2}{N-1}. \tag{3.21}$$

Clearly, in the Kac model with ϱ uniform – as in Kac's original paper, $\gamma = 1/4$, and so (3.21) implies that Δ_N is no larger than $(N+2)/(2(N-1))$. Since for the original Kac model $\lambda_2 = 0$, this upper bound on Δ_N coincides with the lower bound (3.15), and hence this lower bound is sharp.

In fact, the upper bound on Δ_N provided by the trial function (3.18) coincides with the lower bound obtained above whenever $f_2(v_1, v_2) = v_1^4 + v_2^4 - (3/4)$ is such that

$$Qf_2 = \lambda_2 f_2. \tag{3.22}$$

Writing $v_1 = \cos(\theta)$ and $v_2 = \sin(\theta)$, we have

$$f(\cos(\theta), \sin(\theta)) = \frac{\cos(4\theta)}{4}.$$

Hence (3.22) certainly holds whenever (3.16) holds. ■

It is shown in [**15**] that actually in a wide range of circumstances, $-\Delta_N$ is a simple eigenvalue of $\mathcal{L}_N$, and that

$$\mathcal{L}_N f_N = -\Delta_N f_N$$

for all N sufficiently large, even if this is false for, say, $N = 2$. Thus in a great many cases we have the exact computation that

$$\Delta_N = 2\gamma \frac{N+2}{N-1} \tag{3.23}$$

for large N.

3.2 - *From N particles to $N+1$*

We will now explain how to prove Theorems 3.1 and 3.2. The argument turns on several very simple features of the Kac model that occurs in many other models as well. Thus, it is worthwhile to develop the argument in an abstract setting permitting wide application.

The abstract setting is that of a family if continuous time Markov process with a state with state spaces X_N, $n = 2, 3, 4 \ldots$. We think of X_N as the «N particle state space». For each N, we let Q denote the Markov transition operator. We do not subscript it with an N; the number of particles will be clear from the context. Each X_N is equipped with a probability measure μ_N.

In the case of the Kac model X_N is the sphere S^{N-1} and μ_N is the uniform probability measure on the sphere. There is a natural action of Π_N, the symmetric group on N letters, on S^{N-1}: if $\pi \in \Pi_N$, let

$$\pi(\vec{v}) = (v_{\pi(1)}, v_{\pi(2)}, \ldots, v_{\pi(N)}).$$

Clearly, this is a measure preserving transformation. This is the first of four key features of the Kac model that are crucial to the proofs of Theorem 3.1 and 3.2.

Feature 1. *For each $N > 1$ there is measure space $(X_N, \mathcal{S}_N, \mu_N)$, with μ_N a probability measure, on which there is a measure preserving action of Π_N, the*

symmetric group on N letters. We denote

$$\mathcal{H}_N = L^2(X_N, \mu_N). \tag{3.24}$$

We think of the action of Π_N as representing «exchange of particles».

Feature 2. *There is another measure space $(Y_N, \mathcal{C}_N, \nu_N)$ and there are measurable maps $\pi_j: X_N \to Y_N$ for $j = 1, 2, \ldots, N$ such that for all $\sigma \in \Pi_N$, and each j,*

$$\pi_j \circ \sigma = \pi_{\sigma(j)}. \tag{3.25}$$

Moreover, for each j, and all $A \in \mathcal{C}_N$,

$$\nu_N(A) = \mu_N(\pi_j^{-1}(A)). \tag{3.26}$$

We denote

$$\mathcal{K}_N = L^2(Y_N, \nu_N). \tag{3.27}$$

Think of $\pi_j(x)$ as specifying the «state of the j-th particle given that the N particle system is in state x». For example, in the Kac model on the unit sphere, $Y_N = [-1, 1]$ and

$$\pi_j(v_1, v_2, \ldots, v_N) = v_j \in [-1, 1]. \tag{3.28}$$

Here Y_N does not depend on N, and it may appear odd for the single particle state space to depend on N. However, the methods permit this generality, and it turns out to be useful.

Notice that once Y_N and the π_j are given, ν_N is specified through (3.26). In the Kac model we therefore have

$$\nu_N = \frac{|S^{N-2}|}{|S^{N-1}|}(1 - v^2)^{(N-3)/2}\,\mathrm{d}v\,. \tag{3.29}$$

Feature 3. *For each $N \geq 3$ and each $j = 1, 2, \ldots, N$, there is a map*

$$\phi_j: (X_{N-1} \times Y_N) \to X_N \tag{3.30}$$

so that

$$\pi_j(\phi_j(x, y)) = y \tag{3.31}$$

for all $j = 1, \dots, N$ *and all* $(x, y) \in X_{N-1} \times Y_N$. *Moreover,* ϕ_j *has the property that for all* $A \in \mathcal{S}_N$,

$$[\mu_{N-1} \otimes \nu_N](\phi_j^{-1}(A)) = \mu_N(A), \tag{3.32}$$

or equivalently, for all bounded measurable functions f *on* X_N,

$$\int_{X_N} f \, d\mu_N = \int_{Y_N} \left[\int_{X_{N-1}} f(\phi_j(x, y)) \, d\mu_{N-1}(x) \right] d\nu_N(y). \tag{3.33}$$

In the Kac model, for any $\tilde{v} \in X_{N-1} = S^{N-2}$ and any $v \in Y_N = [-1, 1]$ we put

$$\phi_N(\tilde{v}, v) = (\sqrt{1 - v^2}\tilde{v}, v). \tag{3.34}$$

Finally we come to the dynamics; i.e., the Markov transition operator

Feature 4. *For each* $N \geq 2$, *there is a self-adjoint and positivity preserving operator* Q *on* $\mathcal{H}_N$ *such that* $Q1 = 1$. *These operators are related to one another by: for each* $N \geq 3$, *each* $j = 1, 2, \dots, N$, *and each square integrable function* f *on* X_N,

$$\langle f, Qf \rangle_{\mathcal{H}_N} = \frac{1}{N} \sum_{j=1}^{N} \int_{Y_N} (\langle f_{j,y}, Qf_{j,y} \rangle_{\mathcal{H}_{N-1}}) \, d\nu_N(y) \tag{3.35}$$

where for each j *and each* $y \in Y_N$,

$$f_{j,y}(\cdot) = f(\phi_j(\cdot, y)). \tag{3.36}$$

This feature is easily verified for the Kac model.

Definition. A *Kac System* is a system of probability spaces $(X_N, \mathcal{S}_N, \mu_N)$ and $(Y_N, \mathcal{C}_N, \nu_N)$ for $N \in \mathbb{N}$, $N \geq 2$, together with, for each N, maps π_j and ϕ_j, $j = 1, 2, \dots, N$, a measure preserving action of Π_N on $(X_N, \mathcal{S}_N, \mu_N)$, and a Markov transition operator Q on $\mathcal{H} = L^2(X_N, \mu_N)$, related to one another in such a way that they possess all of the properties specified in Features 1 through 4 above.

In our analysis of the spectral gaps of the Markov transition operators Q in Kac systems, certain other operators related to conditional expectations will be important:

Suppose that $(X_N, \mathcal{S}_N, \mu_N)$, $(Y_N, \mathcal{C}_N, \nu_N)$, π_j, and ϕ_j, are defined and related as specified above. For each $j = 1, 2, \dots, N$, let P_j be the orthogonal projection onto the subspace of $\mathcal{H}_N$ consisting of functions of the form $g \circ \pi_j$ for some $g \in \mathcal{K}_N$. Then, with $y = \pi_j(x)$ and $f_{j,y}$ given by (3.36),

$$P_j f(x) = g(\pi_j(x)) \quad \text{where} \quad g(y) = \int_{X_{N-1}} f_{j,y}(\tilde{x})\, \mathrm{d}\mu_{N-1}(\tilde{x}). \tag{3.37}$$

Next, define

$$P = \frac{1}{N} \sum_{j=1}^{N} P_j \tag{3.38}$$

which is clearly a positive contraction on $\mathcal{H}_N$. Define a contraction K on $\mathcal{K}_N$ by

$$(Kg) \circ \pi_N = P_N(g \circ \pi_{N-1}). \tag{3.39}$$

Evidently, K is a self adjoint contraction, and $K1 = 1$.

There is nothing special here about $N - 1$ and N: by the permutation symmetry, $P_i(g \circ \pi_j) = (Kg) \circ \pi_i$ for all $i \neq j$. It follows from the definition of P that

$$\langle gKg \rangle_{\mathcal{H}_N} = \int_{X_N} g(\pi_1)\, g(\pi_2)\, \mathrm{d}\mu_N$$

and so in the Kac model, the quantity κ_N defined in (3.11) is evidently the second largest eigenvalue of K. That is, $1 - \kappa_N$ is the spectral gap of K.

Combining (3.36), (3.37) and (3.39), we obtain

$$Kg(y) = \int_{X_{N-1}} g(\pi_{N-1}(\phi_N(\tilde{x}, y)))\, \mathrm{d}\mu_{N-1}(\tilde{x}) \tag{3.40}$$

which provides an explicit form for the operator K. In the case of the Kac model,

we obtain

$$
\begin{aligned}
Kg(v)) &= \int_{X_{N-1}} g(\sqrt{1-v^2}w_{N-1})\, d\mu_{N-1}(w) \\
&= \int_{-1}^{1} g(\sqrt{1-v^2}w)\, d\nu_{N-1}(w) \qquad (3.41) \\
&= \frac{|S^{N-3}|}{|S^{N-2}|} \int_{-1}^{1} g(\sqrt{1-v^2}w)(1-w^2)^{(N-4)/2}\, dw
\end{aligned}
$$

from (3.29) and (3.40).

Theorem 3.4. *For any Kac system, let P and K be defined by* (3.38) *and* (3.39). *Define μ_N, κ_N and β_N by*

$$
\begin{aligned}
\mu_N &= \sup\{\langle f, Pf\rangle_{\mathcal{H}_N} : \|f\|_{\mathcal{H}_N} = 1 \text{ and } \langle 1, f\rangle_{\mathcal{H}_N} = 0\} \\
\kappa_N &= \sup\{\langle g, Kg\rangle_{\mathcal{H}_N} : \|g\|_{\mathcal{H}_N} = 1 \text{ and } \langle 1, g\rangle_{\mathcal{H}_N} = 0\} \qquad (3.42) \\
\beta_N &= \frac{1}{N-1} \left| \inf\{\langle g, Kg\rangle_{\mathcal{H}_N} | : \|g\|_{\mathcal{H}_N} = 1\} \right| .
\end{aligned}
$$

Suppose, moreover, that the operator P has pure point spectrum. Then, either $\mu_N = 0$ or

$$
\mu_N = \max\left\{ \frac{1}{N}(1 + (N-1)\kappa_N),\ \frac{1}{N}(1 + (N-1)\beta_N) \right\}. \qquad (3.43)
$$

Proof. Let f be an eigenfunction of P with eigenvalue ν. Because P commutes with permutations, either f is invariant under permutations, or else there is some transposition, that we may as well take to be $\sigma_{1,2}$, such that $f \circ \sigma_{1,2} = -f$.

First suppose that f is symmetric. Then for some $h \in \mathcal{H}_N$ independent of k, $P_k f = h \circ \pi_k$, and so

$$
\nu f = Pf = \frac{1}{N} \sum_{k=1}^{N} P_k f = \frac{1}{N} \sum_{k=1}^{N} h \circ \pi_k, \qquad (3.44)
$$

where ν is the eigenvalue, and h is some function on $\boldsymbol{R}$. Applying P_1 to both sides

of (3.44), $\nu P_1 f = \frac{1}{N}\sum_{k=1}^{N} P_1(h\circ\pi_k)$, which can be easily simplified to

$$\nu h\circ\pi_1 = \frac{1}{N}\left(h\circ\pi_1 + (N-1)\,Kh\circ\pi_1\right).$$

If $h=0$, then by (3.44). $\nu = 0$ Otherwise, h is an eigenfunction of K with eigenvalue $\tilde{\nu}$ so that

$$\nu = \frac{1}{N}\left(1 + (N-1)\,\tilde{\nu}\right). \tag{3.45}$$

It might happen that $\nu = 0$ without $h = 0$ being the case. But then h must be an eigenfunction of K with eigenvalue $-1/(N-1)$. In any case, $\nu = 0$ or otherwise equation (3.45) must hold.

Next, consider the remaining case:

$$f\circ\sigma_{1,2} = -f.$$

Note that $P_k(f\circ\sigma_{1,2}) = P_k f = 0$ whenever k is different from both 1 and 2, and hence

$$\frac{1}{N}\sum_{k=1}^{N} P_k f = \frac{1}{N}\left(P_1 f + P_2 f\right).$$

Again, applying P_1 to both sides of the equation and keeping in mind that when $P_1 f = h\circ\pi_1$ then $P_2 f = -h\circ\pi_2$ we get

$$\nu h\circ\pi_1 = \frac{1}{N}\left(h\circ\pi_1 - Kh\circ\pi_1\right).$$

In case h vanishes identically, the eigenvalue ν must vanish also, otherwise h is an eigenfunction of K with eigenvalue $\tilde{\nu}$ such that

$$\nu = \frac{1-\tilde{\nu}}{N}. \tag{3.46}$$

Since by the definition (3.42), $-(N-1)\,\beta_N$ is the most negative eigenvalue of K. The formula (3.43) of μ_N now follows from (3.45) and (3.46). ■

A stronger version which determined the multiplicities is proved in [**15**]. We now come to the key inductive argument that leads directly to the proof of Theorem 3.1.

Theorem 3.5. *For any Kac system, let P and μ_N be defined by* (3.38) *and* (3.42). *Define λ_N by*

$$\lambda_N = \sup\{\langle f, Qf\rangle_{\mathcal{H}_N} \mid \|f\|_{\mathcal{H}_N} = 1, \langle f, 1\rangle_{\mathcal{H}_N} = 0\}. \tag{3.47}$$

Then

$$\lambda_N \leqslant (\lambda_{N-1} + (1-\lambda_{N-1})\mu_N). \tag{3.48}$$

Proof. We start from (3.35), taking any function $f \in \mathcal{H}_N$ satisfying the conditions imposed in (3.47).

$$\begin{aligned}
\langle f, Qf\rangle_{\mathcal{H}_N} &= \frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}\langle f_{j,y}, Qf_{j,y}\rangle_{\mathcal{H}_{N-1}}\mathrm{d}\nu_N(y)\\
&= \frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}\langle [f_{j,y}-P_jf(y)]+P_jf(y), Q([f_{j,y}-P_jf(y)]+P_jf(y))\rangle_{\mathcal{H}_{N-1}}\mathrm{d}\nu_N(y)\\
&= \frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}\langle [f_{j,y}-P_jf(y)], Q[f_{j,y}-P_jf(y)]\rangle_{\mathcal{H}_{N-1}}\mathrm{d}\nu_N(y)\\
&+ \frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}|P_jf(y)|^2\mathrm{d}\nu_N(y),
\end{aligned}$$

since each $P_jf(y)$ is constant on X_{N-1} and so on $\mathcal{H}_{N-1}$, $QP_jf(y) = P_jf(y)$, and

$$\langle [f_{j,y} - P_jf(y)], P_jf(y)\rangle_{\mathcal{H}_{N-1}} = 0.$$

But

$$\frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}|P_jf(y)|^2\mathrm{d}\nu_N(y) = \langle f, Pf\rangle_{\mathcal{H}_N},$$

and hence

$$\begin{aligned}
\langle f, Qf\rangle_{\mathcal{H}_N} &= \frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}\langle [f_{j,y} - P_jf(y)], Q[f_{j,y} - P_jf(y)]\rangle_{\mathcal{H}_{N-1}}\mathrm{d}\nu_N(y)\\
&+ \langle f, Pf\rangle_{\mathcal{H}_N}.
\end{aligned} \tag{3.49}$$

Now since $\langle [f_{j,y} - P_j f(y)], 1\rangle_{\mathfrak{H}_{N-1}} = 0$ for each y and j,

$$\langle [f_{j,y} - P_j f(y)], Q[f_{j,y} - P_j f(y)]\rangle_{\mathfrak{H}_{N-1}} \leq \lambda_{N-1} \|f_{j,y} - P_j f(y)\|^2_{\mathfrak{H}_{N-1}}$$
$$\equiv \lambda_{N-1}(\|f_{j,y}\|^2_{\mathfrak{H}_{N-1}} - |P_j f(y)|^2_{\mathfrak{H}_{N-1}}).$$

Averaging over j and integrating over y,

$$\frac{1}{N}\sum_{j=1}^{N}\int_{Y_N}\langle [f_{j,y} - P_j f(y)], Q[f_{j,y} - P_j f(y)]\rangle_{\mathfrak{H}_{N-1}} d\nu_N(y) \leq \lambda_{N-1}(\|f\|^2_{\mathfrak{H}_N} - \langle f, Pf\rangle_{\mathfrak{H}_N}).$$

From this and (3.49), (3.48) follows, since f itself is an admissible trial function for μ_N. ∎

We have the following immediate corollary:

Corollary. *With κ_N and β_N defined as in* (3.42), *define $\Delta_N = N(1 - \lambda_N)$. Then*

$$\Delta_N \geq (1 - \max\{\kappa_N, \beta_N\})\,\Delta_{N-1} \tag{3.50}$$

for all $N \geq 3$, and hence for all $N > 2$,

$$\Delta_N \geq \prod_{j=3}^{N}(1 - \max\{\kappa_j, \beta_j\})\,\Delta_2. \tag{3.51}$$

We see that a sufficient condition for $\liminf_{N\to\infty} \Delta_N > 0$ is $\Delta_2 > 0$ and

$$\prod_{N=3}^{\infty}(1 - \max\{\kappa_N, \beta_N\}) > 0\,.$$

To prove Theorem 3.1 and Theorem 3.2, we just need to show that for the Kac model, $\kappa_N > \beta_N$, and that κ_N is given by the formula in Theorem 3.2.

Theorem 3.6. *For the Kac model, the operator K possesses a complete orthonormal set $\{g_n\}$, $n \geq 0$, of eigenfunctions where g_n is a polynomial of degree n and the corresponding eigenvalue α_n is zero if n is odd, and if $n = 2k$, α_n is given by*

$$\alpha_{2k} = (-1)^k \frac{|S^{N-3}|}{|S^{N-2}|}\int_0^{\pi}(1 - \sin^2\theta)^k \sin^{N-3}\theta\, d\theta\,.$$

Before proving the theorem, we remark that it is clear from this formula that

$$|\alpha_{2(k+1)}| < |\alpha_{2k}|$$

for all k. Since α_2 is negative, the second largest eigenvalue is

$$\alpha_4 = \frac{3}{N^2-1},$$

hence $\kappa_N = 3/(N^2-1)$, and this proves Theorem 3.2.

The smallest is

$$\alpha_2 = -\frac{1}{N-1}.$$

Hence $\beta_N = 1/(N(N-1))$, and this proves $\beta_N < \alpha_N$ for all $N \geq 3$, and Theorem 3.1 is proved.

Proof of Theorem 3.6. We start from the explicit form (3.41) for K. By a simple change of variables, we rewrite it as

$$Kg(v) = \frac{|S^{N-3}|}{|S^{N-2}|} \int_0^\pi g(\sqrt{1-v^2}\cos\theta) \sin^{N-3}\theta \, d\theta .$$

The right hand side is an even function of v. Since the operator K preserves parity, K must annihilate odd functions. Hence we may assume that g is even.

Because $\left(\sqrt{1-v^2}\right)^{2k} = (1-v^2)^k$ is a polynomial of degree $2k$ in v, the space of polynomials of degree $2n$ or less is an invariant subspace of K. This implies that the eigenvectors are even polynomials, and that there is exactly one such eigenvector for each degree $2k$.

Next, suppose that g_{2k} is the degree $2k$ eigenvector of K, and let α_{2k} be the corresponding eigenvalue. We may normalize g_{2k} so that the leading coefficient is 1, and we then have

$$g_{2k} = v^{2k} + h(v)$$

where $h(v)$ is an even polynomial in v of degree no more than $2k-2$. Thus

$$\alpha_{2k} v^{2k} + \alpha_{2k} h(v) = \alpha_{2k} g_{2k} = Kg_{2k} = Kv^{2k} + Kh(v).$$

This implies that

$$Kv^{2k} = \alpha_{2k} v^{2k} + \text{lower order} .$$

The result now follows directly from the formula for K, the recurrence relation

$$\int_0^\pi \sin^n(\theta)\, \mathrm{d}\theta = \frac{n-1}{n} \int_0^\pi \sin^{n-2}(\theta)\, \mathrm{d}\theta , \tag{3.52}$$

and the fact that $K1 = 1$. Observe that the leading coefficient of v in $(1 - v^2)^k$ is $(-1)^k$. ■

Having explained how our exact determination of the gap for Kac's original model works it is appropriate to compare this approach with Janvresse's [**25**] application of Yau's martingale method [**43**], [**44**] to the same problem. There are similarities between our analysis and Yau's method, in that Yau's martingale method uses induction on N, correlation estimates, and the same conditional expectation operators P_j. There are, however, significant differences, as indicated by the difference between Janvresse's estimate and our exact calculation.

Yau's method turns on an analysis of the spectrum of the P_j operators estimated not in $L^2(S^{N-1})$, but in the Hilbert space whose inner product is $\langle h, (I - Q)\, h\rangle$, the so-called Dirichlet form space associated to Q. Hence using Yau's method, the details of the dynamics enter at each stage of the induction, while in our approach purely geometric estimates relate Δ_N to Δ_{N-1}.

However, Yau's method was designed to handle problems without the permutation symmetry that is present in the class of models considered here. Hence Yau's method is more widely applicable. But in the presence of the permutation symmetry it may be less wide than the method used to obtain the results presented here.

3.3 - *Momentum conserving collisions*

We now consider a master equation leading to a Boltzmann equation that describes physically realistic momentum conserving collisions. Since we may assume that the total momentum is zero, we take the state space to be given by the constraints

$$\sum_{j=1}^N |v_j|^2 = E \quad \text{and} \quad \sum_{j=1}^N v_j = 0 . \tag{3.53}$$

This is a sphere of dimension $3N-4$, however the action of Π_N is crucial for so that the particular embedding in $\boldsymbol{R}^{3N}$ cannot be dispensed with.

Next, consider a pair of identical particles with velocities v_i and v_j in $\boldsymbol{R}^3$. We require the collisions to conserve momentum as well as energy. As in (2.4) we define for any unit vector ω in S^2, and any i and j,

$$v_i^*(\omega) = v_i + \langle \omega, v_j - v_i \rangle \omega \tag{3.54}$$

$$v_j^*(\omega) = v_j - \langle \omega, v_j - v_i \rangle \omega \, . \tag{3.55}$$

These are the post collisional velocities when particles i and j collide with collision parameter ω.

The random collision mechanism is now that we pick a pair i, j, $i<j$, uniformly at random, and then pick an ω in S^2 at random so that

$$(v_1, v_2, \dots, v_N) \rightarrow (v_1, \dots, v_i^*(\omega), \dots, v_j^*(\omega), \dots, v_N),$$

and put

$$R(\omega, i, j)\, \vec{v} = (v_1, \dots, v_i^*(\omega), \dots, v_j^*(\omega), \dots, v_N).$$

We then define the one step transition operator Q by

$$Q_f(\vec{v}) = \binom{N-1}{2}^{-1} \sum_{i<j}^{N} \int_{S^2} f(R(\omega, i, j)\, \vec{v})\, B(\omega \cdot (v_i - v_j)/|v_j - v_j|)\, d\omega \, , \tag{3.56}$$

where B is a non-negative function on $[-1, 1]$ so that

$$2\pi \int_0^{\pi} B(\cos\theta) \sin(\theta)\, d\theta = 1 \, .$$

The function B determines the relative likelihood of the various scattering angles.

Incorporating momentum conservation has not made too much of a difference so far; the formula we have deduced for Q still looks very much like the one for the Kac model. We first check that this Boltzmann collision model is a Kac system.

By choice of units, we may assume that

$$\sum_{j=1}^{N} |v_j|^2 = 1 \quad \text{and} \quad \sum_{j=1}^{N} v_j = 0 \tag{3.57}$$

both hold initially, and hence for all time. Thus our state space X_N is the set of all vectors

$$\vec{v} = (v_1, v_2, \dots, v_N) \in \boldsymbol{R}^{3N}$$

satisfying the constraints in (3.57). Equip X_N with the uniform probability measure inherited from its natural embedding in $\boldsymbol{R}^{3N}$. The symmetric group Π_N acts on X_N as follows: For $\sigma \in \Pi_N$,

$$\sigma(v_1, v_2, \dots, v_N) = (v_{\sigma(1)}, v_{\sigma(2)}, \dots, v_{\sigma(N)}).$$

This action is clearly measure preserving. We note that X_N is geometrically equivalent to the unit sphere S^{3N-4} in $\boldsymbol{R}^{3N-3}$, but apart from identifying normalization factors in our probability measures, this identification is not conducive to efficient computation because any embedding in $\boldsymbol{R}^{3N-3}$ obscures the action of the symmetric group.

The single particle state space Y_N can be identified with the unit ball in $\boldsymbol{R}^3$ as follows: first note that because of the momentum constraint, $\sum_{j=1}^{N-1} v_j = -v_N$, to maximize $|v_N|$, minimize the energy in the first $N-1$ particles. By convexity,

$$\inf\left\{\sum_{j=1}^{N-1} |v_j|^2 : \sum_{j=1}^{N-1} v_j = -v_N\right\}$$

is attained at

$$(v_1, v_2, \dots, v_{N-1}) = -\frac{1}{N-1}(v_N, v_N, \dots, v_N),$$

which leads directly to

$$\sup\{|v_N| \,|\, (v_1, v_2, \dots, v_N) \in X_N\} = \frac{N-1}{N}. \tag{3.58}$$

Because the momentum constraint prevents all of the energy from belonging to a single particle, each v_j lies in the ball of radius $\sqrt{(N-1)/N}$ in $\boldsymbol{R}^3$.

We could take Y_N to be the ball of radius $\sqrt{(N-1)/N}$ in $\boldsymbol{R}^3$, for $N > 3$, which would then depend on N. However, calculations will work out more easily if we rescale and take Y_N to be the unit ball in $\boldsymbol{R}^3$, independent of N. Therefore, we define, for $N \geq 3$,

$$Y_N = \{v \in \boldsymbol{R}^3 \,|\, |v| \leq 1\}$$

and let $\mathcal{C}_N$ be the corresponding Borel field. We take Y_2 to be the unit sphere in $\boldsymbol{R}^3$. We are then led to define $\pi_j: X_N \to Y_N$ by

$$\pi_j(v_1, v_2, \dots, v_N) = \left(\frac{N}{N-1}\right)^{1/2} v_j. \tag{3.59}$$

The measure ν_N is now determined through (3.26).

Before deducing an explicit formula for it, we introduce the maps $\phi_j: X_{N-1} \times Y_N \to X_N$. This is useful in working out the formula for ν_N.

Consider any fixed $N \geq 3$, so that X_{N-1} is non empty. Fix a point $\vec{w} = (w_1, w_2, \dots, w_{N-1}) \in X_{N-1}$, and a point $v \in Y_N$. In order that we have

$$\pi_N(\phi_N(\vec{w}, v)) = v\,,$$

the N-th component of $\phi_N(\vec{w}, v)$ must be $\sqrt{(N-1)/N}v$. Now observe that for any $\alpha \in \mathbb{R}$,

$$\vec{v} = (v_1, v_2, \dots, v_N) = \left(\alpha w_1 - \frac{1}{\sqrt{N^2-N}} v, \dots, \alpha w_{N-1} - \frac{1}{\sqrt{N^2-N}} v, \sqrt{\frac{N-1}{N}}\, v\right)$$

satisfies $\sum_{j=1}^{N} v_j = 0$, and

$$\sum_{j=1}^{N} |v_j|^2 = \alpha^2 + |v|^2,$$

since $\sum_{j=1}^{N-1} |w_j|^2 = 1$ and $\sum_{j=1}^{N-1} v_j = 0$. Therefore, define

$$\alpha^2(v) = 1 - |v|^2 \tag{3.60}$$

and

$$\phi_N((w_1, w_2, \dots, w_{N-1}), v) =$$

$$\left(\alpha(v)\, w_1 - \frac{1}{\sqrt{N^2-N}} v, \dots, \alpha(v)\, w_{N-1} - \frac{1}{\sqrt{N^2-N}} v, \sqrt{\frac{N-1}{N}}\, v\right), \tag{3.61}$$

and we have that $\phi_N: X_{N-1} \times Y_N \to X_N$. For $j = 1, \dots, N-1$, let $\sigma_{j,N}$ be the pair permutation exchanging j and N, and define $\phi_j = \sigma_{j,N} \circ \phi_N$. We now show that with these definitions (3.32) holds, and in the process, obtain an explicit formula for ν_N.

The result [**15**] is contained in the following lemma:

Lemma. *For $N \geq 3$, the measure ν_N induced on Y_N through* (3.26) *for the Boltzmann collision model is*

$$\mathrm{d}\nu_N(v) = \frac{|S^{3N-7}|}{|S^{3N-4}|}(1-|v|^2)^{(3N-8)/2}\mathrm{d}v\,. \tag{3.62}$$

In the case $N=2$, ν_2 is the uniform probability measure on $S^2 = Y_2$. Moreover, for these measures ν_N, and with ϕ_j defined as above, (3.32) *holds for the Boltzmann collision model for all $N \geq 3$.*

Given these formulas, it is easy to check that we have in fact defined a Kac system. The proof of the following Lemma may be found in [**15**].

Lemma. *The Boltzmann collision model, consisting of $(X_N, \mathcal{S}_N, \mu_N)$, $(Y_N, \mathcal{C}_N, \nu_N)$, π_j, ϕ_j, $j = 1, \ldots, N$, and Q as specified in this section constitute a Kac system.*

Now to prove the analog of the Kac conjecture for physically realistic momentum conserving collisions, all we have to do is to analyze the spectrum of the corresponding K operator, and then to show that $\Delta_2 > 0$.

The explicit form of K for the Boltzmann collision model is easily obtained from (3.40): for all functions g on Y_N, the unit ball in $\boldsymbol{R}^3$, and all $N>3$,

$$\begin{aligned} Kg(v) &= \int_{X_{N-1}} g\left(\sqrt{\frac{N}{N-1}}\left(\sqrt{1-|v|^2}w_{N-1} - \frac{1}{\sqrt{N^2-N}}v\right)\right)\mathrm{d}\mu_{N-1}(w) \\ &= \int_{Y_{N-1}} g\left(\sqrt{\frac{N}{N-1}}\sqrt{1-|v|^2}\sqrt{\frac{N-2}{N-1}}y - \frac{1}{N-1}v\right) \\ &\qquad\qquad \cdot(1-|y|^2)^{(3N-11)/2}\mathrm{d}\nu_{N-1}(y) \\ &= \frac{|S^{3N-10}|}{|S^{3N-7}|}\int_{|y|\leq 1} g\left(\frac{\sqrt{N^2-2N}}{N-1}\sqrt{1-|v|^2}y - \frac{1}{N-1}v\right) \\ &\qquad\qquad \cdot(1-|y|^2)^{(3N-11)/2}\mathrm{d}y\,. \end{aligned} \tag{3.63}$$

Several properties of K are evident from (3.63). First, K commutes with rotations in $\boldsymbol{R}^3$. That is, if $R : \boldsymbol{R}^3 \to \boldsymbol{R}^3$ is a rotation, then clearly

$$K(g \circ R) = (Kg) \circ R\,.$$

For this reason, we may restrict our search for eigenfunctions of K to functions of

the form

$$g(v) = h(|v|)\,|v|^{\ell}\,\mathcal{Y}_{\ell,m}(v/|v|)$$

where h is a function on $\mathbb{R}_+$, and $\mathcal{Y}_{\ell,m}$ is a spherical harmonic.

Second, the space of polynomials of degree n is, for each n, invariant under K. Indeed, any monomial in $\sqrt{1-|v|^2}w$ that is of odd degree is annihilated when integrated against $(1-|w|^2)^{(3N-11)/2}\,\mathrm{d}w$, and any even monomial in $\sqrt{1-|v|^2}w$ is a polynomial in v.

Together, these two observations show that K has a complete basis of eigenfunctions of the form

$$g_{n,\ell,m}(v) = h_{n,\ell}(|v|^2)\,|v|^{\ell}\,\mathcal{Y}_{\ell,m}(v/|v|) \tag{3.64}$$

where $h_{n,\ell}$ is a polynomial of degree n.

A third observation identifies these polynomials and gives us a formula for the eigenfunctions: suppose that $Kg(v) = \lambda g(v)$. Let $\hat{e}$ be any unit vector in $\boldsymbol{R}^3$. Then since g is a polynomial and hence continuous,

$$\lim_{t\to 1} Kg(t\hat{e}) = \lim_{t\to 1} \frac{|S^{3N-10}|}{|S^{3N-7}|} \int_{Y_{N-1}} g\left(\frac{\sqrt{N^2-2N}}{N-1}\sqrt{1-t^2}w - \frac{1}{N-1}\,\hat{e}\right)\cdot$$

$$(1-|w|^2)^{(3N-11)/2}\,\mathrm{d}w = g\left(-\frac{1}{N-1}\,\hat{e}\right),$$

since $K1 = 1$. Then since $Kg(v) = \lambda g(v)$, we have

$$\lambda g(\hat{e}) = g\left(-\frac{1}{N-1}\,\hat{e}\right). \tag{3.65}$$

This leads to what is *on principle* an explicit formula for the eigenvalues of K. Consider any eigenfunction $g_{n,\ell,m}$ of the form given in (3.64), and let $\lambda_{n,\ell}$ be the corresponding eigenvalue, which will not depend on m. Then taking any $\hat{e}$ so that $\mathcal{Y}_{\ell,m}(\hat{e}) \neq 0$, we have from (3.65) that

$$\lambda_{n,\ell} = \frac{h_{n,\ell}(1/(N-1)^2)}{h_{n,\ell}(1)}\left(-\frac{1}{N-1}\right)^{\ell}. \tag{3.66}$$

This becomes quite explicit when we identify the polynomials $h_{n,\ell}$ as Jacobi polynomials. For all distinct positive integers n and p, the eigenfunctions $g_{n,\ell,m}$

and $g_{p,\ell,m}$ are orthogonal in $\mathcal{K}_N$. Hence for each ℓ, and for $n \neq p$,

$$\int_{|v| \leq 1} h_{n,\ell}(|v|^2)\, h_{p,\ell}(|v|^2)(1 - |v|^2)^{(3N-8)/2} |v|^{2\ell}\, \mathrm{d}v = 0 .$$

Taking $r = |v|^2$ as a new variable,

$$\int_0^1 h_{n,\ell}(r)\, h_{p,\ell}(r)(1-r)^{(3N-8)/2} r^{\ell+1/2}\, \mathrm{d}r = 0 .$$

This can be recognized as the orthogonality relation for a family of Jacobi polynomials in one standard form, and this identifies the polynomials $h_{n,\ell}$. A more familiar form is obtained using the change of variable $t = 2r - 1$ with t ranging over $[-1, 1]$. Then for $\alpha, \beta > -1$, $J_n^{(\alpha,\beta)}(t)$ is the orthogonal n-th degree polynomial for the weight $(1-t)^\alpha (1+t)^\beta$. Then with the variables t and $|v|^2$ related as above; i.e.,

$$t = 2|v|^2 - 1 , \tag{3.67}$$

$$h_{n,\ell}(|v|^2) = J_n^{(\alpha,\beta)}(t) \tag{3.68}$$

for

$$\alpha = \frac{3N-8}{2} \quad \text{and} \quad \beta = \ell + \frac{1}{2} . \tag{3.69}$$

The particular normalization of the Jacobi polynomials is irrelevant here, as we shall be concerned with ratios of the form $J_n^{(\alpha,\beta)}(t)/J_n^{(\alpha,\beta)}(1)$. Indeed, notice that from (3.67) when $|v|^2 = 1$, $t = 1$, and when $|v|^2 = 1/(N^2 - N)$, $t = -1 + 2/(N^2 - N)$. Hence from (3.68) and (3.66), we see that

$$\lambda_{n,\ell} = \frac{J_n^{(\alpha,\beta)}(-1 + 2/(N^2-N))}{J_n^{(\alpha,\beta)}(1)} \left(-\frac{1}{N-1}\right)^\ell . \tag{3.70}$$

This explicit formula leads to a proof of the Kac conjecture for momentum conserving collisions. For the details, see [**15**]. As it turns out, the details are more formidable than one might hope. The point is that it is not obvious from (3.70) that *in general*

$$|\lambda_{n,\ell+1}| \leq |\lambda_{n,\ell}|$$

or that

$$|\lambda_{n+1,\ell}| \leqslant |\lambda_{n,\ell}| .$$

There are explicit formulas for the Jacobi polynomials, so there is no difficulty in checking these relations for any particular n and ℓ. But the lack of a clear general monotonicity argument means that one must slog through a lot of details to arrive at a rigorous estimate of the κ_N in this case.

Again, for the details, see [**15**].

3.4 - *Entropy production in the Kac model*

The fact that the spectral gap of $\mathcal{L}_N$ is bounded away from zero as N tends to infinity implies a uniform exponential rate of approach to equilibrium in the L^2 sense. While the L^2 setting is suitable for studying solutions that are already close to equilibrium, entropy estimates are required to treat the wider class of all physical solutions, and are required to draw conclusion about the non linear Boltzmann equation.

To explain, let F be a probability density on S^{N-1}. Define the entropy of F, $S(F)$, by

$$S(F) = \int_{S^{N-1}} F \ln F(\vec{v}) \, \mathrm{d}\mu .$$

Being a convex functional of F, the entropy decreases along the evolution described by the master equation. A calculation gives $\dfrac{\mathrm{d}}{\mathrm{d}t} S(F) = \displaystyle\int_{S^{N-1}} \ln F \, \mathcal{L}_N F \, \mathrm{d}\mu$. Define the quantity Γ_N by

$$\Gamma_N = \sup_F \left\{ \frac{-\displaystyle\int_{S^{N-1}} \ln F \, \mathcal{L}_N F \, \mathrm{d}\mu}{S(F)} \right\}$$

where the supremum is taken over all densities F that are chaotic in the sense of Kac – that is, the ones leading to solutions of the Master equation that are relevant to the study of the non linear equation (3.6). For the original Kac model as described above, we make the entropic analog of the Kac conjecture:

$$\liminf_{N \to \infty} \Gamma_N = D > 0 .$$

This would imply that then entropy of chaotic solutions of the master equation decays to zero at a uniform exponential rate. This would have an important conse-

quence for the non linear equation (3.6): for a probability density f on the real line with zero mean and unit variance, define its relative entropy with respect to M, $H(f|M)$, by

$$H(f|M) = \int_R (\ln f - \ln M) f(v)\, dv\,.$$

Then the entropic version of the Kac conjecture would imply that for solutions $f(t)$ of (3.6)

$$H(f(t)|M) \leq Ce^{-Dt}. \tag{3.71}$$

What is presently known, for the Kac's model equation (3.6) and for the actual Boltzmann equation as well, is that for every $\alpha > 0$,

$$H(f(t)|M) \leq C_\alpha t^{-\alpha}, \tag{3.72}$$

where C_α grows rapidly with α.

Estimates of the type (3.72) play an important role in understanding the behavior of the spatially inhomogeneous Boltzmann equation, and a better estimate, such as (3.71), would be more incisive. For technical reasons, the route from the master equation to (3.72) must be slightly more complicated in the case of the full momentum conserving Boltzmann equation than what we have sketched above for Kac's model. Nonetheless, we still conjecture that estimates of the type (3.72) can be proved for the true Boltzmann equation via the analysis of entropy production for the master equation.

Efforts to prove (3.71) for the Kac model have already led to interesting results in analysis, and further work can be expected to yield even more.

In a direct adaptation of the L^2 strategy to the entropic setting, one is led to seek the best constant in an inequality for subadditivity of the entropy on S^{N-1}. Specifically, let F be any probability density on S^{N-1}, and let f_j be the corresponding marginal density for the j-th velocity. As proved in Carlen, Lieb and Loss,

$$\sum_{j=1}^{N} \int_{S^{N-1}} f_j \ln f_j \, d\mu \leq 2 \int_{S^{N-1}} F \ln F \, d\mu\,. \tag{3.73}$$

The constant 2 is best possible, and it is remarkable that this is independent of N. Indeed, viewing μ as a measure on R^N concentrated on $S^{N-1}(\sqrt{N})$, recall that μ

$\approx \gamma_N$ where

$$\gamma_N = \left(\frac{1}{2\pi}\right)^{N/2} e^{-|\vec{v}|^2/2} \mathrm{d}^N v \,.$$

The fact that $\mu \approx \gamma_N$; ie.e, that γ_N is concentrated more and more $S^{N-1}(\sqrt{N})$ with increasing N is known as the *equivalence of ensembles* in statistical mechanics. If one replaces μ by γ_N, then the coordinate functions v_j are independent as random variables, and the inequality (3.73) holds with the 2 replaced by (1). We conjecture that for an appropriately defined class of chaotic densities F, there is less dependence of the coordinate functions and that (3.73) holds with the constant 2 replaced by $1 + \mathcal{O}(1/N)$. This would lead to (3.71).

Very recently, Villani [**42**] has proved such an inequality for the case of *super hard* sphere.

References

[**1**] F. Abrahamsson, *Strong L^1 convergence to equilibrium without entropy conditions for the Boltzmann equation*, Comm. Partial Differential Equations **24** (1999), 1501-1535.

[**2**] R. Alexandre and C. Villani, *On the Boltzmann equation for long-range interactions*, Comm. Pure Appl. Math. **55** (2002), 30-70.

[**3**] R. Alexandre, L. Desvillettes, C. Villani and B. Wennberg, *Entropy dissipation and long-range interactions*, Arch. Ration. Mech. Anal. **152** (2000), 327-355.

[**4**] L. Arkeryd, *Intermolecular forces of infinite range and the Boltzmann equation*, Arch. Rational Mech. Anal. **77** (1981), 11-21.

[**5**] L. Arkeryd, *Stability in L^1 for the spatially homogeneous Boltzmann equation*, Arch. Rat. Mech. Anal. **103** (1988), 151-168.

[**6**] A. V. Bobylev, *Exact solutions of the Boltzmann equation*, Dokl. Akad. Nauk. SSSR **225** (1975), 1296-1299.

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

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

[**9**] F. Bouchut and L. Desvillettes, *A proof of the smoothing properties of the positive part of Boltzmann's kernel*, Rev. Mat. Iberoamericana **14** (1998), 47-61.

[**10**] T. Carleman, *Sur la solution de l'équation intégrodifférentielles de Boltzmann*, Acta Math. **60** (1933), 91-146.

[11] E. A. CARLEN and M. C. CARVALHO, *Strict entropy production bounds and stability of the rate of convergence to equilibrium for the Boltzmann equation*, J. Statist. Phys. **67** (1992), 575-608.

[12] E. A. CARLEN and M. C. CARVALHO, *Entropy production estimates for Boltzmann equations with physically realistic collision kernels*, J. Statist. Phys. **74** (1994), 743-782.

[13] E. A. CARLEN, M.C. CARVALHO and E. GABETTA, *Central limit theorem for Maxwellian molecules and truncation of the Wild expansion*, Comm. Pure Appl. Math. **53** (2000), 370-397.

[14] E. A. CARLEN, M. C. CARVALHO and E. GABETTA, *On the optimal rate of convergence in the Central Limit Theorem for Maxwellian Molecules*, preprint 2003.

[15] E. A. CARLEN, M. CARVALHO and M. LOSS, *The spectral gap in Kac's master equation and related spectral evolutions*, Acta Math. **191** (2003), 1-53.

[16] E. A. CARLEN, E. GABETTA and G. TOSCANI, *Propagation of smoothness and the rate of exponential convergence to equilibrium for a spatially homogeneous Maxwellian gas*, Comm. Math. Phys. **199** (1999), 521-546.

[17] E. A. CARLEN and XUGUANG LU, *Fast and slow convergence to equilibrium for Maxwellian molecules via Wild sums*, J. Statist. Phys. **112** (2003), 59-134.

[18] C. CERCIGNANI, *Theory and application of the Boltzmann equation*, Elsevier, New York 1975.

[19] T. ELMROTH, *Global boundedness of moments of solutions of the Boltzmann equation for forces of infinite range*, Arch. Rational Mech. Anal. **82** (1983), 1-12.

[20] W. FELLER, *Completely monotone functions and sequences*, Duke Math. J. **5** (1939), 661-674

[21] E. GABETTA, G. TOSCANI and B. WENNBERG, *Metrics for probability distributions and the trend to equilibrium for solutions of the Boltzmann equation*, J. Statist. Phys. **81** (1995), 901-934.

[22] T. GOUDON, *On Boltzmann equations and Fokker-Planck asymptotics: influence of grazing collisions*, J. Statist. Phys. **89** (1997), 751-776.

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

[24] E. IKENBERRY and C. TRUESDELL, *On the pressure and flux of energy in a gas according to Maxwell's kinetic theory*, J. Rational Mech. Analysis. Phys. **5** (1956), 1-54.

[25] E. JANVRESSE, *Spectral gap for Kac's model of Boltzmann equation*, Ann. Probab. **29** (2001), 288-304.

[26] P. L. LIONS, *Compactness in Boltzmann's equation via Fourier integral operators and applications I*, J. Math. Kyoto Univ. **34** (1994), 391-427.

[27] X. G. LU, *A direct method for the regularity of the gain term in the Boltzmann equation*, J. Math. Anal. Appl. **228** (1998), 409-435.

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

[29] H. P. MCKEAN Jr, *An exponential formula for solving Boltzmann's equation for a Maxwellian gas*, J. Combinatorial Theory **2** (1967), 358-382.

[30] H. MURATA and H. TANAKA, *An inequality for a certain functional of a multidimensional probability distributions*, Hiroshima Math J. **4** (1974), 75-81.

[31] H. TANAKA, *An inequality for a functional of a probability distributions and it's application to Kac's one dimensional model of a Maxwellian gas*, Z. Wahrsch. Verw. Gebiete **27** (1973), 47-52.

[32] H. TANAKA, *Probabilistic treatment of the Boltzmann equation for Maxwellian molecules*, Z. Wahrsch. Verw. Gebiete **46** (1978/79), 67-105.

[33] G. TOSCANI and C. VILLANI, *Probability metrics and uniqueness of the solution to the Boltzmann equation for a Maxwell gas*, J. Statist. Phys. **94** (1999), 619-637.

[34] G. TOSCANI and C. VILLANI, *Sharp entropy dissipation bounds and explicit rate of trend to equilibrium for the spatially homogeneous Boltzmann equation*, Comm. Math. Phys. **203** (1999), 667-706.

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

[36] B. WENNBERG, *Stability and exponential convergence in L^p for the spatially homogeneous Boltzmann equation*, Nonlinear Anal. **20** (1993), 935-961.

[37] B. WENNBERG, *Regularity in the Boltzmann equation and the Radon transform*, Comm. Partial Differential Equations **19** (1994), 2057-2074.

[38] B. WENNBERG, *Entropy dissipation and moment production for the Boltzmann equation*, J. Statist. Phys. **86** (1997), 1053-1066.

[39] E. WILD, *On Boltzmann's equation in the kinetic theory of gases*, Proc. Cambridge Philos. Soc. **47** (1951), 602-609.

[40] C. VILLANI, *On a new class of weak solutions to the spatially homogeneous Boltzmann and Landau equations*, Arch. Rational Mech. Anal. **143** (1998), 273-307.

[41] C. VILLANI, *A review of mathematical topics in collisional kinetic theory*, Handbook of mathematical fluid dynamics, Vol. I, 71-305, North-Holland, Amsterdam 2002.

[42] C. VILLANI, *Cercignani's conjecture is sometimes true, and always almost true*, Comm. Math. Phys. **234** (2003), 455-490.

[43] E. H. T. YAU, *Logarithmic Sobolev inequality for generalized simple exclusion processes*, Probab. Theory Related Fields **109** (1997), 507-538.

[44] E. H. T. YAU, *The logarithmic Sobolev inequality for lattice gasses with mixing conditions*, Comm. Math. Phys. **181** (1976), 367-408.

Summary

We present a review of probabilistic methods in kinetic theory focusing on problems concerning the spatially homogeneous Boltzmann equation. The material was presented by the first author at the Porto Ercole Summer School in June 2002.

* * *