ON THE RELATION BETWEEN RATES OF RELAXATION AND CONVERGENCE OF WILD SUMS FOR SOLUTIONS OF THE KAC EQUATION

E.A. Carlen[1]

School of Mathematics, Georgia Institute of Technology
Atlanta, Georgia 30332, USA

M.C. Carvalho[2]

Departamento de Matematica, Universidade de Lisboa
Lisboa codex 1699, PORTUGAL

E. Gabetta[2,3]

Dipartimento di Matematica - Università di Pavia
via Abbiategrasso 215, 27100 Pavia, ITALY

Abstract: In the case of Maxwellian molecules, the Wild summation formula gives an expression for the solution of the spatially homogeneous Boltzmann equation in terms of its initial data F as a sum $f(v,t) = \sum_{n=0}^{\infty} e^{-t}(1-e^{-t})^n Q_n^+(F)(v)$. Here, $Q_n^+(F)$ is an average over n–fold iterated Wild convolutions of F. If M denotes the Maxwellian equilibrium corresponding to F, then it is of interest to determine bounds on the rate at which $\|Q_n^+(F) - M\|_{L^1(I\!R)}$ tends to zero. In the case of the Kac model, we prove that for every $\epsilon > 0$, if F has moments of every order and finite Fisher information, there is a constant C so that for all n, $\|Q_n^+(F) - M\|_{L^1(I\!R)} \leq Cn^{\Lambda+\epsilon}$ where Λ is the least negative eigenvalue for the linearized collision operator. We show that Λ is the best possible exponent by relating this estimate to a sharp estimate for the rate of relaxation of $f(\cdot,t)$ to M. A key role in the analysis is played by a decomposition of $Q_n^+(F)$ into a smooth part and a small part. This depends in an essential way on a probabilistic construction of McKean. It allows us to circumvent difficulties stemming from the fact that the evolution does not improve the qualitative

[1] Work partially supported by U.S. National Science Foundation grants DMS 03-00349 and G-37-X43.
[2] Work partially supported the European Community Human Potential Program under contract HPRN–CT–2002–00282 (HYKE)
[3] Work supported by the Ministry of University (MURST) "Problemi matematica delle teorie cinetiche"

regualrity of the initial data.

1. Introduction

The Kac equation is a model Boltzmann equation due to Mark Kac [15]. It describes the evolution of the probability density $f(v,t)$ for the velocities in a gas of particles moving in one dimension. It has the form:

$$\frac{\partial}{\partial t} f(v,t) = \int_{-\pi}^{\pi} \left(\int_R \left[f(v^*(\theta),t) f(w^*(\theta),t) - f(v,t)f(w,t) \right] \mathrm{d}w \right) \rho(\theta)\mathrm{d}\theta \ . \tag{1.1}$$

Here, ρ is an even probability measure on $[-\pi,\pi]$, and the post-collisional velocities are given by

$$v^*(\theta) = v\cos(\theta) + w\sin(\theta) \qquad \text{and} \qquad w^*(\theta) = -v\sin(\theta) + w\cos(\theta) \ . \tag{1.2}$$

The requirement that ρ be even reflects microscopic reversibility; the time reversal of the collision in (1.2) should have the same probability.

In addition, we require that when ρ is extended periodically,

$$\rho(\theta + \pi/2) = \rho(\theta) \tag{1.3}$$

for all θ. Under this condition together with the condition that ρ be even, if $E = v^2 + w^2$, and a is any number with $0 \le a \le 1$, each of the post collisional outcomes

$$(v^*, w^*) = (\pm\sqrt{aE}, \pm\sqrt{(1-a)E}) \qquad \text{and} \qquad (v^*, w^*) = (\pm\sqrt{(1-a)E}, \pm\sqrt{aE})$$

is equally likely. This is natural for the Kac model, in which the collisions conserve energy and mass but not momentum.

Finally, for technical reasons we make the requirement that ρ is uniformly bounded; i.e., that for some finite constant B, $\rho(\theta) \le B$. We refer to probability densities ρ satisfying all of these conditions as *regular*. In the theorems below, it is always assumed ρ is regular.

Because the underlying collisions conserve mass and energy, so does the evolution described by (1.1). That is, for any solution f, $\int_R f(v,t)\mathrm{d}v$ and $\int_R v^2 f(v,t)\mathrm{d}v$ are constant. By a choice of units, we may assume without loss of generality that

$$\int_R f(v,0)\mathrm{d}v = 1 \qquad \text{and} \qquad \int_R v^2 f(v,0)\mathrm{d}v = 1 \ . \tag{1.4}$$

We shall assume this throughout the paper. The other moments of f are not conserved. In particular, the total momentum $\int_R v f(v,t)\mathrm{d}v$ is not constant for the Kac model, unless it happens to be zero.*

Kac introduced his model to study rates of relaxation to equilibrium in kinetic theory. There is only one equilibrium solution to (1.1) for initial data satisfying (1.4). This is the normalized Maxwellian

$$M(v) = \frac{1}{\sqrt{2\pi}} e^{-v^2/2} \ . \tag{1.5}$$

Linearizing the Kac equation about $M(v)$ by considering initial data of the form

$$f(v) = M(v)(1 + h(v))$$

with $h(v)$ "small" leads to the linearized Kac equation:

$$\frac{\partial}{\partial t} h(v,t) = \mathcal{L} h(v,t) \tag{1.6}$$

where

$$\begin{aligned}\mathcal{L}h(v) = &\int_{-\pi}^{\pi} \left(\int_R M(w) \left[h(w^*(\theta)) + h(v^*(\theta)) \right] \mathrm{d}w \right) \rho(\theta) \mathrm{d}\theta \\ &- \left(\int_R h(w) M(w) \mathrm{d}w \right) - h(v) \ .\end{aligned}$$

The first term in $\mathcal{L}$ can be recognized as being composed of averages of Mehler kernels [16], which implies that all of the eigenfunctions are Hermite polynomials. The eigenvalue λ_n corresponding to the nth degree Hermite polynomial for the operator $\mathcal{L}$ is readily worked out in terms of the eigenvalues of the Mehler kernel, and found [16] to be

$$\lambda_n = \int_{-\pi}^{\pi} \left(\sin^n(\theta) + \cos^n(\theta) - 1 \right) \rho(\theta) \mathrm{d}\theta \qquad \text{for} \qquad n \geq 1 \ , \tag{1.7}$$

and $\lambda_0 = 0$, which is a simple special case. Conservation of energy is reflected in the fact that $\lambda_2 = 0$, and conservation of mass is reflected in the fact that $\lambda_0 = 0$.

As for the non zero eigenvalues, notice that since ρ is even, $\int \sin^n(\theta)\rho(\theta)\mathrm{d}\theta = 0$ for all odd values of n. Under the additional condition (1.3), the same is true of

* With one dimensional velocities in a pair collision, conservation of both momentum and energy would leave only exchange of the two velocities as kinematically possible collision result, and that would trivialize the equation.

$\int \cos^n(\theta)\rho(\theta)\mathrm{d}\theta$. Hence, under our conditions $\lambda_n = -1$ for all odd values of n. For even values of n; i.e., $n = 2k$, it is clear that the integrand in (1.7) is monotone decreasing in k. Therefore, the largest non zero eigenvalue of $\mathcal{L}$ is λ_4. That is, if we define

$$\Lambda = \sup_{n\neq 0,2} \left\{ \int_{-\pi}^{\pi} \left(\sin^n(\theta) + \cos^n(\theta) - 1\right) \rho(\theta)\mathrm{d}\theta \right\} \ , \tag{1.8}$$

which is the second largest; i.e., the *least negative*, eigenvalue of $\mathcal{L}$, we have $\Lambda = \lambda_4$.

The *Wild convolution* $f \circ g$ of two probability densities f and g on R is defined by

$$f \circ g(v) = \int_{-\pi}^{\pi} \left(\int_R f(v^*(\theta))g(w^*(\theta))\mathrm{d}w \right) \rho(\theta)\mathrm{d}\theta \ .$$

This adapts to the Kac model the original definition of Wild [19], which was made in the context of Maxwellian molecules. While this product is commutative under the condition (1.3), it is never associative: Even in the uniform case $\rho(\theta) = 1/(2\pi)$ originally considered by Kac, the product is not associative. This point is important in what follows, and we shall return to it.

For now, observe that using the Wild convolution, (1.1) can be written in the form

$$\frac{\partial}{\partial t} f(v,t) = f \circ f(v,t) - f(v,t) \ . \tag{1.9}$$

If the initial condition is $f(v,0) = F(v)$, then Wild's expression [19] for the solution is

$$f(v,t) = \sum_{n=1}^{\infty} e^{-t}(1-e^{-t})^{n-1} Q_n^+(F)(v) \tag{1.10}$$

where $Q_n^+(F)$ is a certain explicitly described average over all the n–fold Wild convolutions of F. (Since the product is not associative, in general $F \circ (F \circ (F \circ F)) \neq (F \circ F) \circ (F \circ F)$). Specifically, we have the inductive definition

$$Q_n^+(F) = \frac{1}{n-1} \sum_{j=1}^{n-1} Q_{n-j}^+(F) \circ Q_j^+(F) \qquad n \geq 2 \tag{1.11}$$

starting from $Q_1^+(F) = F$.

There is an alternate expression for $Q_n^+(F)$ due to McKean [17], and it plays an essential role in what follows. His formula expresses $Q_n^+(F)$ directly as a weighted average over the various n–fold Wild convolutions of F. Since the Wild convolution is not associative, there are as many of these as there are binary bracketings of n

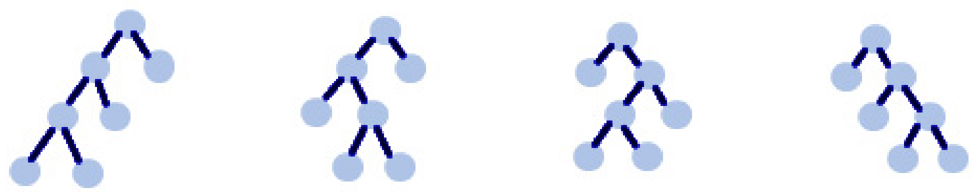
ordered factors. A classic result of Catalan is that there are $c_n = \frac{1}{n}\binom{2n-2}{n-1}$ of these. For example, for $n = 4$, $c_n = 5$, and there are five such convolutions:

$$((F \circ F) \circ F) \circ F \qquad (F \circ (F \circ F)) \circ F \qquad F \circ ((F \circ F) \circ F) \qquad F \circ (F \circ (F \circ F)) \tag{1.12}$$

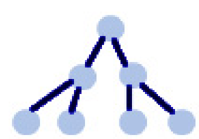
and

$$(F \circ F) \circ (F \circ F) \ . \tag{1.13}$$

These binary bracketings can be put in one to one correspondence with certain tree graphs. Here are the graphs corresponding to the four convolutions in (1.12).



They are arranged in the same order, left to right. Next, here is the graph corresponding to the balanced convolution (1.13):



To go from a graph to the corresponding convolution, label each of the leaves with F. Then start erasing pairs of leaves. This exposes a node as a new leaf. Label that new leaf with $(G \circ H)$ where G is the label on the left leaf that was erased, and H is the label on the right leaf that was erased. When all pairs of leaves are erased, one has the convolution. Let Γ_n denote the set of all graphs of this type with n leaves. For any $\gamma \in \Gamma_n$, let $C_\gamma(F)$ denote the corresponding convolution of F. McKean's formula then is

$$Q_n^+(F) = \sum_{\gamma \in \Gamma_n} P(\gamma) C_\gamma(F) \ , \tag{1.14}$$

where $P(\gamma)$ is the probability that a certain random walk on graphs passes through γ. See [17] or [4] for more details.

The direct averaging formula (1.14) has a significant advantage over the recursive formula (1.11). As we shall show, it can be used to generate decompositions of $Q_n^+(F)$ into "good" and "bad" pieces, whose relative sizes may be estimated through a probabilistic analysis of the random walk that define the weights $P(\gamma)$.

These decompositions shall be generated by *partitioning* Γ_n into two appropriately chosen subsets That is, suppose that $\Gamma_n = A_n \cup B_n$ with $A_n \cap B_n = \emptyset$. Then with $p_n = \sum_{\gamma \in A_n} P(\gamma)$, we put

$$G_n = \frac{1}{p_n} \sum_{\gamma \in A_n} P(\gamma) C_\gamma(F) \qquad \text{and} \qquad H_n = \frac{1}{1-p_n} \sum_{\gamma \in B_n} P(\gamma) C_\gamma(F) \ .$$

Then G_n and H_n are probability densities, and

$$Q_n^+(F) = p_n G_n + (1-p_n) H_n \ .$$

We shall show that if we take A_n to be the set of graphs in which every leaf has a certain minimum depth, then G_n will be smooth, and when n is large, p_n will be very close to 1. This smoothness result is crucial in what follows, and without some sort of decomposition, we would not have it. It simply is not the case that $Q_n^+(F)$ becomes progressively more smooth with increasing n. If the initial data F does not posses a certain degree of smoothness, then in general, neither will $Q_n^+(F)$, no matter how large n is. However, what *does* happen is almost as good: A large piece of it – $p_n G_n$ – becomes smooth, and the L^1 norm of the remainder, which is $(1-p_n)$, shrinks to zero at a very rapid rate.

This smoothing result will be applied to study the relation between the rates of convergence in

$$\lim_{n \to \infty} \| Q_n^+(F) - M \|_{L^1(R)} = 0 \tag{1.15}$$

and

$$\lim_{t \to \infty} \| f(\cdot, t) - M \|_{L^1(R)} = 0 \ . \tag{1.16}$$

The idea that one might obtain precise information on the rate in the second limit by a study of the first limit was developed by McKean [16]. He conjectured that (1.15) should result from a sort of a "central limit theorem", and that it should imply an exponential rate of convergence in (1.16). A previous paper [4] has verified McKean's conjecture.

Further work on the relation between (1.15) and (1.16) has been done in [6]. There, the emphasis was on initial data with "long tails" so that while

$$\lim_{R \to \infty} \int_{|v| > R} v^2 F(v) \mathrm{d}v = 0 \ , \tag{1.17}$$

the rate of decrease is very slow. It was proved in [6] for Maxwellian molecules that the rate of convergence in (1.17) essentially determines the rate of decay in (1.16), which can be arbitrarily slow if $\int_{|v|>R} v^2 F(v)\mathrm{d}v$ decreases to zero sufficiently slowly in R. This analysis produced the first example of solutions of the Boltzmann equation that relax to equilibrium at a sub exponential rate.

Here, our emphasis is instead on initial data possessing moments of every order so that the convergence in (1.17) is very fast. Then the decay in (1.16) will be exponential. Indeed, by a result in [6], this is ensured if $\int_R |v|^{2+\epsilon} F(v)\mathrm{d}v < \infty$ for some $\epsilon > 0$.

Our goal here is to determine the precise exponential rate for such data. We shall do this by determining the precise polynomial rate in (1.15). Our main result is that for reasonable initial data, both rates are governed by Λ as defined in (1.8).

Before stating the theorems, we recall that the Fisher information $I(F)$ of a probability density F on $I\!R$ is defined by

$$I(F) = 4\int_{I\!R} |(F^{1/2}(v))'|^2 \mathrm{d}v = \int_{I\!R} \frac{|F'(v)|^2}{F(v)} \mathrm{d}v \ .$$

The Fisher information is closely related to the entropy, and has played a fundamental role in the study of the Kac equation since the work of McKean [16]. Two fundamental properties of the Fisher information are that *(i)* $I(F)$ is a convex functional of F, and *(ii)*, for any two probability densities F and G,

$$I(F \circ G) \leq \frac{1}{2}\left(I(F) + I(G)\right) \ .$$

This leads directly to the fact that

$$I(C_\gamma(F)) \leq I(F) \ .$$

Likewise, it follows that the solution $f(v,t)$ of the Kac equation with initial data F satisfies $I(f(\cdot,t)) \leq I(F)$ for all t; i.e., the Fisher information of a solution of the Kac equation is monotone decreasing, as McKean discovered.

Theorem 1.1 *Let F be a probability density with finite Fisher information such that $\int_R v^2 F(v)\mathrm{d}v = 1$, and possessing moments of every order. Then for any $\epsilon > 0$, there is a finite constant C depending only on the behavior of the moments of F and on ϵ so that for all $n \geq 1$,*

$$\|Q_n^+(F) - M\|_{L^1(R)} \leq C n^{\Lambda+\epsilon} \ . \tag{1.18}$$

Theorem 1.2 *Let F satisfy the hypotheses of Theorem 1.1. Let $\epsilon > 0$ and C be the constants in (1.18), and let $f(v,t)$ be the solution of (1.1) with initial data F. Then*

$$\|f(\cdot,t) - M\|_{L^1(R)} \leq C\frac{2+\Lambda}{1+\Lambda} e^{(\Lambda+\epsilon)t} \ . \tag{1.19}$$

It is a relatively easy task to prove Theorem 1.2 given Theorem 1.1. We do this now as it explains the relation between (1.15) and (1.16).

Proof of Theorem 1.2: The starting point is the Wild sum (1.10). To use this, note that $M = \sum_{n=1}^{\infty} e^{-t}(1-e^{-t})^{n-1}M$, which is a special case (1.10) since $Q_n^+(M) = M$ for all n. Then by the Minkowski inequality and Theorem 1.1,

$$\begin{aligned} \|f(\cdot,t) - M\|_{L^1(R)} &\leq \sum_{n=1}^{\infty} e^{-t}(1-e^{-t})^{n-1}\|Q_n^+(F) - M\|_{L^1(R)} \\ &\leq \sum_{n=1}^{\infty} e^{-t}(1-e^{-t})^{n-1}Cn^{\Lambda+\epsilon} \ . \end{aligned} \tag{1.20}$$

Fix a positive integer N, and split the sum at the Nth term. Estimating the tail,

$$\sum_{n=N+1}^{\infty} e^{-t}(1-e^{-t})^{n-1}Cn^{\Lambda+\epsilon} \leq CN^{\Lambda+\epsilon} \ , \tag{1.21}$$

since $\sum_{n=N+1}^{\infty} e^{-t}(1-e^{-t})^{n-1} \leq 1$, and since $n^{\Lambda+\epsilon}$ is monotone decreasing.

As for the first n terms, note that for any p with $0 < p < 1$,

$$\sum_{n=1}^{N} n^{-p} = 1 + \sum_{n=2}^{N} n^{-p} \leq 1 + \int_1^N x^{-p}\mathrm{d}x \leq \frac{1}{1-p}N^{1-p} \ .$$

Applying this with $-p = \Lambda + \epsilon$,

$$\sum_{n=1}^{N} e^{-t}(1-e^{-t})^{n-1}Cn^{\Lambda+\epsilon} \leq e^{-t}C\sum_{n=1}^{N} n^{\Lambda+\epsilon} \leq e^{-t}\frac{C}{1+\Lambda+\epsilon}N^{1+\Lambda+\epsilon} \leq \ . \tag{1.22}$$

Combining (1.20), (1.21) and (1.22), we have that

$$\|f(\cdot,t) - M\|_{L^1(R)} \leq CN^{\Lambda+\epsilon} + e^{-t}C\frac{1}{1+\Lambda+\epsilon}N^{1+\Lambda+\epsilon} \ .$$

Choosing $N = e^t$, this is

$$\begin{aligned} \|f_t - M\|_{L^1(R)} &\leq Ce^{(\Lambda+\epsilon)t} + e^{-t}C\frac{1}{1+\Lambda+\epsilon}e^{(1+\Lambda+\epsilon)t} \\ &\leq C\frac{2+\Lambda+\epsilon}{1+\Lambda+\epsilon}e^{(\Lambda+\epsilon)t} \leq C\frac{2+\Lambda}{1+\Lambda}e^{(\Lambda+\epsilon)t}. \end{aligned}$$

■

This proves that the bound obtained in Theorem 1.1 is the best possible in its dependence on n, since any better exponent in (1.18) would lead to a rate of relaxation that would be inconsistent with the bounds provided by the linearized Boltzmann equation.

To prove Theorem 1.1, we first prove an analog for a weaker norm. Then we use the existence of higher order moments and a result on production of smoothness by repeated Wild convolutions to obtain the decay in the L^1 norm. The next section introduces the weak norm, and proves the weak norm analog of Theorem 1.

To pass from the weak norm to the strong norm, we can use interpolation methods, provided we have smoothness and moment bounds. Moment bounds are physically reasonable on the initial data, but smoothness bounds are less so. Therefore, it is fortunate that for large n we can decompose $Q_n^+(F)$ into two pieces: One will be smooth, and the other will be small in the L^1 norm. This is proved in Theorem 3.1 using the decomposition strategy outlined above.

Section 3 of the paper is devoted to this decomposition theorem. Such a decomposition is necessary if one is to prove any sort of "production of smoothness" result for solutions of the Kac equation, or for that matter, the Boltzmann equation for Maxwellian molecules, for which there is also a Wild sum. This can be clearly seen from (1.10). The first term in the Wild sum is $e^{-t}F$, and so the solution will never belong to any Sobolev space to which F itself does not already belong.

While the Kac equation is not regularizing in the manner of a parabolic equation, the intuition that the collisions must be doing some sort of smoothing is correct. This can be seen most directly by looking at $Q_n^+(F)$.

Our decomposition of $Q_n^+(F)$ is induced by a partition of the set Γ_n in (1.14). We shall show that graphs in which every leaf has a certain minimum depth make a smooth contribution to the sum, and that for n large, most of the graphs are of this nice type. This last fact is established by a probabilistic analysis afforded by McKean's representation. It is not clear how one might do this using only (1.11), which does suffice for many purposes; In [6], the Wild sum estimates were all based on the simpler (1.11). For the present purposes, (1.14) seems more incisive. One might well regard the results in Section 3 as the main results of the paper.

In section 4, we prove the interpolation bounds, and in Section 5 we prove Theorem 1.1. We close the introduction with some final remarks on the production of smoothness in kinetic equations.

There is a considerable literature on production of smoothness for the Boltzmann equation, starting with work of Desvillettes [8] on the Kac Model. His results concern the *non cut–off* Kac equation, which means that the density $\rho(\theta)$ is unbounded near $\theta = 0$. This results in a great many "grazing collisions" in which the velocities are barely changed. In this case, the smoothing properties of the evolution are more like those of a parabolic equation. Indeed, in the "grazing collision limit" [13], the Boltzmann equation becomes a non linear parabolic equation known as the Landau equation whose smoothing properties have been investigated by Desvillettes and Villani [9]. When ρ is sufficiently singular at $\theta = 0$, the smoothing behavior is parabolic, with the solution possessing derivatives of every order for all $t > 0$.

When ρ is not singular at $\theta = 0$, then parabolic like smoothing will not occur. Therefore, the very nature of "production of smoothing" results must be different in our setting. It is quite satisfactory that a natural decomposition can be used to formulate a "production of smoothing" result here.

It is important to be able to deal with the case in which ρ is not singular at $\theta = 0$. First, screening effects will likely eliminate the long range interaction needed to produce the singularities in ρ in many physical settings. Second, the singularities in ρ that are necessary for parabolic— like production of smoothness are rather severe, so that only weak solutions [18] of these kinetic equations can be produced.

2. The sharp rate in a weak norm

In this section, we refine a method developed in [4] for estimating $Q_n^+(F) - M$. We do this in a weak norm that permits close estimation, and will later use "production of smoothness" estimates from the next section to draw conclusions about the rate in the strong L^1 norm.

The basic strategy is to find a convex functional Φ on probability densities with the property that there is a constant $c < 1$ such that for any two probability densities F and G

$$\Phi(F \circ G) \leq \frac{c}{2}\left(\Phi(F) + \Phi(G)\right) \ . \tag{2.1}$$

Then Theorem 1.9 of [4] gives us the following estimate: For all $\epsilon > 0$, there is a finite constant A_ϵ

$$\Phi(Q_n^+(F)) \leq A_\epsilon n^{c-1+\epsilon} \ . \tag{2.2}$$

Suppose also that there is some norm $|\!|\!| \cdot |\!|\!|$ so that for all $n > 1$,

$$|\!|\!| Q_n^+(F) - M |\!|\!| \leq \Phi(Q_n^+(F)) \ . \tag{2.3}$$

Then we conclude the analog of Theorem 1.1 for the $|\!|\!| \cdot |\!|\!|$:

$$|\!|\!| Q_n^+(F) - M |\!|\!| \leq A_\epsilon n^{c-1+\epsilon} \ . \tag{2.4}$$

We now explain how to do this, first introducing the norm $\|\cdot\|$.

Let $\mathcal{M}_4$ be the space of $L^1(R)$ functions g such that

$$\int_R |v|^4 |g(v)| \mathrm{d}v < \infty \qquad \text{and} \qquad \int_R (1, v, v^2, v^3) g(v) \mathrm{d}v = (0,0,0,0) \ . \tag{2.5}$$

Let $\widehat{g}$ denote the Fourier transform of g. Then under (2.5), the first four terms in the Taylor expansion of $\widehat{g}$ at the origin all vanish:

$$\widehat{g}(\xi) = \mathcal{O}(|\xi|^4) \ .$$

Therefore, it makes sense to equip $\mathcal{M}_4$ with the norm $\|\cdot\|$ where

$$\|g\| = \sup_{\xi \neq 0} \frac{\widehat{g}(\xi)}{|\xi|^4} \ . \tag{2.6}$$

We now claim that for $n > 1$, $Q_n^+(F) - M$ belongs to $\mathcal{M}_4$. The key is that since ρ is regular

$$\int_\pi^\pi (\cos(\theta) v + \sin(\theta) w) \rho(\theta) \mathrm{d}\theta = \int_\pi^\pi (\cos(\theta) v + \sin(\theta) w)^3 \rho(\theta) \mathrm{d}\theta = 0$$

and so for $n > 1$

$$\int_{I\!R} v Q_n^+(F) \mathrm{d}v = \int_{I\!R} v^3 Q_n^+(F) \mathrm{d}v = 0 \ . \tag{2.7}$$

Of course since M is even, its first and third moments vanish as well. Then since $Q_n^+(F)$ and M are both probability measures with the same variance, their zeroth and second moments agree as well. Hence for $n > 1$, $Q_n^+(F) - M$ belongs to $\mathcal{M}_4$. (The fact that for regular ρ, all odd moments of $F \circ G$ vanish for any two probability densities F and G has been pointed out and exploited in [7]).

We would take our functional Φ to be $\Phi(F) = \|F - M\|$ except for the fact that F itself can have non zero first and third moments, in which case it will not belong to $\mathcal{M}_4$, and $\|F - M\|$ is not defined.

To correct for this, we make a small correction to F so that $\widehat{F}$ has a suitable Taylor expansion at the origin. This could be avoided in the Kac model by just ignoring $n = 1$ in (2.3) and making some corresponding modifications to the arguments in [4]. However, a simple correction that works also for Maxwellian molecules has already been devised in [4] , which in turn draws on ideas of [12] . Here is how it is done:

For any probability density F, let $m_k(F)$ denote the kth moment of F: $m_k(F) = \int_R v^k F(v) \mathrm{d}v$. Define $\mu(F)$ by

$$\mu(F) = \left(m_1^2(F) + m_3^2(F)\right)^{1/2} \ . \tag{2.8}$$

We define a function P_F by

$$\widehat{P}_F(\xi) = \left(i m_1(F)\xi - \frac{i m_3(F)}{6}\xi^3 \right) \chi(|\xi|) \ , \tag{2.9}$$

where χ is a C^∞ monotone decreasing function on R_+ such that for $r \le L_1$, $\chi(r) = 1$, while for $\chi(r) = 0$ for $r \ge L_2$, and where $0 < L_1 < L_2$ will be specified below.

Observe that by the Schwarz inequality and the support properties of χ,

$$|\widehat{P}_F(\xi)| \le |m_1(F)|L_2 + |m_3(F)|L_2^3 \le \mu(F)(L_2^2 + L_2^6)^{1/2} \ , \tag{2.10}$$

so $\mu(F)$ controls the size of $\widehat{P}_F$.

Notice that for any probability density F with a finite fourth moment, $(F - P_F) - M \in \mathcal{M}_4$. We now define a functional Φ on such densities as follows:

$$\Phi(F) = |||(F - P_F) - M||| + K\mu(F) \ . \tag{2.11}$$

The constant K will be chosen below.

Theorem 2.1 *For any $\epsilon > 0$, and all probability densities F and G satisfying*

$$\int_R v^4 F(v)\mathrm{d}v \le C \qquad \text{and} \qquad \int_R v^4 G(v)\mathrm{d}v \le C \tag{2.12}$$

there is finite value, depending only on ϵ and C, of the constant K in (2.11) so that

$$\Phi(F \circ G) \le \frac{\Lambda + 1 + \epsilon}{2}\left(\Phi(F) + \Phi(G)\right) \ . \tag{2.13}$$

Finally, for any $\epsilon > 0$, there is a constant A_ϵ depending only on ϵ so that for all n,

$$\Phi(Q_n^+(F)) \le A_\epsilon n^{\Lambda + \epsilon} \ . \tag{2.14}$$

To extract from (2.14) a result on $|||Q_n^+(F) - M|||$, notice that (2.7) implies that for all $n > 1$, $\mu(Q_n^+(F)) = 0$ and $P_{Q_n^+(F)} = 0$, and hence from (2.11),

$$\Phi(Q_n^+(F)) = |||Q_n^+(F) - M||| \ . \tag{2.15}$$

Therefore, we have the following result:

Theorem 2.2 *For any $\epsilon > 0$, and all probability densities F satisfying $\int_R v^4 F(v)\mathrm{d}v < \infty$ there is finite a constant A_ϵ depending only on ϵ so that for all n,*

$$|||Q_n^+(F) - M||| \le A_\epsilon n^{\Lambda + \epsilon} \ . \tag{2.16}$$

Proof of Theorem 2.1: Comparing (2.1) and (2.2) with (2.13), (2.14) follows once we have (2.13).

First, $|m_1(F)| \leq C^{1/4}$ and $|m_3(F)| \leq C^{3/4}$, so $\mu(F)$ is bounded, with the bound depending only on C. We now choose L_2 small enough so that (2.10) ensures that

$$\|\widehat{P}_F\|_\infty \leq \epsilon \ . \tag{2.17}$$

Since the bounds obtained above depend on F only through C, they apply for G as well.

Next, recall that since $F \circ G$ is even, $P_{F \circ G} = 0$. Adding and subtracting, we have

$$\begin{aligned} F \circ G &= ((F - P_F) + P_F) \circ ((G - P_G) + P_G) \\ &= (F - P_F) \circ (G - P_G) + (F - P_F) \circ P_G \\ &+ P_F \circ (G - P_G) + P_F \circ P_G \ . \end{aligned} \tag{2.18}$$

We need to show that the last three terms are small in the $\|\!| \cdot |\!\|$ norm. Consider the first of these, $(F - P_F) \circ P_G$, and consider ξ with $|\xi| \leq L_1$. We will compute the Wild convolutions with the Bobylev formula [1]

$$\widehat{F \circ G}(\xi) = \int_{-\pi}^{\pi} \widehat{F}(\cos(\theta)\xi)\widehat{G}(\sin(\theta)\xi)\rho(\theta)\mathrm{d}\theta \ . \tag{2.19}$$

It will be convenient to denote the right hand side by $\widehat{F} \circ \widehat{G}$, and we do this below.

Note that when $|\xi| \leq L_1$, so are $|\sin(\theta)\xi|$ and $|\cos(\theta)\xi|$, and hence for such ξ, $\chi(|\sin(\theta)\xi|) = 1$, where χ is the cut–off function in (2.9). Then, by the definition (2.9),

$$\widehat{P}_G(\sin(\theta)\xi) = i\left(m_1(G)\xi\sin(\theta) - \frac{m_3(G)}{6}\xi^3\sin^3(\theta)\right) \ .$$

By Taylor's theorem with remainder, for such ξ,

$$\left|(\widehat{F} - \widehat{P}_F)(\cos(\theta)\xi) - \left(1 - \frac{m_2(F)}{2}\cos^2(\theta)\xi^2\right)\right| \leq R|\xi|^4$$

where $|R|$ has a bound depending only on C. In particular, it is independent of θ and ξ. Since

$$\int_{-\pi}^{\pi} \rho(\theta)\cos^k(\theta)\sin(\theta)\mathrm{d}\theta = 0$$

for all integers $k > 0$, as a consequence of our stipulation that ρ is even,

$$|(\widehat{F} - \widehat{P}_F) \circ \widehat{P}_G(\xi)| \leq \int_{-\pi}^{\pi} |R||P_G(\xi)|\mathrm{d}\theta \leq D\mu(G)|\xi|^4 \ ,$$

where D is a finite constant depending only on C. By symmetry in F and G, we also have, increasing D if need be,

$$|(\widehat{G} - \widehat{P}_G) \circ \widehat{P}_F(\xi)| \leq D\mu(F)|\xi|^4 \ .$$

Next note that for $|\xi| \leq L_1$,

$$\begin{aligned}
&\widehat{P}_F \circ \widehat{P}_G(\xi) = \\
&= \int_{-\pi}^{\pi} \rho(\theta) \left(m_1(F)\xi\cos(\theta) - \frac{m_3(F)}{6}\xi^3\cos^3(\theta) \right) \left(m_1(G)\xi\sin(\theta) - \frac{m_3(G)}{6}\xi^3\sin^3(\theta) \right) \mathrm{d}\theta \\
&= 0 \ .
\end{aligned}$$

Hence, there is a constant D depending only on C so that

$$\begin{aligned}
&\sup_{|\xi|\leq L_1} |\xi|^{-4} \left(|(\widehat{F} - \widehat{P}_F) \circ \widehat{P}_G(\xi)| + |(\widehat{G} - \widehat{P}_G) \circ \widehat{P}_F(\xi)| + |\widehat{P}_F \circ \widehat{P}_G(\xi)| \right) \\
&\leq D(\mu(F) + \mu(G)) \ .
\end{aligned} \tag{2.20}$$

For $|\xi| \geq L_1$, there is the trivial bound $|\xi|^{-4} \leq L_1^{-4}$, together with the bounds $\|\widehat{M}\|_\infty = 1$ and

$$\|(\widehat{G} - \widehat{P}_G)\|_\infty \leq \|\widehat{G}\|_\infty + \|\widehat{P_G}\|_\infty \leq 1 + \epsilon \ , \tag{2.21}$$

we obtain

$$\begin{aligned}
&\sup_{|\xi|\geq L_1} |\xi|^{-4} \left(|(\widehat{F} - \widehat{P}_F) \circ \widehat{P}_G(\xi)| + |(\widehat{G} - \widehat{P}_G) \circ \widehat{P}_F(\xi)| + |\widehat{P}_F \circ \widehat{P}_G(\xi)| \right) \\
&\leq L_1^{-4} \left[(1+\epsilon) \left(\|\widehat{P}_G\|_\infty + \|\widehat{P}_F\|_\infty \right) + \|\widehat{P}_F\|_\infty \|\widehat{P}_G\|_\infty \right] .
\end{aligned} \tag{2.22}$$

Now using the arithmetic geometric mean and then (2.10) and (2.17),

$$\begin{aligned}
\|\widehat{P}_F\|_\infty \|\widehat{P}_G\|_\infty &\leq \frac{1}{2} \left(\|\widehat{P}_F\|_\infty^2 + \|\widehat{P}_G\|_\infty^2 \right) \\
&\leq \frac{\epsilon}{2} \left(\|\widehat{P}_F\|_\infty + \|\widehat{P}_G\|_\infty \right) \\
&\leq \frac{\epsilon(L_2^2 + L_2^6)^{1/2}}{2} (\mu(F) + \mu(G))
\end{aligned}$$

Using this and (2.10) and (2.17) once more in (2.22), we obtain that for a constant D depending only on C,

$$\sup_{|\xi|\geq L_1} |\xi|^{-4}\left(|(\widehat{F}-\widehat{P}_F)\circ \widehat{P}_G(\xi)| + |(\widehat{G}-\widehat{P}_G)\circ P_F(\xi)| + |\widehat{P}_F\circ \widehat{P}_G(\xi)|\right)$$
$$\leq D(\mu(F)+\mu(G)) \ .$$

Taking the larger of the values of D from here or in (2.20), we have that

$$|||(F-P_F)\circ P_G + (G-P_G)\circ P_F + P_F\circ P_G||| \leq D(\mu(F)+\mu(G)) \ . \tag{2.23}$$

Next, we bound $|||(F-P_F)\circ(G-P_G)-M|||$.

$$\begin{aligned}
&(\widehat{F}-\widehat{P}_F)(\cos(\theta)\xi)(\widehat{G}-\widehat{P}_G)(\sin(\theta)\xi) - \widehat{M}(\xi)\\
&= (\widehat{F}-\widehat{P}_F)(\cos(\theta)\xi)(\widehat{G}-\widehat{P}_G)(\sin(\theta)\xi) - \widehat{M}(\cos(\theta)\xi)\widehat{M}(\sin(\theta)\xi)\\
&= \left[(\widehat{F}-\widehat{P}_F)(\cos(\theta)\xi) - \widehat{M}(\cos(\theta)\xi)\right](\widehat{G}-\widehat{P}_G)(\sin(\theta)\xi)\\
&+ \left[(\widehat{G}-\widehat{P}_G)(\sin(\theta)\xi) - \widehat{M}(\sin(\theta)\xi)\right]\widehat{M}(\cos(\theta)\xi)
\end{aligned}$$

Again using the bounds $\|\widehat{M}\|_\infty = 1$ and (2.21),

$$\begin{aligned}
&|\xi|^{-4}\left|(\widehat{F}-\widehat{P}_F)(\cos(\theta)\xi)(\widehat{G}-\widehat{P}_G)(\sin(\theta)\xi) - \widehat{M}(\xi)\right|\\
&\leq (1+\epsilon)\cos^4(\theta)\left[\frac{(\widehat{F}-\widehat{P}_F)(\cos(\theta)\xi) - \widehat{M}(\cos(\theta)\xi)}{\cos^4(\theta)\xi^4}\right]\\
&+ \sin^4(\theta)\left[\frac{(\widehat{G}-\widehat{P}_G)(\sin(\theta)\xi) - \widehat{M}(\sin(\theta)\xi)}{\sin^4(\theta)\xi^4}\right]\\
&\leq (1+\epsilon)\cos^4(\theta)|||F-P_F-M||| + \sin^4(\theta)|||G-P_G-M|||
\end{aligned}$$

Recall that

$$\int_{-\pi}^{\pi}\cos^4(\theta)\rho(\theta)\mathrm{d}\theta = \int_{-\pi}^{\pi}\sin^4(\theta)\rho(\theta)\mathrm{d}\theta$$

and since

$$\Lambda = \int_{-\pi}^{\pi}(\cos^4(\theta)+\sin^4(\theta))\rho(\theta)\mathrm{d}\theta - 1 \ ,$$

we have

$$|||(F-P_F)\circ(G-P_G)-M||| \leq (1+\epsilon)\frac{\Lambda+1}{2}\left(|||(F-P_F)-M||| + |||(G-P_G)-M|||\right) \ .$$

This together with (2.18) and (2.23) gives us

$$\|F\circ G - M\| \le \frac{(1+\epsilon)(\Lambda+1)}{2}\left(\|(F-P_F)-M\| + \|(G-P_G)-M\|\right) + D(\mu(F)+\mu(G)) \ .$$

Now define

$$K = \left(\frac{(1+\epsilon)(\Lambda+1)}{2}\right)^{-1} D\ .$$

Then, recalling that $P_{F\circ G}=0$, and that $\mu(F\circ G)=0$, we have

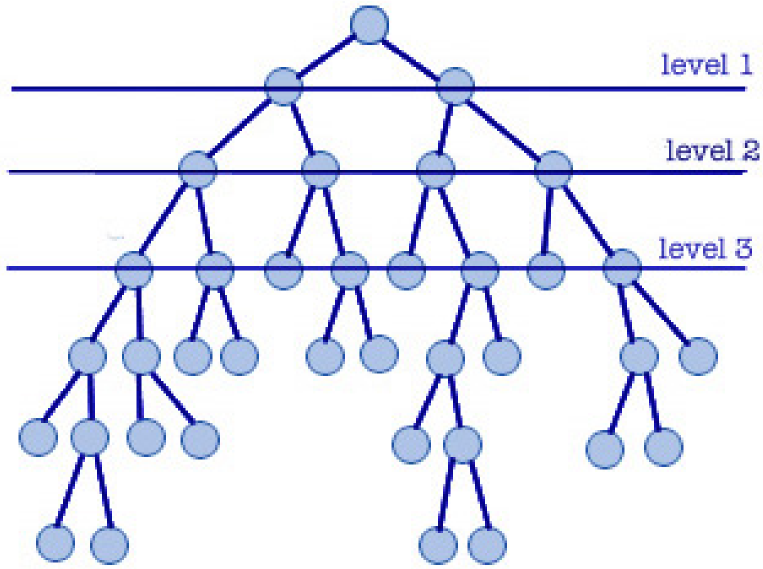
$$\Phi(F\circ G) \le \frac{(1+\epsilon)(\Lambda+1)}{2}\left(\Phi(F)+\Phi(G)\right)\ .$$

This gives us (2.1) with $c=(1+\epsilon)(\Lambda+1)$. As explained in the beginning of the section this gives us the bound stated in Theorem 2.1 after adjusting ϵ. ■

3. Smoothing Properties of the Kac equation

The main result in this section is a decomposition of $Q_n^+(F)$ into two pieces: a "beautiful" piece and an "ugly" piece. The beautiful piece will be very smooth, and the ugly piece will be very small. This decomposition is based in an essential way on the McKean walk representation of $Q_n^+(F)$ as a weighted sum of iterated Wild convolutions $C_\gamma(F)$; i.e., (1.14).

We shall show that if every leaf in a graph γ has a depth k, then $C_\gamma(F)$ has $k/2$ weak derivatives in L^2. How "beautiful" a graph is in the context of smoothness then depends on the minimal depth of the leaves. Here is an example of a graph in which every leaf is at depth of *at least* 3:



While most of the leaves are at a greater depth, it is the minimal depth that counts. There are graphs of every size in which there is one leaf whose depth is only 1. For such graphs, $C_\gamma(F)$ will have only the minimal smoothness inherited from F through one Wild convolution. Under the hypotheses below, this will be only one half of a weak derivative in L^2.

On the other hand, when n is large compared to k, there is a high probability that a graph γ chosen randomly from Γ_n according to the probability law $P(\gamma)$ in (1.14), has minimal depth k. The main result in this section then rests on two supports: We must determine precisely how this probability depends on k and n, and we must also show that there is an incremental improvement of one half of a weak derivative with each Wild convolution.

Theorem 3.1 *Let F be any probability density on $I\!R$ with finite Fisher information $I(F)$. Let c be any number with $0 < c < 1/2$. Then for any positive integer k and and any $n \geq 2^k$, $Q_n^+(F)$ can be decomposed as a convex combination of probability densities $B_{n,k}(F)$ and $U_{n,k}(F)$*

$$Q_n^+(F) = (1 - p_{n,k})B_{n,k}(F) + p_{n,k}U_{n,k}(F) \tag{3.1}$$

where for some constant C depending only on k and $I(F)$,

$$\|B_{n,k}(F)\|_{H^{k/2}(I\!R)} \leq C \ . \tag{3.2}$$

Moreover, $U_{n,k}(F)$ satisfies

$$\|U_{n,k}(F) - M\| \leq \Phi(F) \ , \tag{3.3}$$

and there is a finite number A depending only on c so that

$$p_{n,k} \leq \left(\frac{A}{(c/2)^{k-1}} \right) n^{-(1-2c)} \ . \tag{3.4}$$

Remark: In the opening paragraphs of this section, we discussed the probability that a graph γ chosen randomly from Γ_n according to the probability law $P(\gamma)$ in (1.14), has minimal depth k. We will see that this is $1 - p_{n,k}$. For fixed k, we can choose c arbitrarily close to zero. Then the constant multiplying $n^{-(1-2c)}$ is large but finite, and we see that for every $\epsilon > 0$,

$$1 - p_{n,k} = 1 - \mathcal{O}\left(\frac{1}{n^{1-\epsilon}} \right) \ .$$

This is the quantitative version of what we meant by a "high probability" of choosing a graph with minimal depth k from Γ_n.

Turning now to the lemmas needed to prove Theorem 3.1, we begin with an investigation of the incremental smoothness produced by one Wild convolution. The first is an analog of a result [2] of Bouchut and Desvillettes for the Boltzmann equation. In the case of the Kac equation, the mechanism in the proof is simpler and somewhat different.

In this section, we make use of the assumption that $\rho(\theta)$ is uniformly bounded by a finite constant B, which was part of our definition of a regular density ρ. For any square integrable function f on R, and any positive number s we define the Sobolev norm

$$\|f\|^2_{H^s(I\!R)} = \int_R |\hat{f}(\xi)|^2 |\xi|^{2s} \mathrm{d}\xi \ ,$$

where $\hat{f}$ denotes the Fourier transform of f.

Lemma 3.2 *For $\rho(\theta) \leq B$ for all θ,*

$$\|f \circ g\|^2_{H^s(I\!R)} \leq 2^{s+1/2} B \left[\|f\|^2_{H^{s-1/2}(I\!R)} \|g\|^2_{L^2(I\!R)} + \|g\|^2_{H^{s-1/2}(I\!R)} \|f\|^2_{L^2(I\!R)} \right] \ ,$$

for all f and g in $L^2(I\!R) \cap H^{s-1/2}(I\!R)$.

Proof: By Jensen's inequality and the Bobylev formula,

$$\|f \circ g\|^2_{H^s(I\!R)} le \int_{-\pi}^{\pi} \int_{I\!R} |\hat{f}(\cos(\theta)\xi)|^2 |\hat{g}(\sin(\theta)\xi)|^2 |\xi|^{2s} \mathrm{d}\xi \rho(\theta) \mathrm{d}\theta$$

$$\leq B \int_{-\pi}^{\pi} \int_{I\!R} |\hat{f}(\cos(\theta)\xi)|^2 |\hat{g}(\sin(\theta)\xi)|^2 |\xi|^{2s} \mathrm{d}\xi \mathrm{d}\theta \ .$$

We now make the change of variables $\eta = \cos(\theta)\xi$, which leads to

$$\|f \circ g\|^2_{H^s(I\!R)} \leq B \int_{I\!R} |\hat{f}(\eta)|^2 |\eta|^{2s} \left[\int_{-\pi}^{\pi} |\hat{g}(\tan(\theta)\eta)|^2 \sec^{2s+1}(\theta) \mathrm{d}\theta \right] \mathrm{d}\eta \ .$$

Now consider the inner integral, and make the change of variables $y = \tan(\theta)\eta$, where η is regarded as a fixed parameter for the time being. For any value of η, as θ varies between $-\pi$ and π, y covers the real line twice. Since

$$\sec^2(\theta)\mathrm{d}\theta = \frac{\mathrm{d}y}{\eta} \qquad \text{and} \qquad \sec(\theta) = \sqrt{1 + \left(\frac{y}{\eta}\right)^2} \ ,$$

$$\begin{aligned} \int_{-\pi}^{\pi} |\eta|^{2s} |\hat{g}(\tan(\theta)\eta)|^2 \sec^{2s+1}(\theta) \mathrm{d}\theta &= 2 \int_{I\!R} |\hat{g}(y)|^2 |\eta|^{2s-1} \left(1 + \left(\frac{y}{\eta}\right)^2\right)^{(2s-1)/2} \mathrm{d}y \\ &= 2 \int_{I\!R} |\hat{g}(y)|^2 | \left(|\eta|^2 + |y|^2\right)^{(2s-1)/2} \mathrm{d}y \ . \end{aligned}$$

Using the inequality $(a+b)^p \leq 2^p(a^p + b^p)$ for $a, b, p > 0$, we obtain

$$\|f \circ g\|^2_{H^{s/2}(I\!R)} \leq 2^{(2s+1)/2} B \left[\int_{I\!R} \int_{I\!R} |\hat{g}(y)|^2 |\hat{f}(\eta)|^2 \left(|\eta|^{2s-1} + |y|^{2s-1}\right) \mathrm{d}\eta \mathrm{d}y \right] \ .$$

■

The next estimate provides a uniform $L^2(I\!R)$ bound on all $C_\gamma(F)$. It relies on the fact, mentioned above, that for any McKean graph γ, $I(C_\gamma(F)) \leq I(F)$. The lemma translates this into an L^2 bound.

Lemma 3.3 *Let f be any probability density on $I\!R$ such that $I(f)$ is finite. Then*

$$\|f\|^2_{L^2(I\!R)} \leq 2(1 + I(f)) \ .$$

Proof: It is shown in Lemma 2.3 of [5] that $|\hat{f}(\xi)| \leq \dfrac{\sqrt{I(f)}}{|\xi|}$. Hence

$$\int_{|\xi|\geq 1} |\hat{f}(\xi)|^2 \mathrm{d}\xi \leq \int_{|\xi|\geq 1} \frac{I(f)}{|\xi|^2} \mathrm{d}\xi = 2I(f) \ .$$

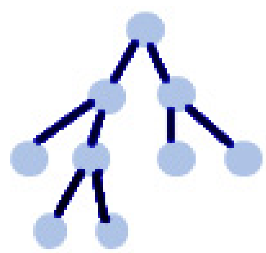
Also, since f is a probability measure $|\hat{f}(\xi)| \le 1$ so that $\int_{|\xi|\le 1} |\hat{f}(\xi)|^2 \mathrm{d}\xi \le 2$. Combining the estimates, we have the result. ■

Next we apply these lemmas to show that if every leaf in a McKean graph γ is of depth k or greater, then $C_\gamma(F)$ has $k/2$ weak derivatives. Toward this end, fix any k and consider any McKean graph $\gamma \in \Gamma_n$ with $n \ge 2^k$. We say that γ is *k–beautiful* if the depth of each of its leaves is at least k, and otherwise we say it is *k–ugly*. Let $\mathcal{B}_{n,k}$ be the subset of all of the k–beautiful McKean graphs in Γ_n, and let $\mathcal{U}_{n,k}$ be the subset of all of the k–ugly McKean graphs in Γ_n. Clearly,

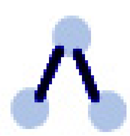
$$\mathcal{B}_{n,k} \cup \mathcal{U}_{n,k} = \Gamma_n \qquad \text{and} \qquad \mathcal{B}_{n,k} \cap \mathcal{U}_{n,k} = \emptyset \ .$$

Now if $\gamma \in \mathcal{B}_{n,k}$, then γ is completely filled in down to level k, and has some other, possibly empty, McKean graphs appended to the 2^k nodes at level k. Let γ_j be the McKean graph appended to the the jth node from left to right at the kth level.

For example, consider the graph $\gamma \in \Gamma_{19}$ in the diagram at the beginning of this section. Level 3 is the deepest filled level, and there are $8 = 2^3$ nodes left to right at level 3. In the diagram, you see



appended below the first node on the left. Hence this is γ_1. Likewise, γ_2 is the graph



since this is what you see appended below the second node at level 3.

Let $G_j = C_{\gamma_j}(F)$ with the understanding that $G_j = F$ if there is no additional graph appended at the jth node. In the example, $G_1 = (F \circ (F \circ F)) \circ (F \circ F)$, $G_2 = F \circ F$, and $G_3 = F$. The point of these definitions is that $C_\gamma(F)$ can be written as a Wild convolution of the G_j:

$$C_\gamma(F) = ((G_1 \circ G_2) \circ (G_3 \circ G_4)) \circ ((G_5 \circ G_6) \circ (G_7 \circ G_8)) \ .$$

By Lemma 3.3, we have an *a–priori* bound on $\|G_j\|_{L^2(I\!R)}$, say $\|G_j\|^2_{L^2(I\!R)} \le C$, uniformly in j.

By Lemma 3.2, we then have

$$\|G_1 \circ G_2\|^2_{H^{1/2}(I\!R)} \le 2B(2C^2) = 2^2BC^2 \ .$$

Of course, we still have $\|G_1 \circ G_2\|^2_{L^2(I\!R)} \le C$ since $G_1 \circ G_2 = C_\delta(F)$ for some McKean graph δ.

The same estimates apply to $G_3 \circ G_4$. Therefore, apply Lemma 3.2 again,

$$\|((G_1 \circ G_2) \circ (G_3 \circ G_4))\|_{H^1(I\!R)} \le 2^{3/2}B2(2^2BC^2C) = 2^{9/2}B^2C^3 \ .$$

Continuing, we finally get

$$\|C_\gamma(F)\|_{H^{3/2}(I\!R)} \le 2^2 2B(2^{9/2}B^2C^4) = 2^{15/2}B^3C^4 \ .$$

This analysis can be extended easily to higher values of k, and we have:

Lemma 3.4 *Let F be any probability density on $I\!R$ such that $I(f)$ is finite. Let k be any fixed positive integer, and n any integer with $n \ge 2^k$. Then there is a constant C depending only on k and $I(F)$ so that*

$$\|C_\gamma(F)\|_{H^{k/2}(I\!R)} \le C \ . \tag{3.5}$$

We next show that when n is large compared to 2^k, the likelihood of "drawing" a beautiful graph in Γ_n is very high. A method developed in [4] is perfectly suited to this task.

Let the function W on Γ_n be defined by

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

where c is some number with $0 < c < 1$, and $d(j)$ is the depth of the jth leaf. Let $P(\gamma)$ be the probability that the McKean walk passes through γ at the nth step. Then, as shown in Lemma 1.4 of [4], if p is any number with $p < 1 - c$, there is a constant A so that

$$\sum_{\gamma \in \Gamma_n} W(\gamma)P(\gamma) \le An^{-p} \ .$$

Now if $\gamma \in \mathcal{U}_{n,k}$, the $W(\gamma) > (c/2)^{k-1}$ since there is at least one leaf of depth no greater than $k - 1$. Therefore,

$$\sum_{\gamma\in\mathcal{U}_{n,k}} P(\gamma) \le \frac{1}{(c/2)^{k-1}} \sum_{\gamma\in\mathcal{U}_{n,k}} W(\gamma)P(\gamma)$$
$$\le \frac{1}{(c/2)^{k-1}} \sum_{\gamma\in\Gamma_n} W(\gamma)P(\gamma)$$
$$\le \frac{An^{-p}}{(c/2)^{k-1}} \ .$$

We now define numbers $p_{n,k}$ by $p_{n,k} = \sum_{\gamma\in\mathcal{U}_{n,k}} P(\gamma)$. Clearly, $p_{n,k}$ is the probability that the McKean walk passes through a k–ugly graph at the nth step. Letting c be any number with $0 < c < 1/2$, the estimate we have just derived gives us the bound

$$p_{n,k} \le \left(\frac{A}{(c/2)^{k-1}}\right) n^{-(1-2c)} \tag{3.6}$$

where A is independent of k and n.

Next define two probability densities $B_{n,k}(F)$ and $U_{n,k}(F)$ by

$$B_{n,k}(F) = \frac{1}{1-p_{n,k}} \sum_{\gamma\in\mathcal{B}_{n,k}} P(\gamma)C_\gamma \qquad \text{and} \qquad U_{n,k}(F) = \frac{1}{p_{n,k}} \sum_{\gamma\in\mathcal{U}_{n,k}} P(\gamma)C_\gamma \ .$$

Since $\Phi(C_\gamma(F) \le \Phi(F)$ for all γ, we have that

$$\Phi(U_{n,k}(F)) = \Phi\left(\frac{1}{p_{n,k}} \sum_{\gamma\in\mathcal{U}_{n,k}} P(\gamma)C_\gamma\right) \le \Phi(F) \ . \tag{3.7}$$

And since *all* of the odd moments of every $C_\gamma(F)$ vanish for $\gamma \in \Gamma_n$, $n > 1$, $\mu(U_{n,k}(F)) = 0$, so that

$$\Phi(U_{n,k}(F)) = \|U_{n,k}(F) - M\| \ . \tag{3.8}$$

We finally arrive at the proof of the main theorem:

Proof of Theorem 3.1: The decomposition (3.1) holds by the definitions of the quantities on the right. The fact that $B_{n,k}(F)$ satisfies (3.2) follows from (3.5) and the convexity of the norm. The fact that $U_{n,k}(F)$ satisfies (3.3) follows from (3.7) and (3.8). Finally, (3.4) has been established in (3.6). ■

4. Interpolation bounds

Here we recall some interpolation inequalities that allow us to pass from the $\|\!| \cdot |\!\|$ norm to the $\| \cdot \|_{L^1(I\!R)}$ norm.

Lemma 4.1 *Let $0 < r < 1$ be given. Then there is a constant C depending only on r so that*

$$\|f\|^2_{L^2(I\!R)} \leq C \|\!| f |\!\|^{2(1-r)} \left(\|f\|^2_{\mathrm{H}^M} + \|f\|^2_{\mathrm{H}^{M+r/2}} \right)^r$$

with $M = 4(1-r)/r$.

Proof: For any r with $0 < r < 1$,

$$\begin{aligned}
\|f\|^2_{L^2(I\!R)} &= \int_{I\!R} |\hat{f}(\xi)|^2 \mathrm{d}\xi \\
&= \int_{I\!R} \left(\frac{|\hat{f}(\xi)|}{|\xi|^4} \right)^{2(1-r)} |\hat{f}(\xi)|^{2r} |\xi|^{8(1-r)} (1+|\xi|^r)^{1/r} (1+|\xi|^r)^{-1/r} \mathrm{d}\xi \\
&\leq \|\!| f |\!\|^{2(1-r)} \int_{I\!R} |\hat{f}(\xi)|^{2r} |\xi|^{8(1-r)} (1+|\xi|^r)^{1/r} (1+|\xi|^r)^{-1/r} \mathrm{d}\xi \\
&\leq \|\!| f |\!\|^{2(1-r)} \left(\int_{I\!R} |\hat{f}(\xi)|^2 |\xi|^{8(1-r)/r} (1+|\xi|^r) \mathrm{d}\xi \right)^r \times \\
&\left(\int_{I\!R} (1+|\xi|^r)^{-1/r(1-r)} \mathrm{d}\xi \right)^{1-r} ,
\end{aligned}$$

where in the last inequality we used Hölder's inequality with exponents $1/r$ and $1/(1-r)$.

Clearly, $\int_{I\!R} (1+|\xi|^r)^{-1/r(1-r)} \mathrm{d}\xi = C < \infty$, and

$$\int_{I\!R} |\hat{f}(\xi)|^2 |\xi|^{8(1-r)/r} (1+|\xi|^r) \mathrm{d}\xi \leq \|f\|^2_{\mathrm{H}^M} + \|f\|^2_{\mathrm{H}^{M+r/2}} .$$

■

The next inequality shows that control of sufficiently many moments and control on the L^2 norm together control the L^1 norm.

Lemma 4.2 *Let f be an integrable function on $I\!R$. Then for all $k > 0$, there is a constant C depending only on k so that*

$$\int_{I\!R} |f(v)| \mathrm{d}v \leq C \left(\int_{I\!R} |f(v)|^2 \mathrm{d}v \right)^{(8k+1)/(8k+2)} (m_{2k}(f))^{1/(8k+2)} .$$

Proof: We may assume that f is non–negative.

Let $L > 0$ be chosen. Then

$$\begin{aligned}\int_{I\!R} f(v)\mathrm{d}v &= \int_{|v|\le L} f(v)\mathrm{d}v + \int_{|v|\ge L} f(v)\mathrm{d}v \le \\ &\le (2L)^{1/2}\|f\|_{L^2(I\!R)} + L^{-2k}\int_{I\!R}|v|^{2k}f(v)\mathrm{d}v \\ &= (\sqrt{2}\|f\|_{L^2(I\!R)})L^{1/2} + m_{2k}(F)L^{-2k}\ .\end{aligned}$$

Choosing $L = 2k(m_{2k}(f)/\|f\|_{L^2(I\!R)})^{2/(4k+1)}$ now yields the result. ■

5. Proof of the main theorem

Proof of Theorem 1.1 Consider the decomposition

$$Q_n^+(F) = (1-p_{n,k})B_{n,k}(F) + p_{n,k}U_{n,k}(F) \tag{5.1}$$

given by Theorem 3.1 for any k and any $n \ge 2^k$. Then by the Minkowski inequality,

$$\begin{aligned}\|Q_n^+(F) - M\|_{L^1(I\!R)} \le (1-p_{n,k})&\|B_{n,k}(F) - M\|_{L^1(I\!R)} \\ &+ p_{n,k}\|U_{n,k}(F) - M\|_{L^1(I\!R)}\ .\end{aligned} \tag{5.2}$$

Since $\|U_{n,k}(F) - M\|_{L^1(I\!R)} \le 2$, (3.4) of Theorem 3.1 gives us the bound

$$p_{n,k}\|U_{n,k}(F) - M\|_{L^1(I\!R)} \le \left(\frac{2A}{(c/2)^{k-1}}\right)n^{-(1-2c)}\ ,$$

where c is any number with $0 < c < 1/2$, and A depends only on c. Since $\Lambda > -1$, we are free to choose c so that

$$-(1-2c) \le \Lambda\ . \tag{5.3}$$

Doing so, and combining results, we see that there is a constant C so that

$$p_{n,k}\|U_{n,k}(F) - M\|_{L^1(I\!R)} \le Cn^{\Lambda} \tag{5.4}$$

Estimating $\|B_{n,k}(F)-M\|_{L^1(I\!R)}$ is a bit more work. We first note that by the Minkowski inequality once more,

$$(1-p_{n,k})|\!|\!|B_{n,k}(F)-M|\!|\!| \le |\!|\!|Q_n^+(F)|\!|\!| + p_{n,k}|\!|\!|U_{n,k}(F)-M|\!|\!| \ .$$

By Theorem 2.2, we have that for any $\tilde{\epsilon}>0$, there is a finite constant $A_{\tilde{\epsilon}}$ so that

$$|\!|\!|Q_n^+(F)-M|\!|\!| \le A_{\tilde{\epsilon}} n^{\Lambda+2\tilde{\epsilon}} \ .$$

By Theorem 3.1, we have $|\!|\!|U_{n,k}(F)-M|\!|\!| \le \Phi(F)$ and $p_{n,k} \le \frac{An^{-(1-2c)}}{(c/2)^{k-1}}$ where A depends only on c. Again making the choice of c in (5.3), we see that there is a constant C so that

$$(1-p_{n,k})|\!|\!|B_{n,k}(F)-M|\!|\!| \le Cn^{\Lambda+2\tilde{\epsilon}} \ . \tag{5.5}$$

Since M is smooth and has moments of every order, $B_{n,k}(F)-M$ is smooth and has moments of every order. In particular, fix any $r>0$, and suppose that $k>16(1-r)/r+r$. By Theorem 3.1 we have a bound on $\|B_{n,k}(F)-M\|_{\mathrm{H}^k}$ that is uniform in n. Hence by Lemma 4.1, there is a constant C so that

$$\|B_{n,k}(F)-M\|_{L^2(I\!R)} \le C\left(|\!|\!|B_{n,k}(F)-M|\!|\!|\right)^{1-r} \ . \tag{5.6}$$

uniformly in n.

Next, since for any ℓ, the 2ℓth moment $m_{2\ell}(B_{n,k}(F)-M)$ is bounded uniformly in n, we can choose ℓ large enough that $1/(8\ell+2)<r$. Then Lemma 4.2 tell us that there is a finite constant C so that for all n,

$$\|B_{n,k}(F)-M\|_{L^1(I\!R)} \le C\left(\|B_{n,k}(F)-M\|_{L^2(I\!R)}\right)^{1-r} \ . \tag{5.7}$$

Combining (5.6) and (5.7), we have that there is a constant C so that for all n,

$$\|B_{n,k}(F)-M\|_{L^1(I\!R)} \le C\left(|\!|\!|B_{n,k}(F)-M|\!|\!|\right)^{(1-r)^2} \ . \tag{5.8}$$

Now combining (5.5) and (5.8), we obtain that there is a constant C so that for all n,

$$(1-p_{n,k})\|B_{n,k}(F)-M\|_{L^1(I\!R)} \le C\left(Cn^{\Lambda+2\tilde{\epsilon}}\right)^{(1-r)^2} \ . \tag{5.9}$$

Combining (5.2) with (5.4) and (5.9) yields

$$\|Q_n^+(F)-M\|_{L^1(I\!R)} \le C\left(Cn^{\Lambda+2\tilde{\epsilon}} + \left(Cn^{\Lambda+2\tilde{\epsilon}}\right)^{(1-r)^2}\right) \ .$$

No matter how small $\epsilon > 0$, we can choose $\tilde{e} > 0$ and $r > 0$ sufficiently small that for another constant C,

$$\|Q_n^+(F) - M\|_{L^1(I\!R)} \le Cn^{\Lambda+\epsilon}$$

for all n. ■

Bibliography

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

[2] Bouchut, F. and Desvillettes, L.: Averaging lemmas without time Fourier transform and application to discretized kinetic equations. Proc. Roy. Soc. Edinburgh Sect. A 129 (1999), no. 1, pp. 19–36.

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

[4] Carlen, E. A., Carvalho, M. C. and Gabetta, E.: *Central Limit Theorem for Maxwellian Molecules and Truncation of the Wild Expansion*, Comm Pure and Appl. Math., Vol.**LIII** (2000) pp. 370–397.

[5] Carlen, E. A., Gabetta, E. and Toscani, G.: Propagation of smoothness and the sharp exponenetial rate of relaxation to equilibrium for Maxwellian molecules, *Comm. Math. Phys.* **199** no. 3, (1999) pp. 521–546.

[6] Carlen, E. A. and Lu, Xuguang.: *Fast and slow convergence to equilibrium for Maxwellian molecules via Wild sums*, Joour. Stat. Phys **112** 1/2, (2003) pp. 59–134

[7] Carlen, E. A., Salvarani, F.: On the optimal choice of coefficients in a truncated Wild sum and approximate solutions for the Kac equation, *Jour. Stat. Phys.* **109** no. 1/2, (2002) pp. 261–277

[8] Desvillettes, L.: About the regularizing properties of the non–cut–off Kac equation, *Comm. Math. Phys.*, **168**, (1995) pp. 417–440

[9] Desvillettes, L. and Villani, C.: On the spatially homogenous Landau equation for hard potentials. Part I: Existence, uniqueness and smoothness, *Comm. P. D. E.*, **25**, (1995) pp. 179–259

[10] Elmroth, T.: Global boundedness of moments of solutions of the Boltzmann equation for forces of infinite range, *Arch. Rat. Mech. Analysis*, **82**, (1983) pp. 1-12

[11] Gabetta, E., Pareschi, L. and Toscani, G: Relaxation schemes for nonlinear kinetic equations, *SIAM J. Num. Analysis.* **34** (1997) pp. 2168–2194

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

[13] Goudon, T: On Boltzmann equations and Fokker–Planck assymptotics: influence of grazing collisions. *J. Stat. Phys.* **80** (1997) pp 751–776

[14] Inkenberry, E and Truesdell C.: On the pressure and the flux of energy according to Maxwell's kinetic energy I, *J. Rat. Mech. Anal.*, **5** (1956) pp 1–54

[15] Kac, M., *Foundations of kinetic theory*, Proc. 3rd Berkeley symp. Math. Stat. Prob., J. Neyman, ed. Univ. of California, vol 3, pp. 171–197, (1956).

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

[17] McKean, H. P. Jr.: An exponential formula for solving Boltzmann's equation for a Maxwellian gas, *J. Comb. Theory*, **2**, (1967) pp. 358–382.

[18] Villani, C.: On a new class of weak solutions for the spatially homogeneous Boltzmann and Landau equations, *Arch. Rat. Mech. Anal.* **143** (1998) pp. 273–307

[19] Wild, I.: On Boltzmann's equation in the kinetic theory of gasses, *Proc. Camb. Phil. Soc.* **47** (1951) pp. 602–609