Cumulants, their evolution hierarchies, and how to use them to control chaos for kinetic theory

Draft lecture notes – Toulouse 2026

Jani Lukkarinen

e-mail: jani.lukkarinen@helsinki.fi

Last updated: June 15, 2026

Preface

These lecture notes are supporting the minicourse for the *Summer School EUR MINT 2026 on Particle systems and PDEs*, given at the University of Toulouse in June 2026.

Notations and structure of the material

The following **mathematical notations** will be used throughout the text:

$[N]$	If $N \in \mathbb{N}_+$, we use this notation for the sequence of first N integers. Explicitly, $[N] := \{1, 2, \ldots, N\}$
$[a, b]$	Closed interval from a to b, i.e., the collection of real numbers x for which $a \leqslant x \leqslant b$
$]a, b[$	Open interval from a to b, i.e., the collection of real numbers x for which $a < x < b$
$\mathbb{R}$	The collection of real numbers, denoted also $\mathbb{R} =]-\infty, \infty[$
$[a, b[$, $]-\infty, b]$, etc.	Obvious variants of the above
$\mathbb{R}^d$	Euclidean d-dimensional space. $x \in \mathbb{R}^d$ can also be denoted by $(x_1, x_2, \ldots, x_d)$ where $x_i \in \mathbb{R}$ is the i:th component of x. In particular, in the physical space $d = 3$, we identify its vectors by using a boldface notation: $\mathbf{x} \in \mathbb{R}^3$ corresponds to $\mathbf{x} = (x_1, x_2, x_3)$
$\mathbb{R}_+$	Non-negative real numbers $[0, \infty[$
$\mathbb{N}$	Natural numbers $\{1, 2, \ldots\}$
$\mathbb{N}_0$	Natural numbers including zero, i.e., $\mathbb{N}_0 := \{0, 1, 2, \ldots\}$
$\mathbb{Z}$	Integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
i	Imaginary unit
$\mathbb{C}$	Set of complex numbers $\{x + \mathrm{i}y \mid x, y \in \mathbb{R}\}$
$\varnothing$	Empty set
$A \cup B$	Union of A and B, defined as $\{x \mid x \in A \text{ or } x \in B\}$
$A \cap B$	Intersection of A and B, defined as $\{x \mid x \in A \text{ and } x \in B\}$
$A \backslash B$	Complement of B relative to A or "A minus B", defined as the set $\{x \mid x \in A \text{ and } x \notin B\}$
$A \subset B$	A is a subset of B, i.e., $x \in A \Rightarrow x \in B$
$\mathbb{1}_{\{P\}}$	Generic characteristic function of the condition P. Defined by setting $\mathbb{1}_{\{P\}} = 1$ if P is true, and $\mathbb{1}_{\{P\}} = 0$ if P is false. For instance, $\mathbb{1}_{\{0 \leqslant x \leqslant 1\}}$ as a function of the real variable x, defines a map $\mathbb{R} \to \mathbb{R}$ which is one on the interval $[0, 1]$ and zero outside of the interval, i.e., the characteristic function of the closed interval $[0, 1]$
$\mathbb{P}_\mu(A)$	Probability of the event A under the probability measure μ. If the measure is obvious from the context, its name will be dropped from the notation
$\mathbb{E}_\mu[F]$	Expectation value of a function $F : \Omega \to \mathbb{C}$ under a probability measure μ on the set Ω. If the measure is obvious from the context, its name will be dropped from the notation

Contents

Chapter 1

Introduction

(See the PDF slides for the introduction, more motivation can be found from the Introduction of [1].)

1.1 Tools in probability

In this section, we consider analysis of properties of some given probability measure μ on the space Ω.[1] The structure of probability measures arising for instance from ergodic limits is usually intractably complex to study in all detail. It is nevertheless often possible to control the salient features of time-evolution of such measures by suitable choice of random variables and related observables. Here we present some tools which will be needed later for examples in kinetic theory.

A real-valued *random variable* X is a measurable function from Ω to $\mathbb{R}$ and a complex-valued random variable is a measurable function from Ω to $\mathbb{C}$. An *observable* of a collection $\mathbf{X} = (X_1, \ldots, X_n)$ of random variables is a function $f : \mathbb{R}^n \to \mathbb{C}$ such that the function[2] $f(\mathbf{X}) : \Omega \to \mathbb{C}$ is measurable. The observable is *integrable* if $\mathbb{E}[|f(\mathbf{X})|] < \infty$ and in this case $\mathbb{E}[f(\mathbf{X})] \in \mathbb{C}$ is always well-defined.

For instance, in the cases considered here, all continuous functions f are observables, as are all pointwise limits of a sequence of continuous functions. If f is also bounded, it is even an integrable observable.

1.1.1 One real-valued random variable

Let us first consider analysis of just one real random variable X. Often the distribution of X is determined by a probability density function $\Phi : \mathbb{R} \to \mathbb{R}_+$ and the expectation value of an integrable observable $f : \mathbb{R} \to \mathbb{C}$ is given by the integral

$$\mathbb{E}_\mu[f(X)] = \int_{-\infty}^{\infty} \mathrm{d}x\, \Phi(x) f(x)\,. \tag{1.1.1}$$

The *characteristic function* of X is defined via the expectation values

$$\varphi_X(s) := \mathbb{E}_\mu[\mathrm{e}^{\mathrm{i}sX}]\,, \qquad s \in \mathbb{R}\,. \tag{1.1.2}$$

[1]Unless mentioned otherwise, the measures μ we consider here are either defined on a finite Euclidean space $\Omega = \mathbb{R}^d$ for some $d < \infty$ and have the same σ-algebra as the Lebesgue measure on $\mathbb{R}^d$, or the space Ω has a locally compact Hausdorff topology and μ is a regular Borel measure on Ω, for instance, obtained by application of the Riesz–Markov–Kakutani representation theorem [2]. In both cases, all continuous functions on Ω are measurable w.r.t. μ, as well as any function $f : \Omega \to \mathbb{C}$ which is a pointwise convergent limit of a sequence of continuous functions on Ω. In particular, if $A \subset \Omega$ is open, closed, or a countable union of such, its characteristic function $f(x) := \mathbb{1}_{\{x \in A\}}$ is measurable w.r.t. μ and its integral is equal to the probability $\mathbb{P}(x \in A)$.

[2]The function "$f(X)$" here denotes the composite map $f \circ X$.

Since $f(X) = \mathrm{e}^{\mathrm{i}sX}$ is a bounded continuous function of X, it is an integrable observable, and the formula always defines a map $\varphi_X : \mathbb{R} \to \mathbb{C}$. The full distribution of X can always be recovered from φ_X. For instance, for any sufficiently regular function[3]

$$\mathbb{E}[f(X)] = \int_{-\infty}^{\infty} \mathrm{d}p \, \varphi_X(2\pi p) \widehat{f}(p) \,, \tag{1.1.3}$$

where $\widehat{f}(p) = \int \mathrm{d}x \, \mathrm{e}^{-\mathrm{i}2\pi p x} f(x)$ denotes the Fourier transform of f.

A related concept is the *moment generating function* G_{mom} of X which is defined by

$$G_{\text{mom}}(\eta) := \mathbb{E}[\mathrm{e}^{\eta X}] \,, \tag{1.1.4}$$

for those $\eta \in \mathbb{R}$ for which the observable $\mathrm{e}^{\eta X}$ is integrable. Unlike the characteristic function, the moment generating function is not necessarily well defined for all η; in fact, it can happen that it is only defined at $\eta = 0$.

One case in which the moment generating function becomes a convenient tool, is when the random variable X has *exponential moments*. This terminology is not entirely standardized, and here we say that this occurs whenever one can find some $\varepsilon > 0$ for which

$$\mathbb{E}[\mathrm{e}^{\varepsilon |X|}] < \infty \,.$$

To check the finiteness of the above integral, it is sometimes easier to check that $\mathbb{E}[\mathrm{e}^{\varepsilon X}], \mathbb{E}[\mathrm{e}^{-\varepsilon X}] < \infty$. Namely, if this is the case, then also $\mathbb{E}[\mathrm{e}^{\varepsilon |X|}] < \infty$, since $0 \leqslant \mathrm{e}^{\varepsilon |X|} \leqslant \mathrm{e}^{\varepsilon X} + \mathrm{e}^{-\varepsilon X}$ for all $X \in \mathbb{R}$.

If the random variable X has exponential moments, the following simplifying properties hold:

1. X has moments of all orders. In other words, each of the observables X^n, $n \in \mathbb{N}$, is integrable.
2. The characteristic function and the moment generating function have an analytic continuation near the origin, at least into the ball of radius ε in the complex plane. In addition, for these analyticly continued functions we have the relation
$$G(\eta) = \varphi_X(-\mathrm{i}\eta) = \mathbb{E}[\mathrm{e}^{\eta X}] \,, \qquad |\eta| < \varepsilon \,.$$
3. The n:th moment of X may be obtained by differentiation from both the characteristic function and the moment generating function. Namely,
$$\left.\frac{\mathrm{d}^n}{\mathrm{d}\eta^n} G_{\text{mom}}(\eta)\right|_{\eta=0} = \mathbb{E}[X^n] = (-\mathrm{i})^n \left.\frac{\mathrm{d}^n}{\mathrm{d}s^n} \varphi_X(s)\right|_{s=0} .$$
4. Also the *cumulant generating function* defined by
$$g_{\mathrm{c}}(\eta) := \ln G_{\text{mom}}(\eta) \,, \tag{1.1.5}$$
is then analytic in some neighbourhood of the origin, and the *n:th cumulant of* X can be obtained as the value of its derivative at the origin, namely,
$$\kappa_n[X] = \left.\frac{\mathrm{d}^n}{\mathrm{d}\eta^n} g_{\mathrm{c}}(\eta)\right|_{\eta=0} . \tag{1.1.6}$$

Cumulants may be defined also without reference to the cumulant generating function in (1.1.5). We discuss the definition in detail later, in Section 1.1.2. By the results of that Section, in the case of just one random variable, it suffices that the moment X^N is integrable for some N, and then one can define all cumulants $\kappa_n[X]$ up to order $n \leqslant N$ by using the recursion formula

$$\kappa_n[X] = \mathbb{E}[X^n] - \sum_{j=1}^{n-1} \binom{n-1}{j-1} \kappa_j[X] \, \mathbb{E}[X^{n-j}] \,. \tag{1.1.7}$$

[3] For "sufficiently regular" it is enough that the inverse formula of Fourier transform is absolutely integrable and works at every point. Hence, f can be any Schwartz test function or a smooth compactly supported function.

In other words, one begins with $\kappa_1[X] = \mathbb{E}[X]$, and then uses (1.1.7) to define κ_n once all κ_j, $j < n$, have been fixed. Note that since X^N is integrable and $|X|^n \leqslant 1 + |X|^N$ for $n \leqslant N$, also each X^n, $1 \leqslant n \leqslant N$, is integrable.

Example 1.1.1 (Gaussian random variable) If X has a probability density function

$$\Phi(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2\sigma^2}(x-x_0)^2}\,, \qquad x \in \mathbb{R}\,,$$

where $x_0 \in \mathbb{R}$ and $\sigma > 0$ are given, then X is a Gaussian random variable, or a normally distributed random variable, with mean x_0 and standard deviation σ. It has exponential moments, since by completing the square one can explicitly compute that

$$\int_{-\infty}^{\infty} \mathrm{d}x\, \Phi(x) \mathrm{e}^{rx} = \mathrm{e}^{x_0 r + \frac{1}{2}\sigma^2 r^2}\,, \qquad r \in \mathbb{R}\,,$$

and hence $\mathbb{E}[\mathrm{e}^{rX}] < \infty$ for any $r \in \mathbb{R}$. Therefore,

$$G_{\mathrm{mom}}(\eta) = \mathrm{e}^{x_0\eta + \frac{1}{2}\sigma^2\eta^2}\,, \quad \varphi_X(s) = \mathrm{e}^{\mathrm{i}x_0 s - \frac{1}{2}\sigma^2 s^2}\,, \quad g_{\mathrm{c}}(\eta) = x_0\eta + \frac{1}{2}\sigma^2\eta^2\,,$$

which yields the cumulants

$$\kappa_1[X] = x_0\,, \qquad \kappa_2[X] = \sigma^2\,, \qquad \kappa_n[X] = 0\,,\ n \geqslant 3\,.$$

Also the moments $\mathbb{E}[X^n]$ could be computed by differentiation, either by recognizing a connection to Hermite polynomials or by iteration. Alternatively, one could use the moment-to-cumulants formula of Section 1.1.2 and the easy expressions for cumulants. The outcome of these computations can be found for instance in Wikipedia, but it is not central to our discussion, so we do not reproduce it here.

1.1.2 Joint properties of finitely many real random variables

Let us next consider the case of a finite collection of random variables X_j, $j \in J$, where J is some fixed nonempty index set. For their joint probabilistic properties we use here notations which were found convenient in [3]. We use sequences of indices, $I = (i_1, i_2, \ldots, i_n) \in J^n$, to label monomials of the above random variables, with the following shorthand notation

$$X^I := X_{i_1} X_{i_2} \cdots X_{i_n} = \prod_{k=1}^{n} X_{i_k} =: \prod_{j \in I} X_j\,. \tag{1.1.8}$$

We also set $X^{\varnothing} := 1$ if I is the empty sequence. Since all X_j commute with each other, we have $X^I = X^{I'}$ for any two sequences I, I' which differ by a permutation. The results in the previous subsection can be reproduced by choosing $J = \{1\}$ and $X_1 = X$.

We say that the collection has *finite moments up to order n*, if the monomials X^I are integrable for all $I \in J^m$ with $m \leqslant n$. Similarly, we say the collection has *joint exponential moments* if there is $\varepsilon > 0$ such that $\exp(\varepsilon \sum_{j \in J} |X_j|)$ is integrable. The validity of this condition is sometimes easier to check by showing that $\exp(\varepsilon \sum_{j \in J} \sigma_j X_j)$ is integrable for all sign-combinations $\sigma_j = \pm 1$, $j \in J$.

The *joint characteristic function* of the collection is defined as

$$\varphi(\mathbf{s}) := \mathbb{E}_\mu[\mathrm{e}^{\mathrm{i}\mathbf{s}\cdot\mathbf{X}}]\,, \qquad \mathbf{s} \in \mathbb{R}^J\,, \tag{1.1.9}$$

where $\mathbf{s}\cdot\mathbf{X} := \sum_{j \in J} s_j X_j$ denotes the scalar product in $\mathbb{R}^J$. Similarly, the *joint moment generating function* is defined by

$$G_{\mathrm{mom}}(\eta) := \mathbb{E}_\mu[\mathrm{e}^{\sum_j \eta_j X_j}]\,, \tag{1.1.10}$$

for those $\eta \in \mathbb{R}^J$ for which the integral converges. If the collection has joint exponential moments, again the joint moment generating function is everywhere defined and all finite polynomials of X_j, $j \in J$, are integrable. Then all moments are finite and may be computed from the generating functions analogously to the one-variable case: for any $I \in J^n$ with $n \in \mathbb{N}$, we have

$$\mathbb{E}[X^I] = \partial_\eta^I G_{\text{mom}}(0) = (-\mathrm{i}\partial)_s^I \varphi(0)\,, \tag{1.1.11}$$

where "∂_η^I" is a shorthand notation for the partial derivative $\partial_{\eta_{i_1}} \partial_{\eta_{i_2}} \cdots \partial_{\eta_{i_n}}$.

To deal with cumulants, we need to operate not only with such sequences but also with their subsequences and "partitions". This will be done by choosing a distinct label for each member of the sequence and collecting these into a set. How the labelling is done is not important, as long as one takes care when combining two "labelled" sets. For instance, it is safe to use the position index in the vector I as a label, i.e., once $I \in J^n$ is fixed, it can be replaced by the set $[n] := \{1, 2, \ldots, n\}$, with the tacit understanding that $i \in [n]$, refers to the random variable X_{I_i}, $I_i \in J$. Since elements may be repeated in the sequence I, it is possible that $I_i = I_{i'}$ even if $i' \neq i$.

For the sequence $I \in J^n$, a collection π of the subsets of $[n]$ is a *partition* if each $A \in \pi$ is non-empty, and every element in $[n]$ is included in exactly one of the sets $A \in \pi$. We denote the *collection of all partitions* of a sequence I by $\mathcal{P}(I)$. We do not make any difference in the notation between $A \subset [n]$ and the corresponding subsequence $A \subset I$. In particular, if $\pi \in \mathcal{P}(I)$ and $A \in \pi$, then $X^A := \prod_{i \in A} X_{I_i}$ can also be written as $X^A := \prod_{j \in A} X_j$. Similarly, we use the shorthand notation

$$X_A := (X_j)_{j \in A}\,,$$

for the vector of random variables which is obtained by removing from the original sequence those random variables X_i for which $i \notin A$. This notation will mainly be used for cumulants. Let us record below various alternative notations for cumulants found in the literature,

$$\kappa[X_I] = \kappa_\mu[X_I] = \mathbb{E}[X_{i_1}; X_{i_2}; \cdots; X_{i_n}] = \kappa(X_{i_1}, X_{i_2}, \cdots, X_{i_n}) = \mathbb{E}[X^I]^{\text{trunc}}\,, \tag{1.1.12}$$

where the last one refers to "truncated moments". Whenever set operations (union, intersection, complement, etc.) are used on subsets of I, we refer to those operations performed on their identifications in the labelling $[n]$.

To avoid separate treatment of expressions involving empty sets and conditions, we employ here the following standard conventions: if the condition P is false, we define

$$\sum_P (\cdots) := 0, \quad \prod_P (\cdots) := 1, \quad \text{and set also} \quad \mathcal{P}(\varnothing) := \{\varnothing\}\,. \tag{1.1.13}$$

We define cumulants here using a recursion relation which avoids the assumption of integrability of exponential moments. Let us assume instead that the collection has finite moments up to order n. Then for any sequence $I \in J^m$ of length $m \leqslant n$, we require that

$$\mathbb{E}[X^I] = \sum_{E: i \in E \subset I} \mathbb{E}[X^{I \setminus E}] \kappa[X_E]\,, \tag{1.1.14}$$

where i denotes the last element in the sequence I. There is only one term on the right hand side which has a cumulant of order m or higher, namely, the term with $E = I$. Hence, if we begin by setting $\kappa[X_j] = \mathbb{E}[X_j]$ for all $j \in J$, the formula is correct for all sequences of length $m = 1$. We can then define all cumulants of length $m = 2$ by using the following formula, equivalent to (1.1.14)

$$\kappa[X_I] = \mathbb{E}[X^I] - \sum_{E: i \in E \subsetneq I} \mathbb{E}[X^{I \setminus E}] \kappa[X_E]\,.$$

This procedure can then be iterated until the final order $m = n$ has been reached, showing also that the requirement (1.1.14) determines the cumulants uniquely once the moments have been fixed.

It is somewhat more straightforward to define cumulants when the collection of random variables has exponential moments. Then we can the *joint cumulant generating function* $g_c(\eta) :=$

$\ln G_{\mathrm{mom}}(\eta)$, $\eta \in \mathbb{R}^J$, to define the cumulants via the formula $\kappa[X_I] = \partial_\eta^I g_{\mathrm{c}}(0)$. These numbers turn out to be identical to the above recursively defined cumulants $\kappa[X_I]$. To see this, consider an arbitrary $I \in J^m$ and compute the first partial derivative corresponding to the last index i in I using the chain rule in the identity $G_{\mathrm{mom}}(\eta) = \exp(g_{\mathrm{c}}(\eta))$. This yields $\partial_{\eta_i} G_{\mathrm{mom}}(\eta) = G_{\mathrm{mom}}(\eta)\partial_{\eta_i} g_{\mathrm{c}}(\eta)$. The remaining derivatives can be distributed between the two factors using the Leibniz rule, and once the result is evaluated at $\eta = 0$, we find that (1.1.14) holds. As pointed out above, the solution to these constraints is unique and this implies that the two definitions must agree.

Although the empty cumulant $\kappa[X_\varnothing]$ never appears in the above procedure, to be consistent with derivatives of the cumulant generating function, we set $\kappa[X_\varnothing] := 0$.

The following known properties of cumulants can then be derived directly from the recursive definition using induction in m. First, by iterating the defining equation (1.1.14) until only expectations of empty powers ($\mathbb{E}[X^\varnothing] = 1$) remain, one can identify each of the terms in the remaining sum with a unique partition of the original index set. As a result, we arrive at the highly useful *moment-to-cumulants* formula

$$\mathbb{E}[X^I] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} \kappa[X_A]\,, \tag{1.1.15}$$

where we have used the notations introduced in the beginning of this subsection. Moreover, the following properties, which are obviously true for monomials, are inherited by the cumulants.

1. The cumulants are permutation invariant: if I' is a permutation of I, then $\kappa[X_{I'}] = \kappa[X_I]$.

2. The cumulants are multilinear in each of their arguments. More precisely, if $a, b \in \mathbb{R}$ and Y, Y' are random variables, then

$$\kappa[X_1, \ldots, aY + bY', \ldots, X_m] = a\,\kappa[X_1, \ldots, Y, \ldots, X_m] + b\,\kappa[X_1, \ldots, Y', \ldots, X_m]\,.$$

   It should be stressed that this formula needs to be properly applied for cumulants of just one random variable discussed earlier. Namely, we have $\kappa_n[X] = \kappa[X, X, \ldots, X]$, and thus κ_n is *not* usually a linear function of X. Instead, it is a homogeneous function of degree n, since $\kappa_n[aX] = a^n \kappa_n[X]$ by multilinearity.

The main reason why it often is better (and sometimes easier) to study cumulants instead of moments is in the following two properties which are not true for the corresponding moments (apart from some degenerate examples):

1. If the joint distribution of the random variables is *Gaussian*, then all cumulants of order 3 or higher are zero.

2. If *any* one of the random variables is *independent* from all the others, then all proper joint cumulants involving this random variable are zero. For example, if X_1 is independent from $X_2, \ldots, X_N$, then

$$\kappa[X_I] = 0\,, \qquad \text{if } 1 \in I \text{ and there is } j \in I \text{ with } j \neq 1\,.$$

   Note that we can still have $\kappa[X_I] \neq 0$, if I does not contain 1 or if it is a repetition of 1 only.

This makes cumulants a good quantitative tool to analyse random variables which are expected to be nearly independent and/or Gaussian since then some of their cumulants should be small.

Although seldom used, there is also a formula which expresses cumulants directly in terms of moments (*cumulant-to-moments* formula). Namely,

$$\kappa[X_I] = \sum_{\pi \in \mathcal{P}(I)} (|\pi| - 1)!\,(-1)^{|\pi|-1} \prod_{A \in \pi} \mathbb{E}[X^A]\,, \tag{1.1.16}$$

where $|\pi|$ denotes the number of clusters in the partition π. The alternating signs and large factorial prefactors often make this formula of little use in practice.

Remark 1.1.2 (Statistical independence) Independence in probability relates to factorization of the joint measure of the random variables; Wikipedia has a nice summary of statistical independence of events and random variables, `https://en.wikipedia.org/wiki/Independence_(probability_theory)`. For example, two real-valued random variables X and Y are independent if $\mathbb{P}(X \leqslant a \text{ and } Y \leqslant b) = \mathbb{P}(X \leqslant a)\,\mathbb{P}(Y \leqslant b)$ for all a, b. In general, for the type of probability measures studied here, this occurs if and only if for all bounded continuous functions f, g

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\,\mathbb{E}[g(Y)]\,.$$

One can prove the above independence statement directly from this property. Namely, if X_1 is independent from $X_2, \ldots, X_N$, then

$$\mathbb{E}[\mathrm{e}^{\mathrm{i}\mathbf{s}\cdot\mathbf{X}}] = \mathbb{E}[\mathrm{e}^{\mathrm{i}s_1 X_1}]\,\mathbb{E}[\mathrm{e}^{\mathrm{i}\sum_{i=2}^{N} s_i X_i}]\,.$$

Thus the cumulant generating function $\ln \mathbb{E}[\mathrm{e}^{\mathrm{i}\mathbf{s}\cdot\mathbf{X}}]$ is a sum of two terms, the first of which depends only on s_1 and the second does not depend on s_1. Hence, any derivative over both s_1 and some other variable yields zero.

1.1.3 Complex-valued random variables(*)

If $\Psi : \Omega \to \mathbb{C}$ is a complex random variable, all of the above method generalize straightforwardly by using a two-component vector $X = (\operatorname{Re}\Psi, \operatorname{Im}\Psi)$ of real random variables. Note that then

$$\mathbf{s} \cdot \mathbf{X} = s_1 \operatorname{Re}\Psi + s_2 \operatorname{Im}\Psi = 2\operatorname{Re}(z^*\Psi) = z^*\Psi + \Psi^* z\,, \qquad z := \frac{1}{2}(s_1 + \mathrm{i}s_2)\,.$$

So we may equally well consider the complex characteristic function

$$\varphi(\mathbf{z}) := \mathbb{E}_\mu\left[\mathrm{e}^{2\mathrm{i}\operatorname{Re}\langle \mathbf{z}, \mathbf{\Psi}\rangle}\right] = \mathbb{E}_\mu\left[\mathrm{e}^{\mathrm{i}(\langle \mathbf{z}, \mathbf{\Psi}\rangle + \langle \mathbf{\Psi}, \mathbf{z}\rangle)}\right]\,, \qquad \mathbf{z} \in \mathbb{C}^J\,, \tag{1.1.17}$$

for a vector $\mathbf{\Psi}$ of complex random variables and using the scalar product $\langle \mathbf{z}, \mathbf{\Psi}\rangle := \sum_{j\in J} z_j^* \Psi_j$. The Wirtinger derivatives for parametrization $z = x + \mathrm{i}y$, $x = \operatorname{Re} z$, $y = \operatorname{Im} z$, are given by

$$\frac{\partial}{\partial z} := \frac{1}{2}\left(\frac{\partial}{\partial x} - \mathrm{i}\frac{\partial}{\partial y}\right)\,, \qquad \frac{\partial}{\partial z^*} := \frac{1}{2}\left(\frac{\partial}{\partial x} + \mathrm{i}\frac{\partial}{\partial y}\right)\,. \tag{1.1.18}$$

Using these definitions, we also find

$$(-\mathrm{i}\partial_{z^*})^I \varphi(\mathbf{0}) = \mathbb{E}[\Psi^I]\,, \quad (-\mathrm{i}\partial_z)^I \varphi(\mathbf{0}) = \mathbb{E}[(\Psi^*)^I]\,,$$

under the same conditions as discussed before for the real random variables. Similarly, mixed derivatives $(-\mathrm{i}\partial_{z^*})^I(-\mathrm{i}\partial_z)^{I'}\varphi(\mathbf{0})$ yield all mixed moments of Ψ and Ψ^*.

1.2 Uniform distribution on the sphere and δ-function constraints

The stationary state of the Kac model with N one-dimensional particles and under standard choice of its parameters and initial data, is the uniform distribution on the sphere $S^{N-1}(\sqrt{N})$. Let us use this case to illustrate the definition of a measure via "δ-function constraints".

Technically, what we mean by the uniform distribution on the sphere $S^{N-1}(R)$, $R > 0$, is a singular Radon probability measure on $\mathbb{R}^N$. The "uniformity" refers here to rotations on the sphere, i.e., we assume that the measure remains invariant under any rotation of $\mathbb{R}^N$, that is, under the transformation $v \mapsto Ov$ where $O \in \mathbb{R}^{N\times N}$ is an orthogonal matrix. We later use the notation $\nu_{N,R}$ for this measure but, for brevity, let us simply denote it by ν in this section.

Remark 1.2.1 (Invariance of a measure) We recall that a measure μ on X is invariant under a measurable transformation $f : X \to X$ if $\mu(f^{\leftarrow}A) = \mu(A)$ for all measurable sets A and that this occurs if and only $\mathbb{E}_\mu[g \circ f] = \mathbb{E}_\mu[g]$ for all integrable functions g (see `https://en.wikipedia.org/wiki/Invariant_measure`, `https://en.wikipedia.org/wiki/Pushforward_measure`). For measures defined by the RMK theorem and continuous transformations f which do not destroy compactness of the support, it thus suffices to check that $\mathbb{E}_\mu[g \circ f] = \mathbb{E}_\mu[g]$ for all testfunctions g. In particular, this applies for measures on $\mathbb{R}^N$ if f is a rotation, as above.

In the δ-function notation commonly used in physics literature, one can in fact denote

$$\nu(\mathrm{d}v) := \frac{1}{Z_R}\delta(|v| - R)\mathrm{d}^N v$$

where $|v|^2 = \sum_{i=1}^N v_i^2$ corresponds to the standard Euclidean norm and $Z_R > 0$ is a normalization constant so that $\mu(\mathbb{R}^N) = 1$. The precise meaning of the above δ-function notation is determined by replacing it by a suitable sequence of approximants. Although the measure ν is simple enough to be defined directly via polar coordinates, let us use it to illustrate the approximant procedure.

The goal is to define a Radon probability measure ν on the locally compact space $\mathbb{R}^N$ via the RMK representation theorem 1.4.1. To this end, first choose a sequence $\varepsilon_n > 0$, $n \in \mathbb{N}$, for which $\varepsilon_n \to 0$. Then, define a sequence of approximants by replacing the δ-function by a Gaussian function with a standard deviation ε_n, i.e., we first set for $y \in \mathbb{R}$

$$G_n(y) := (2\pi\sigma^2)^{-\frac{1}{2}}\mathrm{e}^{-\frac{1}{2\sigma^2}X^2}\,, \qquad \sigma := \varepsilon_n\,,$$

for which, formally, $G_n(y) \to \delta(y)$, as $n \to \infty$. We then define for any continuous test-function with a compact support, i.e., whenever $f \in C_c(\mathbb{R}^N)$,

$$\Lambda_n[f] := \int_{\mathbb{R}^N} \mathrm{d}^N v\, f(v) G_n(|v| - R)\,.$$

Since f is a bounded continuous function, the above integral is always finite. In addition, if $f \geqslant 0$, also $\Lambda_n[f] \geqslant 0$.

We next show that the limit $\lim_{n\to\infty}\Lambda_n[f]$ exists for all $f \in C_c(\mathbb{R}^N)$. Rewriting the integral in spherical coordinates, we obtain

$$\begin{aligned}\Lambda_n[f] &= \int_0^\infty \mathrm{d}r\, r^{N-1} G_n(r - R)\int_{S^{N-1}} \mathrm{d}\Omega\, f(r\Omega)\\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \mathrm{d}x\,\mathrm{e}^{-\frac{1}{2}x^2}\mathbb{1}_{\{x\geqslant -R/\varepsilon_n\}}(R + \varepsilon_n x)^{N-1}\int_{S^{N-1}} \mathrm{d}\Omega\, f((R + \varepsilon_n x)\Omega)\,,\end{aligned}$$

where we have made a change of variables from r to $x = \frac{r-R}{\varepsilon_n}$. Then the x-integrand is bounded by the n-independent, integrable function $\mathrm{e}^{-\frac{1}{2}x^2}(R+M|x|)^{N-1}|S^{N-1}|\|f\|_\infty$, where $M = \sup_n \varepsilon_n < \infty$, and as $n \to \infty$, it converges to

$$\mathrm{e}^{-\frac{1}{2}x^2} R^{N-1}\int_{S^{N-1}} \mathrm{d}\Omega\, f(R\Omega)\,,$$

for all $x \in \mathbb{R}$, using continuity and boundedness of f. Hence, we can use the dominated convergence theorem to conclude that for all test-functions f,

$$\lim_{n\to\infty}\Lambda_n[f] = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^\infty \mathrm{d}x\,\mathrm{e}^{-\frac{1}{2}x^2}R^{N-1}\int_{S^{N-1}}\mathrm{d}\Omega\, f(R\Omega) = R^{N-1}\int_{S^{N-1}}\mathrm{d}\Omega\, f(R\Omega)\,.$$

Indeed, the outcome is what one would expect from the formal integral

$$\int_{\mathbb{R}^N}\mathrm{d}^N v\,\delta(|v| - R) f(v)\,,$$

using the spherical coordinate representation.

We then choose $Z_R := R^{N-1}|S^{N-1}|$, and define

$$\Lambda[f] := \frac{1}{Z_R} \lim_{n\to\infty} \Lambda_n[f] , \qquad f \in C_c(\mathbb{R}^N) .$$

Then the functional Λ satisfies all the assumptions necessary for using the RMK theorem, and we can conclude that there is a unique positive Radon measure ν such that

$$\Lambda[f] = \int_{\mathbb{R}^N} \nu(\mathrm{d}v) f(x) .$$

Considering a sequence of test-functions $g_n \in C_c(\mathbb{R}^N, [0,1])$ for which $g_n(x) = 1$ whenever $|x| \leqslant n$, we find that $\Lambda[g_n] \to \frac{1}{Z_R} R^{N-1}|S^{N-1}| = 1$. Therefore, ν is a probability measure.

Many properties of such measures defined by δ-constraints are easier to check using the approximant representation for them. For example, it is now straightforward to prove that ν is invariant under all rotations: Suppose $O \in \mathbb{R}^{N\times N}$ is an orthogonal matrix for which $|Ov| = |v|$ for all $v \in \mathbb{R}^N$. If f is a test-function, also $f \circ O$ is a test-function, and thus $\mathbb{E}_\nu[f \circ O] = \lim_{n\to\infty} \Lambda_n[f \circ O]$. However,

$$\begin{aligned}\Lambda_n[f \circ O] &= \int_{\mathbb{R}^N} \mathrm{d}^N v\, f(Ov) G_n(|v| - R) = \int_{\mathbb{R}^N} \mathrm{d}^N v' \,|\det O^{-1}| f(v') G_n(|O^{-1}v'| - R) \\ &= \int_{\mathbb{R}^N} \mathrm{d}^N v'\, f(v') G_n(|v'| - R) = \Lambda_n[f] ,\end{aligned}$$

where we have made a change of variables $v' = Ov$. Taking $n \to \infty$ from both sides of the equality shows that $\mathbb{E}_\nu[f \circ O] = \mathbb{E}_\nu[f]$ and thus any orthogonal transformation O leaves the measure ν invariant, as claimed.

As an corollary, we now also find that the measure ν is *label permutation invariant*. Label permutation transformations on $\mathbb{R}^N$ are defined by permutations π of the label set $[N]$, i.e., by bijetions $\pi : [N] \to [N]$. Given such π, the corresponding transformation is

$$(Tv)_i := v_{\pi(i)} , \qquad i = 1, 2, \ldots , N .$$

Clearly, T is a linear map and $|Tv|^2 = |v|^2$, so T is an example of an orthogonal transformation. Therefore, $T_*\nu = \nu$, as well. We will use the term *symmetric* for general label permutation invariant measures on $S^{N-1}(R)$.

One final matter to clarify is to check that indeed ν is a "measure on the sphere $S^{N-1}(R)$." More precisely, we wish to check that the support of this Borel measure is contained in $S^{N-1}(R)$. To do this, we need to show that if $x \in \mathbb{R}^N \setminus S^{N-1}(R)$, then there is a radius $r > 0$ such that the measure of the ball $B(x, r)$ is zero.

For this, consider the open set $E := \{v \in \mathbb{R}^n \,|\, |v| \neq R\}$, and define for all $j \in \mathbb{N}_+$ the closed sets $E_j := \left\{v \in \mathbb{R}^n \,\middle|\, ||v| - R| \geqslant \frac{1}{j}\right\}$. Clearly, $\cup_j E_j = E$, and by Urysohn's lemma to each j there exists a continuous function f_j such that $f_j(v) = 1$ if $v \in E_j$, and $f_j(v) = 0$ if $v \notin E$. Now $\int_{\mathbb{R}^n} \mu(\mathrm{d}v)\, f_j(v) = \frac{1}{|S^{N-1}|} \int_{S^{N-1}} \mathrm{d}\Omega\, f_j(R\Omega)$. Since $f_j(R\Omega) = 0$ for all Ω, it follows that

$$0 \leqslant \mu(E) \leqslant \sum_j \int_{\mathbb{R}^n} \nu(\mathrm{d}v)\, f_j(v) = 0 .$$

Therefore, $\nu(E) = 0$ and $|v|^2 = R$ almost surely under ν.

1.3 Cumulants of generic symmetric measures on $S^{N-1}(\sqrt{N})$

(NB: Some of the notations used here are defined only later, in Chapter 3.)

Proposition 1.3.1 *Let F^N be a symmetric measure on $S^{N-1}(\sqrt{N})$ with $N \geqslant 2$. The joint energy cumulants of order $n \leqslant \frac{N}{2}+1$ satisfy a bound*

$$|\kappa[e_J]| \leqslant 4^{n-1}(n-1)!N^{n-\mathrm{len}(J)}\,, \qquad |J| = n\,. \tag{1.3.1}$$

where $\mathrm{len}(J)$ *denotes the number of distinct labels in* J.

Proof Denote by $\mathbb{E}[\phi]$ the expectation of ϕ with respect to the measure F^N.

If $n = 1$, the first order cumulant equals expectation value which by symmetry is equal to $\mathbb{E}[e_1] = \frac{1}{N}\sum_{i=1}^N \mathbb{E}[e_i] = 1$, since $\sum_{i=1}^N e_i = N$ almost surely. Hence, the stated bound holds for $n = 1$.

We next establish an estimate

$$0 \leqslant \mathbb{E}[e_1 e_2 \cdots e_n] \leqslant 2^{n-1}\,, \quad 1 \leqslant n \leqslant \frac{N}{2}+1\,. \tag{1.3.2}$$

The estimate clearly holds for $n = 1$, and we claim that for other n it satisfies

$$0 \leqslant \mathbb{E}[e_1 e_2 \cdots e_n] \leqslant 2\mathbb{E}[e_1 e_2 \cdots e_{n-1}]\,. \tag{1.3.3}$$

Combining these results proves the claim by induction.

To prove (1.3.3), we first note that

$$\begin{aligned} N\mathbb{E}[e_1 \cdots e_{n-1}] &= \sum_{i=1}^N \mathbb{E}[e_1 e_2 \cdots e_{n-1} e_i] = (N-(n-1))\mathbb{E}[e_1 \cdots e_n] + (n-1)\mathbb{E}[e_1^2 \cdots e_{n-1}] \\ &\geqslant (N-(n-1))\mathbb{E}[e_1 \cdots e_n] \end{aligned} \tag{1.3.4}$$

since the measure is symmetric and we have $e_i \geqslant 0$ and $\sum_{i=1}^N e_i = N$ almost surely. Therefore,

$$\mathbb{E}[e_1 \cdots e_n] \leqslant 2\mathbb{E}[e_1 \cdots e_{n-1}]\,, \tag{1.3.5}$$

using the assumption $n \leqslant \frac{N}{2}+1$. Hence, (1.3.3) holds.

Let us then assume $2 \leqslant n \leqslant \frac{N}{2}+1$. By symmetry, we only need to prove the bound (1.3.1) for $J = J_r$ with $r \in \mathcal{C}_n$. Denote $k := \mathrm{len}(J)$ for which $1 \leqslant k \leqslant n$, $\sum_{\ell=1}^k r_\ell = n$, and $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_k \geqslant 1$. Then, for each $\ell \in [k]$ we may use the trivial energy bound $e_\ell \leqslant N$, which holds almost surely, and the estimate in (1.3.3) to conclude that

$$0 \leqslant \mathbb{E}[e^J] \leqslant N^{\sum_{\ell=1}^k (r_\ell - 1)}\mathbb{E}[e_1 \cdots e_k] \leqslant N^{n-k}2^{k-1}\,. \tag{1.3.6}$$

We next use the general iteration relation satisfied by cumulants

$$\mathbb{E}[e^J] = \sum_{(1,J_1)\in I\subseteq J} \mathbb{E}[e^{I\setminus J}]\kappa[e_J].$$

Combined with the already established energy moment bounds this yields an estimate

$$\begin{aligned} |\kappa[e_J]| &\leqslant \left|\mathbb{E}[e^J]\right| + \sum_{(1,J_1)\in I\subsetneq J} \left|\mathbb{E}[e^{J\setminus I}]\right| |\kappa[e_I]| \\ &\leqslant N^{n-k}2^{k-1} + \sum_{(1,J_1)\in I\subsetneq J} |\kappa[e_I]|\, N^{n'-k'}2^{k'-1}|_{n'=|J\setminus I|,\, k'=\mathrm{len}(J\setminus I)}\,. \end{aligned} \tag{1.3.7}$$

In this sum, $1 \leqslant |I|, n' \leqslant n-1$ and $1 \leqslant k' \leqslant k, n'$. In addition, $n' + |I| = n$ and $k' + \mathrm{len}(I) \geqslant k$.

Let us then denote $A_{m,k} := \max_{I:|I|=m,\mathrm{len}(I)=k} |\kappa[e_I]| N^{1-m}$ for $1 \leqslant m \leqslant \frac{N}{2}+1$ and $1 \leqslant k \leqslant m$. This yields a sequence which, by the above bound, satisfies

$$A_{n,k} \leqslant (2/N)^{k-1} + \sum_{(1,J_1)\in I\subsetneq J} A_{n-n',len(I)} N^{-k'} 2^{k'-1}|_{n'=|J\setminus I|,\, k'=\mathrm{len}(J\setminus I)} \,. \tag{1.3.8}$$

Here, $n' = |J\setminus I|$ satisfies $1 \leqslant n' \leqslant n-1$ and for each n' there are at most $\binom{n-1}{n'}$ terms with $|J\setminus I| = n'$ in the sum. Therefore, $B_n := \max_{1\leqslant m\leqslant n,\, 1\leqslant k\leqslant m} \left(A_{m,k}\frac{(N/2)^{k-1}}{(m-1)!}\right)$, satisfy an inequality

$$B_n \leqslant \frac{1}{(n-1)!} + B_{n-1}\frac{1}{2}\sum_{n'=1}^{n-1}\binom{n-1}{n'}\frac{(n-n'-1)!}{(n-1)!}\,, \tag{1.3.9}$$

for all $2 \leqslant n \leqslant \frac{N}{2}+1$. Since $B_1 = A_{1,1} = 1$ we obtain $B_2 \leqslant 1 + \frac{1}{2} < 2$. For $3 \leqslant n \leqslant \frac{N}{2}+1$, it follows that

$$B_n \leqslant \frac{1}{2}\left(1 + B_{n-1}(\mathrm{e}-1)\right)\,. \tag{1.3.10}$$

Therefore, by induction, we can conclude that

$$B_n \leqslant 2^{n-1}\,, \qquad 1 \leqslant n \leqslant \frac{N}{2}+1\,.$$

Hence, for all such n and whenever $|J| = n$ and $\mathrm{len}(J) = k$, we have

$$|\kappa[e_J]| \leqslant N^{n-1}A_{n,k} \leqslant N^{n-k}2^{k-1}(n-1)!B_n \leqslant N^{n-k}2^{n+k-2}(n-1)!\,.$$

Since $k \leqslant n$, (1.3.1) also holds. □

1.3.1 Extreme case: Symmetrized deterministic initial data

In this subsection, we construct an example of initial data on $S^{N-1}(\sqrt{N})$ which is symmetric but highly non-chaotic, in the sense that it saturates the power of N in the generic bound in (1.3.1).

Pick $\bar{v} \in S^{N-1}(\sqrt{N})$. Define a measure

$$\mu := \frac{1}{N!}\sum_{\sigma\in\mathfrak{S}_N} f_{\sigma,*}[\delta_{\bar{v}}], \tag{1.3.11}$$

on $S^{N-1}(\sqrt{N})$ with Borel σ-algebra. The measure $f_{\sigma,*}[\delta_{\bar{v}}]$ is obtained from the Dirac measure at $\bar{v}$ by pushing it forward by a component permutation map $f_\sigma \colon S^{N-1}(\sqrt{N}) \to S^{N-1}(\sqrt{N})$:

$$f_{\sigma,*}[\delta_{\bar{v}}](B) = \delta_{\bar{v}}(f_\sigma^{-1}(B)) \tag{1.3.12}$$

with

$$f_\sigma^{-1}(v_1,\ldots,v_N) = (v_{\sigma^{-1}(i)})_{i=1}^N. \tag{1.3.13}$$

Then for any continuous function $\phi \in C(S^{N-1}(\sqrt{N})) = C_c(S^{N-1}(\sqrt{N}))$, we have

$$\int_{S^{N-1}(\sqrt{N})} \mu(\mathrm{d}v)\phi(v) = \frac{1}{N!}\sum_{\sigma\in\mathfrak{S}_N}\phi(\bar{v}_{\sigma(i)})_{i=1}^N\,, \tag{1.3.14}$$

which also could serve as a definition of the measure μ, by the Riesz–Markov–Kakutani representation theorem. It is clearly a symmetric measure.

The expectation of energy with respect to this measure are fixed by symmetry to be 1. Since the energy observable is continuous, we can also easily compute all moments from (1.3.14). For example, the non-repeating second moments are given by

$$\mathbb{E}_\mu(e_1 e_2) = \int_{S^{N-1}(\sqrt{N})} e_1(v)e_2(v)\mu(\mathrm{d}v) = \frac{1}{N!}\sum_{\sigma\in\mathfrak{S}_N} \bar{v}^2_{\sigma(1)}\bar{v}^2_{\sigma(2)} = \frac{1}{N(N-1)}\sum_{i,j=1;j\neq i}^{N} \bar{v}_i^2\bar{v}_j^2\,, \tag{1.3.15}$$

and the second cumulant, which is equal to covariance, is

$$\kappa_\mu[e_1,e_2] = \frac{1}{N(N-1)}\sum_{i,j=1;j\neq i}^{N} \bar{v}_i^2\bar{v}_j^2 - 1 = \frac{1}{N}\sum_{i=1}^{N}\bar{v}_i^2\frac{1-\bar{v}_i^2}{N-1}\,.$$

Similarly, we can compute the variance to be

$$\kappa_\mu[e_1,e_1] = \frac{1}{N!}\sum_{\sigma\in\mathfrak{S}_N}\bar{v}^4_{\sigma(1)} - 1 = \frac{1}{N}\sum_{i=1}^{N}\bar{v}_i^2\left(\bar{v}_i^2-1\right)\,.$$

Let us consider a simple example of extreme initial data starting from a case where all the energy lies in the first particle: set $\bar{v} = (\sqrt{N},0,0,\ldots,0)$, and consider the corresponding symmetrized measure μ. By the above formula, its covariance is given by $\kappa_\mu[e_1,e_2] = -1$ and variance by $\kappa_\mu[e_1,e_1] = N-1$. Thus the upper bound in (1.3.1) is saturated for large N and $n=2$, apart from the prefactor 4^{n-1}.

The moment generating function for energy is explicitly given by

$$\mathbb{E}_\mu[\mathrm{e}^{e\cdot\xi}] = \frac{1}{N}\sum_{i=1}^{N}\mathrm{e}^{N\xi_i} = 1 + \frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{e}^{N\xi_i}-1\right) = 1 + X(\xi)\,,$$

and the cumulant generating function is thus

$$g_c(\xi) = \ln\mathbb{E}_\mu[\mathrm{e}^{e\cdot\xi}] = \sum_{m=1}^{\infty}\frac{(-1)^{m-1}}{m}X(\xi)^m\,.$$

Here $X(0)=0$, and we obtain a formula for cumulants with $J = J_r$, $r\in\mathcal{C}_n$,

$$\kappa[e_J] = \sum_{m=1}^{n}\frac{(-1)^{m-1}}{m}\sum_I\prod_{\ell=1}^{m}\partial_\xi^{I_\ell}X(0)\,.$$

The sum $\sum_I$ here denotes a sum over all partitions of the sequence J into a sequence containing m subsequences, $I = (I_\ell)_{\ell=1,\ldots,m}$. If $\mathrm{len}(I_\ell) > 1$, then $\partial_\xi^{I_\ell}X(0) = 0$, and if $\mathrm{len}(I_\ell) = 1$ and $p_\ell = |I_\ell|$, then $\partial_\xi^{I_\ell}X(0) = N^{p_\ell-1}$. Denoting $k = \mathrm{len}(J)$, we may thus conclude that the non-zero terms in this sum must have $m \geqslant k$, and then

$$\prod_{\ell=1}^{m}\partial_\xi^{I_\ell}X(0) = \prod_{\ell=1}^{m}N^{p_\ell-1} = N^{n-m}\,.$$

Hence, the dominant power of N occurs at $m=k$, in which case there are $k!$ choices for I which yield a non-zero contribution. We find that

$$\kappa[e_J] = (-1)^{k-1}(k-1)!N^{n-k} + O(N^{n-k-1})\,,$$

which shows that the power law of N in the bound in (1.3.1) is saturated for large N also for any finite $n > 2$.

1.4 Appendix: RMK representation theorem

To put the above discussion on a firm mathematical footing, we will need to following central result from measure theory (the proof of this theorem can be found from a textbook by W. Rudin: Real and Complex Analysis [2], Theorems 2.14 and 6.19, and the approximation result is contained in the proof of Lusin's theorem, Theorem 2.24).

Theorem 1.4.1 (Riesz–Markov–Kakutani (RMK) representation theorem) *Suppose X is a locally compact Hausdorff space. Let $C_c(X)$ denote the collection of all continuous functions $f : X \to \mathbb{C}$ which have a compact support. Suppose Λ is a linear functional on $C_c(X)$, i.e., $\Lambda : C_c(X) \to \mathbb{C}$ is a linear map.*

1. *If Λ is* positive *($\Lambda[f] \geqslant 0$ whenever $f \geqslant 0$), then there is a unique non-negative Radon measure μ on X such that*

$$\Lambda[f] = \int_X f(x)\mu(\mathrm{d}x)\,, \quad f \in C_c(X)\,. \tag{1.4.1}$$

   If A is a finite measurable set ($\mu(A) < \infty$), then its characteristic function can be approximated by a sequence of test functions: there are $g_n \in C_c(X, [0,1])$, $n \in \mathbb{N}$, and a set $E \subset X$, such that $\mu(E) = 0$ and $\lim_n g_n(x) = \mathbb{1}_{\{x \in A\}}$ for all $x \notin E$.

   Note that it is possible that $\int_X \mu(\mathrm{d}x) = \infty$. However, if one can find $g_n \in C_c(X, [0,1])$, $n \in \mathbb{N}$, such that $\lim_n g_n(x) = 1$ for (almost) every $x \in X$ and $\Lambda[g_n] \to 1$ as $n \to \infty$, then μ is a Borel probability measure on X.

2. *If Λ is* bounded *(there is $M > 0$ such that $|\Lambda[f]| \leqslant M\|f\|_\infty$ for all $f \in C_c(X)$, where $\|f\|_\infty := \sup_{x \in X} |f(x)|$), then there is a unique bounded complex Radon measure μ on X such that* (1.4.1) *holds. In this case, the characteristic function of any measurable set A can be approximated by a sequence (g_n) of test functions, as in the previous item.*

   (This item is usually summarized as a duality statement: $C_c(X)^* = \mathcal{M}_b(X)$.)

One application of the theorem is to uniquely define measures μ by merely stating how they are supposed to act on a restricted class of "nice" functions. Note that once we have the measure μ, the integrals on the right hand side of (1.4.1) are defined for any non-negative function which is a pointwise limit of continuous functions (since these are always Borel measurable), including characteristic functions of open and closed sets. All usual tools from measure theory (dominated convergence, etc.) are available, as well. For applications, it is good to recall that all open and closed subsets of $\mathbb{R}^d$ are locally compact Hausdorff spaces, and that a continuous function defined on $\mathbb{R}^d$ has a compact support[4] if and only if it is zero outside some ball.

Remark 1.4.2 Here "Radon measure" is a technical term for a measure on a topological space which is "compatible" with the topology. For example, restrictions of the Lebesgue measure on Borel σ-algebras of any open or closed subset of $\mathbb{R}^d$ are Radon measures. The precise definition is given below.

Definition 1.4.3 (Radon measure) *Suppose X is a Hausdorff topological space and consider its Borel σ-algebra (the smallest σ-algebra containing all open sets). A measure μ on the Borel σ-algebra is called a Radon measure if it is finite on all compact sets, outer regular on all Borel sets, and inner regular on open sets. (See e.g. Wikipedia for precise definitions.)*

[4] The support of a function $f : X \to \mathbb{C}$ is defined here to be the closure of the values where f is not zero, i.e., $\operatorname{supp} f := \overline{\{x \in X \mid f(x) \neq 0\}}$.

Chapter 2

Evolution hierarchies of cumulants

The second lecture concerns evolution of cumulants for certain class of dynamical systems. To obtain directly a hierarchy which depends only on the cumulants, one needs to have "rate functions" which depend polynomially on the evolving variables. However, both deterministic and stochastic evolution are possible.

2.1 General structure

Particularly convenient derivation of the evolution hierarchy is possible if we can use the real moment generating function G_{mom} to compute the cumulants. In essence, we will require that exponential moments exist uniformly on the time-interval we wish to study. Although a restriction, the assumptions are sufficiently loose to cover, for example, the evolution of the Kac model.

More precisely, we assume that there is a stochastic process such that for every t on some fixed time-interval T_I its trajectories $X(t) \in \mathbb{R}^N$, $t \in T_I$, have the following properties:

1. (*uniform exponential moments*) There are $\varepsilon_0, M_0 > 0$ such that for all $t \in T_I$

$$\mathbb{E}[\mathrm{e}^{\varepsilon_0 |X(t)|}] \leqslant M_0 \,.$$

   Then the function $G_{\mathrm{mom}}(\eta; t) := \mathbb{E}[\mathrm{e}^{\eta \cdot X(t)}]$ is well-defined and analytic in the open ball of $\mathbb{C}^N$ with $|\eta| < \varepsilon_0$.

2. (*polynomial rate function*) There is a function $\mathcal{R}(X, \eta)$ such that for all $t \in T_I$ the time-derivative of G_{mom} exists and satisfies

$$\partial_t G_{\mathrm{mom}}(\eta; t) = \mathbb{E}[\mathcal{R}(X(t), \eta)\mathrm{e}^{\eta \cdot X(t)}] \,.$$

   We assume that the rate function $\mathcal{R}(X, \eta)$ is analytic in η and exponentially bounded, in the following sense: there are $C_1, \delta_1 > 0$ such that $|\mathcal{R}(X, \eta)| \leqslant C_1 \mathrm{e}^{\frac{\varepsilon_0}{2}|X|}$ for all $|\eta| \leqslant \delta_1$ and all X. In addition, we assume that the derivatives of $\mathcal{R}$ at $\eta = 0$ are polynomials of X, i.e., for any multi-index I, the derivative $P_I(X) := \partial_\eta^I \mathcal{R}(X, \eta)\big|_{\eta=0}$ defines a polynomial of $X \in \mathbb{R}^N$.

Under the above assumptions, we may conclude that the cumulants of X satisfy for all $t \in T_I$ the evolution equation

$$\partial_t \kappa[X(t)_I] = \sum_{J \subset I} \mathbb{E}[P_{I \setminus J}(X(t)) \, {:}X(t)^J{:}] \,, \tag{2.1.1}$$

where ${:}X(t)^J{:}$ denotes the Wick polynomial of order J of the random variables $X(t)$. As we will see soon, this implies that the right hand side of (2.1.1) is a finite order polynomial of the cumulants $\kappa[X(t)_\cdot]$. Hence, the collection of these equations over all choices of I results in an evolution hierarchy for the cumulants.

The derivation of this equation is fairly straightforward, assuming sufficient regularity so that, for example, all partial derivatives commute with themselves and with taking the expectation. Namely, since $g_c = \ln G_{\rm mom}$, chain rule can be applied, to yield

$$\partial_t g_c(\eta;t) = \frac{\partial_t G_{\rm mom}(\eta;t)}{G_{\rm mom}(\eta;t)} = {\rm e}^{-g_c(\eta;t)}\mathbb{E}[\mathcal{R}(X(t),\eta){\rm e}^{\eta\cdot X(t)}] = \mathbb{E}[\mathcal{R}(X(t),\eta)G_{\rm wick}(X(t);\eta)]$$

using the definition of the generating function of Wick polynomials,

$$G_{\rm wick}(X;\eta) := {\rm e}^{\eta\cdot X - g_c(\eta)}\,,$$

where g_c is the cumulant generating function of the random variables X. Therefore, by the Leibniz rule for differentiation of a product, we obtain

$$\begin{aligned}\partial_t \kappa[X(t)_I] &= \partial_\eta^I \partial_t g_c(\eta;t)\big|_{\eta=0} = \partial_\eta^I \mathbb{E}[\mathcal{R}(X(t),\eta)G_{\rm wick}(X(t);\eta)]\big|_{\eta=0} \\ &= \sum_{J\subset I}\mathbb{E}[\partial_\eta^{I\setminus J}\mathcal{R}(X(t),\eta)\partial_\eta^J G_{\rm wick}(X(t);\eta)]\Big|_{\eta=0} = \sum_{J\subset I}\mathbb{E}[P_{I\setminus J}(X(t))\,{:}X(t)^J{:}]\,.\end{aligned}$$

The above assumptions are one possible choice which would allow everywhere changing the order of derivatives and integrations above. The reason why analyticity is particularly simple for this, is the fact that then all derivatives may be expressed as integrals and, hence, changing the order with other integrals is possible, by relying on Fubini's theorem.

Remark 2.1.1 The basic tools for completing the details of the above argument from complex analysis are Morera's theorem and Cauchy's integral formula for derivatives. The first says that a continuous, complex valued function $f(z)$ is analytic on an open subset of the plane, if for every closed piecewise C^1 curve γ inside the open set, the contour integral $\oint_\gamma {\rm d}z f(x) = 0$. The assumption about uniform exponential moments first allows using dominated convergence theorem to conclude that the map is continuous in η. Furthermore, it can then be used to justify exchanging the contour integral over any η_i and the expectation $\mathbb{E}$. Since $\eta_i \mapsto {\rm e}^{\eta_i X_i}$ is entire, the resulting integral yields zero, and thus by Morera's theorem, the function $G_{\rm mom}$ is analytic in each of its arguments. Cauchy's integral formula then implies that for each i and any order $k_i \in \mathbb{N}_0$, we have

$$\partial_{\eta_i}^{k_i} G_{\rm mom}(\eta|_{\eta_i\to 0}) = k_i!\oint_{\gamma_i}\frac{{\rm d}z_i}{2\pi{\rm i}}\frac{G_{\rm mom}(\eta|_{\eta_i\to z_i})}{z_i^{k_i+1}}$$

where γ_i is a contour which goes once around a circle centred at the origin and of sufficiently small radius in the positive direction.

In the derivation of the cumulant hierarchy, one needs to apply the same argument to the product $\mathcal{R}(X,\eta)G_{\rm wick}(X,\eta)$. This requires checking analyticity of $g_c(\eta)$ which, however, also follows from the exponential moment bounds, after possibly reducing the radius ε_0 so that always $|G_{\rm mom}(\eta) - 1| \leqslant \frac{1}{2}$.

2.2 Wick polynomials

(This section is adapted from [3] which also contains more details and discussion about Wick polynomials.)

Let us first recall the moments-to-cumulants formula which holds for any index sequence I as long as all moments X^A, $A \subset I$, belong to $L^1(\mu)$:

$$\mathbb{E}[X^I] = \sum_{\pi\in\mathcal{P}(I)}\prod_{A\in\pi}\kappa[X_A]\,, \tag{2.2.1}$$

where $\mathcal{P}(I)$ denotes the collection of partitions of the set I. For consistency with the notations in [3], let us for this Section denote the collection of index sequences by $\mathcal{I}$. For a partition $\pi \in \mathcal{P}(I)$, let us call the subsets $A \in \pi$ *clusters* or *blocks*. Let us also recall that the cumulants are multilinear, i.e., they are separately linear in each of the variables X_j, $j \in I$.

We next show that it is possible to choose a subset of the indices and remove all its "internal clusters" from the moments-to-cumulants formula by replacing the corresponding power with a polynomial of the same order. This will be achieved by using the following recursive definition.

Definition 2.2.1 *Suppose that $I_0 \in \mathcal{I}$ is such that $\mathbb{E}[|X^I|] < \infty$ for all $I \subset I_0$. We define polynomials $\mathcal{W}[X_I] := \sum_{E\subset I} c_E[X^I]\, X^E$ for $I \subset I_0$ inductively in $|I|$ using the following rule: set $\mathcal{W}[X_\varnothing] := 1$, and for $I \neq \varnothing$ use*

$$\mathcal{W}[X_I] := X^I - \sum_{\varnothing\neq E\subset I} \mathbb{E}[X^E]\,\mathcal{W}[X_{I\setminus E}]\,. \tag{2.2.2}$$

The definition makes sense since the $\mathcal{W}$-terms on the right hand side all have an order lower than $|I|$. It also implies that indeed each $\mathcal{W}[X_I]$ is a polynomial of order $|I|$, with only the term X^I being of the highest order. It is also straightforward to prove by induction that the coefficients $c_E[X^I]$ can be chosen so that they only depend on $\mathbb{E}[X^A]$ with $A \subset I$. Therefore, the definition of $\mathcal{W}[X_I]$ is independent of I_0 in the following sense: if $I_0, I_0' \in \mathcal{I}$ are such that $\mathbb{E}[|X^I|] < \infty$ for every $I \subset I_0$ and $I \subset I_0'$, then for all $I \subset I_0 \cap I_0'$ we have $\mathcal{W}[X_I; I_0] = \mathcal{W}[X_I; I_0']$.

The following theorem shows that these polynomials indeed have the promised truncated moments-to-cumulants expansion. We also see that the polynomials are essentially uniquely defined by this property. What is perhaps surprising is that the coefficients of the polynomial can be chosen depending only on the moments of its constituent random variables. This implies that the same polynomial can be used for many different probability distributions, as long as the marginal distributions for the constituent random variables are the same.

Theorem 2.2.2 *Assume that the measure μ has all moments of order N, i.e., suppose that $\mathbb{E}[|X^I|] < \infty$ for all $I \in \mathcal{I}$ with $|I| \leqslant N$. Use Definition 2.2.1 to define $\mathcal{W}[X_I]$ for every such I.*

Then replacing X^I by $\mathcal{W}[X_I]$ removes all terms with clusters internal to I: the following truncated moments-to-cumulants formula holds for every $I' \in \mathcal{I}$ with $|I'| + |I| \leqslant N$

$$\mathbb{E}\Big[\mathcal{W}[X_I]X^{I'}\Big] = \sum_{\pi\in\mathcal{P}(I+I')} \mathbb{1}_{\{A\cap I'\neq\varnothing \text{ for all } A\in\pi\}} \prod_{A\in\pi} \kappa[X_A]\,. \tag{2.2.3}$$

In particular, $\mathbb{E}[\mathcal{W}[X_I]] = 0$ if $I \neq \varnothing$.

In addition, if $I \in \mathcal{I}$ with $|I| \leqslant N/2$ and $\mathcal{W}'$ is a polynomial of order at most $|I|$ such that (2.2.3) holds for all I' with $|I'| \leqslant N - |I|$, then $\mathcal{W}'$ is μ-almost surely equal to $\mathcal{W}[X_I]$.

As proven in [3], $\mathcal{W}[X_I]$ are μ-almost surely unique polynomials of order $|I|$ such that (2.2.3) holds for every $I' \in \mathcal{I}$. They can also be obtained from a generating function, for example if exponential moments are available. Namely, given sufficiently small $\eta \in \mathbb{R}^N$, we set for all $X \in \mathbb{R}^N$,

$$G_{\text{wick}}(\eta; X) = \mathrm{e}^{\eta\cdot X - g_c(\eta)} = \frac{\exp\left(\sum_{i=1}^N \eta_i X_i\right)}{\mathbb{E}[\exp\left(\sum_{i=1}^N \eta_i X_i\right)]}\,, \tag{2.2.4}$$

and then for any index sequence I we have

$$:\!X^I\!: = \partial_\eta^I G_{\text{wick}}(\eta; X)\big|_{\eta=0}\,, \tag{2.2.5}$$

which results in a polynomial of X. From now on, we will use the standard notation $:\!X^I\!:$ for Wick polynomials.

The next Proposition collects some of the most important properties of Wick polynomials.

Proposition 2.2.3 *The following statements hold for any $I \subset I_0$:*

1.

$$X^I = \sum_{U \subset I} {:}X^U{:}\, \mathbb{E}[X^{I \setminus U}] = \sum_{U \subset I} {:}X^U{:} \sum_{\pi \in \mathcal{P}(I \setminus U)} \prod_{A \in \pi} \kappa[X_A] \,. \tag{2.2.6}$$

2. Wick polynomials are argument permutation invariant: if I' is a permutation of I, then ${:}X^{I'}{:} = {:}X^I{:}$.

3.

$${:}X^I{:} = \sum_{U \subset I} X^U \sum_{\pi \in \mathcal{P}(I \setminus U)} (-1)^{|\pi|} \prod_{A \in \pi} \kappa[X_A] \,. \tag{2.2.7}$$

4. If $I' := \widehat{I}^{(1)}$ denotes the sequence obtained by cancelling the first element of I,

$${:}X^I{:} = X_{i_1} {:}X^{I'}{:} - \sum_{(1,i_1) \in V \subset I} \kappa[X_V] \,{:}X^{I \setminus V}{:} = X_{i_1} {:}X^{I'}{:} - \sum_{U \subset I'} \kappa[X_{(i_1)+U}] \,{:}X^{I' \setminus U}{:} \,. \tag{2.2.8}$$

5. The Wick polynomials are multilinear, i.e., if α, β are constants such that $X_j = \alpha X_i + \beta X_{i'}$ for some $j, i, i' \in J$, then, whenever I and k are such that $i_k = j$, we have

$${:}X^I{:} = \alpha \,{:}X^{\widehat{I}^{(k)}+(i)}{:} + \beta \,{:}X^{\widehat{I}^{(k)}+(i')}{:} \,.$$

Example 2.2.4 Written is terms of cumulants, the Wick polynomials of lowest order are

$$\begin{aligned}
{:}X{:} &= X - \kappa(X)\,, \\
{:}X_1X_2{:} &= X_1X_2 - \kappa(X_1,X_2) - \kappa(X_1)X_2 - \kappa(X_2)X_1 + \kappa(X_1)\kappa(X_2)\,, \\
{:}X_1X_2X_3{:} &= X_1X_2X_3 - \kappa(X_1,X_2,X_3) + \kappa(X_1,X_2)\kappa(X_3) + \kappa(X_1,X_3)\kappa(X_2) + \kappa(X_2,X_3)\kappa(X_1) \\
&\quad -\kappa(X_1)\kappa(X_2)\kappa(X_3) - \kappa(X_1,X_2)X_3 - \kappa(X_1,X_3)X_2 - \kappa(X_2,X_3)X_1 \\
&\quad +\kappa(X_1)\kappa(X_2)X_3 + \kappa(X_1)\kappa(X_3)X_2 + \kappa(X_2)\kappa(X_3)X_1 \\
&\quad -\kappa(X_1)X_2X_3 - \kappa(X_2)X_1X_3 - \kappa(X_3)X_1X_2\,.
\end{aligned}$$

The following result extends the earlier theorem and shows that multiple application of Wick polynomial replacements continues to simplify the moments-to-cumulants formula by removing all terms with any internal clusters.

Proposition 2.2.5 *Assume that the measure μ has all moments of order N, i.e., suppose that $\mathbb{E}[|X^I|] < \infty$ for all $I \in \mathcal{I}$ with $|I| \leqslant N$. Suppose $L \geqslant 1$ is given and consider a collection of $L+1$ index sequences $J', J_\ell \in \mathcal{I}$, $\ell = 1, \ldots, L$, such that $|J'| + \sum_\ell |J_\ell| \leqslant N$. Then for $I := \sum_{\ell=1}^L J_\ell + J'$ (with the implicit identification of J_ℓ and J' with the set of its labels in I) we have*

$$\mathbb{E}\Big[\prod_{\ell=1}^{L} {:}X^{J_\ell}{:}\, X^{J'} \Big] = \sum_{\pi \in \mathcal{P}(I)} \prod_{A \in \pi} \big(\kappa[X_A] \mathbb{1}_{\{A \not\subset J_\ell \ \forall \ell\}} \big) \,. \tag{2.2.9}$$

2.3 Deterministic dynamical systems

One case in which the above assumptions used for the derivation of the cumulant hierarchy are valid, is given by a dynamical system generated by a polynomial vector field. Explicitly, assuming that

$$\partial_t X(t) = \mathcal{P}(X(t))\,,$$

where each $\mathcal{P}_i(X)$, $i = 1, 2, \ldots, N$, is a polynomial in $X \in \mathbb{R}^N$, implies that

$$\partial_t (\mathrm{e}^{\eta \cdot X(t)}) = \eta \cdot \mathcal{P}(X(t))\, \mathrm{e}^{\eta \cdot X(t)}\,,$$

and thus we should set

$$\mathcal{R}(X,\eta) := \eta \cdot \mathcal{P}(X(t)) .$$

Then $\mathcal{R}(X,\eta)$ is analytic in η, and for any $\varepsilon_0, \delta_1 > 0$ it can be bounded by

$$|\mathcal{R}(X,\eta)| \leqslant C|X|^m |\eta| \leqslant C m! \delta_1 (2/\varepsilon_0)^m e^{\frac{\varepsilon_0}{2}|X|} ,$$

whenever $|\eta| \leqslant \delta_1$. Here, m is the largest degree of the polynomials $\mathcal{P}_i(X)$ and C does not depend on X nor on η. In addition, computing the derivatives with respect to η is simple, yielding

$$P_I(X) := \partial_\eta^I \mathcal{R}(X,\eta)\big|_{\eta=0} = \begin{cases} \mathcal{P}_i(X) , & i \in [N] \text{ and } I = (i) , \\ 0 , & \text{otherwise} , \end{cases}$$

all of which are polynomials in X.

Then, given some time-interval T_I, we need to choose the distribution of the initial data $X(0)$ in such a way that the solutions $X(t)$, $t \in T_I$, exist and do not increase too much, so that

$$\sup_{t \in T_I} \mathbb{E}[e^{\varepsilon_0 |X(t)|}] \leqslant M_0$$

for some M_0. This occurs, for instance, if the trajectories are almost surely confined into a fixed ball. The above bounds imply that $|\partial_t (e^{\eta \cdot X(t)})| \leqslant C_1 e^{\varepsilon_0 |X(t)|}$ for $|\eta| \leqslant \frac{\varepsilon_0}{2}$, so we may also use the exponential function as a majorant for application of the dominated convergence theorem which shows that for all t

$$\partial_t G_{\text{mom}}(\eta; t) = \mathbb{E}\big[\partial_t (e^{\eta \cdot X(t)})\big] = \mathbb{E}[\mathcal{R}(X(t),\eta) e^{\eta \cdot X(t)}] .$$

Hence, in this case all the assumptions from the first Section are satisfied, and we can conclude that the cumulants satisfy an evolution hierarchy

$$\partial_t \kappa[X(t)_I] = \sum_{J \subset I} \mathbb{E}[P_J(X(t)) :X(t)^{I \setminus J}:] = \sum_{\ell \in I} \mathbb{E}[\mathcal{P}_\ell(X(t)) :X(t)^{I \setminus \{\ell\}}:] .$$

The right hand side can be expanded using the rules of Wick polynomials which results in a polynomial of the cumulants at most of the order of m. Hence, the order of nonlinearity of the rate function for the hierarchy remains uniformly bounded, even though one usually needs to look at the full infinite hierarchy (in I) to close the evolution equations.

Hamiltonian evolution equations given by polynomial Hamiltonian functions with compact level sets often are amenable to the above treatment. For example, then one could choose initial data supported on one of these level sets, and it would follow that $X(t)$ are then almost surely bounded, even for the maximal choice of the time-interval, $T_I = \mathbb{R}$. Without such conservation laws, the analysis of which initial data distributions satisfy the assumptions is more delicate.

An explicit example of this argument is given by the discrete Nonlinear Schrödinger equation for which the real and imaginary part of the complex field form an Hamiltonian system with compact level sets. The details are worked out in Ref. [3].

2.4 Stochastic Kac model

For stochastic evolution corresponding to a Feller process, its generator L is an operator for which

$$\partial_t \mathbb{E}[f(X(t))] = \mathbb{E}[(Lf)(X(t))] , \tag{2.4.1}$$

for any observable f in the domain of the generator. If $f_\eta(X) = e^{\eta \cdot X}$ belongs to the domain of L, this allows defining

$$\mathcal{R}(X,\eta) := e^{-\eta \cdot X} (L f_\eta)(X) ,$$

and then one needs to check if this function $\mathcal{R}$ has all the properties listed above for applying the earlier general results. Since f_η does not vanish at infinity, this is not usually immediately the case, but needs a regularization argument using a mollifier. For example, one can sometimes add a factor $e^{-\frac{1}{2}\varepsilon^2 X^2}$, use the generator result, and then take $\varepsilon \to 0$ to conclude which "LF" appears on the right hand side of (2.4.1).

The stochastic Kac model was introduced by Kac in his 1956 article [4]. We discuss here the usual conventions and choose the parameters of this model based on computational convenience. Changing some of the parameters may be implemented by straightforward scalings, but some changes are not so innocuous. For example, some changes to the collision rate function are expected to speed up the generation chaos time; this will be discussion briefly at the end.

Take a number $N \geqslant 2$ of particles, which we label by $[N] := \{1, \dots, N\}$. At each moment in time, the particle with label $i \in [N]$ is associated with one-dimensional velocity $v_i \in \mathbb{R}$, and the collective velocities of all the particles form a velocity vector $v = (v_1, \dots, v_N) \in \mathbb{R}^N$. Each velocity vector constitutes a configuration of the system. Originally, Kac consider the physical case with three-dimensional velocities, $v_i \in \mathbb{R}^3$, but this would lead to more complex collision rules and indexing of the configuration vectors. Here, as in most mathematical work on this model, we focus on the one-dimensional case. As noted earlier, we do not track the positions of the particles and accordingly they are assumed to have no influence on the time-evolution of the velocities.

The configurations are updated as follows. Assume that the velocities are at some time given by $v \in \mathbb{R}^N$. Collisions occur at times determined by a Poisson process. When the Poisson clock rings, we update the velocity vector by picking, uniformly at random, two particles i and j, with $j \neq i$, and a "collision" angle $\theta \in (-\pi, \pi]$, all independent from each other. We then update the velocity vector to

$$v^* := R_{i,j}(\theta) v \in \mathbb{R}^N .$$

Here the matrix $R_{i,j}(\theta)$ is obtained as a permutation of the following $N \times N$-matrix

$$R_{1,2}(\theta) := \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \oplus I_{N-2} , \tag{2.4.2}$$

where I_n denotes the n-dimensional identity operator. The non-zero elements of the matrix $R_{i,j}(\theta)$ are given by

$$\begin{aligned} (R_{i,j}(\theta))_{\ell,\ell} &= 1, \quad \ell \neq i, j \\ (R_{i,j}(\theta))_{i,i} &= \cos(\theta) \\ (R_{i,j}(\theta))_{i,j} &= \sin(\theta) \\ (R_{i,j}(\theta))_{j,i} &= -\sin(\theta) \\ (R_{i,j}(\theta))_{j,j} &= \cos(\theta) . \end{aligned}$$

For any fixed i, j and θ, the matrix $R_{i,j}(\theta)$ is orthogonal. Thus the Euclidean norm of the velocity vectors is preserved,

$$|R_{i,j}(\theta) v| = |v| . \tag{2.4.3}$$

Therefore, we can consider the dynamics as taking place on a (hyper-)sphere $S^{N-1}(r)$, with the radius r specified by the Euclidean norm of the initial configuration.

The above specification of the dynamics on the configuration space lifts to the space observables $\phi \in C_c(S^{N-1}(r))$ in terms of the transition operator, which specifies the conditional expectation of $\phi(v^*)$ based on the previous configuration:

$$(Q_N \phi)(v) := \frac{1}{N(N-1)} \sum_{i,j=1}^{N} \mathbb{1}_{\{i \neq j\}} \int_{-\pi}^{\pi} \frac{d\theta}{2\pi} \phi(R_{i,j}(\theta) v).$$

Kac's stochastic model is now obtained by assuming that the above collision times are given by a Poisson process. As usual, we want to consider time scales at which the average rate of collisions experienced by one particle remains order one, i.e., it is bounded from below and above by some uniform constants for all relevant N. If the Poisson process has rate one, this can be achieved by speeding up time by a factor of N. For the sake of convenience, we follow the usual convention for time-scales in this model and choose the scale in which the average collision rate per particle is equal to one.

Since each $R_{i,j}(\theta)$ is orthogonal and leaves the sphere $S^{N-1}(r)$ invariant, the operator Q_N is a self-adjoint operator on the Hilbert space $L^2(S^{N-1}(r), \nu_{N,r})$, where $\nu_{N,r}$ is the uniform probability measure on $S^{N-1}(r)$; recall that $\nu_{N,r}$ is also invariant under rotations. Since we are interested in the properties of the system in the thermodynamic limit, we will pick $r = \sqrt{N}$, and ν_N stands for the corresponding uniform measure $\nu_{N,\sqrt{N}}$. This choice yields configurations for which the kinetic energy per particle, $\frac{1}{N}\sum_{i=1}^N v_i^2$, is almost surely equal to one; note that we choose each particle to have the same mass, equal to two, so that $\frac{m}{2} = 1$.

Implementing the above choices, we can now state the evolution equation for probability densities $f_t^N(v)$, $v \in S^{N-1}(\sqrt{N})$, $t \geqslant 0$. Namely, if $f_0^N \in L^2(S^{N-1}(\sqrt{N}), \nu_N)$ is non-negative and integrates to one, we can consider initial data for the stochastic Kac process given by the measure $f_0^N \nu_N$. Then, the distribution of $v(t)$, $t > 0$, is given by $f_t^N \nu_N$ where also $f_t^N \in L^2(S^{N-1}(\sqrt{N}), \nu_N)$ and it can be solved from the following Cauchy problem:

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t} f_t^N & = N(Q_N - I) f_t^N \\ f_t^N|_{t=0} & = f_0^N . \end{cases} \tag{2.4.4}$$

More details about the derivation of this equation can be found from [5, 6] and the references therein.

Then for any bounded continuous observable $\phi \in C_b(S^{N-1}(\sqrt{N}))$ we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle \phi \rangle_{F_t^N} = \langle N(Q_N - I)^* \phi \rangle_{F_t^N} = \langle N(Q_N - I)\phi \rangle_{F_t^N} ,$$

using the self-adjointness of the operator Q_N. Due to the fact that the Kac model initial data have support on $S^{N-1}(\sqrt{N})$, continuity is the only requirement here.

One simplifying feature of the above formulation of Kac model is that it has exponential moments and a moment generating function which is entire. We will consider both the velocity random variables v_i and the corresponding kinetic energies $e_i := (v_i)^2$. Since for any $v \in S^{N-1}(\sqrt{N})$

$$\sum_{i=1}^N e_i = |v|^2 = N ,$$

it follows that $0 \leqslant e_i(t) \leqslant N$ and $|v_i(t)| \leqslant \sqrt{N}$ almost surely for any $t \geqslant 0$, and thus the functions

$$\phi_1(v;\eta) := \mathrm{e}^{\eta\cdot v} , \qquad \phi_2(v;\eta) := \mathrm{e}^{\sum_{i=1}^N \eta_i (v_i)^2} ,$$

are bounded and continuous on $S^{N-1}(\sqrt{N})$ for any $\eta \in \mathbb{C}^N$.

Then, explicit computations yield

$$\mathcal{R}_{\mathrm{v}}(v,\eta) = \frac{1}{N-1}\sum_{i,j=1}^N \mathbb{1}_{\{i\neq j\}} \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \left(\mathrm{e}^{(R_{i,j}(\theta)-I)v\cdot\eta} - 1 \right) , \quad v \in \mathbb{R}^N$$

for the case of velocity cumulants, and

$$\mathcal{R}(e,\eta) = \frac{1}{N-1}\sum_{i,j=1}^N \mathbb{1}_{\{i\neq j\}} \sum_{\varnothing\neq J\subseteq I} \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \left(\mathrm{e}^{\eta_i P_\theta(v_i,v_j) + \eta_j Q_\theta(v_j,v_i)} - 1 \right) , \quad e \in \mathbb{R}_+^N$$

for the energy cumulants, where

$$\begin{aligned} P_\theta(v_i, v_j) &= -\sin(\theta)^2 v_i^2 + 2\cos(\theta)\sin(\theta) v_i v_j + \sin(\theta)^2 v_j^2 \\ Q_\theta(v_j, v_i) &= -\sin(\theta)^2 v_j^2 - 2\cos(\theta)\sin(\theta) v_i v_j + \sin(\theta)^2 v_i^2 . \end{aligned} \tag{2.4.5}$$

Although not immediately apparent, after an expansion of the exponential, one notices that all terms which an odd power of v_i have also a prefactor which is odd in θ and thus integrates to zero. Hence, the formula indeed defines a function of e_i only.

In the final lecture we focus on properties of the evolution hierarchy of the cumulants of the enery variables.

Chapter 3

Generation and propagation of chaos in the Kac model

The detailed description of these results can be found from [1]. Most of the material here is an extract from that reference.

3.1 Analysis of the cumulant hierarchy

3.1.1 Symmetric measures and partition classifiers of their cumulants

Our main result will be to control the generation of chaos in the above Kac model in the sense of asymptotic near-independence of energies of different particles. Similar results could then be derived for the particle velocities, but for the sake of brevity we do not go into this argument in detail here. The control will be given as an explicit bound for the difference between the energy cumulant and its value at equilibrium, i.e., for the uniform distribution on the energy sphere.

Although the bound will be given for fixed system, in particular, for fixed number of particles N, we aim at bounds which are valid for all large enough N. To stress this, we will carefully track the dependence of the constants in the upper bounds on N and on the initial data. Although we do not consider limiting properties of suitably prepared sequences $(F_0^N)_{N \geqslant 2}$ of initial probability measures, as is common in derivations of kinetic equations, the bounds here could be employed to make also such scaling limit statements.

Given a number of particles $N \geqslant 2$, the initial data is taken to be essentially arbitrary, apart from one technically simplifying assumption: we assume that F_0^N is a *symmetric* Radon probability measure on $S^{N-1}(\sqrt{N})$, as explained earlier. Although this assumption hides some features of the dynamics, it is not a substantial restriction. For example, given an arbitrary initial Radon probability measure but only asking for properties of expectation values of label permutation invariant observables, we can replace the initial measure by its symmetric projection without changing any of the expectation values of interest.

The symmetry of the probability measure is reflected in the joint cumulants of the energy variables as an invariance with respect to permutations of the labels of the random variables. Indeed, if $I = \{(1, i_1), \ldots, (n, i_n)\}$ is a sequence of particle labels, then for every permutation $\sigma \in \mathfrak{S}_N$, we have

$$\kappa[e(0)_I] = \kappa[e(0)_{\sigma^{-1}(I)}], \tag{3.1.1}$$

with $\sigma^{-1}(I) = \{(1, \sigma^{-1}(i_1)), \ldots (n, \sigma^{-1}(i_n))\}$.[1]

[1] It is important to keep in mind that this is distinct from the usual permutation symmetry of the arguments of joint cumulants. The label permutation invariance means that, for instance, $\kappa[e_1(0), e_2(0), e_2(0)] = \kappa[e_3(0), e_2(0), e_2(0)]$, while $\kappa[e_1(0), e_2(0), e_2(0)] = \kappa[e_2(0), e_1(0), e_2(0)]$ is something that is true even if F_0 is not a symmetric measure.

The time-evolution commutes with label permutations, so this property of label permutation invariance carries over to the time-evolved cumulants $\kappa[e(t)_I]$. Therefore, the value of $\kappa[e(t)_I]$ is determined simply by how many *different* labels there are in the sequence $I = \{(1, i_1), \ldots, (n, i_n)\}$, and how many times each such label is repeated in the sequence – no matter what the specific label is. This motivates the following definition of *partition classifiers*, which we will use to index the cumulants after the label permutation symmetry has been factored out.

Definition 3.1.1 (Partition classifier) *An n-tuple $r = (r_1, \ldots, r_n) \in \mathbb{N}_0^n$ is called a* partition classifier *(or* classifier*) of order n, in case the following two conditions are met: (i) The components of the vector sum to n, i.e. $\sum_{k=1}^n r_k = n$ and (ii) the components have been arranged in a decreasing order, so that $r_1 \geqslant r_2 \geqslant \cdots \geqslant r_n \geqslant 0$.*

Given a partition classifier r, we reserve the notation $\mathrm{len}(r)$ for the number of its nonzero components. The collection of all partition classifiers of order n is denoted by $\mathscr{C}_n$, and we often drop the trailing zeroes when denoting elements $r \in \mathscr{C}_n$; note that n is uniquely determined by the sum of the (non-zero) elements in r. The classifier corresponding to the label sequence with no repeated labels has a special role in what follows, and we will denote it by $1_n := (1, 1, \ldots, 1) \in \mathbb{N}_0^n$. Finally, $\mathscr{C}'_n := \mathscr{C}_n \backslash \{1_n\}$ denotes the set of all particle classifiers that correspond to a label sequence with at least one repeated label.

Remark 3.1.2 Each partition classifier of order n corresponds bijectively to a number theoretic partition of the natural number n – hence the name. The number of different partition classifiers therefore satisfies $|\mathscr{C}_n| \sim \frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$ as $n \to \infty$.

Any partition classifier defines a multi-index of order n, with the additional special property that the components are decreasing. We therefore need to note how to use multi-indices to label joint cumulants. Given any multi-index $r \in \mathbb{N}_0^n$, we may construct one possible compatible label sequence in the following way. First, we let $\Delta(r) = (0, r_1, r_1 + r_2, \ldots, \sum_{i=1}^n r_i)$ be a sequence of increments. After this, we define a sequence I_r by setting

$$(I_r)_i = \sum_{j=1}^n j \mathbb{1}_{\{i \in (\Delta(r)_j, \Delta(r)_{j+1}]\}}, \quad i = 1, 2, \ldots, n\,.$$

For example, with $r = (3, 1, 1, 0, 0) \in \mathscr{C}'_5$ we thus have $I_r = (1, 1, 1, 2, 3)$ which corresponds to the set $\{(1, 1), (2, 1), (3, 1), (4, 2), (5, 3)\}$ in the notation used in the beginning of this section.

With this in mind, we define cumulants corresponding to partition classifiers or any other multi-indices in the following way.

Notation 3.1.3 *Let $r \in \mathbb{N}_0^n$ be a multi-index. The joint cumulant corresponding to this multi-index is defined as*

$$\kappa[e_r] := \kappa[e_{I_r}]. \tag{3.1.2}$$

If r is a partition classifier of order n and $s \in \mathbb{N}_0^n$ is a multi-index, we say that $r \sim s$, in case there exists a permutation $\sigma \in \mathfrak{S}_n$ such that $r = \sigma(s)$

In the following computations, we will encounter situations where certain repetitions of a random variable are removed when computing a Wick polynomial or cumulant corresponding to a collection of random variables (determined by a sequence, multi-index or a partition classifier, dependent on the context). This motivates the following notation:

Notation 3.1.4 (Removing and adding multi-indices) *Suppose that $r \in \mathbb{N}_0^n$ is a multi-index (possibly a partition classifier). If $r_i \geqslant \ell$, we denote by*

$$r - (\ell \times i) \tag{3.1.3}$$

the multi-index whose components are given by $(r-(\ell\times i))_k = r_k$ *for* $k \neq i$ *and* $(r-(\ell\times i))_i = r_i-\ell$. *Note that even if* r *is a partition classifier, this is usually not a partition classifier.*

Similarly, given any $r \in \mathbb{N}_0^n$,

$$r + (\ell \times i) \tag{3.1.4}$$

is the multi-index obtained by increasing the i*th component of* r *by* ℓ *while keeping other components the same.*

The following notation will become relevant later, when we try to isolate the main contribution in the time-evolution of cumulants.

Definition 3.1.5 *Let* $r \in \mathscr{C}_n$ *be a partition classifier. We define* $\mathtt{break}_\ell(r)$ *to be the set that consist of those* $q \in \mathscr{C}_n$, *for which there exist* $\ell_1, \ell_2 \in [n]$, $\ell_1 \neq \ell_2$, *such that* $r_\ell = q_{\ell_1} + q_{\ell_2}$; *and in addition for which the sequence* r' *obtained from* r *by removing the components* ℓ, n *and shifting, is identical to the sequence* q' *obtained from* q *by removing the components* ℓ_1, ℓ_2 *and shifting.*

The set $\mathtt{break}(r)$ *is the union* $\cup_{\ell=1}^n \mathtt{break}_\ell(r)$.

Example 3.1.6 Let $r = (4,0,0,0)$. Now $(3,1,0,0), (2,2,0,0) \in \mathtt{break}_1(r)$, whereas $(2,1,1,0) \notin \mathtt{break}_1(r)$. Also, $(3,1,0,0) \notin \mathtt{break}_2(r)$.

Definition 3.1.7 *Fix* $n \in \mathbb{N}$. *Define a relation* $\succ$ *on* $\mathscr{C}_n$ *by setting* $s \succ r$, *if and only if* $r \in \mathtt{break}(s)$.

Then, set $\mathscr{C}_{n,1} = \{1_n\} = \{(1,\ldots,1)\}$, *and define iteratively using the above relation for any* i *with* $1 < i \leqslant n$,

$$\mathscr{C}_{n,i} = \{s \in \mathscr{C}_n\colon \text{ there exists } r \in \mathscr{C}_{n,i-1} \text{ such that } s \succ r\}. \tag{3.1.5}$$

Remark 3.1.8 The above relation can be extended to be a strict partial order on $\mathscr{C}_n$, but this fact will not be used in what follows. Clearly $\{\mathscr{C}_{n,i}\}_{i=1}^n$ partitions $\mathscr{C}_n$, and $s \in \mathscr{C}_{n,k}$ precisely when it is a partition classifier that has exactly $n-(k-1)$ non-zero components. Figure 3.1 illustrates these sets and their relation to $\mathscr{C}_n$, when $n = 5$

3.1.2 Main Results: controlling chaos and equilibriation via the cumulant hierarchy

We will quantify the chaoticity of the energy cumulants by studying their norm in the following normed spaces, which we will call the α-chaos spaces.

Definition 3.1.9 (α-chaos space) *Let* $\alpha \in [0,1]$. *We define the normed space* $(X^\alpha_{n,N}, \|\cdot\|_{\alpha,n,N})$ *to consists of vectors* $(\kappa_r)_{r\in\mathscr{C}'_n} \subset \mathbb{R}^{\mathscr{C}'_n}$, *together with the weighted supremum norm*

$$\|\kappa\|_\alpha = \|\kappa\|_{\alpha,n,N} = \sup_{r\in\mathscr{C}'_n} |\kappa_r|\,(N-1)^{\alpha(\mathrm{len}(r)-1)}. \tag{3.1.6}$$

We will mainly use these norms with $0 < \alpha < 1$, but the extremal cases will appear in some of our estimates. The norms are increasing functions of α: For any $r \in \mathscr{C}'_n$ and $0 \leqslant \alpha \leqslant \alpha' \leqslant 1$ we have $\|\kappa\|_{\alpha,n,N} \leqslant \|\kappa\|_{\alpha',n,N}$.

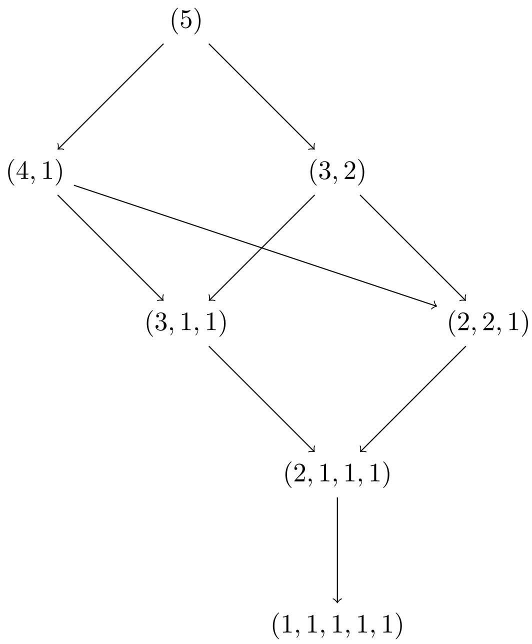

Figure 3.1: Relations between partition classifiers in $\mathscr{C}_5$. In the graph, the nodes are in correspondence with elements in $\mathscr{C}_5$, and there is a directed edge $s \to r$ in the graph precisely when $r \in \operatorname{break}(s)$.

Remark 3.1.10 We have excluded the non-repeated label sequence $1_n = (1, 1, \ldots, 1)$ from the definition of these α-chaos spaces. The reason for this is that the completely non-repeated cumulants will be treated separately, and thus will show up as a source term to the cumulant hierarchy. However, the conservation law implies that if

$$\sup_{r \in \mathscr{C}'_n} |\kappa[e_r]| \, (N-1)^{\alpha(\operatorname{len}(r)-1)} \leqslant C, \tag{3.1.7}$$

then

$$\sup_{r \in \mathscr{C}_n} |\kappa[e_r]| \, (N-1)^{\alpha(\operatorname{len}(r)-1)} \leqslant 4C \tag{3.1.8}$$

whenever $n \leqslant N/2$ (see the proof of Proposition 3.1.16).

Therefore, assumptions pertaining to the joint cumulants indexed by $\mathscr{C}'_n$ will imply similar assumptions for all joint cumulants indexed by $\mathscr{C}_n$ with a slightly larger constant.

If the above norm of a vector $(\kappa_r)_{r \in \mathscr{C}'_n}$ is bounded by a constant C_n, it follows that the individual components are bounded by

$$|\kappa_r| \leqslant \frac{C_n}{(N-1)^{\alpha(\operatorname{len}(r)-1)}}. \tag{3.1.9}$$

Since the elements of $X^{\alpha}_{n,N}$ will be given by the nth order energy cumulants with some repetition, so that $\kappa_r = \kappa^{n,N}[e_r]$, it follows that such uniform bounds imply

$$\limsup_{N \to \infty} \left|\kappa^{n,N}[e_r]\right| \leqslant \mathbb{1}_{\{\operatorname{len}(r)=1\}} C_n, \tag{3.1.10}$$

corresponding to the fact that only cumulants involving one particle can survive to the many-particle limit of any sequence of probability measures whose cumulants have a norm uniformly bounded in N.

When studying the time-evolution of the energy cumulants in the Kac model, we will try to find good bounds for their $X^\alpha_{n,N}$ norms. To this end, it should be noted that the operator norm of an operator $A\colon X^\alpha_{n,N} \to X^\alpha_{n,N}$ satisfies

$$\|A\|_\alpha \leqslant \sup_r \sum_{r'\in\mathscr{C}'_n} |A_{r,r'}|\,(N-1)^{\alpha(\operatorname{len}(r)-\operatorname{len}(r'))}, \tag{3.1.11}$$

a fact that we will need in order to control the linear part of the time-evolution.

We are now ready to state the main theorem of this article. Our results concern the generation of chaos and convergence towards the equilibrium of the energy cumulant hierarchy. The initial data is allowed to have different "orders of chaoticity", and this will affect the corresponding results.

Assumption 3.1.11 (Chaos bounds) *Let $c \geqslant 0$, $N \geqslant 2$, and $n^* \geqslant 1$ be given. Define $\gamma_n = c(n-1)$ for $n \geqslant 1$. We say that $B > 0$ is a constant for the* chaos bound *of type γ_n up to order n^* if for every $n \in [n^*]$ the joint cumulants are bounded in the α-norm as*

$$\|\kappa^{n,N}_0\|_\alpha \leqslant B^{n-1}(n-1)!N^{\gamma_n}. \tag{3.1.12}$$

A sequence of initial data $(F^N_0)_{N\geqslant 2}$ on $S^{N-1}(\sqrt{N})$ is said to satisfy the chaos bound of type γ_n with a constant B if the same B can be used for every N.

Remark 3.1.12 Of course, for any fixed N, c and n^* such a constant B can be found. However, B might need be very large, in particular, it could increase as a power law in N. For the results below to be useful, B should be order one even if N is large. This can always be achieved for chaos bounds with a constant $c = 1$ in the Assumption 3.1.11 by increasing N if needed, at least when considering joint energy cumulants in a symmetric state up to any finite order. Indeed, Proposition 1.3.1 shows that the energy joint cumulants will always satisfy a bound

$$\|\kappa^{n,N}\|_\alpha \leqslant 4^{n-1}(n-1)!N^{n-1}, \quad n \leqslant N/2+1\,. \tag{3.1.13}$$

Theorem 3.1.13 (Generation of chaotic bounds) *Let $c \geqslant 0$ and $\alpha \in (0,1)$. Let $n^* \in \mathbb{N}$ be a maximal order of cumulants. Then there exists a $N_0 = N_0(n^*,\alpha,c) \geqslant 2$ such that the following result holds.*

Consider some fixed $N \geqslant N_0$ and some symmetric initial data F^N_0 on $S^{N-1}(\sqrt{N})$. Denote the corresponding joint cumulants at order $n \in [n^*]$ and at time $t \geqslant 0$ by $\kappa^n_t[e_r] = \kappa[e_{I_r}(t)]$. Let $B \geqslant 1$ be a constant such that the initial values of the cumulants satisfy

$$\max_{1\leqslant n\leqslant n^*}\max_{r\in\mathscr{C}_n}\left(|\kappa^n_0[e_r]|\,N^{\alpha(\operatorname{len}(r)-1)-c(n-1)}/(n-1)!\right)^{1/n} \leqslant B. \tag{3.1.14}$$

In other words, assume that the initial data satisfies assumption 3.1.11. Then there exists a constant C, depending only on B, such that for all $n \leqslant n^*$, the time-evolved cumulants satisfy

$$\|\kappa^{n,N}_t\|_\alpha \leqslant C^{n^2}n!\left(N^{\gamma_n}\mathrm{e}^{-\frac{1}{4}t}+1\right). \tag{3.1.15}$$

We should point out that we have not tried to optimize the bounds in n, the order of the cumulants. Most likely their asymptotic increase is much less than $C^{n^2}n!$, perhaps even as good as $C^{n-1}n!$.

Remark 3.1.14 Under conditions that guarantee that odd joint cumulants involving velocities of particles vanish, one could use Möbius inversion techniques to recover chaotic bounds for the joint cumulants of the *velocities* of the particles. For example, Theorem 2 in Bauer et al. [8] can be used to express $\kappa^{2n,N}_t(v_I)$ in terms of certain energy cumulants $\kappa^{n,N}_t(e_J)$, and it remains to check what kinds of energy cumulants appear on the right hand side of the identity. Some combinatorial losses seem inevitable, however.

Theorem 3.1.15 (Convergence to stationarity) *Let $c \geqslant 0$ and $\alpha \in (0,1)$. Let $\gamma_n = c(n-1)$ and fix $n^* \in \mathbb{N}$ as the maximal order of cumulants. Then there exists $N_0 = N_0(n^*, \alpha, c) \geqslant 2$ such that the following result holds.*

Consider some fixed $N \geqslant N_0$ and some symmetric initial data F_0^N on $S^{N-1}(\sqrt{N})$. Denote the corresponding joint cumulants at order $n \in [n^*]$ and at time $t \geqslant 0$ by $\kappa_t^n[e_r] = \kappa[e_{I_r}(t)]$ and the stationary cumulants at order n, those corresponding to the uniform probability distribution on $S^{N-1}(\sqrt{N})$, by $\bar{\kappa}^n[e_r] = \bar{\kappa}^n[e_{I_r}]$. Let $B \geqslant 1$ be a constant such that the initial values of the cumulants satisfy

$$\max_{1\leqslant n\leqslant n^*} \max_{r\in\mathscr{C}_n} \left(|\kappa_0^n[e_r]|\, N^{\alpha(\mathrm{len}(r)-1)-c(n-1)}/(n-1)!\right)^{1/n} \leqslant B. \tag{3.1.16}$$

Then there exists a constant C, depending only on B, such that the time-evolved energy cumulants satisfy the following bound for all $n \leqslant n^*$ and $t \geqslant 0$

$$\|\kappa_t^{n,N} - \bar{\kappa}^{n,N}\|_\alpha \leqslant n! C^{n^2} N^{\gamma_n} \mathrm{e}^{-\frac{t}{4}}\,. \tag{3.1.17}$$

We have excluded the completely non-repeated cumulants from the α-chaos spaces, since we will treat them as source terms. Fortunately for us, their time-evolution solves a closed hierarchy, so we can first describe their time-evolution and then take it as known when proving results pertaining to the time-evolution of the other cumulants. The following stronger results holds for the non-repeated energy cumulants.

Proposition 3.1.16 (Generation of chaos for non-repeated cumulants) *Let $\alpha \in (0,1)$ and choose some $c \geqslant 0$ and a maximal order of cumulants $n^* \in \mathbb{N}$. Then there is $N_0 = N_0(n^*, \alpha, c) \geqslant 2$ such that the following result holds.*

Consider some fixed $N \geqslant N_0$ and some symmetric initial data. Denote the corresponding non-repeated cumulants at order $n \in [N]$ and time $t \geqslant 0$ by $\kappa_t^{n,nr} := \kappa[e_1(t), e_2(t), \ldots, e_n(t)]$. Let $B \geqslant 1$ be a constant such that the initial values of the non-repeated cumulants satisfy

$$\max_{1\leqslant n\leqslant n^*} \left(|\kappa_0^{n,nr}| N^{(\alpha-c)(n-1)}/(n-1)!\right)^{\frac{1}{n}} \leqslant B\,. \tag{3.1.18}$$

Then there is a constant $C \geqslant 1$, which depends only on B, such that the time-evolved non-repeated cumulants satisfy

$$N^{\alpha(n-1)}\, |\kappa_t^{n,nr}| \leqslant C^n (n-1)! (\mathrm{e}^{-\frac{n}{4}t} N^{c(n-1)} + 1)\,, \tag{3.1.19}$$

for all $n \in [1, n^*]$ and $t \geqslant 0$.

3.2 Implications for the kinetic theory of the Kac model

Having fixed the number of particles N and a symmetric initial data F_0^N we obtain the measures F_t^N, $t \geqslant 0$, from the Kac process. Suppose we first wait a time $t_0 \geqslant 0$ and then use the first marginal of $F_{t_0}^N$ to define an initial measure μ_0 for the Boltzmann–Kac equation for which $T = t - t_0 \geqslant 0$ serves as the time parameter. We can then use these measures to define the symmetric product measures $\tilde{F}_T^N := \otimes_{i=1}^N \mu_T$ on $\mathbb{R}^N$, and ask the question: How close are the energy cumulants of $\tilde{F}_T^N$ to those of $F_{t_0+T}^N$? The following result shows that then up to any finite order and for all large enough systems, the two cumulants remain very close to each other, if the Boltzmann equation is started with initial data from a chaotic state, i.e., with suitable large enough t_0. As shown in the theorem, any t_0 for which $N^{c(n^*-1)}\mathrm{e}^{-\frac{1}{4}t_0} \leqslant 1$, is sufficient for cumulants of order $n \leqslant n^*$. If $c = 0$, we can use $t_0 = 0$; hence the name "chaotic initial data" for these. If $c > 0$, we can set $t_0 = 4c(n^* - 1)\ln N$ which is $O(\ln N)$ as $N \to \infty$.

Without going into details, let us also point out that these estimates imply concrete upper bounds also for expectation values of much larger class of observables g. For example, if g is any function which can be approximated by polynomials of energies of a fixed number of particles, we would still have $\mathbb{E}_{F^N_{t_0+T}}[g] \approx \mathbb{E}_{\tilde{F}^N_T}[g]$ if N is large enough, with an explicit bound for the error obtained by applying moments-to-cumulants formula to the expectation of the polynomial approximation.

Theorem 3.2.1 (Accuracy of the "Boltzmann–Kac" hierarchy) *Let $c \geqslant 0$ and $\alpha \in (0,1)$. Let $n^* \in \mathbb{N}$ be a maximal order of cumulants. Consider $N \geqslant N_0$ and some symmetric initial data F^N_0 on $S^{N-1}(\sqrt{N})$, and suppose $N_0 \geqslant 2$ and $B \geqslant 1$ are constants for which Theorem 3.1.13 holds. Denote the corresponding joint cumulants at order $n \in [n^*]$ and at time $t \geqslant 0$ by $\kappa^n_t[e_r] = \kappa[e_{I_r}(t)]$.*

Pick some $t_0 \geqslant 0$ for which $N^{c(n^*-1)}\mathrm{e}^{-\frac{1}{4}t_0} \leqslant 1$. Let μ_0 be the first first marginal of $F^N_{t_0}$, and let μ_T denote the corresponding weak solution to the Boltzmann–Kac equation. Denote the cumulants of the product measure $\otimes^N_{i=1}\mu_T$ on $\mathbb{R}^N$ by $\tilde{\kappa}^{n,N}_T[e_r]$. Then, for all $T \geqslant 0$, we have

$$|\kappa^{n,N}_{t_0+T}[e_r] - \tilde{\kappa}^{n,N}_T[e_r]| \leqslant 2(N-1)^{-\alpha}C^{n^2}n! = O(N^{-\alpha}). \tag{3.2.1}$$

3.3 Some details of the proof

Most of the structure of the proof is given on the slides of the last lecture, and all the details can be found from [1]. However, to illustrate the structure and highlight one trick which is particular to the energy cumulants of the Kac model, let us look look at closure of the fully non-repeating hierarchy separately here.

If $s = (s_1, \ldots, s_n)$ is a partition classifier, then the earlier derived cumulant hierarchy may be computed as

$$\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t[e_s] = T^{\mathrm{break}}_t(s) + T^{\mathrm{fuse}}_t(s) + T^{\mathrm{ex}}_t(s), \tag{3.3.1}$$

where we first set $L = L(s) := \mathrm{len}(s) \leqslant n$ and then define

$$\begin{aligned} T^{\mathrm{break}}_t(s) = {} & \frac{1}{N-1}\sum_{i=1}^{L}\sum_{j=L+1}^{N}\sum_{\ell=1}^{s_i}\binom{s_i}{\ell}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{2\pi}\langle :e_{s-(\ell\times i)}:P_\theta(v_i,v_j)^\ell\rangle \\ & + \frac{1}{N-1}\sum_{j=1}^{L}\sum_{i=L+1}^{N}\sum_{\ell=1}^{s_j}\binom{s_j}{\ell}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{2\pi}\langle :e_{s-(\ell\times j)}:Q_\theta(v_j,v_i)^\ell\rangle, \end{aligned} \tag{3.3.2}$$

$$\begin{aligned} T^{\mathrm{fuse}}_t(s) = {} & \frac{1}{N-1}\sum_{i=1}^{L}\sum_{j=1}^{L}\mathbb{1}_{\{j\neq i\}}\sum_{\ell=1}^{s_i}\binom{s_i}{\ell}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{2\pi}\langle :e_{s-(\ell\times i)}:P_\theta(v_i,v_j)^\ell\rangle \\ & + \frac{1}{N-1}\sum_{j=1}^{L}\sum_{i=1}^{L}\mathbb{1}_{\{j\neq i\}}\sum_{\ell=1}^{s_j}\binom{s_j}{\ell}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{2\pi}\langle :e_{s-(\ell\times j)}:Q_\theta(v_j,v_i)^\ell\rangle, \end{aligned} \tag{3.3.3}$$

$$T^{\mathrm{ex}}_t(s) = \frac{1}{N-1}\sum_{i,j=1}^{L}\mathbb{1}_{\{i\neq j\}}\sum_{\ell=1}^{s_i}\sum_{\ell'=1}^{s_j}\binom{s_i}{\ell}\binom{s_j}{\ell'}\int_{-\pi}^{\pi}\frac{\mathrm{d}\theta}{2\pi}(-1)^{\ell'}\langle :e_{s-(\ell\times i)-(\ell'\times j)}:P_\theta(v_i,v_j)^{\ell+\ell'}\rangle. \tag{3.3.4}$$

Note that for any $s \in \mathscr{C}_n$ we have $s_i \geqslant 1$ if $i \leqslant L(s)$, and $s_i = 0$ if $L(s) < i \leqslant n$. The names of these terms stem from the type of partition classifiers that contribute to the linear operator arising from each term: for the first term, the contribution comes from classifiers where one of the components has been "broken" into two pieces, for the second term, the contribution comes from "fused" components, and for the third term, the contribution comes for classifiers where some components have "exchanged" particles with each other.

To compute the contributions from each term on the evolution of the cumulants, we recall the explicit form of P_θ and Q_θ given in (2.4.5). For any $m \in \mathbb{N}$, we can employ the multinomial theorem and conclude

$$P_\theta(v_i, v_j)^m = \sum_{a,b,c=0}^{m} \mathbb{1}_{\{a+b+c=m\}} \frac{m!}{a!b!c!}(-1)^a(\sin(\theta))^{2a}e_i^a 2^b(\cos(\theta))^b(\sin(\theta))^b v_i^b v_j^b(\sin(\theta))^{2c}e_j^c \,.$$

Therefore,

$$\int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} P_\theta(v_i, v_j)^m = \sum_{k,h=0}^{m} \mathbb{1}_{\{k+h\leqslant m,\, h \text{ even}\}} \frac{m!}{k!h!(m-k-h)!}(-1)^k 2^h e_i^{k+\frac{h}{2}} e_j^{m-k-\frac{h}{2}} I_{m-\frac{h}{2},\frac{h}{2}} \,, \tag{3.3.5}$$

where we have defined

$$I_{a,b} := \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} (\sin(\theta)^2)^a (\cos(\theta)^2)^b \;, \qquad a, b \in \mathbb{N}_0 \,. \tag{3.3.6}$$

Clearly, $0 \leqslant I_{a,b} \leqslant 1$, and these integrals can also be evaluated explicitly in terms of factorials. Similar computation for $Q_\theta(v_j, v_i)$ yields

$$\int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} Q_\theta(v_j, v_i)^m = \sum_{k,h=0}^{m} \mathbb{1}_{\{k+h\leqslant m,\, h \text{ even}\}} \frac{m!}{k!h!(m-k-h)!}(-1)^k 2^h e_j^{k+\frac{h}{2}} e_i^{m-k-\frac{h}{2}} I_{m-\frac{h}{2},\frac{h}{2}} \,. \tag{3.3.7}$$

Let us then look at the fully non-repeating case which has $s = 1_n = (1, 1, \ldots, 1)$. In the integral over the random angle θ, any odd function of θ evaluates to zero. In particular, this is the case for the trigonometric term $2\cos(\theta)\sin(\theta)$, and thus the first term in the sum may be simplified to

$$\begin{aligned} &\frac{1}{N-1} \sum_{i,j=1}^{N} \mathbb{1}_{\{i\neq j\}} \sum_{k=1}^{n} \mathbb{1}_{\{k=i\}} \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \langle :e_{1_n-(1\times k)}: P_\theta(v_i, v_j)\rangle \\ &= \frac{1}{N-1} \sum_{i=1}^{n}\sum_{j=1}^{N} \mathbb{1}_{\{i\neq j\}} \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \langle :e_{1_n-(1\times i)}: P_\theta(v_i, v_j)\rangle \\ &= \frac{1}{N-1} \sum_{i=1}^{n}\sum_{j=1}^{N} \mathbb{1}_{\{i\neq j\}} \int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \left((-\sin(\theta)^2)\langle :e_{1_n-(1\times i)}: e_i\rangle + \sin(\theta)^2 \langle :e_{1_n-(1\times i)}: e_j\rangle\right) \,. \end{aligned} \tag{3.3.8}$$

In the remaining sum, we distinguish between two different cases. If $n < j \leqslant N$, we can use the symmetry to conclude that $\langle :e_{1_n-(1\times i)}: e_j\rangle = \kappa_t[e_{1_n}]$ since we can swap the labels for i and j. Since also $\langle :e_{1_n-(1\times i)}: e_i\rangle = \kappa_t[e_{1_n}]$, these values of j will not contribute to the sum.

If $1 \leqslant j \leqslant n$, $j \neq i$, we will get a cumulant involving a repetition of the index j. We use the symmetry to reorder the label sequence so that i is moved to the end (position n) and j is moved to the beginning (position 1). This yields $\langle :e_{1_n-(1\times i)}: e_j\rangle = \kappa_t[e_1, e_{1_{n-1}}]$.

We recall that $\int_{-\pi}^{\pi} \frac{\mathrm{d}\theta}{2\pi} \sin(\theta)^2 = \frac{1}{2}$ which allows to conclude that the right hand side of (3.3.8) is equal to

$$-\frac{n(n-1)}{2(N-1)} \left(\kappa_t[e_{1_n}] - \kappa_t[e_1, e_{1_{n-1}}]\right) \,. \tag{3.3.9}$$

Similar evaluation of the remaining terms yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t^{n,\mathrm{nr}} = -\frac{3n(n-1)}{4(N-1)}\kappa_t^{n,\mathrm{nr}} + \frac{n(n-1)}{4(N-1)}\kappa_t[e_1, e_1, e_2, \ldots, e_{n-1}] + \mathcal{N}_{<n}(t). \tag{3.3.10}$$

As indicated in the introduction, this evolution equation is linear for the order n cumulants and has a non-linear source term involving the lower order cumulants. However, we can simplify the linear part further, by using the conservation law.

Namely, by the permutation symmetry of F_t^N, we have

$$\begin{aligned}\kappa_t[N, e_2, \ldots, e_n] &= \sum_{i=1}^{N} \kappa_t[e_i, e_2, \ldots, e_n] \\ &= (n-1)\kappa_t[e_1, e_1, \ldots, e_{n-1}] + (N-(n-1))\kappa_t^{n,\mathrm{nr}}. \end{aligned} \tag{3.3.11}$$

Any joint cumulant involving a constant random variable and some other random variables vanishes. In particular, $\kappa_t[N, e_2, \ldots, e_n] = 0$ above. Therefore, we have the following relationship between the completely non-repeated energy cumulants and energy cumulants with one repeated particle label:

$$\kappa_t[e_1, e_1, \ldots, e_{n-1}] = -\frac{N-(n-1)}{n-1}\kappa_t^{n,\mathrm{nr}}\,. \tag{3.3.12}$$

Plugging this into (3.3.10), we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\kappa_t^{n,\mathrm{nr}} = -D_{n,N}\kappa_t^{n,\mathrm{nr}} + \mathcal{N}_{<n}(t), \quad 2 \leqslant n \leqslant N\,, \tag{3.3.13}$$

where the dissipation constant is given by

$$D_{n,N} = \frac{n}{4} + \frac{2n^2-n}{4(N-1)}\,, \quad 2 \leqslant n \leqslant N \tag{3.3.14}$$

To solve the system of equations (3.3.13), we define a family of new dynamic variables by setting

$$h_n^\alpha(t) = h_n^{\alpha,N}(t) := (-1)^{n-1}\frac{(N-1)^{\alpha(n-1)}}{(n-1)!}\kappa_t^{n,\mathrm{nr}}. \tag{3.3.15}$$

Our goal is to prove (3.1.19) which is equivalent to the statement

$$|h_n^\alpha(t)| \leqslant C^{n-1}(\mathrm{e}^{-\frac{n}{4}t}N^{\gamma_n} + 1)\,, \tag{3.3.16}$$

in the new variables.

Evaluating also the nonlinear terms explicitly shows that

$$\frac{\mathrm{d}}{\mathrm{d}t}h_n^\alpha(t) = -D_{n,N}h_n^\alpha(t) + \frac{n}{2(N-1)^{1-\alpha}}\sum_{m=1}^{n-1} h_m^\alpha(t)h_{n-m}^\alpha(t). \tag{3.3.17}$$

This results implies the Duhamel formula

$$h_n^\alpha(t) = \mathrm{e}^{-D_{n,N}t}h_n^\alpha(0) + \frac{n}{2(N-1)^{1-\alpha}}\sum_{m=1}^{n-1}\int_0^t \mathrm{e}^{-D_{n,N}(t-s)}h_m^\alpha(s)h_{n-m}^\alpha(s)\mathrm{d}s\,.$$

The final part of the proof consists of using the "goal" (3.3.16) as an induction assumption, then checking that the constant C can be adjusted so that the assumption propagates.

Bibliography

[1] Jani Lukkarinen and Aleksis Vuoksenmaa. Generation of chaos in the cumulant hierarchy of the stochastic Kac model, 2024. Arxiv preprint, `https://arxiv.org/abs/2407.17068`.

[2] Walter Rudin. *Real and Complex Analysis.* McGraw-Hill, 3rd edition, 1987.

[3] Jani Lukkarinen and Matteo Marcozzi. Wick polynomials and time-evolution of cumulants. *J. Math. Phys.*, 57(8):083301:1–27, 2016.

[4] Mark Kac. Foundations of Kinetic Theory. *Proceedings of The third Berkeley symposium on mathematical statistics and probability*, 3:171–197, 1956.

[5] Eric A. Carlen, Maria C. Carvalho, Jonathan Le Roux, Michael Loss, and Cédric Villani. Entropy and chaos in the Kac model. *Kinetic & Related Models*, 3(1):85–122, 2010.

[6] Stéphane Mischler and Clément Mouhot. Kac's program in kinetic theory. *Inventiones mathematicae*, 193(1):1–147, July 2013.

[7] Edwin Hewitt and Leonard J. Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80(2):470–501, 1955.

[8] Michel Bauer, Denis Bernard, Philippe Biane, and Ludwig Hruza. Bernoulli Variables, Classical Exclusion Processes and Free Probability. *Annales Henri Poincaré*, 25(1):125–172, January 2024.