

CENTRAL LIMIT THEOREMS FOR GROUP ACTIONS WHICH ARE EXPONENTIALLY MIXING OF ALL ORDERS

MICHAEL BJÖRKLUND AND ALEXANDER GORODNIK

ABSTRACT. Let H be a locally compact group equipped with a proper left-invariant metric d and a left-invariant Haar measure m_H that has sub-exponential volume growth. We consider a measure-preserving action of H on a probability space (X, μ) , and show that if the action is exponentially mixing of all orders (with respect to the metric d) for test functions belonging to a given H -invariant sub-algebra \mathcal{A} of $L^\infty(X, \mu)$, then for every right Følner sequence (F_T) in H with $m_H(F_T) \rightarrow \infty$ and $f \in \mathcal{A}$,

$$\frac{1}{m_H(F_T)^{1/2}} \int_{F_T} (f \circ h^{-1} - \mu(f)) dm_H(h) \implies N(0, \sigma_f^2) \quad \text{as } T \rightarrow \infty, \quad (\text{CLT})$$

where \implies denotes convergence in the sense of distributions, and $N(0, \sigma_f^2)$ denotes the Gaussian distribution on the real line with mean zero and variance

$$\sigma_f^2 = \int_H (\langle f, f \circ h^{-1} \rangle - \mu(f)^2) dm_H(h).$$

In particular, we deduce that if G is a connected semisimple Lie group with finite center, Γ is a lattice in G with strong spectral gap, H is a closed subgroup of a Cartan group of G , and F_T 's are centered Euclidean balls in H of increasing radii, then (CLT) holds for every smooth function f with compact support on the quotient space $X = \Gamma \backslash G$. As another application of our techniques, we also show that if $u(t)$ is a non-trivial *unipotent* one-parameter subgroup of G acting $\Gamma \backslash G$, and (λ_k) is a lacunary sequence in \mathbb{R}_+ , then for all smooth, compactly supported functions f on $\Gamma \backslash G$,

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N (f \circ u(\lambda_k) - \mu(f)) \implies N(0, \|f\|_{L^2}^2) \quad \text{as } N \rightarrow \infty.$$

1. INTRODUCTION

1.1. Central limit theorems in dynamics

One of the fundamental problem in the theory of dynamical systems is to understand whether a sequence of observables of a chaotic dynamical system computed along generic orbits behaves similarly to a sequence of independent identically distributed random variables. More precisely, given a measure-preserving transformation T of a probability space (X, μ) and a measurable function f on X , one is interested in analysing statistical properties of the sequence

$$f(Tx), f(T^2x), \dots, f(T^kx), \dots \quad (1.1)$$

where $x \in X$ is distributed according to the measure μ . One says that ‘‘Central Limit Theorem’’ (CLT) holds if there exists $\sigma_f \geq 0$ such that

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N (f \circ T^k - \mu(f)) \implies N(0, \sigma_f^2) \quad \text{as } N \rightarrow \infty, \quad (1.2)$$

where $N(0, \sigma_f^2)$ denotes the Gaussian distribution with variance σ_f^2 , and \implies denotes convergence in the sense of distributions. Explicitly this means that for every interval (a, b) ,

$$\mu \left(\left\{ x \in X : a < \frac{1}{\sqrt{N}} \sum_{k=1}^N (f(T^k x) - \mu(f)) < b \right\} \right) \longrightarrow \frac{1}{\sqrt{2\pi\sigma_f}} \int_a^b e^{-u^2/(2\sigma_f)} du$$

as $N \rightarrow \infty$. Due to deterministic nature of the sequence (1.1), it is known that CLT cannot hold for general functions in $L^1(X, \mu)$, and one is interested in finding a “large” subspace \mathcal{A} of functions satisfying (1.2). In applications, the space X is often assumed to be a finite-volume Riemannian manifold, $T : X \rightarrow X$ a smooth map which preserves the volume measure on X , and \mathcal{A} is some linear subspace of continuous functions with prescribed regularity (Hölder, \mathcal{C}^r for some r , etc.). Starting with the pioneering work of Sinai [28] who established the Central Limit Theorem for geodesic flows on compact manifolds with constant negative curvature, this problem has been extensively studied for transformations satisfying some hyperbolicity assumptions [24, 27, 5, 19, 12, 23, 6, 25, 32, 20, 21, 10, 26, 13, 17]. We refer to [11, 12, 18, 22, 30] for surveys of this area of research.

More generally, we consider a measure-preserving action of a group H on a probability space (X, μ) . Given a measurable function f on X , we obtain a collection of observables

$$\{f(h^{-1}x) : h \in H\} \quad \text{with } x \in X. \quad (1.3)$$

The aim of this paper is to establish a general Central Limit Theorem for the averages

$$\frac{1}{|F_T|^{1/2}} \sum_{h \in F_T} (f \circ h^{-1} - \mu(f)), \quad F_T \subset H,$$

for actions of higher-dimensional (possibly non-commutative) groups. Previously, CLT was established in [7, 8, 9] for actions of the group $H = (\mathbb{Z}^d, +)$ by automorphisms of compact abelian groups. Our result holds for general actions that are “sufficiently chaotic”, which is manifested by quantitative estimates on the higher-order correlations

$$\int_X f_1(h_1^{-1}x) \cdots f_r(h_r^{-1}x) d\mu(x)$$

for $h_1, \dots, h_r \in H$ and $f_1, \dots, f_r : X \rightarrow \mathbb{R}$.

The multi-parameter averages of this type naturally arise in number-theoretic problems. In [3] we apply the developed techniques to analyse discrepancy of distribution for values of products of linear forms.

1.2. The main result

Let H be a locally compact group equipped with a left-invariant Haar measure m_H and a proper left-invariant metric d . We say that (H, d) has *sub-exponential growth* if

$$\overline{\lim}_{r \rightarrow \infty} \frac{\log m_H(B_d(e, r))}{r} = 0, \quad (1.4)$$

where $B_d(e, r)$ denotes the d -ball around the identity element e in G of radius r . It is not hard to show that every locally compact second countable abelian group has sub-exponential growth; a bit more work is required to show that every locally compact second countable nilpotent group has sub-exponential growth. Furthermore, it is well-known that groups of sub-exponential

growth are amenable, and thus possess right Følner sequences. We recall that a sequence (F_T) of compact subsets with non-empty interiors in H is *right Følner* if

$$\lim_{T \rightarrow \infty} \frac{m_H(F_T \Delta F_T h)}{m_H(F_T)} = 0, \quad \text{for all } h \in H,$$

where Δ denotes the symmetric difference of sets. It is easy to show that if $H = (\mathbb{R}^n, +)$, then any sequence of Euclidean balls of increasing radii forms a Følner sequence in H . More generally, balls in groups of sub-exponential growth also form Følner sequences.

Suppose that the group H acts jointly measurably by measure-preserving maps on a probability measure space (X, μ) . We shall assume that there exists an H -invariant sub-algebra \mathcal{A} of $L^\infty(X, \mu)$ and a family $\mathcal{N} = (N_s)$ of semi-norms on \mathcal{A} , which satisfy some technical conditions spelled out in Section 2 (see (2.3)–(2.6) below), such that the H -action is *exponentially mixing of all orders* with respect to d and $(\mathcal{A}, \mathcal{N})$ (in the sense of Definition 2.1 below). Roughly speaking, this property requires that

$$\int_X f_1(h_1^{-1}x) \cdots f_r(h_r^{-1}x) d\mu(x) = \left(\int_X f_1 d\mu \right) \cdots \left(\int_X f_r d\mu \right) + o(1)$$

with $f_1, \dots, f_r \in \mathcal{A}$ and $h_1, \dots, h_r \in H$ and an explicit error term depending on the quantities $\min\{d(h_i, h_j) : i \neq j\}$ and $N_s(f_i)$. We refer to Section 2 for the required definitions and notation.

Our main result now reads as follows.

Theorem 1.1 (Sub-exponential growth and CLT). *Let (H, d) have sub-exponential growth and (F_T) be a right Følner sequence in H such that $m_H(F_T) \rightarrow \infty$. Suppose that the action of H on (X, μ) is exponentially mixing of all orders on an H -invariant sub-algebra \mathcal{A} of $L^\infty(X, \mu)$. Then, for any $f \in \mathcal{A}$ with $\mu(f) = 0$,*

$$\frac{1}{m_H(F_T)^{1/2}} \int_{F_T} (f \circ h^{-1}) dm_H(h) \implies N(0, \sigma_f^2) \quad \text{as } T \rightarrow \infty, \quad (1.5)$$

where

$$\sigma_f^2 = \int_H \langle f, f \circ h^{-1} \rangle dm_H(h). \quad (1.6)$$

A more general version of this theorem, which does not assume sub-exponential growth, will be stated in Theorem 1.5.

Remark 1.2. Exponential 2-mixing, combined with the sub-exponential growth of H , will ensure that σ_f given by (1.6) is finite for all $f \in \mathcal{A}$ (see Section 3.2). However, we stress that $\sigma_f = 0$ is definitely possible for non-zero f ; indeed, it is not hard to show that this happens if $f = g - g \circ h_0$ for some $h_0 \in H$ and $g \in \mathcal{A}$.

1.2.1. A sample application: CLT for Cartan actions on homogeneous spaces

Let L be a connected Lie group, Γ a lattice in L and μ_Y the unique L -invariant probability measure on $Y = \Gamma \backslash L$. Let G be a semisimple Lie subgroup of L with finite center, and assume that the G -action on (Y, μ_Y) has strong spectral gap. Let $A < G$ denote a Cartan subgroup of G , and fix a closed subgroup H of A . Let (B_T) be a sequence of strictly increasing balls in H with respect to some left-invariant Riemannian metric on G restricted to H . One readily checks that (B_T) forms a right Følner sequence in H .

In recent joint work [2] with M. Einsiedler, the authors showed that if one takes \mathcal{A} to be the algebra of smooth functions on Y with compact supports, and \mathcal{N} a family of certain Sobolev norms on \mathcal{A} , then the assumptions of Theorem 1.5 are satisfied for the H -action on (Y, μ_Y) with respect to $(\mathcal{A}, \mathcal{N})$, which leads to the following corollary of Theorem 1.1.

Corollary 1.3. *For every real-valued, compactly supported smooth function f on Y with $\mu_Y(f) = 0$,*

$$\frac{1}{m_H(B_T)^{1/2}} \int_{B_T} (f \circ h^{-1}) dm_H(h) \implies N(0, \sigma_f^2),$$

where σ_f is given by (1.6).

For example, Corollary 1.3 applies when

$$X = \mathrm{SL}_m(\mathbb{Z}) \backslash \mathrm{SL}_m(\mathbb{R}) \quad \text{or} \quad X = (\mathbb{Z}^m \rtimes \mathrm{SL}_m(\mathbb{Z})) \backslash (\mathbb{R}^m \rtimes \mathrm{SL}_m(\mathbb{R})),$$

and H is a closed subgroup of the group of diagonal matrices in $\mathrm{SL}_m(\mathbb{R})$.

Remark 1.4. The first author and G. Zhang [4] have recently provided, under some technical assumptions on G and H , lower (positive) bounds on the variance σ_f , whenever f is non-zero and invariant under a maximal compact subgroup of G .

1.3. Connections to earlier works

A number of different approaches have been developed for proving dynamical Central Limit Theorems for one-parameter actions. An influential approaches based on a martingale approximation originated in the work of Gordin [16]. We refer to the survey [22] by Le Borgne for an overview of this technique, as well as for an extensive list of references. The martingale approximation method becomes harder to implement already for actions of $H = (\mathbb{Z}^d, +)$, $d \geq 2$, let alone for actions by non-commutative groups; see for instance [31] for some recent developments in this direction when $H = (\mathbb{Z}^d, +)$. Other approaches to the Central Limit Theorem involve Markov approximations [27, 5, 6, 13] and spectral analysis of transfer operators [19, 18], and it is also not clear how to implement them for multi-parameter actions.

In this paper, we use an alternative approach to the Central Limit Theorem based on the classical *method of cumulants*, due to Fréchet and Shohat [15] (see Section 5 for an outline of this method). Roughly speaking, this method is equivalent to the more well-known *method of moments*, but is better tailored for approximations to Gaussian laws. An important novelty in our work is the systematic use of *conditional cumulants* (see Section 8), which greatly simplifies the estimation of cumulants in the presence of exponential mixing of all orders.

The method of cumulants has been recently used by Cohen and Conze in [7, 8, 9] to establish Central Limit Theorems for multiple mixing actions by $H = (\mathbb{Z}^d, +)$ by automorphisms of a compact abelian group X — but only for functions which are finite sums of characters on X . It seems that it is still not known whether in this setting the CLT also holds for all smooth functions.

1.4. A general Central Limit Theorem

Our method allows to prove the Central Limit Theorem for actions of general groups and for more general averaging schemes provided that some technical conditions are verified. Let us now assume that H is a locally compact second countable group equipped with a left-invariant metric d .

The following is the main technical result of this paper.

Theorem 1.5 (General CLT). *Let (ν_T) be a sequence of positive and finite Borel measures on H such that $\|\nu_T\| \rightarrow \infty$, and for any integer $r \geq 3$ and real number $c > 0$,*

$$\lim_{T \rightarrow \infty} \int_H \nu_T(B_d(h, c \log \|\nu_T\|))^{r-1} d\nu_T(h) = 0. \quad (1.7)$$

Suppose that the action of H on (X, μ) is exponentially mixing of all orders on an H -invariant subalgebra \mathcal{A} of $L^\infty(X, \mu)$. Let $f \in \mathcal{A}$ with $\mu(f) = 0$ and suppose that the limit

$$\sigma_f := \lim_{T \rightarrow \infty} \|\nu_T * f\|_L^2 < \infty \quad (1.8)$$

exists. Then

$$\nu_T * f \implies N(0, \sigma_f^2) \quad \text{as } T \rightarrow \infty.$$

In Section 3 below we show that these conditions are satisfied for the averages

$$\nu_T = \frac{1}{m_H(F_T)^{1/2}} \int_{F_T} \delta_h dm_H(h),$$

when (H, d) has sub-exponential growth, and (F_T) is a right Følner sequence in H , and we identify the limit in (1.8). Hence, Theorem 1.1 will be deduced from Theorem 1.5.

1.4.1. A sample application: CLT for unipotent flows sampled at lacunary times

Let us now retain the notation from Section 1.2.1, and let $u(t)$ be a non-trivial one-parameter unipotent subgroup G . Although actions of Cartan subgroups of G on $Y = \Gamma \backslash L$ are exponentially mixing of all orders, actions of *unipotent* subgroups are not (although they are *polynomially mixing of all orders*). Sinai raised the question whether the Central Limit Theorem still holds for unipotent flows on Y ; this was later answered in the negative by Flaminio and Forni in [14, Cor. 1.6]. However, our next corollary shows that if one is willing to “speed up” unipotent flows by sampling them at lacunary times, then the CLT does hold.

Corollary 1.6. *For any lacunary sequence $(\lambda_k) \subset \mathbb{R}_+$ and a real-valued, compactly supported smooth function f on Y with $\mu_Y(f) = 0$,*

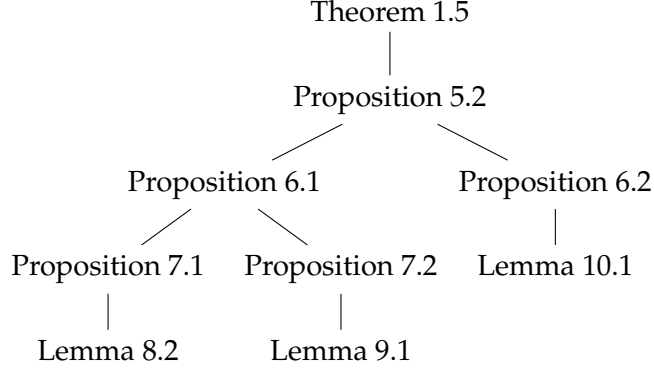
$$\frac{1}{\sqrt{N}} \sum_{k=1}^N f \circ u(\lambda_k) \implies N(0, \|f\|_{L^2}^2) \quad \text{as } N \rightarrow \infty. \quad (1.9)$$

Remark 1.7. Contrary to Theorem 1.1, we note that in this setting the variance σ_f is always positive whenever f has mean zero and does not vanish identically on Y .

1.5. Structure of the paper

In Sections 3 and 4 below we shall assume Theorem 1.5 and show how Theorem 1.1 and Corollary 1.6 respectively can be deduced from it.

The rest of the paper is then devoted to the proof of Theorem 1.5, which we shall break down into several propositions, whose proofs, in turn, will be divided into further lemmas and propositions. The following tree graphically represents this break-down:



2. DEFINITIONS AND STANDING TECHNICAL ASSUMPTIONS

Let us here in this section give the definitions of some notions used in the introduction, and collect some of the technical assumptions which are necessary for our analysis.

Throughout the paper, we shall fix a locally compact second countable group H , and a left-invariant metric d on H . We assume that H acts jointly measurably by measure preserving maps on a probability measure space (X, μ) . In particular, H also acts (weakly continuously) by isometries on $L^\infty(X, \mu)$ via

$$h \cdot f = f \circ h^{-1}, \quad \text{for } f \in L^\infty(X, \mu) \text{ and } h \in H.$$

We fix an H -invariant sub-algebra \mathcal{A} of $L^\infty(X, \mu)$, and a family $\mathcal{N} = (N_s)$ of semi-norms on \mathcal{A} , indexed by positive integers s .

Definition 2.1 (Exponential mixing of all orders). Let $r \geq 2$ be an integer. We say that the H -action on (X, μ) is *exponentially mixing of order r* , with respect to d and $(\mathcal{A}, \mathcal{N})$, if there exist $\delta_r > 0$ and an integer $s_r > 0$ such that for all $s > s_r$ and $f_1, \dots, f_r \in \mathcal{A}$,

$$\left| \mu \left(\prod_{i=1}^r h_i \cdot f_i \right) - \prod_{i=1}^r \mu(f_i) \right| \ll_{r,s} e^{-\delta_r d_r(\underline{h})} \prod_{i=1}^r N_s(f_i), \quad (2.1)$$

for all $\underline{h} = (h_1, \dots, h_r) \in H^r$, where

$$d_r(\underline{h}) = \min_{i \neq j} d(h_i, h_j). \quad (2.2)$$

We refer to δ_r as the *rate of r -mixing*. The notation $A \ll_{r,s} B$ here means that there exists a constant c , which is allowed to depend on r and s such that $A \leq c B$.

We shall from now on always assume that the H -action on (X, μ) is exponentially mixing of all orders with respect to d and $(\mathcal{A}, \mathcal{N})$; in particular, from now on, the meanings of the numbers s_r and δ_r have been fixed. Furthermore, we shall require the following four technical conditions on the family $\mathcal{N} = (N_s)$ (with all implicit constants depending only on s):

- **Monotonicity:** For all $s \geq 1$ and $f \in \mathcal{A}$,

$$N_s(f) \ll_s N_{s+1}(f). \quad (2.3)$$

- **Sobolev embedding:** For all $s \geq 1$ and $f \in \mathcal{A}$,

$$\|f\|_{L^\infty} \ll_s N_s(f). \quad (2.4)$$

- **H-boundedness:** For all $s \geq 1$, there exists $\sigma_s > 0$ such that for all $f \in \mathcal{A}$ and $h \in H$,

$$N_s(h \cdot f) \ll_s e^{\sigma_s d(h,e)} N_s(f). \quad (2.5)$$

- **Almost multiplicative:** For all $s \geq 1$ and $f_1, f_2 \in \mathcal{A}$,

$$N_s(f_1 f_2) \ll_s N_{s+1}(f_1) N_{s+1}(f_2). \quad (2.6)$$

We may further, and shall, throughout the paper, assume that the sequence (σ_s) increases with s , and the sequence (δ_r) from Definition 2.1 decreases with r . Also, without any loss of generality, we may assume that $\delta_r < r\sigma_s$ for all r and s .

3. PROOF OF THEOREM 1.1 ASSUMING THEOREM 1.5

Recall our standing assumptions on $H, d, (X, \mu)$ and $(\mathcal{A}, \mathcal{N})$ from Section 2. Let us further assume that (H, d) has sub-exponential growth, and (F_T) is a right Følner sequence in H . We shall apply Theorem 1.5 to the sequence (ν_T) of positive and finite measures on H defined by

$$\nu_T = \frac{1}{m_H(F_T)^{1/2}} \int_{F_T} \delta_h dm_H(h). \quad (3.1)$$

One readily checks that $\|\nu_T\| = m_H(F_T)^{1/2}$ for all T . In particular, $\|\nu_T\| \rightarrow \infty$ as $T \rightarrow \infty$. In order to prove Theorem 1.1, it suffices to verify conditions (1.7) and (1.8).

3.1. Checking condition (1.7)

Writing out condition (1.7) explicitly in our setting, we see that we must prove that for every integer $r \geq 3$ and real number $c > 0$,

$$\frac{1}{m_H(F_T)^{r/2}} \int_{F_T} m_H(F_T \cap B_d(h, \frac{c}{2} \log m_H(F_T)))^{r-1} dm_H(h) \rightarrow 0,$$

as $T \rightarrow \infty$. The integral is bounded from above by

$$m_H(F_T) m_H(B_d(e, \frac{c}{2} \log m_H(F_T)))^{r-1}.$$

Hence it suffices to show that for every $r \geq 3$ and real number $c > 0$,

$$\frac{m_H(B_d(e, \frac{c}{2} \log m_H(F_T)))^{r-1}}{m_H(F_T)^{r/2-1}} \rightarrow 0,$$

as $T \rightarrow \infty$. Since $m_H(F_T) \rightarrow \infty$ and $r/2 > 1$, this readily follows from the sub-exponential growth of (H, d) , see (1.4).

3.2. Calculating the variance

Upon expanding (1.8) for our choice of (ν_T) and a fixed $f \in \mathcal{A}$ with $\mu(f) = 0$, we see that one has to show that

$$\frac{1}{m_H(F_T)} \int_H \int_H \chi_{F_T}(h_1) \chi_{F_T}(h_2) \phi_f(h_1^{-1} h_2) dm_H(h_1) dm_H(h_2) \rightarrow \int_H \phi_f(h) dm_H(h), \quad (3.2)$$

where $\phi_f(h) = \langle f, h \cdot f \rangle$. Using left- H -invariance of m_H , the left-hand side can be re-written as

$$\int_H \frac{m_H(F_T \cap F_T h^{-1})}{m_H(F_T)} \phi_f(h) dm_H(h). \quad (3.3)$$

Since (F_T) is a right Følner sequence in H ,

$$\lim_{T \rightarrow \infty} \frac{m_H(F_T \cap F_T h^{-1})}{m_H(F_T)} = 1, \text{ for all } h \in H,$$

whence (3.2) follows from the Dominated Convergence Theorem if we can show that ϕ_f belongs to $L^1(H)$. Since the H -action on (X, μ) is exponentially mixing of order two with respect to d and $(\mathcal{A}, \mathcal{N})$, we have

$$|\phi_f(h)| \ll_s e^{-\delta_2 d(h,e)} N_s(f)^2 \quad \text{for all } h \in H \text{ and } s > s_2,$$

where the implicit constant is independent of h . Hence, the following lemma finishes the proof.

Lemma 3.1. *If ϕ is a complex-valued measurable function on H such that for some $C, \alpha > 0$,*

$$|\phi(h)| \leq C e^{-\alpha d(h,e)}, \quad h \in H, \quad (3.4)$$

then ϕ belongs to $L^1(H)$.

Proof. Using (3.4), we obtain

$$\int_{H \setminus \{0\}} |\phi(h)| dm_H(h) = \sum_{n \geq 0} \int_{B_d(e, n+1) \setminus B_d(e, n)} |\phi(h)| dm_H(h) \leq C \sum_{n \geq 0} \beta_n e^{-\alpha n},$$

where $\beta_n = m_H(B_d(e, n+1))$. Since (H, d) has sub-exponential growth, $\beta_n^{1/n} \rightarrow 1$, which readily implies that the series converges. \square

4. PROOF OF COROLLARY 1.6 ASSUMING THEOREM 1.5

Recall our assumptions on L, Γ, G, Y and μ_Y from Corollary 1.6, and let $u(t)$ be a non-trivial unipotent one-parameter subgroup of G , and (λ_k) a sequence in \mathbb{R}_+ such that for some $\theta > 1$,

$$\lambda_{k+1} \geq \theta \lambda_k, \quad \text{for all } k. \quad (4.1)$$

We shall apply Theorem 1.5 to the sequence (ν_N) of positive and finite measures on H defined by

$$\nu_N = \frac{1}{\sqrt{N}} \sum_{k=1}^N \delta_{u(\lambda_k)^{-1}}. \quad (4.2)$$

One readily checks that $\|\nu_N\| = \sqrt{N}$.

We shall crucially use the easily checkable fact the distance along unipotent subgroups grows at least logarithmically (see, for instance, [2, Lem. 2.1]): for every fixed choice of such a unipotent subgroup $u(t)$, there exist $c_1, c_2 > 0$ such that

$$d(u(t), e) \geq c_1 \log |t| - c_2 \quad \text{for all } t \neq 0. \quad (4.3)$$

In particular, it follows from the exponential mixing property that there exists $p > 0$ such that for every $s > s_2$ and $f_1, f_2 \in \mathcal{A}$ satisfying $\mu_Y(f_1) = \mu_Y(f_2) = 0$,

$$\langle f_1, u(t) \cdot f_2 \rangle \ll_s |t|^{-p} N_s(f_1) N_s(f_2) \quad \text{for all } t \neq 0. \quad (4.4)$$

4.1. Checking condition (1.7)

If we write out condition (1.7) explicitly in our setting, we see that we must show that for all $r \geq 3$ and $c > 0$,

$$\frac{1}{N^{r/2}} \sum_{n=1}^N |\{m = 1, \dots, N : d(u(\lambda_m), u(\lambda_n)) \leq c \log \sqrt{N}\}|^{r-1} \rightarrow 0,$$

as $N \rightarrow \infty$. We deduce from (4.3) that for $m \neq n$,

$$d(u(\lambda_m), u(\lambda_n)) = d(u(\lambda_m - \lambda_n), e) \geq c_1 \log |\lambda_m - \lambda_n| - c_2.$$

Hence, it suffices to show that for every $c, \theta > 0$,

$$\frac{1}{N^{r/2}} \sum_{n=1}^N |\{m = 1, \dots, N : |\lambda_m - \lambda_n| \leq c N^\theta\}|^{r-1} \rightarrow 0. \quad (4.5)$$

Using (4.1), it is not hard to show that

$$|\{m \geq 1 : |\lambda_m - \lambda_n| \leq c N^\theta\}| \ll \log N,$$

where the implied constant is independent of n . Then

$$\sum_{n=1}^N |\{m = 1, \dots, N : |\lambda_m - \lambda_n| \leq c N^\theta\}|^{r-1} \ll N(\log N)^{r-1},$$

and (4.5) is immediate since $r \geq 3$.

4.2. Calculating the variance

Take $f \in \mathcal{A}$ with $\mu_Y(f) = 0$. If we expand $\|\nu_N * f\|_{L^2}^2$, we get

$$\|f\|_{L^2}^2 + \frac{2}{N} \sum_{1 \leq m < n \leq N} \langle f, u(\lambda_n - \lambda_m) \cdot f \rangle.$$

We wish to prove that the second term tends to zero as $N \rightarrow \infty$. By (4.4), we have for all $s > s_2$,

$$\langle f, u(\lambda_n - \lambda_m) \cdot f \rangle \ll_s \frac{1}{|\lambda_n - \lambda_m|^p} N_s(f)^2.$$

It thus suffices to show that

$$\sum_{1 \leq m < n < \infty} \frac{1}{|\lambda_n - \lambda_m|^p} = \sum_{m=1}^{\infty} \sum_{k=1}^{\infty} \frac{1}{|\lambda_{m+k} - \lambda_m|^p} < \infty.$$

Since $|\lambda_{m+k} - \lambda_m| \geq \lambda_m(\theta^k - 1)$ by (4.1), this follows from the finiteness of the series

$$\sum_{m=1}^{\infty} \frac{1}{\lambda_m^p} \quad \text{and} \quad \sum_{k=1}^{\infty} \frac{1}{(\theta^k - 1)^p}.$$

5. AN OUTLINE OF THE PROOF OF THEOREM 1.5

Our proof of Theorem 1.5 makes use of the classical *cumulant method*, in essence due to Fréchet and Shohat in [15]. We shall briefly summarize its main steps below.

Let (X, μ) be a probability measure space and $r \geq 2$ an integer. Denote by $[r]$ the set $\{1, \dots, r\}$. A *cyclically ordered partition* \mathcal{P} of the set $[r]$ is a partition $\{I_1, \dots, I_k\}$ of $[r]$ into non-empty subsets I_1, \dots, I_k , where the cyclic order of I_1, \dots, I_k is also taken into account; for instance, if $[r] = I_1 \sqcup I_2 \sqcup I_3$, then $\{I_1, I_2, I_3\}$ and $\{I_2, I_3, I_1\}$ are viewed as the same *cyclically ordered* partition, while $\{I_1, I_2, I_3\}$ and $\{I_2, I_1, I_3\}$ are viewed as different partitions, since the associated orders (123) and (213) are not cyclic permutations of each other. We denote by $\mathfrak{P}_{[r]}$ the set of all cyclically ordered partitions of $[r]$.

Given an r -tuple (f_1, \dots, f_r) in $L^\infty(X, \mu)$ and a subset $I \subset [r]$, we define

$$f_I = \prod_{i \in I} f_i, \quad \text{for } \emptyset \neq I \subset [r],$$

and the *joint cumulant* $\text{cum}_{[r]}(f_1, \dots, f_r)$ of order r by

$$\text{cum}_{[r]}(f_1, \dots, f_r) = \sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \prod_{I \in \mathcal{P}} \mu(f_I), \quad (5.1)$$

where $|\mathcal{P}|$ denotes the number of partition elements in \mathcal{P} . If $f \in L^\infty(X, \mu)$, we define $\text{cum}_r(f)$, the *cumulant of f of order r* , to be

$$\text{cum}_r(f) = \text{cum}_{[r]}(f, \dots, f). \quad (5.2)$$

The utility of cumulants for problems pertaining to Central Limit Theorems is well-known; we shall use the following classical criterion, which can be deduced from the results in [15].

Proposition 5.1 (Cumulants and CLT). *Let (Z_T) be a sequence of real-valued, bounded and measurable functions on (X, μ) satisfying $\mu(Z_T) = 0$. If*

$$\lim_{T \rightarrow \infty} \text{cum}_r(Z_T) = 0, \quad \text{for all } r \geq 3, \quad (5.3)$$

and the limit

$$\sigma^2 := \lim_{T \rightarrow \infty} \|Z_T\|_{L^2}^2 < \infty \quad (5.4)$$

exists, then

$$Z_T \implies N(0, \sigma^2) \quad \text{as } T \rightarrow \infty.$$

5.1. Main proposition

Recall our standing assumptions concerning $H, d, (X, \mu), \mathcal{A}$ and the norms (N_s) from Subsection 2. In particular, the meaning of the numerical sequences (s_r) and (δ_r) has been fixed. Let (ν_T) be a sequence of positive and finite measures on H . Given $f \in \mathcal{A}$, thanks to Proposition 5.1, the proof of the CLT for the sequence $Z_T := \nu_T * f$ is essentially reduced to the asymptotic vanishing of $\text{cum}_r(\nu_T * f)$ for $r \geq 3$. Using our assumptions on (ν_T) in Theorem 1.5, this will be deduced from the following proposition.

Proposition 5.2 (Estimating cumulants). *For all $r \geq 3$ and $s > s_r + r$, there exists $c_{r,s} > 0$ such that for all $\gamma > 0, T > 0$ and $f \in \mathcal{A}$,*

$$|\text{cum}_r(\nu_T * f)| \ll_{r,s} \left(\int_H \nu_T(B_d(h, c_{r,s}\gamma))^{r-1} d\nu_T(h) + e^{-\delta_r \gamma} \|\nu_T\|^r \right) N_s(f)^r \quad (5.5)$$

where the implicit constant depends only on r and s .

5.2. Proof of Theorem 1.5 assuming Theorem 5.2

Fix $f \in \mathcal{A}$. By Proposition 5.1, applied to $Z_T = \nu_T * f$, we need to show that the limit

$$\sigma_f^2 = \lim_{T \rightarrow \infty} \|\nu_T * f\|_{L^2}^2$$

exists, and

$$\lim_{T \rightarrow \infty} \text{cum}_r(\nu_T * f) = 0, \quad \text{for all } r \geq 3.$$

The existence of the first limit is assumed in Theorem 1.5, so we only need to consider the second kind of limits. Fix $r \geq 3$, and choose $\gamma = \varepsilon \log \|\nu_T\|$ in Proposition 5.2 for some $\varepsilon > r/\delta_r$. Then, (5.5) yields for all $s > s_r + r$,

$$|\text{cum}_r(\nu_T * f)| \ll_{r,s} \left(\int_H \nu_T(B_d(h, c_{r,s} \varepsilon \log \|\nu_T\|))^{r-1} d\nu_T(h) + \|\nu_T\|^{-(\delta_r \varepsilon - r)} \right) N_s(f)^r.$$

Applying our assumption (1.7) with $c = c_{r,s} \varepsilon$, we conclude that this quantity tends to zero as $T \rightarrow \infty$, since $\|\nu_T\| \rightarrow \infty$.

6. AN OUTLINE OF THE PROOF OF PROPOSITION 5.2

Throughout this section, we retain our assumptions on $H, d, (X, \mu), \mathcal{A}$ and the norms (N_s) . We shall further fix an integer $r \geq 3$, and write $[r]$ for the set $\{1, \dots, r\}$.

6.1. Rewriting cumulants

Let (ν_T) be a sequence of positive and finite measures on H . For any $f \in \mathcal{A}$, we see that

$$\text{cum}_r(\nu_T * f) = \int_{H^r} \text{cum}_{[r]}(h_1 \cdot f, \dots, h_r \cdot f) d\nu_T^{\otimes r}(h_1, \dots, h_r).$$

Furthermore, for every fixed $\underline{h} = (h_1, \dots, h_r) \in H^r$, we can write every joint cumulant of the form $\text{cum}_{[r]}(h_1 \cdot f, \dots, h_r \cdot f)$ as

$$\sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \prod_{I \in \mathcal{P}} \psi_{f, \underline{h}}(I),$$

where

$$\psi_{f, \underline{h}}(I) = \mu\left(\prod_{i \in I} h_i \cdot f\right), \quad \text{for } I \subset [r].$$

We shall now adopt the following notational convention. If ψ is a real-valued function defined on the set $2^{[r]}$ of all subsets of $[r]$ with $\psi(\emptyset) = 1$, then we define its *cumulant* as

$$\text{cum}_{[r]}(\psi) = \sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \prod_{I \in \mathcal{P}} \psi(I), \quad (6.1)$$

so that $\text{cum}_{[r]}(\psi_{f, \underline{h}}) = \text{cum}_{[r]}(h_1 \cdot f, \dots, h_r \cdot f)$. With this convention, we now have

$$\begin{aligned} \text{cum}_r(\nu_T * f) &= \int_{H^r} \left(\sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \prod_{I \in \mathcal{P}} \psi_{f, \underline{h}}(I) \right) d\nu_T^{\otimes r}(\underline{h}) \\ &= \int_{H^r} \text{cum}_{[r]}(\psi_{f, \underline{h}}) d\nu_T^{\otimes r}(\underline{h}). \end{aligned} \quad (6.2)$$

In what follows, we shall estimate $\text{cum}_{[r]}(\psi_{f, \underline{h}})$ for “well-separated” r -tuples $\underline{h} = (h_1, \dots, h_r)$, and we shall also show that “most” r -tuples are well-separated on suitable scales. In order to make the notions of “well-separateness” and “most” more precise, we must first introduce some additional notation.

6.2. Well-separated r -tuples

If $I, J \subset [r]$ and $\underline{h} = (h_1, \dots, h_r) \in H^r$, we set

$$d^I(\underline{h}) = \max \{d(h_i, h_j) : i, j \in I\}, \quad (6.3)$$

and

$$d_{I, J}(\underline{h}) = \min \{d(h_i, h_j) : i \in I, j \in J\}.$$

If \mathcal{Q} is a partition of $[r]$, we set

$$d^{\mathcal{Q}}(\underline{h}) = \max \{d^I(\underline{h}) : I \in \mathcal{Q}\} \quad (6.4)$$

and

$$d_{\mathcal{Q}}(\underline{h}) = \min \{d_{I, J}(\underline{h}) : I \neq J, I, J \in \mathcal{Q}\}.$$

Two extremal cases of this notation will be of special interest. We note that if we write \mathcal{K}_r for the partition of $[r]$ into points, then

$$d_{\mathcal{K}_r}(\underline{h}) = \min \{d(h_i, h_j) : 1 \leq i \neq j \leq r\},$$

which we have previously also denoted by $d_r(\underline{h})$. At the other extreme, if $\{[r]\}$ denotes the partition into one single block, then

$$d^{\{[r]\}}(\underline{h}) = \max \{d(h_i, h_j) : 1 \leq i, j \leq r\}.$$

In order to ease the somewhat heavy notation, we set $d^r = d^{\{[r]\}}$. For $\beta > 0$, we set

$$\Delta(\beta) = \{\underline{h} \in H^r : d^r(\underline{h}) \leq \beta\}.$$

We note that for all $h_1 \in H$, the set

$$\Delta(\beta)_{h_1} := \{(h_2, \dots, h_r) : (h_1, \dots, h_r) \in \Delta(\beta)\}$$

satisfies

$$\Delta(\beta)_{h_1} \subset B_d(h_1, \beta)^{r-1}. \quad (6.5)$$

Given a partition \mathcal{Q} of $[r]$, and $0 \leq \alpha < \beta$, we define

$$\Delta_{\mathcal{Q}}(\alpha, \beta) = \{\underline{h} \in H^r : d^{\mathcal{Q}}(\underline{h}) \leq \alpha, \text{ and } d_{\mathcal{Q}}(\underline{h}) > \beta\}. \quad (6.6)$$

We shall think of the elements in $\Delta_{\mathcal{Q}}(\alpha, \beta)$ for some partition \mathcal{Q} with $|\mathcal{Q}| \geq 2$ and $0 \leq \alpha < \beta$ as being “well-separated”, while we think of the elements in $\Delta(\beta)$ as being “clustered”.

6.3. Main propositions

Our first proposition roughly asserts that the joint cumulants $\text{cum}_{[r]}(h_1 \cdot f, \dots, h_r \cdot f)$ are “small” for all “well-separated” r -tuples $\underline{h} = (h_1, \dots, h_r)$.

Proposition 6.1 (Separated tuples). *Let \mathcal{Q} be a partition of $[r]$ with $|\mathcal{Q}| \geq 2$, and fix $0 \leq \alpha < \beta$ and an integer $s > s_r + r$. Then, for every $\underline{h} \in \Delta_{\mathcal{Q}}(\alpha, \beta)$ and $f \in \mathcal{A}$, we have*

$$|\text{cum}_{[r]}(\psi_{f, \underline{h}})| \ll_{r,s} e^{-(\beta \delta_r - r \alpha \sigma_s)} N_s(f)^r, \quad (6.7)$$

where the implicit constant depends only on r and s .

Our second proposition roughly shows that we have a lot of flexibility in setting up the thresholds for the notions of “well-separated” and “clustered”.

Proposition 6.2 (Exhausting H^r). *For every sequence (β_j) with $\beta_0 = 0$ and*

$$0 < \beta_1 < 3\beta_1 < \beta_2 < \dots < \beta_{r-1} < 3\beta_{r-1} < \beta_r, \quad (6.8)$$

we have

$$H^r = \Delta(\beta_r) \cup \left(\bigcup_{j=0}^{r-1} \bigcup_{|\mathcal{Q}| \geq 2} \Delta_{\mathcal{Q}}(3\beta_j, \beta_{j+1}) \right). \quad (6.9)$$

6.4. Proof of Proposition 5.2 assuming Proposition 6.1 and Proposition 6.2

Fix $s > s_r + r$ and $f \in \mathcal{A}$, and pick a sequence (β_j) such that $\beta_0 = 0$, and

$$0 < \beta_1 < 3\beta_1 < \beta_2 < 3\beta_2 < \beta_3 < \dots < \beta_{r-1} < 3\beta_{r-1} < \beta_r. \quad (6.10)$$

By (6.2) and (6.9), we have that for all $T > 0$,

$$\begin{aligned} |\text{cum}_r(\nu_T * f)| &\ll_r \nu_T^{\otimes r}(\Delta(\beta_r)) \|f\|_{L^\infty}^r \\ &+ \max_j \max_{|\mathcal{Q}| \geq 2} \int_{\Delta_{\mathcal{Q}}(3\beta_j, \beta_{j+1})} |\text{cum}_{[r]}(\psi_{f, \underline{h}})| d\nu_T^{\otimes r}(\underline{h}), \end{aligned} \quad (6.11)$$

where the second maximum is taken over all partitions of $[r]$ with at least two partition elements. Recall from our standing assumptions in Subsection 2 that $\|f\|_\infty \ll_s N_s(f)$. Hence, using the inclusion (6.5) for the first term, we see that

$$\begin{aligned} |\text{cum}_r(\nu_T * f)| &\ll_r \left(\int_{\mathbf{H}} \nu_T(B(\mathbf{h}, \beta_r))^{r-1} d\nu_T(\mathbf{h}) \right) N_s(f)^r \\ &+ \max_j \max_{|\mathcal{Q}| \geq 2} \int_{\Delta_{\mathcal{Q}}(3\beta_j, \beta_{j+1})} |\text{cum}_{[r]}(\psi_{f, \underline{h}})| d\nu_T^{\otimes r}(\underline{h}). \end{aligned} \quad (6.12)$$

We stress that this inequality is valid for any $f \in \mathcal{A}$ and sequence (β_j) satisfying (6.10). It remains to choose a sequence (β_j) so that the second term is as small as possible.

Let us now fix $\gamma > 0, T > 0$ once and for all, and set $\beta_0 = 0$. For every $j \geq 0$, we pick recursively β_{j+1} so that

$$\beta_{j+1} \delta_r - 3r\beta_j \sigma_s = \delta_r \gamma. \quad (6.13)$$

Recall from Subsection 2 that we assume that $\delta_r < r\sigma_s$, and thus (6.13) in particular implies that $3\beta_j < \beta_{j+1}$, that is to say, (β_j) thus constructed satisfies (6.10). By induction, (6.13) also implies that

$$\beta_r \leq \gamma \sum_{j=0}^{r-1} \left(\frac{3r\sigma_s}{\delta_r} \right)^j =: \gamma c_{r,s}. \quad (6.14)$$

In what follows, we can choose any sequence (β_j) as in (6.13) with $\beta_r \leq \gamma c_{r,s}$.

We fix a partition \mathcal{Q} of $[r]$ with $|\mathcal{Q}| \geq 2$ and an index j . By Proposition 6.1, we know that for all “well-separated” r -tuples $\underline{h} \in \Delta_{\mathcal{Q}}(3\beta_j, \beta_{j+1})$,

$$|\text{cum}_{[r]}(\psi_{f, \underline{h}})| \ll_{r,s} e^{-(\beta_{j+1} \delta_r - 3r\beta_j \sigma_s)} N_s(f)^r = e^{-\delta_r \gamma} N_s(f)^r,$$

where the last equality follows from (6.13). We stress that the right hand side is independent of both \mathcal{Q} and j , and thus it follows from (6.12) that

$$|\text{cum}_r(\nu_T * f)| \ll_{r,s} \left(\int_{\mathbf{H}} \nu_T(B(\mathbf{h}, c_{r,s} \gamma))^{r-1} d\nu_T(\mathbf{h}) + e^{-\delta_r \gamma} \right) N_s(f)^r,$$

which finishes the proof.

7. PROOF OF PROPOSITION 6.1

We retain the notation from the previous section. In particular, an integer $r \geq 3$ has been fixed, and we write $[r]$ for the set $\{1, \dots, r\}$, and $\mathfrak{P}_{[r]}$ for the set of cyclically ordered partitions of $[r]$.

Recall from Subsection 6.1 that if $\psi : 2^{[r]} \rightarrow \mathbb{R}$ with $\psi(\emptyset) = 1$, then its cumulant $\text{cum}_{[r]}(\psi)$ is defined by

$$\text{cum}_{[r]}(\psi) = \sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \prod_{I \in \mathcal{P}} \psi(I).$$

We can extend ψ to a function $\tilde{\psi} : \mathfrak{P}_{[r]} \rightarrow \mathbb{R}$ by

$$\tilde{\psi}(\mathcal{P}) = \prod_{I \in \mathcal{P}} \psi(I), \quad \text{for } \mathcal{P} \in \mathfrak{P}_{[r]}, \quad (7.1)$$

so that

$$\text{cum}_{[r]}(\psi) = \sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \tilde{\psi}(\mathcal{P}).$$

Finally, given a partition \mathcal{Q} of $[r]$, we set

$$\psi^{\mathcal{Q}}(I) = \prod_{J \in \mathcal{Q}} \psi(I \cap J) \quad \text{and} \quad \tilde{\psi}^{\mathcal{Q}}(\mathcal{P}) = \prod_{I \in \mathcal{P}} \psi^{\mathcal{Q}}(I). \quad (7.2)$$

7.1. Estimating cumulants

Recall that if $f \in \mathcal{A}$ and $\underline{h} = (h_1, \dots, h_r) \in \mathbb{H}^r$, then $\psi_{f, \underline{h}} : 2^{[r]} \rightarrow \mathbb{R}$ is defined by

$$\psi_{f, \underline{h}}(I) = \mu\left(\prod_{i \in I} h_i \cdot f\right), \quad \text{for } \emptyset \neq I \subset [r],$$

and $\psi_{f, \underline{h}}(\emptyset) = 1$. Our first proposition asserts that in order to estimate $\text{cum}_{[r]}(\psi_{f, \underline{h}})$ from above, it suffices to estimate all differences of the form $|\tilde{\psi}_{f, \underline{h}}(\mathcal{P}) - \tilde{\psi}_{f, \underline{h}}^{\mathcal{Q}}(\mathcal{P})|$, where \mathcal{Q} varies over all possible partitions of $[r]$ with at least two blocks. These differences will be estimated below using our assumption that the \mathbb{H} -action on (X, μ) is exponentially mixing of all orders.

Proposition 7.1. *For any partition \mathcal{Q} of $[r]$ with $|\mathcal{Q}| \geq 2$, $\underline{h} \in \mathbb{H}^r$ and $f \in \mathcal{A}$,*

$$|\text{cum}_{[r]}(\psi_{f, \underline{h}})| \ll_r \max \left\{ |\tilde{\psi}_{f, \underline{h}}(\mathcal{P}) - \tilde{\psi}_{f, \underline{h}}^{\mathcal{Q}}(\mathcal{P})| : \mathcal{P} \in \mathfrak{P}_{[r]} \right\}. \quad (7.3)$$

Given this result, which will be established in Section 8 below, Proposition 6.1 follows immediately from the following proposition, which will be established in Section 9.

Proposition 7.2 (Estimating the effect of conditioning). *Fix $0 \leq \alpha < \beta$ and an integer $s > s_r + r$. Then for any partition \mathcal{Q} of $[r]$, $\underline{h} \in \Delta_{\mathcal{Q}}(\alpha, \beta)$ and $f \in \mathcal{A}$,*

$$|\tilde{\psi}_{f, \underline{h}}(\mathcal{P}) - \tilde{\psi}_{f, \underline{h}}^{\mathcal{Q}}(\mathcal{P})| \ll_{r, s} e^{-(\beta \delta_r - r \alpha \sigma_s)} \mathbf{N}_s(f)^r, \quad (7.4)$$

where the implicit constant depends only on r and s .

8. PROOF OF PROPOSITION 7.1

Throughout this section, let $r \geq 3$ be an integer, and write $[r]$ for the set $\{1, \dots, r\}$.

If $\psi : 2^{[r]} \rightarrow \mathbb{R}$ is a set function with $\psi(\emptyset) = 1$, recall the definition of its cumulant $\text{cum}_{[r]}(\psi)$ from (6.1), and if \mathcal{Q} is a partition of $[r]$, recall the definition of the ‘‘conditional’’ set function $\psi^{\mathcal{Q}}$ from (7.2), and the definitions of the ‘‘extended’’ versions $\tilde{\psi}$ and $\tilde{\psi}^{\mathcal{Q}}$ from (7.1). It follows immediately from the definition of $\text{cum}_{[r]}(\psi)$ that

$$|\text{cum}_{[r]}(\psi) - \text{cum}_{[r]}(\psi^{\mathcal{Q}})| \ll_r \max \left\{ |\tilde{\psi}(\mathcal{P}) - \tilde{\psi}^{\mathcal{Q}}(\mathcal{P})| : \mathcal{P} \in \mathfrak{P}_{[r]} \right\}. \quad (8.1)$$

We recall that the cumulant of random variables X_1, \dots, X_r vanish provided that there exists a non-trivial partition $[r] = I \sqcup J$ such that $(X_i : i \in I)$ and $(X_j : j \in J)$ are independent (see, for instance, [29, Lem. 4.1]). The following proposition is a combinatorial version of this property. It can be proved by modifying the argument from [29] (see also Theorem 2 in [1]). We include a proof for completeness.

Proposition 8.1. *For any partition \mathcal{Q} of $[r]$ with $|\mathcal{Q}| \geq 2$ and $\psi : 2^{[r]} \rightarrow \mathbb{R}$, we have*

$$\text{cum}_{[r]}(\psi^{\mathcal{Q}}) = 0.$$

We note that Proposition 7.1 is a direct consequence of Proposition 8.1 and estimate (8.1).

Let us briefly explain the driving mechanism in the proof of this proposition. Recall that $\mathfrak{P}_{[r]}$ denotes the set of all cyclically ordered partitions of the set $[r]$. Let $\psi : 2^{[r]} \rightarrow \mathbb{R}$ be a set function and suppose that there exists a bijection $\tau : \mathfrak{P}_{[r]} \rightarrow \mathfrak{P}_{[r]}$ such that

$$|\tau(\mathcal{P})| = |\mathcal{P}| + 1 \pmod{2} \quad \text{and} \quad \tilde{\psi}(\tau(\mathcal{P})) = \tilde{\psi}(\mathcal{P}), \quad (8.2)$$

for all $\mathcal{P} \in \mathfrak{P}_{[r]}$. Then,

$$\text{cum}_{[r]}(\psi) = \sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\tau(\mathcal{P})|-1} \tilde{\psi}(\tau(\mathcal{P})) = - \sum_{\mathcal{P} \in \mathfrak{P}_{[r]}} (-1)^{|\mathcal{P}|-1} \tilde{\psi}(\mathcal{P}) = -\text{cum}_{[r]}(\psi),$$

and thus $\text{cum}_{[r]}(\psi) = 0$.

The next lemma shows that one can produce, for every partition \mathcal{Q} of $[r]$ with at least two partition elements, a bijection $\tau : \mathfrak{P}_{[r]} \rightarrow \mathfrak{P}_{[r]}$ such that (8.2) holds for $\tilde{\psi}^{\mathcal{Q}}$, for every choice of set function $\psi : 2^{[r]} \rightarrow \mathbb{R}$. In particular, by the comment above, this shows that $\text{cum}_{[r]}(\tilde{\psi}^{\mathcal{Q}}) = 0$, which finishes the proof of Proposition 7.1.

Lemma 8.2. *For any partition \mathcal{Q} of $[r]$ with $|\mathcal{Q}| \geq 2$, there exists a bijection $\tau : \mathfrak{P}_{[r]} \rightarrow \mathfrak{P}_{[r]}$ such that*

- for every $\mathcal{P} \in \mathfrak{P}_{[r]}$, we have $|\tau(\mathcal{P})| = |\mathcal{P}| + 1 \pmod{2}$, and
- for every $\psi : 2^{[r]} \rightarrow \mathbb{R}$, we have $\tilde{\psi}^{\mathcal{Q}} \circ \tau = \tilde{\psi}^{\mathcal{Q}}$.

8.1. Proof of Lemma 8.2

If $\mathcal{Q} = (I_1, \dots, I_n)$ is a cyclically ordered partition with $n \geq 2$, we set

$$M = I_1 \quad \text{and} \quad N = \bigsqcup_{k=2}^n I_k,$$

so that $[r] = M \sqcup N$, and we note if $\psi : 2^{[r]} \rightarrow \mathbb{R}$, then the function $\phi := \psi^{\mathcal{Q}}$ satisfies

$$\phi(I \sqcup J) = \phi(I) \phi(J), \quad \text{for all } I \subset M \text{ and } J \subset N. \quad (8.3)$$

We shall use the decomposition $[r] = M \sqcup N$ to construct a bijection $\tau : \mathfrak{P}_{[r]} \rightarrow \mathfrak{P}_{[r]}$ such that

$$|\tau(\mathcal{P})| = |\mathcal{P}| + 1 \pmod{2} \quad \text{and} \quad \tilde{\phi} \circ \tau = \tilde{\phi},$$

for all cyclically ordered partitions \mathcal{P} of $[r]$ and all functions $\phi : 2^{[r]} \rightarrow \mathbb{R}$ which satisfy (8.3).

To this end, we choose once and for all an element $y_o \in M$. Given a cyclically ordered partition $\mathcal{P} = (P_1, \dots, P_k)$ of $[r]$, let i be the unique index such that $y_o \in P_i$, and pick the first index j following i (in the cyclic ordering of \mathcal{P}) such that $P_j \cap N \neq \emptyset$. We now set

$$\tau(\mathcal{P}) = \begin{cases} (P_1, \dots, P_{j-1}, P_j \cap M, P_j \cap N, \dots) & \text{if } P_j \cap M \neq \emptyset, \\ (P_1, \dots, P_{j-2}, P_{j-1} \sqcup P_j, P_{j+1}, \dots) & \text{if } P_j \cap M = \emptyset. \end{cases}$$

Let $\phi : 2^{[r]} \rightarrow \mathbb{R}$ be a function which satisfies (8.3). If $P_j \cap M \neq \emptyset$, we see that $|\tau(\mathcal{P})| = |\mathcal{P}| + 1$ and by (8.3),

$$\tilde{\phi}(\tau(\mathcal{P})) = \phi(P_1) \cdots \phi(P_j \cap M) \phi(P_j \cap N) \cdots \phi(P_k) = \phi(P_1) \cdots \phi(P_k) = \tilde{\phi}(\mathcal{P}).$$

If $P_j \cap M = \emptyset$, we see that $|\tau(\mathcal{P})| = |\mathcal{P}| - 1$. In this case, we observe that $P_j \subset N$ and $i < j$ because $y_o \in P_i$, so that $P_{j-1} \cap N = \emptyset$ and $P_{j-1} \subset M$. Hence, by (8.3),

$$\tilde{\phi}(\tau(\mathcal{P})) = \phi(P_1) \cdots \phi(P_{j-2}) \phi(P_{j-1} \sqcup P_j) \cdots \phi(P_k) = \phi(P_1) \cdots \phi(P_k) = \tilde{\phi}(\mathcal{P}).$$

It is clear that the map $\tau : \mathfrak{P}_{[r]} \rightarrow \mathfrak{P}_{[r]}$ constructed in this manner satisfies $\tau \circ \tau = \text{id}$, which finishes the proof.

9. PROOF OF PROPOSITION 7.2

Lemma 9.1 (Estimating local effects of conditioning). *Fix $0 \leq \alpha < \beta$ and an integer $s > s_r + r$. Then, for any partition \mathcal{Q} of $[r]$, $\underline{h} \in \Delta_{\mathcal{Q}}(\alpha, \beta)$ and $f \in \mathcal{A}$,*

$$|\psi_{f, \underline{h}}(I) - \psi_{f, \underline{h}}^{\mathcal{Q}}(I)| \ll_{r, s} e^{-(\beta \delta_r - r \alpha \sigma_s)} N_s(f)^{|I|}, \quad \text{for all } I \subset [r],$$

where the implicit constant depends only on r and s .

9.1. Proof of Proposition 7.2 assuming Lemma 9.1

Fix a partition \mathcal{P} of $[r]$. Given $I \in \mathcal{P}$, we set

$$A_I = \psi_{f, \underline{h}}(I) - \psi_{f, \underline{h}}^{\mathcal{Q}}(I) \quad \text{and} \quad B_I = \psi_{f, \underline{h}}^{\mathcal{Q}}(I),$$

so that we can write

$$\tilde{\psi}_{f, \underline{h}}(\mathcal{P}) = \prod_{I \in \mathcal{P}} (A_I + B_I) \quad \text{and} \quad \tilde{\psi}_{f, \underline{h}}^{\mathcal{Q}}(\mathcal{P}) = \prod_{I \in \mathcal{P}} B_I.$$

We claim that

$$\left| \prod_{I \in \mathcal{P}} (A_I + B_I) - \prod_{I \in \mathcal{P}} B_I \right| \leq 2^r C, \quad (9.1)$$

where

$$C = \max \left\{ \prod_{I \in S} \prod_{J \in T} A_I B_J : \mathcal{P} = S \sqcup T, S \neq \emptyset \right\}.$$

Indeed, if one expands the first product in (9.1), one ends up with 2^r terms, one of which equals the product of all of the B_I 's. All other terms contains at least one A_I for some $I \in \mathcal{P}$ in them, and thus their absolute values are trivially estimated from above by C .

By Lemma 9.1, we have for all $s > s_r + r$,

$$|A_I| \ll_{r, s} e^{-(\beta \delta_r - r \alpha \sigma_s)} N_s(f)^{|I|},$$

and by (2.3) and (2.4),

$$|B_I| \ll_s \prod_{J \in \mathcal{Q}} \|f\|_{L^\infty}^{|\mathcal{I} \cap J|} = \|f\|_{L^\infty}^{|\mathcal{I}|} \ll_s N_s(f)^{|I|},$$

for every $I \in \mathcal{P}$. Hence,

$$A_I B_J \ll_{r, s} e^{-(\beta \delta_r - r \alpha \sigma_s)} N_s(f)^{|I|+|J|}, \quad \text{for all } I, J \subset [r] \text{ with } I \neq \emptyset.$$

Note that the bound in Proposition 7.2 is trivial (and useless) if $\beta \delta_r - r \alpha \sigma_s < 0$, so let us henceforth assume that $\beta \delta_r - r \alpha \sigma_s \geq 0$. We then get that

$$|C| \ll_{r, s} e^{-(\beta \delta_r - r \alpha \sigma_s)} N_s(f)^r,$$

which finishes the proof.

9.2. Proof of Lemma 9.1

We assume that $0 \leq \alpha < \beta$ have been fixed once and for all, as well as a partition \mathcal{Q} of $[r]$, along with a subset $I \subset [r]$. Pick $f \in \mathcal{A}$ and a tuple $\underline{h} = (h_1, \dots, h_r) \in \Delta_{\mathcal{Q}}(\alpha, \beta)$. We recall that the latter means that

$$d^{\mathcal{Q}}(\underline{h}) \leq \alpha \text{ and } d_{\mathcal{Q}}(\underline{h}) > \beta, \quad (9.2)$$

where $d^{\mathcal{Q}}$ and $d_{\mathcal{Q}}$ are defined in Subsection 6.2. Let

$$W_I = \{J \in \mathcal{Q} : I \cap J \neq \emptyset\},$$

and choose, for every $J \in W_I$, an index $i_J \in I \cap J$. For every $J \in W_I$, we now set

$$f_J = \prod_{j \in I \cap J} h_{i_j}^{-1} h_j \cdot f,$$

and note that

$$\psi_{f, \underline{h}}(I \cap J) = \mu(h_{i_J} \cdot f_J) = \mu(f_J) \quad \text{and} \quad \psi_{f, \underline{h}}(I) = \mu\left(\prod_{J \in W_I} h_{i_J} \cdot f_J\right).$$

In particular, by our convention that $\psi_{f, \underline{h}}(\emptyset) = 1$,

$$\psi_{f, \underline{h}}^{\mathcal{Q}}(I) = \prod_{J \in \mathcal{Q}} \psi_{f, \underline{h}}(I \cap J) = \prod_{J \in W_I} \mu(f_J).$$

Since the action $H \curvearrowright (X, \mu)$ is assumed to be exponentially mixing of all orders, we conclude by (2.1) with $k = |W_I|$ and $\underline{h}^{W_I} = (h_j)_{J \in W_I}$, that for all $s' > s_k$,

$$\begin{aligned} |\psi_{f, \underline{h}}(I) - \psi_{f, \underline{h}}^{\mathcal{Q}}(I)| &= \left| \mu\left(\prod_{J \in W_I} h_{i_J} \cdot f_J\right) - \prod_{J \in W_I} \mu(f_J) \right| \\ &\ll_{k, s'} e^{-\delta_k d_k(\underline{h}^{W_I})} \prod_{J \in W_I} N_{s'}(f_J), \end{aligned} \quad (9.3)$$

where

$$d_k(\underline{h}^{W_I}) = \min \{d(h_J, h_{J'}) : J, J' \in W_I, J \neq J'\},$$

and the implicit constants depend only on k and s' .

Let us now estimate the norms $N_{s'}(f_J)$ for $J \in W_I$. We fix $J \in W_I$, and suppose that $I \cap J$ contains at least two elements so that we can write $I \cap J = J' \sqcup \{j\}$ for some $j \in I \cap J$ and $J' \subset I \cap J$. Then, by (2.6) and (2.5), we have

$$\begin{aligned} N_{s'}(f_J) &= N_{s'}(f_{J'}(h_{i_j}^{-1} h_j \cdot f)) \ll_{s'} N_{s'+1}(f_{J'}) N_{s'+1}(h_{i_j}^{-1} h_j \cdot f) \\ &\ll_{s'} \exp(\sigma_{s'+1} d(h_{i_j}, h_j)) N_{s'+1}(f_{J'}) N_{s'+1}(f). \end{aligned}$$

If we iterate this argument as many times as there are elements in $I \cap J$, we arrive at the bound

$$N_{s'}(f_J) \ll_{s'} \exp(|I \cap J| \sigma_{s'+|J|} d^J(\underline{h})) N_{s'+|J|}(f)^{|I \cap J|},$$

where we used that (σ_s) is increasing, and thus, by (2.3),

$$\prod_{J \in W_I} N_{s'}(f_J) \ll_{s'} \exp\left(r \sigma_{s'+r} d^{\mathcal{Q}}(\underline{h})\right) N_{s'+r}(f)^{|I|}, \quad (9.4)$$

where d^J and d^Q are as in (6.3) and (6.4) respectively. Going back to (9.3), and using our assumption from Subsection 2 that the sequences (σ_s) and (s_k) are increasing and that the sequence (δ_k) is decreasing, we conclude from (9.4) that for all $s > s_r + r$, we have

$$|\psi_{f, \underline{h}}(I) - \psi_{f, \underline{h}}^Q(I)| \ll_{r,s} e^{-(\delta_r d_k(\underline{h}^{W_1}) - r\sigma_s d^Q(\underline{h}))} N_s(f)^{|I|}.$$

We have assumed that $\underline{h} \in \Delta_Q(\alpha, \beta)$, and thus

$$d^Q(\underline{h}) \leq \alpha \quad \text{and} \quad d_k(\underline{h}^{W_1}) \geq d_Q(\underline{h}) > \beta,$$

whence

$$\delta_r d_k(\underline{h}^{W_1}) - r\sigma_s d^Q(\underline{h}) > \delta_r \beta - r\sigma_s \alpha,$$

which finishes the proof.

10. PROOF OF PROPOSITION 6.2

We retain the conventions and notations which were set up in Subsection 6.2. In particular, we fix an integer $r \geq 3$ throughout the section, and write $[r]$ for the set $\{1, \dots, r\}$.

10.1. Passing to coarser partitions

If Q and \mathcal{R} are partitions of $[r]$, we say that \mathcal{R} is **coarser** than Q if every partition element in \mathcal{R} is a union of partition elements in Q , and **strictly coarser** if \mathcal{R} also has fewer partition elements than Q . In other words, \mathcal{R} is strictly coarser than Q if at least one partition element in \mathcal{R} is the union of at least two partition elements from Q . In particular, the partition $\{[r]\}$ into one single block is strictly coarser than any other partition of $[r]$, and every partition of $[r]$ with strictly less than r partition elements is strictly coarser than the partition \mathcal{K}_r of $[r]$ into points.

The following lemma summarizes the main inductive step in the proof of Proposition 6.2.

Lemma 10.1 (Passing to coarser partitions). *Let Q be a partition of $[r]$ with $|Q| \geq 2$. Fix $0 \leq \alpha < \beta$, and suppose that $\underline{h} \in H^r$ satisfies*

$$d^Q(\underline{h}) \leq \alpha \quad \text{and} \quad d_Q(\underline{h}) \leq \beta.$$

Then there exists a partition \mathcal{R} of $[r]$, strictly coarser than Q , such that

$$d^{\mathcal{R}}(\underline{h}) < 3\beta.$$

10.2. Proof of Proposition 6.2 assuming Proposition 10.1

Let us fix a sequence

$$0 = \beta_0 < \beta_1 < 3\beta_1 < \beta_2 < 3\beta_2 < \beta_3 < \dots < \beta_{r-1} < 3\beta_{r-1} < \beta_r.$$

Pick an element $\underline{h} = (h_1, \dots, h_r) \in H^r$. We wish to prove that either

- $\max_{i,j} d(h_i, h_j) \leq \beta_r$, or
- there exist $0 \leq k < r-1$ and a partition Q of $[r]$ with $|Q| \geq 2$ such that $\underline{h} \in \Delta_Q(3\beta_k, \beta_{k+1})$.

This will be done in several steps. We first check whether $\underline{h} \in \Delta_{\mathcal{K}_r}(0, \beta_1)$. If not, then

$$d^{\mathcal{K}_r}(\underline{h}) \leq 0 \quad \text{and} \quad d_{\mathcal{K}_r}(\underline{h}) \leq \beta_1,$$

and thus Lemma 10.1 (applied to $\alpha = 0$ and $\beta = \beta_1$) implies that there exists a strictly coarser partition Q_1 than \mathcal{K}_r such that $d^{Q_1}(\underline{h}) \leq 3\beta_1$. If $|Q_1| = 1$ (that is, $Q_1 = \{[r]\}$), then

$$\max_{i,j} d(h_i, h_j) \leq 3\beta_1 < \beta_r,$$

and we are done, so let us assume that $|\mathcal{Q}_1| \geq 2$. We now check whether $\underline{h} \in \Delta_{\mathcal{Q}_1}(3\beta_1, \beta_2)$. If not, then

$$d^{\mathcal{Q}_1}(\underline{h}) \leq 3\beta_1 \quad \text{and} \quad d_{\mathcal{Q}_1}(\underline{h}) \leq \beta_2,$$

and since $|\mathcal{Q}_1| \geq 2$, Lemma 10.1 (applied to $\alpha = 3\beta_1$ and $\beta = \beta_2$) implies that there exists a strictly coarser partition \mathcal{Q}_2 than \mathcal{Q}_1 such that $d^{\mathcal{Q}_2}(\underline{h}) \leq 3\beta_2$. If $|\mathcal{Q}_2| = 1$, then we again can conclude the argument as before, so we may assume that $|\mathcal{Q}_2| \geq 2$.

If we continue like this, then we will have produced a chain $\mathcal{K}_r, \mathcal{Q}_1, \dots, \mathcal{Q}_m$ of strictly coarser partitions of $[r]$, which eventually must terminate at the trivial partition $\{[r]\}$ in no more than r steps. At the k -th step, we check whether \underline{h} belongs to $\Delta_{\mathcal{Q}_k}(3\beta_k, \beta_{k+1})$. If this check fails for every k , then we conclude that $\max_{i,j} d(h_i, h_j) \leq \beta_r$.

10.3. Proof of Lemma 10.1

Let \mathcal{Q} be a partition of $[r]$ with $|\mathcal{Q}| \geq 2$ and $\underline{h} \in H^r$ satisfy $d^{\mathcal{Q}}(\underline{h}) \leq \alpha$ and $d_{\mathcal{Q}}(\underline{h}) \leq \beta$. Since $|\mathcal{Q}| \geq 2$, it follows from the second inequality that there exist atoms $I \neq J$ in \mathcal{Q} such that $d_{I,J}(\underline{h}) \leq \beta$. We consider the partition \mathcal{R} consisting of $I \cup J$ and $K \in \mathcal{Q} \setminus \{I, J\}$ which is strictly coarser than \mathcal{Q} . Since $d_{I,J}(\underline{h}) \leq \beta$, there exist $i_0 \in I$ and $j_0 \in J$ such that $d(h_{i_0}, h_{j_0}) \leq \beta$. Moreover, since $d^{\mathcal{Q}}(\underline{h}) \leq \alpha$, we have $d^I(\underline{h}) \leq \alpha$ and $d(h_i, h_{i_0}) \leq \alpha$ for all $i \in I$. Similarly, we conclude that $d(h_j, h_{j_0}) \leq \alpha$ for all $j \in J$. Hence, it follows that for all $i \in I$ and $j \in J$,

$$d(h_i, h_j) \leq d(h_i, h_{i_0}) + d(h_{i_0}, h_{j_0}) + d(h_{j_0}, h_j) \leq \beta + 2\alpha < 3\beta.$$

This proves that $d^{I \cup J}(\underline{h}) < 3\beta$. Additionally, for $K \in \mathcal{Q} \setminus \{I, J\}$,

$$d^K(\underline{h}) \leq d^{\mathcal{Q}}(\underline{h}) \leq \alpha < 3\beta.$$

Hence, we conclude that $d^{\mathcal{R}}(\underline{h}) < 3\beta$, as required.

REFERENCES

1. Ahlback, C.; Pippenger N.; Usatine, J. *A combinatorial interpretation of the joint cumulant*. Preprint, <https://arxiv.org/pdf/1211.0652.pdf>.
2. Björklund, M.; Einsiedler, M.; Gorodnik, A. *Quantitative multiple mixing*. Preprint, <https://arxiv.org/abs/1701.00945>
3. Björklund, M.; Gorodnik, A. *Central limit theorems in the geometry of numbers*. Preprint.
4. Björklund, M.; Zhang, G. *Positivity of integrals of spherical functions over Cartan subgroups*. In preparation.
5. Bowen, R. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Lecture Notes in Mathematics, Vol. 470. Springer-Verlag, Berlin-New York, 1975.
6. Chernov, N. I. *Limit theorems and Markov approximations for chaotic dynamical systems*. Prob. Th., Rel. Fields 101 (1995), 321–363.
7. Cohen G.; Conze, J-P. *CLT for Random Walks of Commuting Endomorphisms on Compact Abelian Groups*. J. Theoret. Probab. 30 (2017), no. 1, 143–195.
8. Cohen G.; Conze, J-P. *Central limit theorem for commutative semigroups of toral endomorphisms*. Preprint, <https://arxiv.org/abs/1304.4556>.
9. Cohen G.; Conze, J-P. *Almost mixing of all orders and CLT for some \mathbb{Z}^d -actions on subgroups of $\mathbb{F}_p^{\mathbb{Z}^d}$* . Preprint, <https://arxiv.org/abs/1609.06484>.
10. Conze, J.-P.; Le Borgne, S. *Méthode de martingales et flot géodésique sur une surface de courbure constante négative*. Ergodic Theory Dynam. Systems 21 (2001), no. 2, 421–441.
11. Derriennic, Y. *Some aspects of recent works on limit theorems in ergodic theory with special emphasis on the “central limit theorem”*. Discrete Contin. Dyn. Syst. 15 (2006), no. 1, 143–158.
12. Denker, M. *The central limit theorem for dynamical systems*. Banach Center Pub. 23 (1989), 33–61.

13. Dolgopyat, D. *Limit theorems for partially hyperbolic systems*.
Trans. Amer. Math. Soc. 356 (2004), no. 4, 1637–1689.
14. Flaminio, L.; Forni, G. *Invariant distributions and time averages for horocycle flows*.
Duke Math. J. 119 (2003), no. 3, 465–526.
15. Fréchet, M.; Shohat, J. *A proof of the generalized second-limit theorem in the theory of probability*.
Trans. Amer. Math. Soc. 33 (1931), no. 2, 533–543.
16. Gordin, M. I. *The central limit theorem for stationary processes*.
Dokl. Akad. Nauk SSSR 188 (1969), 739–741; English transl., Soviet Math. Dokl. 10 (1969), 1174–1176.
17. Gorodnik, A.; Spatzier, R. *Exponential mixing of nilmanifold automorphisms*.
J. Anal. Math. 123 (2014), 355–396.
18. Gouëzel, S. *Limit theorems in dynamical systems using the spectral method*.
Hyperbolic dynamics, fluctuations and large deviations, 161–193, Proc. Sympos. Pure Math., 89, Amer. Math. Soc., Providence, RI, 2015.
19. Guivarc’h Y.; Hardy, J. *Théorèmes limites pour une classe de chaînes de Markov et applications aux difféomorphismes d’Anosov*. Ann. Inst. H. Poincaré Probab. Statist. 24 (1988), no. 1, 73–98.
20. Le Borgne, S. *Limit theorems for non-hyperbolic automorphisms of the torus*.
Israel J. Math. 109 (1999), 61–73.
21. Le Borgne, S. *Principes d’invariance pour les flots diagonaux sur $SL(d, \mathbb{R})/SL(d, \mathbb{Z})$* .
Ann. Inst. H. Poincaré Probab. Statist. 38 (2002), no. 4, 581–612.
22. Le Borgne, S. *Martingales in hyperbolic geometry*.
Analytic and probabilistic approaches to dynamics in negative curvature, 1–63, Springer INdAM Ser., 9, Springer, Cham, 2014.
23. Le Jan, Y. *The central limit theorem for the geodesic flow on noncompact manifolds of constant negative curvature*.
Duke Math. J. 74 (1994), no. 1, 159–175.
24. Leonov, V. P. *On the central limit theorem for ergodic endomorphisms of compact commutative groups*. Dokl. Akad. Nauk SSSR 135 (1960), 258–261.
25. Liverani, C. *Central limit theorem for deterministic systems*.
in International conference on dynamical systems (Eds. F. Ledrappier, J. Lewowics & S. Newhouse), Pitman Res. Notes 363 (1996) 56–75.
26. Melbourne, Ian; Török, A. *Central limit theorems and invariance principles for time-one maps of hyperbolic flows*.
Comm. Math. Phys. 229 (2002), no. 1, 57–71.
27. Ratner, M. *The central limit theorem for geodesic flows on n -dimensional manifolds of negative curvature*.
Israel J. Math. 16 (1973), 181–197.
28. Sinai, Y. G. *The central limit theorem for geodesic flows on manifolds of constant negative curvature*.
Soviet Math. Dokl. 1 (1960), 938–987.
29. Speed, T. P. *Cumulants and Partition Lattices*.
Austral. J. Statist., 25(2) (1983), 378–388.
30. Viana, M. *Stochastic dynamics of deterministic systems*.
Lecture Notes XXI Bras. Math. Colloq. IMPA, Rio de Janeiro, 1997.
31. Volný, D. *A central limit theorem for fields of martingale differences*.
Preprint, <https://arxiv.org/abs/1504.02439>.
32. Young, L.-S. *Statistical properties of dynamical systems with some hyperbolicity*.
Ann. Math. 147 (1998), 585–650.

DEPARTMENT OF MATHEMATICS, CHALMERS, GOTHENBURG, SWEDEN
E-mail address: micbjo@chalmers.se

UNIVERSITY OF BRISTOL, BRISTOL, UK
E-mail address: a.gorodnik@bristol.ac.uk