SHARP ERGODIC THEOREMS FOR GROUP ACTIONS AND STRONG ERGODICITY

ALEX FURMAN AND YEHUDA SHALOM

Abstract. Let μ be a probability measure on a locally compact group G, and suppose G acts measurably on a probability measure space (X, m), preserving the measure m. We study ergodic theoretic properties of the action along μ-i.i.d. random walks on G. It is shown that under a (necessary) spectral assumption on the μ-averaging operator on $L^2(X, m)$, almost surely the mean and the pointwise (Kakutani's) random ergodic thorems have roughly $n^{-\frac{1}{2}}$ rate of convergence. We also prove a central limit theorem for the pointwise convergence. Under a similar spectral condition on the diagonal G-action on $(X \times X, m \times m)$, an almost surely exponential rate of mixing along random walks is obtained.

The imposed spectral condition is shown to be connected to a strengthening of the ergodicity property, namely, the uniqueness of m-integration as a G-invariant mean on $L^\infty(X, m)$. These related conditions, as well as the presented sharp ergodic theorems, never occur for amenable G. Nevertheless, we bring many natural examples, among them automorphism actions on tori and actions on Lie groups' homogeneous spaces, for which our results can be applied.

1. Introduction and statement of the main results

Throughout this paper G denotes a locally compact group, acting measurably on a probability space $(X, \mathcal{B}, m)$, preserving the measure m. The action is assumed to be ergodic, i.e. there are no G-invariant measurable subsets of X, except subsets of measure zero or one.

Classical ergodic theory, which studies the behavior of a single measure preserving transformation, has been widely extended through the last two decades to actions of amenable groups (see e.g. [OW1-3]). The existence of Főlner sets in amenable groups enables natural averaging, which makes these groups good candidates for such an extension. However, for general locally compact groups one should seek a different approach, and the one we shall explore in the present paper is based upon the concept of a random walk on the group.

Let μ be a probability measure on G. We shall consider the probability space of the random walk $(\Omega, \mathcal{P}) = (G^{\mathbb{N}}, \mu^{\mathbb{N}})$, and for $\omega = (\omega_1, \omega_2 ...) \in \Omega$ refer to the sequence of products $g_n^\omega = \omega_n \cdots \omega_1$ as a sequence of *random products*. In the sequel we shall impose different conditions on a measure μ on the group G: μ is said to be *symmetric* if $\check{\mu} = \mu$ (here $\check{\mu}(E) = \mu(E^{-1})$ for measurable $E \subset G$); μ is said to be *generating* if it is not supported on a proper closed subgroup of G; and is said to be *aperiodic* if it is not supported on a coset of a proper closed subgroup of G. Note that μ is aperiodic iff $\check{\mu} * \mu$ is generating.

Date: April 1, 1998.

Theorem (Kakutani, [Ka]**).** *Suppose G acts ergodically on (X, m), and assume that μ is a generating probability measure on G. Then for any function $f \in L^p(X, m)$, $p \geq 1$, for $\mathcal{P}$-a.e. $\omega \in \Omega$:*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(g_i^{\omega}\, x) = \int_X f\, dm \tag{1.1}$$

for m-a.e. $x \in X$.

The statement follows from Birkhoff's ergodic theorem, applied to the single transformation T, acting on the product space $(\Omega \times X, \mathcal{P} \times m)$ by $T(\omega,\, x) = (\theta\omega,\, \omega_1 x)$ where $\theta : \Omega \to \Omega$ is the shift: $(\theta\omega)_i = \omega_{i+1}$. Under the assumptions of the theorem, the skew-product $(\Omega \times X, \mathcal{P} \times m, T)$ is known to be ergodic. In fact, under some mild conditions on the distribution μ, this skew-product is an exact transformation. This (incidental) result is proved in Appendix B.

An intrinsic feature of the ergodic theory of amenable actions is the lack of any prescribed *rate of convergence*, valid for *all* bounded functions ([Kr], [DJR]). In the present paper we shall present a family of non-amenable actions, where Kakutani's random ergodic theorem can be sharpened: a *universal* rate of convergence holds for *all* L^2-functions. To formulate the precise result, we shall need some notations: consider the unitary G-representation π on the space $L^2_0(X, m)$, defined by $\pi(g)f(x) = f(g^{-1}x)$, where $g \in G$, $f \in L^2_0(X, m)$. Hereafter $L^2_0(X, m)$ denotes the space of zero-mean functions in $L^2(X, m)$. Given a probability distribution μ on G, recall that the μ-convolution operator $\pi(\mu)$ is defined on $L^2_0(X, m)$ by:

$$\langle \pi(\mu)u,\, v\rangle = \int_G \langle \pi(g)u,\, v\rangle\, d\mu(g). \tag{1.2}$$

Observe that its operator norm always satisfies $\|\pi(\mu)\| \leq 1$, and we shall be interested in the situation where $\|\pi(\mu)\| < 1$. This spectral gap condition is typical for actions of *non-amenable* groups; in fact, it never holds for actions of amenable groups, while for discrete groups with Kazhdan's property (T) (see Definition 6.2), it is always satisfied (assuming μ is aperiodic on G). Let us postpone further discussion to section 6, where precise statements and other examples are considered, and concentrate on some consequences of this spectral condition:

Theorem 1.1 (Rate of L^2-Convergence). *Suppose G acts ergodically on a probability space (X, m), let μ be a probability distribution on G, and assume that $\|\pi(\mu)\| < 1$. Let $\{a_n\}_{n=1}^{\infty}$ be a decreasing sequence with $\sum_{n=1}^{\infty} a_n^2 < \infty$. Then for any function $f \in L^2_0(X, m)$, for $\mathcal{P}$-a.e. $\omega \in \Omega$:*

$$\lim_{n\to\infty} a_n \cdot \|\sum_{i=1}^{n} f(g_i^{\omega} x)\| = 0.$$

In particular, for any function $f \in L^2(X, m)$ and every $\epsilon > 0$, for $\mathcal{P}$-a.e. $\omega \in \Omega$:

$$\|\frac{1}{n} \sum_{i=1}^{n} f(g_i^{\omega} x) - \int_X f\, dm\| = o\left(\frac{\log^{\frac{1}{2}+\epsilon} n}{\sqrt{n}}\right).$$

For a proof see section 2. We remark that a spectral assumption is necessary for *any* rate of convergence to hold for *all* functions in $L^2(X,m)$ (see Remark 2.4). Moreover, the rate of convergence in the above result is essentially optimal, for by the Law of the Iterated Logarithm there exist actions for which the averages exceed $c\,(\sqrt{\log\log n}/\sqrt{n})$ infinitely many times.

By a completely different method, a similar rate of convergence to the one asserted in Theorem 1.1 is established in section 3 for the pointwise convergence:

Theorem 1.2 (Rate of Pointwise Convergence). *Let G,μ and (X,m) be as in 1.1 and let $\{a_n\}_{n=1}^{\infty}$ be a monotone sequence with $\sum_{n=1}^{\infty}|a_n\log n|^2<\infty$. Then for all $f\in L^2_0(X,m)$ for $\mathcal{P}$-a.e. $\omega\in\Omega$ and for m-a.e. $x\in X$:*

$$\lim_{n\to\infty} a_n\cdot\sum_{i=1}^{n} f(g_i^{\omega}x)=0$$

In particular, for any $f\in L^2(X,m)$ and any $\epsilon>0$, for $\mathcal{P}$-a.e. $\omega\in\Omega$ and m-a.e. $x\in X$ one has:

$$\left|\frac{1}{n}\sum_{i=1}^{n} f(g_i^{\omega}\,x)-\int_X f\,dm\right|=o\left(\frac{\log^{\frac{3}{2}+\epsilon} n}{\sqrt{n}}\right).$$

Theorems 1.1-1.2 show that, assuming $\|\pi(\mu)\|<1$, the rate of convergence in (1.1) is roughly $n^{-\frac{1}{2}}$. This suggests to estimate the rate of convergence, normalized by $\sqrt{n}$. In this context we obtain the following:

Theorem 1.3 (Functional Central Limit Theorem). *Let G,μ and (X,m) be as in Theorem 1.2 and let $0\neq f\in L^p(X,m)$, $p>2$, be a real valued function. Define*

$$S_n(\omega,x)=\sum_{i=1}^{n} f(g_i^{\omega}\,x),\qquad \sigma_n=\|S_n\|_2$$

on $(\Omega\times X,\mathcal{P}\times m)$. Then $\sigma=\lim_{n\to\infty}n^{-1/2}\cdot\sigma_n$ is non zero, the sequence of random processes $\sigma_n^{-1}\cdot S_{[nt]}$, $t\in[0,1]$ converges in the Skorohod topology to the standard Brownian Motion $W(t)$ on $[0,1]$, and the following estimate holds

$$\sup_{a\in[0,1]}\left|\mathcal{P}\times m\left(\{\sigma_n^{-1}\cdot S_n<a\}\right)\;-\;\frac{1}{\sqrt{2\pi}\cdot\sigma}\int_{-\infty}^{a} e^{-t^2/2\sigma^2}\,dt\right|=O\left(\frac{\log^{\frac{p}{2}} n}{n^{\delta}}\right),$$

where $\delta=\min\{\frac{1}{2},\frac{p-1}{2}\}$.

Let us discuss now mixing properties along random products. Obviousely we have to assume that $(\pi,L^2_0(X,m))$ has no finite dimensional sub-representations, which is known to be equivalent to the condition $1_G\not\subseteq\pi\otimes\pi$, i.e. ergodicity of the G-action on $(X\times X,m\times m)$. Motivated by Theorems 1.1-1.3, we shall impose the spectral condition $\|\pi\otimes\pi\,(\mu)\|<1$ on the G-action on (X,m). Again we remark, that this condition never holds for actions of amenable groups, while for discrete groups with Kazhdan's property (T), it is equivalent to weak mixing (assuming μ is aperiodic on G).

Theorem 1.4. *Suppose G acts on a probability space $(X, \mathcal{B}, m)$, μ is a probability distribution on G, and that $\|\pi \otimes \pi(\mu)\| < 1$. Then there exists a conull set $\Omega' \subset \Omega$, such that for each $\omega \in \Omega'$ the sequence of random products $\{g_n^\omega\}_{n=1}^\infty$ is mixing, namely, for any $A, B \in \mathcal{B}$ one has $m(g_n^\omega A \cap B) \to m(A) \cdot m(B)$. Moreover, for any λ with $\sqrt{\|\pi \otimes \pi(\mu)\|} = \lambda_0 < \lambda \leq 1$, and any $A, B \in \mathcal{B}$, there is a full measure subset $\Omega_{A,B} \subset \Omega$, such that for all $\omega \in \Omega_{A,B}$ one has:*

$$\lim_{n \to \infty} \frac{m(g_n^\omega A \cap B) - m(A) \cdot m(B)}{\lambda^n} = 0$$

The assumption made in the sharp ergodic theorems is a spectral one. However it is very closely related to a natural strengthening of ergodicity:

Definition 1.5. Let G be a locally compact group acting measurably on a probability measure space (X, m), where m is G-invariant. The G-action on (X, m) is said to be *strongly ergodic* if m-integration is the unique G-invariant mean (i.e. continuous positive normalized functional) on $L^\infty(X, m)$.

This definition is slightly stronger than the one given by A. Connes and B. Weiss in [CW] (K. Schmidt showed in [Sc] that they are not equivalent). Notice that the ergodicity assumption is equivalent to the assertion that m-integration is the unique G-invariant mean on $L^p(X, m)$ for any $1 \leq p < \infty$.

A probability measure μ on a locally compact group G is said to be *measurably aperiodic* if μ is not supported on a coset of a proper measurable subgroup of G. For example, if μ is absolutely continuous with respect to the Haar measure on a connected locally compact group G, then it is measurably aperiodic. The following result, which is proved in section 5, connects strong ergodicity and our spectral assumption.

Theorem 1.6. *Assume that G and (X, m) are as in 1.5 and let μ be a* measurably aperiodic *probability distribution on G. If the G-action is strongly ergodic, then $\|\pi(\mu)\| < 1$.*

For countable G Theorem 1.6 was proved in [Sc]. In fact the two conditions are then equivalent, but this is not the case in general. Thus the question of whether an action is strongly ergodic has ergodic theoretic implications. This problem was however investigated in its own right by many authors (e.g. in relation to the Ruziewicz problem). In section 6 we shall recall briefly several known examples of such actions, and present some new ones:

Theorem 1.7. *Let $G \subseteq \mathrm{SL}_d(\mathbb{Z})$ and suppose that G acts irreducibly on $\mathbb{R}^d$ and that G does not contain an abelian subgroup of finite index. Then the G-action on the d-torus $\mathbb{R}^d/\mathbb{Z}^d$ is strongly ergodic.*

Theorem 1.8. *Let $G = \Pi G_i$ be a semisimple Lie group with finite center, $\Gamma \subset G$ be a lattice, and let $H \subseteq G$ be a subgroup. Suppose that the projection of H to every simple factor G_i does not have an amenable closure. Then the H-action on G/Γ is strongly ergodic.*

Acknowledgments. We would like to extend our deep gratitude to Hillel Furstenberg and Benjamin Weiss for many helpful discussions.

2. Sharp Mean Ergodic Theorem

In this section we prove Theorem 1.1. In fact, it will be natural to consider this result in the general framework of unitary representations:

Theorem 2.1. *Let $(\pi, \mathcal{H})$ be a unitary G-representation, μ a probability distribution on G, and assume that $\|\pi(\mu)\| < 1$. Then for any decreasing l^2 sequence $\{a_n\}$, and any vector $u \in \mathcal{H}$, there exists a conull set $\Omega_u \subset \Omega$, such that for all $\omega \in \Omega_u$:*

$$\lim_{n\to\infty} a_n \cdot \|\sum_{i=1}^{n} \pi(g_i^\omega)\, u\| = \lim_{n\to\infty} a_n \cdot \|\sum_{i=1}^{n} \pi(g_i^\omega)^{-1}\, u\| = 0.$$

We shall consider both the action of random products by $\pi(g_i^\omega)$, and the action of their inverses by $\pi(g_i^\omega)^{-1}$, because the former is natural in the abstract unitary representation setup, while the latter appears in representations coming from G-actions (cf. Theorem 1.1): $\pi(g_i^\omega)^{-1} f(x) = f(g_i^\omega x)$.

For a fixed $u \in \mathcal{H}$, define the random variables S_k and S_k^* on $(\Omega, \mathcal{P})$ by:

$$S_k(\omega) = \|\sum_{i=1}^{k} \pi(g_i^\omega) u\|, \qquad S_k^*(\omega) = \|\sum_{i=1}^{k} \pi(g_i^\omega)^{-1} u\|$$

Our theorem will be deduced from the following estimates:

Lemma 2.2. *There exists a constant C_1, such that for every $N \geq 1$*

$$\mathbf{E}\left(S_N^2\right) \leq C_1 \cdot N, \qquad \mathbf{E}\left((S_N^*)^2\right) \leq C_1 \cdot N. \tag{2.1}$$

Lemma 2.3. *There exists a constant C_2, such that for any $k < N$ the conditional expectations satisfy:*

$$\mathbf{E}\left(S_N^2 \;\middle|\; \omega_1, \ldots, \omega_k\right) \geq S_k^2 - C_2 \cdot S_k \tag{2.2}$$

and a similar inequality for S_k^*. For S_k (resp. S_k^*) sufficiently large, the right hand side dominates $S_k^2/2$ (resp. $(S_k^*)^2/2$).

Proof of Theorem 2.1 based on 2.2 and 2.3. Fix some $\epsilon > 0$ and let $G_{k,\epsilon}$, $G_{k,\epsilon}^*$ denote the events:

$$G_{k,\epsilon} = \{\omega \mid a_k\, S_k(\omega) > \epsilon\}, \qquad G_{k,\epsilon}^* = \{\omega \mid a_k\, S_k^*(\omega) > \epsilon\}.$$

We will show that with probability 1 the events $G_{k,\epsilon}$, $G_{k,\epsilon}^*$ occur only finitely many times. Since $\epsilon > 0$ is arbitrary, this will prove the theorem. Hereafter we focus on $G_{k,\epsilon}$, but the same arguments apply to $G_{k,\epsilon}^*$.

Consider the intervals $2^n \leq k \leq 2^{n+1}$, and define the events

$$E_{n,\epsilon} = \bigcup_{k=2^n}^{2^{n+1}} G_{k,\epsilon} = \dot{\bigcup_{k=2^n}^{2^{n+1}}} F_{k,\epsilon}, \quad \text{where} \quad F_{k,\epsilon} = G_{k,\epsilon} \setminus \bigcup_{j=2^n}^{k-1} G_{j,\epsilon}$$

It is enough to show that with probability 1 the events $\{E_{n,\epsilon}\}_n$ occur only finitely many times, and the latter will follow from the Borel-Cantelli Lemma, as soon as we prove:

$$\sum_{n=1}^{\infty} \mathcal{P}(E_{n,\epsilon}) < \infty. \tag{2.3}$$

Observe that $F_{k,\epsilon}$ is measurable with respect to $\omega_1, \ldots, \omega_k$, and $\omega \in F_{k,\epsilon}$ implies $S_k(\omega) > \epsilon/a_k$. Hence S_k is large on $F_{k,\epsilon}$, for large k. Therefore (2.2) gives, using the fact that $\{a_i\}$ is monotonic:

$$\frac{1}{\mathcal{P}(F_{k,\epsilon})} \int_{F_{k,\epsilon}} S^2_{2^{n+1}} \, d\mathcal{P} = \mathbf{E} \left(S^2_{2^{n+1}} \;\middle|\; F_{k,\epsilon} \right) \geq \frac{\epsilon^2}{2\, a_k^2} \geq \frac{\epsilon^2}{2\, a_{2^n}^2}.$$

Since $E_{n,\epsilon}$ is a disjoint union of the $F_{k,\epsilon}$-s, we obtain, using (2.1):

$$\begin{aligned}
\mathcal{P}(E_{n,\epsilon}) &= \sum_{k=2^n}^{2^{n+1}} \mathcal{P}(F_{k,\epsilon}) \leq \frac{2\, a_{2^n}^2}{\epsilon^2} \sum_{k=2^n}^{2^{n+1}} \int_{F_{k,\epsilon}} S^2_{2^{n+1}} \, d\mathcal{P} \\
&\leq \frac{2\, a_{2^n}^2}{\epsilon^2} \, \mathbf{E} \left(S^2_{2^{n+1}} \right) \leq \frac{2\, a_{2^n}^2}{\epsilon^2} \, C_1 \, 2^{n+1} = O(2^n \cdot a_{2^n}^2)
\end{aligned}$$

The monotone series $\sum a_n^2$ converges together with $\sum 2^n \, a_{2^n}^2$, thereby proving (2.3). This completes the proof of Theorem 2.1, and thus also Theorem 1.1. □

It remains to prove Lemmas 2.2 and 2.3.

Proof of Lemma 2.2. We estimate the expectation $\mathbf{E}\left(S_N^2\right) = \sum_{i,j}^{N} \mathbf{E} \langle \pi(g_i^\omega)u, \, \pi(g_j^\omega)u \rangle$ by:

$$\begin{aligned}
\mathbf{E}\left(S_N^2\right) &\leq \sum_{j=1}^{N} \|\pi(g_j^\omega)\| + 2 \sum_{j=1}^{N} \sum_{r=1}^{N-j} \left| \mathbf{E} \langle \pi(\omega_{j+r} \cdots \omega_{j+1}) \pi(g_j^\omega) u, \, \pi(g_j^\omega) u \rangle \right| \\
&\leq N \cdot \|u\|^2 + 2 \sum_{j=1}^{N} \sum_{r=1}^{N-j} \|\pi(\mu)\|^r \, \|u\|^2 \leq C_1 \cdot N
\end{aligned}$$

where $C_1 = (1 + 2\, \|\pi(\mu)\| + 2\, \|\pi(\mu)\|^2 + \ldots) \cdot \|u\|^2 < \infty$.

To estimate $\mathbf{E}\left((S_N^*)^2\right)$ we use similar inequalities, replacing $\left|\mathbf{E} \langle \pi(g_{j+r}^\omega)u, \, \pi(g_j^\omega)u \rangle\right|$ by:

$$\begin{aligned}
\left| \mathbf{E} \langle \pi(g_{j+r}^\omega)^{-1} u, \, \pi(g_j^\omega)^{-1} u \rangle \right| &= \left| \mathbf{E} \langle \pi(g_j^\omega)^{-1} \pi(\omega_{j+1}^{-1} \cdots \omega_{j+r}^{-1}) u, \, \pi(g_j^\omega)^{-1} u \rangle \right| \\
\left| \mathbf{E} \langle \pi(\omega_{j+1}^{-1} \cdots \omega_{j+r}^{-1}) u, \, u \rangle \right| &= \left| \langle \pi(\mu)^r u, \, u \rangle \right| \leq \|\pi(\mu)\|^r \cdot \|u\|^2
\end{aligned}$$

thereby obtaining $\mathbf{E}\left((S_N^*)^2\right) \leq C_1 \cdot N$. This proves Lemma 2.2. □

Proof of Lemma 2.3. Let us fix some $k < N$ and define the random vectors:

$$Z = Z(\omega) = \sum_{j=1}^{k} \pi(g_j^\omega) u, \quad W = W(\omega) = \sum_{r=1}^{N-k} \pi(g_{k+r}^\omega) u = \sum_{r=1}^{N-k} \pi(\omega_{k+r} \cdots \omega_{k+1}) \, \pi(g_k^\omega) u.$$

Note that Z is measurable with respect to $\omega_1, \ldots, \omega_k$ and $\|Z\| = S_k$. Therefore:

$$\begin{aligned}
\mathbf{E}\left(S_N^2 \mid \omega_1, \ldots, \omega_k\right) &= \mathbf{E}\left(\|Z\|^2 + \|W\|^2 + 2\, Re\langle W, Z\rangle \mid \omega_1, \ldots, \omega_k\right) \\
&\geq S_k^2 - 2 \sum_{r=1}^{N-k} \left|\mathbf{E}\left(\langle \pi(\omega_{k+r} \cdots \omega_{k+1})\, \pi(g_k^\omega)u,\, Z\rangle \mid \omega_1, \ldots, \omega_k\right)\right| \\
&\geq S_k^2 - 2 \sum_{r=1}^{N-k} \|\pi(\mu)\|^r \cdot \|u\| \cdot \|Z\| \geq S_k^2 - C_2 \cdot S_k
\end{aligned}$$

where $C_2 = 2 \cdot (\|\pi(\mu)\| + \|\pi(\mu)\|^2 + \ldots) \cdot \|u\|$.

Estimating $\mathbf{E}\left((S_N^*)^2 \mid \omega_1, \cdots, \omega_k\right)$, we use the random vectors:

$$Z = \sum_{j=1}^{k} \pi(g_j^\omega)^{-1} u, \quad W = \sum_{r=1}^{N-k} \pi(g_{k+r}^\omega)^{-1} u = \sum_{r=1}^{N-k} \pi(g_k^\omega)^{-1}\, \pi(\omega_{k+1}^{-1} \cdots \omega_{k+r}^{-1})u,$$

and use similar inequalities, replacing $\left|\mathbf{E}\left(\langle \pi(g_{k+r}^\omega \cdots \omega_{k+1})u,\, Z\rangle \mid \omega_1, \ldots, \omega_k\right)\right|$ by:

$$\begin{aligned}
&\left|\mathbf{E}\left(\langle \pi(g_k^\omega)^{-1} \pi(\omega_{k+1}^{-1} \cdots \omega_{k+r}^{-1})u,\, Z\rangle \mid \omega_1, \ldots, \omega_k\right)\right| \\
&= \left|\mathbf{E}\left(\langle \pi(\omega_{k+1}^{-1} \cdots \omega_{k+r}^{-1})u,\, \pi(g_k^\omega)Z\rangle \mid \omega_1, \ldots, \omega_k\right)\right| \\
&\leq |\langle \pi(\check{\mu})^r u,\, \pi(g_k^\omega)Z\rangle| \leq \|\pi(\check{\mu})\|^r \cdot \|u\| \cdot \|Z\|.
\end{aligned}$$

Since $\|\pi(\check{\mu})\| = \|\pi(\mu)^*\| = \|\pi(\mu)\| < 1$ we get the same estimation:

$$\mathbf{E}\left((S_N^*)^2 \mid \omega_1, \cdots, \omega_k\right) \geq (S_k^*)^2 - C_2 \cdot S_k^*.$$

The proof of Lemma 2.3 is completed. □

Remark 2.4. The assumption $\|\pi(\mu)\| < 1$ is essentially necessary to achieve any rate of convergence. More precisely, in [FS] the following statement is proved: suppose that for all $u \in \mathcal{H}$: $\|n^{-1} \cdot \sum_{i=1}^n \pi(g_i^\omega)u\| = o(a_n)$ for some fixed sequence $a_n \to 0$, then $1_G \not\prec \pi$ (see definition 6.1). The conditions $1_G \not\prec \pi$ and $\|\pi(\mu)\| < 1$ are closely related, and are often equivalent (see Theorems 6.3 and **??** below).

3. Sharp Pointwise Ergodic Theorem and CLT

Sharp Pointwise Ergodic Theorem. In section 2 a prescribed rate of convergence was established for the *mean* ergodic theorem. In this section we prove a pointwise ergodic theorem, with essentially the same rate of convergence, but using a completely different approach.

Theorem 1.1 states, roughly, that if G, μ and (X, m) satisfy $\|\pi(\mu)\| < 1$, then almost surely the norm $\| \sum_{i=1}^n f(g_i^\omega x)\|$ is of order $\sqrt{n}$, which intimates that the vectors $\{\pi(g_n^\omega) f\}_n$ are "almost" orthogonal. This property suggests we apply an argument, due to Kac, Salem and Zygmund, on pointwise convergence of *quasi-orthogonal* sequences, which we shall describe below.

Definition 3.1. A sequence $\{\phi_n\}_{n=1}^{\infty}$ in a Hilbert space $\mathcal{H}$ is said to be *quasi-orthogonal*, if the following equivalent conditions are satisfied:

(i) The bi-infinite matrix $\{\langle\phi_i,\, \phi_j\rangle\}_{i,j}$ has a bounded norm.

(ii) There exists a constant $L < \infty$, so that for any $\psi \in \mathcal{H}$ the following Bessel's type inequality is satisfied: $\sum_n |\langle\phi_n,\, \psi\rangle|^2 \leq L\,\|\psi\|^2$.

(iii) For every l^2 sequence $\{c_n\}_{n=1}^{\infty}$, the series $\sum_n c_n\,\phi_n$ converges in $\mathcal{H}$.

The equivalence of the above conditions is a standard exercise in bi-linear forms on Hilbert spaces. One also easily verifies that a sequence of vectors $\{\phi_n\}$ satisfying

$$|\langle\phi_{n+k},\, \phi_n\rangle| \leq C\cdot\lambda^k \qquad n,k\geq 1 \tag{3.1}$$

for some $C < \infty$ and $\lambda < 1$, has property (iii), and is therefore quasi-orthogonal.

Proof of Theorem 1.2. Following Kac-Salem-Zygmund ([KSZ]), we observe that if $\{\phi_n(y)\}$ is a quasi-orthogonal sequence in $L^2(Y,\nu)$, where (Y,ν) is a probability space, then property (ii) in Definition 3.1 suffices to prove the following beautiful Rademacher-Menshov maximal inequality (which was originally stated for *orthonormal* sequences):

Theorem (see [Zy] **pp. 189–193).** *There exists a constant $K < \infty$, depending only on $\{\phi_n(y)\}_{n=1}^{\infty}$, so that for any sequence $\{a_n\}_{n=1}^{\infty}$:*

$$\int \sup_{1\leq k<\infty} \left|\sum_{n=1}^{k} a_n\,\phi_n(y)\right|^2 d\nu(y) \leq K\cdot\sum_{n=1}^{\infty}|a_n\,\log n|^2. \tag{3.2}$$

It is not difficult to verify that (3.2) implies the ν-a.e. convergence of the series $\sum_{n=1}^{\infty} a_n\,\phi_n(y)$ given any sequence $\{a_n\}$ with $\sum_n |a_n\,\log n|^2 < \infty$.

Now, let $\{a_n\}$ be a monotone positive sequence with $\sum_n |a_n \log n|^2 < \infty$, and $f \in L_0^2(X,m)$. Considering f as a function on the skew-product $(Y,\nu) = (\Omega\times X, \mathcal{P}\times m)$, we assert that the sequence

$$\phi_n(\omega,x) = f\circ T^n(\omega,x) = f(\omega_n\cdots\omega_1 x)$$

satisfies estimation (3.1), and is therefore quasi-orthogonal. Indeed

$$\begin{aligned} |\langle\phi_{n+k},\, \phi_n\rangle| &= |\langle f\circ T^k,\, f\rangle| = \left|\int_X\int_\Omega f(x)\cdot\bar{f}(\omega_k\cdots\omega_1 x)\,d\mathcal{P}(\omega)\,dm(x)\right| \\ &= |\langle\pi(\mu)^k f,\, f\rangle| \leq \|\pi(\mu)\|^k\cdot\|f\|^2 \end{aligned}$$

Therefore the functions ϕ_n have exponential decay of correlations (3.1), with $\lambda = \|\pi(\mu)\| < 1$.

Next, applying inequality (3.2) to $\phi_n = f\circ T^n$, we conclude that for $\mathcal{P}$-a.e. $\omega\in\Omega$ and m-a.e. $x\in X$ the series: $\sum a_n\,\phi_n(\omega,x) = \sum a_n\,f(g_n^\omega\,x)$ converges. Finally, recall Kronecker's lemma: if $\{a_n\}$ is monotonic and $\sum_n a_n\,s_n$ converges, then
$\lim_{n\to\infty} a_n\,|\sum_{i=1}^n s_i| = 0$. Hence we obtain for $\mathcal{P}$-a.e. $\omega\in\Omega$ and m-a.e. $x\in X$:

$$\lim_{n\to\infty} a_n\left|\sum_{i=1}^{n} f(g_i^\omega\,x)\right| = 0$$

thereby proving Theorem 1.2. □

Two auxiliary results for the CLT. For the proof of the Central Limit Theorem 1.3, we shall need two results on the structure of the skew-product $(\Omega \times X, \mathcal{P} \times m, T)$, which may have an independent interest.

Consider the the space $\bar\Omega$ of bi-infinite sequences with the product measure $\bar{\mathcal{P}} = \mu^{\mathbb{Z}}$ and the shift θ, which is the natural extension of $(\Omega, \mathcal{P}, \theta)$. The skew-product $(\bar\Omega \times X, \bar{\mathcal{P}} \times m, T)$, defined by $T(\omega, x) = (\theta\omega, \omega_1 x)$ is the natural extension of $(\Omega \times X, \mathcal{P} \times m, T)$. Consider the σ-algebra $\mathcal{B}$ of all measurable sets of the form $\bar\Omega \times B$, where $B \subset X$ is measurable.

Theorem 3.2. *Let G, μ and (X, m) be as in Theorem 1.2. Then the σ-algebra $\mathcal{B}$ has the following property: The coefficients $\rho(n)$ defined by*

$$\rho(n) = \sup \left\{ \frac{|\langle f, g\rangle|}{\|f\| \cdot \|g\|} \;\mid\; f \in L^2_0(\bigvee_{i=1}^{\infty} T^i\mathcal{B}),\; g \in L^2_0(\bigvee_{-\infty}^{-n} T^i\mathcal{B}) \right\}, \tag{3.3}$$

decay exponentially fast: $\rho(n) \le \|\pi(\mu)\|^n$.

The theorem is a direct consequence of the following:

Lemma 3.3. *Let $f, g \in L^2_0(\bar\Omega \times X, \bar{\mathcal{P}} \times m)$ be such that $f \in L^2_0(\bigvee_{i=1}^{\infty} T^i\mathcal{B})$ and $g \in L^2_0(\bigvee_{-\infty}^{-n} T^i\mathcal{B})$, where $n \ge 1$. Then*

$$|\langle f, g\rangle| \le \|\pi(\check\mu)^n P_X f\| \cdot \|g\| \le \|\pi(\mu)\|^n \cdot \|f\| \cdot \|g\|. \tag{3.4}$$

where $P_X : L^2_0(\bar\Omega \times X, \bar{\mathcal{P}} \times m) \to L^2_0(X, m)$ denotes the orthogonal projection:

$$(P_X f)(x) = \int_{\bar\Omega} f(\omega,\, x)\, d\bar{\mathcal{P}}(\omega).$$

Proof of Lemma 3.3. Since f is measurable with respect to $x, \omega_1, \omega_2, \ldots$, we can write $f(\omega, x) = f(\omega_1, \omega_2, \ldots, x)$, and therefore $f \circ T^n(\omega, x) = f(\omega_{n+1}, \omega_{n+2}, \ldots, \omega_n \cdots \omega_1 x)$. Let $h = g \circ T^n \in T^n L^2_0(\bigvee_{-\infty}^{-n} T^i\mathcal{B}) = L^2_0(\bigvee_{-\infty}^{0} T^i\mathcal{B})$. The function h is measurable with respect to x, ω_i, $i \le 0$, and hence $(g \circ T^n)(\omega, x) = h(\omega, x) = h(\ldots, \omega_0, x)$. Now compute

$$\begin{aligned} |\langle f, g\rangle| &= |\langle f \circ T^n, g \circ T^n\rangle| = |\langle f \circ T^n, h\rangle| \\ &= \left| \int_X \int_{\bar\Omega} f(\omega_{n+1}, \ldots, \omega_n \cdots \omega_1 x) \cdot \bar h(\ldots, \omega_0, x)\, d\bar{\mathcal{P}}(\omega)\, dm(x) \right| \end{aligned}$$

integrating the last expression with respect to variables ω_i with $i \le 0, i > n$, we continue

$$\begin{aligned} &= \left| \int_X \int_{\bar\Omega} P_X f(\omega_n \cdots \omega_1 x) \cdot P_X \bar h(x)\, d\mu(\omega_1) \cdots d\mu(\omega_n)\, dm(x) \right| \\ &= |\langle \pi(\check\mu)^n P_X f,\, P_X h\rangle| \le \|\pi(\check\mu)^n P_X f\| \cdot \|P_X h\|\,. \end{aligned}$$

Since $\|P_X h\| \le \|h\| = \|g\|$ and $\|P_X f\| \le \|f\|$, we have

$$\|\pi(\check\mu)^n P_X f\| \le \|\pi(\check\mu)\|^n \cdot \|P_X f\| \le \|\pi(\mu)\|^n \cdot \|f\|$$

which implies (3.4). ☐

For the proof of Theorem 1.3 we shall also need:

Proposition 3.4. *Let G act ergodically on a probability space (X,m), and let μ be a measure on G. Suppose that a function $f \in L^2_0(X,m)$ is an L^2-coboundary on $(\bar{\Omega}\times X, \bar{\mathcal{P}}\times m, T)$, i.e. $f(x) = h(\theta\omega, \omega_1 x) - h(\omega, x)$ with $h \in L^2_0(\bar{\Omega}\times X)$. Then $h(\omega,x) = h(x)$ does not depend on ω. Furthermore, if μ is aperiodic on G, or satisfies $\|\pi(\mu)\| < 1$, then necessarily $f = 0$.*

The proof of a more general statement, namely, that $f \neq 0$ is not a *measurable* coboundary (rather than merely an L^2-coboundary), is given in appendix A.

Proof of Theorem 1.3. Theorem 3.2 states that S_n is a sum of n random variables $f \circ T^i$, where f is an L^p bounded ($p > 2$) random variable, which is measurable with respect to a ρ-mixing σ-subalgebra $\mathcal{B}$. In this situation, proposition 3.4 is known to imply that $\sigma = \lim_{n\to\infty} n^{-1/2}\cdot\sigma_n$ is non zero, and therefore S_n satisfy the assumptions of [OY]. The second statement, concerning the rate of convergence, follows from the *exponential* ρ-mixing as was shown in [Ti]. ☐

4. Mixing Properties

In this section we discuss mixing properties of group actions along random products. In finite dimensional representations no mixing can occur, hence, seeking such phenomena, we need to exclude the existence of finite dimensional sub-representations in $L^2_0(X,m)$. In general, it is well known that a unitary representation π contains a finite dimensional sub-representation iff $\pi\otimes\pi^c$ contains the trivial sub-representation 1_G. Here π^c denotes the contragredient representation to π, i.e. the natural representation on the dual space. If π arises from a measure preserving G-action on (X,m), then $\pi^c \cong \pi$, so the condition $1_G \not\subseteq \pi\otimes\pi^c$ is equivalent to $1_G \not\subseteq \pi\otimes\pi$, itself equivalent to the ergodicity of the diagonal G-action on $(X\times X, m\times m)$. For proofs of these and further results in this direction see [BR]. In light of this, and motivated by Theorems 1.1-1.2, it only seems natural now to impose the spectral condition $\|\pi\otimes\pi\,(\mu)\| < 1$, for an appropriate measure μ on G, where the representation π is as in the previous sections. We remark that in many natural examples the condition $\|\pi\otimes\pi\,(\mu)\| < 1$ is equivalent to $\|\pi(\mu)\| < 1$ (see section 6), but this is not the case in general.

Our first goal is to prove Theorem 1.4. To emphasize the Hilbertian nature of the statement, we reformulate the theorem:

Theorem 4.1. *Let G act on $(X,\mathcal{B},m)$, and μ be a probability distribution on G, such that $\|\pi\otimes\pi\,(\mu)\| < 1$. Then there exists a full measure subset $\Omega' \subset \Omega$, so that for each $\omega\in\Omega'$ the sequence of random products $\{g_n^\omega\}_{n=1}^\infty$ is mixing: for all $\phi,\psi \in L^2_0(X,m)$*

$$\lim_{n\to\infty}\langle\phi,\,\pi(g_n^\omega)\psi\rangle = \lim_{n\to\infty}\int_X \phi(g_n^\omega x)\cdot\bar{\psi}(x)\,dm(x) = 0. \tag{4.1}$$

Moreover, for any λ with $\lambda_0 = \sqrt{\|\pi \otimes \pi\,(\mu)\|} < \lambda \leq 1$, and any $\phi, \psi \in L_0^2(X, m)$, there is a subset of full measure $\Omega_{\phi,\psi} \subset \Omega$, such that for any $\omega \in \Omega_{\phi,\psi}$:

$$\lim_{n\to\infty} \frac{\langle \phi,\ \pi(g_n^\omega)\psi\rangle}{\lambda^n} = 0. \tag{4.2}$$

For the proof of Theorem 4.1 we shall need the following

Lemma 4.2. *Let G act on (X, m), and suppose that $\|\pi \otimes \pi\,(\mu)\| < 1$ for some probability distribution μ on G. Let $\lambda_0 = \sqrt{\|\pi \otimes \pi\,(\mu)\|} < 1$, then for any $\phi, \psi \in L_0^2(X, m)$:*

$$\mathbf{E}\left|\int_X \phi(g_n^\omega x)\cdot\bar{\psi}(x)\,dm(x)\right| \leq \lambda_0^n \cdot \|\phi\|\cdot\|\psi\|, \tag{4.3}$$

Proof. Consider the functions $\Phi = \phi\otimes\bar{\phi}$ and $\Psi = \psi\otimes\bar{\psi}$ in $L_0^2(X\times X, m\times m)$.

$$\begin{aligned}
\mathbf{E}\ \ & \left|\int_X \phi(g_n^\omega x)\cdot\bar{\psi}(x)\,dm(x)\right| \leq \left(\int_G \left|\int_X \phi(gx)\cdot\bar{\psi}(x)\,dm(x)\right|^2 d\mu^{(n)}(g)\right)^{1/2}\\
= \ \ & \left(\int_G\int_X\int_X \phi(gx)\cdot\bar{\phi}(gy)\cdot\bar{\psi}(x)\cdot\psi(y)\,dm(x)\,dm(y)\,d\mu^{(n)}(g)\right)^{1/2}\\
= \ \ & \left(\int_{X\times X}\left(\int_G \Phi(g\,(x,y))\,d\mu^{(n)}(g)\right)\cdot\bar{\Psi}(x,y)\,dm\times m(x,y)\right)^{1/2}\\
\leq \ \ & \|\pi\otimes\pi\,(\mu)\|^{n/2}\cdot\|\Phi\|^{1/2}\cdot\|\Psi\|^{1/2}
\end{aligned}$$

Since $\|\Phi\| = \|\phi\|^2$ and $\|\Psi\| = \|\psi\|^2$, this completes the proof of the Lemma. $\square$

Proof of Theorem 4.1. Choose some $\lambda > \lambda_0$, and observe that the non-negative series

$$S(\omega) = \sum_{n=1}^{\infty}\frac{1}{\lambda^n}\left|\int_X \phi(g_n^\omega x)\cdot\psi(x)\,dm(x)\right|$$

has finite expectation, for by Lemma 4.2: $\mathbf{E}\,(S) \leq \sum_1^\infty (\lambda_0/\lambda)^n < \infty$. Therefore for a full measure subset $\Omega_{\phi,\psi}\subset\Omega$ we have:

$$\lim_{n\to\infty}\frac{1}{\lambda^n}\left|\int_X \phi(g_n^\omega x)\cdot\psi(x)\,dm(x)\right| = 0$$

thereby proving statement (4.2) in the theorem. Taking $\Omega' = \bigcap_{i,j}\Omega_{\phi_i,\phi_j}$, where $\{\phi_i\}_{i\in\mathbb{N}}$ is a dense countable sequence in $L_0^2(X, m)$, we obtain (4.1). $\square$

Note that the probabilistic nature of (4.2) is intrinsic: one cannot expect any deterministic rate of mixing for all $A, B \in \mathcal{B}$. In fact no *fixed* sequence $\{g_n\}$ can provide any specific *rate* of mixing for *all* the functions, as the following general assertion holds:

Remark 4.3. Given any sequence $\{U_n\}_{n=1}^{\infty}$ of unitary operators on a Hilbert space $\mathcal{H}$, any $v \neq 0$, and any sequence $a_n \to 0$, there exists $u \in \mathcal{H}$ such that $|\langle U_n\, u,\, v\rangle| > a_n\, \|u\| \cdot \|v\|$. This follows from Banach's uniform boundness principle, applied to the sequence of linear functionals $f_n(x) = a_n^{-1} \cdot \langle U_n\, x,\, v\rangle$ having norms $\|f_n\| = a_n^{-1} \cdot \|v\| \to \infty$.

As we have already mentioned, the condition $\|\pi \otimes \pi\,(\mu)\| < 1$ holds for any weakly mixing action of a discrete group G with Kazhdan's property (T) and aperiodic μ on G. In particular, it is satisfied by lattices in higher rank simple Lie groups.

Corollary 4.4. *Let Γ be a discrete group with Kazhdan's property (T). Suppose Γ acts weakly mixing on a probability space (X, m), and let μ be an aperiodic probability measure on Γ. Then with probability 1 the sequence of transformations $\{g_n^{\omega}\}$ is mixing, and in fact an exponential rate of mixing holds for a dense set of L^2-functions.*

Theorem 4.1 as well as Corollary 4.4, apply to any unitary representation. This can be considered a "random product" analog of Howe-Moore's theorem concerning the vanishing of matrix coefficients, which applies to semi-simple Lie groups, but does not hold for their lattices. It is known, that for lattices (e.g., $\mathrm{SL}_3(\mathbb{Z})$) the notions of ergodicity, weak mixing, mixing and 3-mixing are all distinct. In particular, there exist *non mixing* actions, which satisfy $\|\pi \otimes \pi\,(\mu)\| < 1$, and hence are *mixing on random products* (see [FS]).

Remark 4.5. Given G, μ and (X, m) such that $\|\pi \otimes \pi\,(\mu)\| < 1$, one can consider the critical exponential rate of mixing for the action, namely the greatest lower bound λ_c for all λ satisfying (4.2). Theorem 4.1 gives $\lambda_c \leq \sqrt{\|\pi \otimes \pi\,(\mu)\|}$. On the other hand, it is shown in [FS] that the Kaimanovich-Vershik entropy $h(G, \mu)$ gives a lower bound: $e^{-h(G,\mu)} \leq \lambda_c$. In the case of an automorphism group of a compact abelian group, one can show a stronger estimate: $e^{-h(G,\mu)/2} \leq \lambda_c \leq \|\pi(\mu)\|$. In the particular case of the two-dimensional torus and $G \subset \mathrm{SL}_2(\mathbb{Z})$ it can further be shown that $\|\pi(\mu)\| = \|\rho(\mu)\|$ where $(\rho, L^2(G))$ denotes the regular representation. For instance, for the free group

$$G = F_2 = \left\langle \begin{pmatrix} 1 & \pm 2 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ \pm 2 & 1 \end{pmatrix} \right\rangle \subset \mathrm{SL}_2(\mathbb{Z})$$

acting on $\mathbb{R}^2/\mathbb{Z}^2$, where μ is equally distributed on the four generators, one obtains

$$0.76 \approx \frac{1}{3^{1/4}} \leq \lambda_c \leq \frac{\sqrt{3}}{2} \approx 0.86$$

5. Strong Ergodicity and the Spectral Condition

In this section we prove Theorem 1.6, regarding the connection between strong ergodicity and the spectral gap condition $\|\pi(\mu)\| < 1$. It is easy to check that $\|\pi(\mu)\| < 1$ implies $1_G \not\prec \pi$. Thus for measurably aperiodic probability distribution μ on a locally compact group G, acting ergodically on (X, m), one has:

$$(i)\ G \text{ is strongly ergodic on } (X, m) \quad \Rightarrow \quad (ii)\ \|\pi(\mu)\| < 1 \quad \Rightarrow \quad (iii)\ 1_G \not\prec \pi.$$

If G were discrete, the three conditions are in fact equivalent, as is shown in Theorem 6.3 below. For general locally compact groups (iii) does not imply (i), as the example of $G = S^1$ acting on itself shows (see the next section for a further discussion).

The proof of Theorem 1.6 follows the ideas of the discrete case in [Sc]. We shall also use the opportunity to close a gap in [Sc]. We first need the following two lemmas:

Lemma 5.1. *Let G act ergodically on (X, m). Let $g \in G$ and suppose there exists a sequence of measurable sets $C_n \subseteq X$, satisfying $m(gC_n \triangle C_n)/m(C_n) \to 0$. Then every limit point in the weak-* topology on $L^\infty(X, m)^*$ of the sequence of means ϕ_n, given by*

$$\phi_n(F) = \frac{1}{m(C_n)} \int_{C_n} F\, dm, \qquad F \in L^\infty(X, m)$$

is a non zero g-invariant mean on $L^\infty(X, m)$. Moreover, if $\sum m(C_n) < 1$ then every such limit mean is different from m.

Proof. Denoting $h_n = m(C_n)^{-1}\, 1_{C_n}$ we note that

$$\|h_n \circ g - h_n\|_1 = \frac{m(gC_n \triangle C_n)}{m(C_n)} \to 0.$$

Therefore, if ϕ is a weak-* limit of $\{\phi_{n_k}\}$, then for any $F \in L^\infty(X, m)$:

$$|g\,\phi(F) - \phi(F)| = \lim_{k\to\infty} \left| \int_X (h_{n_k} \circ g - h_{n_k}) \cdot F\, dm \right| \leq \|F\|_\infty \cdot \lim_{k\to\infty} \|h_{n_k} \circ g - h_{n_k}\|_1 = 0$$

Thus ϕ is g-invariant. If $\sum m(C_n) < 1$, then $m(F) < 1 = \phi(F)$ for $F = \sum 1_{C_n}$. □

Lemma 5.2. *Let ν be a probability distribution on G, which is not supported on a proper measurable subgroup in G. Suppose that $C_n \subseteq X$ satisfy $m(C_n) \to t$ with $0 < t < 1$, and that for ν-a.e. $g \in G$: $m(gC_n \triangle C_n) \to 0$. Then m is not unique as a G-invariant mean on $L^\infty(X, m)$.*

Proof. We shall assume that m is unique and show first, that under these conditions, there exists a sequence $D_n \subset X$, such that $m(D_n) \to t^2$ and for ν-a.e. $g \in G$: $m(g\,D_n \triangle D_n) \to 0$.

Indeed, from Lemma 5.1 and the uniqueness of m, it follows that

$$\lim_{k\to\infty} \frac{m(C_k \cap C_n)}{m(C_k)} = \lim_{k\to\infty} \frac{m(C_k \cap C_n)}{t} = m(C_n) \tag{5.1}$$

The sequence we shall construct is of the form $D_{n,k} = C_n \cap C_k$, for suitable n, k. The argument above, shows that if n is fixed, such that $m(C_n)$ is close to t, then for k large enough $m(D_{n,k})$ is close to t^2. Note also, that if $g \in G$, $m(g\,C_n \triangle C_n) < \epsilon$ and $m(g\,C_k \triangle C_k) < \epsilon$, then

$$m((g\,C_n \cap g\,C_k) \triangle (C_n \cap C_k)) \leq m(g\,C_k \triangle C_k) + m(g\,C_n \triangle C_n) < 2\epsilon. \tag{5.2}$$

Thus, given $\epsilon > 0$, we can choose C_n with $|m(C_n) - t| < \epsilon$ and such that for a measurable set $F \subseteq G$ with $\nu(F) > 1 - \epsilon$, $m(gC_n \triangle C_n) < \epsilon$ for every $g \in F$.

Now, applying (5.1), take k large enough so that $|m(D_{n,k}) - t^2| < 2\epsilon$, and, using (5.2), we have that for a set $F_1 \subseteq G$ with $\nu(F_1) > 1 - 2\epsilon$:

$$g \in F_1 \Rightarrow m(g\, D_{n,k} \triangle D_{n,k}) < 2\epsilon.$$

Letting $\epsilon = 2^{-i} \to 0$, we obtain from Borel-Cantelli a sequence of D's as required. Repeating this process, one obtains for every $i \in \mathbb{N}$, a sequence D_n^i such that $\lim_{n\to\infty} m(D_n^i) = t^{2^i}$ and for ν-a.e. $g \in G$: $\lim_{n\to\infty} m(gD_n^i \triangle D_n^i) = 0$.

For every i, choose n_i large enough so that

(i) $m(D_{n_i}^i) < 2 \cdot t^{2^i} \to 0 \quad as \quad i \to \infty$

(ii) $\exists F_i \subset G$ with $\nu(F_i) > 1 - 2^{-i}$, such that for any $g \in F_i$: $m(gD_{n_i}^i \triangle D_{n_i}^i)/m(D_{n_i}^i) < 1/i$.

By Borel-Cantelli ν-a.e. $g \in G$ is contained in all, but finitely many F_i-s. Since $m(D_{n_i}^i) \to 0$, we may take a subsequence with $\sum_i m(D_{n_i}^i) < 1$. By Lemma 5.1 we obtain a limit mean $\phi \neq m$ on $L^\infty(X, m)$, which is invariant under a measurable set $F = \liminf F_i$ with $\nu(F) = 1$. But ν is not supported by a proper measurable subgroup of G, thus ϕ is G-invariant, and not equal to m. □

Proof of Theorem 1.6. We assume that $\|\pi(\mu)\| = 1$ and show that m is not the unique G-invariant mean on $L^\infty(X, m)$. Consider $\nu = \check{\mu} * \mu$. Since μ is measurably aperiodic, ν is not supported by a proper measurable subgroup of G. The self-adjoint non-negative operator $\pi(\nu) = \pi(\mu)^*\pi(\mu)$ has norm $\|\pi(\nu)\| = \|\pi(\mu)\|^2 = 1$, which can be approximated by a sequence $f_n \in L_0^2(X, m)$ with $\|f_n\|_2 = 1$ and

$$\langle \pi(\nu)f_n,\, f_n\rangle = \int_G \int_X f_n(g^{-1}x) \cdot \bar{f}_n(x)\, dm(x)\, d\nu(g) \to 1 \tag{5.3}$$

Viewing $f_n(g^{-1}x), f_n(x)$ as unit vectors in $L^2(G \times X, \nu \times m)$ we deduce that

$$\int_G \int_X |f_n - \pi(g)\, f_n)|^2\, dm\, d\nu = \int_G \|f_n - \pi(g)\, f_n\|^2\, d\nu \to 0.$$

Following [Sc] we now define a sequence σ_n of probability measures on $\mathbb{R}$, by

$$\sigma_n(D) = m(f_n^{-1}(D)).$$

Observe that $\int t\, d\sigma_n(t) = 0$ and $\int t^2\, d\sigma_n(t) = 1$. The last equality shows that the sequence $\{\sigma_n\}$ is uniformly tight, so without loss of generality, we may assume that σ_n converges weakly on compacta to a probability measure σ on $\mathbb{R}$.

If σ is not concentrated on a single point, we can find $\alpha \in \mathbb{R}$, such that

$$0 < \sigma(-\infty, \alpha) = \sigma(-\infty, \alpha] = t < 1$$

Putting $C_n = f_n^{-1}(-\infty, \alpha)$ it follows that $m(C_n) \to t$, and for any $g \in G$ which satisfies $\|f_n - g\, f_n\|_2 \to 0$, we have $m(gC_n \triangle C_n) \to 0$. Thus the sequence C_n satisfies the conditions of Lemma 5.2 and m is not unique.

We now deal with the more complicated case, where σ is concentrated on one point α_0, starting by showing that $\alpha_0 = 0$ (in [Sc] the proof proceeds without identifying α_0, which causes a gap later). Let $\epsilon > 0$ and take N with $N > \max(|\alpha_0|, 1/\epsilon)$. Let $\rho \in \{\sigma_n\}$, such that

$$\left| \int_{\{|t|\le N\}} t\, d\rho - \int_{\{|t|\le N\}} t\, d\sigma \right| \le \epsilon. \tag{5.4}$$

Notice that

$$N \cdot \int_{\{|t|>N\}} |t|\, d\rho \le \int_{\{|t|>N\}} t^2\, d\rho \le 1.$$

So $|\int_{\{|t|>N\}} t\, d\rho| \le 1/N < \epsilon$, which, together with $\int t\, d\rho = 0$ gives $|\int_{|t|\le N} t\, d\rho \le \epsilon$.

By (5.4) we now deduce that $|\int_{|t|\le N} t\, d\sigma| \le 2\epsilon$. But $\sigma = \delta_{\alpha_0}$, so $|\alpha_0| \le 2\epsilon$, as required. Defining $h_n = f_n^2$ we obtain that $\int h_n\, dm = 1$, and for ν-a.e. g:

$$\|h_n - \pi(g)\, h_n\|_1 \le \|f_n - \pi(g)\, f_n\|_2 \cdot \|f_n + \pi(g)\, f_n\|_2 \to 0$$

Now for every $n \ge 1$ let

$$h_n^*(x) = \begin{cases} h_n(x) & \text{if } h_n(x) \ge 1 \\ 0 & \text{if } h_n(x) < 1 \end{cases}$$

As in [Sc] one observes, that since σ is concentrated at 0, and $\|h_n\|_1 = 1$ we have $\lim_{n\to\infty} \|h_n^*\| = 1$, and for ν-a.e. g: $\lim_{n\to\infty} \|h_n^* - h_n^* \circ g\|_1 = 0$. Using an adaptation, due to A. Connes, of a technique of I. Namioka, we set for every $s \in \mathbb{R}$, $x \in X$, $g \in G$:

$$F_n(g, s, x) = \begin{cases} 1 & \text{if } h_n^*(g^{-1}x) \ge s \\ 0 & \text{if } h_n^*(g^{-1}x) < s \end{cases}$$

and obtain:

$$\begin{aligned} \|h_n^* - h_n^* \circ g\|_1 &= \int_X \int_0^\infty |F_n(e,s,x) - F_n(g,s,x)|\, ds\, dm(x) && (5.5) \\ &= \int_0^\infty \int_X |F_n(e,s,x) - F_n(g,s,x)|\, dm(x)\ ds && (5.6) \end{aligned}$$

$$\|h_n^*\| = \int_X \int_0^\infty F_n(e,s,x)\, ds\, dm(x) = \int_0^\infty \int_X F_n(e,s,x)\, dm(x)\, ds \tag{5.7}$$

Replacing $\{h_n^*\}$ by a subsequence, we may assume that there exists an increasing sequence of Borel sets $E_n \subset G$, with $\nu(\cup E_n) = 1$, such that

$$\forall g \in E_n: \quad \|h_n^* - h_n^* \circ g\|_1 < \frac{\|h_n^*\|_1}{n^2} \tag{5.8}$$

¿From (5.6-5.8) it follows that

$$\int_{E_n} \int_0^\infty \int_X |F_n(e,s,x) - F_n(g,s,x)|\ dm(x)\ ds\, d\nu(g) < \frac{1}{n^2} \int_0^\infty \int_X F_n(e,s,x)\, dm(x)\, ds$$

Changing the order of integration between ds and $d\nu$, in the left hand side, and inserting $1/n^2$ into the right hand side integral, we deduce the existence of $s = s(n) > 0$, such that

$$\int_{E_n} \int_X |F_n(e, s(n), x) - F_n(g, s(n), x)| \, dm(x) \, d\mu(g) < \frac{1}{n^2} \int_X F_n(e, s(n), x) \, dm(x) \tag{5.9}$$

Defining $C_n = \{x \mid h_n^*(x) \geq s(n)\}$ we see that $m(C_n) \to 0$ (as $x \in C_n$ implies $h_n(x) \geq 1$ and σ is concentrated at 0), so (5.9) translates to:

$$\int_{E_n} m(g\, C_n \triangle C_n) \, d\nu(g) < \frac{1}{n^2} m(C_n).$$

Now put

$$\phi_n(g) = \begin{cases} \frac{m(g\, C_n \triangle C_n)}{m(C_n)} & \text{if } g \in E_n \\ 0 & \text{otherwise} \end{cases}$$

Since $\int \phi_n(g) \, d\mu(g) < 1/n^2$, for μ-a.e.g we have $\sum \phi_n(g) < \infty$ and hence $\phi_n(g) \to 0$. From $\mu(E_n) \to 1$, it follows that $m(g\, C_n \triangle C_n)/m(C_n) \to 0$ for μ-a.e. g. Recalling that $m(C_n) \to 0$, we can pass to a subsequence with total measure less than 1. Since the assumption that μ is measurably aperiodic implies that ν is not supported by a proper measurable subgroup of G, it follows from Lemma 5.1 and 5.2 that m is not the unique G-invariant. This proves the theorem. □

6. Strongly Ergodic Actions

In light of the results of the previous sections, we are naturally interested in establishing the spectral gap condition $\|\pi(\mu)\| < 1$, and Theorem 1.6 suggests that the strong ergodicity property should be examined more thoroughly. Let us first recall the following:

Definition 6.1. Let (π, V) be a unitary representation of a locally compact group G. Then π is said to *weakly contain* the trivial representation 1_G (notation: $1_G \prec \pi$), if π admits *almost invariant* vectors, i.e. for any compact subset $K \subset G$ and any $\epsilon > 0$, there exists $v \in V$, such that $\|\pi(g)v - v\| \leq \epsilon \|v\|$ for all $g \in K$.

Definition 6.2. A locally compact group G is said to have *Kazhdan's property* (T), if for any unitary G-representation π with $1_G \prec \pi$, necessarily $1_G \subset \pi$.

It turns out that for a countable group G, there is a clear connection between strong ergodicity and representation theoretical properties related to the action.

Theorem 6.3. *Let G be a countable discrete group, acting on (X, m) preserving the probability measure m, and let μ be an aperiodic probability measure on G. Then the following are equivalent*

(i) The G-action on (X, m) is strongly ergodic.
(ii) $\|\pi(\mu)\| < 1$ (where π is the associated representation, as in the previous section).
(iii) $1_G \not\prec \pi$.

Note the analogy to the well known equivalent conditions, in the case of a measure preserving action of a locally compact group G on (X,m), where μ is a generating probability measure on G: (i) The G-action on (X,m) is ergodic (ii) $\pi(\mu)$ has no fixed vectors in $L^2(X,m)$ (iii) $1_G \not\subseteq \pi$.

Proof of Theorem 6.3. (i) $\Rightarrow$ (ii) was proved in Theorem 1.6 (actually, the implication (i) $\Rightarrow$ (iii) is due to K. Schmidt [Sc], and (iii) $\Rightarrow$ (ii) can be proved by an easy direct argument). The implication (ii) $\Rightarrow$ (iii) is obvious, and (iii) $\Rightarrow$ (i) follows from [Ro].

□

The assumption that G is countable is crucial for the above results. When G is any locally compact group, it is shown in [FS] that if μ is absolutely continuous, and is not supported on a coset of an open subgroup, then the implication (iii) $\Rightarrow$ (ii) above, still holds, and thus (ii) – (iii) in Theorem 6.3 are equivalent. However, in general the situation for non-countable groups can be quite different than the one described in Theorem 6.3, and it is perhaps best illustrated by groups with Kazhdan's property (T). If G is discrete, countable, and has Kazhdan's property (T), then every ergodic action of it is strongly ergodic. However $G = \mathbb{R}/\mathbb{Z}$ is compact (hence has property (T)), but its ergodic action on itself is not strongly ergodic (see e.g. [**?**] 2.2.11). This construction can be extended to any (Kazhdan) group which admits a non-trivial character. We remark that, at least for connected (Kazhdan) Lie groups, it is shown in [Sh2] that admitting such character is the **only** obstacle for strong ergodicity.

Before proceeding to more concrete examples, we should mention here that amenable groups have no strongly ergodic actions, and that these are the **only** groups with this property. On the other extreme, groups with property (T) are exactly those groups for which every ergodic action is strongly ergodic (cf. [**?**] 3.5 and the references therein).

Automorphisms of compact abelian groups. An important class of measure preserving actions are automorphic actions on compact abelian groups, and we shall investigate strong ergodicity for these actions.

Suppose G is a discrete, countable group, acting by automorphisms on a compact abelian group A, with Haar measure m (which is obviously preserved by the action). Denote, as usual, by π the G-representation on $L_0^2(A,m)$. Let $\hat{A}$ be the discrete dual group of unit characters, and $0 \in \hat{A}$ - the trivial character. Recall that G acts on $\hat{A}$ by: $g \cdot \chi = \chi \circ g^{-1}$, fixing 0, and hence G acts also on $\hat{A}\backslash\{0\}$. We denote by $\hat{\pi}$ the G-representation on $L^2(\hat{A}\backslash\{0\})$, and recall that the Fourier transform:

$$f \in L_0^2(A,m) \;\mapsto\; \hat{f} \in L^2(\hat{A} \setminus \{0\}) \quad \text{defined by} \quad \hat{f}(\chi) = \int_A f(x) \cdot \bar{\chi}(x)\, dm(x)$$

is an isomorphism between the G-representations π on $L_0^2(A,m)$ and $\hat{\pi}$ on $L^2(\hat{A} \setminus \{0\})$.

Theorem 6.4. *With the above notations, the following conditions are equivalent:*

1. *The G-action on (A,m) is strongly ergodic.*

2. $1_G \not\prec \pi$.
3. $1_G \not\prec \hat{\pi}$.
4. *There is no G-invariant mean on $L^\infty(\hat{A} \setminus \{0\})$.*
5. *The G-action on $(A \times A, m \times m)$ is strongly ergodic.*
6. $1_G \not\prec \pi \otimes \pi$.
7. $1_G \not\prec \hat{\pi} \otimes \hat{\pi}$.
8. *There is no G-invariant mean on $L^\infty(\hat{A} \setminus \{0\} \ \times \ \hat{A} \setminus \{0\})$.*

Proof. $1 \Rightarrow 2$. Follows from Theorem 6.3.

$2 \Rightarrow 3$. Obvious, since $\pi \cong \hat{\pi}$.

$3 \Rightarrow 4$. This follows from a general result, which is proved in [FS]. For completeness we sketch the proof in this case. Suppose $\phi \in L^\infty(\hat{A} \setminus \{0\})^*$ is a G-invariant mean. Then there exists a sequence $h_n \in L^1(\hat{A} \setminus \{0\}) \subset L^\infty(\hat{A} \setminus \{0\})^*$, with $h_n \to \phi$ weakly. Passing to convex combinations, one obtains such h_n with $\|h_n \circ g - h_n\|_1 \to 0$ for every $g \in G$. We can assume, that $\|h_n\|_1 = 1$ and $h_n \geq 0$ for all n, and define $f_n(x) = \sqrt{h_n(x)}$ - unitary vectors in $L^2(\hat{A} \setminus \{0\})$. Using an elementary inequality: $|\sqrt{a} - \sqrt{b}|^2 \leq |a - b|$, we obtain for every $g \in G$:

$$\|\hat{\pi}(g) f_n - f_n\|_2^2 \leq \|h_n \circ g - h_n\|_1 \to 0$$

and therefore, $1_G \prec \hat{\pi}$, contradicting the assumption.

$4 \Rightarrow 8$. If there was a G-invariant mean Φ on $L^\infty(\hat{A} \setminus \{0\} \ \times \ \hat{A} \setminus \{0\})$ then its projection ϕ on $L^\infty(\hat{A} \setminus \{0\})$, defined by $\phi(f) = \Phi(f \otimes 1)$ would be G-invariant too.

$8 \Rightarrow 7$. Suppose $f_n \in L^2(\hat{A} \setminus \{0\} \ \times \ \hat{A} \setminus \{0\})$ are almost invariant unit vectors. Consider the $L^1(\hat{A} \setminus \{0\} \ \times \ \hat{A} \setminus \{0\})$ functions $h_n = |f_n|^2$ as means on $L^\infty(\hat{A} \setminus \{0\} \ \times \ \hat{A} \setminus \{0\})$. By Cauchy-Schwartz inequality we have for every $g \in G$:

$$\begin{aligned} \|h_n \circ g - h_n(x)\|_1 &= \sum_{x,y} |f_n(g \cdot x, g \cdot y) - f_n(x,y)| \cdot |f_n(g \cdot x, g \cdot y) + f_n(x,y)| \\ &\leq 2 \cdot \left(\sum_{x,y} |f_n(g \cdot x, g \cdot y) - f_n(x,y)|^2 \right)^{1/2} \to 0 \end{aligned}$$

Therefore, any weak limit of the h_n's defines a G-invariant mean.

$7 \Rightarrow 6$. Follows from $\hat{\pi} \otimes \hat{\pi} \cong \pi \otimes \pi$.

$6 \Rightarrow 5$. Notice that the representation on $L^2_0(\hat{A} \setminus \{0\} \ \times \ \hat{A} \setminus \{0\})$ is isomorphic to $\pi \otimes \pi \ \oplus \ \pi \otimes 1_G \ \oplus \ 1_G \otimes \pi$. Since $1_G \not\prec \pi \otimes \pi$, also $1_G \not\prec \pi$, and condition 5 follows from Theorem 6.3.

$5 \Rightarrow 1$. If $\phi \neq m$ is another G-invariant mean on A, then the mean $\phi \times m$ defined by

$$\phi \times m(f) = \phi \left(\int f(x,y) \, dm(y) \right)$$

is a G-invariant mean which differs from $m \times m$. $\square$

Note that from Theorem 6.4 it follows that for automorphism actions strong ergodicity implies strong ergodicity of the diagonal action too, so the mixing theorems (of section 4) apply as well as the ergodic ones.

To show more concrete examples, and the way Theorem 6.4 may be used, we have:

Theorem 6.5. *Let G be a subgroup of $SL_d(\mathbb{Z})$. Then either of the following conditions implies strong ergodicity of the G-action on the torus $\mathbb{T}^d = \mathbb{R}^d/\mathbb{Z}^d$:*

(i) The Zariski closure $\bar{G}$ of G in $GL_n(\mathbb{R})$ admits no invariant probability measures on the projective space $\mathbb{P}^{d-1}(\mathbb{R})$, for the adjoint action, defined by: $g \cdot [u] = [(g^t)^{-1}u]$.

(ii) G acts irreducibly on $\mathbb{R}^n$, and G does not have an abelian subgroup of finite index.

Proof. Assume (i). Since stabilizers in $PGL_d(\mathbb{R})$ of measures on $\mathbb{P}^{d-1}(\mathbb{R})$ are algebraic ([Zi] 3.2.18), condition (i) implies that there is no G-invariant probability measure on $\mathbb{P}^{d-1}(\mathbb{R})$ for the adjoint action. By Theorem 6.4 it is enough to show that there is no G-invariant mean on $L^2(\hat{A} \setminus \{0\})$ for the dual action $g \cdot u = (g^t)^{-1}u$. Suppose that such a mean ϕ exists. Denote by $P : \hat{A} \setminus \{0\} \to \mathbb{P}^{d-1}(\mathbb{R})$ the natural projection, and let $\phi_*(f) = \phi(f \circ P)$. It is easily verified that ϕ_* defines a positive normalized functional on $C(\mathbb{P}^{d-1}(\mathbb{R}))$, and thus corresponds to a probability measure. But, since ϕ is G-invariant, ϕ_* is G-invariant as well, contradicting the assumption. This proves the first statement of the theorem.

Assume (ii). Since G has no invariant subspaces in $\mathbb{R}^d$, the same is true for the adjoint G-representation. By (i), it is enough to show that the adjoint G-action does not preserve any probability measure on $\mathbb{P}^{d-1}(\mathbb{R})$. Suppose G admits such a measure. Passing to a subgroup, we can assume that G is finitely generated and still acts irreducibly. By the descending chain condition there exists a minimal quasi-linear subvariety $S = V_1 \cup \cdots \cup V_k$, which has full measure. By Furstenberg's Lemma [**?**] and the minimallity of S, G has a finite index subgroup $G_1 \subseteq G$, preserving each of the subspaces V_i-s, and its image in every $PGL(V_i)$ is precompact. Since G_1 is linear (over characteristic zero) and finitely generated, it contains a torsion free subgroup G_2 of finite index in G_1. The image of the commutator $[G_2, G_2]$ in each $GL(V_i)$ is contained in $SL(V_i)$, and is precompact there. Since G acts irreducibly, $\mathbb{R}^n = \mathrm{span}\{\cup V_i\}$, so $[G_2, G_2]$ is precompact in $SL_n(\mathbb{R})$. By discreteness $[G_2, G_2]$ is finite, and as it is torsion free, $[G_2, G_2]$ is trivial. □

Remark 6.6. For $G = SL_d(\mathbb{Z})$ Theorem 6.5 was proved originally by J. Rosenblatt [Ro], using a different method.

The fast rate of convergence in the mean and pointwise ergodic theorems (1.1, 1.2) suggests a way to distribute points uniformly on probability spaces, using strongly ergodic group actions. For example, Theorem 6.5 applied to

$$G = \mathrm{SL}_2(\mathbb{Z}) = \left\langle \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \right\rangle$$

yields the following simple random algorithm of distributing points uniformly on the circle (and tori):

Example 6.7. Choose randomly $(x_0, y_0) \in [0,1]^2$ and apply the following random procedure: at the n-th step choose $(x_{n+1},\, y_{n+1})$ to be either $(x_n \pm y_n, y_n)$ or $(x_n, y_n \pm x_n)$ mod 1, with probability 1/4 each. Then, with probability 1, the sequence (x_n, y_n) is distributed uniformly in the torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$, averaging any given L^2-function with rate $o(\log^{\frac{3}{2}+\epsilon} n/\sqrt{n})$ and satisfying the CLT estimates.

As a different example, consider the following compact, connected, abelian group (which is not a Lie group): $A = (\mathbb{R} \times \mathbb{Q}_p/\mathbb{Z}[\frac{1}{\mathrm{p}}])^d$, where the ring $\mathbb{Z}[\frac{1}{\mathrm{p}}]$ is embedded diagonally in the product $\mathbb{R} \times \mathbb{Q}_p$ of the real and p-adic fields. Then the diagonal linear action of $SL_d(\mathbb{Z})$ on $\mathbb{R}^d \times \mathbb{Q}_p^d$ naturally induces an action on A. By Theorem 6.4 and the method of proof of Theorem 6.5 one can then deduce:

Theorem 6.8. *Let G be a subgroup of $SL_d(\mathbb{Z})$, and A be as above. Then under either one of the conditions (i) or (ii) in Theorem 6.5, the G-action on A is strongly ergodic.*

In a way similar to Example 6.7, one can now obtain, using Theorem 6.8, a random algorithm for distributing points uniformly on $\mathbb{Z}_p^d$.

Semisimple groups. As was previously indicated, for countable groups the connection between strong ergodicity and the unitary representation associated with the action is completely understood. For non discrete groups the structure is, in general, more subtle. However for the distinguished class of semisimple groups we do have a much more accurate information, as the following analog of Moore's theorem [Mo] shows:

Theorem 6.9 ([?]). *Let $G = \Pi_1^n G_i$ be a semisimple Lie group with finite center, and $(\pi, \mathcal{H})$ be a unitary G-representation with $1_{G_i} \not\prec \pi|_{G_i}$ for each $1 \le i \le n$. Then*

(i) If $H \subseteq G$ is a closed non-amenable subgroup, then $1_H \not\prec \pi|_H$.

(ii) If μ is a probability distribution on G, which is not supported on a closed amenable subgroup, then the spectral radii satisfy: $r_{sp}(\pi(\mu)) < 1$ and $r_{sp}(\pi \otimes \pi(\mu)) < 1$.

In particular, for simple Lie groups we deduce a result which is stronger than the one for discrete groups (compare with Theorem 6.3):

Theorem 6.10. *Let G be a connected non-compact simple Lie group with finite center, and let μ be a probability measure on G, which is not supported on a coset of a closed amenable subgroup in G. Then for any G-action on (X, m) the following conditions are equivalent:*

(i) $\|\pi(\mu)\| < 1$.

(ii) $\|\pi \otimes \pi(\mu)\| < 1$.

(iii) The G-action on (X, m) is strongly ergodic.

(iv) $1_G \not\prec \pi$.

Moreover, if any of the above conditions hold, then for any closed non-amenable subgroup $H \subseteq G$, the H-action on (X, m) is strongly ergodic as well.

Proof. For any locally compact group, both (i) and (ii) imply (iv). The implication $(iii) \Rightarrow (iv)$, follows from Theorem 1.6. By Theorem **??** (iv) implies (i) and (ii), once we notice that under the assumption on μ, the measure $\check{\mu} * \mu$ is not supported on a closed amenable subgroup, and $r_{sp}(\pi(\check{\mu} * \mu)) = \|\pi(\check{\mu} * \mu)\| = \|\pi(\mu)\|^2$. To prove that (iv) implies (iii), as well as the last assertion, consider any finitely generated subgroup $H_1 \subseteq H$ with non-amenable closure. Then, taking a measure μ supported on generators and using Theorems **??** and 6.3, one obtains that already the H_1-action on (X, m) is strongly ergodic. □

Notice that as in the case of automorphisms of a compact abelian group, the conditions $\|\pi(\mu)\| < 1$ and $\|\pi \otimes \pi\,(\mu)\| < 1$ are equivalent. Thus strong ergodicity implies also the mixing results of section 4.

A primary object in the ergodic theory of semisimple groups is the action on homogeneous spaces G/Γ, where Γ is a lattice (i.e. Γ is a discrete subgroup with finite co-volume in G). If Γ is a uniform (i.e. co-compact) lattice in any locally compact group G, then it is not difficult to show (see e.g. [Ma] III.1.10):

$$1_G \not\prec \pi \quad \text{for the representation} \quad (\pi, L^2_0(G/\Gamma)). \tag{6.1}$$

However, (6.1) does not hold for general non uniform lattices: in [BL] a tree lattice is constructed, which is a counterexample (we thank A. Lubotzky for bringing this fact to our attention). Nevertheless, when G is a connected semisimple Lie group and $\Gamma \subseteq G$ is any (also non-uniform) lattice, (6.1) holds, a fact which relies on rather deep results concerning the geometry of such spaces (for a proof see [Be]). As remarked in [KM], it seems very reasonable that actually for any irreducible lattice Γ in a semisimple Lie group $G = \Pi G_i$, for every i: $1_{G_i} \not\prec \pi|_{G_i}$. This information is required in order to apply Theorem **??**. As in general we know only the weaker property (6.1), the following is needed:

Theorem 6.11 ([?]). *Let $G = \Pi G_i$ be as in Theorem* **??** *and let π be a unitary G-representation, such that $1_G \not\prec \pi$. Then statement 1 in Theorem* **??** *holds, if we assume that the projection of H to* every *simple factor does not have an amenable closure. Statement 2 holds, if we assume that the projection of μ to* every *simple factor is not supported on a closed amenable subgroup.*

Theorem 1.8 of the introduction follows easily from the above result, by taking an appropriate measure μ, exactly as in the proof of 6.9. The case $H = G$ in Theorem 1.8 was treated in [Be].

7. Appendix A: On a coboundary problem

Recall that if G be a locally compact group, acting ergodically on a probability measure space (X, m), then a measurable function $\alpha : G \times X \to H$ with values in a locally compact group H, is said to be a (measurable) *cocycle* if for every $g_1, g_2 \in G$: $\alpha(g_1 g_2,\, x) = \alpha(g_1,\, g_2 x) \cdot \alpha(g_2,\, x)$ for m-a.e. $x \in X$. A cocycle α is said to be a (measurable) *coboundary* in H, if there exists a measurable function $\phi : X \to H$, such that for every $g \in G$ and m-a.e. on X: $\alpha(g,\, x) = \phi(gx)\phi^{-1}(x)$.

Remark 7.1. Let $F = \langle S \rangle$ be a free group (where S is a set of free generators), acting ergodically on a probability space (X, m), and let H be any locally compact group. Then any collection of measurable functions $f_s : X \to H$, $s \in S$, uniquely defines a measurable H-valued cocycles $\alpha : F \times X \to H$, such that $\alpha(s,x) = f_s(x)$.

Our goal is to prove the following:

Proposition 7.2. *Let G be a countable group, acting ergodically on a probability measure space (X,m), μ be a generating probability measure on G, and let f a measurable function on (X,m). Suppose f is a measurable coboundary as an $\mathbb{R}$-function on $(\bar{\Omega} \times X, \bar{\mathcal{P}} \times m, T)$, i.e. there exists a measurable function $h(\omega,x)$ on $\bar{\Omega} \times X$, such that*

$$f(x) = (h \circ T - h)(\omega, x) = h(\theta\omega, \omega_1 x) - h(\omega, x), \quad \bar{\mathcal{P}} \times m - \text{a.e.} \tag{7.1}$$

Then h does not depend on ω, i.e. $h(\omega,x) = \phi(x)$, and $\phi(gx) = \phi(x)$ for $\check{\mu} * \mu$-a.e. $g \in G$.

First observe that Proposition 3.4 follows from 7.2. Indeed, in the case where μ is aperiodic and G acts ergodically, ϕ is a.e. a constant, and hence $f(x) = 0$ a.e., as required. The case where the spectral gap condition $\|\pi(\mu)\| < 1$ is satisfied, together with the assumption that $h \in L^2$ in (7.1) is established as follows: by Proposition 7.2, $h(\omega,x) = \phi(x)$ is in $L^2_0(X,m)$ and $\phi(gx) = \phi(x)$ for $\check{\mu} * \mu$-a.e. $g \in G$. Thus

$$\|\phi\|^2 = \langle \pi(\check{\mu} * \mu)\phi,\, \phi\rangle = \|\pi(\mu)\phi\|^2$$

and $\|\pi(\mu)\| < 1$ implies $f(x) = \phi(x) = 0$.

For the proof of Proposition 7.2 we shall need some auxiliary results:

Lemma 7.3. *Let G be a locally compact group, acting ergodically on a probability space (X,m), and let $\alpha : H \times X \to S^1$ be a measurable cocycle. Suppose that the α-defined skew-product $(X \times_\alpha S^1, G)$, given by $g\,(x,\, e^{it}) = (g\,x, \alpha(g,\, x) \cdot e^{it})$ is not ergodic. Then for some integer $n \geq 1$, α^n is a coboundary, i.e. $\alpha^n(g,x) = \phi(gx)/\phi(x)$ for some measurable $\phi : X \to S^1$.*

Proof. (This is essentially Zimmer's "reduction of the range of a cocycle" argument, see [Zi] 5.2.11). Let $F : X \times S^1 \to \mathbb{R}$ be a non-trivial G-invariant L^2_0-function. By Fubini's theorem, $F(x,\cdot)$ is a family of $L^2(S^1)$-functions, satisfying

$$F(gx,\cdot) = \alpha(g,\, x) \cdot F(x,\cdot). \tag{7.2}$$

Since $\int_0^{2\pi} F(x, e^{it})\, dt$ is a G-invariant zero-mean function on X, we conclude that for m-a.e. $x \in X$, $F(x,\cdot)$ has zero mean on S^1, i.e. $F(x,\cdot) \in L^2_0(S^1)$.

Now consider the space $V = L^2_0(S^1)$ as a Borel space with the natural S^1-action given by the translations. Since S^1 is compact, the S^1-action on V is *smooth*, i.e., the space of S^1-orbits in V is countably separated, in the sense that there exists a countable collection $\{V_i\}$ of Borel S^1-invariant subsets V_i of V, which separate S^1-orbits in V. This holds for any compact group action (see [Zi] 2.1.21), although in our particular case, such separating sets may be constructed explicitly using the Fourier coefficients.

Now, considering the measurable sets $E_i = \{x \in X \mid F(x,\cdot) \in V_i\}$, we note that by (7.2) the sets E_i-s are G-invariant, and therefore have m-measure 0 or 1. Intersecting those

E_i which have the full m-measure and the complements of all other E_i-s, we remain with a single S^1 orbit in V, supporting m-a.e. $F(x,\cdot)$. Hence $F(x,\cdot)=\phi(x)v_0$ for some (measurable) $\phi : X \to S^1$ and $v_0 \in V$. In particular, $\alpha(g,\, x)\phi(x)v_0 = \phi(gx)v_0$, so that $\alpha(g,\, x)\phi(x)/\phi(gx)$ takes values in the stabilizer of v_0. Since the only proper closed subgroups of S^1 are finite, we conclude that for some $n \geq 1$, $\alpha^n(g,\, x) = \phi(gx)/\phi(x)$ is a coboundary. □

We shall need the following application of results due to C. C. Moore and K. Schmidt characterizing coboundaries by their mappings to the circles:

Theorem 7.4 ([MS]). *Let G be a locally compact group acting measurably on a probability space (X,m), and let $a : G\times X \to \mathbb{R}$ be a measurable cocycle. Let $B_a \subseteq \mathbb{R}$ be the set of all λ, such that the cocycle $\alpha_\lambda : G \times X \to S^1$, defined by: $\alpha_\lambda(g,x) = e^{i\cdot\lambda\cdot a(g,x)}$ is a measurable coboundary in S^1, i.e. $\alpha_\lambda(g,x) = \phi_\lambda(gx)/\phi_\lambda(x)$. Then B_a is a Borel subgroup of $\mathbb{R}$, so if it has a positive Lebesgue measure, then $B_a = \mathbb{R}$. Moreover, in that case $a(g,x)$ is a measurable coboundary in $\mathbb{R}$: $a(g,x) = b(gx) - b(x)$ for some measurable $b : X \to \mathbb{R}$.*

Proof. Measurability of B_a is the content of Proposition 4.1 in [MS]. The second statement follows from Theorem 4.3 in [MS].. □

Proof of Proposition 7.2. We first claim that without loss of generality, the group G can be assumed to be free. Indeed, there exists a free group $\tilde G$ with at most a countable set S of generators, a probability measure $\tilde\mu$ with $\mathrm{supp}(\tilde\mu) = S$ and a homomorphism $p : \tilde G \to G$, such that $p(\tilde\mu^{(n)}) = \mu^{(n)}$ (one can take $\mathrm{supp}(\mu)$ as the set S of the free generators for $\tilde G$ with the $\tilde\mu$-weights given by μ). Defining the $\tilde G$-action by $\tilde g\, x = p(\tilde g)\, x$, we may, and shall, replace G by $\tilde G$ and μ by $\tilde\mu$ in the assumptions and conclusions of Proposition 7.2.

Hence, we assume hereafter that the group G is free, and the measure μ is supported on the set S of the free generators. With this assumption the measurable function $f : X \to \mathbb{R}$ uniquely defines a cocycle $a : G \times X \to \mathbb{R}$ by $a(g,x) = f(x)$ for $g \in S$ (see Remark 7.1). For each $\lambda \in \mathbb{R}$ consider the measurable cocycle $\alpha_\lambda : G \times X \to S^1$, given by: $\alpha_\lambda(g,x) = e^{i\cdot\lambda\cdot a(g,x)}$, and the α_λ-defined skew-product $(X \times_{\alpha_\lambda} S^1, G)$.

Observe that the skew-product $(\bar\Omega \times_{\omega_1} (X\times_{\alpha_\lambda} S^1), T_1)$ where:

$$T_1(\omega,\, (x, e^{it})) = (\theta\omega,\, \omega_1(x, e^{it})) = (\theta\omega,\, (\omega_1 x, e^{i(t+\lambda f(x))})) \tag{7.3}$$

with $(\bar\Omega,\theta)$ being the base, and $(X\times_{\alpha_\lambda} S^1, G)$ being the fiber, is naturally isomorphic to the λf-defined skew-product $((\bar\Omega \times_{\omega_1} X) \times_{\lambda\cdot f} S^1, T_2)$ where:

$$T_2((\omega, x),\, e^{it}) = (T(\omega, x),\, e^{i(t+\lambda f(x))}) = ((\theta\omega, \omega_1 x),\, e^{i(t+\lambda f(x))}) \tag{7.4}$$

with $(\bar\Omega \times X, T)$ being the base and S^1 - the fiber (associativity of the skew-product construction).

By the assumption, the function f (as well as $\lambda\cdot f$) forms a coboundary on $(\bar\Omega \times_{\omega_1} X)$. Therefore, for each $\lambda \in \mathbb{R}$ the system (7.4), and hence also (7.3), is not ergodic. Since μ generates G, Kakutani's random ergodic theorem implies that the G-action on $(X\times_{\alpha_\lambda} S^1)$ is not ergodic. By Lemma 7.3, every α_λ has a power n_λ such that $\alpha_\lambda^{n_\lambda} = \alpha_{n_\lambda\cdot\lambda}$ is a coboundary in S^1. This implies that for each λ, there exists n_λ such that $n_\lambda \cdot \lambda \in B_a$ (see the notation of Theorem 7.4). But this implies that B_a has positive Lebesgue measure, and thus, by Theorem 7.4, the cocycle $a(g,x)$ is a measurable coboundary, i.e. for some measurable $\phi : X \to \mathbb{R}$: $a(g,x) = \phi(gx) - \phi(x)$ holds for all $g \in G$. In particular,

$$\forall g \in \mathrm{supp}(\mu): \quad f(x) = \phi(gx) - \phi(x)$$

This shows that (7.1) holds with $h(\omega, x) = \phi(x)$. Since $\bar{\Omega} \times X$ is ergodic, the equation (7.1) determines h uniquely up to a constant, and thus $h(\omega, x) = \phi(x)$. We have for μ-a.e. $g \in G$, $\phi(gx) = \phi(x) + f(x)$, and therefore ϕ is invariant under $\check{\mu} * \mu$-a.e. $g \in G$. □

Remark 7.5. Proposition 7.2 holds for a general locally compact group G. The assumption that G is countable was used to construct a free group $\tilde{G}$ which acts on the $\lambda f(x)$-defined skew-product $(X \times S^1)$. However, the topological structure of the group $\tilde{G}$ is immaterial, since Kakutani's theorem holds for any *measurable family* of transformations, and elements of G with the distribution μ form such a family, each of them giving rise to a measurable $\lambda f(x)$-defined transformation on $(X \times S^1)$.

8. Appendix B: $(\bar{\Omega} \times X, \bar{\mathcal{P}} \times m, T)$ is a K-system

In this section we prove an incidental result, concerning the structure of the skew-product $(\bar{\Omega} \times X, \bar{\mathcal{P}} \times m, T)$. The following theorem is a generalization of the $\mathbb{Z}$-case, proved in [Me].

Theorem 8.1. *Let G be a locally compact group, acting ergodically on a probability measure space (X, m). Let μ be a probability measure on G, such that $\pi(\mu^{(n)}) \to 0$ in the strong operator topology. Then T is an exact endomorphism of $(\Omega \times X, \mathcal{P} \times m)$, and the invertible system $(\bar{\Omega} \times X, \bar{\mathcal{P}} \times m, T)$ is a K-system.*

The assumption $\pi(\mu^{(n)}) \to 0$ is satisfied if either one of the following properties holds:

(i) μ is symmetric *and* aperiodic*, or*

(ii) for some $k \geq 1$, the convolution powers $\mu^{(k)}$ and $\mu^{(k+1)}$ are not mutually singular.

Proof of Theorem 8.1. Assuming that $\lim_{n\to\infty} \|\pi(\mu^{(n)}\,\phi\| = 0$ for every $\phi \in L^2_0(X, m)$, we shall prove that $(\bar{\Omega} \times X, \bar{\mathcal{P}} \times m, T)$ is a K-system. Since an inverse limit of K-systems is a K-system, it is enough to construct a sequence $\{\xi_k\}$ of (finite) measurable partitions, so that $\mathcal{F} = \bigvee \xi_k$ modulo μ, and each ξ satisfies the following "uniform mixing" property:

$$\forall f \in L^2_0(\xi) \quad \lim_{n\to\infty} \sup\{\ |\langle f, g\rangle|\ \mid\ g \in L^2_0(\bigvee_{-\infty}^{-n} T^i \xi),\ \|g\| = 1\ \} = 0 \tag{8.1}$$

We shall use partitions ξ of the form $\xi = \alpha \vee \beta$, where α is any partition of $(\bar{\Omega}, \bar{\mathcal{P}})$, depending on a *finite* number of coordinates ω_i, $|i| \leq N$. The partition β will be any finite measurable partition of X, i.e. $\beta \subset \mathcal{B}$. Since such ξ-s generate all $\mathcal{F}$, it is enough to show (8.1) for

$\xi = \alpha \vee \beta$. Observe that $\xi \subset \bigvee_{-N}^{N} T^i\mathcal{B}$, and therefore:

$$L_0^2(\bigvee_{-\infty}^{-n} T^i\xi) \subset L_0^2(\bigvee_{-\infty}^{-n+N} T^i\mathcal{B}), \qquad L_0^2(\xi) \subset L_0^2(\bigvee_{-N}^{\infty} T^i\mathcal{B})$$

We can now apply the first inequality in (3.4) of Lemma 3.3 to deduce (8.1), thereby proving the first assertion.

We are left with the second part. Assume (i), i.e. μ is symmetric and aperiodic. Then $\pi(\mu)$ is a self-adjoint operator, and thus by the spectral representation

$$\|\pi(\check{\mu})^n\phi\|^2 = \langle \pi(\mu)^{2n}\phi,\, \phi\rangle = \int_{-1}^{1} t^{2n}\, d\nu_\phi(t) \tag{8.2}$$

where ν_ϕ on $[-1,1]$ is the spectral measure corresponding to $\pi(\mu) = \pi(\check{\mu})$ and ϕ. By the dominated convergence theorem the expression in (8.2) converges to 0, as soon as we show that $\nu_\phi\{-1\} = \nu_\phi\{1\} = 0$. If the latter does not hold, there exists $\psi \in L_0^2(X,m)$ with $\pi(\mu)\psi = \pm\psi$ so that $\pi(\check{\mu} * \mu)\psi = \psi$. By convexity, $\pi(g)\psi = \psi$ for $(\check{\mu} * \mu)$-a.e. $g \in G$, and hence, for all the g-s in the smallest closed group $H \subset G$, supporting $\check{\mu} * \mu$. Since μ is aperiodic, $H = G$, and we contradict to the fact that G has no fixed vectors in $L_0^2(X,m)$.

Now assume that condition (ii) is satisfied. Then denoting $S = \pi(\check{\mu})$, the adjoint operator $S^* = \pi(\mu)$. The assumption that $\mu^{(k)}$ and $\mu^{(k+1)}$ are not mutually singular, implies that the operator norms satisfy: $\|S^k + S^{k+1}\| = \|\pi(\check{\mu}^{(k)} + \check{\mu}^{(k+1)})\| < 2$. By the "0-2" law (see [KT]) this implies that for any ψ:

$$\lim_{n\to\infty} \|(S^n - S^{n+1})\psi\| = \lim_{n\to\infty} \|S^n\,(1-S)\psi\| = 0.$$

In our case the operator $(1-S)$ is onto, since the dual operator $(1-S^*) = (1-\pi(\mu))$ is 1-1 (this follows from ergodicity of the G-action and the fact that μ is generating). Therefore for any $\phi \in L_0^2(X,m)$, there exists $\psi \in L_0^2(X,m)$ with $\|S^n\phi\| = \|S^n\,(1-S)\psi\| \to 0$. $\square$

References

[BL] H. Bass and A. Lubotzky, Non uniform tree lattices, *in preparation*.

[Be] M. Bekka, On uniqueness of invariant means, *to appear in Proc. of AMS*.

[BR] V. Bergelson and J. Rosenblatt, Mixing actions of groups, *Ill. J. Math.* **32** (1988), 65–80.

[CW] A. Connes and B. Weiss, Property (T) and asymptotically invariant sequences, *Israel J. Math.* **37 no. 3** (1980), 209–210.

[DJR] A. del-Junco and J. Rosenblatt, Counterexamples in ergodic theory and number theory, *Math. Ann.* **245** (1979), 185–197.

[Dy] H. A. Dye, On the ergodic mixing theorem, *Trans. Amer. Math. Soc.* **118** (1965), 123–130.

[FS] A. Furman and Y. Shalom, Random walks on Hilbert spaces and Lyapunov exponents, *in preparation*.

[Fu] H. Furstenberg, A note on Borel's density theorem, *Proc. AMS* **55** (1976), 209–212.

[HM] R. E. Howe and C. C. Moore, Vanishing of matrix coefficients, *J. of Func. Anal.* **32** (1979), 72–96.

[KSZ] M. Kac, R. Salem and A. Zygmund, A gap theorem, *Trans. of Amer. Math. Soc.*, **63 no 2**, (1948), 235–243.

[Ka] S. Kakutani, Random ergodic theorems and Markoff processes with a stable distribution, Proc. Second Berkeley Symp. (1955), 247–261.
[KT] Y. Katznelson and L. Tzafriri, On power bounded operators, *J. Func. Anal.* **68** no 3, (1986), 313–328.
[KM] D. Y. Kleinbock and G. A. Margulis, Bounded orbits of nonquasiunipotent flows on homogeneous spaces, *preprint.*
[Kr] U. Krengel, On the speed of convergence in the ergodic theorem, *Monatsh. Math.* **86** no. 1 (1979), 3–6.
[Lu] A. Lubotzky, *Discrete Groups, Expanding Graphs and Invariant Measures*, Birkhauser, Progress in Math, 1994.
[LPS] A. Lubotzky, R. Phillips and P. Sarnak, Hecke operators and distributing points on S^2 I., *Comm. Pure and Applied Math.* **39** (1986), S149–186.
[Ma] G. A. Margulis, *Discrete subgroups of semisimple Lie groups*, Springer-Verlag, 1991.
[Me] I. Meilijson, Mixing properties of a class of skew-products, *Isr. J. Math.* **19** (1974), 266–270,
[Mo] C. C. Moore, Ergodicity of flows on homogenneous spaces, *Amer. J. Math.* **88** (1966), 154–178.
[MS] C. C. Moore and K. Schmidt, Coboundaries and homomorphisms for non-singular actions and a problem of H. Helson, *Proc. London Math. Soc.* (3) **40** (1980), 443–475.
[Na] I. Namioka, Folner's conditions for amenable semi-groups, *Math. Scand.* **15** (1964), 18–28.
[OY] H. Oodaira and K-I. Yoshihara, Functional central limit theorems for strictly stationary processes satisfying the strong mixing conditions, *Kodai Math. Sem. Rep.* **24** (1972), 259–269.
[OW1] D. Ornstein and B. Weiss, Entropy and isomorphism theorems for actions of amenable groups, *J. D'Analyse Math.* **48** (1987), 1–141.
[OW2] D. Ornstein and B. Weiss, Ergodic theory of amenable groups actions I. The Rohlin Lemma, *Bull. Am. Math. Soc.* **2** (1980), 161–164.
[OW3] D. Ornstein and B. Weiss, The Shannon-McMillan-Breimann Theorem for a class of amenable groups, *Israel J. of Math.* **44 no. 1** (1983), 53–60.
[Ro] J. Rosenblatt, *Uniqueness of invariant means for measure preserving transformations*, Trans. of Amer. Math. Soc. **265** (1981), 623–636.
[Sc] K. Schmidt, Amenability, Kazhdan's property (T), strong ergodicity and invariaant means for ergodic group actions, *Ergod. Th. Dyn. Syst.* **1** (1981),223 –236.
[Sh1] Y. Shalom, Explicit Kazhdan constants for representations of semisimple groups and their lattices, *preprint.*
[Sh2] Y. Shalom, Invariant measures for algebraic actions, Zariski dense subgroups and Kazhdan's property (T), *to appear in Trans. of AMS.*
[Ti] A. N. Tihomirov, On the convergence rate in the central limit theorem for weakly dependent random variables, *Theor. Prob. and Appl.* **25** no 4 (1980), 790–808. (1980).
[Zi] R. Zimmer, *Ergodic theory and semisimple Lie groups*, Birkhauser, Boston 1984.
[Zy] A. Zygmund, *Trigonometric series* **2**, Cambrodge Univercity Press, 1959.

Penn State University, 218 McAllister Bldg, University Park, PA 16802, US
e-mail: furman@math.psu.edu

Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
e-mail: yehuda@math.huji.ac.il