ILLINOIS JOURNAL OF MATHEMATICS
Volume 43, Number 3, 1999

OPERATORS COMMUTING WITH MIXING SEQUENCES

M. D. HA

ABSTRACT. Let $(X, \mathcal{F}, \mu)$ be a probability space and let $L^2(X, 0)$ be the collection of all $f \in L^2(X)$ with zero integrals. A collection $\mathcal{A}$ of linear operators on $L^2(X)$ is said to satisfy the Gaussian-distribution property (G.D.P.) if $L^2(X, 0)$ is invariant under $\mathcal{A}$ and there exists a constant $C < \infty$ such that the following condition holds:

Whenever $T_1, \ldots, T_k$ are finitely many operators in $\mathcal{A}$, and f is a function in L_2 with zero integral, then, for any required degree of approximation, there is another L_2-function g with $\| g \|_2 \leq C \| f \|_2$, such that all the inner products $(\operatorname{Re} T_i g, \operatorname{Re} T_j g)$ are approximately equal to the corresponding inner products $(\operatorname{Re} T_i f, \operatorname{Re} T_j f)$ for all $1 \leq i, j \leq k$ and such that the joint distribution of the functions $\operatorname{Re} T_1 g, \ldots, \operatorname{Re} T_k g$ is approximately Gaussian.

It has been proved that if $(S_n)_1^\infty$ is a sequence of uniformly bounded linear operators on $L^2(X)$ that satisfies the Bourgain's infinite entropy condition and the G.D.P., then there exists an $h \in L^2(X)$ such that $\lim_{n\to\infty} S_n h$ fails to exist μ-a.e. as a finite limit on X.

The purpose of this paper is to provide sufficient conditions for a collection $\mathcal{A}$ of linear operators on $L^2(X)$ to satisfy the G.D.P.

1. Introduction

The Zygmund-Marcinkiewicz conjecture states that if f is a 1-periodic, bounded measurable function, then the sequence of Riemann sums $(R_n f)_1^\infty$ converges a.e. on [0,1) to $\int_{[0,1)} f \, d\lambda_1$. Here, $R_n f(x) = \frac{1}{n} \sum_{k=1}^n f(x + \frac{k}{n})$, $\forall x \in [0, 1)$ and λ_1 is the Lebesgue measure on $\mathbf{R}$. This was disproved by Rudin in 1962 [9]. The method employed by Rudin essentially made use of only the arithmetic properties of the primes, namely that if $p_1, p_2, \ldots, p_N$ are distinct primes, then for any $1 \leq i \leq N$, p_i does not divide the least common multiple of $p_1, p_2, \ldots, p_{i-1}, p_{i+1}, \ldots, p_N$.

Also, in 1969, Marstrand [7] proved that if $(a_k)_1^\infty$ is any $\mathcal{L}$-sequence in $\mathbb{N}$ (see [7] for the definition), then given any $\epsilon > 0$, there is some open set $\mathcal{O} \subseteq (0, 1)$ for which $\lambda_1(\mathcal{O}) < \epsilon$ and

$$\limsup_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \chi_{\mathcal{O}} (a_k x (\text{mod} 1)) = 1, \quad \forall x \in (0, 1).$$

The sequence $(k)_1^\infty$ was then shown to be an $\mathcal{L}$-sequence, and thus, applying the above, Khintchine's conjecture [7] was settled. Again, the arguments given by Marstrand depend essentially only on the properties of an $\mathcal{L}$-sequence. The approaches taken by both Rudin and Marstrand in the disproof of the respective conjectures were

Received January 14, 1998.
1991 Mathematics Subject Classification. Primary 28D99; Secondary 60F99.

apparently ad hoc and seem to depend strongly on each particular situation being encountered. Therefore, it became clear that a more unified and general principle, one that would somehow allow us to "simultaneously" settle these and other related almost everywhere convergence problems, was very much needed.

Then in 1987, Bourgain [3] established a criterion providing necessary conditions for a.e. convergence in a very general setting (so that many types of operators encountered in ergodic theory are covered). Bourgain's proof of the result above made use of several ideas and theorems in the theory of Gaussian Processes.

2. Some motivations

Let λ_1 be the Lebesgue measure on $\mathbf{R}$. For each real number a, let τ_a: $[0, 1) \to [0, 1)$ be the translation $x \mapsto x + a \pmod 1$. A *weighted averages of translations on the unit interval* is a contraction A defined for all $1 \le p \le \infty$ by

$$A : L^p\Big([0,1),\lambda_1\Big) \to L^p\Big([0,1),\lambda_1\Big)$$
$$Af = \sum_{j=1}^{\infty} \alpha_j f \circ \tau_{a_j} \quad \text{for all } f \in L^p\Big([0,1),\lambda_1\Big),$$

where $(\alpha_j)_1^\infty$ is a sequence of non-negative reals with $\sum_{j=1}^\infty \alpha_j = 1$ and $(a_j)_1^\infty$ is a sequence of real numbers.

In an attempt to simplify and better understand Bourgain's proof of the Entropy Criteria, M. A. Akcoglu, M. D. Ha and R. L. Jones [1] discovered and proved the following.

(i) Every sequence (A_n) of weighted average of translations on the unit interval has an interesting property, which was called the *Gaussian-distribution property*.

(ii) Moreover, if a sequence (A_n) of weighted averages of translations on the unit interval has infinite L^2-entropy, then, using the Gaussian-distribution property of (A_n), it was shown that there exists some $f \in L^\infty([0, 1), \lambda_1)$ such that the a.e. convergence of $(A_n f)_1^\infty$ fails.

The techniques used in [1] were later refined in [2] so that in certain cases, much stronger results can be obtained. The methods employed in [1] and [2] differ substantially from the one used by Bourgain in [3] and the proofs given are self-contained, avoiding specialized estimates used in the theory of probability.

The purpose of this paper is to extend some of the results in [1] by establishing a general criterion whereby various classes of operators in ergodic theory can be shown to also have the Gaussian Distribution Property. This is the contents of Theorems I and II (Section 4) of this paper.

3. Preliminaries

3.1. *Weak convergence of measures.* For any topological space Y, let $\mathcal{B}(Y)$ be the σ-algebra of all Borel sets in Y. We denote by $M(Y)$ the collection of all probability measures on $\mathcal{B}(Y)$. Let $\mathcal{C}_b(Y)$ be the collection of all real-valued bounded continuous functions on Y. Each φ in $\mathcal{C}_b(Y)$ induces a map

$$\begin{aligned} \varphi^* \ &: \ M(Y) \to \mathbf{R} \\ \mu &\mapsto \int_Y \varphi\, d\mu, \ \mu \in M(Y). \end{aligned}$$

The *weak topology* on $M(Y)$ is the smallest topology on $M(Y)$ making each of the maps φ^* above continuous. A sequence $(\mu_n)_1^\infty$ in $M(Y)$ *converges weakly to* $\mu \in M(Y)$ if the convergence takes place in the weak topology on $M(Y)$. Equivalently, $(\mu_n)_1^\infty$ converges weakly to μ iff

$$\lim_{n\to\infty} \int_Y \varphi\, d\mu_n = \int_Y \varphi\, d\mu \quad \text{for all } \varphi \in \mathcal{C}_b(Y).$$

We denote this weak convergence of $(\mu_n)^\infty$ to μ symbolically by

$$\mu_n \rightharpoonup \mu \quad \text{as } n \to \infty.$$

3.2. *Distribution measures.* Let $(X, \mathcal{F}, \mu)$ be a probability space and let Y be a topological space. The *distribution measure* ν of a measurable function $f\colon X \to Y$ is the measure on Y defined by

$$\nu(E) = \mu\left(f^{-1}(E)\right), \ E \in \mathcal{B}(Y).$$

We shall write $\mu \circ f^{-1}$ for the distribution measure of f. If $f_1, \dots, f_N$ are measurable functions from X into Y, $(f_1, \dots, f_N)\colon X \to Y^N$ is the function defined by $(f_1, \dots, f_N)(x) = (f_1(x), \dots, f_N(x))$ for every $x \in X$. It is clear that $(f_1, \dots, f_N)$ is measurable when Y^N is given the product topology. It then makes sense to talk about $\mu \circ (f_1, \dots, f_N)^{-1}$, the *joint-distribution measure of* $f_1, \dots, f_N$.

3.3. *Gauss measures on* $\mathcal{B}(\mathbf{R}^N)$. Unless otherwise stated, $x \in \mathbf{R}^n$ will stand for the point $x = (x_1, \dots, x_N)$. The usual inner-product on $\mathbf{R}^N$ will be denoted by $\langle\ ,\ \rangle$, i.e., $\langle x, y\rangle = \sum_{i=1}^N x_i y_i$ for all $x, y \in \mathbf{R}^N$. Here and elsewhere, $\bigotimes_{j=1}^n \mathcal{F}_j$ is the product σ-algebra of the σ-algebras $\mathcal{F}_j$. We shall identify $\mathbf{R}^{KN}$ with $(\mathbf{R}^N)^K$. Hence, if λ is a probability measure on $\mathcal{B}(\mathbf{R}^N)$ and $H\colon \left((\mathbf{R}^N)^K,\ \bigotimes_{j=1}^K \mathcal{B}(\mathbf{R}^N)\right) \to \mathbf{R}$ is measurable, we will write $\int_{\mathbf{R}^{KN}} H\, d\lambda^K$ for the more cumbersome $\int_{(\mathbf{R}^N)^K} H\, d\lambda^K$.

For each positive integer m, the *standard m-dimensional Gaussian density function*, $\mathcal{G}_m$, is defined to be

$$\mathcal{G}_m\colon \mathbf{R}^m \to \mathbf{R}$$
$$\mathcal{G}_m(x) = \frac{1}{(2\pi)^{m/2}} e^{\frac{-1}{2}\langle x,x\rangle}, \quad x \in \mathbf{R}^m.$$

The above usage of terminology is in agreement with that of a probability theorist concerning density functions of random variables. For a lively account of these matters and more, the reader can consult [4].

Throughout this paper, λ_m denotes the m-dimensional Lebesgue measure restricted to $\mathcal{B}(\mathbf{R}^m)$, the Borel subsets of $\mathbf{R}^m$.

Definition 3.3.1. For each positive integer m, the *standard m-dimensional Gaussian measure* is the probability measure γ_m on $\mathcal{B}(\mathbf{R}^m)$ defined by

$$\gamma_m(E) = \int_E \mathcal{G}_m \, d\lambda_m, \quad E \in \mathcal{B}(\mathbf{R}^m).$$

If N is any positive integer, a *Gauss measure* γ on $\mathcal{B}(\mathbf{R}^N)$ is the distribution measure of a linear transformation

$$L\colon \left(\mathbf{R}^m, \mathcal{B}(\mathbf{R}^m), \gamma_m\right) \to \mathbf{R}^N$$

for some integer m. Here, we do not require L to be one to one.

For each $1 \le i \le N$, we will let $\pi_i\colon \mathbf{R}^N \to \mathbf{R}$ denote the projection map onto the i^{th}-coordinate. In [6], the following result was deduced from the Multi-dimensional Central Limit Theorem (Theorem 11.10 in [4]).

3.4. THE CENTRAL LIMIT THEOREM FOR GAUSSIAN MEASURES. *Fix some positive integer N. Let λ be any probability measure on $\mathcal{B}(\mathbf{R}^N)$ such that for all $1 \le i, j \le N$,*

$$\int_{\mathbf{R}^N} \pi_i \, d\lambda = 0 \quad \text{and} \int_{\mathbf{R}^N} \pi_i \pi_j \, d\lambda < \infty.$$

For each $K \in \mathbf{N}$, define $T_K\colon \mathbf{R}^{KN} \to \mathbf{R}^N$ by

$$T_K(x^1, \ldots, x^K) = \frac{x^1 + \cdots + x^K}{\sqrt{K}}, \quad x^1, \ldots, x^K \in \mathbf{R}^N.$$

Then, there exists a unique Gauss measure γ on $\mathcal{B}(\mathbf{R}^N)$ satisfying

$$\int_{\mathbf{R}^N} \pi_i \pi_j \, d\gamma = \int_{\mathbf{R}^N} \pi_i \pi_j \, d\lambda \quad \text{for all } 1 \le i, j \le N.$$

Moreover,

$$\lambda^K \circ T_K^{-1} \rightharpoonup \gamma \quad \text{as } K \to \infty.$$

From this result, we obtain the following.

COROLLARY. *Let $(X, \mathcal{F}, \mu)$ be a probability space. Let $f_1, \dots, f_N \in L^2(X)$ and $\int_X f_i \, d\mu = 0$ for all $i = 1, 2, \dots, N$. Then, there exists a unique Gauss measure γ on $\mathcal{B}(\mathbf{R}^N)$ such that*

$$\int_{\mathbf{R}^N} \pi_i \pi_j \, d\gamma = \int_X f_i f_j \, d\mu \quad \text{for all } 1 \le i, j \le N.$$

Proof. Let $\Omega = (f_1, \dots, f_N)$ and $\lambda = \mu \circ \Omega^{-1}$. Then

$$\int_{\mathbf{R}^N} \pi_i \, d\lambda = 0 \text{ and } \int_{\mathbf{R}^N} \pi_i \pi_j \, d\lambda = \int_X f_i f_j \, d\mu < \infty \quad \text{for all } 1 \le i, j \le N.$$

Hence, by the theorem immediately above, we are done. □

Definition 3.4.1. The Gauss measure γ in the above corollary is called the *Gauss measure induced by* $(f_1, \dots, f_N)$. We denote this γ by Gauss$(f_1, \dots, f_N)$.

3.4. *Mixing transformations and sequences.* Throughout this discussion, we let $(X, \mathcal{F}, \mu)$ be a fixed probability space. We will first recall some standard terminologies concerning various kinds of transformations on X that are frequently encountered in ergodic theory.

A map $\tau \colon X \to X$ is said to be *measure-preserving* if $\tau^{-1}A \in \mathcal{F}$ and $\mu(\tau^{-1}A) = \mu(A)$ for all $A \in \mathcal{F}$. A measure-preserving map $\tau \colon X \to X$ is said to be *ergodic* if whenever $\tau^{-1}A = A$ for some $A \in \mathcal{F}$, then $\mu(A) = 1$ or $\mu(A) = 0$. A measure-preserving map $\tau \colon X \to X$ is said to be *weakly mixing* if for all $A, B \in \mathcal{F}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \left|\mu\left(A \cap \tau^{-j}B\right) - \mu(A)\,\mu(B)\right| = 0,$$

and is *strongly mixing* if for all $A, B \in \mathcal{F}$,

$$\lim_{n\to\infty} \mu\left(A \cap \tau^{-n}B\right) = \mu(A)\,\mu(B).$$

Also, as is well known, strongly mixing $\Longrightarrow$ weakly mixing $\Longrightarrow$ ergodicity [8]. We shall be interested in sequences of measure-preserving maps on X that has a certain property of mixing as following.

Definition 3.5.1. Let $(\tau_n)_1^\infty$ be a sequence of measure-preserving transformations from X into X. Then $(\tau_n)_1^\infty$ is said to be *mixing of all orders* if for all $K \ge 1$ and all $A_1, A_2, \dots, A_K \in \mathcal{F}$,

$$\lim_{\substack{\inf_{1\le j\le K-1} \frac{m_{j+1}}{m_j} \to\infty \\ \inf_{1\le j\le K} m_j\to\infty}} \mu(\tau_{m_1}^{-1}A_1 \cap \tau_{m_2}^{-1}A_2 \cap \dots \cap \tau_{m_K}^{-1}A_K) = \mu(A_1)\,\mu(A_2)\dots\mu(A_K).$$

Remark 3.5.1. A slightly different notion of mixing of all orders from the one introduced in Definition 3.5.1 is the following. Let (τ_n) be a sequence of measure-preserving transformations on a probability space $(X, \mathcal{F}, \mu)$. Then $(\tau_n)_1^\infty$ is said to be *mixing of all orders* (of type II) if for all $K \geq 1$ and all $A_1, A_2, \ldots, A_K \in \mathcal{F}$,

$$\lim_{\substack{\inf_{1\le j\le K-1} (m_{j+1}-m_j)\to\infty \\ \inf_{1\le j\le K} m_j\to\infty}} \mu(\tau_{m_1}^{-1} A_1 \cap \tau_{m_2}^{-1} A_2 \cap \ldots \cap \tau_{m_K}^{-1} A_K) = \mu(A_1)\, \mu(A_2) \ldots \mu(A_K).$$

It is clear that if a sequence (τ_n) is mixing of all orders of type II, then it is mixing of all orders as defined in our definition above. Thus, mixing of all orders of type II is a more stringent condition imposed on a sequence (τ_n). For the present paper, Definition 3.5.1 suffices. Also, if τ is a measure-preserving map on X such that $(\tau^n)_1^\infty$ is mixing of all orders of type II, then τ is strongly mixing. The converse still remains an open problem for sometime.

3.5. *Examples.* (a) *Bernoulli shifts.* Let us first recall what we mean by Bernoulli shifts. Let X be a topological space, $\mathcal{B}(X)$ being the σ-algebra of its Borel sets, and let μ be a probability measure on $\mathcal{B}(X)$. Let $Y = \prod_{-\infty}^{\infty} X$ be endowed with the product topology. Given $A_0, \ldots, A_m \in \mathcal{B}(X)$ and $j \in \mathbf{Z}$, define a cylinder set in Y as

$$\mathcal{C}(j, A_0, \ldots, A_m) = \{(x_k)_{-\infty}^{\infty} \in Y\colon\ x_j \in A_0, x_{j+1} \in A_1, \ldots, x_{j+m} \in A_m\}.$$

It can be shown that there exists a unique probability measure ν on $\mathcal{B}(Y)$ satisfying

$$\nu\big(\mathcal{C}(j, A_0, \ldots, A_m)\big) = \prod_{i=0}^{m} \mu(A_i)\ , \quad j \in \mathbf{Z}\ , \quad A_i \in \mathcal{B}(X).$$

The map

$$\sigma\colon\ \big(Y, \mathcal{B}(Y), \nu\big) \to \big(Y, \mathcal{B}(Y), \nu\big)$$

defined by

$$\big(\sigma(\theta)\big)(n) = \theta(n+1), \quad n \in \mathbf{Z}\ , \ \theta \in Y$$

(considering an element θ of Y as a map $\theta\colon\ \mathbf{Z} \to X$) is called a *Bernoulli shift* on X. It can be shown that $(\sigma^n)_1^\infty$ is mixing of all orders. We call $\big(Y, \mathcal{B}(Y), \nu, \sigma\big)$ a *Bernoulli scheme.* In the special case when $X = \{1, 2, \ldots, n\}$ is given the discrete topology and $\mu(\{i\}) = p_i$ for each $i = 1, \ldots, n$, where $p_i \geq 0$, $\sum_{i=1}^{n} p_i = 1$, we denote the corresponding Bernoulli scheme as $\mathbf{B}(p_1, p_2, \ldots, p_n)$.

(b) *Continuous ergodic automorphisms of compact abelian groups.* Let X be a compact, abelian group equipped with the normalized Haar measure m defined on the σ-algebra $\mathcal{B}(X)$ of its Borel sets. An *automorphism* $\tau\colon X \to X$ is a 1-1, onto map such that $\tau(xy) = \tau(x)\tau(y)$ for all $x, y \in X$, i.e., a group isomorphism. It can be shown that if $\tau\colon X \to X$ is a continuous, surjective group homomorphism then τ is measure-preserving. It can also be shown, using the fact that the dual space $\widehat{X}$ of X forms an orthonormal basis for $L^2(X, m)$, that a continuous ergodic automorphism of X is strongly mixing. In fact, Lind, Miles and Thomas [8] were able to show that any continuous, ergodic, automorphism of a compact abelian group is isomorphic to a Bernoulli shift on some probability space. Recall that if $(X, \mathcal{F}, \mu)$ and $(Y, \sum, \nu)$ are probability spaces, and if $\tau\colon X \to X$, $\rho\colon Y \to Y$ are measure-preserving transformations, then τ and ρ are said to be *isomorphic* if $(X, \mathcal{F}, \mu, \tau)$ and (Y, Σ, ν, ρ) are isomorphic in the following sense: There exists $X_0 \subseteq X$, $Y_0 \subseteq Y$, $\mu(X_0) = 1 = \nu(Y_0)$ such that

(i) $\tau(X_0) \subseteq X_0$ $\rho(Y_0) \subseteq Y_0$ and
(ii) there exists an invertible, measure-preserving map $\phi\colon X_0 \to Y_0$ such that $\phi(\tau(x)) = \rho(\phi(x))$ for all x in X_0.

(By ϕ an invertible measure-preserving map we mean that both ϕ and ϕ^{-1} are measure-preserving, i.e., $\nu(\phi(A \cap X_0)) = \mu(A \cap X_0)$ and $\mu(\phi^{-1}(E \cap Y_0)) = \nu(E \cap Y_0)$, $A \in \mathcal{F}$, $E \in \sum$.)

It is intuitively clear that if τ and ρ are isomorphic, then $(\tau^n)_1^\infty$ is mixing of all orders if and only if $(\rho^n)_1^\infty$ is.

For an application of the above observation, consider the 2-torus $T^2 = [0, 1)^2$ with the usual Lebesgues measure λ_2 defined on its Borel sets. Consider the "baker's transformation" $b\colon [0, 1)^2 \to [0, 1)^2$ defined by

$$b(x, y) = \begin{cases} (2x, \frac{1}{2}y) & (\text{mod } 1) \quad \text{if } 0 \le x < \frac{1}{2} \\ (2x, \frac{1}{2}(y+1)) & (\text{mod } 1) \quad \text{if } \frac{1}{2} \le x < 1 \end{cases}$$

Then, $(T^2, \mathcal{B}(T^2), \lambda_2, \mathbf{b})$ is isomorphic to the Bernoulli scheme $\mathbf{B}(\frac{1}{2}, \frac{1}{2})$. Hence, $(b^n)_1^\infty$ is mixing of all orders.

Since a continuous ergodic automorphism of a compact abelian group is isomorphic to a Bernoulli shift, we conclude from (a) that *if τ is a continuous ergodic automorphism of a compact abelian group then $(\tau^n)_1^\infty$ is mixing of all orders.* For another example, consider $S^1 = \{z \in \mathbf{C}\colon |z| = 1\}$ and let l_1 be the usual normalized arc-length measure on S^1. Let

$$(S^n, \mathcal{B}(S^n), l_n) = \bigotimes_{k=1}^{n} (S^1, \mathcal{B}(S^1), l_1),$$

the product measure space. Then S^n is a compact abelian group under coordinate-wise multiplication and l_n is the normalized Haar measure on S^n. Let $A = [a_{ij}] \in M_{n\times n}(\mathbf{Z})$, with det $A = \pm 1$ and suppose that A has no eigenvalues which are roots of unity. Define

$$\tau\colon S^n \to S^n,$$
$$\tau(z_1, \dots, z_n) = (z_1^{a_{11}} \dots z_n^{a_{n1}}, \dots, z_1^{a_{n1}} \dots z_n^{a_{nn}}), \quad (z_1, \dots, z_n) \in S^n.$$

Then τ is a continuous ergodic automorphism [8] so that $(\tau^n)_1^\infty$ is mixing of all orders. So, if $\rho\colon T^2 \to T^2$ is given by, say,

$$\rho(x, y) = (5x + 7y, 3x + 4y) \text{ (mod 1)}, \quad (x, y) \in T^2,$$

then, $(\rho^n)_1^\infty$ is mixing of all orders.

(c) *Product of mixing sequences.*

Let $(X, \mathcal{F}, \mu)$ be a probability space. Assume that $(\tau_n)_1^\infty$ and $(s_n)_1^\infty$ are mixing of all orders on X. Define

$$\psi_n\colon (X \times X, \mathcal{F} \otimes \mathcal{F}, \mu \otimes \mu) \to (X \times X, \mathcal{F} \otimes \mathcal{F}, \mu \otimes \mu),$$
$$\psi_n(x, y) = (\tau_n x, s_n y), \quad \forall (x, y) \in X^2, n = 1, 2, \dots$$

Then $(\psi_n)_1^\infty$ is mixing of all orders. An application of this gives the following result.

Let n be a given positive integer. Let $T^n = [0, 1)^n$ be the n-torus, with the usual Lebesgue measure λ_n defined on its σ-algebra $\mathcal{B}(T^n)$ of Borel sets. Define a sequence (ψ_k) of transformations on T^n as follows:

$$\psi_k\colon T^n \to T^n,$$
$$\psi_k(x_1, \dots, x_n) = \big(kx_1 \text{ (mod 1)}, \dots, kx_n \text{ (mod 1)}\big), \quad k = 1, 2, \dots$$

Then, $(\psi_k)_1^\infty$ is mixing of all orders.

To see this, consider the 1-torus $(T^1, \mathcal{B}(T^1), \lambda_1)$. Define the multiplication operators $\mathcal{M}_j$ on T^1 as follows:

$$\mathcal{M}_j(x) = jx \text{ (mod 1)}, j = 1, 2, \dots, x \in T^1.$$

Since $(T^n, \bigotimes_{k=1}^n \mathcal{B}(T^1), \bigotimes_{k=1}^n \lambda_1) = (T^n, \mathcal{B}(T^n), \lambda_n)$, it suffices to show that $(\mathcal{M}_j)_1^\infty$ is mixing of all orders. Simply apply *the approximate independence lemma* [1] which states that if $f_1, \dots, f_K \in L^\infty(T^1)$ and $\epsilon > 0$ are given, there exists M_0 such that whenever $(n_1, \dots, n_K) \in \mathbf{N}^K$ satisfies $\frac{n_{j+1}}{n_j} \ge M_0$ for all $1 \le j \le K - 1$, then

$$\left| \int_{[0,1)} f_1(n_1 x) f_2(n_2 x) \cdots f_K(n_K x)\, d\lambda_1 - \prod_{i=1}^K \int_{[0,1)} f_i(x)\, d\lambda_1 \right| < \epsilon.$$

By putting $f_i = \chi_{A_i}$, $A_i \in \mathcal{B}(T^1)$, $1 \le i \le K$, we see immediately that $(\mathcal{M}_j)_1^\infty$ is mixing of all orders.

(d) If $\tau\colon S^1 \to S^1$ is an ergodic rotation of the unit circle, then τ is not weak mixing. Hence, $(\tau^n)_1^\infty$ cannot be mixing of all orders.

3.6. *The Gaussian distribution property.* For any complex-valued function f, we let $\operatorname{Re} f$ denote the real part of f.

Definition 3.7.1. Let $(X, \mathcal{F}, \mu)$ be a probability space. A finite collection $\Lambda = \{S_1, S_2, \ldots, S_N\}$ of linear transformations on $L^2(X)$ is said to satisfy the *Gaussian distribution property*, abbreviated as G.D.P., if it has the following properties:

(a) $L^2(X, 0)$ is invariant under Λ, i.e., $S\big(L^2(X, 0)\big) \subseteq L^2(X, 0)$ for all $S \in \Lambda$.
(b) There exists a constant $\kappa_\Lambda < \infty$ such that given any $f \in L^2(X, 0)$, we can find a sequence $(g_k)_1^\infty \subseteq L^2(X)$ with $\|g_k\|_2 \leq \kappa_\Lambda \|f\|_2$ for all $k = 1, 2, \ldots$ and such that

$$\mu \circ (\operatorname{Re} S_1 g_k, \ldots, \operatorname{Re} S_N g_k)^{-1} \rightharpoonup \operatorname{Gauss}(\operatorname{Re} S_1 f, \ldots, \operatorname{Re} S_N f) \quad \text{as } k \to \infty.$$

A collection $\mathcal{A}$ of linear transformations on $L^2(X)$ is said to satisfy the G.D.P. if every finite sub-collection Λ of $\mathcal{A}$ does and

$$\kappa := \sup_{\text{finite } \Lambda \subseteq \mathcal{A}} \kappa_\Lambda < \infty.$$

Finally, a collection $(\tau_\alpha)_{\alpha \in J}$ of measure-preserving transformations on X is said to satisfy the G.D.P. if the corresponding collection $(U_{\tau_\alpha})_{\alpha \in J}$ of induced operators on $L^2(X)$ satisfies the G.D.P. Here, each U_{τ_α} is defined by $U_{\tau_\alpha}(f) = f \circ \tau_\alpha$ for all $f \in L^2(X)$.

4. Main results

The following two theorems are the main results of the paper.

THEOREM I. *Let $(X, \mathcal{F}, \mu)$ be a probability space and let Λ be any finite collection of bounded, linear operators on $L^2(X, \mathcal{F}, \mu)$ such that $L^2(X, 0)$ is invariant under Λ. Assume that there exists a sequence $(\tau_n)_1^\infty$: $X \to X$ which is mixing of all orders and satisfies the following approximate commutativity condition*:

Given $f \in L^2(X, \mathcal{F}, \mu)$ and $\epsilon > 0$, there exist infinitely many integers $n \geq 1$ such that

$$\|S(f \circ \tau_n) - (Sf) \circ \tau_n\|_2 < \epsilon \quad \textit{for all } S \in \Lambda.$$

Then Λ satisfies the G.D.P., and moreover, we can take κ_Λ to be 3.

To state our next theorem, we first need to define the following concept of subsets of **N** having certain densities. For each finite set A, let Card (A) be the number of elements in A.

A set $S \subseteq \mathbf{N}$ is said to have *density* α if

$$\lim_{n\to\infty} \frac{\operatorname{Card}\Big(S \cap \{1,2,\dots,n\}\Big)}{n} = \alpha.$$

$S \subseteq \mathbf{N}$ is said to have *positive upper density* if

$$\limsup_{N\to\infty} \frac{\operatorname{Card}\Big(S \cap \{1,2,\dots,N\}\Big)}{N} > 0.$$

Clearly, S has positive density implies S has positive upper density.

THEOREM II. *Let $(X, \mathcal{F}, \mu)$ be a probability space. Let Λ be a finite collection of bounded, linear operators on $L^2(X)$ such that $L^2(X, 0)$ is invariant under Λ. Assume there exists a weakly-mixing transformation $\tau\colon X \to X$ satisfying the following condition:*

For any $f \in L^\infty(X)$ and for any $K, \epsilon > 0$, K an integer, there exists positive integers $j_1 < j_2 < \cdots < j_K$ such that the set

$$\mathcal{D} := \{n \in \mathbf{N}\colon \|S\,(f \circ \tau^{nj_k}) - (Sf) \circ \tau^{nj_k}\|_2 < \epsilon \quad \forall S \in \Lambda, \quad \forall 1 \le k \le K\}$$

has positive upper density.

Then Λ satisfies the G.D.P. with $\kappa_\Lambda = 3$.

Consequently, Λ satisfies the G.D.P. if there exists a weakly mixing $\tau\colon X \to X$ satisfying

$$SU_\tau = U_\tau S \quad \textit{for all } S \in \Lambda.$$

5. Some lemmas

This part of the paper is devoted to statements and proofs of all necessary lemmas that will be needed in the proof of Theorem I.

LEMMA 5.1. *Let $N \geq 1$ be any given integer. Let $a_1, \dots, a_N, b_1, \dots, b_N$ be real numbers with $P = \max\{|a_i|, |b_i|\colon 1 \le i \le N\}$. Then*

$$\left|\prod_{i=1}^N a_i - \prod_{i=1}^N b_i\right| \le P^{N-1} \sum_{i=1}^N |a_i - b_i|.$$

Proof. We have

$$\prod_{i=1}^{N} a_i - \prod_{i=1}^{N} b_i = a_1a_2\cdots a_N - b_1a_2\cdots a_N$$
$$+ b_1a_2\cdots a_N - b_1b_2a_3\cdots a_N$$
$$+ b_1b_2a_3\cdots a_N - b_1b_2b_3a_4\cdots a_N$$
$$+$$
$$\vdots$$
$$+ b_1b_2\cdots b_{N-1}a_N - b_1b_2\cdots b_N.$$

Thus, letting the empty product be 1, we obtain

$$\left|\prod_{i=1}^{N} a_i - \prod_{i=1}^{N} b_i\right| \le \sum_{j=1}^{N}\left(|a_j - b_j|\prod_{i=1}^{j-1}|b_i|\prod_{i=j+1}^{N}|a_i|\right)$$
$$\le P^{N-1}\sum_{j=1}^{N}|a_j - b_j|. \qquad \square$$

LEMMA 5.2. *Let $(X, \mathcal{F}, \mu)$ be a probability space. Let $(\tau_n)_1^\infty\colon X \to X$ be mixing of all orders. Let K be any positive integer and assume $h_j\colon X \to \mathbf{C}$ is in $L^\infty(X)$ for $j = 1, 2, \ldots, K$. Then, for any $\epsilon > 0$, there exists an integer $M \ge 1$ with the following property:*

Whenever $(n_1, \ldots, n_K) \in \mathbf{N}^K$ satisfies

(i) $\dfrac{n_{j+1}}{n_j} \ge M,\ 1 \le j \le K-1$ *and*

(ii) $n_1, \ldots, n_K \ge M$,

then

$$\left|\int_X \prod_{j=1}^{K}(h_j \circ \tau_{n_j})\, d\mu - \prod_{j=1}^{K}\int_X h_j\, d\mu\right| < \epsilon.$$

Proof. This follows from Lemma 1 and a routine approximation argument. $\square$

LEMMA 5.3. *Let $A > 0$. Let $K, N \ge 1$ be integers. Consider the compact set $C = [-A, A]^N \times \cdots \times [-A, A]^N$ in $\mathbf{R}^{KN}$. Let $\mathcal{B}$ be the collection of all real-valued functions on C of the form*

$$\varphi\colon C \to \mathbf{R}$$
$$(x^1, \ldots, x^K) \mapsto \varphi_1(x^1) \times \cdots \times \varphi_K(x^K),$$

where $x^1, \ldots, x^K$ *are points of* $[-A, A]^N$ *and* φ_i *is bounded and continuous on* $[-A, A]^N$ *for each* $1 \leq i \leq K$.

Then lin $\mathcal{B}$ *is dense in the normed space of all continuous real-valued functions on* C *with the sup-norm* $\|\cdot\|_\infty$

Proof. This follows directly from the Stone-Weierstrass theorem. □

We now combine Lemma 2 and Lemma 3 to obtain the following proposition which plays a key role in the proof of Theorem I.

PROPOSITION 5.1. *Let* N, K *be given positive integers. Let* $(X, \mathcal{F}, \mu)$ *be a probability space and suppose* $(\tau_n)_1^\infty\colon X \to X$ *is mixing of all orders. Let* $h_1, \ldots, h_K\colon X \to \mathbf{R}^N$ *be measurable such that* $\pi_i \circ h_j \in L^2(X, \mu)$ *for all* $1 \leq i \leq N$, $1 \leq j \leq K$. *For each* $j = 1, 2, \ldots, K$, *let* $\nu_j = \mu \circ h_j^{-1}$. *Consider the product measure space*

$$\left(\mathbf{R}^{KN}, \mathcal{B}(\mathbf{R}^N) \otimes \cdots \otimes \mathcal{B}(\mathbf{R}^N), \nu\right) = \bigotimes_{j=1}^{K} \left(\mathbf{R}^N, \mathcal{B}(\mathbf{R}^N), \nu_j\right).$$

For $n = (n_1, \ldots, n_K) \in \mathbf{N}^K$, *let* $H_n = (h_1 \circ \tau_{n_1}, \ldots, h_K \circ \tau_{n_K})\colon X \to \mathbf{R}^{KN}$, *and* $\mu_n = \mu \circ H_n^{-1}$. *Let* $\psi\colon \mathbf{R}^{KN} \to \mathbf{R}$ *be bounded and continuous and let* $\eta > 0$ *be given.*

Then there exists an integer $M_0 \geq 1$ *with the following property: If* $n = (n_1, \ldots, n_K) \in \mathbf{N}^K$ *satisfies*

(i) $\dfrac{n_{j+1}}{n_j} \geq M_0$, $j = 1, 2, \ldots, K-1$ *and*

(ii) $n_1, \ldots, n_k \geq M_0$,

then

$$\left| \int_{\mathbf{R}^{KN}} \psi \, d\mu_n - \int_{\mathbf{R}^{KN}} \psi \, d\nu \right| < \eta.$$

Proof. Let $\psi\colon \mathbf{R}^{KN} \to \mathbf{R}$ be bounded and continuous and let $\eta > 0$. We may assume that $\|\psi\|_\infty \leq 1$. Choose $\epsilon > 0$ so that $KN\epsilon + 1 - (1 - N\epsilon)^K < \eta/4$. Since each $\pi_i \circ h_j \in L^2(X)$, we can choose $A > 0$ such that the set $X_{ij} = \{x \in X\colon |\pi_i \circ h_j(x)| > A\}$ has measure $\mu(X_{ij}) < \epsilon$, for all $1 \leq i \leq N$, $1 \leq j \leq K$. Hence, for each j,

$$\{x \in X\colon h_j(x) \notin [-A, A]^N\} \subseteq \bigcup_{i=1}^{N} X_{ij}.$$

Thus, if $\nu_j = \mu \circ h_j^{-1}$, then $\nu_j\left([-A, A]^N\right) \geq (1 - N\epsilon)$ for all j. Let $\nu = \bigotimes_{j=1}^K \nu_j$ and let

$$C \equiv \underbrace{[-A, A]^N \times \cdots \times [-A, A]^N}_{K \text{ times}}.$$

Then, $\nu(C) \geq (1 - N\epsilon)^K$ so that

$$\nu(\mathbf{R}^{KN} - C) \leq 1 - (1 - N\epsilon)^K.$$

Similarly, if $n = (n_1, \dots, n_K) \in \mathbf{N}^K$ and $\mu_n = \mu \circ (h_1 \circ \tau_{n_1}, \dots, h_K \circ \tau_{n_K})^{-1}$, then $\mu_n(C) \geq 1 - KN\epsilon$, so that

$$\mu_n(\mathbf{R}^{KN} - C) \leq KN\epsilon.$$

Let $\mathcal{B}$ be as in Lemma 3. Then there exists $\varphi_0 \in \operatorname{lin} \mathcal{B}$ such that

$$\sup_{x \in C} |\varphi_0(x) - \psi(x)| < \eta/4.$$

Also, for any $n = (n_1, \dots, n_K) \in \mathbf{N}^K$, let $I = \left|\int_{\mathbf{R}^{KN}} \psi\, d\mu_n - \int_{\mathbf{R}^{KN}} \psi d\nu\right|$. Then

$$\begin{aligned} I &= \left|\int_{\mathbf{R}^{KN}-C} \psi\, d\mu_n - \int_{\mathbf{R}^{KN}-C} \psi\, d\nu\right| + \left|\int_C \psi d\mu_n - \int_C \psi\, d\nu\right| \\ &\leq \|\psi\|_\infty \left(KN\epsilon + 1 - (1 - N\epsilon)^K\right) + \left|\int_C \psi\, d\mu_n - \int_C \psi\, d\nu\right|. \end{aligned}$$

Now, if $J = \left|\int_C \psi\, d\mu_n - \int_C \psi\, d\nu\right|$, then

$$\begin{aligned} J &= \left|\int_C (\psi - \varphi_0) d\mu_n - \int_C (\psi - \varphi_0) d\nu\right| + \left|\int_C \varphi_0 d\mu_n - \int_C \varphi_0 d\nu\right| \\ &\leq 2 \sup_{x \in C} |\psi(x) - \varphi_0(x)| + \left|\int_C \varphi_0 d\mu_n - \int_C \varphi_0\, d\nu\right| \\ &\leq \eta/2 + \left|\int_C \varphi_0\, d\mu_n - \int_C \varphi_0\, d\nu\right|. \end{aligned}$$

Therefore, we have

$$\begin{aligned} I &= \left|\int_{\mathbf{R}^{KN}} \psi\, d\mu_n - \int_{\mathbf{R}^{KN}} \psi d\, \nu\right| \\ &\leq KN\epsilon + 1 - (1 - N\epsilon)^K + \eta/2 + \left|\int_C \varphi_0\, d\mu_n - \int_C \varphi_0\, d\, \nu\right| \\ &\leq 3\eta/4 + \left|\int_C \varphi_0\, d\mu_n - \int_C \varphi_0\, d\nu\right|. \end{aligned}$$

Hence, to complete the proof of our proposition, it suffices to show that for any $\varphi \in \mathcal{B}$, there exists an integer $M \geq 1$ such that whenever $n = (n_1, \ldots, n_K) \in \mathbf{N}^K$ satisfies $n_j \geq M$ for each $j = 1, 2 \ldots, K$ and $\frac{n_{j+1}}{n_j} \geq M$, we have

$$\left| \int_C \varphi \, d\mu_n - \int_C \varphi \, d\nu \right| < \eta/4.$$

So, let $\varphi \in \mathcal{B}$, say $\varphi(x^1, \ldots, x^K) = \varphi_1(x^1) \cdots \varphi_K(x^K)$, where each φ_i is a bounded continuous real-valued function on $[-A, A]^N$. Then

$$\begin{aligned} \int_C \varphi \, d\nu &= \int_C (\varphi_1 \times \cdots \times \varphi_K) d(\nu_1 \otimes \cdots \otimes \nu_K) \\ &= \prod_{j=1}^{K} \int_{[-A,A]^N} \varphi_j \, d\nu_j = \prod_{j=1}^{K} \int_{\mathbf{R}^N} (\chi_{[-A,A]^N} \times \varphi_j) \, d\nu_j \\ &= \prod_{j=1}^{K} \int_X \left[(\chi_{[-A,A]^N} \circ h_j) \times (\varphi_j \circ h_j) \right] d\mu \end{aligned} \tag{1}$$

Similarly, if $n = (n_1, \ldots, n_K) \in \mathbf{N}^K$, then

$$\begin{aligned} \int_C \varphi \, d\mu_n &= \int_{\mathbf{R}^{KN}} (\chi_C \times \varphi) d\mu_n \\ &= \int_X \chi_C \circ (h_1 \circ \tau_{n_1}, \ldots, h_K \circ \tau_{n_K}) \\ &\quad \times \varphi_1 \circ (h_1 \circ \tau_{n_1}) \times \ldots \times \varphi_K \circ (h_K \circ \tau_{n_K}) \, d\mu. \end{aligned}$$

Since

$$\chi_C \circ (h_1 \circ \tau_{n_1}, \ldots, h_K \circ \tau_{n_K}) = \prod_{j=1}^{K} (\chi_{[-A,A]^N} \circ h_j) \circ \tau_{n_j},$$

we have

$$\int_C \varphi \, d\mu_n = \int_X \left(\prod_{j=1}^{K} \left[\chi_{[-A,A]^N} \circ h_j) \times (\varphi_j \circ h_j) \right] \circ \tau_{n_j} \right) d\mu \tag{2}$$

For each $j = 1, 2, \ldots, K$, let $g_j = (\chi_{[-A,A]^N} \circ h_j) \times (\varphi_j \circ h_j)$. Then from (1) and (2),

$$\left| \int_C \varphi \, d\mu_n - \int_C \varphi \, d\nu \right| = \left| \int_X \prod_{j=1}^{K} (g_j \circ \tau_{n_j}) \, d\mu - \prod_{j=1}^{K} \int_X g_j \, d\mu \right| \tag{3}$$

Since φ_j is bounded for each j, we have $g_j \in L^\infty(X)$, $j = 1, 2, \ldots, K$. Now apply Lemma 2 to (3) to conclude that there exists an integer $M \geq 1$ with the required property. □

COROLLARY 5.1. *Let $N, K \in \mathbf{N}$. Suppose $(X, \mathcal{F}, \mu)$ is a probability space and $(\tau_n)_1^\infty$: $X \to X$ is mixing of all orders.*

Let $f \in L^2(X, 0)$, $f_1, \ldots, f_N \in L^2(X, 0)$ be real-valued functions. Let $\Omega = (f_1, \ldots, f_N)$: $X \to \mathbf{R}^N$. Put $\lambda = \mu \circ \Omega^{-1}$, and for $n = (n_1, \ldots, n_K) \in \mathbf{N}^K$, let

$$\mu_n = \mu \circ (\Omega \circ \tau_{n_1}, \ldots, \Omega \circ \tau_{n_K})^{-1}.$$

Let ψ: $\mathbf{R}^{KN} \to \mathbf{R}$ be bounded and continuous and let $\eta > 0$ be given. Then there exists an integer $M_0 \geq 1$ with the following property:

Whenever $n = (n_1, \ldots, n_K) \in \mathbf{N}^K$ satisfies

(i) $\frac{n_{j+1}}{n_j} \geq M_0$ for each $j = 1, 2, \ldots, K-1$ and

(ii) $n_1, \ldots, n_K \geq M_0$,

we have

(a) $\left| \int_{\mathbf{R}^{KN}} \psi \, d\mu_n - \int_{\mathbf{R}^{KN}} \psi \, d\lambda^K \right| < \eta$ *and*

(b) $\left| \int_X (f \circ \tau_{n_i})(f \circ \tau_{n_j}) \, d\mu \right| < \eta \quad \forall 1 \leq i \neq j \leq K.$

LEMMA 5.4. *There exists a countable set $\mathcal{H}$ of continuous real-valued functions on $\mathbf{R}^n$ with compact supports satisfying the following condition:*

Let ϕ: $\mathbf{R}^N \to \mathbf{R}$ be continuous with $\sup_{x \in \mathbf{R}^n} |\phi(x)| \equiv \|\phi\|_\infty \leq 1$. Let $K \subseteq \mathbf{R}^N$ be compact and $\epsilon > 0$. Then there is some φ in $\mathcal{H}$ such that

$$\sup_{x \in K} |\varphi(x) - \phi(x)| < \epsilon.$$

Proof. Write $\mathbf{R}^N = \bigcup_{i=1}^\infty K_n$, where $K_1 \subseteq K_2 \subseteq K_3 \subseteq \ldots$ are compact subsets of $\mathbf{R}^N$. Let

$$\mathcal{C}(K_n) = \{f\colon K_n \to \mathbf{R}\colon f \text{ continuous}\}.$$

Since K_n is compact, $\mathcal{C}(K_n)$ has a countable dense subset $\mathcal{G}_n$. Here, the norm of $f \in \mathcal{C}(K_n)$ is $\|f\|_{K_n} \equiv \sup_{x \in K_n} |f(x)|$. For each $\alpha \in \mathcal{G}_n$, let $\widetilde{\alpha}$: $\mathbf{R}^N \to \mathbf{R}$ be a Tietze's extension of α to all of $\mathbf{R}^N$, i.e., $\widetilde{\alpha}$ is continuous of compact support, $\widetilde{\alpha}(x) = \alpha(x)$ for all $x \in K_n$ and $\|\widetilde{\alpha}\|_\infty = \|\alpha\|_{K_n}$. For $n = 1, 2, \ldots$, let

$$\mathcal{H}_n = \{\widetilde{\alpha}\colon \mathbf{R}^N \to \mathbf{R}\colon \alpha \in \mathcal{G}_n\}.$$

Put $\mathcal{H} = \bigcup_{n=1}^\infty \mathcal{H}_n$.

Now, suppose that ϕ: $\mathbf{R}^N \to \mathbf{R}$ is continuous with $\|\phi\|_\infty \leq 1$. Let $\epsilon > 0$ and let $K \subseteq \mathbf{R}^N$ be compact. Then $K \subseteq K_n$ for some n. Hence, there exists $\alpha \in \mathcal{G}_n$ such that

$$\sup_{x\in K_n} \left|\phi(x) - \alpha(x)\right| \leq \epsilon.$$

Let $\widetilde{\alpha} \in \mathcal{H}_n$ be a corresponding Tietze's extension of α to all of $\mathbf{R}^N$. Then

$$\sup_{x\in K} \left|\widetilde{\alpha}(x) - \phi(x)\right| \leq \sup_{x\in K_n} \left|\widetilde{\alpha}(x) - \phi(x)\right| = \sup_{x\in K_n} \left|\alpha(x) - \phi(x)\right| \leq \epsilon. \qquad \square$$

The two lemmas below are very elementary. We omit the proofs.

LEMMA 5.5. *Let $(X, \mathcal{F}, \mu)$ be a probability space. Let N be a positive integer. Let $\mathcal{S} \subseteq L^2(X, \mathcal{F}, \mu)$ be bounded, so that there exists $C < \infty$ with $\|f\|_2 \leq C$ for all $f \in \mathcal{S}$. Then for each $\epsilon > 0$, there exists a compact set $K \subseteq \mathbf{R}^N$ such that, for all $f_1, \ldots, f_N \in \mathcal{S}$,*

$$\mu \circ (f_1, \ldots, f_N)^{-1}(\mathbf{R}^N - K) < \epsilon.$$

LEMMA 5.6. *Let $\xi\colon \mathbf{R}^N \to \mathbf{R}$ be any continuous function of compact support. Let $\eta > 0$. Then there is some $\epsilon > 0$ with the following property:*

Let $(X, \mathcal{F}, \mu)$ be any probaility space. Let

$$\alpha = (\alpha_1, \ldots, \alpha_N)\colon X \to \mathbf{R}^N,\ \beta = (\beta_1, \ldots, \beta_N)\colon X \to \mathbf{R}^N,$$

where α_i, β_i are real-valued measurable functions on X. Assume that

$$\|\alpha_i - \beta_i\|_2 < \epsilon \quad \forall 1 \leq i \leq N.$$

Then

$$\left|\int_{\mathbf{R}^N} \xi \, d(\mu \circ \alpha^{-1}) - \int_{\mathbf{R}^N} \xi \, d(\mu \circ \eta^{-1})\right| < \eta.$$

6. Proof of Theorem I

Proof. Let $(X, \mathcal{F}, \mu)$ be a probability space and let $\Lambda = \{S_1, \ldots, S_N\}$, $(\tau_n)^\infty$ be as in the hypotheses of Theorem I. Assume that $f \in L^2(X, 0)$. Then $\operatorname{Re} S_i f \in L^2_{\mathbf{R}}(X, 0)$ for all $1 \leq i \leq N$.

Let $\{\xi_j\}_1^J$ be a finite collection of compactly-supported, continuous functions from $\mathbf{R}^N$ to $\mathbf{R}$. Let $\gamma = \operatorname{Gauss}(\operatorname{Re} S_1 f, \ldots, \operatorname{Re} S_N f)$. We will show the following:

For any $\eta > 0$, there exists some $g \in L^2(X)$, $\|g\|_2 \leq 2\|f\|_2$ such that if

$$\nu = \mu \circ (\operatorname{Re} S_1 g, \ldots, \operatorname{Re} S_N g)^{-1},$$

then

$$\left|\int_{\mathbf{R}^N} \xi_j \, d\nu - \int_{\mathbf{R}^N} \xi_j \, d\gamma\right| < \eta \quad \forall j = 1, 2, \ldots, J.$$

1. Put $f_i = \operatorname{Re} S_i f$, $1 \le i \le N$. Let $\Omega := (f_1, \dots, f_N)\colon X \to \mathbf{R}^N$, and $\lambda = \mu \circ \Omega^{-1}$. If $\pi_i\colon \mathbf{R}^N \to \mathbf{R}$ is the i^{th}-projection, then for $1 \le i, j \le N$,

$$\int_{\mathbf{R}^N} \pi_i \, d\lambda = \int_X f_i \, d\mu = 0,$$
$$\int_{\mathbf{R}^N} \pi_i \pi_j \, d\lambda = \int_X f_i f_j \, d\mu = \int_X \pi_i \pi_j \, d\gamma < \infty.$$

Thus, by the Central Limit Theorem in Section 3, we can choose an integer $K \ge N$ large enough so that for all $1 \le j \le J$,

$$\left| \int_{\mathbf{R}^{KN}} (\xi_j \circ T) \, d\lambda^K - \int_{\mathbf{R}^N} \xi_j \, d\gamma \right| < \eta, \tag{4}$$

where $T\colon \mathbf{R}^{KN} \to \mathbf{R}$ is defined by

$$T(x^1, \dots, x^K) = \frac{x^1 + \dots + x^K}{\sqrt{K}} \quad \forall x^1, \dots, x^K \in \mathbf{R}^N.$$

2. For $n = (n_1, \dots, n_K) \in \mathbf{N}^K$, let $H_n = (\Omega \circ \tau_{n_1}, \dots, \Omega \circ \tau_{n_K})$ and $\mu_n = \mu \circ H_n^{-1}$. Choose $\epsilon_0 > 0$ small enough so that Lemma 6 holds for each ξ_j in place of ξ, $1 \le j \le J$.

By our assumption on $(\tau_n)_1^\infty$, there exist infinitely many integers $m \ge 1$ such that for all $1 \le k \le N$,

$$\|S_k(f \circ \tau_m) - (S_k f) \circ \tau_m\|_2 < \frac{\epsilon_0}{K} \tag{5}$$

By Corollary 1, we can choose a $p = (p_1, p_2, \dots, p_K) \in \mathbf{N}^K$ satisfying

$$\|S_k(f \circ \tau_{p_i}) - (S_k f) \circ \tau_{p_i}\|_2 < \frac{\epsilon_0}{K} \quad \forall 1 \le k \le N, \quad \forall 1 \le i \le K \tag{6}$$

with

$$\left| \int_{\mathbf{R}^{KN}} (\xi_j \circ T) \, d\mu_p - \int_{\mathbf{R}^{KN}} (\xi_j \circ T) \, d\lambda^K \right| < \eta \quad \forall 1 \le j \le J \tag{7}$$

and so that

$$\left| \int_X (f \circ \tau_{p_i})(f \circ \tau_{p_j}) \, d\mu \right| \le \frac{\|f\|_2^2}{K} \quad \forall 1 \le i \ne j \le K. \tag{8}$$

From (4) and (7),

$$\left| \int_{\mathbf{R}^{KN}} (\xi_j \circ T) \, d\mu_p - \int_{\mathbf{R}^N} \xi_j \, d\gamma \right| < 2\eta \quad \forall 1 \le j \le J. \tag{9}$$

3. For each $i = 1, 2, \dots, N$, let $g_i\colon X \to \mathbf{R}$ be defined by

$$g_i = \frac{f_i \circ \tau_{p_1} + f_i \circ \tau_{p_2} + \cdots + f_i \circ \tau_{p_K}}{\sqrt{K}}.$$

Let $\widetilde{\nu} = \mu \circ (g_1, \dots, g_N)^{-1}$. Then

$$\int_{\mathbf{R}^{KN}} \xi_j \, d\tilde{\nu} = \int_{\mathbf{R}^{KN}} (\xi_j \circ T)\, d\mu_P, \quad j = 1, 2, \dots, J \tag{10}$$

4. Let $g = \frac{1}{\sqrt{K}}\left[f \circ \tau_{p_1} + \cdots + f \circ \tau_{p_K}\right]$. For all $1 \le i \le N$, we have (by (6))

$$\begin{aligned} \|\mathrm{Re}\, S_i g - g_i\|_2 &= \frac{1}{\sqrt{K}} \left\| \sum_{j=1}^{K} \mathrm{Re}\, S_i(f \circ \tau_{p_j}) - (\mathrm{Re}\, S_i f) \circ \tau_{p_j} \right\|_2 \\ &\le \frac{1}{\sqrt{K}} \sum_{j=1}^{K} \|\mathrm{Re}\left(S_i(f \circ \tau_{p_j}) - (S_i f) \circ \tau_{p_j}\right)\|_2 < \epsilon_0 \end{aligned}$$

Consequently, if $\alpha = (\mathrm{Re}\, S_1 g, \dots, \mathrm{Re}\, S_N g)$ and $\beta = (g_1, \dots, g_N)$, then by Lemma 6,

$$\left| \int_{\mathbf{R}^N} \xi_j \, d\nu - \int_{\mathbf{R}^N} \xi_j \, d\widetilde{\nu} \right| < \eta \quad \forall 1 \le j \le J, \tag{11}$$

where $\nu = \mu \circ \alpha^{-1}$, $\tilde{\nu} = \mu \circ \beta^{-1}$.

5. Therefore, by using (9), (10) and (11), for each $j = 1, 2, \dots, J$ we obtain

$$\begin{aligned} \left| \int_{\mathbf{R}^N} \xi_j \, d\nu - \int_{\mathbf{R}^N} \xi_j \, d\gamma \right| &\le \left| \int_{\mathbf{R}^N} \xi_j \, d\nu - \int_{\mathbf{R}^N} \xi_j \, d\widetilde{\nu} \right| + \left| \int_{\mathbf{R}^N} \xi_j \, d\tilde{\nu} - \int_{\mathbf{R}^N} \xi_j \, d\gamma \right| \\ &= \left| \int_{\mathbf{R}^N} \xi_j \, d\nu - \int_{\mathbf{R}^N} \xi_j \, d\widetilde{\nu} \right| + \left| \int_{\mathbf{R}^{KN}} (\xi_j \circ T)\, d\mu_p - \int_{\mathbf{R}^N} \xi_j \, d\gamma \right| \\ &\le 3\eta. \end{aligned}$$

Finally, since $g = \dfrac{1}{\sqrt{K}}\left[f \circ \tau_{p_1} + \cdots + f \circ \tau_{p_K}\right]$, we have

$$\begin{aligned} \|g\|_2^2 &= \frac{1}{K} \sum_{1 \le i, j \le K} \int_X (f \circ \tau_{p_i})(f \circ \tau_{p_j})\, d\mu \\ &= \frac{1}{K} \left[K\|f\|_2^2 + \sum_{1 \le i \ne j \le K} \int_X (f \circ \tau_{p_i})(f \circ \tau_{p_j})\, d\mu \right]. \end{aligned}$$

Thus, by (8),

$$\begin{aligned} \|g\|_2^2 &\le \|f\|_2^2 + \frac{1}{K} \sum_{1 \le i \ne j \le K} \left| \int_X (f \circ \tau_{p_i})(f \circ \tau_{p_j})\, d\mu \right| \\ &\le \|f\|_2^2 + \frac{1}{K} K(K-1) \frac{\|f\|_2^2}{K} \\ &\le 2\|f\|_2^2. \end{aligned}$$

We have completely proved what we set out to prove.

6. Let $\mathcal{H}$ be as in Lemma 5.4, say $\mathcal{H} = \{\varphi_1, \varphi_2, \varphi_3, \dots\}$. It follows from step 5 above that for any positive integer l, there exists $g_l \in L^2(X)$, $\|g_l\|_2^2 \le 2\|f\|_2$, such that if $\nu_j = \mu \circ (\operatorname{Re} S_1 g_j, \dots, \operatorname{Re} S_N g_j)^{-1}$ for each $j \ge 1$, then

$$\left|\int_{\mathbf{R}^N} \varphi_i \, d\nu_l - \int_{\mathbf{R}^N} \varphi_i \, d\gamma\right| < \frac{1}{l} \quad \forall i = 1, 2, \dots, l.$$

Hence, it follows that there exists $(g_j)_1^\infty \subseteq L^2(X)$, $\|g_j\|_2 \le 2\|f\|_2$ for all $j = 1, 2, \dots$ such that

$$\left|\int_{\mathbf{R}^N} \varphi_i \, d\nu_j - \int_{\mathbf{R}^N} \varphi_i \, d\gamma\right| < \frac{1}{j} \quad \forall i = 1, 2, \dots, j.$$

Therefore,

$$\lim_{j\to\infty} \int_{\mathbf{R}^N} \varphi \, d\nu_j = \int_{\mathbf{R}^N} \varphi \, d\gamma \quad \forall \varphi \in \mathcal{H}. \tag{12}$$

7. Let $\xi \colon \mathbf{R}^N \to \mathbf{R}$ be continuous, $\|\xi\|_\infty \le 1$. We will show that

$$\lim_{j\to\infty} \int_{\mathbf{R}^N} \xi \, d\nu_j = \int_{\mathbf{R}^N} \xi \, d\gamma.$$

Fix some small $\delta > 0$. By Lemma 5.5, there exists a compact set $K \subset \mathbf{R}^N$ such that

$$\nu_j(\mathbf{R}^N - K) < \delta \quad \text{for all } j \ge 1 \quad \text{and} \quad \gamma(\mathbf{R}^N - K) < \delta.$$

8. Choose $\varphi \in \mathcal{H}$ so that $\sup_{x\in K} |\varphi(x) - \xi(x)| < \delta$. Then, for all $j \ge 1$,

$$\begin{aligned}
&\left|\int_{\mathbf{R}^N} \varphi \, d\nu_j - \int_{\mathbf{R}^N} \xi \, d\nu_j\right| \\
&\le \left|\int_{\mathbf{R}^N - K} \varphi \, d\nu_j\right| + \left|\int_{\mathbf{R}^N - K} \xi \, d\nu_j\right| + \left|\int_K \varphi \, d\nu_j - \int_K \xi \, d\nu_j\right| \\
&< 2\nu_j(\mathbf{R}^N - K) + \nu_j(\mathbf{R}^N - K) + \delta\, \nu_j(\mathbf{R}^N) < 4\delta.
\end{aligned}$$

Similarly, using the last inequality in step 7, we obtain

$$\left|\int_{\mathbf{R}^N} \varphi \, d\gamma - \int_{\mathbf{R}^N} \xi \, d\gamma\right| < 4\delta.$$

Hence,

$$\begin{aligned}
\left|\int_{\mathbf{R}^N} \xi \, d\nu_j - \int_{\mathbf{R}^N} \xi \, d\gamma\right| \le{}& \left|\int_{\mathbf{R}^N} \xi \, d\nu_j - \int_{\mathbf{R}^N} \varphi \, d\nu_j\right| + \left|\int_{\mathbf{R}^N} \varphi \, d\nu_j - \int_{\mathbf{R}^N} \varphi \, d\gamma\right| \\
&+ \left|\int_{\mathbf{R}^N} \varphi \, d\gamma - \int_{\mathbf{R}^N} \xi \, d\gamma\right|.
\end{aligned}$$

Thus,

$$\left|\int_{\mathbf{R}^N} \xi\, d\nu_j - \int_{\mathbf{R}^N} \xi\, d\gamma\right| \le 8\delta + \left|\int_{\mathbf{R}^N} \varphi\, d\nu_j - \int_{\mathbf{R}^N} \varphi\, d\gamma\right|.$$

By (12),

$$\lim_{j\to\infty} \int_{\mathbf{R}^N} \xi\, d\nu_j = \int_{\mathbf{R}^N} \xi\, d\gamma. \tag{13}$$

9. If $\alpha\colon \mathbf{R}^N \to \mathbf{R}$ is continuous and bounded, let $\xi = \dfrac{\alpha}{\|\alpha\|_\infty + 1}$, so that $\|\xi\|_\infty \le 1$, and then apply (13) to get

$$\lim_{j\to\infty} \int_{\mathbf{R}^N} \alpha\, d\nu_j = \int_{\mathbf{R}^N} \alpha\, d\gamma. \qquad \square$$

Remarks. By making use of Furstenberg's theorem stated below [5] and adapting the techniques employed in the proof of Theorem I, Theorem II can be similarly proved, and therefore we omit the proof.

FURSTENBERG'S RECURRENCE THEOREM FOR WEAKLY-MIXING TRANSFORMATIONS. *Let $(X, \mathcal{F}, \mu)$ be a probability space and let $\sigma\colon X \to X$ be weakly-mixing. Let $K \in \mathbf{N}$. Then, for every $j_1 < j_2 < \dots < j_K$ in $\mathbf{N}$, every $f_1, \dots, f_K \in L^\infty(X)$ and every $\eta > 0$, there exists a set $\mathcal{D} = \mathcal{D}(\eta, j_1, \dots, j_K, f_1, \dots, f_K) \subset \mathbf{N}$ of density 1, such that, for all $n \in \mathcal{D}$, we have*

$$\left|\int_X \prod_{i=1}^{K} (f_i \circ \sigma^{j_i n})\, d\mu - \prod_{i=1}^{K} \int_X f_i\, d\mu\right| < \eta.$$

REFERENCES

1. M. A. Akcoglu, M. D. Ha and R. L. Jones, "Divergence of ergodic averages" in *Topological Vector Spaces, Algebras and related areas*, Pitman Research Notes in Mathematics, vol. 316, Longman, 1994, pp. 175–192
2. ———, *Sweeping-out properties of operator sequences*, Canad. J. Math. **49** (1997), 3–23.
3. J. Bourgain, *Almost sure convergence and bounded entropy*, Israel J. Math. **63** (1988), 79–97.
4. L. Breiman, *Probability*, SIAM, 1993.
5. H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, 1981.
6. M. D. Ha, *Operators with Gaussian distributions, L^2-entropy and almost everywhere convergence*, Ph. D Thesis, The University of Toronto, 1994.
7. J. M. Marstrand, *On Khitnchine's conjecture about strong uniform distribution*, Proc. London Math. Soc. **21** (1970), 540–556.
8. K. Petersen, *Ergodic theory*, Cambridge University Press, 1989.
9. W. Rudin, *An arithmetic property of Riemann sums*, Proc. Amer. Math. Soc. **15** (1964), 321–324.

Department of Mathematics, Middle East Technical University, Ankara, Turkey
hadzung@math.toronto.edu.ca