Averages along cubes for not necessarily commuting m.p.t.

I. Assani

ABSTRACT. We study the pointwise convergence of some weighted averages linked to averages along cubes. We show that if $(X, \mathcal{B}, \mu, T_i)$ are not necessarily commuting measure preserving systems on the same finite measure space and if f_i, $1 \le i \le 6$ are bounded functions then the averages

$$\frac{1}{N^3} \sum_{n,m,p=1}^{N} f_1(T_1^n x) f_2(T_2^m x) f_3(T_3^p x) f_4(T_4^{n+m} x) f_5(T_5^{n+p} x) f_6(T_6^{m+p} x)$$

converge almost everywhere.

1. Introduction

Let $(X, \mathcal{B}, \mu, T_i)$, $1 \le i \le 3$, be three measure preserving systems on the same finite measure space. In [**1**] we proved that if f_i, $1 \le i \le 3$ are three bounded functions then the averages

$$\frac{1}{N^2} \sum_{n=1}^{N} f_1(T_1^n x) f_2(T_2^m x) f_3(T_3^{n+m} x)$$

converge almost everywhere. This is a bit surprising as it is known [**3**] that the averages along diagonal terms such as $\frac{1}{N} \sum_{n=1}^{N} f_1(T_1^n x) f_2(T_2^n x)$ do not converge even in norm when the transformations T_1 and T_2 do not necessarily commute. In the first section of this paper we will extend this result by proving the following theorem.

THEOREM 1. *Let $(X, \mathcal{B}, \mu, T_i)$, $1 \le i \le 6$, be six measure preserving systems on the same finite measure space and consider f_i, $1 \le i \le 6$ bounded functions. Then the averages*

$$\frac{1}{N^3} \sum_{n,m,p=1}^{N} f_1(T_1^n x) f_2(T_2^m x) f_3(T_3^p x) f_4(T_4^{n+m} x) f_5(T_5^{n+p} x) f_6(T_6^{m+p} x)$$

converge almost everywhere and in norm.

The method used to prove this theorem is a combination of the following key estimates obtained in [**1**] and the ergodic decomposition.

Department of Mathematics, UNC Chapel Hill, NC 27599, assani@math.unc.edu.

Keywords: not necessarily commuting m.p.t., averages along the cubes, Wiener Wintner averages.

AMS subject classification 37A05, 37A30.

Lemma 1. *Let a_n, b_n and c_n, $n \in \mathbb{N}$ be three sequences of scalars that we assume for simplicity bounded by one. Then for each N positive integer we have*

$$\Big|\frac{1}{N^2}\sum_{m,n=0}^{N-1} a_n.b_m.c_{n+m}\Big|^2$$
$$\leq \min\left[\sup_t\Big|\frac{1}{N}\sum_{m'=1}^{2(N-1)} c_{m'}e^{2\pi i m' t}\Big|^2, \sup_t\Big|\frac{1}{N}\sum_{n'=1}^{N} a_{n'}e^{2\pi i n' t}\Big|^2, \sup_t\Big|\frac{1}{N}\sum_{n''=1}^{N} b_{n''}e^{2\pi i n'' t}\Big|^2\right]$$

Lemma 2. *Let*

$$M_N(A_1, A_2, ..., A_7) = \frac{1}{N^3}\sum_{p,n,m=0}^{N-1} a_{1,p}a_{2,n}a_{3,p+n}a_{4,m}a_{5,n+m}a_{6,p+m}a_{7,n+m+p}$$

the averages of seven bounded (by one) sequences $A_i = (a_{i,n})$, $1 \leq i \leq 7$. Let us denote by $\mathcal{G}$ the set of couples of integers between 1 and 7, (i, j), which are connected by one of the indices n, m or p. Then for each N positive integer we have

$$\Big|M_N(A_1, A_2, ..., A_7)\Big|^2$$
$$\leq C \min_{(i,j)\in\mathcal{G}}\left[\max\Big[\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\Big|\frac{1}{N}\sum_{m=0}^{N-1} a_{i,m}a_{j,n+m}e^{2\pi i m t}\Big|^2,\right.$$
$$\left.\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\Big|\frac{1}{N}\sum_{m=0}^{2(N-1)} a_{i,m}a_{j,n+m}e^{2\pi i m t}\Big|^2\Big]\right].$$

With these lemmas we will derive the pointwise convergence of Wiener-Wintner types of averages that will lead to the conclusion stated in Theorem 1. These pointwise results extend Wiener-Wintner classical ergodic theorem. (see [**2**], for instance for several proofs of this Wiener Wintner result).

This is done in a first subsection. In a second subsection we will study the problem of recurrence to a single set in the case of three transformations. We will be able to extend Khintchine's recurrence result by studying for any measurable set A with positive measure the positivity of the limit

$$\lim_N \frac{1}{N^2}\sum_{n,m=0}^{N-1}\mu\{A\cap T_1^n A\cap T_2^m A\cap T_3^{n+m}A\} > 0$$

when the transformations are not necessarily commuting.

In the second section of the paper we will look at the convergence of weighted averages. For a measure preserving transformation T we denote by $\mathcal{K}$ the σ-algebra spanned by the eigenfunctions of T. The method used in [**1**] to prove the pointwise convegence of averages along the cubes for the powers of the same measure preserving transformation led to the following results.

Lemma 3. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system and let $f \in \mathcal{K}^\perp$. Then for μ a.e. x for all bounded sequences a_n , b_n , c_n,*

(1) $\lim_N \frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}x) = 0,$

(2) $\lim_N \frac{1}{N^2} \sum_{n,m=0}^{N-1} f(T^n x) b_m c_{n+m} = 0$ *and*

(3) $\lim_N \frac{1}{N^2} \sum_{n,m=0}^{N-1} a_n f(T^m x) c_{n+m} = 0.$

and

Proposition 2. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system and let $f \in L^2(\mu)$. Then for μ a.e. x for all bounded sequences a_n, b_n such that $\frac{1}{N} \sum_{n=0}^{N-1} a_n e^{2\pi int}$ and $\frac{1}{N} \sum_{n=0}^{N-1} b_n e^{2\pi int}$ converge for each t, the sequence*

$$\frac{1}{N^2} \sum_{n=0}^{N-1} a_n b_m f(T^{n+m} x)$$

converges. A similar statement holds if one replaces a_n with $f(T^n x)$ and uses instead b_m and c_{n+m} or if one chooses $b_m = f(T^m x)$ and uses a_n and c_{n+m}.

The intriguing aspect of these results is the fact that the set of convergence for x is independent of the bounded sequences a_n, b_n and c_n. An illustration of such property can be given by taking $a_n = (f_1(T_1^n x))$, $b_m = f_2(T_2^m x)$ with $f_1, f_2 \in L^\infty$. One obtains immediately the almost everywhere convergence of the averages

$$\frac{1}{N^2} \sum_{n,m=1}^{N} f_1(T_1^n x) f_2(T_2^m x) f(T^{n+m} x)$$

if the transformation T is ergodic. Other choices for the sequences a_n and b_n are also possible. For instance one could easily take $a_n = f(T^{p(n)})(x)$ where $p(x)$ is a real polynomial with positive integer coefficients. . Such observation seemed to indicate that the almost everywhere convergence of the averages along the cubes of a single transformation, namely $\frac{1}{N^2} \sum_{n,m=1}^{N} f(T^n x) g(T^m x) h(T^{n+m} x)$, relies more on the underlying arithmetic structure than on its dynamical structure. This is one of the reasons why we asked in [**1**] if the assumption of ergodicity made in Lemma 3 and the Proposition 2 above was necessary. In this paper we will answer in part this question by showing that the ergodicity assumption is indeed necessary in Lemma 3. At the present time we do not know if Proposition 2 is true without ergodicity assumption. With the method used in [**1**] we have the following

Proposition 3. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $f \in L^2(\mu)$. Define the set $\mathcal{WW}_1$ as*

$$\mathcal{WW}_1 = \big\{ a \in l^\infty ; \lim_N \frac{1}{N} \sum_{n=0}^{N-1} a_n e^{2\pi int} \textit{exists for all } t \big\}.$$

If the set

$$D = \big\{ x : \frac{1}{N^2} \sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m} x) \textit{converge for all } (a_n) \in \mathcal{WW}_1,\ (b_m) \in \mathcal{WW}_1 \big\}$$

is measurable then for μ a.e. x for all bounded sequences $(a_n) \in \mathcal{WW}_1$, $(b_m) \in \mathcal{WW}_1$ the averages

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}x)$$

converge.

The currently open question is the measurability of D that we will not address in this paper.

It is worth pointing out that if one looks only at the norm convergence Proposition 2 is true without ergodicity.

We will also look at the higher order averages. We denote by $A_i = (a_{n,i})$ $1 \leq i \leq 6$, six bounded sequences of scalars. We consider the averages

$$M_N(A_1, A_2, ...A_6, f)(x) = \frac{1}{N^3}\sum_{n,m,p=0}^{N-1} a_{1,p}a_{2,n}a_{3,p+n}a_{4,m}a_{5,n+m}a_{6,p+m}f(T^{n+m+p}x).$$

In [**1**] we proved that if $f \in \mathcal{CL}^{\perp}$ then we have a similar result to Lemma 1. More precisely we have;

Proposition 4. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system and let $f \in CL^{\perp}$. Then for μ a.e. x for all bounded sequences $A_i = (a_{i,n})$, $1 \leq i \leq 6$ the sequence*

$$M_N(A_1, A_2, ...A_6, f)(x) = \frac{1}{N^3}\sum_{n,m,p=0}^{N-1} a_{1,p}a_{2,n}a_{3,p+n}a_{4,m}a_{5,n+m}a_{6,p+m}f(T^{n+m+p}x)$$

converge to zero.

A natural question is to find the precise condition on the sequences A_i that will give the almost everywhere convergence of the averages $M_N(A_1, A_2, ...A_6, f)(x)$ when $f \in \mathcal{CL}$. We will show that a condition such as $\lim_N \frac{1}{N}\sum_{n=1}^{N} a_{n,i}e^{2\pi int}$ exists for each $t \in \mathbb{R}$ which is actually necessary and sufficient for the convergence of the weighted averages

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}x)$$

in the universal sense described by Proposition 2 is no longer sufficient for the convergence of the averages $M_N(A_1, A_2, ...A_6, f)(x)$. At the present time sufficient conditions on the sequences that would guarantee the almost convergence are not yet clear to us.

2. Almost everywhere convergence and recurrence for not necessarily commuting measure preserving transformations

2.1. Proof of Theorem 1. We recall that if $(X, \mathcal{B}, \mu, T)$ is a measure preserving dynamical system then the measure μ can be disintegrated in a product so that $d\mu = d\mu_c dc$ and $(X, \mathcal{B}, \mu_c)$ becomes an ergodic dynamical system. This disintegration allows to lift several results from the ergodic case to the not necessarily ergodic one.

The proof of Theorem 1 will be completed after several steps. First we will need a Wiener Wintner strengthening of Theorem 10 in [**1**].

Lemma 4. *Let $(X, \mathcal{B}, \mu, T_i)$ be three measure preserving transformations on the same finite measure space. Consider three bounded functions f_i, $1 \leq i \leq 3$ then for μ a.e. x for all $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ the averages*

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n x) f_2(T_2^m x) f_3(T_3^{n+m} x) e^{2\pi i n\varepsilon_1} e^{2\pi i m\varepsilon_2}$$

converge.

Proof. Without loss of generality we can assume that the functions f_i are bounded by one. We use the ergodic decomposition with respect to T_3 to obtain a disintegration of μ, $d\mu = d\mu_{c,3}dc$ into ergodic components. By the same disintegration and because of the Wiener Wintner theorem for measure preserving transformations for c a.e., for $\mu_{c,3}$ a.e. y, the averages

$$\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n y) e^{2\pi i n\varepsilon_1}$$

and

$$\frac{1}{N}\sum_{m=0}^{N-1} f_2(T_2^m y) e^{2\pi i m\varepsilon_2}$$

converge for all $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$. It is clear that the transformations T_1 and T_2 may no longer be measure preserving with respect to $\mu_{c,3}$ but we are only using here the disintegration of measurable sets of full measure given by Wiener Wintner ergodic theorem. Let us consider the Kronecker factor $\mathcal{K}_{3,c}$ of T_3 with respect to $(X, \mathcal{B}, \mu_{c,3})$ and let us decompose the function f_3 into the sum $f_{3,K_c} + f_{3,K_c^\perp}$ where f_{3,K_c} is its projection onto $\mathcal{K}_{3,c}$. By Bourgain's uniform Wiener Wintner ergodic theorem (see [**2**] for instance for a proof) we have for $\mu_{3,c}$ a.e. y

$$\lim_N \sup_t \Big|\frac{1}{N}\sum_{n=0}^{N-1} f_{3,K_c}(T_3^n y) e^{2\pi i n t}\Big| = 0.$$

Applying Lemma 1 with $a_n = f_1(T_1^n y)e^{2\pi i n\varepsilon_1}$, $b_m = f_2(T_2^m y)e^{2\pi i m\varepsilon_2}$, and $c_k = f_3(T_3^k y)$ we obtain the estimate

$$\sup_{\varepsilon_1,\varepsilon_2}\left|\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n y) f_2(T_2^m y) f_{3,K_c^\perp}(T_3^{n+m} y) e^{2\pi i n\varepsilon_1} e^{2\pi i m\varepsilon_2}\right|$$
$$\leq \sup_t \Big|\frac{1}{N}\sum_{k=0}^{N-1} f_{3,K_c^\perp}(T_3^k y) e^{2\pi i k t}\Big|.$$

As a consequence of the uniform Wiener Wintner theorem we have

$$\lim_N \sup_{\varepsilon_1,\varepsilon_2}\left|\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n y) f_2(T_2^m y) f_{3,K_c^\perp}(T_3^{n+m} y) e^{2\pi i n\varepsilon_1} e^{2\pi i m\varepsilon_2}\right| = 0.$$

The function f_{3,K_c} projects onto the eigenfunctions of T_3 with respect to $\mu_{c,3}$. If $e_{j,3}$ is one of these eigenfunctions with corresponding eigenvalue $e^{2\pi i\theta_j}$ then we have

$$f_{3,K_c} = \Big[\sum_{j=0}^{\infty}\int f_{3,K_c}(y)\overline{e_{j,3}}(y)d\mu_{c,3}(y)e_{j,3}\Big].$$

Hence by linearity and approximation it is enough to consider the case where f_{3,K_c} is one of the eigenfunctions $e_{j,3}$. In this case $f_{3,K_c}(T_3^{n+m}y) = e^{2\pi i(n+m)\theta_j}e_{j,3}$ and the averages become

$$e_{j,3}\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)e^{2\pi i(n+m)\theta_j}e^{2\pi in\varepsilon_1}e^{2\pi im\varepsilon_2}$$
$$= e_{j,3}\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n y)e^{2\pi in(\theta_j+\varepsilon_1)}f_2(T_2^m y)e^{2\pi im(\theta_j+\varepsilon_2)}$$

The convergence can be derived now by the disintegration, done at the beginning of the proof, of the sets where the Wiener Wintner ergodic theorem applied to the functions f_1 and f_2.

As the set of x for which for all $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ the averages

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)e^{2\pi in\varepsilon_1}e^{2\pi im\varepsilon_2}$$

converge is $\mathcal{B}$ measurable we can integrate with respect to $\mu_{c,3}$ and dc to show that this set has full measure.

□

Lemma 5. *Let $(X, \mathcal{B}, \mu, T_i)$ be four measure preserving transformations on the same finite measure space. Then for all bounded functions f_i, $1 \le i \le 4$, the averages*

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)f_4(T_4^p x)$$

converge μ a.e. and in norm.

Proof. We can write these averages as

$$\Big[\frac{1}{N}\sum_{p=0}^{N-1} f_4(T_4^p x)\Big]\Big[\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)\Big].$$

The conclusion now follows from Birkhoff's pointwise ergodic theorem and Theorem 10 in [**1**]. The convergence in norm is an easy consequence of Lebesgue dominated convergence theorem. □

We need now a Wiener Wintner version of Lemma 5.

Lemma 6. *Let $(X, \mathcal{B}, \mu, T_i)$ be four measure preserving transformations on the same finite measure space. Then for all bounded functions f_i, $1 \le i \le 4$, for μ a.e. x, for all $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$ the averages*

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)f_4(T_4^p x)e^{2\pi i(n+p)\varepsilon_1}e^{2\pi i(m+p)\varepsilon_2}$$

converge.

Proof. We can rewrite the averages as

$$\Big[\frac{1}{N}\sum_{p=0}^{N-1} f_4(T_4^p x)e^{2\pi ip(\varepsilon_1+\varepsilon_2)}\Big]\Big[\frac{1}{N^2}\sum_{n,m=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)e^{2\pi in\varepsilon_1}e^{2\pi im\varepsilon_2}\Big].$$

The a.e. convergence is a consequence of the Wiener Wintner ergodic theorem for measure preserving transformations and Lemma 4.

□

LEMMA 7. *Let $(X, \mathcal{B}, \mu, T_i)$ be five measure preserving transformations on the same finite measure space. Then for all bounded functions f_i, $1 \leq i \leq 5$, for μ a.e. x the averages*

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)f_4(T_4^p x)f_5(T_5^{n+p}x)$$

converge.

PROOF. We follow the path of the proof of Lemma 4. The set where the averages converge is $\mathcal{B}$ measurable. We use the ergodic decomposition of $(X, \mathcal{B}, \mu, T_5)$ into ergodic components on $(X, \mathcal{B}, \mu_{c,5})$. We disintegrate the set where the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m}x)f_4(T_4^p x)e^{2\pi i(n+p)\varepsilon}$$

converge for each $\varepsilon \in \mathbb{R}$. We decompose the function f_5 into its projection onto the corresponding Kronecker factor f_{5,K_c} and $f_{5,K_c^\perp}$. By considering first the case of one eigenfunction then by approximation and linearity we obtain for $\mu_{c,5}$ a.e. y the convergence of the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)f_3(T_3^{n+m}y)f_4(T_4^p y)f_{5,K_c}(T_5^{n+p}y).$$

We can dominate the averages with the function $f_{5,K_c^\perp}$ by their absolute value

$$\left|\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)f_3(T_3^{n+m}y)f_4(T_4^p y)f_{5,K_c^\perp}(T_5^{n+p}y)\right|$$

which in turn are bounded by

$$\frac{1}{N^3}\sum_{m=0}^{N-1}|f_2(T_2^m y)|\sum_{n=0}^{N-1}|f_1(T_1^n y)||f_3(T_3^{n+m}y)||\sum_{p=0}^{N-1} f_4(T_4^p y)f_{5,K_c^\perp}(T_5^{n+p}y)|.$$

Using the fact that the functions are uniformly bounded (by one without loss of generality) we get the upper bound

$$\frac{1}{N}\sum_{n=0}^{N-1}\Big|\frac{1}{N}\sum_{p=0}^{N-1} f_4(T_4^p y)f_{5,K_c^\perp}(T_5^{n+p}y)\Big|.$$

We can apply the remark made after the proof of Lemma 5 in [**1**] to obtain the bound

$$\sup_t\Big|\frac{1}{N}\sum_{k=0}^{N-1} f_{5,K_c^\perp}(T_5^k y)e^{2\pi ikt}\Big|$$

which converges to zero by the uniform Wiener-Wintner ergodic theorem. By combining the convergence obtained for functions f_{5,K_c} and $f_{5,K_c^\perp}$ we can reach the

$\mu_{5,c}$ a.e. y convergence of the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)f_3(T_3^{n+m} y)f_4(T_4^p y)f_{5,K_c}(T_5^{n+p} y).$$

The convergence μ a.e. x can be obtained by integration with respect to $d\mu_{5,c}dc$.

□

It remains to add one more transformation and function, namely T_6 and f_6. The path is quite clear . We start with a Wiener Wintner version of the Lemma 7.

LEMMA 8. *Let $(X, \mathcal{B}, \mu, T_i)$ be five measure preserving transformations on the same finite measure space. Then for all bounded functions f_i, $1 \leq i \leq 5$, for μ a.e. x for all $t \in \mathbb{R}$ the averages*

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m} x)f_4(T_4^p x)f_5(T_5^{n+p} x)e^{2\pi i(m+p)t}$$

converge.

PROOF. We reconsider the disintegration of the measure μ into ergodic components with respect to $(X, \mathcal{B}, \mu_{5,c})$. We disintegrate the measurable set where the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m} x)f_4(T_4^p x)e^{2\pi i(n+p)\varepsilon_1}e^{2\pi i(m+p)\varepsilon_2}$$

converge for all $\varepsilon_1, \varepsilon_2 \in \mathbb{R}$. We disintegrate also the measurable set where the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m} x)f_4(T_4^p x)f_5(T_5^{n+p} x)e^{2\pi i(m+p)t}$$

converge for all $t \in \mathbb{R}$. We decompose the function f_5 into its projection onto the corresponding Kronecker factor f_{5,K_c} and $f_{5,K_c^\perp}$. Again by approximation and linearity it is enough to look at the case of an eigenfunction $e_{j,5}$ with eigenvalue $e^{2\pi i\theta_j}$. The averages in this case are equal to

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)f_3(T_3^{n+m} y)f_4(T_4^p y)e^{2\pi i(n+p)\theta_j}e^{2\pi i(m+p)t}$$

and converge $\mu_{c,5}$ a.e. y for all t. We are left with the averages related to the function $f_{5,K_c^\perp}$. By observations similar to those made in Lemma 7 we obtain for each t the upper bound

$$\frac{1}{N}\sum_{n=0}^{N-1}\Big|\frac{1}{N}\sum_{p=0}^{N-1} f_4(T_4^p y)e^{2\pi ipt}f_{5,K_c^\perp}(T_5^{n+p} y)\Big|.$$

This last term is dominated by

$$\sup_s\Big|\frac{1}{N}\sum_{k=0}^{N-1} f_{5,K_c^\perp}(T_5^k y)e^{2\pi iks}\Big|$$

and the convergence follows by the uniform Wiener Wintner ergodic theorem.

□

End of the proof of Theorem 1

We consider $(X, \mathcal{B}, \mu, T_i)$ six measure preserving transformations on the same finite measure space and six bounded functions f_i, $1 \leq i \leq 6$. We want to prove that for μ a.e. x the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^{n+m} x)f_4(T_4^p x)f_5(T_5^{n+p} x)f_6(T_6^{m+p} x)$$

converge.

We use an ergodic decomposition of the measure μ into ergodic components for T_6. As in the previous lemmas this reduces the study of the convergence on these components. The function f_6 is decomposed into the sum f_{6,K_c} and $f_{6,K_c^\perp}$. The convergence $\mu_{6,c}$ a.e. y for f_{6,K_c} is obtained by linearity, approximation and the use of Lemma 8. It remains to prove the convergence for the averages related to $f_{6,K_c^\perp}$. We can use Lemma 2 with the sequence $a_{7,k} = 1$ (see also the proof of Lemma 6 in [**1**]) to obtain the following inequalities:

$$\Big|\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)f_3(T_3^{n+m} y)f_4(T_4^p y)f_5(T_5^{n+p} y)f_{6,K_c^\perp}(T_6^{m+p} y)\Big|$$

$$= \Big|\frac{1}{N^3}\sum_{n,m=0}^{N-1} f_1(T_1^n y)f_2(T_2^m y)f_3(T_3^{n+m} y)\sum_{p=0}^{N-1} f_4(T_4^p y)f_5(T_5^{n+p} y)f_{6,K_c^\perp}(T_6^{m+p} y)\Big|$$

$$\leq \frac{1}{N^3}\sum_{n,m=0}^{N-1}\big|\sum_{p=0}^{N-1} f_4(T_4^p y)f_5(T_5^{n+p} y)f_{6,K_c^\perp}(T_6^{m+p} y)\big|$$

$$\leq \frac{1}{N^2}\Big(\sum_{n,m=0}^{N-1}\big|\sum_{p=0}^{N-1} f_4(T_4^p y)f_5(T_5^{n+p} y)f_{6,K_c^\perp}(T_6^{m+p} y)\big|^2\Big)^{1/2}$$

(by Cauchy-Schwartz inequality)

$$= \frac{1}{N^2}$$

$$\Big(\sum_{n,m=0}^{N-1}\big|\int\Big((\sum_{p=0}^{N-1} f_4(T_4^p y)f_5(T_5^{n+p} y)e^{-2\pi ips})(\sum_{k=0}^{2N-1} f_{6,K_c^\perp}(T_6^k y)e^{2\pi iks})e^{-2\pi ims}\Big)ds\big|^2\Big)^{1/2}$$

$$\leq \frac{1}{N^2}\Big(\sum_{n=0}^{N-1}\int\big|(\sum_{p=0}^{N-1} f_4(T_4^p y)f_5(T_5^{n+p} y)e^{-2\pi ips})(\sum_{k=0}^{2N-1} f_{6,K_c^\perp}(T_6^k y)e^{2\pi iks})\big|^2 ds\Big)^{1/2}$$

(by Parseval 's equality)

$$\leq \sup_s\big|\frac{1}{N}\sum_{k=0}^{2N-1} f_{6,K_c^\perp}(T_6^k y)e^{2\pi iks}\big|\frac{1}{N}\Big(\sum_{n=0}^{N-1}(\int\big|(\sum_{p=0}^{N-1} f_4(T_4^p y)f_5(T_5^{n+p} y)e^{-2\pi ips})\big|^2 ds)\Big)^{1/2}$$

$$\leq \sup_s\big|\frac{1}{N}\sum_{k=0}^{2N-1} f_{6,K_c^\perp}(T_6^k y)e^{2\pi iks}\big|$$

The conclusion of the theorem follows after using the uniform Wiener Wintner ergodic theorem and integration.

2.2. An extension of Khintchine recurrence theorem. Khintchine classical recurrence theorem says that if A is a set of positive measure, T an invertible measure preserving system and $\varepsilon > 0$ the set

$$\{n \in \mathbb{Z} : \int \mathbf{1}_A . \mathbf{1}_A \circ T^n d\mu \geq [\int \mathbf{1}_A d\mu]^2 - \varepsilon\}$$

has bounded gaps. This recurrence result states that for any measurable set A with positive measure its images under the iterates of T come back and overlap the set with bounded gaps. This is a consequence of von Neumann mean ergodic theorem as

$$\lim_{N\to\infty} \int \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_A . \mathbf{1}_A \circ T^n d\mu \geq \mu(A)^2.$$

In this section we study similar recurrence properties with two and three measure preserving transformations that do not necessarily commute. We first give a two dimensional extension of Khintchine's theorem. We can remark that an example given in [**6**] shows that the averages

$$\frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T_1^{-n} A \cap T_2^{-m} A \cap T_1^{-n} T_2^{-m} A)$$

may diverge if T_1 and T_2 do not necessarily commute.

PROPOSITION 5. *Let $(X, \mathcal{B}, \mu)$ be a probability measure space and T_1, T_2 two measure preserving transformations on this measure space. We denote by $\mathcal{I}_1$ and $\mathcal{I}_2$ the σ algebras of the invariant sets for T_1 and T_2. Consider A a set of positive measure. Then*

$$\lim_N \frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T_1^{-n} A \cap T_2^{-n-m} A) = \int_A \mathbb{E}(\mathbf{1}_A, \mathcal{I}_1)(x) . \mathbb{E}(\mathbf{1}_A, \mathcal{I}_2)(x) d\mu.$$

In particular

$$\lim_N \frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T_1^{-n} A \cap T_2^{-n-m} A) \geq \mu(A)^4.$$

PROOF. The averages

$$\frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T_1^{-n} A \cap T_2^{-n-m} A)$$

are the integrals of the functions

$$\frac{1}{N^2} \sum_{n,m=1}^{N} \mathbf{1}_A(x) \mathbf{1}_A(T_1^n x) \mathbf{1}_A(T_2^{n+m} x)$$

with respect to the measure μ. As a particular case of Theorem 1 we have the pointwise convergence of these averages. Thus

$$\lim_N \frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T_1^{-n} A \cap T_2^{-n-m} A)$$

exists after integration. So we just have to prove that

$$\lim_N \frac{1}{N^2}\sum_{n,m=1}^{N}\mathbf{1}_A(T_1^n x)\mathbf{1}_A(T_2^{n+m}x)=\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)(x).\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)$$

in L^2 norm to conclude. For each N we have

$$\frac{1}{N^2}\sum_{n,m=1}^{N}\mathbf{1}_A(T_1^n x)\mathbf{1}_A(T_2^{n+m}x)$$

$$=\frac{1}{N^2}\sum_{n,m=1}^{N}\mathbf{1}_A(T_1^n x)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)+\frac{1}{N^2}\sum_{n,m=1}^{N}\mathbf{1}_A(T_1^n x)[\mathbf{1}_A(T_2^{n+m}x)-\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)]$$

The first term of the last equation converges by Birkhoff's pointwise ergodic theorem to $\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)(x).\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)$. Noticing that the function $\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)$ is T_2 invariant we can bound the L^2 norm of the second term by

$$\|\frac{1}{N}\sum_{n=1}^{N}|\frac{1}{N}\sum_{m=1}^{N}[\mathbf{1}_A\circ T_2^m-\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)]\circ T_2^n\|_2.$$

This term is less than

$$\frac{1}{N}\sum_{n=1}^{N}\||\frac{1}{N}\sum_{m=1}^{N}[\mathbf{1}_A\circ T_2^m-\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)]\|_2$$

which is equal to

$$\|\frac{1}{N}\sum_{m=1}^{N}[\mathbf{1}_A\circ T_2^m-\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)]\|_2$$

This last term tends to zero by the mean ergodic theorem applied to T_2. This proves that $\lim_N \|\frac{1}{N^2}\sum_{n,m=1}^{N}\mathbf{1}_A(T_1^n x)\mathbf{1}_A(T_2^{n+m}x)-\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)(x).\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)\|_2=0$.
It remains to show that

$$\int_A \mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)(x).\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)d\mu\geq\mu(A)^4.$$

We have

$$\int_A \mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)(x).\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)(x)d\mu=\int\mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1),\mathcal{I}_2)\big)\mathbf{1}_A d\mu$$

$$=\int\mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1),\mathcal{I}_2)\big)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu$$

(and as $\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)(x)\leq 1$ we have $\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)\geq\mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1),\mathcal{I}_2\big)$

$$\geq\int\Big(\mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1),\mathcal{I}_2)\big)\Big)^2 d\mu$$

$$\geq\Big(\int\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)d\mu\Big)^2=\Big(\int\big(\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)\big)^2 d\mu\Big)^2$$

$$\geq\Big(\int\mathbf{1}_A d\mu\Big)^4=\mu(A)^4$$

□

The study of the case of three measure preserving transformations seems much more complex.

Lemma 9. *Let $(X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system on a finite measure space, $\mathcal{K}$ the σ algebra spanned by the eigenfunctions of T and f a bounded function. Let us denote by X_f the set of full measure given by the Wiener Wintner ergodic theorem such that for each $x \in X_f$ the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)e^{-2\pi int}$$

converge for each t. For each $t \in \mathbb{R}$ let us denote by $E_t(f)$ the limit function of these averages.

(1) *$E_t(f)$ is the projection of the function f onto the eigenspace of T corresponding to the eigenvalue $e^{2\pi it}$. In particular $E_0(f)$ is equal to $\mathbb{E}(f, \mathcal{I})$ the conditional expectation with respect to the σ algebra of invariant sets for T.*
(2) *If $t \neq s$ we have $\int E_t(f)\overline{E_s(f)}d\mu = 0$.*
(3) *If $e^{2\pi i\theta_k}$ is the countable sequence of eigenvalues for T and $e^{2\pi it_k}$ any countable set of distinct complex numbers then*

$$\sum_{k=0}^{\infty} \|E_{t_k}(f)\|_2^2 \leq \sum_{k=0}^{\infty} \|E_{\theta_k}(f)\|_2^2 \leq \|E(f, \mathcal{K})\|_2^2$$

Proof. This is a simple consequence of the spectral theorem. If we denote by $P_t(f)$ the projection onto the eigenspace corresponding to the eigenvalue $e^{2\pi it}$ and by $\sigma_{f-P_t(f)}$ the spectral measure of the function $f - P_t(f)$ then we have

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)e^{-2\pi int} = \frac{1}{N}\sum_{n=1}^{N} P_t(f)(T^n x)e^{-2\pi int} + \frac{1}{N}\sum_{n=1}^{N}[f - P_t(f)](T^n x)e^{-2\pi int}$$

$$= P_t(f)(x) + \frac{1}{N}\sum_{n=1}^{N}[f - P_t(f)](T^n x)e^{-2\pi int}$$

As

$$\lim_N \|\frac{1}{N}\sum_{n=1}^{N}[f - P_t(f)](T^n x)e^{-2\pi int}\|_2^2 = \int \left|\frac{1}{N}\sum_{n=1}^{N} e^{2\pi in(\theta - t)}\right|^2 d\sigma_{(f-P_t(f))}(\theta)$$

$$= \sigma_{f-P_t(f)}(\{0\}) = 0,$$

we can conclude that $P_t(f) = E_t(f)$.

From this identification the remaining parts of the lemma follow without difficulty. For the last part of the lemma we just need to observe that $E_t(f) = 0$ if $e^{2\pi it}$ is not an eigenvalue of T. □

Remark It is worth noticing that there is a key difference at the pointwise level between $E_t(f)(x)$ and $P_t(f)(x)$. This difference highlights the difficulty one faces when dealing with ergodic versus not necessarily ergodic transformations. The function $E_t(f)$ is defined off a single set of measure zero for ALL $t \in \mathbb{R}$. For each $t \in \mathbb{R}$ it is almost everywhere equal to the function $P_t(f)(x)$ and so for each t the L^2 functions $P_t(f)$ and $E_t(f)$ are equal. However we can not claim that there is a universal null set off which one could write that $E_t(f)(x) = P_t(f)(x)$ for all $t \in \mathbb{R}$. One can look at the example given in Proposition 7 below.

PROPOSITION 6. *Let $(X, \mathcal{B}, \mu, T_i)$, $1 \leq i \leq 3$, be three measure preserving systems on the same finite measure space. There exists a constant $0 < \delta < 1$ (independent from the T_i) such that for all measurable set A with measure $\mu(A) > 1 - \delta$ we have*

$$\lim_N \frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T_1^{-n} A \cap T_2^{-m} A \cap T_3^{-n-m} A) \geq \frac{1}{2} \mu(A)^8$$

PROOF. We consider three measure preserving transformations T_j, $1 \leq j \leq 3$ and a measurable set A with positive measure. We list the following notations and properties.

(1) We denote by $E_t^j(\mathbf{1}_A)(x)$ the limit of
$$\frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_A(T_j^n x) e^{-2\pi i n t}$$
for all $t \in \mathbb{R}$ off a single set of measure zero.
(2) For each $1 \leq j \leq 3$ we consider the universal sets $X_{\mathbf{1}_A}^j$ such that $E_t^j(\mathbf{1}_A)(x)$ exists for all $t \in \mathbb{R}$.
(3) We consider an ergodic decomposition of $(X, \mathcal{B}, \mu, T_3)$ with the measures μ_c where $d\mu = d\mu_c dc$.
(4) We call $\mathcal{K}_c$ the Kronecker factor of T_3 relative to the measure space $(X, \mathcal{B}, \mu_c)$. The basis of eigenfunctions of T_3 relative to μ_c is denoted by $e_{k,c}$. The constant function 1 corresponds to $e_{0,c}$.
(5) The eigenvalue corresponding to the eigenfunction $e_{k,c}$ is $e^{-2\pi i \theta_{k,c}}$.
(6) By Birkhoff pointwise ergodic theorem combined with the disintegration of μ we have for a.e. c for μ_c a.e. y $\lim_N \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_A(T_3^n y) = E(\mathbf{1}_A, \mathcal{I})(y) = E_0^3(\mathbf{1}_A)(y)$, where $\mathcal{I}$ denotes the σ algebra of T_3 invariant sets with respect to μ.

As a consequence of the ergodicity of T_3 with respect to μ_c we have

$$\mathbb{E}(\mathbf{1}_A, \mathcal{K}_c) = \sum_{k=0}^{\infty} \Big(\int \mathbf{1}_A \overline{e_{k,c}} d\mu_c \Big) e_{k,c}.$$

We disintegrate the measurable sets $X_{\mathbf{1}_A}^j$ with respect to the measure $d\mu_c$. We obtain for μ_c a.e. y for all $t \in \mathbb{R}$ the pointwise convergence of the averages

$$\frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_A(T_j^n y) e^{-2\pi i n t}.$$

This is crucial for our method as with respect to the measure μ_c the transformations T_1 and T_2 are not necessarily measure preserving.

For each eigenfunction $e_{k,c}$ we have

$$\lim_N \frac{1}{N^2}\sum_{n,m=1}^N \mathbf{1}_A(T_1^n y)\mathbf{1}_A(T_2^m y)e_{k,c}(T_3^{n+m}y)$$
$$= e_{k,c}(y)\lim_N \frac{1}{N^2}\sum_{n,m=1}^N \mathbf{1}_A(T_1^n y)e^{-2\pi i n\theta_{k,c}}\mathbf{1}_A(T_2^m y)e^{-2\pi i m\theta_k}$$
$$= e_{k,c}(y)E^1_{\theta_{k,c}}(\mathbf{1}_A)(y)E^2_{\theta_{k,c}}(\mathbf{1}_A)(y)$$

We can observe that Lemma 3 implies that

$$\lim_N \frac{1}{N^2}\sum_{n,m=1}^N \int \mathbf{1}_A(T_1^n y)\mathbf{1}_A(T_2^m y)\big(\mathbf{1}_A - E(\mathbf{1}_A)(T_3^{n+m}y)d\mu_c = 0$$

for almost every c. As a consequence we have

$$\lim_N \frac{1}{N^2}\sum_{n,m=1}^N \mu_c\big(A\cap T_1^n A\cap T_2^m A\cap T_3^{n+m}A\big)$$
$$= \sum_{k=0}^{\infty}\big(\int \mathbf{1}_A\overline{e_{k,c}}d\mu_c\big)\big(\int \mathbf{1}_A e_{k,c}E^1_{\theta_{k,c}}(\mathbf{1}_A)E^2_{\theta_{k,c}}(\mathbf{1}_A)d\mu_c\big)$$
$$= \mu_c(A)\big(\int \mathbf{1}_A E^1_0(\mathbf{1}_A)E^2_0(\mathbf{1}_A)d\mu_c\big)$$
$$+\sum_{k=1}^{\infty}\big(\int \mathbf{1}_A\overline{e_{k,c}}d\mu_c\big)\big(\int \mathbf{1}_A e_{k,c}E^1_{\theta_{k,c}}(\mathbf{1}_A)E^2_{\theta_{k,c}}(\mathbf{1}_A)d\mu_c\big)$$

The first term of the previous line is for a.e. c equal to

$$\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\int \mathbf{1}_A(y)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu_c.$$

In view of the constance of $\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)$ with respect to μ_c, this last term can be written as

$$\int \mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\mathbf{1}_A(y)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu_c.$$

Integrating with respect to dc and using properties of the conditional expectation we get

$$\int \mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\mathbf{1}_A(y)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu_c dc = \int \mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu(x)$$
$$= \int \mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2),\mathcal{I}_1\big)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_1)d\mu$$
$$\geq \int \big(\mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2),\mathcal{I}_1\big)\big)^2 d\mu \geq \bigg(\int \mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu\bigg)^2$$
$$= \bigg(\int \mathbb{E}\big(\mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3),\mathcal{I}_2\big)\mathbb{E}(\mathbf{1}_A,\mathcal{I}_2)d\mu\bigg)^2$$
$$= \bigg(\int \big(\mathbb{E}\big(\mathbf{1}_A,\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3),\mathcal{I}_2\big)\big)^2 d\mu\bigg)^2 \geq \bigg(\int \mathbf{1}_A\mathbb{E}(\mathbf{1}_A,\mathcal{I}_3)d\mu\bigg)^4 \geq \mu(A)^8.$$

This shows that after integration with respect to $dc\,\mu_c(A)\big(\int \mathbf{1}_A E^1_0(\mathbf{1}_A)E^2_0(\mathbf{1}_A)d\mu_c\big)$ is bounded below by $\mu(A)^8$. Our proof will be complete if one can show that if $\mu(A) > 1-\delta$ for some universal $0<\delta<1$ then

(1)

$$\int |(I)_c|dc = \int \Big| \int \sum_{k=1}^{\infty} \big(\int \mathbf{1}_A \overline{e_{k,c}} d\mu_c\big)\big(\int \mathbf{1}_A e_{k,c} E^1_{\theta_{k,c}}(\mathbf{1}_A)E^2_{\theta_{k,c}}(\mathbf{1}_A)d\mu_c\big)\Big| dc \le \frac{1}{2}\mu(A)^8.$$

By Cauchy-Schwartz's inequality we have

$$|(I)_c| \le \Big(\sum_{k=1}^{\infty} \big|\int \mathbf{1}_A \overline{e_{k,c}} d\mu_c\big|^2\Big)^{1/2}\Big(\sum_{k=1}^{\infty} \big|\int \mathbf{1}_A e_{k,c} E^1_{\theta_{k,c}}(\mathbf{1}_A)E^2_{\theta_{k,c}}(\mathbf{1}_A)d\mu_c\big|^2\Big)^{1/2}.$$

The vectors $e_{k,c}$ form an orthonormal basis of $L^2(X,\mathcal{K}_c,\mu_c)$ because T_3 on this space is ergodic. Thus we have

(2)

$$\Big(\sum_{k=1}^{\infty} \big|\int \mathbf{1}_A \overline{e_{k,c}} d\mu_c\big|^2\Big)^{1/2} = \Big(\int \big|\mathbb{E}(\mathbf{1}_A,\mathcal{K}_c)\big|^2 d\mu_c - \mu_c(A)^2\Big)^{1/2} \le \big(\mu_c(A)-\mu_c(A)^2\big)^{1/2}$$

The second term $(II)_c = \Big(\sum_{k=1}^{\infty} \big|\int \mathbf{1}_A e_{k,c} E^1_{\theta_{k,c}}(\mathbf{1}_A)E^2_{\theta_{k,c}}(\mathbf{1}_A)d\mu_c\big|^2\Big)^{1/2}$ can also be bounded above by

$$\Big(\sum_{k=1}^{\infty}\Big(\int \mathbf{1}_A \big|E^1_{\theta_{k,c}}(\mathbf{1}_A)\big|^2 d\mu_c\Big)^2\Big)^{1/4}\Big(\sum_{k=1}^{\infty}\Big(\int \mathbf{1}_A \big|E^2_{\theta_{k,c}}(\mathbf{1}_A)\big|^2 d\mu_c\Big)^2\Big)^{1/4}.$$

Using Lemma 9 part (3), this last term is bounded above by

$$\Big(\int |\mathbb{E}(\mathbf{1}_A,\mathcal{K}_c)|^2 d\mu_c - \mu_c(A)^2\Big)^{1/4}\Big(\int |\mathbb{E}(\mathbf{1}_A,\mathcal{K}_c)|^2 d\mu_c - \mu_c(A)^2\Big)^{1/4}$$

which is equal to

$$\Big(\int |\mathbb{E}(\mathbf{1}_A,\mathcal{K}_c)|^2 d\mu_c - \mu_c(A)^2\Big)^(1/2) \le (\mu_c(A)-\mu_c(A)^2)^(1/2) \tag{3}$$

Combining the bounds found in (2) and in (3) we get

$|(I)_c| \le (\mu_c(A)-\mu_c(A)^2)$. As $\int \mu_c(A)^2 dc \ge (\int \mu_c(A)dc)^2$, integrating with respect to c we obtain

$$\int |(I)_c|dc \le \int (\mu_c(A)-\mu_c(A)^2)dc \le \mu(A)-\mu(A)^2.$$

Going back to (1) we will reach our conclusion if we can find $0<\delta<1$ such that

$$\mu(A)-\mu(A)^2 \le \frac{1}{2}\mu(A)^8,$$

for all measurable set A with measure greater or equal to $1-\delta$. This is an easy consequence of the uniqueness of the root for the polynomial $1/2x^7+x-1$ on $(0,1)$.

Remark The constant $\frac{1}{2}$ for the lower bound $\frac{1}{2}\mu(A)^8$ is certainly not optimal. Following the same path one can show that

$$\lim_N \frac{1}{N^2}\sum_{n,m=1}^{N} \mu(A\cap T_1^{-n}A\cap T_2^{-m}A\cap T_3^{-n-m}A) > 0$$

for all measurable set A when $\mu(A) > \beta$ where β is the root of $x^7 + x - 1$ on $(0,1)$. ☐

3. On the almost everywhere convergence of weighted averages

3.1. The averages $\frac{1}{N^2}\sum_{n,m=1}^{N} a_n b_m f(T^{n+m}x)$. Our goal is to prove first that the ergodicity assumptions are necessary in Lemma 3. We recall that we denote by $\mathcal{K}$ the σ algebra generated by the eigenfunctions of a measure preserving transformation. Even without the ergodicity assumption this σ-algebra is well defined.

PROPOSITION 7. *There exists a non ergodic measure preserving system $(Y, \mathcal{B}, \nu, S)$, a function $f \in L^\infty(\nu) \cap \mathcal{K}^\perp$ such that for ν a.e. y we can find bounded sequences a_n and b_n such that the averages*

$$\frac{1}{N^2}\sum_{n,m=1}^{N} a_n b_m f(S^{n+m}y)$$

do not converge when N tends to ∞. In other words Lemma 3 is false if we remove the ergodicity assumption.

PROOF. Let $S(x,y) = (x+\alpha, x+y)$ be the ergodic measure preserving transformation defined on the two Torus where α is an irrational number. We consider the measure preserving transformation $T = S \times S$ on $\mathbb{T}^4$ defined as

$$T(x_1, x_2, x_3, x_4) = (x_1 + \alpha, x_1 + x_2, x_3 + \alpha, x_3 + x_4).$$

The transformation T is not ergodic and the Kronecker factor (σ algebra spanned by the eigenfunctions of T) corresponds to the functions depending on the first and third coordinates x_1 and x_3. This is because the eigenfunctions of S depend on their first coordinates (see also Lemma 4.18 in [**5**] on the way in general the eigenfunctions of T are created from those of S.) Consider the function $f(x_1, x_2, x_3, x_4) = e^{-2\pi i x_2} e^{2\pi i x_4}$. This function belongs to $\mathcal{K}^\perp$. We have $f(T^{n+m}(x_1, x_2, x_3, x_4)) = e^{2\pi i (x_4 - x_2 + (n+m)(x_3 - x_1))}$. Let us assume that Lemma 3 was true without ergodicity assumption. Then we could find a set of full measure such that for a.e. x_1, x_2, x_3, x_4 in this set and for all bounded sequences a_n and b_n we would have

$$\lim_N \frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}(x_1, x_2, x_3, x_4))$$
$$= \lim_N \frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m e^{2\pi i (x_4 - x_2 + (n+m)(x_3 - x_1))} = 0.$$

To disprove this we can take a bounded sequence v_n such that the averages $\frac{1}{N}\sum_{n=0}^{N-1} v_n$ diverge. Then we can take $a_n = v_n e^{-2\pi i n (x_3 - x_1)}$, and $b_m = e^{-2\pi i m (x_3 - x_1)}$. As

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}(x_1, x_2, x_3, x_4)) = \frac{1}{N}\sum_{n=0}^{N-1} v_n e^{2\pi i (x_4 - x_2)},$$

this shows that Lemma 3 is false once we remove the ergodicity assumption. This ends the proof of Proposition 7. ☐

Remarks 1

(1) Proposition 7 shows that Lemma 3 as stated is quite sharp as one can not even expect to have the convergence of the averages $\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}x)$ as in this example they are equal to $\frac{1}{N}\sum_{n=1}^{N} v_n.e^{2\pi i(x_4-x_2)}$

(2) The same measure preserving system can be used to show that the uniform Wiener Wintner ergodic theorem is no longer valid if T is not ergodic. By this we mean that if we denote by $\mathcal{K}$ the σ algebra spanned by the eigenfunctions of T then we do not have in general for functions $f \in \mathcal{K}^{\perp}$,

$$\lim_N \sup_t \left|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)e^{2\pi int}\right| = 0.$$

(3) As indicated earlier the norm convergence holds without difficulty as the next proposition shows. We give the proof just for the sake of completeness and to show the difference between the pointwise and norm convergence.

Definition 1. *We will denote by* $\mathcal{WW}_1$ *the set of bounded sequences* $a = (a_n)$ *of scalars such that* $\lim_N \frac{1}{N}\sum_{n=1}^{N} a_n e^{2\pi int}$ *exists for each* $t \in \mathbb{R}$.

Proposition 8. *Let* T *be a unitary operator and let* $a = (a_n)$ *and* $b = (b_m)$ *be bounded sequences. Then the averages*

$$\frac{1}{N^2}\sum_{n,m=1}^{N} a_n b_m T^{n+m}$$

converge in norm if a *and* b *belong to* $\mathcal{WW}_1$

Proof. It is a simple consequence of the spectral theorem. If we denote by σ_f the spectral measure of the function f with respect to T then we have

$$\left\|\frac{1}{N^2}\sum_{n,m=1}^{N} a_n b_m T^{n+m} - \frac{1}{M^2}\sum_{n,m=1}^{M} a_n b_m T^{n+m}\right\|^2$$
$$= \int \left|\frac{1}{N^2}\sum_{n,m=1}^{N} a_n b_m e^{2\pi i(n+m)t} - \frac{1}{M^2}\sum_{n,m=1}^{M} a_n b_m e^{2\pi i(n+m)t}\right|^2 d\sigma_f(t)$$
$$= \int \left|\frac{1}{N}\sum_{n=1}^{N} a_n e^{2\pi int}\frac{1}{N}\sum_{m=1}^{N} b_m e^{2\pi imt} - \frac{1}{M}\sum_{n=1}^{M} a_n e^{2\pi int}\frac{1}{M}\sum_{m=1}^{M} b_m e^{2\pi imt}\right|^2 d\sigma_f(t)$$

which easily shows that the averages form a Cauchy sequence. □

3.2. Higher order averages. Proposition 2 shows that if the transformation T is ergodic and the function $f \in L^2$ then for μ a.e. x the averages

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}x)$$

converge for all sequences $a = (a_n), b = (b_n)$ that belong to $\mathcal{WW}_1$.

The next proposition shows that the class $\mathcal{WW}_1$ does not characterize those bounded sequences for which the similar averages for seven terms converge a.e. even under the condition of ergodicity of the transformation

PROPOSITION 9. *There exists an ergodic dynamical system* $(X, \mathcal{A}, \mu, T)$ *and a function* $f \in L^\infty(\mu)$ *such that for* μ *a.e.* x *we can find bounded sequences* $A_i = (a_{n,i}) \in \mathcal{WW}_1$ *for which the averages*

$$M_N(A_1, A_2, ...A_6, f)(x) = \frac{1}{N^3}\sum_{n,m,p=0}^{N-1} a_{1,p}a_{2,n}a_{3,p+n}a_{4,m}a_{5,n+m}a_{6,p+m}f(T^{n+m+p}x)$$

do not converge.

PROOF. We consider the sequence v_n with values 1 or -1 such that the averages $\frac{1}{N}\sum_{n=1}^N v_n$ diverge. The sequence is built from longer and longer stretches of 1 and -1 so that the averages get close to 1 then close to -1 and so on. We extend v_n to negative indices by putting $v_{-n} = v_n$. We can observe that this sequence has a correlation in the sense that for any $h \in \mathbb{Z}$ the averages $\displaystyle\frac{1}{N}\sum_{n=1}^N v_n\overline{v_{n+h}}$ converge to a scalar $\gamma(h)$. Simple considerations show that the limit for all h is equal to one. The quantity $\gamma(h)$ represents the h Fourier coefficients of a positive measure σ that is equal then to the Dirac measure at zero, δ_0, a discrete measure.

We take now an irrational number α . We claim that the sequence $a_n = v_n e^{2\pi i n^2\alpha}$ belongs to $\mathcal{WW}_1$. To see this first one can observe that the sequence $e^{2\pi i n^2\alpha}e^{2\pi i n t}$ does have a correlation; for each $h \in \mathbb{Z}$ the limit of

$$\frac{1}{N}\sum_{n=1}^N e^{2\pi i[(n^2-(n+h)^2]\alpha}e^{2\pi i(nt-(n+h)t)}$$

is equal to zero for $h \neq 0$ and one otherwise. Therefore the measure associated with these Fourier coefficients is Lebesgue measure, m. As a consequence of the Affinity principle [**4**] the measures m and δ_0 being orthogonal we have for each $t \in \mathbb{R}$

$$\lim_N \frac{1}{N}\sum_{n=1}^N v_n e^{2\pi i n^2\alpha}e^{2\pi i n t} = 0.$$

Thus we have shown that the sequence $a_n = v_n e^{2\pi i n^2\alpha}$ belongs to $\mathcal{WW}_1$.

We consider the ergodic measure preserving transformation $S(x, y) = (x + \alpha, x + y)$ defined on the two Torus where α is the irrational number used to define the sequence a_n. Our goal is to prove that for the function $f(x, y) = e^{4\pi i y}$ it is impossible to find a set of full measure off which for all six bounded sequences $A_i = (a_{i,n})$, $1 \leq i \leq 6$ the averages

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} a_{1,p}a_{2,n}a_{3,p+n}a_{4,m}a_{5,n+m}a_{6,p+m}f(S^{n+m+p}(x, y))$$

converge. To reach this conclusion we can use the simple equality

$$(n + m + p)^2 = (n + m)^2 + (n + p)^2 + (m + p)^2 - n^2 - p^2 - m^2.$$

We have $f(T^{n+m+p}(x, y)) = e^{4\pi i(y+(n+m+p)(x-\alpha/2))}.e^{2\pi i(n+m+p)^2\alpha}$. As a consequence if we take $a_{1,p} = e^{2\pi i p^2\alpha}$, $a_{2,n} = v_n e^{2\pi i n^2\alpha}$, $a_{3,p+n} = e^{-2\pi i(p+n)^2\alpha}e^{-2\pi i(p+n)(x-\alpha/2)}$,

$a_{4,m} = e^{2\pi i m^2 \alpha}$, $a_{5,n+m} = e^{-2\pi i (n+m)^2 \alpha} e^{-2\pi i (n+m)(x-\alpha/2)}$ and
$a_{6,p+m} = e^{-2\pi i (p+m)^2 \alpha} e^{-2\pi i (p+m)(x-\alpha/2)}$, then

$$\frac{1}{N^3} \sum_{n,m,p=0}^{N-1} a_{1,p} a_{2,n} a_{3,p+n} a_{4,m} a_{5,n+m} a_{6,p+m} f(S^{n+m+p}(x,y)) = \frac{1}{N} \sum_{n=0}^{N-1} v_n e^{4\pi i y}.$$

Therefore the averages

$$\frac{1}{N^3} \sum_{n,m,p=0}^{N-1} a_{1,p} a_{2,n} a_{3,p+n} a_{4,m} a_{5,n+m} a_{6,p+m} f(S^{n+m+p}(x,y))$$

do not converge. (The arguments in the previous paragraphs can also be used to show that each sequence $A_i \in \mathcal{WW}_1$.) □

References

[1] **I. Assani**:“Pointwise convergence of ergodic averages along cubes”, available at www.math.unc.edu/Faculty/assani/PCACubesAdMath3.pdf

[2] **I. Assani**:“Wiener Wintner Ergodic theorems”, World Scientific Pub Co, (2003), 228 pages,ISBN: 9810244398.

[3] **D. Berend**:“Joint ergodicity and mixing”, *J. d’Anal. Math.*, 45,(1985), 255-284.

[4] **J. Coquet, T. Kamae and M. Mendes-France**:“Sur la mesure spectrale de certaines suites arithmetiques”, *Bull. Soc. Math. France*, 105, (1977), 369-384.

[5] **H. Furstenberg**:“Recurrence in Ergodic Theory and Combinatorial Number Theory”, *Princeton University Press*, (1981).

[6] **A. Leibman**:“Lower bounds for ergodic averages”, *Ergodic Theory and Dynamical Systems*, 22 (2002), 863-872.