POINTWISE CONVERGENCE OF ERGODIC AVERAGES ALONG CUBES

I. ASSANI

ABSTRACT. Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. We prove the pointwise convergence of ergodic averages along cubes of $2^k - 1$ bounded and measurable functions for all k. We show that this result can be derived from estimates about bounded sequences of real numbers. We apply these estimates to establish the pointwise convergence of some weighted ergodic averages and ergodic averages along cubes for not necessarily commuting measure preserving transformations.

1. INTRODUCTION

Let $(X, \mathcal{B}, \mu, T)$ be a dynamical system on a finite measure space, where $T : X \to X$ is a measure preserving transformation i.e. $\mu(T^{-1}A) = \mu(A)$ for all measurable subsets of $\mathcal{B}$. We will assume that T is invertible. A factor of the system $(X, \mathcal{B}, \mu, T)$ is a sub-σ algebra invariant under T. For convenience we shall denote by the same letter a factor $\mathcal{Z}$ and the L^2 space built on this invariant sub-σ algebra.

We will assume in some of the statements that T is ergodic. This means that the only invariant functions for T are the constant functions. As we look for pointwise results we

---

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

Keywords: characteristic factors, ergodic averages along the cubes, Wiener Wintner averages.

AMS subject classification 37A05, 37A30, 47A35.

will use the ergodic decomposition to lift some results obtained for ergodic maps to general measure preserving transformations.

Theorem 1. *(A. Y. Khintchine* [12]*) For any invertible measure preserving system and any set $A \in \mathcal{B}$ and any $\varepsilon > 0$ 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 theorem 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 follows from the uniform version of von Neumann ergodic theorem as

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

V. Bergelson [5] considered a generalization of Khintchine's recurrence theorem with the expressions

$$\mu((A \cap T^n A) \cap T^m(A \cap T^n A)) = \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A).$$

These expressions can be viewed as discrete versions of continuous averages introduced by T. Gowers [10]. These averages played a key role in T. Gowers's proof of Szemeredi's theorem on the existence of arbitrary long arithmetic progressions in set of integers with positive upper density. Bergelson proved the following convergence result.

Theorem 2. *(V. Bergelson* [5]*)*

Consider L^∞ functions, f, g and h. The averages $\frac{1}{N^2} \sum_{n,m=1}^{N} f(T^n x) g(T^m x) h(T^{n+m} x)$

converge in L^2 *norm. Furthermore for any measurable set* A *with* $\mu(A) > 0$, *we have* $\lim_N \frac{1}{N^2} \sum_{n,m=1}^{N} \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A) \geq \mu(A)^4$.

The averages $\frac{1}{N^2} \sum_{n,m=1}^{N} f(T^n x) g(T^m x) h(T^{n+m} x)$ are now called the averages along the cubes for 3 terms. These averages are also called non conventional in comparison with the ergodic averages $\frac{1}{N} \sum_{n=1}^{N} f(T^n x)$. Observe that if one integrates these averages for $f = g = h = \mathbf{1}_A$ with respect to the measure $\mathbf{1}_A d\mu$ one obtains averages of the expressions $\mu(A \cap T^n A \cap T^m A \cap T^{n+m} A)$.

The averages along the cubes of seven functions are defined as

$$\frac{1}{N^3} \sum_{n,m,p=0}^{N} f_1(T^n x) f_2(T^m x) f_3(T^p x) f_4(T^{n+m} x) f_5(T^{p+n} x) f_6(T^{p+m} x) f_7(T^{n+m+p} x).$$

One can define similarly the averages of $2^k - 1$ bounded functions. We will denote them by $M_N(f_1, f_2, ..., f_{2^k-1})$.

In [11] B. Host and B. Kra proved that the averages of $2^k - 1$ bounded functions converge in L^2 norm. To achieve this result they identified increasing factors Z_k, $k = 0, 1, 2, ...$ of ergodic dynamical systems. Each of these factors is isomorphic to an inverse limit of nilsystems. We refer to [11] for the definition of these nilsystems (and for a nilsystem in general). The same factors were identified independently by a different method by T. Ziegler [17]. For ergodic dynamical systems, B. Host and B. Kra showed the following

(1) The averages of $2^k - 1$ bounded functions converge a.e. if each function belongs to the factor Z_{k-1}. The pointwise convergence on the factors can be viewed as a consequence of the following result of A. Leibman [14].

Theorem 3. *(A. Leibman [14]) Let G/Λ be a nilmanifold and let $t_1, ..., t_l$ be commuting elements of G. Then for any continuous function f on X the averages*

$$\prod_{i=1}^{k} \frac{1}{N_i - M_i} \sum_{M_1 \leq n_1 < N_1, \cdots, M_k \leq n_k < N_k} f(t_1^{n_1} ... t_k^{n_k} x)$$

converge everywhere on X when $N_1 - M_1, ..., N_k - M_k$ tend to infinity.

Indeed one can decompose any function f into the sum of its projection on the factor Z_j, $E[f|Z_j]$ and $f - E[f|Z_j]$. By Theorem 3, one can establish the pointwise convergence for nilsystems. The result for the inverse limit of nilsystems which are the factors Z_j can be obtained by approximation.

(2) For ergodic systems the averages of $2^k - 1$ functions converge in L^2 norm to zero if one of the functions is orthogonal to the factor Z_{k-1}.

We recall that a factor is characteristic for some averages if the limit behavior of these averages remains unchanged if each function is replaced by its conditional expectation on this factor. When the limit is studied with respect to the L^2 norm then we will speak of a factor characteristic for the L^2 norm. When the limit is studied with respect to the pointwise convergence (mainly the case in this paper) we will speak of a factor characteristic for the pointwise convergence. The notion of characteristic factor is due to H. Furstenberg and can be found explicitly stated in [9]. One consequence of Host and Kra results is that for each k the factor Z_{k-1} is characteristic for the L^2 norm of the averages of $2^k - 1$ functions. Actually they show that these factors are characteristic for the L^2 norm of the averages where one sums from M to N and take the limit when $(M - N)$ tends to ∞. Thus in the case of seven functions they consider the averages

$$\frac{1}{(N-M)^3}\sum_{n,m,p=M}^{N} f_1(T^n x) f_2(T^m x) f_3(T^p x) f_4(T^{n+m} x) f_5(T^{p+n} x) f_6(T^{p+m} x) f_7(T^{n+m+p} x).$$

Let us note that Z_1 is the Kronecker factor and Z_2 the Conze-Lesigne factor. When the system is ergodic the limit in L^2 norm of the averages along the cubes has been identified in [11] (see theorem 13.1). From this they derived (theorem 1.3) the following inequality for the averages of seven functions

$$\lim_N \frac{1}{N^3}\sum_{n,m,p=0}^{N-1} \mu\big(A \cap T^n A \cap T^m A \cap T^p A \cap T^{n+m} A \cap T^{p+n} A \cap T^{n+m+p} A\big) \geq \mu(A)^8.$$

A combinatorial interpretation for subsets of $\mathbb{Z}$ with positive upper density is given in the same paper with theorem 1.5.

Our present paper answers the question raised by Host and Kra [11] about the pointwise convergence of such averages. A second motivation comes from the fact that very little is known on the pointwise convergence of nonconventional ergodic averages for general measure preserving systems ([6]). The paper is divided into two parts. In the first part we focus on the averages along the cubes and prove the following result.

Theorem 4. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system. Then for each positive integer k the averages along the cubes of $2^k - 1$ functions converge almost everywhere. If the system $(X, \mathcal{B}, \mu, T)$ is ergodic then for each $k \geq 1$ the factor Z_{k-1} is characteristic for the pointwise convergence of the averages along the cubes of $2^k - 1$ bounded and measurable functions.*

The pointwise limit of the averages is of course the same as the L^2 limit identified in [11]. Therefore the second statement of the previous theorem is a consequence of the pointwise convergence of these averages when the system is ergodic and the fact that the averages are characteristic in norm.

More general averages along the cubes were considered and proved to converge in norm in [5] and [11]. Not all of them converge almost everywhere and so for them the factors Z_{k-1} are characteristic in norm but not pointwise. To see it one can take for instance averages of the form $\frac{1}{N-M}\sum_{n=M}^{N} f(T^n x)$ that do not converge a.e. for some function $f \in L^\infty$. There exist integers N and M such that this function f can be found in any ergodic system. In particular for weakly mixing system such function f always exist. We recall that for an ergodic dynamical system $(X, \mathcal{B}, \mu, T)$ the Kronecker factor is the invariant $\sigma-$ algebra spanned by the eigenfunctions of T. It is the factor Z_1. For the particular case of weakly mixing systems the Kronecker factor is trivial (and coincides with the factors Z_k for $k \geq 2$). Therefore for these averages we always have a.e. convergence for functions in the Kronecker factor. But as the averages of the function f do not converge a.e. the limit behavior of the averages is not the same as the one observed on the Kronecker factor. Therefore the Kronecker factor is not pointwise characteristic for these averages. However the Kronecker factor is characteristic for the norm convergence. Thus the notion of being characteristic in norm is not equivalent to being pointwise characteristic.

In the second part we give applications of the method we used in the first part. With key estimates on bounded sequences of scalars we derive pointwise convergence results for weighted ergodic averages and for averages along cubes for *not necessarily commuting* measure preserving transformations that extend part of the results obtained in the first part. In the case of not necessarily commuting transformations we prove the following result.

Theorem 5. *Let $(X, \mathcal{B}, \mu)$ be a probability measure space and T_1, T_2 , T_3 three not necessarily commuting measure preserving transformations on $(X, \mathcal{B}, \mu)$. Then for all bounded functions f_i, $1 \leq i \leq 3$ the averages*

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

converge a.e. and in L^2 norm.

Some of the results presented in this particular section were announced in [1]. In [3] we extended Theorem 5 to the averages of six not commuting measure preserving transformations. Our method highlights the fact that for convergence results averages along cubes rely more on the underlying arithmetic structure of the sequence $f(T^n x)$ than its dynamical structure. We wrote this section in such a way that readers interested mainly in those results would not need to read the previous sections. We hope this makes the paper more readable.

Remarks and Acknowledgements. The current paper is a slightly edited/updated version of a 2004 paper that was posted at that time on our web site. Pointwise results for averaging processes for powers of the same transformation were afterwards obtained by C. Demeter, T. Tao and C. Thiele [7].

We thank a previous referee for his/her comments and for suggesting the study of the weighted averages in section 3. We thank also the current referee of this paper for his/her comments and for drawing our attention to the generalized Kronecker factor.

2. POINTWISE CONVERGENCE OF THE AVERAGES ALONG CUBES FOR A SINGLE TRANSFORMATION

In the subsequent inequalities the constant C may change from one line to the other. It will depend only at time on the L^∞ norm of the functions f_j. We will first prove the almost everywhere convergence of the averages of three then seven functions. This will explain the first induction step in our proof and our method.

2.1. Pointwise convergence for the averages of three functions.

We start by proving the pointwise convergence of the averages for three functions.

$$M_N(f_1, f_2, f_3)(x) = \frac{1}{N^2} \sum_{n,m=0}^{N-1} f_1(T^n x) f_2(T^m x) f_3(T^{n+m} x)$$

for f_i bounded and measurable functions.

We recall Bourgain's uniform Wiener Wintner ergodic result announced in [6].

Lemma 1. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system and f a function in the ortho-complement of the Kronecker factor. Then for a.e. x we have $\lim_N \sup_t |\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) e^{2\pi i n t}| = 0$.*

Using this lemma we can prove the following proposition. It is inspired from the computations made on the third page of [6].

Proposition 6. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and f_i, $1 \leq i \leq 3$ three bounded functions then the averages*

$$M_N(f_1, f_2, f_3)(x) = \frac{1}{N^2} \sum_{n,m=0}^{N-1} f_1(T^n x) f_2(T^m x) f_3(T^{n+m} x)$$

converge a.e.

Proof. It is enough to show this pointwise convergence result for ergodic measure preserving systems. The general result will follow by using the ergodic decomposition. We have the following inequalities.

$$
\begin{aligned}
&|M_N(f_1,f_2,f_3)(x)|^2\\
&\le \|f_1\|_\infty^2\Big(\frac{1}{N}\sum_{n=0}^{N-1}\big|\frac{1}{N}\sum_{m=0}^{N-1} f_2(T^m x)f_3(T^{n+m}x)\big|^2\Big), \text{by Cauchy-Schwarz' inequality,}\\
&\le \|f_1\|_\infty^2\frac{1}{N}\sum_{n=0}^{N-1}\Big|\int\big(\sum_{m=0}^{N-1} f_2(T^m x)e^{-2\pi imt}\big)\big(\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\big).e^{-2\pi int}dt\Big|^2\\
&\le \|f_1\|_\infty^2\frac{1}{N}\int\Big|\sum_{m=0}^{N-1} f_2(T^m x)e^{-2\pi imt}\Big|^2\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\Big|^2 dt, \text{by Parseval's inequality,}\\
&\le \frac{C}{N}\sup_t\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\Big|^2\int\big|\sum_{m=0}^{N-1} f_2(T^m x)e^{-2\pi imt}\big|^2 dt\\
&\le C\sup_t\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\Big|^2\frac{1}{N}N\|f_2\|_\infty^2.
\end{aligned}
$$

With the help of lemma 1 we can conclude that for f_3 in the orthocomplement of the Kronecker factor the averages $M_N(f_1,f_2,f_3)(x)$ converge a.e. to zero.

If f_3 is one of the eigenfunctions for T with eigenvalue $e^{2\pi i\theta}$ then

$$M_N(f_1,f_2,f_3)(x)=f_3(x)\big(\frac{1}{N}\sum_{n=0}^{N-1} f_1(T^n x)e^{2\pi in\theta}\big)\big(\frac{1}{N}\sum_{m=0}^{N-1} f_2(T^m x)e^{2\pi im\theta}\big).$$

The convergence in this case follows from Birkhoff 's theorem applied to the product of T and the rotation θ. The pointwise convergence for a finite linear combination of eigenfunctions in the Kronecker factor follows now by linearity. To establish the same result for a general

L^∞ function f_3 in the same factor, one can use the following inequalities.

$$0 \le A(g)(x) = \limsup_N M_N(f_1, f_2, g)(x) - \liminf_N M_N(f_1, f_2, g)(x)$$

$$\le 2 \sup_N |M_N(f_1, f_2, g)(x)| \le \|f_1\|_\infty \|f_2\|_\infty M^*[M^*[|g|]].$$

We denoted by $M^*(|g|) = \sup_N \frac{1}{N} \sum_{n=1}^{N} |g|(T^n x)$ the maximal ergodic function associated with the Cesaro averages of T. The maximal inequality in L^2 tells us that for all function $g \in L^2$ we have

$$\|M^*(|g|)\|_2 \le 2\|g\|_2.$$

Applying this maximal inequality twice we conclude that

$$\|A(g)\|_2 \le 4\|g\|_2.$$

Now consider a function f_3 in the Kronecker factor. There exists a sequence g_i of finite linear combination of eigenfunctions that converge in L^2 norm to f_3. As the averages $M_N(f_1, f_2, g_i)(x)$ converge a.e. we have by linearity for a.e. x

$$A(f_3)(x) = \limsup_N M_N(f_1, f_2, f_3 - g_i)(x) - \liminf_N M_N(f_1, f_2, f_3 - g_i)(x).$$

Hence we have

$$\|A(f_3)\|_2 \le 4\|f_3 - g_i\|_2$$

for each i. Taking the limit when i tends to infinity we obtain $\|A(f_3)\|_2 = 0$ which implies that $A(f_3)(x) = 0$ a.e., and the sequence $M_N(f_1, f_2, f_3)(x)$ converge a.e. $\square$

Remarks 1

- The proof of proposition 6 shows that if f_1 and f_2 are bounded functions and $P_{\mathcal{K}}$ denotes the projection onto the Kronecker factor of T then

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1}\Big|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{m+n}x)\Big|^2 = \limsup_N \frac{1}{N}\sum_{n=0}^{N-1}\Big|\frac{1}{N}\sum_{m=0}^{N-1} P_{\mathcal{K}}(f_1)(T^m x)P_{\mathcal{K}}(f_2)(T^{m+n}x)\Big|^2. \tag{1}$$

- The proof of this proposition actually shows that

$$\left(\frac{1}{N}\sum_{n=0}^{N-1}\Big|\frac{1}{N}\sum_{m=0}^{N-1} f_2(T^m x)f_3(T^{n+m}x)\Big|^2\right) \le C\sup_t\left|\frac{1}{N}\sum_{m'=0}^{N-1} f_3(T^{m'}x)e^{2\pi i m' t}\right|^2 \|f_2\|_\infty^2. \tag{2}$$

  A similar estimate can be obtained with $\sup_t\left|\frac{1}{N}\sum_{m'=0}^{N-1} f_2(T^{m'}x)e^{2\pi i m' t}\right|^2$ if we focus instead on the function f_2.

2.2. **Pointwise convergence for the averages of seven functions.** In [11] it is shown that Z_2, the Conze-Lesigne factor, is characteristic for the convergence in L^2 norm of the averages of seven functions. Functions in this factor are characterized by the seminorm $|||.|||_3$ such that

$$|||f|||_3^8 = \lim_H \frac{1}{H}\sum_{h=0}^{H-1} |||f.f\circ T^h|||_2^4 \tag{3}$$

where

$$|||f|||_2^4 = \lim_H \frac{1}{H}\sum_{h=0}^{H-1} \Big|\int f.f(T^h)d\mu\Big|^2. \tag{4}$$

A function $f \in {Z_2}^\perp$ if and only $|||f|||_3 = 0$. More generally they showed that for each positive integer k we have

$$|||f|||_{k+1}^{2^{k+1}} = \lim_H \frac{1}{H}\sum_{h=0}^{H-1} |||f.f\circ T^h|||_k^{2^k},$$

with the condition that $f \in Z_{k-1}$ if and only if $|||f|||_k = 0$.

Lemma 2. *Let* $(X, \mathcal{B}, \mu, T)$ *be an ergodic dynamical system and* $f \in L^\infty(\mu)$ *then for all* H *positive integer we have*

$$\limsup_N \sup_t \big|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)e^{2\pi int}\big|^2 \leq C\bigg(\frac{1}{H} + \frac{1}{H}\sum_{h=1}^{H}\big|\int f.\overline{f \circ T^h}d\mu\big|\bigg).$$

In particular we have

$$\limsup_N \sup_t \big|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)e^{2\pi int}\big|^2 \leq C|||f|||_2^2. \tag{5}$$

Proof. Without loss of generality we can assume that the function f takes only real values. We apply van der Corput's inequality ([13]) . Because of this inequality for $H < N$ we get

$$\sup_t \big|\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)e^{2\pi int}\big|^2 \leq C\bigg(\frac{1}{H} + \frac{1}{H}\sum_{h=1}^{H}\bigg|\frac{1}{N}\sum_{n=0}^{N-h} f(T^n x)f(T^{n+h}x)\bigg|\bigg).$$

(The factor $\dfrac{N+H}{N}$ that normally appears on the right side of van der Corput's inequality being less than 2 has been "swallowed" in the constant C.) Birkhoff's pointwise ergodic theorem allows us to obtain the first part of the lemma. For the second part we can use Cauchy Schwarz inequality to write that

$$\frac{1}{H}\sum_{h=1}^{H}\big|\int f.f \circ T^h d\mu\big| \leq \bigg(\frac{1}{H}\sum_{h=1}^{H}\big|\int f.f \circ T^h d\mu\big|^2\bigg)^{1/2}.$$

Now using the definition of $|||f|||_2$, (see (4)), we can end the proof of this lemma. ☐

The lemma that replaces the uniform Wiener Wintner ergodic theorem in the case of the averages of seven functions is the following.

Lemma 3. *If $(X, \mathcal{B}, \mu, T)$ is an ergodic dynamical system and f_1 or f_2 is in $Z_2^{\perp}$ then for a.e. x*

$$\lim_N \frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{n+m}x)e^{2\pi i m t}\right|^2 = 0. \tag{6}$$

Proof. We can assume without loss of generality that the functions are uniformly bounded by one. We use again van der Corput's inequality, [13]. For $(H+1)^2 < N$ we get

$$\begin{aligned}&\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{n+m}x)e^{2\pi i m t}\right|^2\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\left|\frac{1}{N}\sum_{m=0}^{N-h-1} f_1(T^m x) f_2(T^{m+n}x)\overline{f_1(T^{m+h}x) f_2(T^{m+n+h}x)}\right|.\end{aligned}$$

So recalling that the constant C may change from one line to another but remains an absolute constant we have,

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{n+m}x)e^{2\pi i m t}\right|^2\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-h-1} f_1(T^m x) f_2(T^{m+n}x)\overline{f_1(T^{m+h}x) f_2(T^{m+n+h}x)}\right|\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{m+n}x)\overline{f_1(T^{m+h}x) f_2(T^{m+n+h}x)}\right.\\ &\left.-\sum_{m=N-h}^{N-1} f_1(T^m x) f_2(T^{m+n}x)\overline{f_1(T^{m+h}x) f_2(T^{m+n+h}x)}\right|\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{m+n}x)\overline{f_1(T^{m+h}x) f_2(T^{m+n+h}x)}\right| + \frac{C}{H}\sum_{h=1}^{H}\frac{1}{N}\sum_{n=0}^{N-1}\frac{h}{N}\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{m+n}x)\overline{f_1(T^{m+h}x) f_2(T^{m+n+h}x)}\right|.\end{aligned}$$

Thus using the inequality (or Cauchy Schwarz's inequality)

$$|\frac{1}{P}\sum_{p=1}^{P} u_p| \leq (\frac{1}{P}\sum_{p=1}^{P}|u_p|^2)^{1/2} \tag{7}$$

we obtain

$$\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{n+m}x)e^{2\pi imt}\right|^2$$
$$\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}\left(\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{m+n}x)\overline{f_1(T^{m+h}x)f_2(T^{m+n+h}x)}\right|^2\right)\right)^{1/2}.$$

Finally by applying the inequality (2) made after the Remarks 1 to the function $f_1.\overline{f_1 \circ T^h}$ we get

$$\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{n+m}x)e^{2\pi imt}\right|^2$$
$$\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}\left(\sup_t\left|\frac{1}{N}\sum_{m'=0}^{N-1}(f_1.\overline{f_1\circ T^h})(T^{m'}x)e^{2\pi im't}\right|^2\right)\right)^{1/2}.$$

Now by using Lemma 2 and the inequality $\frac{1}{H}\sum_{h=1}^{H}|u_h|^2 \leq (\frac{1}{H}\sum_{h=1}^{H}|u_h|^4)^{1/2}$ we obtain

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{n+m}x)e^{2\pi imt}\right|^2$$
$$\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}\limsup_N\sup_t\left|\frac{1}{N}\sum_{m'=0}^{N-1}(f_1.\overline{f_1\circ T^h})(T^{m'}x)e^{2\pi im't}\right|^2\right)^{1/2}$$
$$\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}|||f_1.\overline{f_1\circ T^h}|||_2^2\right)^{1/2}$$
$$\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}|||f_1.\overline{f_1\circ T^h}|||_2^2\right)^{1/2}$$
$$\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}|||f_1.\overline{f_1\circ T^h}|||_2^4\right)^{1/4}.$$

Taking now the limit when H tends to ∞ we get the following estimate

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{n+m}x)e^{2\pi imt}\right|^2 \leq C|||f_1|||_3^2. \tag{8}$$

Thus if we assume that $f_1 \in Z_2^{\perp}$ then $|||f_1|||_3 = 0$ and we obtain the equation (7). We have the same conclusion if one assumes that $f_2 \in Z_2^{\perp}$. Indeed using again the inequality (2) made after Remarks 1 but this time to the function $f_2.\overline{f_2 \circ T^h}$ we would obtain

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{n+m}x)e^{2\pi imt}\right|^2\\ &\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}\left(\sup_t\left|\frac{1}{N}\sum_{m'=0}^{N-1}(f_2.\overline{f_2\circ T^h})(T^{m'}x)e^{2\pi im't}\right|^2\right)^{1/2}.\end{aligned}$$

The same computations made above would led us to the estimate

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} f_1(T^m x)f_2(T^{n+m}x)e^{2\pi imt}\right|^2 \leq C|||f_2|||_3^2.$$

□

Using Lemma 3 we can now give a proof of the almost convergence of the averages of seven functions.

Proposition 7. *The averages along the cubes of seven functions converge almost surely.*

Proof. First by using the ergodic decomposition we can assume that the system $(X, \mathcal{B}, \mu, T)$ is ergodic.

$$|M_N(f_1, f_2, ..., f_7)(x)|^2$$

$$= \Big| \frac{1}{N^3} \sum_{p=0}^{N-1} f_1(T^p x) \sum_{n=0}^{N-1} f_2(T^n x) f_3(T^{p+n} x) \big(\sum_{m=0}^{N-1} f_4(T^m x) f_5(T^{n+m} x) f_6(T^{p+m} x) f_7(T^{n+m+p} x) \big) \Big|^2$$

$$\leq \frac{1}{N^2} \sum_{p=0}^{N-1} \sum_{n=0}^{N-1} \|f_1\|_\infty^2 \|f_2\|_\infty^2 \|f_3\|_\infty^2 \Big| \frac{1}{N} \sum_{m=0}^{N-1} f_4(T^m x) f_5(T^{n+m} x) f_6(T^{p+m} x) f_7(T^{p+n+m} x) \Big|^2$$

$$= \frac{1}{N^2} \prod_{i=1}^{3} \|f_i\|_\infty^2 .$$

$$\sum_{n=0}^{N-1} \sum_{p=0}^{N-1} \Big| \int \big(\sum_{m=0}^{(N-1)} f_4(T^m x) f_5(T^{n+m} x) e^{-2\pi i m t} \big) \big(\frac{1}{N} \sum_{m'=0}^{2(N-1)} f_6(T^{m'} x) f_7(T^{n+m'} x) e^{2\pi i m' t} \big) . e^{-2\pi i p t} dt \Big|^2$$

$$\leq \frac{1}{N^2} \prod_{i=1}^{3} \|f_i\|_\infty^2 \sum_{n=0}^{N-1} \int \Big| \sum_{m=0}^{N-1} f_4(T^m x) f_5(T^{n+m} x) e^{-2\pi i m t} \big) \big(\frac{1}{N} \sum_{m'=0}^{2(N-1)} f_6(T^{m'} x) f_7(T^{n+m'} x) e^{2\pi i m' t} \big) \Big|^2 dt$$

$$\leq \frac{C}{N^2} \prod_{i=1}^{3} \|f_i\|_\infty^2 \sum_{n=0}^{N-1} \sup_t \Big| \frac{1}{N} \sum_{m'=0}^{2(N-1)} f_6(T^{m'} x) f_7(T^{n+m'} x) e^{2\pi i m' t} \Big|^2 N \prod_{j=4}^{5} \|f_j\|_\infty^2$$

$$= C \prod_{i=1}^{5} \|f_i\|_\infty^2 \frac{1}{N} \sum_{n=0}^{N-1} \sup_t \Big| \frac{1}{N} \sum_{m'=0}^{2(N-1)} f_6(T^{m'} x) f_7(T^{n+m'} x) e^{2\pi i m' t} \Big|^2 .$$

With the help of the lemma 3 one can conclude that if f_6 or f_7 belong to $Z_2^{\perp}$ then the averages of these seven functions converge to zero. By using a very similar argument one can see that the averages will converge to zero if one of the functions $f_i \in Z_2^{\perp}$, $1 \leq i \leq 7$.

It remains then to establish the pointwise convergence when all functions are measurable with respect to Z_2. As we observed in the introduction the pointwise convergence is a consequence of Leibman's result and the nature of the factor Z_2 being an inverse limit of

2-step nilsystems. Theorem 3 [14] gives us the pointwise convergence for nilsystems. The result for inverse limit of nilsystems follows by approximation. □

Remarks 2

- The last steps of the proof of proposition 6 show that for bounded functions f_i, $4 \leq i \leq 7$ if we denote by $P_{Z_2}(f_i)$ their projection onto the Conze-Lesigne factor then we have

$$\begin{aligned}
&\limsup_N \frac{1}{N^2} \sum_{n,p=0}^{N-1} \left| \frac{1}{N} \sum_{m=0}^{N-1} f_4(T^m x) f_5(T^{n+m} x) f_6(T^{p+m} x) f_7(T^{p+n+m} x) \right|^2 \\
&= \limsup_N \frac{1}{N^2} \sum_{n,p=0}^{N-1} \left| \frac{1}{N} \sum_{m=0}^{N-1} P_{Z_2}(f_4)(T^m x) P_{CL}(f_5)(T^{n+m} x) P_{Z_2}(f_6)(T^{p+m} x) P_{Z_2}(f_7)(T^{p+n+m} x) \right|^2 .
\end{aligned} \tag{9}$$

- The proof of lemma 3 gives the following estimate

$$\limsup_N \frac{1}{N} \sum_{n=0}^{N-1} \sup_t \left| \frac{1}{N} \sum_{m=0}^{N-1} f_1(T^m x) f_2(T^{n+m} x) e^{2\pi i m t} \right|^2 \leq C \min[|||f_1|||_3^2, |||f_2|||_3^2]. \tag{10}$$

2.3. Proof of Theorem 4.

We will prove Theorem 4 by induction on k. In the previous sections we proved that the averages of three then seven functions converge a.e. We showed that the Z_2 factor was characteristic for the pointwise convergence of averages of seven functions while the Kronecker factor is characteristic for the pointwise convergence of averages of three functions . This established the first steps of the induction process. Even if for the induction purposes we only needed to establish the case of three functions we proved these two cases to show how one could move from one step to another.

We will use the same notation and some of the remarks made in these previous sections.

- For each $k \geq 4$ we denote by $M_N(f_1, f_2, ..., f_{2^k-1})$ the averages of $2^k - 1$ bounded functions. Without loss of generality we assume that the functions are bounded by 1 in absolute value.
- The functions f_j are listed in such a way that those depending on the index i_k are indexed by those j , $2^{k-1} \leq j \leq 2^k - 1$. The product of these terms depending on i_k is denoted by $S_{N,(i_1,i_2,...,i_k)}(f_{2^{k-1}}, ..., f_{2^k-1})(x)$. Each term $S_{N,(i_1,i_2,...,i_k)}(f_{2^{k-1}}, ..., f_{2^k-1})(x)$ is the product of two groups of 2^{k-2} functions denoted by

  $$A_{N,(i_1,i_2,...,i_k)}(f_{2^{k-1}}, f_{2^{k-1}+1}, ..., f_{3.2^{k-2}})(x)$$

  and

  $$B_{N,(i_1,i_2,...,i_k)}(f_{3.2^{k-2}+1}, ..., f_{2^k-1})(x)$$

  where the powers of T associated with each function in the second group are those appearing in the first group shifted by the index i_1. We have

$$B_{N,(i_1,i_2,...,i_k)}(f_{3.2^{k-2}+1}, ..., f_{2^k-1})(x) = A_{N,(i_1,i_2,...,i_k)}(f_{3.2^{k-2}+1}, ..., f_{2^k-1})(T^{i_1}x).$$

- We have also the inequality

$$\begin{aligned} &|M_N(f_1, f_2, ..., f_{2^k-1})(x)|^2 \\ &\leq \prod_{j=1}^{2^{k-1}-1} \|f_j\|_\infty^2 \frac{1}{N^{k-1}} \sum_{i_1,...,i_{k-1}=0}^{N-1} \Big|\frac{1}{N}\sum_{i_k=0}^{N-1} S_{N,(i_1,i_2,...,i_k)}(f_{2^{k-1}}, ..., f_{2^k-1})(x)\Big|^2. \end{aligned} \tag{11}$$

Induction Assumption

We make the following assumption (for $k - 1$).

For all bounded functions g_j, $3.2^{k-2}+1 \leq j \leq 2^k-1$ we have

$$\limsup_N \frac{1}{N^{k-2}} \sum_{i_1,\dots,i_{k-2}=0}^{N-1} \left| \frac{1}{N} \sum_{i_k=0}^{N-1} A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(g_{3.2^{k-2}+1}\cdot,\dots,g_{2^k-1})(x) \right|^2 \tag{12}$$
$$\leq C. \min_{\{3.2^{k-2}+1\leq j\leq 2^k-1\}} |||g_j|||_{k-1}^2.$$

As indicated above this assumption is shown to be true for $k=3,4$. We want to show that it also holds for k. To this end we have the following extension of lemmas 2 and 3.

Lemma 4. *Let $(X,\mathcal{B},\mu,T)$ be an ergodic dynamical system. If one of the 2^{k-2} functions f_j, $3.2^{k-2}+1\leq j\leq 2^k-1$ is in $Z_{k-1}^{\perp}$ then*

$$\lim_N \frac{1}{N^{k-2}} \sum_{i_1,\dots,i_{k-2}=0}^{N-1} \sup_t \left| \frac{1}{N} \sum_{i_k=0}^{N-1} A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1},\dots,f_{2^k-1})(x) e^{2\pi i i_k t} \right|^2 = 0 \tag{13}$$

Proof. With van der Corput inequality applied to each term

$$\sup_t \left| \frac{1}{N} \sum_{i_k=0}^{N-1} A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1},\dots,f_{2^k-1})(x) e^{2\pi i i_k t} \right|^2,$$

we have then for each $(H+1)^2 << N$

$$\frac{1}{N^{k-2}}\sum_{i_1,\dots,i_{k-2}=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{i_k=0}^{N-1}A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1},\dots,f_{2^k-1})(x)e^{2\pi i i_k t}\right|^2$$

$$\leq C.\Bigg(\frac{1}{H}+\frac{1}{H}\sum_{h=1}^{H}\frac{1}{N^{k-2}}\sum_{i_1,\dots,i_{k-2}=0}^{N-1}$$

$$\left|\frac{1}{N}\sum_{i_k=0}^{N-h-1}A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1}.f_{3.2^{k-2}+1}\circ T^h,\dots,f_{2^k-1}.f_{2^k-1}\circ T^h)(x)\right|\Bigg)$$

$$\leq C.\Bigg(\frac{1}{H}+\frac{1}{H}\sum_{h=1}^{H}\frac{1}{N^{k-2}}\sum_{i_1,\dots,i_{k-2}=0}^{N-1}$$

$$\left|\frac{1}{N}\sum_{i_k=0}^{N-1}A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1}.f_{3.2^{k-2}+1}\circ T^h,\dots,f_{2^k-1}.f_{2^k-1}\circ T^h)(x)\right|\Bigg)$$

$$\leq C.\Bigg(\frac{1}{H}+\Bigg(\frac{1}{H}\sum_{h=1}^{H}\frac{1}{N^{k-2}}\sum_{i_1,\dots,i_{k-2}=0}^{N-1}$$

$$\left|\frac{1}{N}\sum_{i_k=0}^{N-1}A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1}.f_{3.2^{k-2}+1}\circ T^h,\dots,f_{2^k-1}.f_{2^k-1}\circ T^h)(x)\right|^2\Bigg)^{1/2}\Bigg).$$

So by the induction assumption we have

$$\limsup_N\frac{1}{N^{k-2}}\sum_{i_1,\dots,i_{k-2}=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{i_k=0}^{N-1}A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1},\dots,f_{2^k-1})(x)e^{2\pi i i_k t}\right|^2$$

$$\leq C.\Bigg(\frac{1}{H}+\Bigg(\frac{1}{H}\sum_{h=1}^{H}\limsup_N\frac{1}{N^{k-2}}\sum_{i_1,\dots,i_{k-2}=0}^{N-1}$$

$$\left|\frac{1}{N}\sum_{i_k=1}^{N-1}A_{N,(i_1,i_2,\dots,i_{k-2},i_k)}(f_{3.2^{k-2}+1}.f_{3.2^{k-2}+1}\circ T^h,\dots,f_{2^k-1}.f_{2^k-1}\circ T^h)(x)\right|^2\Bigg)^{1/2}\Bigg)$$

$$\leq C.\Bigg(\frac{1}{H}+\Bigg(\frac{1}{H}\sum_{h=1}^{H}\min_{\{3.2^{k-2}+1\leq j\leq 2^k-1\}}|||f_jf_j\circ T^h|||_{k-1}^2\Bigg)^{1/2}\Bigg).$$

By using the monotonicity in α of the fractions $\Big(\frac{1}{H}\sum_{h=1}^{H}|u_h|^{\alpha}\Big)^{1/\alpha}$, we have

$$\limsup_N \frac{1}{N^{k-2}}\sum_{i_1,\ldots,i_{k-2}=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{i_k=0}^{N-1}A_{N,(i_1,i_2,\ldots,i_{k-2},i_k)}(f_{3.2^{k-2}+1},\ldots,f_{2^k-1})(x)e^{2\pi i i_k t}\right|^2$$
$$\leq C.\left(\frac{1}{H}+\left(\frac{1}{H}\sum_{h=1}^{H}\min_{\{3.2^{k-2}+1\leq j\leq 2^k-1\}}|||f_j f_j\circ T^h|||_{k-1}^{2^{k-1}}\right)^{1/2^{k-1}}\right).$$

By taking now the $\limsup_H$ of the last term we get

$$\begin{aligned}&\limsup_N \frac{1}{N^{k-2}}\sum_{i_1,\ldots,i_{k-2}=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{i_k=0}^{N-1}A_{N,(i_1,i_2,\ldots,i_{k-2},i_k)}(f_{3.2^{k-2}+1},\ldots,f_{2^k-1})(x)e^{2\pi i i_k t}\right|^2\\ &\leq C.\min_{\{3.2^{k-2}+1\leq j\leq 2^k-1\}}|||f_j|||_k^2\end{aligned}\tag{14}$$

Thus if one of the functions f_j belongs to $Z_{k-1}^{\perp}$ then the limit in the equation (13) is equal to zero. □

End of the proof of Theorem 4.

We just need to finish the induction process by proving the induction assumption for k . We consider the averages of 2^k-1 functions f_j, $M_N(f_1,f_2,\ldots,f_{2^k-1})(x)$. With the inequality (11) we have

$$|M_N(f_1,f_2,\ldots,f_{2^k-1})(x)|^2$$
$$\leq \prod_{j=1}^{2^{k-1}-1}\|f_j\|_\infty^2\frac{1}{N^{k-1}}\sum_{i_1,\ldots,i_{k-1}=0}^{N-1}\Big|\frac{1}{N}\sum_{i_k=0}^{N-1}S_{N,(i_1,i_2,\ldots,i_k)}(f_{2^{k-1}},\ldots,f_{2^k-1})(x)\Big|^2.$$

By using the same method used to derive (2) and (10) we get

$$\prod_{j=1}^{2^{k-1}-1} \|f_j\|_\infty^2 \frac{1}{N^{k-1}} \sum_{i_1,\ldots,i_{k-1}=0}^{N-1} \Big|\frac{1}{N}\sum_{i_k=0}^{N-1} S_{N,(i_1,i_2,\ldots,i_k)}(f_{2^{k-1}},\ldots,f_{2^k-1})(x)\Big|^2$$

$$\leq C\frac{1}{N^{k-2}} \sum_{i_2,\ldots,i_{k-1}=0}^{N-1} \sup_t \Big|\frac{1}{N}\sum_{i'_k=0}^{N-1} A_{N,(i_2,\ldots,i_{k-1},i'_k)}(f_{3.2^{k-2}+1},\ldots,f_{2^k-1})(x)e^{2\pi i'_k t}\Big|^2.$$

By using lemma 4 and (12) one concludes that

$$\limsup_N \frac{1}{N^{k-1}} \sum_{i_1,\ldots,i_{k-1}=0}^{N-1} \Big|\frac{1}{N}\sum_{i_k=0}^{N-1} S_{N,(i_1,i_2,\ldots,i_k)}(f_{2^{k-1}},\ldots,f_{2^k-1})(x)\Big|^2$$

$$\leq C\frac{1}{N^{k-2}} \sum_{i_2,\ldots,i_{k-1}=0}^{N-1} \sup_t \Big|\frac{1}{N}\sum_{i'_k=0}^{N-1} A_{N,(i_2,\ldots,i_{k-1},i'_k)}(f_{3.2^{k-2}+1},\ldots,f_{2^k-1})(x)e^{2\pi i'_k t}\Big|^2$$

$$\leq C. \min_{\{3.2^{k-2}+1\leq j\leq 2^k-1\}} |||f_j|||_k^2.$$

By a similar argument for the indices $i_1, i_2, \ldots, i_k$ one obtains the following inequality for the 2^{k-1} functions f_j

$$\limsup_N \frac{1}{N^{k-1}} \sum_{i_1,\ldots,i_{k-1}=0}^{N-1} \Big|\frac{1}{N}\sum_{i_k=0}^{N-1} S_{N,(i_1,i_2,\ldots,i_k)}(f_{2^{k-1}},\ldots,f_{2^k-1})(x)\Big|^2$$

$$\leq C \min_{\{2^{k-1}\leq j\leq 2^k-1\}} |||f_j|||_k^2$$

By applying this last inequality to any set of 2^{k-1} functions functions g_j that we can label from $3.2^{k-1}+1$ to $2^{k+1}-1$ instead of 1 to 2^k-1 we obtain our induction assumption for k. Thus the averages $M_N(f_1, f_2, \ldots, f_{2^k-1})(x)$ converge a.e. to zero if one of the functions $f_j \in Z_{k-1}^{\perp}$ (using very similar arguments). We can finish the proof of Theorem 4 in a similar way as we did in the case of seven functions. One projects each function onto the factor Z_{k-1} using the conditional expectation associated with this factor. The factor being an inverse limit of nilsystems one can use Leibman's result [14] which gives us the pointwise convergence for nilsystems and extend it by approximation to the the factor Z_{k-1}.

3. WEIGHTED AVERAGES AND ERGODIC AVERAGES ALONG CUBES FOR NOT NECESSARILY COMMUTING TRANSFORMATIONS.

In this section we are applying the method we used in the previous section to some weighted averages and ergodic averages along cubes for several transformations. We establish first key estimates on bounded sequences of scalars. We use these estimates to derive pointwise convergence results that are in most cases stronger than those stated in the previous section.

Lemma 5. *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*

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

Proof. We denote by $M_N(a,b,c)$ the quantity $\dfrac{1}{N^2}\sum\limits_{n,m=0}^{N-1} a_n.b_m.c_{n+m}$. The steps are similar to those given in the proof of Proposition 6 so we only sketch them. We have

$$\begin{aligned}
&|M_N(a,b,c)|^2\\
&\leq \|a\|_\infty^2\left(\frac{1}{N}\sum_{n=0}^{N-1}\Big|\frac{1}{N}\sum_{m=0}^{N-1} b_m c_{n+m}\Big|^2\right) \text{ by Cauchy-Schwarz's inequality}\\
&\leq \|a\|_\infty^2\frac{1}{N}\sum_{n=0}^{N-1}\left|\int\Big(\sum_{m=0}^{N-1} b_m e^{-2\pi imt}\Big)\Big(\frac{1}{N}\sum_{m'=0}^{2(N-1)} c_{m'}e^{2\pi im't}\Big).e^{-2\pi int}dt\right|^2\\
&\leq \|a\|_\infty^2\frac{1}{N}\int\left|\sum_{m=0}^{N-1} b_m e^{-2\pi imt}\right|^2\left|\frac{1}{N}\sum_{m'=0}^{2(N-1)} c_{m'}e^{2\pi im't}\right|^2 dt \text{ by Parseval's inequality}\\
&\leq \|a\|_\infty^2\|b\|_\infty^2\sup_t\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} c_{m'}e^{2\pi im't}\Big|^2.
\end{aligned}$$

This provides a first bound for $|M_N(a,b,c)|^2$. To obtain the second bound we can start instead in the following manner.

$$\begin{aligned}
&|M_N(a,b,c)|^2\\
&\leq \|b\|_\infty^2\frac{1}{N}\sum_{m=0}^{N-1}\left|\int\Big(\frac{1}{N}\sum_{n=0}^{N-1} a_n e^{-2\pi int}\Big)\Big(\sum_{n'=0}^{2(N-1)} c_{n'}e^{2\pi in't}\Big)e^{2\pi mt}dt\Big)\right|^2.
\end{aligned}$$

From these last steps by using a similar path we obtain the second bound. The same idea gives the third bound. □

Remarks 3:

(1) The proof shows a little more than what it stated in this lemma. For instance if one focus on the sequence c_n one does not need to assume that the sequence c_n is also

bounded as we have the estimate

$$|M_N(a,b,c)|^2 \leq \|a\|_\infty^2 \|b\|_\infty^2 \sup_t \Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} c_{m'} e^{2\pi i m' t}\Big|^2.$$

(2) A second look at the proof shows that

$$|M_N(a,b,c)|^2 \leq \|a\|_\infty^2 \frac{1}{N}\sum_{n=0}^{N-1} \Big|\frac{1}{N}\sum_{m=0}^{N-1} b_m c_{m+n}\Big|^2 \leq \|a\|_\infty^2 \|b\|_\infty^2 \sup_t \Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} c_{m'} e^{2\pi i m' t}\Big|^2.$$

We will use these remarks later.

We denote by

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

the averages of seven bounded sequences $A_i = (a_{i,n})$, $1 \leq i \leq 7$. We denote similarly by $M_N(A_1, A_2, ..., A_6, f)$ the sequence where $a_{7,n+m+p} = f(T^{n+m+p}x)$. (Even if we focus later only on this sequence the interested reader will verify that the conclusions reached for this sequence also hold when any one of the bounded sequences A_i is replaced with f.) We also define 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. Thus $(1,2)$ is not a connected couple of integers but $(2,3)$ is because of the terms $a_{2,n}$ and $a_{3,p+n}$ appearing in the numerator of the averages. One can observe that for all integer i, $1 \leq i \leq 7$ there exists an integer j so that (i,j) is connected.

Lemma 6. *Let $A_i = (a_{i,n})$, $1 \leq i \leq 7$, $n \in \mathbb{N}$ be seven bounded sequences that we assume for simplicity to be bounded by one. Then for each N positive integer we have*

$$\big|M_N(A_1, A_2, ..., A_7)\big|^2$$

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

Proof.

$$
\begin{aligned}
&|M_N(A_1, A_2, ..., A_7)(x)|^2 \\
&= \left|\frac{1}{N^3}\sum_{p=0}^{N-1} a_{1,p}\sum_{n=0}^{N-1} a_{2,n}a_{3,p+n}\big(\sum_{m=0}^{N-1} a_{4,m}a_{5,n+m}a_{6,p+m}a_{7,n+m+p}\big)\right|^2 \\
&\leq \frac{1}{N^2}\sum_{p=0}^{N-1}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} a_{4,m}a_{5,n+m}a_{6,p+m}a_{7,p+n+m}\right|^2 \\
&= \sum_{n=0}^{N-1}\sum_{p=0}^{N-1}\left|\int\big(\sum_{m=0}^{(N-1)} a_{4,m}a_{5,n+m}e^{-2\pi imt}\big)\big(\frac{1}{N}\sum_{m'=0}^{2(N-1)} a_{6,m'}a_{7,n+m'}e^{2\pi im't}\big).e^{-2\pi ipt}dt\right|^2 \\
&\leq \sum_{n=0}^{N-1}\int\left|\big(\sum_{m=0}^{N-1} a_{4,m}a_{5,n+m}e^{-2\pi imt}\big)\big(\frac{1}{N}\sum_{m'=0}^{2(N-1)} a_{6,m'}a_{7,n+m'}e^{2\pi im't}\big)\right|^2 dt \\
&\leq \frac{C}{N^2}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m'=0}^{2(N-1)} a_{6,m'}a_{7,n+m'}e^{2\pi im't}\right|^2 N \\
&= C\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m'=0}^{2(N-1)} a_{6,m'}a_{7,n+m'}e^{2\pi im't}\right|^2.
\end{aligned}
$$

If we had bounded above

$$
\sum_{p=0}^{N-1}\left|\int\big(\sum_{m=0}^{(N-1)} a_{4,m}a_{5,n+m}e^{-2\pi imt}\big)\big(\frac{1}{N}\sum_{m'=0}^{2(N-1)} a_{6,m'}a_{7,n+m'}e^{2\pi im't}\big).e^{-2\pi ipt}dt\right|^2
$$

by

$$
\frac{C}{N^2}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} a_{4,m}a_{5,n+m}e^{-2\pi imt}\right|^2 N
$$

then we would have obtained instead the upper bound

$$
C\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} a_{4,m}a_{5,n+m}e^{-2\pi imt}\right|^2.
$$

By a similar argument we obtain the bounds listed in the lemma. □

Generalizations of these lemmas to averages along the cubes of 2^k-1 bounded sequences $a_{i,n}$, $1 \leq i \leq 2^k-1$ can also be obtained. But for simplicity we only state and prove the cases of three and seven sequences.

3.1. **Pointwise convergence of weighted averages along the cubes.** We recall the definition of the generalized Kronecker factor of a measure preserving system. Given a measure preserving system $(X, \mathcal{B}, \mu, T)$, we denote by $\tilde{\mathcal{K}}$ the closed linear span in L^2 of functions f that satisfy $f(Tx) = g(x)f(x)$ for some T-invariant function g with $|g(x)| = 1$.

For ergodic systems we have $\tilde{\mathcal{K}} = \mathcal{K}$. We have the following characterization of functions orthogonal to the generalized Kronecker factor.

Proposition 8. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system on. Consider the generalized Kronecker factor $\tilde{K}$. For a function f consider the splitting $f = f_1 + f_2$ where $f_1 \in \tilde{K}$ and $f_2 \in \tilde{K}^{\perp}$. Then the following are equivalent*

(1) $f = f_2$

(2) *For μ a.e. x*

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

Proof. If f is orthogonal to the "generalized Kronecker" factor then using the ergodic decomposition of T for almost every component T_s of the transformation T the function f is orthogonal to the (regular) Kronecker factor of the transformation T_s. Therefore for a.e. s for μ_s a.e. y applying the uniform Wiener Wintner theorem one gets

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

From this one obtains condition (2) by integration. This proves one implication.

For the reverse implication consider f satisfying condition (2). Let us take a function H verifying $H(Tx) = g(x)H(x)$ with $|g(x)| = 1$. We have

$$\int f\overline{H}d\mu = \int f(T^n x)\overline{H(T^n x)}d\mu = \int f(T^n x)g(x)^n\overline{H(x)}d\mu.$$

Therefore for each integer N we have

$$|\int f\overline{H}d\mu| = |\int \frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(x)^n\overline{H(x)}d\mu| \leq \int \sup_t \left|\frac{1}{N}\sum_{n=1}^{N} f(T^n y)e^{2\pi int}\right| |\overline{H(x)}|d\mu.$$

Taking the limit when N tends to infinity one obtains $\int f\overline{H}d\mu = 0$ and this proves condition (1).

□

Lemma 7. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $f \in \tilde{\mathcal{K}}^{\perp}$. 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) $\lim_N \frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n f(T^m x)c_{n+m} = 0.$

Proof. We only establish the first universal (the null set is independent of the bounded sequences) limit. We take a function $f \in \tilde{\mathcal{K}}^{\perp}$. We choose x in the set of full measure for which by the previous proposition $\lim_N \sup_t |\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)e^{2\pi int}| = 0$. This set is independent of any other bounded sequence a_n or b_n. Applying lemma 5 to the sequence $c_n = f(T^n x)$ we obtain

$$\limsup_N \Big|\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f(T^{n+m}x)\Big|^2 \leq C \limsup_N \sup_t \Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f(T^{m'}x)e^{2\pi i m' t}\Big|^2 = 0.$$

Here we used the first remark made after the proof of Lemma 5. The sequence c_n is not uniformly bounded but a_n and b_n are. A similar argument gives a proof of the other limits. □

Proposition 9. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and let $f \in L^2(\mu)$. Then for μ a.e. x for all bounded sequences a_n, b_n such that $\frac{1}{N}\sum_{n=0}^{N-1} a_n e^{2\pi i n t}$ and $\frac{1}{N}\sum_{n=0}^{N-1} b_n e^{2\pi i n t}$ converge for each t, the sequence*

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

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}.

Proof. We only give the proof for the convergence of the sequence

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

The convergence of the other sequences can be obtained in a similar way. We give the details of the proof in order to keep track of the null sets and show that the set of convergence is truly independent of the sequences.

Let us take $f \in L^2(\mu)$. We decompose this function into the sum $f_1 + f_2$ where $f_1 \in \tilde{\mathcal{K}}$ and $f_2 \in \tilde{\mathcal{K}}^{\perp}$. By lemma 7 for μ a.e. x for all bounded sequences a_n and b_n we have $\lim_N M_N(a, b, f_2)(x) = 0$. It remains to establish the convergence of $M_N(a, b, f_1)(x)$. This convergence follows easily from the assumptions made on a_n and b_n when the function f_1 is a finite linear combination of functions Δ satisfying $\Delta(Tx) = g(x)\Delta(x)$ for some invariant function g with $|g(x)| = 1$. We denote by $\mathcal{W}$ the set of bounded sequences w_n for which $\lim_N \frac{1}{N}\sum_{n=0}^{N-1} w_n e^{2\pi i n t}$ exists for each $t \in \mathbb{R}$. We want to show that

$$\sup_{a,b\in\mathcal{W}} \big[\limsup_N M_N(a,b,f)(x) - \liminf_N M_N(a,b,f)(x)\big] = 0.$$

We consider a sequence F_i of finite linear combinations of eigenfunctions converging in norm to f_1. We have for μ a.e x,

$$\begin{aligned} &\sup_{a,b\in\mathcal{W}} \big[\limsup_N M_N(a,b,f)(x) - \liminf_N M_N(a,b,f)(x)\big] \\ &= \sup_{a,b\in\mathcal{W}} \big[\limsup_N M_N(a,b,f_1 - F_i)(x) - \liminf_N M_N(a,b,f_1 - F_i)(x)\big] \\ &\leq 2\|a\|_\infty \|b\|_\infty M^*[M^*(|f_1 - F_i|)](x) \end{aligned}$$

where as in the proof of proposition 6 we denote by M^* the maximal operator associated with the ergodic averages.

As $\|M^*[M^*(|f_1 - F_i|)]\|_2 \leq 2\|f_1 - F_i\|_2$ and $\lim_i \|f_1 - F_i\|_2 = 0$ we have

$$\liminf_i M^*[M^*(|f_1 - F_i|)](x) = 0.$$

As $\sup_{a,b\in\mathcal{W}} \big[\limsup_N M_N(a,b,f)(x) - \liminf_N M_N(a,b,f)(x)\big]$ does not depend on i we conclude that μ a.e. x we have $\sup_{a,b\in\mathcal{W}} \big[\limsup_N M_N(a,b,f)(x) - \liminf_N M_N(a,b,f)(x)\big] = 0$. This proves the proposition.

$\square$

We recall that Z_2 is the Conze-Lesigne factor.

Lemma 8. *Let* $(X, \mathcal{B}, \mu, T)$ *be an ergodic dynamical system and* $f \in {Z_2}^{\perp}$ *then for* μ *a.e.* x, *for all bounded sequences* a_n *we have*

$$\lim_N \frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} a_m f(T^{n+m}x)e^{2\pi imt}\right|^2 = 0 \textit{ and}$$

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

Proof. We only give a proof for the first limit. The second limit can be established similarly. We can assume that the sequence a_n is real and bounded by one. We follow the steps of the proof of lemma 3. As the arguments are similar we skip some of the steps. By van der Corput's inequality for $(H+1)^2 < N$ we get

$$\begin{aligned}&\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} a_m f(T^{n+m}x)e^{2\pi imt}\right|^2\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\left|\frac{1}{N}\sum_{m=0}^{N-h-1} a_m a_{m+h} f(T^{m+n}x)\overline{f(T^{m+n+h}x)}\right|\end{aligned}$$

and

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} a_m f(T^{n+m}x)e^{2\pi imt}\right|^2\\ &\leq \frac{C}{H} + \frac{C}{H}\sum_{h=1}^{H}\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} a_m a_{m+h} f(T^{m+n}x)\overline{f(T^{m+n+h}x)}\right|.\end{aligned}$$

Thus using the equation (7) we obtain

$$\begin{aligned}&\frac{1}{N}\sum_{n=0}^{N-1}\sup_t\left|\frac{1}{N}\sum_{m=0}^{N-1} a_m f(T^{n+m}x)e^{2\pi imt}\right|^2\\ &\leq \frac{C}{H} + \left(\frac{C}{H}\sum_{h=1}^{H}\left(\frac{1}{N}\sum_{n=0}^{N-1}\left|\frac{1}{N}\sum_{m=0}^{N-1} a_m a_{m+h} f(T^{m+n}x)\overline{f(T^{m+n+h}x)}\right|^2\right)\right)^{1/2}.\end{aligned}$$

Finally by applying the second part of the remarks 3 to the sequences $b_m = a_m.a_{m+h}$ and $c_{n+m} = f(T^{m+n}x)\overline{f(T^{m+n+h}x)}$ and Lemma 2 we get the following estimate

$$\limsup_N \frac{1}{N}\sum_{n=0}^{N-1} \sup_t \left|\frac{1}{N}\sum_{m=0}^{N-1} a_m f(T^{n+m}x)e^{2\pi imt}\right|^2 \leq C|||f|||_3^2. \tag{15}$$

Thus if $f \in {Z_2}^{\perp}$ then $|||f|||_3 = 0$ and the limit is equal to zero. An examination of the proof shows that the set of convergence of full measure is independent of the sequence a_n. This ends the proof of the lemma. □

Proposition 10. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system and let $f \in {Z_2}^{\perp}$. 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.

Proof. This is a simple consequence of Lemma 6 and Lemma 8. □

The same method shows that proposition 10 holds also for the sequences $M_N(f, A_2, A_3, ..., A_7)(x)$, $M_N(A_1, f, A_3, ..., A_7)(x)$,..., where for instance with $M_N(f, A_2, A_3, ..., A_7)(x)$ the bounded sequence A_1 is replaced with the sequence $(f(T^nx))$. We can extend the method by induction on k for higher order averages. We just state the theorem. Once again for the sake of simplicity we only write one of the sequences for which the pointwise convergence holds. The proof follows a similar induction step as in the first part of this paper. The previous propositions and lemmas show how to make the induction work.

Remark 4

We showed in [3] that if we only assume that the limit of $\frac{1}{N}\sum_{n=0}^{N-1} a_{k,n}e^{2\pi int}$ exists for each t and for each $1 \leq k \leq 6$ then the averages

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

may diverge for functions $f \in Z_2$.

Theorem 11. *Let* $(X, \mathcal{B}, \mu, T)$ *be an ergodic dynamical system and let* $f \in Z_{k-1}^{\perp}$. *For* μ *a.e.* x *for all bounded sequences* $A_i = (a_{i,n})$, $1 \leq i \leq 2^k - 2$ *the averages along the cubes* $M_N(A_1, A_2, ..., A_{2^k-2}, f)(x)$ *converge to zero.*

Remark 5. If $f \in Z_{k-1}$ it is not true that we have convergence of the relevant averages along cubes. Already for $k = 2$ the averages $M_N(a, b, f)$ do not converge for all bounded sequence if f is the constant function 1. We can easily find bounded sequences for which the averages $\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m$ do not converge.

3.2. **Pointwsise convergence for averages along cubes for not necessarily commuting measure preserving systems.** In this subsection we will be interested in averages along the cubes for not necessarily commuting measure preserving transformations. Thus in the case of three functions we will look at the averages

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

where T_i are measure preserving transformations on the same measure space. The averages of seven functions are defined as

$$\frac{1}{N^3}\sum_{n,m,p=0}^{N-1} f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^p x)f_4(T^{n+m}x)f_5(T_5^{n+p}x)f_6(T_6^{m+p}x)f_7(T_7^{n+m+p}x).$$

Generalizations to higher order averages are clear. One can observe that if T_1 and T_2 do not necessarily commute then the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T_1^n x)g(T_2^n x)$$

may diverge [4]. Also an example given in [15] 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 also diverge if T_1 and T_2 do not necessarily commute. However we recall Theorem 5 already quoted in the introduction which is apparently new even for the norm convergence. We give now a proof of this theorem.

Theorem 5

Let $(X, \mathcal{B}, \mu)$ be a probability measure space and T_1, T_2 , T_3 three not necessarily commuting measure preserving transformations on $(X, \mathcal{B}, \mu)$. Then for all bounded functions f_i, $1 \leq i \leq 3$ the averages

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

converge a.e. and in L^2 norm.

Proof. This is a consequence of Proposition 9 and the Wiener Wintner ergodic theorem. We consider the null set of points x off which the sequences

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T_1^n x)e^{2\pi int}$$

and

$$\frac{1}{N}\sum_{n=1}^{N} f_2(T_2^n x)e^{2\pi int}$$

converge for every t. This null set is given by the Wiener Wintner ergodic theorem. For the function f_3 and the transformation T_3 we consider the null set of points x given by Proposition 9 off which the averages

$$\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n b_m f_3(T_3^{n+m}x)$$

converge 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. Deleting these two null sets we obtain for $a_n = f_1(T_1^n x)$ and $b_n = f_2(T_2^n x)$ the convergence of the averages

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

□

Remarks 6

(1) One of the interests of Theorem 5 is that it highlights the different behavior between the averages along cubes and the diagonal averages for not necessarily commuting transformations. The diagonal averages do not necessarily converge even in norm [4].

(2) We have extended in [3] Theorem 5 to the averages of six functions.

(3) If each transformation is ergodic then we can identify the limit. If we write the projection of each function f onto the Kronecker factor as $E(f_i|\mathcal{K}) = \sum_{j=0}^{\infty}(\int f_i\overline{e_{j,i}}d\mu)e_{j,i}$, where $e_{j,i}$ is an eigenfunction corresponding to the eigenvalue $\lambda_{j,i}$, then simple computations show that the limit is equal to $\sum_{\mathcal{J}}\prod_{i=1}^{3}(\int f_i\overline{e_{j,i}}d\mu)e_{j,i}$. The set $\mathcal{J}$ denotes the set of eigenvalues of T_3 which are common to T_2 and T_1.

(4) At the present time we do not know if the pointwise convergence holds for averages along the cubes of 2^k-1 functions for $k>2$ for not necessarily commuting measure preserving transformations. However if the transformations T_i, $1\leq i\leq k$ are weakly mixing then we can establish the pointwise convergence of the averages for all positive integer k and identify the limit.

Theorem 12. *Let* T_i, $1\leq i\leq 2^k-1$ *be ergodic measure preserving transformations on the measure space* $(X,\mathcal{B},\mu)$ *and let* f_i , $1\leq i\leq 2^k-1$ *be bounded functions. If we denote by* $Z_{k-1,i}$ *the corresponding* Z_{k-1} *factor for* T_i *then if one of the functions* $f_i\in Z_{k-1,i}^{\perp}$ *then for* μ *a.e.* x

$$\lim_N M_N(f_1,f_2,...,f_{2^k-1})(x)=0$$

where in the case of seven functions

$$M_N(f_1,...,f_7)(x)$$

$$=\frac{1}{N^3}\sum_{n,m,p=0}^{N-1}f_1(T_1^n x)f_2(T_2^m x)f_3(T_3^p x)f_4(T^{n+m}x)f_5(T_5^{n+p}x)f_6(T_6^{m+p}x)f_7(T_7^{n+m+p}x).$$

Proof. It is a consequence of Theorem 11 as the set of convergence to zero is independent of the sequences A_i. This allows us to take $A_i=(a_{i,n})=(f_i(T_i^n x))$. □

Remarks 7

(1) We do not need in Theorem 12 to have each transformation to be ergodic to have the limit equals zero. For instance for the case of seven bounded functions f_i if one assumes that the transformations T_i for $1\leq i\leq 6$ are simply measure preserving and T_7 is ergodic then if $f_7\in Z_{27}^{\perp}$ then for μ a.e. x $\lim_N M_N(f_1,f_2,...,f_7)(x)=0$. This is a consequence of Theorem 11.

(2) At the present time it does not seem simple to control the averages when each function $f_i \in Z_{k-1,i}$. The factors have no reason to be the same for these transformations.

In this direction we have the following results.

Theorem 13. *Let $(X, \mathcal{B}, \mu)$ be a probability measure space and T_i ergodic commuting measure preserving transformation on this measure space. Then for each k positive integer and for every bounded function f_i, $1 \leq i \leq 2^k - 1$ the averages $M_N(f_1, f_2, ..., f_{2^k-1})$ converge a.e. as N tends to infinity.*

Proof. We denote by $Z_{k-1,i}$ the Z_{k-1} factor for the ergodic transformation T_i. By Theorem 12 the averages along the cubes converge a.e to zero when one of the functions f_i belongs to $Z^{\perp}_{k-1,i}$. The key observation with the assumptions made on the transformations T_i (being ergodic and commuting) is the fact that in this case $Z_{k-1,i} = Z_{k-1,j}$ for all $1 \leq i \leq j \leq 2^k - 1$. This observation was also made independently by N. Frantzikinakis and B. Kra in [8]. Therefore one is left with proving the a.e. convergence when the commuting action of the maps T_i is reduced to this common factor. As pointed by the referee, as done in [8], one can observe that the isomorphism between the common factor Z_{k-1} and the inverse limit of nilsystems can be taken to be the same for all T_i. The factor maps can also be taken to be the same for all T_i. But in this case one can apply Leibman 's result (Theorem 3) mentioned in the introduction and uses an approximation to reach all bounded functions measurable with respect to this inverse limit of nilsystems. □

The next result eliminates the commuting assumption of the transformations but adds a more restrictive assumption on the transformations T_i.

Proposition 14. *Let $(X, \mathcal{B}, \mu)$ be a probability measure space and T_i weakly mixing transformations (not necessarily commuting) on this measure space. Then the averages along the cubes applied to the bounded functions f_i, $1 \leq i \leq 2^k - 1$ converge a.e. to* $\prod_{i=1}^{2^k-1} \int f_i d\mu$.

Proof. When each transformation is weakly mixing the factors $Z_{k-1,i}$ are all reduced to the trivial one. Hence the projection on this factor is just the integral $\int f_i d\mu$. The a.e. convergence and the identification of the limit follows then from Theorem 11 and this observation. □

References

[1] **I. Assani**:"Pointwise convergence along cubes for measure preserving systems", (preprint 2003) available at www.arxiv.org. math.DS/0311274.

[2] **I. Assani**:" Wiener Wintner ergodic theorems", World Scientific Publ. Co; 2003, ISBN: 9810244398.

[3] **I. Assani**:"Averages along cubes for not necessarily commuting measure preserving transformations", Contemporary Math., vol.430 1-19, 2007.

[4] **D. Berend**:"Joint ergodicity and mixing", *J.d'Anal. Math.*, 45,(1985), 255-284.

[5] **V. Bergelson**:"The multifarious Poincare Recurrence theorem," *Descriptive Set Theory and Dynamical Systems,*, Eds M. Foreman, A.S. Kechris, A. Louveau, B. Weiss. Cambridge University Press, New York (2000), 31-57.

[6] **J. Bourgain**:"Double recurrence and almost sure convergence," *J. für die Reine und Angewandte Mathematik*, 404, 140–161, 1990.

[7] **C. Demeter, T. Tao and C. Thiele**:"Maximal multilinear operators", *Trans. Amer. Math. Soc.* , to appear.

[8] **N. Frantzikinakis and B. Kra**:"Convergence of multiple ergodic averages for some commuting transformations", *Erg. Th and Dyn. Syst.* 25, Issue 3, (2005), 799-809.

[9] **H. Furstenberg and B. Weiss**:"A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2} x)$",*Convergence in Ergodic Theory and Probability*, Eds: Bergelson/March/Rosenblatt, Walter de Gruyter-Co, Berin, New York (1996), 193-227.

[10] **W.T. Gowers**:" A new proof of Szemerédi's theorem," *Geom. Funct. Anal.*, 11, (2001), no. 3, 465–588.

[11] **B. Host and B. Kra**: "Nonconventional ergodic averages and nilmanifolds", Annals of Math., 161,1, 397-488 (2005)

[12] **A. Y. Khintchine**:"Eine Verscharfung des Poincareschen "Wiederkehrsatzes"," *Comp. Math.*, 1, (1934), 177-179.

[13] **L. Kuipers and H. Niederreiter**: *Uniform Distribution of Sequences.* John Wiley & Sons, 1974.

[14] **A. Leibman**:"Pointwise convergence of ergodic averages for polynomial sequences of translations of a nilmanifold":*Erg. Th. and Dyn. Syst.* 25 (2005), no. 1, 201-113.

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

[16] **M. Schwartz**: "Polynomially moving ergodic averages, " *Proc. Amer. Math. Soc.*,vol. 103,no. 1, pp. 252-254, 1988.

[17] **T. Ziegler**:"Universal characteristic factors and Furstenberg averages", *J. Amer. Math. Soc.*, 20, (2007), 5397.