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 [10]) 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 von Neumann ergodic theorem as

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

V. Bergelson [6] introduced a generalization of Khintchine's recurrence theorem by considering the expressions

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

He proved the following convergence result.

Theorem 2. *(V. Bergelson [6])*
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 $M_N(f_1, f_2, ..., f_{2^k-1})$. Averages of this form 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 [9].

In [4] 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 and 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 a result of A. Liebman [13].
(2) 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}.

One consequence of their method 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 CL factor. When the system is ergodic the limit in L^2 norm of the averages along the cubes has been identified in [4] (see theorem 13.1). From this they derived (theorem 1.3) the following inequality for the averages of seven functions

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

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

The notion of characteristic factor is due to H. Furstenberg and can be found explicitly stated in [8].

This paper answers the question raised by Host and Kra [4] 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 ([7]). The current paper is divided into two parts. In the first part we focus on the averages along the cubes and prove the following results.

Theorem 3. *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.*

The pointwise limit of the averages is of course the same as the L^2 limit identified in [4]. More general averages along the cubes were considered and proved to converge in norm in [6] and [4]. Not all of them converge almost everywhere and so for them the factors Z_{k-1} are characteristic in norm but not pointwise.

Theorem 4. *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic dynamical system. 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. However the Kronecker factor which is characterisitic in L^2 norm for the averages*

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

is not in general pointwise characteristic for these averages.

In the second part we give applications of the method we used in the first part. With key estimates on bounded sequences of scalars already announced in [1], we derive pointwise convergence results for weighted averages (suggested by an earlier referee) and for averages along cubes for not necessarily commuting measure preserving transformations. These results partially extend those obtained in the previous sections. Some of the results presented in this section were also announced in [1].

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 the 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 [7].

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

Using this lemma we can prove the following proposition. It is inspired from the computations made on the third page of [7].

Proposition 5. *Let $(X, \mathcal{B}, \mu, T)$ be a measure preserving system and f_i, $1 \le i \le 3$ 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 (using the ergodic decomposition). We have the following inequalities.

$$|M_N(f_1,f_2,f_3)(x)|^2$$

$$\leq \|f_1\|_\infty^2 \Big(\frac{1}{N}\sum_{n=0}^{N-1}\big|\frac{1}{N}\sum_{m=0}^{N-1} f_2(T^m x) f_3(T^{n+m}x)\big|^2\Big), \text{by Cauchy-Schwarz' inequality,}$$

$$\leq \|f_1\|_\infty^2 \frac{1}{N}\sum_{n=0}^{N-1}\Big|\int \big(\sum_{m=0}^{N-1} f_2(T^m x)e^{-2\pi imt}\big)\big(\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\big).e^{-2\pi int}dt\Big|^2$$

$$\leq \|f_1\|_\infty^2 \frac{1}{N}\int\Big|\sum_{m=0}^{N-1} f_2(T^m x)e^{-2\pi imt}\Big|^2\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\Big|^2 dt, \text{by Parseval's inequality,}$$

$$\leq \frac{C}{N}\sup_t\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\Big|^2\int|\sum_{m=0}^{N-1} f_2(T^m x)e^{-2\pi imt}|^2 dt$$

$$\leq C\sup_t\Big|\frac{1}{N}\sum_{m'=0}^{2(N-1)} f_3(T^{m'}x)e^{2\pi im't}\Big|^2\frac{1}{N}N\|f_2\|_\infty^2$$

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.

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

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

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

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

Applying this maximal inequality twice we conclude that

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

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

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

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. □

Remarks 1

- The proof of proposition 5 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

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

- The proof of this proposition actually shows that

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

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 [4] it is shown that the CL 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 $f\in CL^\perp$ 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*

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

In particular we have

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

Proof. Without loss of generality we can assume that the function f takes only real values. We apply van der Corput's inequality ([12]) . Because of this inequality for $H < N$ we get

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

(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

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

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 f_1 or f_2 is in $CL^{\perp}$ then for a.e. x*

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

Proof. We can assume without loss of generalities that the functions are uniformly bounded by one. We use again van der Corput's inequality, [12]. For $(H+1)^2 < N$ we get

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

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

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

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 $\frac{1}{H}\sum_{h=1}^{H}|u_h|^2 \leq \big(\frac{1}{H}\sum_{h=1}^{H}|u_h|^4\big)^{1/2}$ 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 $f_1 \in CL^{\perp}$ then $|||f_1|||_3 = 0$ and we obtain the equation (7). We have the same conclusion if one assumes that $f_2 \in CL^{\perp}$. □

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

Proposition 6. *The averages along the cubes of seven functions converge almost surely.*

Proof.

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

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

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

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

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

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

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

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

With the help of the lemma 3 one can conclude that if f_6 or f_7 belong to $CL^{\perp}$ then the averages of these seven functions converge to zero. By using the symmetry of the averages with respect to n, m and p one can see that the averages will converge to zero if one of the functions $f_i \in CL^{\perp}$, $1 \leq i \leq 7$.

□

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 $P_{CL}(f_i)$ their projection onto the CL factor then we have

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

- The proof of lemma 3 gives the following estimate

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

2.3. **Proof of Theorem 3 for the case of three and seven functions.** The proof is a consequence of the path used in establishing the propositions 5 and 6. We have shown that if one of the functions $f_i \in CL^{\perp}$, $1 \leq i \leq 7$, then the averages converge pointwise to zero. This shows that the CL factor is characteristic for the pointwise convergence. For the averages of three functions the Kronecker factor is characteristic for the pointwise convergence for the same reason.

2.4. **Proof of theorem 3.** We will prove theorem 3 by induction on k. In the previous sections we proved that the averages of seven functions converge a.e. We showed that the $Z_2 = CL$ factor was characteristic for the pointwise convergence of such averages. This established the first step of the induction process. 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

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

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. *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 $(H+1) << N$

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

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

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

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

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

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

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

So by the induction assumption we have

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

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

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

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

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

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

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

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

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 3.

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, $M_N(f_1, f_2, \dots, f_{2^k-1})(x)$. With the inequality (11) we have

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

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

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

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

By using lemma 4 and (12) one concludes that

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

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

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

By symmetry on the indices $i_1, i_2, ..., i_k$ one obtains the following inequality for the 2^{k-1} functions f_j

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

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

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 $M_N(f_1, f_2, ..., f_{2^k-1})(x)$ converge a.e. to zero if one of the functions $f_j \in Z_{k-1}^\perp$ (using the symmetry of the indices). Combining this result with the pointwise convergence when all functions are in Z_{k-1} mentioned in [4], (see also [13]) this ends the proof of theorem 1.

2.5. **Proof of theorem 4.** The proof of the first part of theorem 4 follows from the path we used. We showed that if one of the functions f_j is in the orthocomplement of the Z_{k-1} factor then the averages of these $2^k - 1$ functions converge a.e to zero. Thus the limit is given by the pointwise convergence when all functions are in the factor Z_{k-1}.

If one considers instead the averages

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

where $(N-M)$ tends to ∞ then we do not have a.e. convergence in general while as shown in [6] and [4] we do have convergence in L^2 norm. For instance it is shown in [15] that for $\beta \geq 3$ the averages

$$\frac{1}{N^{\beta-1}} \sum_{n=N^\beta}^{(N+1)^\beta} f(T^n x)$$

do not converge a.e. even if f is the characteristic function of a set of positive measure. So in this case the Kronecker factor is characteristic for the L^2 norm but not for the pointwise convergence.

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 when they are applied to functions in the orthocomplement of the factors $\mathcal{Z}_k$. However the difficulty in this case compared to the first section is the convergence on the factors. More precisely we could use [13] to obtain the convergence on the factors in theorem 3 when dealing only with bounded functions. Here with only bounded sequences of scalars the situation is more complicated. With the ergodicity assumption one can still get some control in the orthocomplement of the factors (see Theorem 9). But without ergodicity identifying the proper pointwise characteristic factors is not easy. Furthermore the convergence on the factors requires some conditions on the sequences (see Remark 5). The approach we use here is a combination of the estimates on bounded sequences of scalars and the ergodic decomposition. This provides us with pointwise results such as Theorem 10 with no clear identification of the limit and the pointwise characteristic factor.

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*

$$\Big|\frac{1}{N^2}\sum_{m,n=0}^{N-1} a_n.b_m.c_{n+m}\Big|^2$$

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

Proof. We denote by $M_N(a,b,c)$ the quantity $\frac{1}{N^2}\sum_{n,m=0}^{N-1} a_n.b_m.c_{n+m}$. The steps are similar to those given in the proof of Proposition 5 so we only sketch them. We have

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

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.

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

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 bounded by one. Then for each N positive integer we have*

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

Proof.

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

If we had bounded above

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

by

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

then we would have obtained instead the upper bound

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

By using the symmetry of the indices 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.**

Lemma 7. *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) $\lim_N \frac{1}{N^2} \sum_{n,m=0}^{N-1} a_n f(T^m x) c_{n+m} = 0.$

Proof. We only establish the first universal (the null set is independent of the bounded sequences) limit. We take a function $f \in \mathcal{K}^{\perp}$. We choose x in the set of full measure for which by lemma 1 $\lim_N \sup_t |\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) e^{2\pi i n t}| = 0$. 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. ☐

Remark As shown in [3] the ergodicity assumption in Lemma 7 is necessary.

Proposition 7. *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}.

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

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

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. (see also Remark 4 below)

Let us take $f \in L^2(\mu)$. We decompose this function into the sum $f_1 + f_2$ where $f_1 \in \mathcal{K}$ and $f_2 \in \mathcal{K}^\perp$. 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 eigenfunctions. We denote by $\mathcal{W}$ the set of bounded sequences w_n for which $\lim_N \frac{1}{N}\sum_{n=0}^{N-1} w_n e^{2\pi int}$ exists for each $t \in \mathbb{R}$. We want to show that

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

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,

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

$$= \sup_{a,b\in\mathcal{W}} \big[\limsup_N M_N(a,b,f_1-F_i)(x) - \liminf_N M_N(a,b,f_1-F_i)(x)\big]$$

$$\leq 2\|a\|_\infty\|b\|_\infty M^*[M^*(|f_1-F_i|](x)$$

where as in the proof of proposition 5 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.

Remark 4. We currently do not know if the proposition 7 can be obtained without any ergodicity assumption. The obstacle following the path above seems to be the measurability of the set of x for which the averages $M_N(a,b,f)(x)$ converge for all bounded sequences a_n and b_n. This measurability seems to be needed in order to use the ergodic decomposition. □

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

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

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

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

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

and

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

Thus using the equation (7) we obtain

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

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 $f \in CL^{\perp}$ 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 8. *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 $A_i = (a_{i,n})$, $1 \le i \le 6$ 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 8 holds also for the sequences $M_N(f, A_2, A3, ..., A7)(x)$, $M_N(A_1, f, A_3, ..., A_7)(x)$,..., where for instance with $M_N(f, A_2, A3, ..., A7)(x)$ the bounded sequence A_1 is replaced with the sequence $(f(T^n x))$. 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 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 works.

Remark

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 $1 \le k \le 6$ 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 $f \in CL$.

Theorem 9. *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. Theorem 9 does not hold when the function $f \in Z_{k-1}$. 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

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

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 [5]. Also an example given in [14] 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 have the following result which is apparently new even for the norm convergence.

Theorem 10. *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, $1\le i\le 3$ 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. As the L^2 norm follows from the dominated convergence and the a.e. convergence, it is enough to prove the a.e. convergence. One can observe that if T_3 is ergodic and we take x in the set of full measure where by the Wiener Wintner ergodic theorem for measure preserving transformations (see for instance [2] for a proof of this statement) the sequences $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n x)e^{2\pi int}$ and $\frac{1}{N}\sum_{n=0}^{N-1} f_2(T_2^n x)e^{2\pi int}$ converge for each $t\in\mathbb{R}$ then the sequence

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

will converge. We just need to take $a_n = f_1(T_1^n x)$, $b_m = f_2(T_2^m x)$ and apply the proposition 7. This proves this theorem when T_3 is ergodic. To reach the same conclusion without ergodicity assumption on $(X,\mathcal{B},\mu,T_3)$, we consider an ergodic decomposition $\mu_{c,3}$ for T_3 on $(X,\mathcal{B},\mu)$. This means that on $(X,\mathcal{B},\mu_{c,3})$ the transformation T_3 is measure preserving and ergodic. Furthermore $\mu_{c,3}$ is a disintegration of μ, i.e. for each integrable function $f \in L^1(\mu)$ we have $\int f(x)d\mu(x) = \int f(y)d\mu_{c,3}(y)dP(c)$ where P is a probability measure. We apply this disintegration to the sets where the functions $f_1 \circ T_1^n$ and $f_2 \circ T_2^n$ are bounded and to those where the sequences $\frac{1}{N}\sum_{n=0}^{N-1} f_1(T_1^n x)e^{2\pi int}$ and $\frac{1}{N}\sum_{n=0}^{N-1} f_2(T_2^n x)e^{2\pi int}$ converge for each $t \in \mathbb{R}$. The pointwise convergence established for ergodic dynamical system tells us then that for $\mu_{c,3}$ a.e. y the sequence

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

converge. As the set S of x where the sequence

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

converge is μ measurable and its disintegration is given by sets of measure one with respect to the measures $\mu_{c,3}$ we can conclude that this set S has full measure. This ends the proof of this theorem. □

Remarks 6:

(1) If each transformation is ergodic then we can identify the limit. Writing for each function its projection 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.

(2) 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, identify the limit and obtain some recurrence property.

Theorem 11. *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 9 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) One of the interests of theorem 10 is that it highlights the different behaviors between the averages along cubes and the diagonal averages for not necessarily commuting transformations. The diagonal averages do not necessarily converge even in norm [5].

(2) We do not need in theorem 11 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 $f_7 \in CL_7^{\perp}$ then for μ a.e. x $\lim_N M_N(f_1, f_2, ..., f_7)(x) = 0$. This is a consequence of theorem 9.

(3) 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. So far we can only establish the a.e. convergence when each $f_i \in Z_{k-1,i}$ if one assumes that the transformations T_i are also commuting. One of the main reasons why is because in this case $Z_{k-1,i} = Z_{k-1,j}$ for all $1 \leq i \leq j \leq 2^k - 1$. This fact was observed independently by B. Kra and N. Frantzikinakis in [11].

Corollary 1. *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 $\prod_{i=1}^{2^k-1} \int f_i d\mu$. In particular for $k = 7$ we have for all measurable set A,*

$$\lim_N \frac{1}{N^3} \sum_{n,m,p=0}^{N-1} \mu[A \cap T_1^n A \cap T_2^m A \cap T_3^p A \cap T_4^{n+m} A \cap T_5^{p+n} A \cap T_6^{m+p} A \cap T_7^{n+m+p} A] \geq \mu(A)^8.$$

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. The last part of the corollary is a simple consequence of integration in the particular case of seven transformations applied to the characteristic function of the set A. □

Remarks More on the recurrence property (extension of Khintchine's theorem (cf. theorem 1)) is done in [3].

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 and weighted averages",(in preparation)

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

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

[6] **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.

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

[8] **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.

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

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

[11] **B. Kra and N. Frantzikinakis**:"Convergence of multiple ergodic averages for some commuting transformations", *Erg. Th and Dyn. Syst.* to appear, available at http://www.math.psu.edu/nikos/publications.html

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

[13] **A. Leibman**: *Pointwise convergence of ergodic averages for polynomial sequences of translations of a nilmanifold*: http://www.math.ohio-state.edu/ leibman/preprints/, to appear in Eg. Th. and Dyn. Syst.

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

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