January 3, 2005

POINTWISE CONVERGENCE OF NONCONVENTIONAL AVERAGES

BY I. ASSANI

Abstract. We answer a question of H. Furstenberg on the pointwise convergence of the averages

$$\frac{1}{N}\sum_{n=1}^{N} U^n(f\cdot R^n(g)),$$

where U and R are positive operators. We also study the pointwise convergence of the averages

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

when T and S are measure preserving transformations.

Introduction

Throughout this paper we will denote by $(X, \mathcal{B}, \mu)$ a finite Lebesgue measure space. An operator V on $L^p(\mu)$ is said to be positive if $f \geq 0$ implies $Vf \geq 0$. We denote by U and R two positive linear operators on $L^p(\mu)$, $1 < p < \infty$, such that $\sup_{n\in\mathbb{Z}} \|U^n\|_p < \infty$ and $\sup_{n\in\mathbb{Z}} \|R^n\|_p < \infty$. We assume that their inverses U^{-1} and R^{-1} are also positive. This implies (see [Ka]) that there are nonsingular transformations ϕ and θ and functions ω and Δ such that $U(f)(x) = \omega \cdot f(\phi)(x)$ and $R(f)(x) = \Delta \cdot f(\theta)(x)$ for all functions f in $L^p(\mu)$. We will also consider $(X, \mathcal{B}, \mu, T)$ and $(X, \mathcal{B}, \mu, S)$ two dynamical systems where $T : X \to X$ and

---

Department of Mathematics, UNC Chapel Hill, NC 27599

keywords:nonconventional averages, Kronecker factor

AMS subject classification 37A05, 37A30, 47A35

email: assani@math.unc.edu.

$S : X \to X$ are measure preserving transformations. The present paper is motivated by the following two questions of H. Furstenberg, [F]:

Question 1. *Do we have the pointwise convergence of the averages*

$$\frac{1}{N}\sum_{n=1}^{N} V^n[f \cdot H^n g](x)$$

for all bounded functions f and for all positive operators V and H?

Question 2. *Do we have the pointwise convergence of the averages*

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

for all bounded functions f and g and all measure preserving transformations T and S ?

The averages $\frac{1}{N}\sum_{n=1}^{N} U^n[f \cdot R^n g](x)$ were proposed by H. Furstenberg as a natural generalization at the operator level of the nonconventional averages $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n x)$. When U and R are not necessarily measure-preserving functions, the averages

$$\frac{1}{N}\sum_{n=1}^{N} U^n(f)(x) \cdot R^n(g)(x)$$

may not even be integrable for all functions f and g in $L^p(\mu)$. However, $f \cdot R^n g$ is in $L^p(\mu)$, so using $U^n[f \cdot R^n]$ will give integrable averages to evaluate. The assumption of positivity of the operators in the first question is essential as there are examples of unitary operators (see [Kr] p. 191) on L^2 for which the averages already fail to converge a.e. for some $g \in L^2$ and $f = 1$. As pointed out by one of the referees the averages $\frac{1}{N}\sum_{n=1}^{N} V^n[f \cdot V^n g]$ do not necessarily converge in L^2 norm for V unitary, [Boi]. However we will show in the first part of the paper that the averages $\frac{1}{N}\sum_{n=1}^{N} V^n[f \cdot V^n g]$ converge a.e. when V is a positive contraction in L^p, $1 < p < \infty$, $f \in L^\infty$ and $g \in L^p$.

The present paper is divided in the following way. In the first part, we will focus on the pointwise convergence of the averages

$$\frac{1}{N}\sum_{n=1}^{N} U^n[f \cdot R^n g](x),$$

for positive operators such that $\sup_{n\in\mathbb{Z}} \|U^n\|_p < \infty$ and $\sup_{n\in\mathbb{Z}} \|R^n\|_p < \infty$. We will answer the first question of Furstenberg by showing that the averages do not converge pointwise or even weakly when R is a negative power of U, namely U^{-1}. However, if R is a positive power of U, we do have a.e. convergence for functions in L^p. As indicated earlier we will also show that for a positive contraction V in L^p, $1 < p < \infty$ the averages $\frac{1}{N}\sum_{n=1}^{N} V^n[f \cdot V^n g]$ converge almost everywhere.

The second part of the paper deals with the second question of Furstenberg. We will study first the case when T and S commute. Included in this section are several remarks on the almost everywhere double recurrence theorem of J. Bourgain [Bou]. Lastly, we will study the case where T and S do not necessarily commute.

Acknowledgements

The author thanks H. Furstenberg for bringing these questions to his attention during his stay at the MSRI in 1992. He also thanks the referees of this paper for their remarks.

1. CONVERGENCE OF $\frac{1}{N}\sum_{n=1}^{N} U^n[f\cdot R^n g](x)$

In this section we consider positive linear operators, U and R on $L^p(\mu)$ such that

$$\sup_{n\in\mathbb{Z}}\|U^n\|_p<\infty \text{ and } \sup_{n\in\mathbb{Z}}\|R^n\|_p<\infty.$$

They are called invertible power bounded Lamperti operators. First, we will show that even when R is a power of U, the convergence does not necessarily hold. To do this we will need the following lemma. We sketch a proof for positive contractions on L^p in order to make the paper self-contained. More about such decomposition for L^1 contraction can be found in [Kr]. The conjugate of p is denoted by q.

Lemma 1. *Let V be an invertible power bounded Lamperti operator or a positive contraction on $L^p(\mu)$, $1<p<\infty$. There exists a decomposition of the space X into two subsets C and D such that:*

(1) *There exists $v_0^*\in L^p$ such that supp$(v_0^*)=C$, $V(v_0^*)=v_0^*$ and $V^*(v_0^{*p-1})=v_0^{*p-1}$. Furthermore C is the maximal support of any invariant function for V or its adjoint V^*.*

(2) *If supp$(f)\subseteq C$, then supp$(Vf)\subseteq C$. If supp$(f)\subseteq D$, then supp$(Vf)\subseteq D$.*

Proof. Let v_0^* be the limit in norm of the constant function **1**. which exists by the mean ergodic theorem. The function v_0^* is then V invariant. We denote by C its support and by D the complement of C. As

$$\|v_0^*\|_p^p=\int v_0^*V^*[(v_0^*)^{p-1}]d\mu=\|v_0^*\|_p\|(v_0^*)^{q(p-1)}\|_q,$$

we have the equality for Hölder's inequality. Thus we must have $V^*[(v_0^*)^{p-1}]=(v_0^*)^{p-1}$.

If f has its support in D then

$$0 = \int (v_0^*)^{p-1} f d\mu = \int V^*[(v_0^*)^{p-1}] f.d\mu = \int (v_0^*)^{p-1} V f d\mu.$$

This shows that Vf has also its support in D. To prove that, if a function f has its support in C, then Vf has also its support in C, it is enough to consider functions f such that $f \leq kv_0^*$ for some positive constant k. Functions f with this property are dense in the set of functions with support in C. If $f \leq kv_0^*$ then $Vf \leq kV(v_0^*) = kv_0^*$ and Vf has its support in C. For the case of power bounded Lamperti operators U the proof is similar. One can use Theorem 4.2 in [Ka] to obtain the precise connection between U and its adjoint: if $Uf(x) = \omega \cdot f(\phi)(x)$ then $U^*g(x) = \omega^{1-p} \circ \phi^{-1}(x) g \circ \phi^{-1}(x) D(\omega)(x)$. $\square$

Theorem 2. *There exists an invertible power bounded Lamperti operator U on L^2 (actually an isometry) and functions $f \in L^\infty$ and $g \in L^2$, such that the averages $\frac{1}{N}\sum_{n=1}^{N} U^n(fU^{-n}g)$ do not converge in either norm or almost everywhere.*

Proof. The proof relies on the existence of nonsingular transformations ϕ for which we do not have the convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} \mu(\phi^{-n}(A))$ for all measurable sets A. Take a nonsingular invertible transformation ϕ on $[0,1]$ with no finite invariant measure equivalent to Lebesgue measure μ (D. Ornstein,[O], A. Brunel [Br]). We define an invertible isometry on $L^2(\mu)$ by $Uf = \omega \cdot f \circ \phi$. Using the previous lemma, the lack of a finite invariant measure equivalent to Lebesgue measure implies that $\mu(D) > 0$.

Otherwise we would have $\mu(C) = 1$ and if we take any invariant function v_0^* for U then the measure m defined as $m(A) = \int_A v_0^{*p} d\mu$ would be invariant and equivalent to μ. This would be because $v_0^* > 0$, $\omega^p v_0^{*p} \circ \phi = v_0^{*p}$ and

$$\begin{aligned} m(\phi^{-1}(A)) & \\ = & \int \mathbf{1}_A \circ \phi v_0^{*p} d\mu = \int \mathbf{1}_A \circ \phi \omega \omega^{p-1} v_0^{*p} \circ \phi d\mu \\ = & \int U(\mathbf{1}_A v_0^*) v_0^{*p-1} \circ \phi d\mu = \int \mathbf{1}_A v_0^* v_0^{*p-1} d\mu = m(A). \end{aligned}$$

We claim that the averages

$$\frac{1}{N+1} \sum_{n=0}^{N} \mu(\phi^{-n}(A))$$

cannot converge for all measurable sets $A \subseteq D$.

If the above average converges for all measurable sets A in D, the limit for each A would define an invariant finite measure absolutely continuous with respect to Lebesgue measure. Its Radon Nykodim derivative would then contradict the maximality of C.

Thus there exists a set A for which the averages

$$\frac{1}{N+1} \sum_{n=0}^{N} \mu(\phi^{-n}(A))$$

do not converge. This implies that the averages

$$\frac{1}{N+1} \sum_{n=0}^{N} \chi_A \circ \phi^n = \frac{1}{N+1} \sum_{n=0}^{N} U^n(\chi_A U^{-n} f)$$

do not converge in either norm or almost everywhere. (Note that the averages

$$\frac{1}{N+1} \sum_{n=0}^{N} \chi_A \circ \phi^n$$

are uniformly bounded by 1 and that

$$\frac{1}{N+1} \sum_{n=0}^{N} \mu(\phi^{-n}(A)) = \int \frac{1}{N+1} \sum_{n=0}^{N} \chi_A \circ \phi^n d\mu.)$$

$\square$

Even though the previous result gives a negative answer to the first Furstenberg question, there exist some positive results in the direction of characterizing those operators U and V for which the averages

$$\frac{1}{N}\sum_{n=1}^{N} V^n(fV^n g)$$

converge a.e.

Theorem 3. *Let V be a positive contraction on L^p, $1 < p < \infty$ and f a bounded function. For all $g \in L^p$ the averages*

$$\frac{1}{N}\sum_{n=1}^{N} V^n(fV^n g)$$

converge a.e.

Proof. By Lemma 1 one has to consider only two cases: the case where all the functions involved are supported on D and the other when all functions are supported on C. The first case can be solved by the pointwise ergodic theorem for positive contractions on L^p (see [Kr] p. 189-190). We have for $f \in L^\infty$ and $g \in L^p$ with support in D

$$\frac{1}{N}\sum_{n=1}^{N} V^n(fV^n g)(x) \le \|f\|_\infty \frac{1}{N}\sum_{n=1}^{N} V^{2n} g(x) \to 0, a.e.$$

For the case where all functions have support in C, we can introduce the operator W on $L^\infty(\nu)$ where $d\nu = [v_0^*]^p d\mu$ by the formula

$$W(f) = \frac{V(v_0^* f)}{v_0^*}.$$

Simple computations using Lemma 1 show that $W(\mathbf{1}_C) = 1_C = W^*(\mathbf{1}_C)$, $W^n(g) = \frac{V^n(v_0^* g)}{v_0^*}$ and for $f \in L^\infty(\mu) = L^\infty(\nu)$ we have $W^n(fW^n g) = \frac{V^n[fV^n(gv_0^*)]}{v_0^*}$. Thus W is a Markov operator which extends to a contraction on $L^1(\nu)$ and $L^\infty(\nu)$. So we are left with proving that the averages

$$\frac{1}{N}\sum_{n=1}^{N} W^n(fW^n g)$$

converge a.e. for a Markov positive operator W for which $W1 = 1$ and $W^*1 = 1$. We recall a few properties on the dilation of such operators that we will need. They were used in [Boi]. In order to simplify the notations we assume that W is defined on $(Y, \mathcal{B}, \nu)$.

(1) There exists a probability measure space $(Y \times Z, \mathcal{B} \times \mathcal{P}, \Delta)$ which naturally extends $(Y, \mathcal{B}, \nu)$.

(2) There exists a measure preserving transformation $\Theta : Y \times Z \to Y \times Z$ which is $\mathcal{B} \times \mathcal{P}$ measurable such that if we denote by T the operator $Th = h \circ \Theta$ then for all $h \in L^p(\Delta)$, $1 \leq p \leq \infty$ and all $n \in \mathbb{N}$ we have

$$E[h \circ \Theta^n|\mathcal{B}] = W^n(E[h|\mathcal{B}]).$$

For $f, g \in L^\infty(Y, \mathcal{B}, \nu)$ we can express the quantities $W^n(fW^n g)$ in terms of the operator T;

$$W^n(fW^n g) = W^n[f.E[g \circ \Theta^n|\mathcal{B}]] = W^n\big[E[fg \circ \Theta^n|\mathcal{B}]\big] = E\big[f \circ \Theta^n g \circ \Theta^{2n}|\mathcal{B}\big].$$

Thus the averages $\frac{1}{N}\sum_{n=1}^{N} W^n(fW^n g)$ can be written as $E\big[\frac{1}{N}\sum_{n=1}^{N} f \circ \Theta^n g \circ \Theta^{2n}\big|\mathcal{B}\big]$. The a.e double recurrence result, [Bou], gives the pointwise convergence of $\frac{1}{N}\sum_{n=1}^{N} f \circ \Theta^n(x) g \circ \Theta^{2n}(x)$ for a.e. x. As the maximal function $\sup_N |\frac{1}{N}\sum_{n=1}^{N} f \circ \Theta^n g \circ \Theta^{2n}|$ is integrable, the pointwise convergence still holds with the application of the conditional expectation $E(.|\mathcal{B})$. This ends the proof of this theorem. $\square$

Theorem 4. *Let U be an invertible power bounded Lamperti operator on L^p, $1 < p < \infty$ and f a bounded function. For all $g \in L^p$ the averages*

$$\frac{1}{N}\sum_{n=1}^{N} U^n(fU^n g)$$

converge a.e.

Proof. Because of Lemma 1, we need only look at the following two cases:

a: f and g have their support in C,

b: f and g have their support in D.

Case a: If the support of f and g is in C then we can restrict ourselves to the measure space $(C, \mathcal{B} \cap C, \mu_C)$. The measure m defined by $m(A) = \int_A v_0^{*p} d\mu$ is invariant with respect to ϕ. It is also equivalent to μ. If we denote by $U^n f = \omega_n \cdot f \circ \phi^n$ for each positive integer n, then simple computations lead to the equation

$$U^n(fU^n g)(x) = v_0^* \cdot f \circ \phi^n \cdot (\frac{g}{v_0^*}) \circ \phi^{2n}.$$

An application of the a.e double recurrence theorem of Bourgain [Bou] allows us to conclude that the average $\frac{1}{N}\sum_{n=1}^{N} U^n(fU^n g)$ converges a.e..

Case b: If the support of f and g is in D then the average $\frac{1}{N}\sum_{n=1}^{N} U^{2n}(g)(x)$ converges to 0 a.e . This is because the Cesaro averages $\frac{1}{N}\sum_{n=1}^{N} U^n(|g|)(x)$ converge a.e. to zero as the functions g and $U^n(|g|)$ have their support in D and there is no invariant function with support in D by Lemma 1. The inequality

$$|\frac{1}{N}\sum_{n=1}^{N} U^n(fU^n g)(x)| \leq \frac{\|f\|_\infty}{N}\sum_{n=1}^{N} U^{2n}(|g|)(x)$$

allows us to reach the same conclusion for the averages

$$\frac{1}{N}\sum_{n=1}^{N} U^n(fU^n g)(x).$$

□

2. CONVERGENCE OF $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n x)$

In this section, T and S are measure preserving transformations on the same Lebesgue measure space $(X, \mathcal{B}, \mu)$. If T is ergodic then the Kronecker factor is defined as the σ-algebra generated by the eigenfunctions of T. We will denote this σ-algebra by $\mathcal{K}$. We will use the same notation to denote the space $L^2(X, \mathcal{K}, \mu)$. In this section we offer a contribution to the Furstenberg second question which is based on spectral theory and the analysis of the speed of convergence in the Wiener Wintner ergodic theorem developed in [A1] and [A2]. As pointed out to us by one of the referees recent results on the pointwise convergence of $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n x)$ have appeared in [LRR]. In this later paper the notion of weak disjointness of two systems is introduced. Following the referees suggestions we will compare this property to the tools and results we used in our earlier version. Thus we will show in the case where T and S commute that our Proposition 5 is not covered by the weak disjointness property. We will provide examples (2 and 4) showing that this proposition holds for systems that are not weakly disjoint. Lastly we look at the case where T and S do not necessarily commute.

Case 1: T and S commute.

The commutativity of T and S is easily checked when one is a power of the other. In response to the original question of H. Furstenberg, J. Bourgain proved that if T is ergodic then the averages

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

converge a.e. for all bounded functions f and g and all integers a. We used this result in Section 1 during the proof of theorems 3 and 4. Before proceeding further, we would like to make several remarks on his proof.

Remarks

The study of the speed of convergence in the Wiener Wintner ergodic theorem [A1], [A2] leads to a simplification of Bourgain's proof for a large class of dynamical systems. One can show (as in [A3] for $p = 2$) that for every ergodic dynamical system and for each p with ($2 \leq p < \infty$) there exists a continuous increasing function G with $\lim_N G(N) = \infty$ and a dense set of functions $\mathcal{F}$ in $L^p \cap \mathcal{K}^{\perp}$, such that for each $f \in \mathcal{F}$

$$\sup_{\epsilon} \|\frac{1}{N}\sum_{n=1}^{N} f \circ T^n e^{2\pi i n\epsilon}\|_p \leq \frac{C_f}{G(N)}.$$

For those dynamical systems for which $G(N) \geq (\frac{N^\alpha}{\log(N)})$ where $p > \frac{1}{\alpha}$ the proof of the a.e double recurrence can be greatly simplified. For such dynamical systems, the speed of convergence of the averages

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

can also be estimated for a dense set of functions. Examples of dynamical systems with this last property are K-automorphisms , Abramov systems [A2], and some transformations with singular spectrum .

The study in L^2 of the "good" function G leads to classes of dynamical systems which are not characterized by the entropy or the spectrum of the transformation. There are weakly mixing dynamical systems [A2] called Gaussian Dynamical Systems, for which $G(N) \geq$

$(\log(N+1))^{(1+\alpha)}$ but $G(N) \leq N^{\beta}$ for infinitely many N for any $\beta > 0$. We are not aware of examples for which $G(N)$ goes to infinity much faster than the logarithmic rate $(\log(N+1))^{(1+\alpha)}$.

The result in [Bou] covers the case where the centralizer of one transformation is only composed of its powers. Such is the case of Chacon's transformation. We are interested in situations where the centralizer contains more than the powers of the transformation.

The next result is a contribution to the convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n x)$ when T and S commute. It is motivated by the results in [A2] and obtained with a simple application of van der Corput's Lemma [KN].

We recall that if U is a unitary operator, then $\hat{\sigma}_{f,g}(n) = \langle U^n f, g\rangle$ is the n-th Fourier coefficient of the measure $\sigma_{f,g}$. We showed in [A2] (Proposition 3) that if for all positive integers N we have

$$\sup_{\epsilon} \| \frac{1}{N}\sum_{n=1}^{N} F(T^n x)e^{2\pi i n\epsilon}\|_2 \leq \frac{C_F}{(\log(N+1))^{1+\gamma}},$$

then we have

$$\sup_{\|g\|_2\leq 1} \frac{1}{H}\sum_{h=1}^{H} |\hat{\sigma}_{F,g}(n)| \leq \frac{C_F}{(\log(H+1))^{\frac{1+\gamma}{2}}}.$$

We recall also that if T and S are two commuting ergodic transformations, then they have the same Kronecker factor. In particular, if T and S commute then T, S and TS^{-1} have the same Kronecker factor $\mathcal{K}$ when T, S and TS^{-1} are ergodic. In this case, we denote by r_j the orthonormal basis of $\mathcal{K}$ of eigenfunctions all with modulus 1 which correspond to the eigenvalues $e^{2\pi i\theta_j}$. The function r_1 is the constant function 1.

Proposition 5. *Assume the following:*

(1) *T, S and TS^{-1} are ergodic,*

(2) *there exist a constant α and a dense set of bounded functions $\mathcal{F}$ in the orthocomplement of the Kronecker factor such that for all positive integers N, all $F \in \mathcal{F}$ and all $H = O(N^d)$ for some d with $0 < d < 1$ we have*

$$\frac{1}{H}\sum_{h=1}^{H} \|\frac{1}{N}\sum_{n=1}^{N}(F.\overline{F \circ T^h}))(TS^{-1})^n x)\|_2 \leq \frac{C_F}{(\log(N+1))^{1+\alpha}}.$$

Then the averages

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

converge a.e. for all functions f and g in $L^2(\mu)$.

Proof. It is enough to prove that the averages

$$A_N(F,g)(x) = \frac{1}{N}\sum_{n=1}^{N} F(T^n x)\cdot g(S^n x)$$

converge for all $g \in L^\infty$ as the result would follow from this by approximation and the use of maximal inequalities.

Our goal will be to show that under the assumptions made, we have for all N,

$$\|\frac{1}{N}\sum_{n=1}^{N} F(T^n x)\cdot g(S^n x)\|_2 \leq \frac{C}{(\log(N+1))^{1+\gamma}} \tag{1}$$

for some $\gamma > 0$. By using sequences of the form $N+1 = [\rho^M]$ where $1 < \rho < \infty$ we can get the convergence of the sequence $A_N(F,g)(x)$. Note that this assumption implies the convergence of the series $\sum_{N=1}^{\infty} \|\frac{1}{N}\sum_{n=1}^{N} F(T^n x)\cdot g(S^n x)\|_2^2$ which implies the a.e convergence to zero of the averages $\frac{1}{N}\sum_{n=1}^{N} F(T^n x)\cdot g(S^n x)$. Without loss of generality, we can assume that the functions F and g are all bounded by 1.

To estimate the L^2 norm of $A_N(F,g)$ we will use van der Corput Lemma. We have for all $1 < H < [N/2]$,

$$\|A_N(F,g)\|_2 \leq C\cdot\left(\frac{1}{H} + \left|\frac{1}{(H+1)^2}\sum_{h=1}^{H}(H+1-h)\frac{1}{N}\sum_{n=1}^{N-h}\int F(T^n x)g(S^n x)\overline{F(T^{n+h}x)g(S^{n+h}x)}d\mu\right|\right)$$

Denoting by R the transformation TS^{-1}, we can rewrite the second term on the right of this inequality as

$$\left|\frac{1}{(H+1)^2}\sum_{h=1}^{H}(H+1-h)\int g(x)\overline{g(S^h x)}\frac{1}{N}\sum_{n=1}^{N-h}F(R^n x)\overline{F\circ T^h(R^n x)}d\mu\right|.$$

By using the fact that g is uniformly bounded and that $h \leq [N/2]$ we can dominate the last average by

$$C\cdot\frac{1}{H}\sum_{h=1}^{H}\|\frac{1}{N}\sum_{n=1}^{N}(F\cdot\overline{F\circ T^h})(R^n x)\|_2.$$

With assumption (2) this gives us equation (1) for $H = O([N^d])$. □

Remarks

In reference to the assumptions for Proposition 5, we do not need to assume that the functions F in $\mathcal{F}$ are bounded. If F is just in L^2, then the second assumption can be replaced by the following

$$\frac{1}{H}\sum_{h=1}^{H}\|\frac{1}{N}\sum_{n=1}^{N}(F\cdot\overline{F\circ T^h}))(TS^{-1})^n x)\|_1 \leq \frac{C_F}{(\log(N+1))^{1+\alpha}}.$$

Secondly, we can decompose the function $F\cdot\overline{F\circ T^h}$ into the sum of three orthogonal functions $\int F\cdot\overline{F\circ T^h}d\mu$, $P_{\mathcal{K}^\perp}(F\cdot\overline{F\circ T^h})$ and $P_{\mathcal{K}'}(F\cdot\overline{F\circ T^h})$ where $\mathcal{K}'$ denotes the closed linear span of the functions r_j , $2 \leq j < \infty$. The second assumption is then equivalent to the set of the following three statements (with the same condition on H as before);

$$\frac{1}{H}\sum_{h=1}^{H}\left|\int F\cdot(\overline{F\circ T^h})d\mu\right| \leq \frac{C_F}{(\log(N+1))^{1+\alpha}}. \tag{2}$$

$$\frac{1}{H}\sum_{h=1}^{H}\|\frac{1}{N}\sum_{n=1}^{N}P_{\mathcal{K}^\perp}(F\cdot\overline{F\circ T^h})(TS^{-1})^n x)\|_2 \leq \frac{C_F}{(\log(N+1))^{1+\alpha}}. \tag{3}$$

$$\frac{1}{H}\sum_{h=1}^{H}\|\frac{1}{N}\sum_{n=1}^{N}P_{\mathcal{K}'}(F\cdot\overline{F\circ T^h})(TS^{-1})^n x)\|_2 \leq \frac{C_F}{(\log(N+1))^{1+\alpha}}. \tag{4}$$

The study made in [A2] shows that, for many dynamical systems, we can find a set of bounded functions $\mathcal{F}$ whose linear span is dense in $\mathcal{K}^\perp$ and a positive number $\gamma > 1$ such that

$$\sup_{\epsilon}\|\frac{1}{H}\sum_{h=1}^{H}F\circ T^h e^{2\pi i h\epsilon}\|_2 \leq \frac{C_F}{(\log(H+1))^{1+\gamma}} \tag{5}$$

for all F in $\mathcal{F}$ and all positive integers H. By Proposition 3 in [A2] this last equation implies

$$\frac{1}{H}\sum_{h=1}^{H}\left|\int F\cdot\overline{F\circ T^h})d\mu\right| \leq C\cdot\sup_{\|g\|_2\leq 1}\frac{1}{H+1}\sum_{h=1}^{H}|\hat{\sigma}_{F,g}(h)|.$$

The same proposition shows that the last term on the right is then less than

$$\frac{C_F}{(\log(H+1))^{\frac{1+\gamma}{2}}}.$$

There are examples of ergodic transformations which are not necessarily weakly mixing for which the second assumption in the proposition is true. We will check this by using the set of the three inequalities (2), (3) and (4) given above.

Example 1:

The transformations

$$T : (x, y) \to (x + \alpha, 2x + y)$$

and

$$S : (x, y) \to (x + \frac{\alpha}{2}, x + y)$$

on the 2-torus where α is irrational with finite type η. They commute and neither is a power of the other. If we let

$$R = TS^{-1} : (x, y) \to (x + \frac{\alpha}{2}, x + y - \frac{\alpha}{2}),$$

then simple computations show that

$$R^n(x, y) = (x + n\alpha/2, nx + y + n(n-3)\alpha/4)$$

for all n and that

$$T^h(x, y) = (x + h\alpha, 2hx + h(h-1)\alpha + y).$$

The Kronecker factor is the set of functions depending on the first coordinate x. Hence, $\mathcal{K}^{\perp}$ is spanned by the functions $F(x, y) = e^{2\pi ipx} \cdot e^{2\pi iqy}$ where q is not equal to zero. One can see that the functions $F(x, y) \cdot \overline{(F \circ T^h)}(x, y)$ belong to $\mathcal{K}$ as

$$F(x, y) \cdot \overline{(F \circ T^h)}(x, y) = e^{-2\pi i(ph+qh(h-1)\alpha}e^{-2\pi iq2hx}.$$

This trivially shows that inequality (3) is satisfied.

For the same function F, $\int F(x, y) \cdot \overline{(F \circ T^h)}(x, y) dm \times m = 0$, so inequality (2) is true.

Finally, it remains to check inequality (4). We have

$$|\frac{1}{N}\sum_{n=1}^{N} P_{\mathcal{K}'}(F \cdot F \circ T^h)(R^n x)| = |\frac{1}{N}\sum_{n=1}^{N} e^{2\pi iqhn\alpha}|.$$

The second term is dominated by

$$C \cdot \frac{1}{N}\frac{1}{\sin(\pi\langle qh\alpha\rangle)}.$$

Using the fact that α is of finite type η and the estimate on [KN] p. 123, we have (see the proof of Theorem 2 in [AN])

$$\frac{1}{H}\sum_{h=1}^{H}\frac{1}{N}\frac{1}{\sin(\pi\langle qh\alpha\rangle)} \leq \frac{C}{N^t}$$

for some t where $0 < t < 1$ for an appropriate choice of H ($H = [N^r]$ with $0 < r < \min\{1, \frac{1}{\eta-1}\}$). This proves that inequality (4) is satisfied.

The conditions being compatible, we have an example of transformations T and S satisfying the assumptions of Proposition 5. We can observe that the transformations having quasi discrete spectrum the pointwise convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n x)$ follows also from the results in [LRR]. This example shows that Proposition 5 holds for some transformations with zero entropy.

If the transformations are weakly mixing, then the only assumption that matters is the second and in this case out of the three inequalities (2), (3) and (4), only (2) and (3) will count. Such is the case in the next example.

Example 2:

We consider a weakly mixing dynamical system T on the probability measure space $(X, \mathcal{B}, \mu)$ of logarithmic class α ([A2]) and positive entropy (for instance a K-automorphism). This means that we can find a dense set $\mathcal{F}$ of bounded functions F in $\mathcal{K}^{\perp} = \mathbb{C}^{\perp}$ such that

$$\sup_{\epsilon}\|\frac{1}{N}\sum_{n=1}^{N} F(T^n x)e^{2\pi i n\epsilon}\|_2 \leq \frac{C_F}{(\log(N+1))^{1+\alpha}}.$$

We define two commuting invertible measure preserving transformations on $(X, \mathcal{B}, \mu)^{\mathbb{Z}}$. The first one is simply the forward shift S with the product measure $\mu^{\mathbb{Z}}$. The second $\hat{T}$ is defined pointwise as $(\hat{T}(x_n))_n = (T(x_n))_n$. (Let us note that the system $((X, \mathcal{B}, \mu)^{\mathbb{Z}}, \hat{T})$, as well as the K-automorphism $((X, \mathcal{B}, \mu)^{\mathbb{Z}}, S)$, is also of logarithmic class α. To show that the assumptions of Proposition 5 apply to this commuting system we just need to prove that functions of the form

$$F(x) = \prod_{i=1}^{L} F_i(x_i)$$

satisfy the second assumption of Proposition 5. In this product of functions the F_i are either constant or equal to one of the functions in $\mathcal{F}$. At least one of the functions F_i belongs to $\mathcal{F}$. Let us denote it by F_{i_0}. We need to compute

$$\frac{1}{H}\sum_{h=1}^{H} \| \frac{1}{N}\sum_{n=1}^{N} (F \cdot \overline{F \circ T^h}))(TS^{-1})^n x)\|_2.$$

We have

$$(F \cdot \overline{F \circ T^h})((TS^{-1})^n x) = \prod_{i=1}^{L} F_i(T^n(x_{i-n}))\overline{F_i}(T^{h+n})(x_{i-n}).$$

An application of van der Corput Lemma leads to the estimate

$$C(\frac{1}{M} + \frac{1}{M}\sum_{m=1}^{M} \left| \int \prod_{i=1}^{L} F_i(x_i)\overline{F_i(T^h(x_i))}\,\overline{F_i}(T^m(x_{i-m}))F_i(T^{h+m}(x_{i-m}))d\mu^{\mathbb{Z}} \right|$$

where $m \leq M \leq N$. Assuming (without loss of generality) that $M > 2L$, in the sum from 1 to M we can focus only on the terms from $2L + 1$ to M. In this case, the integral

$$\int \prod_{i=1}^{L} F_i(x_i)\overline{F_i(T^h(x_i))}\,\overline{F_i}(T^m(x_{i-m})) \cdot F_i(T^{h+m}(x_{i-m}))d\mu^{\mathbb{Z}}$$

is equal to the product of integrals

$$\left(\int \prod_{i=1}^{L} F_i(x_i)\overline{F_i(T^h(x_i))}d\mu^{\mathbb{Z}}\right) \cdot \left(\int \overline{F_i}(T^m(x_{i-m}))F_i(T^{h+m}(x_{i-m}))d\mu^{\mathbb{Z}}\right)$$

because of the independence of the variables x_i , $1 \leq i \leq L$ and x_{i-m}. This product is equal to

$$\left|\int \prod_{i=1}^{L} F_i(x_i)\overline{F_i(T^h(x_i))}d\mu^{\mathbb{Z}}\right|^2$$

which is less than

$$C \cdot |\int F_{i_0}(x_{i_0})\overline{F_{i_0}}(T^h(x_{i_0}))d\mu|^2.$$

Averaging on h and using once more Proposition 3 in [A2], we obtain the second assumption in Proposition 5.

As the systems have positive entropy they are not weakly disjoint ([LRR], Proposition 5.2). Thus Proposition 5 is not covered by this property. Another example of this kind is given below (see example 4).

Example 3: K-systems

We recall that T and S generate a K-system (see [C] for the notation we use below) if there exists a measurable partition $\mathcal{A}$ of X such that

(1) $\displaystyle\bigvee_{(n,p)\in\mathbb{Z}^2} T^nS^p(\mathcal{A}) = \mathcal{B}$.
(2) For the lexicographic order on $\mathbb{Z}^2$, $\leq$, we have $(n',p') \leq (n,p) \Rightarrow T^{n'}S^{p'}\mathcal{A} \leq T^nS^p\mathcal{A}$.
(3) $\displaystyle\bigwedge_n S^{-n}\mathcal{A} = T^{-1}\mathcal{A}_S$.
(4) $\displaystyle\bigwedge_n T^{-n}\mathcal{A}_S = \{X, \emptyset\}$.

It is enough to show that each function $\mathbf{1}_A - \mu(A)$ for $A \in T^kS^l\mathcal{A}$ can be approximated in L^2 norm by functions F_A that satisfy the second assumption of Proposition 5. Our goal is to show that functions of the form

$$F_A = \mathbf{1}_A - \mathbb{E}(\mathbf{1}_A | T^t S^s \mathcal{A})$$

for $(t, s) \in \mathbb{Z}^2$ work. The method is similar to the one used in [A1]. As in the 1-dimensional case, one can see that

$$\mathbb{E}[\mathbf{1}_A | T^t S^s \mathcal{A}] \circ (T^m S^{-m}) = \mathbb{E}[\mathbf{1}_A \circ (T^m S^{-m}) | T^{t-m} S^{s+m} \mathcal{A}].$$

For simplicity we assume that $A \in \mathcal{A}$. Applying van der Corput Lemma, we estimate

$$\frac{1}{H} \sum_{h=1}^{H} \| \frac{1}{N} \sum_{n=1}^{N} (F_A \cdot \overline{F_A \circ T^h}))(TS^{-1})^n x) \|_2 \tag{6}$$

We just need to concentrate on

$$C \cdot \frac{1}{M} \sum_{m=1}^{M} \left| \int (F_A).(F_A \circ T^h)(F_A \circ T^m S^{-m})(F_A \circ T^{h+m} S^{-m}) d\mu \right| \tag{7}$$

for all $M < N$. The functions F_A, $F_A \circ T^h$, $F_A \circ T^m S^{-m}$, $F_A \circ T^{h+m} S^{-m}$ are respectively $\mathcal{A}$, $T^{-h}\mathcal{A}$, $T^{-m} S^m \mathcal{A}$ and $T^{-(h+m)} S^m \mathcal{A}$ measurable. We distinguish two cases.

Case I: Assume $t \geq 0$.

$$\int (F_A) \cdot (F_A \circ T^h)(F_A \circ T^m S^{-m})(F_A \circ T^{h+m} S^{-m}) d\mu$$

$$= \int \left(\mathbb{E}(\mathbf{1}_A | \mathcal{A}) - \mathbb{E}\left[\mathbb{E}(\mathbf{1}_A | T^t S^s \mathcal{A}) | \mathcal{A} \right] \right) \cdot (F_A \circ T^h)(F_A \circ T^m S^{-m})(F_A \circ T^{h+m} S^{-m}) d\mu$$

Because in this case $\mathcal{A} \subset T^t S^s \mathcal{A}$, the above integral is equal to zero.

Case II: Assume $t < 0$. By assuming that $M = h$, which is possible as M is any positive integer less than N, we know that $T^{-h}\mathcal{A} \subset T^{-m} S^m \mathcal{A}$. If we take the conditional expectation with respect to $T^{-m} S^m \mathcal{A}$ of F_A, we obtain

$$\mathbb{E}(\mathbf{1}_A | T^{-m} S^m \mathcal{A}) - \mathbb{E}[\mathbb{E}(\mathbf{1}_A | T^t S^s \mathcal{A}) | T^{-m} S^m \mathcal{A}].$$

Thus, if $m > -t$, this difference is zero. Therefore in this case, we have at most $|t|$ non zero terms.

Combining Cases I and II, we see that equation (7) is dominated by $C \cdot \frac{1}{h} \cdot |t|$. Summing on h we get for equation (6) the upper bound $C \cdot |t| \frac{1}{H} \sum_{h=1}^{H} \frac{1}{h}$ which is less than $C \cdot |t| \cdot \frac{\log H}{H}$. For $H = [N^d]$ for any d with $0 < d < 1$ we can claim that the second assumption of Proposition 5 is true for a K-system.

Example 4:

Consider a finite space $X = \{x_1, x_2, \cdots, x_r\}$ with masses p_i, $1 \leq i \leq r$, $\sum_{i=1}^{r} p_i = 1$ We consider the Bernoulli shift U based on $X^{\mathbb{Z}}$ and m the countable product of the measure defined on X. We take $T = U^2$ and $S = U$. Then TS^{-1} is equal to U.. The functions F that depend on finitely many coordinates and with zero integral form a dense set of functions in $C^{\perp}$. We need to check the conditions (2), (3) and (4) for such functions that we can assume real and bounded. For h large enough the functions F and $F \circ T^h$ are independent so the integral $\int F.F \circ T^h dm = 0$. Hence (2) is true. As $\mathcal{K}' = \mathbb{C}$ for the same reason (4) is also true. It remains to verify (3).

For each N we have

$$\|\frac{1}{N} \sum_{n=1}^{N} F.F \circ T^h\|_2 = \frac{1}{N^2} \sum_{n=1}^{N} \int F^2.[F \circ U^{2h}]^2 dm + 2 \frac{1}{N^2} \sum_{j<l} \int (F.F \circ U^{2h})((F.F \circ U^{2h}) \circ U^{l-j} dm.$$

The first term is bounded by $\frac{C}{N}$. For h large and $l - j$ large enough the functions F, $F \circ U^{2h}$, $F \circ U^{l-j}$ and $F \circ U^{2h+l-j}$ are independent. Thus many of the integrals

$$\int (F.F \circ U^{2h})((F.F \circ U^{2h}) \circ U^{l-j} dm$$

are in fact equal to zero. Putting these remarks together one can see that (3) is also true. Thus we have found two systems of positive entropy that satisfy Proposition 5. However there are not weakly disjoint because they have positive entropy.

Case 2: T and S do not necessarily commute.

In [Be] it is shown that we cannot expect the weak convergence of the averages

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

when T and S do not commute (see Example 7.1). When the systems are disjoint, the convergence has also been studied in [Be]. Here we answer a question raised by one of the referees.

Theorem 6. *Assume that T and S are ergodic and the spectrum of their restrictions to the orthocomplement of their maximal quasi-discrete factors is mutually singular. Then the averages*

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

converge a.e. for all functions f and g in $L^2(\mu)$.

We do not give a proof of this result as it is a consequence of a more general result that we prove below. First we introduce the following definition motivated by [LRR].

Definition 1. *Let $(X, \mathcal{B}, \mu, T)$ and $(X, \mathcal{B}, \mu, S)$ be two measure preserving systems on the same finite measure space. We will say that a function $f \in L^2(\mu)$ has the weak disjointness property with respect to (T, S) if for each function $g \in L^2(\mu)$ there exists a set A of full* *measure and a set B of full measure such that the averages*

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(S^n y)$$

converge for each $x \in A$ and $y \in B$.

Examples of functions with the weak disjointness property are the eigenfunctions of T. Indeed if $f(Tx) = e^{2\pi i\lambda} f(x)$, then the averages $\frac{1}{N}\sum_{n=1}^{N} e^{2\pi in\lambda} g(S^n x)$ can be shown to converge by applying Birkhoff's pointwise ergodic theorem to the product of the rotation $R_\lambda \times S$. We denote by $\mathcal{WD}_T$ the closed linear space in $L^2(\mu)$ of the functions with the weak disjointness property with respect to (T, S). In the same way we denote by $\mathcal{WD}_S$ the closed linear space in $L^2(\mu)$ of the functions with the weak disjointness property with respect to (S, T). It is simple to check that $\mathcal{WD}_T$ and $\mathcal{WD}_S$ are respectively T and S invariant. The following example motivates our next definition.

Example 5: The weak disjointness property was introduced without assuming the commutativity of the transformations T and S. However if one looks at the situation where T and S are the same Bernoulli transformation then we have a system which is not weakly selfdisjoint because of the entropy. However the averages

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

converge a.e.. This is a consequence of a simple application of Birkhoff's theorem to the function $f \times g$. So we have the pointwise convergence for all function $f, g \in L^2$ even if the systems are not weakly disjoint.

Definition 2. *Let $(X, \mathcal{B}, \mu, T)$ and $(X, \mathcal{B}, \mu, S)$ be two measure preserving systems on the same finite measure space. We will say that a function $f \in L^2(\mu)$ has the property A_T with* *respect to* (T, S) *if for each function* $g \in L^2(\mu)$ *the averages*

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

converge a.e.

It is simple to check that the closed linear span A_T (resp. A_S) of functions $f \in L^2$ with the property A_T (resp. A_S) is T (resp. S) invariant. Furthermore $\mathcal{WD}_T$ is by definition a subset of A_T. The example 5 shows that $\mathcal{WD}_T$ can be a strict subset of A_T. In that example we actually have $A_T = L^2$ while $\mathcal{WD}_T \neq L^2$.

Now we can state the following result.

Theorem 7. *Assume that* T *and* S *are ergodic and the spectrum of their restrictions to the orthocomplement of* A_T *and* A_S *is mutually singular. Then the averages*

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

converge a.e for all functions f *and* g *in* $L^2(\mu)$*.*

Proof. We can decompose each function f and g into the sum of two functions $f = f_1 + f_2$ and $g = g_1 + g_2$, with $f_1 \in A_T$, $f_2 \in A_T^{\perp}$, $g_1 \in A_S$ and $g_2 \in A_S^{\perp}$. By definition the averages

$$\frac{1}{N}\sum_{n=1}^{N} f_1(T^n x) g(S^n x) \text{ and } \frac{1}{N}\sum_{n=1}^{N} f(T^n x) g_1(S^n x)$$

converge a.e..

Let us show why the averages $\frac{1}{N}\sum_{n=1}^{N} f_2(T^n x) g_2(S^n x)$ converge a.e..

The convergence follows from the Affinity Principle [CKM]. More precisely, for μ a.e x the sequences $a_n = f_2(T^n x)$ and $b_n = g_2(S^n x)$ have a correlation in a sense that for each h the averages

$$\frac{1}{N}\sum_{n=1}^{N} a_n.\overline{a_{n+h}} \text{ and } \frac{1}{N}\sum_{n=1}^{N} b_n.\overline{b_{n+h}}$$

converge. Because of the ergodicity of T and S, these limits are the h-th Fourier coefficients of the spectral measures σ_{f_2} and σ_{g_2} of f_2 and g_2. As these measures are mutually singular, we have

$$\lim_N |\frac{1}{N}\sum_{n=1}^{N} f_2(T^n x) g_2(S^n x)| = 0.$$

□

Remarks

(1) One can conclude from (4.1) and Proposition 4.1 in [LRR] that the maximal quasi discrete factor for each ergodic dynamical system T is contained in $\mathcal{WD}_T$ and hence in A_T no matter what ergodic transformation one takes for S. Thus Theorem 6 is a consequence of Theorem 7.

(2) An example of transformations satisfying the assumption of Theorem 7 can be given by the product of the map

$$T(x, y) = (x + \alpha, x + y)$$

on the 2-torus and a K automorphism. For the second map we can take the product of the map

$$S(x, y) = (x + \alpha, y + \beta x)$$

with any weakly mixing transformation with singular spectrum. On the 2-torus the dynamical systems associated with the maps T and S are not disjoint because of their common non-trivial Kronecker factor given by the set of functions depending only on x. As indicated in [M] if β is irrational and α is irrational with unbounded partial quotients, then S is ergodic with singular spectrum in the orthocomplement of the Kronecker factor. The map T has Lebesgue spectrum in the orthocomplement of the functions depending on the first coordinate. This last known statement can be shown by computing directly the Fourier coefficients of the spectral measure of functions of the form $e^{2\pi ipx}e^{2\pi iqy}$ where p and q are integers and q is not equal to zero.

REFERENCES

[A1] **I. Assani**: "Wiener Wintner Dynamical Systems," *Erg. Th. and Dyn. Syst.*, 23, 1637-1654, 2003.

[A2] **I. Assani**:"Spectral characterization of Wiener Wintner Dynamical Systems," *details of [A4], Erg. Th. and Dyn. Syst.*24, 2, 347-365, 2004.

[A3] **I. Assani**:"Spectral characterization of Ergodic Dynamical Systems," *Contemporary Mathematics*,vol. 284, 13-22, 2001.

[A4] **I. Assani**:"Spectral characterization of Wiener Wintner Dynamical Systems," *Note de Comptes Rendus de l'Acad. des Sciences Paris*,t.332, Serie I, 321-324, 2001

[AN] **I. Assani and K. Nicolaou**:"Properties of Wiener-Wintner Dynamical Systems," *Bull. Soc. Math. France*, 129, (3), 361-377, 2001.

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

[Boi] **D. Boivin**:"Récurrence multiple et contractions linéaires," *C. R. Acad. Sci. Paris Ser. I Math.*, 305, 653-656,1987.

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

[Br] **A. Brunel**:"Sur les mesures invariantes," *Zeitschr. Wahrs.* , 5, 300-303, 1966.

[C] **J.P. Conze**:"Entropie d'un groupe abélien de transformations," *Zeitschr. Wahrs.*, 25, 11-30, 1972.

[CKM] **J. Coquet, T. Kamae, M. Mendes France**:"Sur la mesure spectrale de certaines suites arithmétiques," *Bull. Soc. Math. France*, 105, 369-384, 1977.

[F] **H. Furstenberg**: *Personal Communication*

[Ka] **C. Huen Kan**:"Ergodic properties for Lamperti operators," *Can J. Math*, 30, 1206-1214, 1978.

[KN] **L. Kuipers and H. Niederreiter**:"Uniform distribution of sequences," *John Wiley and Sons*, 1974.

[Kr] **U. Krengel**:"Ergodic Theorems," *De Gruyter, 1985.*

[LRR] **E. Lesigne, B. Rittaud and T. de la Rue**:"Weak disjointness of measure-preserving dynamical systems," *Erg. Theor. and Dyn. Syst.*, 23, 1173-1198, 2003.

[M] **H. Medina**:"Spectral types of unitary operators arising from irrational rotations on the circle group," *Mich. Math. J.*, 41, 39-49, 1994.

[O] **D. Ornstein**:"On invariant measures," *Bull. Amer. Math. Soc.*, 66, 297-300, 1960.