ON THE POINTWISE CONVERGENCE OF MULTIPLE ERGODIC AVERAGES

E. H. EL ABDALAOUI

Abstract. It is shown that the multiple ergodic averages of commuting invertible measure preserving transformations of a Lebesgue probability space converge almost everywhere provided that the maps are weakly mixing with an ergodic extra condition.

1. Introduction

The purpose of this note is to give a partial affirmative answer to the well-known open problem of the pointwise convergence of the Furstenberg ergodic averages. This problem can be formulated as follows: let \( k \geq 2 \), \((X, \mathcal{B}, \mu, T_i)_{i=1}^k\) be a finite family of dynamical systems where \( \mu \) is a probability measure, \( T_i \) are commuting invertible measure preserving transformations and \( f_1, f_2, \ldots, f_k \) a finite family of bounded functions. Does the following averages

\[
\frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k f_i(T_i^n x)
\]

convergence almost everywhere?

The classical Birkhoff theorem correspond to the case \( k = 1 \). The case \( k = 2 \) with \( T_i = T^{p_i} \), and \( p_i \in \mathbb{N}^\ast \) for each \( i \), is covered by Bourgain double ergodic theorem [11].

The \( L^2 \) version of this problem has been intensively studied and the topics is nowadays very rich. These studies were originated in the seminal work of Furstenberg on Szemeredi’s theorem in [15]. Furstenberg-Katznelson-Ornstein in [16] proved that the \( L^2 \)-norm convergence holds for \( T_i = T^p \) and \( T \) weakly mixing. Twenty three years later, Host and Kra [19] and independently T. Ziegler [31] extended Furstenberg-Katznelson-Ornstein result by proving that for any transformation preserving measure, the \( L^2 \)-norm convergence for \( T_i = T^p \) holds. In 1984, J. Conze and E. Lesigne in [12] gives a positive answer for the case \( k = 2 \). Under some extra ergodicity assumptions, Conze-Lesigne result was extended to the case \( k = 3 \) by Zhang in [30] and for any \( k \geq 2 \) by Frantzikinakis and Kra in [14]. Without these assumptions, this result was proved by T. Tao in [29]. Subsequently, T. Austin in [6] gives a joining alternative proof of Tao result, and recently, M. Walsh

Date: June 11, 2014.

2010 Mathematics Subject Classification. Primary 37A05, 37A30, 37A40; Secondary 42A05, 42A55.

Key words and phrases. ergodic theorems, non-singular maps, Furstenberg ergodic averages, Bourgain double recurrence theorem, Birkhoff theorem, Frobenius-Perron operator, Koopman operator, Lebesgue probability space, strictly ergodic topological model, joining.
extended Tao result by proving that the \( L^2 \)-norm convergence holds for the maps \((T_i)^k_{i=1}\) generate a nilpotent group \([33]\). This solves a Bergelson-Leibman conjecture stated in \([10]\). Therein, the authors produced a counter-examples of maps generated a solvable group for which the \( L^2 \)-norm convergence doesn’t holds.

For the pointwise convergence, partial results were obtained in \([10]\) and \([33]\) under some ergodic and spectral assumptions. In a very recent preprint \([5]\), I. Assani observe that the action of the maps \( T_i \) induced a dynamical system \((X^k, \otimes^k_{i=1} B, \nu, \phi)\) where \( \phi(x_1, \ldots, x_k) = (T_i x_i)^k_{i=1} \) is a non-singular map with respect to the probability measure \( \nu \) given by

\[
\nu(A_1 \times A_2 \times \cdots \times A_k) = \frac{1}{3} \sum_{n \in \mathbb{Z}} \frac{1}{2^{\lvert n \rvert}} \mu_\Delta(\phi^{-n}(A_1 \times A_2 \times \cdots \times A_k)),
\]

with \( \mu_\Delta \) is a diagonal probability measure on \( X^k \) define on the rectangle \( A_1 \times A_2 \times \cdots \times A_k \) by

\[
\mu_\Delta(A_1 \times A_2 \times \cdots \times A_k) = \mu(A_1 \cap A_2 \cap \cdots \cap A_k).
\]

It is easy to see that the Radon-Nikodym derivative of the pushforward measure of \( \nu \) under \( \phi \) verify

\[
\frac{1}{2} \leq \frac{d\nu \circ \phi}{d\nu} \leq 2.
\]

We associate to \( \phi \) the Koopman operator \( U_{\phi} \) defined by \( U_{\phi}(\mathcal{F}) = \overline{\mathcal{F}} \circ \phi \), where \( \overline{\mathcal{F}} \) is a measurable function on \( X^k \). Since \( U_{\phi} \) maps \( L^\infty \) on \( L^\infty \) and \( \phi \) is non-singular, the adjoint operator \( U_{\phi}^* \) acting on \( L_1 \) can be defined by the relation

\[
\int U_{\phi}(\mathcal{F}) d\nu = \int \mathcal{F} U_{\phi}^*(\overline{\mathcal{F}}) d\nu,
\]

for any \( \mathcal{F} \in L^\infty \) and \( \overline{\mathcal{F}} \in L^1 \). For simplicity of notation, we write \( \phi^* \) instead of \( U_{\phi}^* \) and \( \phi \) instead of \( U_{\phi} \) when no confusion can arise.

\( \phi^* \) is often referred to as the Perron-Frobenius operator or transfer operator associated with \( \phi \). The basic properties of \( \phi^* \) can be found in \([21]\) and \([1]\). It is well known that the pointwise ergodic theorem for \( \phi \) can be characterized by \( \phi^* \). Indeed, Y. Ito established that the validity of the \( L^1 \)-mean theorem for \( \phi^* \) implies the validity of the pointwise ergodic theorem for \( \phi \) from \( L^1 \) to \( L^1 \) \([20]\). Moreover, the subject has been intensively studied by many authors (Ryll-Nardzewski \([20]\), Assani \([2]\), Assani-Woś \([3]\), Ortega Salvador \([23]\), R. Sato \([28]\, \[27]\)). Here, we will use and adapt the Ryll-Nardzewski approach \([20]\).

Notice that the \( L^2 \)-norm convergence implies that for any \( k \)-uple of Borel set \( (A_i)_{i=1}^k \), we have

\[
\lim_{N \to +\infty} \frac{1}{N} \sum_{n=1}^N \mu_\Delta(\phi^{-n}(\prod_{i=1}^k A_i)) = \mu^F(\prod_{i=1}^k A_i),
\]

where \( \mu^F \) is an invariant probability measure under the actions of \( (T_i) \). Following T. Austin, \( \mu^F \) is called a Furstenberg self-joining of \( (T_i) \) \([9]\, \[11]\, \[8]\). Therein, T. Austin stated the multiple recurrence theorem in the following form

\[
\forall(A_i)_{i=1}^k \in \mathcal{B}^k, \quad \mu^F(A_1 \times A_2 \times \cdots \times A_k) = 0 \implies \mu_\Delta(A_1 \times A_2 \times \cdots \times A_k) = 0,
\]
ON THE POINTWISE CONVERGENCE OF MULTIPLE ERGODIC AVERAGES

We further point out that if the maps $T_i = T^i$ and $T$ is a weakly mixing map, then $\mu^F = \otimes_{i=1}^k \mu_i$, that is, $\mu^F$ is singular with respect to $\mu_\Delta$ and this is not incompatible with the multiple recurrence theorem, since the absolutely continuity holds only one the sub-algebra generated by the rectangle Borel sets.

The proof of T. Austin is based on the description of $\mu^F$ in terms of the measure $\mu$ and various partially-invariant factors of $B$. This is done using a suitable extension of the originally-given system $(X, B, \mu, T_i)$. Here, we will further need the method used by Hansel-Raoult in [13] and its generalization to $\mathbb{Z}^d$ action obtain by B. Weiss in [32]. Therein, the authors gives a generalization of the Jewett theorem using the Stone representation theorem combined with some combinatorial arguments. The proof given by Weiss yields a “uniform” extension of Rokhlin towers lemma. This allows us, under a suitable assumption, to produce a strictly ergodic topological model for which we are able to show under some condition that the Furstenberg averages converge almost everywhere.

In summary, our proof is essentially based on two kind of arguments. On one hand, the Assani observation [5] combined with Ryll-Nardzewski theorem [26] and some basic fact from the non-singular ergodic theory, and on the other hand on the $C$-method introduced by Austin [6] combined with Hansel-Raoult-Weiss procedure [18].

Let us mention that it is obvious that our result is not yet a generalization of classical ergodic Birkhoff theorem and cannot be considered as a generalization of the Bourgain double ergodic theorem. The paper is organized as follows.

2. Main result

Let $(X, B, \mu)$ be a Lebesgue probability space, that is, $X$ is a Polish space (i.e. metrizable separable and complete), whose Borel $\sigma$-algebra $B$ is complete with respect to the probability measure $\mu$ on $X$. The notion of Lebesgue space is due to Rokhlin [25], and it is well known [22] that a Lebesgue probability space is isomorphic (mod 0) to ordinary Lebesgue space $([0, 1], C, \lambda)$ possibly together with countably many atoms, that is, there are $x_0, x_1, \cdots, \in X, X_0 \subset X, Y_0 \subset [0, 1)$, and $\phi : X_0 \cup \{x_i\} \to Y_0$, which is invertible such that the pushforward measure of $\mu$ under $\phi$ is $\lambda$ with $\mu(X_0 \cup \{x_i\}) = 1 = \lambda(Y_0) = 1$. Here, we will deal only with non-atomic Lebesgue space.

A dynamical system is given by $(X, B, \mu, T)$ where $(X, B, \mu)$ is a Lebesgue space and $T$ is an invertible bi-measurable transformation which preserves the probability measure $\mu$.

In this context, we state our main result as follows
Theorem 2.1. Let $k \in \mathbb{N}^*$ and $(X, B, \mu, T_i)_{i=1}^k$ be a finite family of dynamical systems where $(X, \mathcal{B}, \mu)$ is Lebesgue probability space, and assume that $T_1, T_2, \ldots, T_k$ are commuting weakly mixing transformations on $X$ such that for any $i \neq j$, the map $T_i \circ T_j^{-1}$ is ergodic. Then, for every $f_i \in L^\infty(\mu), i = 1, \ldots, k$, the averages
\[
\frac{1}{N} \sum_{n=1}^N \prod_{i=1}^k f_i(T_i^n x)
\]
converge almost everywhere to
\[
\prod_{j=1}^k \int_{T_j} f_j d\mu.
\]

The proof will follow from the Ryll-Nardzewski theorem \[26\] combined with the machinery of $C$-systems introduced by T. Austin. We remind that this machinery allows T. Austin to obtain a joining proof of the Tao theorem on the $L^2$-norm convergence of the Furstenberg ergodic averages. We will use the $C$-systems machinery in the section 3. Here, we recall the Ryll-Nardzewski theorem \[26\].

Theorem 2.2 (Ryll-Nardzewski \[26\]). Let $(X, B, \nu)$ be a probability measure and $\phi$ an invertible map on $X$ such that the pushforward measure of $\nu$ under $\phi$ is absolutely continuous with respect to $\nu$. Then the following conditions are equivalent.

(a) The operator $Tf = f \circ \phi$ satisfies the pointwise ergodic theorem from $L^1(\nu)$ to $L^1(\nu)$, that is, for any function $f \in L^1(\nu)$ there is a function $g \in L^1(\nu)$ such that
\[
\lim_{N \to +\infty} \frac{1}{N} \sum_{n=1}^N (T^nf)(x) = g(x) \nu\text{-a.e.}
\]

(b) There is a constant $K$ such that
\[
\limsup_{N \to +\infty} \frac{1}{N} \sum_{n=1}^N \nu(\phi^{-n} A) \leq K \nu(A),
\]
for each Borel set $A$.

The condition (b) is called Hartman condition, and the proof of Theorem \[2.2\] is essentially based on the notion of Mazur-Banach limit. Precisely, using the Mazur-Banach limit, Ryll-Nardzewski proved that there is a finite measure $\rho$ such that, for any Borel set $A$, we have

(i) $0 \leq \rho(A) \leq K \nu(A)$;
(ii) if $A = \phi^{-1} A$ then $\rho(A) = \nu(A)$;
(iii) $\rho(\phi^{-1} A) = \rho(A)$.

This insure that $\rho$ is $\phi$-invariant and we can apply the Birkhoff ergodic theorem to conclude. For more details and the rest of the proof, we refer the reader to the Ryll-Nardzewski paper \[26\]. Let us further point out that therein, Ryll-Nardzewski produce a counter-example for which the pointwise ergodic theorem in $L^1(\nu)$ doesn’t imply the ergodic theorem in $L^2(\nu)$. We remind that the ergodic theorem in $L^1(\nu)$ holds, if for any $f \in L^1(\nu)$, there is a function $g \in L^1(\nu)$ such that
\[
\left\| \frac{1}{N} \sum_{n=1}^N f(\phi^n(x)) - g(x) \right\|_1 \to 0.
\]
We notice that under our assumption, and following Ryll-Nardzewski ideas, one may choose a non-decreasing sequence of Borel set \((A_m)\) such that \(\lim A_m = X\). Hence, for any Borel set, the sequence \(\nu_m(A)\) defined by
\[
\nu_m(A) = \frac{1}{N} \sum_{n=0}^{N-1} \nu(\phi^n(A) \cap A_m)
\]
is a bounded sequence. But, it is an easy exercise to see that we can extend the operator \(\lim\) on the space of real bounded sequences to obtain a bounded operator on \(\ell^\infty\) by Hahn-Banach theorem. We denote such operator by \(\text{MBlim}\). This allows us to define a sequence of probability measure \(\nu_m\) on \(X\) given by
\[
\nu_m(A) = \text{MBlim} \left( \frac{1}{N} \sum_{n=0}^{N-1} \nu(\phi^n(A) \cap A_m) \right).
\]

It follows that if \(A = \phi(A)\) then, for any \(m \in \mathbb{N}\), we have \(\nu_m(A) = \nu(A \cap A_m)\). We further have, for any Borel set \(A\),
\[
\nu_m(A) \leq K \nu(A),
\]
and
\[
\nu_m(\phi(A)) = \text{MBlim} \left( \frac{1}{N} \sum_{n=0}^{N-1} \nu(\phi^{n+1}(A) \cap A_m) \right)
\]
\[
= \text{MBlim} \left( \frac{1}{N} \sum_{n=0}^{N-1} \nu(\phi^n(A) \cap A_m) - \frac{\nu(A \cap A_m)}{N} + \frac{\nu(\phi^N(A) \cap A_m)}{N} \right)
\]
\[
= \text{MBlim} \left( \frac{1}{N} \sum_{n=0}^{N-1} \nu(\phi^n(A) \cap A_m) \right)
\]
\[
= \nu_m(A),
\]
since
\[
\text{MBlim} \left( - \frac{\nu(A \cap A_m)}{N} + \frac{\nu(\phi^N(A) \cap A_m)}{N} \right) = 0.
\]
That is, \(\nu_m\) is invariant under \(\phi\). Now, the sequence \((\nu_m(A))\) is a bounded non-decreasing sequence. Therefore, we can put
\[
\rho(A) = \lim_{m \to +\infty} \nu_m(A),
\]
and it is easy to check that \(i), (ii)\) and \(iii)\) holds.

Remark 2.3. Ryll-Nardzewski proved Theorem 2.2 in the more general setting where \(\nu\) is a \(\sigma\)-finite measure.

3. \(L^2\)-norm convergence of Furstenberg averages

We start by proving the following

Theorem 3.1. Let \(k \in \mathbb{N}^\ast\) and \((X,B,\mu,T_i)_{i=1}^k\) be a finite family of dynamical systems where \(\mu\) is a probability measure space, and \(T_1,T_2,\ldots,T_k\) are commuting
weakly mixing transformations on $X$ such that for any $i \neq j$, the map $T_i \circ T_j^{-1}$ is ergodic. Then, for every $f_i \in L^\infty(\mu)$, $i = 1, \cdots, k$, the averages

$$\frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{k} f_i(T_i^n x)$$

converge in $L^2(X, \mu)$ to $\prod_{j=1}^{k} \int f_j d\mu$.

The proof is based on van der Corput trick and the $C$-sated systems method introduced by T. Austin. In our case the $C$-sated systems are trivial, and the van der Corput lemma can be stated in the following form

**Lemma 3.2** (van der Corput [9]). Let $(u_n)$ be a bounded sequence in a Hilbert space. Then,

$$\limsup \left\| \frac{1}{N} \sum_{n=0}^{N-1} u_n \right\|^2 \leq \limsup \frac{1}{H} \sum_{h=0}^{H-1} \limsup \left| \sum_{n=0}^{N} \langle u_{n+h}, u_n \rangle \right|.$$

As a simple consequence it follows that if

$$\lim_{H \to +\infty} \lim_{N \to +\infty} \frac{1}{H} \sum_{h=0}^{H-1} \sum_{n=0}^{N} \langle u_{n+h}, u_n \rangle = 0,$$

then

$$\left\| \frac{1}{N} \sum_{n=0}^{N-1} u_n \right\| \to 0.$$

We remind that if $(T_i)_{i=1}^{k}$ are a commuting maps on $X$ then the associated $C$-systems are the dynamical systems $(X, C, \mu, T_i)$, $i = 1, \cdots, k$, for which

$$C = \mathcal{I}_{T_1} \lor \mathcal{I}_{T_2T_1^{-1}} \lor \mathcal{I}_{T_3T_2^{-1}} \lor \cdots \lor \mathcal{I}_{T_kT_{k-1}^{-1}},$$

where, for any transformation $S$, $\mathcal{I}_S$ is the factor $\sigma$-algebra of $S$-invariant Borel sets, that is,

$$\mathcal{I}_S = \{ A : \mu(A \Delta S^{-1}A) = 0 \}.$$

Notice that under our assumption the $\sigma$-algebra $C$ is trivial. The key notion in the Austin proof is the notion of $C$-sated system defined as follows [7]

**Definition 3.3.** Let $k$ be a integer such that $k \geq 2$. The system $(X, \mathcal{B}, (T_i)_{i=1}^{k})$, $k \geq 2$ is $C$-sated if any joining $\lambda$ of $X$ with any $C$-system $Y$ is relatively independent over the largest $C$-factor $X_C$ of $X$, that is, for any bounded measurable function $f$ on $X$, we have

$$E_\lambda(f(x)|Y) = E_\lambda(E_X(f(x)|X_C)|Y),$$

Where $E_\lambda(\cdot|\cdot)$ is a conditional expectation operator.

We denote the expectation operator by

$$E(f) = \int f d\mu, \ f \in L^2(X).$$

Following this setting, the Tao $L^2$-norm convergence theorem can be stated as follows
Theorem 3.4 (Austin-Tao [6, 29]). Let \((X, \mathcal{B}, T_i)\), \(i = 1, \ldots, k\), be a \(C\)-sated system. Then, for any \(f_1, f_2, \ldots, f_k\) functions in \(L^\infty\),

\[
\mathbb{E}_X(f_1|X_x) = 0 \implies \left\| \frac{1}{N} \sum_{n=1}^{N} f_1 \circ T_1 \cdots f_k \circ T_k \right\|_2 \xrightarrow{N \to +\infty} 0.
\]

Now, we are able to give the proof of Theorem 3.1.

Proof of Theorem 3.1. We use induction on \(k\) to prove the theorem. The statement is obvious for \(k = 1\). For the case \(k = 2\), by our assumption the \(C\)-system is trivial and we can write

\[
\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^n_1 x) f_2(T^n_2 x) - \mathbb{E}(f_1) \mathbb{E}(f_2) \right\|_2 \\
\leq \left\| \frac{1}{N} \sum_{n=1}^{N} (f_1 - \mathbb{E}(f_1)) (T^n_1 x) f_2(T^n_2 x) + \frac{1}{N} \sum_{n=1}^{N} f_2(T^n_2 x) \mathbb{E}(f_1) - \mathbb{E}(f_1) \mathbb{E}(f_2) \right\|_2 \\
\leq \left\| \frac{1}{N} \sum_{n=1}^{N} (f_1 - \mathbb{E}(f_1)) (T^n_1 x) f_2(T^n_2 x) \right\|_2 + \left\| \frac{1}{N} \sum_{n=1}^{N} f_2(T^n_2 x) \mathbb{E}(f_1) - \mathbb{E}(f_1) \mathbb{E}(f_2) \right\|_2
\]

Hence, by the \(L^2\)-norm convergence and the von Neumann ergodic theorem, we have

\[
\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^n_1 x) f_2(T^n_2 x) - \mathbb{E}(f_1) \mathbb{E}(f_2) \right\|_2 \xrightarrow{N \to +\infty} 0.
\]

In the general case, applying the van der Corput Lemma 3.7, the \(L^2\) convergence of Furstenberg average of order \(k\), can be reduced in the class of \(C\)-systems to the case of \(k - 1\) commuting maps. In this case, the \(L^2\) convergence gives

\[
\mathbb{E}(f_1) = 0 \implies \left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^n_1 x) f_2(T^n_2 x) \cdots f_k(T^n_k x) \right\|_2 \xrightarrow{N \to +\infty} 0.
\]

Therefore, suppose that the result holds for some integer \(k \geq 1\), and assume that \((X, \mu, T_1, \cdots, T_l)\) is a system of order \(k + 1\). Then,

\[
\left\| \frac{1}{N} \sum_{n=1}^{N} f_1(T^n_1 x) f_2(T^n_2 x) \cdots f_{k+1}(T^n_{k+1} x) - \mathbb{E}(f_1) \mathbb{E}(f_2) \cdots \mathbb{E}(f_{k+1}) \right\|_2
\]

\[
= \left\| \frac{1}{N} \sum_{n=1}^{N} (f_1 - \mathbb{E}(f_1)) (T^n_1 x) f_2(T^n_2 x) \cdots f_{k+1}(T^n_{k+1} x) + \mathbb{E}(f_1) \frac{1}{N} \sum_{n=1}^{N} f_2(T^n_2 x) \cdots f_{k+1}(T^n_{k+1} x) - \mathbb{E}(f_1) \mathbb{E}(f_2) \cdots \mathbb{E}(f_{k+1}) \right\|_2
\]

\[
\leq \left\| \frac{1}{N} \sum_{n=1}^{N} (f_1 - \mathbb{E}(f_1)) (T^n_1 x) f_2(T^n_2 x) \cdots f_{k+1}(T^n_{k+1} x) \right\|_2 + \mathbb{E}(f_1) \left\| \frac{1}{N} \sum_{n=1}^{N} f_2(T^n_2 x) \cdots f_{k+1}(T^n_{k+1} x) - \mathbb{E}(f_2) \cdots \mathbb{E}(f_{k+1}) \right\|_2 \xrightarrow{N \to +\infty} 0.
\]

This complete the proof of the theorem. \(\square\)
As a consequence we deduce the following result

**Corollary 3.5.** Let \( k \in \mathbb{N}^* \) and \( (X, \mathcal{B}, \mu, T_i)_{i=1}^k \) be a finite family of dynamical systems where \( \mu \) is a probability measure space, assume that \( T_i, i = 1 \cdot \cdot \cdot k \) are commuting weakly mixing transformations on \( X \). Then, for every \( A_i \in \mathcal{B}, i = 1, \cdot \cdot \cdot, k \), the averages

\[
\frac{1}{N} \sum_{n=1}^N \mu_\Delta(T_1^{-n} A_1 \times T_2^{-n} A_2 \times \cdots \times T_k^{-n} A_k)
\]

converge to \( \mu(A_1)\mu(A_2)\cdots \mu(A_k) \).

Form this we deduce the following crucial lemma

**Lemma 3.6.** Let \( k \in \mathbb{N}^* \) and \( (X, \mathcal{B}, \mu, T_i)_{i=1}^k \) be a finite family of dynamical systems where \( \mu \) is a probability measure space, and assume that \( T_1, T_2, \cdots, T_k \) are commuting weakly mixing transformations on \( X \). Then, for every \( f_i \in L^\infty(\mu), i = 1, \cdots, k \), the averages

\[
\frac{1}{N} \sum_{n=1}^N \int \prod_{i=1}^k f_i(T_i^n x) d\nu
\]

converge to \( \prod_{i=1}^k \mu(f_i) \).

**Proof.** It is suffice to prove the lemma for any finite family of Borel set \( (A_i)_{i=1}^k \). Indeed, by Corollary 3.5, we have

\[
\frac{1}{N} \sum_{n=1}^N \mu_\Delta(\phi^{-n}(\overline{A}))
\]

converge to \( \bigotimes_{i=1}^k \mu(A_i) \), where \( \overline{A} = A_1 \times A_2 \cdots \times A_k \). This gives also that the convergence holds when replacing \( \mu_\Delta \) by \( \mu_{\phi^j} \Delta \), for any \( j \in \mathbb{Z} \), where \( \mu_{\phi^j} \Delta \) is the pushforward measure of \( \mu_\Delta \) under \( \phi^j \). Therefore, for any \( M \in \mathbb{N} \), the convergence holds for \( \nu_M \) where

\[
\nu_M(\overline{A}) = \frac{1}{3} \sum_{|n| \leq M} \frac{1}{2^{|n|}} \mu_{\phi^n} \Delta(\overline{A}).
\]

Now, let \( \varepsilon > 0 \) then there is a positive integer \( M_0 \) such that for any \( M \geq M_0 \), we have

\[
|\nu_M(\overline{A}) - \nu(\overline{A})| \leq \varepsilon, \forall \overline{A} \in \mathcal{A}^k.
\]

Hence, for any \( N \in \mathbb{N} \),

\[
\left| \frac{1}{N} \sum_{n=1}^N \nu_M(\phi^{-n}(\overline{A})) - \frac{1}{N} \sum_{n=1}^N \nu(\phi^{-n}(\overline{A})) \right| < \varepsilon.
\]

By letting \( N \) and \( M \) goes to infinity we obtain

\[
\bigotimes_{i=1}^k \mu(\overline{A}) - \varepsilon \leq \liminf_{N \to +\infty} \frac{1}{N} \sum_{n=1}^N \nu(\phi^{-n}(\overline{A})) \leq \limsup_{N \to +\infty} \frac{1}{N} \sum_{n=1}^N \nu(\phi^{-n}(\overline{A})) \leq \bigotimes_{i=1}^k \mu(\overline{A}) + \varepsilon.
\]
Since $\varepsilon$ was chosen arbitrarily, we conclude that
\[
\frac{1}{N} \sum_{n=1}^{N} \nu(\phi^{-n}(A)) \xrightarrow{N \to +\infty} \bigotimes_{i=1}^{k} \mu(A).
\]
This proves the lemma. \qed

4. Stone representation theorem and Furstenberg averages

Let us consider a dynamical system $(X, B, \mu, T)$. Then, there exists a countable algebra $A$ dense in $B$ for the pseudo-metric $d(A, B) = \mu(A \Delta B)$, $A$ and $B$ in $A$. We further have that $A$ separates the points of $X$, that is, for each $x, y \in X$ with $x \neq y$, there is $A \in A$ such that either $x \in A, y \notin A$ or else $y \in A, x \notin A$. Hence, by the Stone representation theorem, we associate to $A$ a Stone algebra $\hat{A}$ on the set $\hat{X}$ of all ultrafilter on $X$ such that for any $A \in A$, $\hat{A} = \{U_A \in \hat{X}/A \in U_A\}$. Consequently $\hat{A} = \{\hat{A}, A \in A\}$ is algebra of subsets of $\hat{X}$, which is isomorphic to $A$. $\hat{A}$ is called the Stone algebra.

Assuming that $T$ is ergodic, Hansel and Raoult proved in [18] that there is a dense and invariant countable algebra $A_T$ such that for any $A$ of $A_T$ we have
\[
\left\| \frac{1}{N} \sum_{n=1}^{N} 1_A(T^n x) - \mu(A) \right\|_\infty \xrightarrow{N \to +\infty} 0,
\]
that is, $A$ is a uniform ergodic set.

Their result was extended to the ergodic $\mathbb{Z}^d$-action by B. Weiss in [32]. Indeed, B. Weiss proved that if $\overline{T} = (T_i)_{i=1}^d$ is a generator of ergodic $\mathbb{Z}^d$-action then there is a dense and $\overline{T}$-invariant countable algebra $\hat{A}_T$ such that all elements $A$ in $\hat{A}_T$ are uniform ergodic sets, that is,
\[
\left\| \frac{1}{|R_n|} \sum_{i \in R_n} 1_A(\overline{T} i x) - \mu(A) \right\|_\infty \xrightarrow{n \to +\infty} 0,
\]
where $R_n$ is the square $\{i \in \mathbb{Z}^d : |i|_\infty \leq n\}$.

Applying a Stone representation theorem to Hansel-Raoult-Weiss algebra, and letting $\hat{X}$ be equipped with the topology which has $\hat{A}$ as a base of clopen sets. It follows that $\hat{X}$ is metrizable space (since $\hat{A}$ is countable), compact (by the standard ultrafilter lemma) and totally disconnected. $\hat{X}$ is called the Stone space of $A$. Furthermore, the probability $\mu$ induces a mapping $\mu'$ from $\hat{A}$ into $[0,1]$, such that
(1) for any $A \in \hat{A} \setminus \{\emptyset\}$, $\mu'(A) > 0$,
(2) $\mu'$ is $\sigma$-additive on $\hat{A}$, and $\mu'(\hat{X}) = 1$.

Hence, by Carathéodory’s extension theorem, $\mu'$ has an unique extension as a probability measure $\hat{\mu}$ on the Borel $\sigma$-algebra $\hat{B}$ generate by $\hat{A}$. We further have that for every non-empty open subset $O \subset \hat{X}$, $\hat{\mu}(O) > 0$. This gives in particular that for any $\hat{A} \in \hat{B}$, if $\hat{\mu}(\hat{A}) = 1$ then $\hat{A}$ is a dense in $\hat{X}$.

The $\mathbb{Z}^d$-action with generators $\overline{T}$ induces a $\mathbb{Z}^d$-action with generators $\widehat{T} = (\widehat{T}_i)_{i=1}^d$ such that, by construction and due to intrinsic properties of Lebesgue space, the
two action are isomorphic and for each \(i\), \(\hat{T}_i\) is a homeomorphism on \(\hat{X}\). We further have that for any \(\hat{A}\), the sequence of continuous function

\[
\frac{1}{|R_n|} \sum_{\pi \in R_N} 1_{\hat{A}} \circ \hat{T}^\pi
\]

converges uniformly to \(\hat{\mu}(\hat{A})\). Then, it follows by Birkhoff’s ergodic theorem that the \(\mathbb{Z}^d\)-action with generators \(\hat{T} = (\hat{T}_i)^d_{i=1}\) is strictly ergodic, that is, \(\hat{\mu}\) is a unique probability measure \(\hat{T}\)-invariant with \(\hat{\mu}(O) > 0\) for every nonempty open set \(O \subset \hat{X}\). The topological model \((\hat{X}, \hat{\mathcal{B}}, \hat{T})\) is called the Stone-Jewett-Weiss topological model. For the case \(d = 1\), we call it the Stone-Jewett-Hansel-Raoult topological model.

Finally, let us denote by \(\mathcal{B}(\mathcal{A})\) the Banach space of all scalar-valued functions that are uniform limits of sequences of \(\mathcal{A}\)-measurable step functions, equipped with the supremum norm. Then, \(\mathcal{B}(\mathcal{A})\) is isometrically isomorphic to \(\mathcal{C}(\hat{X})\), the Banach space of all continuous functions on the Stone space \(\hat{X}\).

5. Proof of the Main Result

Let \((\hat{X}, \hat{\mathcal{B}}, \hat{\mu}, \hat{T}) = (\hat{T}_i)_{i=1}^k\) be the Stone-Weiss topological model associated to \((X, \mathcal{B}, \mu, T)\), and for any \(A_1 \times A_2 \times \cdots A_k \in \mathcal{B}_k\), put

\[
\tilde{\nu}(A_1 \times A_2 \times \cdots A_k) = \frac{1}{3} \sum_{n \in \mathbb{Z}} \frac{1}{2^{|n|}} \hat{\mu}(\hat{T}^n(A_1 \times A_2 \times \cdots A_k)),
\]

where \(\hat{T}^n = (\hat{T}_i^n)_{i=1}^k\) and \(\hat{\mu}\) is the diagonal measure on \(\hat{X}^k\) associated to \(\hat{\mu}\).

From this, let us consider the non-singular dynamical system \((\hat{X}, \hat{\mathcal{B}}, \hat{\lambda}, \hat{T})\), where

\[
\hat{\lambda} = \tilde{\nu} + \bigotimes_{i=1}^k \hat{\mu}.
\]

Hence, by Lemma 3.6, for any \(A_1 \times A_2 \times \cdots A_k\), we have

\[
\int \frac{1}{N} \sum_{n=0}^{N-1} 1_{A_1 \times A_2 \times \cdots A_k} \circ \hat{T}^n \hat{\lambda} \rightarrow \otimes_{i=1}^k \hat{\mu}(A_1 \times A_2 \times \cdots A_k).
\]

Therefore, for any \(f_1 \otimes f_2 \otimes \cdots \otimes f_k \in \mathcal{C}(\hat{X})^k\), we get

\[
\int \frac{1}{N} \sum_{n=0}^{N-1} f_1 \otimes f_2 \otimes \cdots \otimes f_k \circ \hat{T}^n \hat{\lambda} \rightarrow \int f_1 \otimes f_2 \otimes \cdots \otimes f_k d(\otimes_{i=1}^k \hat{\mu}).
\]

This gives, by the standard argument, that for any continuous functions \(f\) on \(\hat{X}^k\)

\[
\int \frac{1}{N} \sum_{n=0}^{N-1} \hat{T}^n f d\hat{\lambda} \rightarrow \int \frac{d(\otimes_{i=1}^k \hat{\mu})}{d\hat{\lambda}} d\hat{\lambda}.
\]

Applying again a standard density argument, it follows that for any \(f \in L^\infty(\hat{X}, \hat{\lambda})\), we have

\[
\int \frac{1}{N} \sum_{n=0}^{N-1} \hat{T}^n f d\hat{\lambda} \rightarrow \int \frac{d(\otimes_{i=1}^k \hat{\mu})}{d\hat{\lambda}} d\hat{\lambda}.
\]
Hence, for any Borel $A$ in $\hat{X}^k$, we get
\[
\frac{1}{N} \sum_{n=0}^{N-1} \hat{\lambda}(T^n A) \xrightarrow{N \to \infty} \otimes_{i=1}^k \mu(A) \leq 2\hat{\lambda}(A).
\]

Consequently, the Hartman condition holds for our dynamical system $(\hat{X}, \hat{B}, \hat{\lambda}, \hat{T})$. Therefore, by Ryll-Nardzewski theorem (Theorem 2.2), $\hat{T}$ satisfy the pointwise ergodic theorem. We thus get that the pointwise ergodic theorem holds for $\hat{\nu}$, which yields that the pointwise ergodic theorem holds for $\hat{\mu}_\Delta$ since $\hat{\mu}_\Delta$ is absolutely continuous with respect to $\hat{\nu}$.

\textbf{Remark 5.1.} Notice that our proof yields that $\hat{T}$ satisfy the pointwise ergodic theorem with respect to the product measure, which follows also by the Birkhoff ergodic theorem. Moreover, our last density argument is based essentially on the classical Lusin theorem [24, p.53]. We further notice that our proof gives more, namely if $f_i \in L^k(\mu)$, then by generalized Hölder inequality, $f_1 \otimes f_2 \otimes \cdots \otimes f_k \in L^1(\hat{\lambda})$. Hence, the pointwise ergodic theorem holds.

In view of this last observation, it is natural to ask

\textbf{Question.} If $(X, \mathcal{A}, \mu, T_i)_{i=1}^k$, $k \geq 2$ is a finite family of Lebesgue dynamical systems where $\mu$ is a probability measure space, and if $(T_i)$ are commuting weakly mixing transformations on $X$. Does, for any $p \geq 1$ and for every functions $(f_i)_{i=1}^k \in (L^p(X))^k$, the following averages
\[
\frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^k f_i(T_i^n x)
\]

converge a.e.?

Notice that the case $k = 1$ follows from Dunford-Schwartz theorem [17, p.27] and the case $k = 2$ with $T_i = T^{p_i}$, $p_i \in \mathbb{N}^*$ follows from the Bourgain double ergodic theorem and the following observation

\textbf{Fact.} Assume that the pointwise convergence of Furstenberg average holds for $L^\infty$-functions then the pointwise convergence holds for $L^p$-functions by $L^\infty$ approximations combined with the Marcinkiewicz interpolation theorem [17, p.25] and the maximal ergodic inequality.

\textbf{Acknowledgment 1.} I would like to express my warmest thanks to Professor XiangDong Ye for his invitation to Hefei University of Technology and for bringing to my attention the problem of the pointwise convergence of Furstenberg average and the arXiv-paper of Professor Assani [5]. My thanks go also to Professor Idris Assani for his invitation to the Ergodic Workshop in Chapel Hill and for the conversations on Bourgain double ergodic theorem and his recent work with Ryo Moore [4]. Finally, my posthumous thanks to Professor Antoine Brunel who introduce me to the work of R. Sato.

\*A joint work on the extension of their result is in preparation.
References

[1] J. Aaronson, An introduction to infinite ergodic theory, Mathematical Surveys and Monographs, 50. American Mathematical Society, Providence, RI, 1997.
[2] I. Assani, Quelques résultats sur les opérateurs positifs à moyennes bornées dans $L_p$, (French) [Some results on $L_p$ mean bounded positive operators] Ann. Sci. Univ. Clermont-Ferrand II Probab. Appl., No. 3 (1985), 65-72.
[3] I. Assani, Multiple recurrence and almost sure convergence for weakly mixing dynamical systems, Israel J. Math., 103 (1998), 111-124.
[4] I. Assani & J. Wós, J. An equivalent measure for some nonsingular transformations and application, Studia Math., 97 (1990), no. 1, 1-12.
[5] I. Assani, A.e. Multiple recurrence for weakly mixing commuting actions, available on arXiv:1312.5270.
[6] T. Austin, On the norm convergence of non-conventional ergodic averages, Ergodic Theory Dynam. Systems, 30 (2010), 321-338.
[7] T. Austin, Pleasant extensions subject to some algebraic constraints, and applications, Preliminary notes available on arXiv:0905.0518, 2009.
[8] T. Austin, Topics in Ergodic Theory, Notes 8: Multiple Recurrence III: structure of the Furstenberg self-joining, notes available on http://www.math.brown.edu/~timaustin.
[9] V. Bergelson, Weakly mixing PET, Ergodic Theory Dynam. Systems, 7 (1987), 337-349.
[10] V. Bergelson and A. Leibman, A nilpotent Roth theorem, Invent. Math., 147 (2002), 429-470.
[11] J. Bourgain, Double recurrence and almost sure convergence, J. Reine Angew. Math., 404 (1990), 140-161.
[12] J.-P. Conze and E. Lesigne, Théorèmes ergodiques pour des mesures diagonales, Bull. Soc. Math. France, 112 (1984), 143-175.
[13] J-M. Derrien and E. Lesigne, Un théorème ergodique polynomial ponctuel pour les endomorphismes exacts et les K-systemes, Ann. Inst. H. Poincare Probab. Statist., 32 (1996), no. 6, 765-778.
[14] N. Frantzikinakis and B. Kra, Convergence of multiple ergodic averages for some commuting transformations, Ergodic Theory Dynam. Systems, 25 (2005), 799-809.
[15] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, J. Analyse Math., 31 (1977), 204-256.
[16] H. Furstenberg, Y. Katznelson, D. Ornstein, The ergodic theoretical proof of Szemerédi’s theorem, Bull. Amer. Math. Soc., (N.S.) 7 (1982), no. 3, 527-552.
[17] A. Garcia, Topics in almost everywhere convergence, Markham Publishing Company, Chicago, 1970.
[18] G. Hansel, J. P. Raoult, Ergodicity, uniformity and unique ergodicity, Indiana Univ. Math. J., 23 (1973/74), 221-237.
[19] B. Host and B. Kra, Nonconventional ergodic averages and nilmanifolds, Ann. of Math., 161 (2005), 397-488.
[20] Y. Ito, Uniform integrability and the pointwise ergodic theorem, Proc. Amer. Math. Soc., 16 (1965), 222-227.
[21] U. Krengel, Ergodic theorems, With a supplement by Antoine Brunel, de Gruyter Studies in Mathematics, 6. Walter de Gruyter & Co., Berlin, 1985.
[22] K. Petersen, Ergodic theory. Cambridge Studies in Advanced Mathematics 2., Cambridge University Press, Cambridge, 1983.
[23] P. Ortega Salvador, Weights for the ergodic maximal operator and a.e. convergence of the ergodic averages for functions in Lorentz spaces, Tohoku Math. J., (2) 45 (1993), no. 3, 437-446.
[24] W. Rudin, Real and complex analysis, Third edition. McGraw-Hill Book Co., New York, 1987.
[25] V. A. Rohklin, On the fundamental ideas of measure theory, Amer. Math. Soc. Translation, (1952), 55 pp.
[26] C. Ryll-Nardzewski, On the ergodic theorems, I, Generalized ergodic theorems, Studia Math., 12, (1951), 65-73.
[27] R. Sato, Pointwise ergodic theorems in Lorentz spaces $L(p, q)$ for null preserving transformations, Studia Math., 114 (3) (1995), 227-236.
[28] Sato, Ryotaro Pointwise ergodic theorems for functions in Lorentz spaces $L_{pq}$ with $p \neq \infty$, Studia Math., 109 (1994), no. 2, 209-216.

[29] T. Tao, Norm convergence of multiple ergodic averages for commuting transformations, Ergodic Theory Dynam. Systems, 28 (2008), 657-688.

[30] Q. Zhang, On convergence of the averages $(\frac{1}{N}) \sum_{n=1}^{N} f_1(R^n x) f_2(S^n x) f_3(T^n x)$, Monatsh. Math., 122 (1996), 275-300.

[31] T. Ziegler, Universal characteristic factors and Furstenberg averages, J. Amer. Math. Soc., 20 (2007), 53-97.

[32] B. Weiss, Strictly ergodic models for dynamical systems, Bull. Amer. Math. Soc., (N.S.) 13 (1985), no. 2, 143-146.

[33] M. Walsh, Norm convergence of nilpotent ergodic averages, Annals of Math, Volume 175 (2012), Issue 3, 1667-1688.

Normandie Université, Université de Rouen-Mathématiques, Labo. de Maths Raphael SALEM UMR 60 85 CNRS Avenue de l’Université, BP.12 76801 Saint Etienne du Rouvray - France
E-mail address: elhoucin.elabdalaoui@univ-rouen.fr