Journal of Nonlinear Modeling and Analysis
Volume 1, Number 4, December 2019, 513–525
http://jnma.ca; http://jnma.ijournal.cn
DOI:10.12150/jnma.2019.513

Ergodic Behaviour of Nonconventional Ergodic Averages for Commuting Transformations*

Xia Pan[1,†], Zuohuan Zheng[2,3,4] and Zhe Zhou[2]

Abstract Based on T. Tao's celebrated result on the norm convergence of multiple ergodic averages for commuting transformations, we find that there is a subsequence which converges almost everywhere. Meanwhile, we obtain the ergodic behaviour of diagonal measures, which indicates the time average equals the space average. According to the classification of transformations, we also give several different results. Additionally, on the torus $\mathbb{T}^d$ with special rotation, we prove the pointwise convergence in $\mathbb{T}^d$, and get a result for ergodic behaviour.

Keywords Commuting transformation, convergence almost everywhere, ergodic behaviour, time average, space average.

MSC(2010) 37A20, 28A35, 47E05.

1. Introduction

In 2008, T. Tao proved a convergence result for several commuting transformations:

Theorem 1.1. [14] Let $d \geq 1$ be an integer. Assume that $T_1, T_2, \ldots, T_d : X \to X$ are commuting invertible measure-preserving transformations of a measure space $(X, \mathcal{B}, \mu)$. Then, for any $f_1, f_2, \ldots, f_d \in L^\infty(X, \mathcal{B}, \mu)$, the averages

$$\frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^n x) \ldots f_d(T_d^n x) \tag{1.1}$$

are convergent in $L^2(X, \mathcal{B}, \mu)$.

[†]the corresponding author.
Email address:panxia13@mails.ucas.ac.cn(X. Pan), zhzheng@amt.ac.cn(Z. Zheng),zzhou@amss.ac.cn(Zh. Zhou)
[1]College of Liberal Arts and Sciences, National University of Defense Technology, Changsha, Hunan 410073, China
[2]Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[3]School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China
[4]College of Mathematics and Statistics, Hainan Normal University, Haikou, Hainan 571158, China
*The authors were supported by 1. National Natural Science Foundation of China(NSFC) under grant 11671382 and 11301512; 2. Key Laboratory of Random Complex Structures and Data Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences(No.2008DP173182); 3. CAS Key Project of Frontier Sciences(No.QYZDJ-SSW-JSC003); 4. National Center for Mathematics and Interdisciplinary Sciences.

Soon after, H. Towsner [15], B. Host [9] and T. Austin [2] gave proofs of Theorem 1.1 from different viewpoints. T. Tao's approach was combinatorial and finitary, inspired by the hypergraph regularity and removal lemmas. H. Towsner used nonstandard analysis, whereas T. Austin and B. Host all exploited ergodic methods, building an extension of the original system with good properties.

There is a rich history towards Theorem 1.1. For $d = 1$, it reduces to the classical mean ergodic theorem. When $T_1 = T, T_2 = T^2, \ldots, T_d = T^d$, Furstenberg studied such averages originally in his proof of Szemerédi's theorem [7], where T is weakly mixing or T is general but $d = 2$. For higher d, various special cases have been shown by Conze and Lesigne [4, 5], Furstenberg and Weiss [8], Host and Kra [10], and Ziegler [19]. Finally, it was totally proved by Host and Kra [11] for arbitrary d, and independently by Ziegler [20].

When $T_1, T_2, \ldots, T_d$ are commuting measure-preserving transformations with some hypothesis on the transformations, Zhang [18] gave a proof for $d = 3$ and Frantzikinakis and Kra [6] for general d. Without those assumptions, the L^2-convergence of the averages (1.1) was established by Tao. As we have mentioned above, it possesses four different proofs. When $T_1, T_2, \ldots, T_d$ belongs to nilpotent group, it was proved by Miguel N. Walsh [16].

Although most people believe the existence of the averages (1.1) almost everywhere, the cases in which one knows the answer are scarce. In this paper, With the fact that the averages (1.1) have a subsequence which converges almost everywhere, the ergodic behaviour of diagonal measure is proved. Furthermore, on the torus $\mathbb{T}^d$ with special rotation, say, $R_{\alpha_1,\ldots,\alpha_d} : \mathbb{T}^d \to \mathbb{T}^d$, where $1, \alpha_1, \ldots, \alpha_d$ are rationally independent, the convergence of the averages (1.1) for every point in $\mathbb{T}^d$ is obtained, and a result for ergodic behaviour is presented.

Before launching into the main result, we first remind the reader some elements of the measure theory and the ergodic theory in Section 2. With sufficient preparation, we give a proof of the ergodic behaviour of Theorem 1.1, and give a classification of $T_1, T_2, \ldots, T_d$, in case 1: all the T_i are pairwise different, i.e., $T_i \neq T_j,\ i \neq j$, and in case 2: there is k, with $1 \leq k \leq d$, such that $T_{i_1} = T_{i_2} = \cdots = T_{i_k}$ in Section 3. In Section 4, we will employ the result obtained in Section 3 to the special case in which the space is the torus $\mathbb{T}^d$, and transformations $R_{\alpha_1}, \ldots, R_{\alpha_d} : \mathbb{T} \to \mathbb{T}$, satisfying that $1, \alpha_1, \cdots, \alpha_d$ are rationally independent. In Section 5, we give several examples to show that each alternative in Section 3 and Section 4 really occurs.

2. Preliminary

Let us first recall from [13, 17] some basic facts on measure theory and ergodic theory.

2.1. Measure Theory

In this section, X will be an arbitrary measure space equipped a positive measure μ.

Definition 2.1. [13] Let μ be a positive measure on X. A sequence $\{f_n\}$ of measurable functions on X is said to convergence in measure to the measurable function f if for every $\epsilon > 0$ there corresponds an N such that

$$\mu(\{x : |f_n(x) - f(x)| > \epsilon\}) < \epsilon \tag{2.1}$$

for all $n > N$.

Definition 2.2. [13] If $1 \le p < \infty$ and f is a measurable function on X, define

$$\|f\|_p = \left\{ \int_X |f|^p \, \mathrm{d}\mu \right\}^{\frac{1}{p}} \tag{2.2}$$

and let $L^p(\mu)$ consist of all f for which

$$\|f\|_p < \infty. \tag{2.3}$$

We call $\|f\|_p$ the L^p-norm of f.

If $f, f_1, \ldots, f_n, \ldots \in L^p(\mu)$ with $\lim_{n\to\infty} \|f_n - f\|_p = 0$, we say that $\{f_n\}$ converges to f in the mean of order p, or that $\{f_n\}$ is L^p-convergent to f.

Theorem 2.3. *Assuming $\mu(X) < \infty$, we have the following statements:*

1. *If $f_n \in L^p(\mu)$ and $\|f_n - f\|_p \to 0$, then $f_n \to f$ in measure; here $1 \le p < \infty$.*
2. *If $f_n \to f$ in measure, then $\{f_n\}$ has a subsequence $\{f_{n_i}(x)\}$ which converges to f almost everywhere, i.e.,*

$$\lim_{i\to\infty} f_{n_i}(x) = f(x), a.e..$$

2.2. Ergodic Theory

Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ be a measure-preserving transformation. T is called ergodic if $T^{-1}B = B$ for $B \in \mathcal{B}$ satisfy $\mu(B) = 0$ or $\mu(B) = 1$. The next theorem gives another form of the definition of ergodicity.

Theorem 2.4. [17] *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be a measure-preserving transformation. Then T is ergodic iff $\forall A, B \in \mathcal{B}$,*

$$\frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}A \cap B) \to \mu(A)\mu(B). \tag{2.4}$$

In order to describe the proof of our main result, it will be convenient to reformulate Theorem 2.4 in terms of functions.

Corollary 2.5. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $T : X \to X$ be a measure-preserving transformation. Then T is ergodic iff $\forall f, g : X \to \mathbb{R}, f, g \in L^\infty(X, \mathcal{B}, \mu)$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int f(T^i(x)) g(x) \, \mathrm{d}\mu = \int f(x) \, \mathrm{d}\mu \int g(x) \, \mathrm{d}\mu. \tag{2.5}$$

Proof Let us first prove the “ only if ” part. Assume that T is ergodic, from Theorem 2.4, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int \chi_A(T^i(x)) \chi_B(x) \, \mathrm{d}\mu = \int \chi_A(x) \, \mathrm{d}\mu \int \chi_B(x) \, \mathrm{d}\mu, \tag{2.6}$$

The functions in L^∞ can be approximated by some simple functions, based on the function approximation theory. we obtain the desired result (2.5).

Now let us prove the " if " part. Let $f = \chi_A, g = \chi_B$, we can get (2.4) easily. From Theorem 2.4, T is ergodic. □

The first major result in ergodic theory was proved in 1931 by G.D. Birkhoff. The explicit form is stated below.

Theorem 2.6. [17] *Suppose $T : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ is measure-preserving and $f \in L^1(\mu)$. Then $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$ converges almost everywhere to a function $f^* \in L^1(\mu)$. Also $f^* \circ T = f^*$ almost everywhere and if $\mu(X) < \infty$, then $\int f^* \,\mathrm{d}\mu = \int f \,\mathrm{d}\mu$.*

Remark 2.7. If T is ergodic then f^* is constant almost everywhere and so if $\mu(X) < \infty$, $f^* = (1/\mu(X)) \int f \,\mathrm{d}\mu$ almost everywhere. If $(X, \mathcal{B}, \mu)$ is a probability space and T is ergodic we have $\forall f \in L^1(\mu)$, $\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \int f(T^n(x)) \,\mathrm{d}\mu = \int f(x) \,\mathrm{d}\mu$ almost everywhere.

Unique ergodicity can get much stronger behaviour in the ergodic theorem, now let us recall the definition of unique ergodicity and its behaviour of those averages in Theorem 2.6.

Definition 2.8. [17] A continuous transformation $T : X \to X$, where X is a compact metrisable space, is called uniquely ergodic if there is only one T invariant Borel probability measure on X.

Theorem 2.9. [17] *Let $T : X \to X$ be a continuous transformation of a compact metrisable space X. The following statements are equivalent:*

- *For every $f \in C(X)$, $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$ converges uniformly to a constant.*
- *For every $f \in C(X)$, $\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x)$ converges pointwise to a constant.*
- *There is a unique probability measure on X which is invariant under T such that for all $f \in C(X)$ and all $x \in X$,*

$$\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \to \int f \,\mathrm{d}\mu.$$

- *T is uniquely ergodic.*

When X is a compact metrisable space, let $\mathcal{B}(X)$ be the Borel σ-algebra on X, $T : X \to X$ be a continuous transformation. We shall denote $\mathcal{M}(X), \mathcal{M}(X, T)$ as

$$\mathcal{M}(X) = \{\mu : \mathcal{B}(X) \to [0, 1] \mid \mu(X) = 1\},$$

$$\mathcal{M}(X, T) = \{\mu \in \mathcal{M}(X) \mid \mu(T^{-1}B) = \mu(B), B \in \mathcal{B}(X)\}.$$

$\mathcal{M}(X)$ is convex and compact in the weak*-topology [17, Theorem 6.5], and $\mathcal{M}(X, T)$ is a convex and compact subset of $\mathcal{M}(X)$ [17, Theorem 6.5].

Corollary 2.10. *If $T_1, T_2, \ldots, T_d : X \to X$ are pairwise commuting continuous maps of a compact space X, then they possess a common invariant probability measure.*

3. Ergodic behaviour

Ergodic theory is the study of statistical properties of dynamical systems related to a measure on the underlying space of the dynamical system. The name comes from classical statistical mechanics, where the "ergodic hypothesis" asserts that, asymptotically, the time averages of an observable is equal to the space average. Now, we have known the "ergodic hypothesis" happens if the system is ergodic(see Remark 2.7).

Before the main result of this paper is given, let us introduce the concept of an irreducible dynamical system as a preliminary.

Definition 3.1. Given a probability space $(X, \mathcal{B}, T_1, T_2, \ldots, T_d, \mu)$, where $T_1, T_2 \ldots,$ $T_d : X \to X$ are commuting invertible measure-preserving transformations. If for any $f_1, f_2, \ldots, f_d \in L^\infty(X, \mathcal{B}, \mu)$,

$$A = \{x : \lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^n x) \ldots f_d(T_d^n x) < \alpha\},$$

$$B = \{x : \lim_{n\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(T_1^{n+1} x) \ldots f_d(T_d^{n+1} x) < \alpha\},$$

$\forall \alpha \in \mathbb{R}$, we have $A = B$, then the set A is called 'invariant' with respect to the measure μ.

Definition 3.2. The probability space $(X, \mathcal{B}, T_1, T_2, \ldots, T_d, \mu)$, where $T_1, T_2, \ldots, T_d :$ $X \to X$ are commuting invertible measure-preserving transformations, is called 'irreducible' with respect to the measure μ, if it is impossible to represent X as the sum of two measurable 'invariant' sets of positive measure without common points.

Theorem 3.3. *Let $d \geq 1$ be an integer. Assume that $T_1, T_2, \ldots, T_d : X \to X$ are commuting invertible measure-preserving transformations of a 'irreducible' space $(X, \mathcal{B}, \mu)$, $T_i T_j^{-1}, i \neq j$ are ergodic. Then, for any $f_1, f_2, \ldots, f_d \in L^\infty(X, \mathcal{B}, \mu)$, $T_i \neq T_j, i \neq j$, there is a subsequence $\{N_k\}$ such that*

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T_1^n x) \ldots f_d(T_d^n x) = \int_X f_1 \,\mathrm{d}\mu \ldots \int_X f_d \,\mathrm{d}\mu \qquad a.e.. \tag{3.1}$$

Proof According to Theorem 1.1, the averages (1.1) converges in L^2, According to Theorem 2.3, the averages (1.1) has a subsequence which converges almost everywhere.

First of all, We shall prove that under the condition of irreducible, the subsequence converges to a constant almost everywhere. Using the method of proof by contradiction, we denote by M the least upper bound of $f(x)$ over X computed on neglecting a set of measure zero and analogously we denote by m the greatest lower bound of the function $f(x)$ on neglecting a set of measure zero. From the assumption there follows $M > m$.
Let α satisfy the inequalities $m < \alpha < M$. We obtain

$$\mu\{p : f(p) < \alpha\} = \mu E_\alpha > 0,$$

and

$$\mu(X - E_\alpha) = \mu\{p : f(p) \geq \alpha\} > 0,$$

then we decompose X into two sets of positive measure which contradicts the condition of 'irreducibility'.
For $d = 1$, the statement is the famous mean ergodic theorem.
For $d = 2$, almost every $x \in X$

$$\begin{aligned}\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)f_2(T_2^n x) &= \int_X \lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)f_2(T_2^n x)\,\mathrm{d}\mu \\ &= \lim_{k\to\infty}\int_X \frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)f_2(T_2^n x)\,\mathrm{d}\mu \\ &= \lim_{k\to\infty}\int_X \frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(x)f_2((T_2T_1^{-1})^n x)\,\mathrm{d}\mu \\ &= \lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1}\int_X f_1(x)f_2((T_2T_1^{-1})^n x)\,\mathrm{d}\mu \\ &= \int_X f_1\,\mathrm{d}\mu\int_X f_2\,\mathrm{d}\mu. \qquad \text{(by Corollary 2.5)}\end{aligned}$$

In the second equality, we used the Lebesgue's Dominated Convergence Theorem [13, p. 26].

Putting $f_1 = \chi_A, f_2 = \chi_B$ in the situation $d = 2$, then

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1}\chi_A(T_1^n x)\chi_B(T_2^n x) = \int_X \chi_A\,\mathrm{d}\mu\int_X \chi_B\,\mathrm{d}\mu \qquad a.e..$$

Multiplying the both sides by χ_C,

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1}\chi_{T_1^{-n}A\cap T_2^{-n}B}\chi_C = \int_{\mathbb{T}}\chi_A\,\mathrm{d}\mu\int_{\mathbb{T}}\chi_B\,\mathrm{d}\mu\cdot\chi_C \qquad a.e.,$$

thus the dominated convergence theorem implies

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1}\mu(T_1^{-n}A\cap T_2^{-n}B\cap C) = \mu(A)\mu(B)\mu(C). \tag{3.2}$$

For $d = 3$, almost every $x \in X$

$$\begin{aligned}&\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)f_2(T_2^n x)f_3(T_3^n x) \\ =&\int_X \lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)f_2(T_2^n x)f_3(T_3^n x)\,\mathrm{d}\mu \\ =&\lim_{k\to\infty}\int_X \frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)f_2(T_2^n x)f_3(T_3^n x)\,\mathrm{d}\mu \\ =&\lim_{k\to\infty}\int_X \frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(x)f_2((T_2T_1^{-1})^n x)f_3((T_3T_1^{-1})^n x)\,\mathrm{d}\mu\end{aligned}$$

$$= \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} \int_X f_1(x) f_2((T_2T_1^{-1})^n x) f_3((T_3T_1^{-1})^n x)\, \mathrm{d}\mu$$
$$= \int_X f_1 \,\mathrm{d}\mu \int_X f_2 \,\mathrm{d}\mu \int_X f_3 \,\mathrm{d}\mu.$$

In the same way, after d steps, we obtain

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T_1^n x) \dots f_d(T_d^n x) = \int_X f_1 \,\mathrm{d}\mu \dots \int_X f_d \,\mathrm{d}\mu.$$

□

Corollary 3.4. *Let $d \geq 1$ be an integer. Assume that $T_1, T_2, \dots, T_d : X \to X$ are commuting invertible measure-preserving transformations of a irreducible space $(X, \mathcal{B}, \mu)$, with $T_{i_1} = T_{i_2} = \dots = T_{i_k}$, $1 \leq k \leq d$. Without loss of generality, let $i_1 = 1, i_2 = 2, \dots, i_k = k$. $T_iT_j^{-1}, i \neq j$ are ergodic. Then, for any $f_1, f_2, \dots, f_d \in L^\infty(X, \mathcal{B}, \mu)$,*

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T_1^n x) \dots f_d(T_d^n x) = \int_X f_1 f_2 \dots f_k \,\mathrm{d}\mu \int_X f_{k+1} \,\mathrm{d}\mu \dots \int_X f_d \,\mathrm{d}\mu \quad a.e.. \tag{3.3}$$

Proof As the proof of Theorem 3.3, the subsequence converges to a constant almost everywhere.
For $d = 1$, the statement is the mean ergodic theorem.
For $d = 2$, if $T_1 \neq T_2$, it is conformity with Theorem3.3.
For $d = 2$, and $T_1 = T_2 = T$, for almost every $x \in X$

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T_1^n x) f_2(T_2^n x) &= \int_X \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T^n x) f_2(T^n x) \,\mathrm{d}\mu \\ &= \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} \int_X f_1(T^n x) f_2(T^n x) \,\mathrm{d}\mu \\ &= \int_X f_1 f_2 \,\mathrm{d}\mu. \qquad (\mu \text{ is invariant}) \end{aligned}$$

In the procedure of the proof, we used the Lebesgue's Dominated Convergence Theorem [13, p. 26].
For $d = 3$, if $T_1 \neq T_2 \neq T_3$, it is conformity with Theorem3.3.
For $d = 3$, and $T_1 = T_2 = T \neq T_3$, for almost every $x \in X$,

$$\begin{aligned} \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T_1^n x) f_2(T_2^n x) f_3(T_3^n x) &= \int_X \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_1(T^n x) f_2(T^n x) f_3(T_3^n x) \,\mathrm{d}\mu \\ &= \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} \int_X f_1(T^n x) f_2(T^n x) f_3(T_3^n x) \,\mathrm{d}\mu \\ &= \lim_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} \int_X f_1(x) f_2(x) f_3((T_3T^{-1})^n x) \,\mathrm{d}\mu \end{aligned}$$

$$=\int_X f_1 f_2 \,\mathrm{d}\mu \int_X f_3 \,\mathrm{d}\mu.$$

For $d=3$, and $T_1=T_2=T_3=T$, for almost every $x \in X$,

$$\begin{aligned}\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x) f_2(T_2^n x) f_3(T_3^n x) &=\int_X \lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T^n x) f_2(T^n x) f_3(T^n x)\,\mathrm{d}\mu\\ &=\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1}\int_X f_1(T^n x) f_2(T^n x) f_3(T^n x)\,\mathrm{d}\mu\\ &=\int_X f_1 f_2 f_3 \,\mathrm{d}\mu.\end{aligned}$$

In the same way, after d steps, we obtain

$$\lim_{k\to\infty}\frac{1}{N_k}\sum_{n=0}^{N_k-1} f_1(T_1^n x)\dots f_d(T_d^n x) = \int_X f_1 f_2 \dots f_k \,\mathrm{d}\mu \int_X f_{k+1}\,\mathrm{d}\mu \dots \int_X f_d \,\mathrm{d}\mu \quad a.e..$$

□

Remark 3.5. If the open problem of pointwise almost everywhere convergence of the averages (1.1) is solved, our methods can be applied to the multiple averages directly.

4. Special case

Consider the torus $\mathbb{T}^d$ with special rotation, we can not only get the convergence of the averages (1.1) for every point in $\mathbb{T}^d$, although almost everywhere convergence of the averages (1.1) still unknown, but also get the result of Theorem 3.3.

At first, let us recall the concept of rationally independent.

Definition 4.1. The real numbers $1, \alpha_1, \dots, \alpha_d$ are rationally independent, if there is no $k_0, k_1, \dots, k_d \in Z^{d+1}\backslash\{0\}$ such that $k_0 + k_1\alpha_1 + \dots k_d\alpha_d = 0$.

The translation

$$R_\alpha = R_{\alpha_1,\dots,\alpha_d}(x_1,\dots,x_d) = (x_1+\alpha_1,\dots,x_d+\alpha_d)$$

induces the rotation $R_\alpha = R_{\alpha_1,\dots,\alpha_d} : \mathbb{T}^d \to \mathbb{T}^d$.

From next lemma, the rationally independent rotation $R_\alpha = R_{\alpha_1,\dots,\alpha_d} : \mathbb{T}^d \to \mathbb{T}^d$ is uniquely ergodic.

Lemma 4.2. *Let $\alpha = (\alpha_1,\dots,\alpha_d)$ with $1,\alpha_1,\dots,\alpha_d$ rationally independent. The Haar measure is the only probability measure which is invariant by $R_\alpha : \mathbb{T}^d \to \mathbb{T}^d$.*

Analogy to Theorem 3.3, the following statements are valid:

Theorem 4.3. *Let $d \geq 1$ be an integer and $\alpha = (\alpha_1,\dots,\alpha_d)$ with $1,\alpha_1,\dots,\alpha_d$ rationally independent, $R_{\alpha_1},\dots,R_{\alpha_d} : \mathbb{T}\to\mathbb{T}$, $f_1,\dots,f_d : \mathbb{T}\to\mathbb{R}$, $f_1, f_2, \dots, f_d \in C(\mathbb{T})$. Assuming μ is the common invariant probability measure of $R_{\alpha_1},\dots,R_{\alpha_d}$. Then*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} f_1(R_{\alpha_1}^n x)\dots f_d(R_{\alpha_d}^n x) = \int_{\mathbb{T}} f_1\,\mathrm{d}\mu \dots \int_{\mathbb{T}} f_d\,\mathrm{d}\mu. \tag{4.1}$$

Proof From Lemma 4.2, for $\forall f \in C(\mathbb{T}^d), 1 \le p < \infty$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(R_\alpha^n(y)) = \int_{\mathbb{T}^d} f \,\mathrm{d}\nu, \forall y \in \mathbb{T}^d, \tag{4.2}$$

where $\nu = \mu \times \ldots \times \mu$.
Let $f(x_1, x_2, \ldots, x_d) = f_1(x_1) \cdot f_2(x_2) \cdot \ldots \cdot f_d(x_d)$, then

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(R_{\alpha_1}^n x) \ldots f_d(R_{\alpha_d}^n x) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(R_\alpha^n(x, \ldots, x)) \\ &= \int_{\mathbb{T}^d} f \,\mathrm{d}\nu \\ &= \int_{\mathbb{T}} f_1 \,\mathrm{d}\mu \ldots \int_{\mathbb{T}} f_d \,\mathrm{d}\mu.\end{aligned}$$

□

Corollary 4.4. *Let $d \ge 1$ be an integer and $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $1, \alpha_1, \ldots, \alpha_d$ rationally independent, $R_{\alpha_1}, \ldots, R_{\alpha_d}, S : \mathbb{T} \to \mathbb{T}$, $f_1, \ldots, f_d, g : \mathbb{T} \to \mathbb{R}$, $f_1, f_2, \ldots, f_d, g \in C(\mathbb{T})$, S is a periodic transformation with $S^k x = x$. Assuming μ is the common invariant probability measure of $R_{\alpha_1}, \ldots, R_{\alpha_d}$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(R_{\alpha_1}^n x) \ldots f_d(R_{\alpha_d}^n x) g(S^n x) = \int_{\mathbb{T}} f_1 \,\mathrm{d}\mu \ldots \int_{\mathbb{T}} f_d \,\mathrm{d}\mu \frac{1}{k} \sum_{r=0}^{k-1} g(S^r x). \tag{4.3}$$

Proof When $N = kp$, then

$$\begin{aligned}\lim_{p\to\infty} \frac{1}{kp} \sum_{n=0}^{kp-1} f_1(R_{\alpha_1}^n x) \ldots f_d(R_{\alpha_d}^n x) g(S^n x) &= \lim_{p\to\infty} \frac{1}{kp} \sum_{r=0}^{k-1} \sum_{i=0}^{p-1} f_1(R_{\alpha_1}^{ik+r} x) \ldots f_d(R_{\alpha_d}^{ik+r} x) g(S^r x) \\ &= \int_{\mathbb{T}} f_1 \,\mathrm{d}\mu \ldots \int_{\mathbb{T}} f_d \,\mathrm{d}\mu \frac{1}{k} \sum_{r=0}^{k-1} g(S^r x).\end{aligned}$$

When $N = kp + m$, $1 \le m < k$, then

$$\begin{aligned}&\lim_{p\to\infty} \frac{1}{kp+m} \sum_{n=0}^{kp+m-1} f_1(R_{\alpha_1}^n x) \ldots f_d(R_{\alpha_d}^n x) g(S^n x) \\ &= \lim_{p\to\infty} \frac{1}{kp+m} \{ \sum_{r=0}^{k-1} \sum_{i=0}^{p-1} f_1(R_{\alpha_1}^{ik+r} x) \ldots f_d(R_{\alpha_d}^{ik+r} x) g(S^r x) \\ &\quad + \sum_{j=1}^{m} f_1(R_{\alpha_1}^{pk+j} x) \ldots f_d(R_{\alpha_d}^{pk+j} x) g(S^j x) \} \\ &= \lim_{p\to\infty} \frac{kp}{kp+m} \frac{1}{kp} \sum_{r=0}^{k-1} \sum_{i=0}^{p-1} f_1(R_{\alpha_1}^{ik+r} x) \ldots f_d(R_{\alpha_d}^{ik+r} x) g(S^r x) \\ &\quad + \lim_{p\to\infty} \frac{1}{kp} \sum_{j=1}^{m} f_1(R_{\alpha_1}^{pk+j} x) \ldots f_d(R_{\alpha_d}^{pk+j} x) g(S^j x) \\ &= \int_{\mathbb{T}} f_1 \,\mathrm{d}\mu \ldots \int_{\mathbb{T}} f_d \,\mathrm{d}\mu \frac{1}{k} \sum_{r=0}^{k-1} g(S^r x).\end{aligned} \tag{4.4}$$

5. Examples

Now we present several examples in which each alternative in Section 3 and 4 occurs. Although in the uniquely ergodic theorem, we need the function f is continuous, in our examples, we can get the result with function is not continuous.

Proposition 5.1. *Let $\{x\}$ denote the decimal part of x, $f(x) : S^1 \to \mathbb{R}$, $f(x) = \{x\}$, giving measure-preserving transformation $T : S^1 \to S^1$ is ergodic, μ is the Haar measure. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \frac{1}{2}. \tag{5.1}$$

Proof Let $g(x), h(x)$ be

$$g(x) = \begin{cases} f(x) & 0 \le x \le 1 - \frac{1}{m}, \\ (1-m)(x-1) & 1 - \frac{1}{m} \le x \le 1. \end{cases}$$

$$h(x) = \begin{cases} (1-m)x + 1 & 0 \le x \le \frac{1}{m}, \\ f(x) & \frac{1}{m} \le x \le 1. \end{cases}$$

Obviously, $g(x), h(x)$ have the following properties:

1. $g(x), h(x)$ are continuous functions,
2. $g(x) \le f(x) \le h(x), \quad \forall x \in S^1$,
3. $\lim_{m\to\infty} g(x) = \lim_{m\to\infty} h(x) = f(x)$,
4. $\int_{S^1} g(x)\,\mathrm{d}\mu = (m-1)/(2m)$,
5. $\int_{S^1} h(x)\,\mathrm{d}\mu = (m+1)/(2m)$.

By Theorem 2.9,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} g(T^n x) = \int_{S^1} g(x)\,\mathrm{d}\mu. \tag{5.2}$$

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} h(T^n x) = \int_{S^1} h(x)\,\mathrm{d}\mu. \tag{5.3}$$

From property 2, we have

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} g(T^n x) \le \liminf_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \le \limsup_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} f(T^n x) \le \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} h(T^n x).$$

According to all the properties and equations listed above, we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) = \frac{1}{2}.$$

□

Example 5.2. Let $\{x\}$ denote the decimal part of x.

1. $\lim_{N\to\infty} \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-1}\{x+\sqrt{2}n\}\{x+\sqrt{3}n\} = \dfrac{1}{4}$.

2. $\lim_{N\to\infty} \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-1}\{x+\sqrt{2}n\}\{x+\sqrt{2}n\} = \dfrac{1}{3}$.

3. $\lim_{N\to\infty} \dfrac{1}{N}\displaystyle\sum_{n=0}^{N-1}\{x+\sqrt{2}n\}\left\{x+\dfrac{n}{k}\right\} = \dfrac{1}{2k}\{kx\} + \dfrac{k-1}{4k}$.

Proof Let $f_1, f_2 : S^1 \to \mathbb{R}$, with $f_1(x) = f_2(x) = \{x\}$

$$\begin{aligned} T_{\sqrt{2}} &: S^1 \to S^1 \\ & x \mapsto \sqrt{2} + x \ mod\ 1. \\ T_{\sqrt{3}} &: S^1 \to S^1 \\ & x \mapsto \sqrt{3} + x \ mod\ 1. \end{aligned}$$

- Proof of Example 1

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\{x+\sqrt{2}n\}\{x+\sqrt{3}n\} &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f_1(T_{\sqrt{2}}^n x) f_2(T_{\sqrt{3}}^n x) \\ &= \int_{S^1} f_1 \,\mathrm{d}\mu \int_{S^1} f_2 \,\mathrm{d}\mu \\ &= \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}. \end{aligned}$$

- Proof of Example 2

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1}\{x+\sqrt{2}n\}\{x+\sqrt{2}n\} &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f_1(T_{\sqrt{2}}^n x) f_2(T_{\sqrt{2}}^n x) \\ &= \int_{S^1} f_1^2 \,\mathrm{d}\mu \\ &= \int_{S^1} x^2 \,\mathrm{d}\mu \\ &= \frac{1}{3}. \end{aligned}$$

- Proof of Example 3

Assertion:

$$\sum_{r=0}^{k-1}\left\{x+\frac{r}{k}\right\} = \{kx\} + \frac{k-1}{2}, \quad x \in [0,1]. \tag{5.4}$$

Let us prove the assertion firstly:

Proof When $\frac{i}{k} \le x \le \frac{i+1}{k}$, $i = 0,1,\ldots,k-1$. We have $x + \frac{r}{k} \le 1$, $r = 0,1,\ldots,k-i-1$, and $1 \le x + \frac{r}{k} \le 2$, $r = k-i, k-i+1, \ldots, k-1$, then

$$\sum_{r=0}^{k-1}\left\{x+\frac{r}{k}\right\} = \sum_{r=0}^{k-i-1}(x+\frac{r}{k}) + \sum_{r=k-i}^{k-1}(x+\frac{r}{k}-1)$$

$$
\begin{aligned}
&= kx + \frac{k-1}{2} - i \\
&= \{kx\} + \frac{k-1}{2}.
\end{aligned}
$$

□

Let $g : S^1 \to \mathbb{R}$, with $g(x) = \{x\}$,

$$
\begin{aligned}
T : S^1 &\to S^1 \\
x &\mapsto x + \frac{1}{k} \; mod \; 1.
\end{aligned}
$$

$$
\begin{aligned}
\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \{x+\sqrt{2}n\} \left\{x + \frac{n}{k}\right\} &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} f_1(T^n_{\sqrt{2}} x) g(T^n x) \\
&= \int_{S^1} f_1 \, \mathrm{d}\mu \cdot \frac{1}{k} \sum_{r=0}^{k-1} g(T^r x) \\
&= \frac{1}{2k} \sum_{r=0}^{k-1} \left\{x + \frac{r}{k}\right\} \\
&= \frac{1}{2k} \{kx\} + \frac{k-1}{4k}.
\end{aligned}
$$

Acknowledgements: The authors thank to Dr. Xu Xia for valuable suggestions and Dr. Yuanyuan Mao for modifying the article. Recently, we have noticed that E. H. EL Abdalaoui proved that there exist a subsequence for which the multiple ergodic averages of commuting invertible measure preserving transformations of a Lebesgue probability space converge almost provided that the maps are weakly mixing with extra condition [1], the authors are also thankful to Professor H. EL Abdalaoui for sharing his results with us.

References

[1] E.el Abdalaoui, *On the pointwise convergence of multiple ergodic averages and non-singular dynamical systems.* Preprint, available online at arXiv.org:1406.2608v3[math.DS].

[2] T. Austin, *On the norm convergence of non-conventional ergodic averages*, Ergodic Theory and Dynamical Systems, 2010, 30, 321–338.

[3] V. Bergelson, *Weakly mixing PET*, Ergodic Theory and Dynamical Systems, 1987, 7, 337–349.

[4] J.P. Conze and E. Lesigne, *Théorèmes ergodiques pour des mesures diagonales*, bulletin de la Société Mathématique de France, 1984, 112, 143–175.

[5] J.P. Conze and E. Lesigne, *Sur un théorèmes ergodique pour des mesures diagonales*, Publi. Inst. Rech. Math. Rennes, Probabilitès, 1987.

[6] N. Frntzikinakis and B. Kra, *Convergence of multiple ergodic averages for some commuting transformations*, Ergodic Theory Dynamical Systems, 2005, 25, 799–809.

[7] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Anal. Math, 2009, 31, 204–256.

[8] H. Furstenberg and B. Weiss, *A mean ergodic theorem for* $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$, in Convergence in Ergodic Theory and Probability(V. Bergelson, P.March and J.M.Rosenblatt, eds.), Walter de Gruyter, Berlin, 1996, 193–227.

[9] B. Host, *Ergodic seminorm for commutting transformations and applications*, Studia Mathematica, 2009, 195(1), 31–49.

[10] B. Host and B. Kra, *Convergence of Conze-Lesigne averages*, Ergodic Theory and Dynamical System, 2001, 21, 493–509.

[11] B. Host and B. Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math., 2005, 161, 397–488.

[12] Leibman, *Convergence of multiple ergodic averages along polynomials of several variables*, Israel J. Math., 2005, 146, 303–315.

[13] Walter Rudin, *Real and Complex Analysis*, China Machine Press, Beijing, 2004.

[14] T. Tao, *Norm convergence of multiple ergodic averages for commuting transformations*, Ergodic Theory and Dynamical Systems, 2008, 28, 657–688.

[15] H. Towsner, *Convergence of diagonal ergodic averages*, Ergodic Theory and Dynamical Systems, 2009, 29, 1309–1326.

[16] Miguel N.Walsh, *Norm convergence of nilpotent ergodic averages*, Ann. of Math., 2012, 175, 1667–1688.

[17] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, New York/Berlin, 1982.

[18] 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., 1996, 122, 275–300.

[19] T. Ziegler, *Nonconventional ergodic averages*, Ph.D. Thesis, Technion, 2002.

[20] T. Ziegler, *Universal characteristic factors and Furstenberg averages*, J. Amer. Math. Soc., 2007, 20, 53–97.