Mixing properties for a class of skew products and uniform convergence in ergodic theorems

Julia Brettschneider

EURANDOM, P.O.Box 513, 5600 MB Eindhoven, The Netherlands

j.brettschneider@tue.nl[1]

Abstract

Ergodicity and mixing properties with respect to the product measure for a class of skew products are discussed. We prove that the properties ergodic, weakly mixing, or strongly mixing, are passed on from the transformation in the base to the skew product provided that the semigroup of transformations in the fibre fulfills a suitable mixing conditions.

We study the convergence of ergodic theorems along skew products for the case when the transformation in the base is ergodic, and when it is periodic. Under suitable conditions, we show uniform convergence with respect to the first parameter, and L^1-convergence with respect to the second parameter. A P-a.s. version is derived as well.

1 Introduction

Two measure-preserving transformations on different probability spaces may be joined to form a measure-preserving transformation of the product space, by letting each one act on its own component. They can also be linked by letting the transformation in the the second argument depend on the value of the first argument. The resulting transformation is called a *skew product* of transformations. The independent transformation is called the base, the dependent one the fibre.

Ergodicity, and other mixing properties of skew products with respect to the product measure, have been studied by a number of authors. The case of a Bernoulli-shift in the base and an ergodic transformation in the fibre was introduced by Kakutani [17]. He showed that the skew product is ergodic if and only if the transformation in the fibre is ergodic. Other mixing properties were investigated, e.g., by Meilijson [20], den Hollander and Keane [10], den Hollander [9], and Georgii [14]. Adler and Shields (cf. [1] and [2]) consider a translation on the torus for the fibre. Anzai [3] introduced skew products of two translations on the torus, and derived a criteria for ergodicity. Furstenberg [12] studied unique ergodicity. Zhang [25] investigated this for a translation on a higher dimensional torus in the fibre. A torus translations in the base can also be combined with the translation on $\mathbb{R}$ by the value of a real function of the argument in the base. Skew products of this type are called real extensions of torus translations, and they were explored in Oren [22], Hellekalek and Larcher [15], [16], and Pask [23].

The class of skew products investigated in this paper is given by

$$S(t, \omega) = (\tau(t), \theta_{\kappa(t)}\omega) \qquad (t \in M, \omega \in \Omega), \tag{1}$$

---

where τ is an ergodic transformation on a probability space $(M, \mathcal{B}, \mu)$, $(\theta_k)_{k \in K}$ is a mixing semi-group of measure-preserving transformations on $(\Omega, \mathcal{F}, P)$, and κ is a K-valued $\mathcal{B}$-measurable function on M.

Such a skew product occurs naturally in the following example: Consider a $\mathbb{Z}^2$-indexed random field P, i.e., a stationary probability measure on $\Omega = \Upsilon^{\mathbb{Z}^2}$, where Υ is some suitable state space. Let $\lambda \in [0, 1]$ be the line with slope λ and y-intercept t, and let $L_{\lambda,t}(z) = (z, [\lambda z + t])$ $(z \in \mathbb{Z})$ be its approximation in the lattice $\mathbb{Z}^2$. Look at the ergodic averages

$$\frac{1}{n} \sum_{i=1}^{n} f(\omega_{L_{\lambda,t}(i)}) \quad (n \in \mathbb{N}) \tag{2}$$

of a function $f \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ along this lattice approximation. Ergodic averages of this type are the key in the proof of a directional Shannon-MacMillan theorem for a $\mathbb{Z}^2$-indexed random field (cf. [6], [7]).

What can we say about P-almost sure or $\mathcal{L}^1(P)$-convergence of the sequence (2)? To make this problem accessible to ergodic theorems we have to find a transformation which captures the stair climbing pattern along the lattice approximation of the line. If the slope is rational, the steps become periodic, and we proceed by combining a finite number of different transformations. In the case of an irrational slope, this method fails. Here, we need to keep track not only of the integer part but also of the fractional part $\{\lambda z + t\}$ in each step. This suggests the skew-product transformation

$$S(t, \omega) := (\tau_\lambda(t), \vartheta_{(1, [\lambda z + t])} \omega) \qquad (t \in \mathbb{T}, \omega \in \Omega),$$

where $\mathbb{T}$ is the one-dimensional torus, equipped with the Borel σ-algebra and the Haar measure, and τ_λ is the translation by λ.

The first natural question is about conditions for the skew product to be ergodic. Applying a standard ergodic theorem would then yield convergence to a constant. However, this shows convergence for *almost all* but not for *all* $t \in \mathbb{T}$, and we want the sequence (2) to converge for any $t \in \mathbb{T}$. Further considerations regarding *sure* convergence with respect to t are needed.

Inspired by the above example, this paper deals with the following two questions: Is the skew product (1) ergodic with respect to the product measure?, and Can we prove, under suitable conditions, that the convergence of the ergodic averages along the skew product is uniform with respect to the first parameter?

The answer to the first question is summarized the results in Theorem 2.5. Under suitable mixing conditions ((C1) and (C2)) we prove that the properties ergodic, weakly mixing, or strongly mixing, are passed on from the transformation in the base to the skew product. As an explicit example we study the case when P is a random field and $(\theta_k)_{k \in K}$ is a group of shift transformations. In this case, (C2) can be insured by assuming tail-trivial for P and a growth condition for the sequence $\left(\sum_{i=0}^{n} \kappa \circ \tau^i\right)_{n \in \mathbb{N}}$ (see Corollary 2.6).

To tackle our second question, we combine two different methods: in the first component we follow the lines of Weyl's and Oxtoby's theorems, and for the second component we use techniques inspired by ergodic theorems along subsequences.

Recall Weyl's classical theorem: The ergodic averages of continuous functions along translations modulo 1 by an irrational number converge uniformly. Oxtoby proved that this holds for

all uniquely ergodic transformations. Weyl's theorem can actually be extended to the class of Riemann-integrable functions, and in Corollary 4.8 we are able to extend Oxtoby's statement to this class as well.

Fixing the first component in our skew product, we obtain an ergodic average along a subsequence. Almost-sure convergence in ergodic theorems along subsequences turned out to be a very subtle question (cf. [4] for an overview), but convergence in $\mathcal{L}^1$ was established under fairly convenient conditions by Blum and Hanson [5]. Extending their theorem to the d-parameter case, and combining it with the uniform-convergence results for the first parameter, we obtain Theorem 4.13. It states uniform convergence in the first and $\mathcal{L}^1(P)$-convergence in the second component, provided that τ is continuous and uniquely ergodic, P is strongly mixing, and a certain technical condition on κ holds. Finally, Theorem 4.15 is a result about uniform convergence in the first, and almost-sure convergence in the second component. The proof is based on an Arzela-Ascoli approach, introduced by Krengel in his short proof of Weyl's theorem (cf. Theorem 2.6 from Chapter 1 of [19]).

Maker (cf. Theorem 7.4 in Chapter 1 of [19]) replaced the function in Birkhoff's ergodic theorem by an approximating sequence of functions. The same can be done for the ergodic theorems for skew products mentioned above, and for the uniform convergence result 4.13. Since the statements and proofs are rather straight forward extensions of what is presented here, we refer the reader to the Corollaries 2.2.3, 2.2.4, 2.2.10, 2.3.1, and 2.3.15 in [6] for details.

Ouline of the paper: The next section starts out reviewing some basic mixing properties and defining the class of skew products considered in this paper. Then, we give the result on ergodicity and mixing properties of the skew product with respect to the product measure. We have a closer look on the case when the second transformation is a shift operator for a random field. Section 3 discusses ergodic theorems along skew products in the case when the first transformation is ergodic, and in the case when it is periodic. Our results about uniform convergence with repect to the first component can be found in the last section. It begins with a closer look on Weyl type theorems, and on ergodic theorems along subsequences. We conclude by applying our results to Example (2).

2 Definition and mixing properties

In this paper, τ is a measure-preserving transformation on a probability space $(M, \mathcal{B}, \mu)$. Let $(\Omega, \mathcal{F}, P)$ be another probability space. The product measure $\overline{P} := \mu \otimes P$ is a probability measure of the product space $\overline{\Omega} := M \times \Omega$ furnished with the product σ-algebra $\overline{\mathcal{F}} := \mathcal{B} \otimes \mathcal{F}$. Let $(\theta_k)_{k \in \mathbb{N}_0^d}$ be a semigroup of measure-preserving transformations on $(\Omega, \mathcal{F}, P)$, i.e., each of the transformations preserves the measure P,

$$\theta_0 = \mathrm{Id}, \quad \text{and } \theta_k \circ \theta_l = \theta_{k+l} \text{ for all } k, l \in \mathbb{N}_0^d. \tag{3}$$

$(\theta_k)_{k \in \mathbb{Z}_0^d}$ is called a d-parameter group if $(\theta_k)_{k \in \mathbb{N}_0^d}$ is a semigroup and $\theta_{-k} = \theta_k^{-1}$ for all $k \in \mathbb{Z}^d$. The following two examples will be used frequently in our settings. To simplify the notation, the two-parameter case is given here; the generalization to d parameters is obvious.

Example 2.1. Let σ_1 and σ_2 be measure-preserving transformations on $(\Omega, \mathcal{F}, P)$. Assume further

that they commute, i.e., $\sigma_1 \circ \sigma_2 = \sigma_2 \circ \sigma_1$. Then

$$\theta_k := \sigma_1^{k^{(1)}} \circ \sigma_2^{k^{(2)}} \qquad \text{for} \quad k = \left(k^{(1)}, k^{(2)}\right) \in \mathbb{N}_0^2 \tag{4}$$

defines a two-paramter semigroup $(\theta_k)_{k \in \mathbb{N}_0^2}$ of measure-preserving transformations on $(\Omega, \mathcal{F}, P)$. Requiring in addition that σ_1 and σ_2 are invertible this construction extends to a two-paramter group $(\theta_k)_{k \in \mathbb{Z}^2}$.

To define the skew product, we still need a function that controls the action of the group or semigroup on $\omega \in \Omega$, depending on $t \in M$. Let κ be a $\mathcal{B}$-measurable function on M with values in K. Then

$$S(t, \omega) = \left(\tau(t), \theta_{\kappa(t)}\omega\right) \qquad (t \in M, \omega \in \Omega),$$

defines a skew product on the product space. In particular, chosing $\kappa \equiv k_0$ for some $k_0 \in K$ yields the uncoupled product of τ and θ_{k_0}. Obviously, S is measurable with respect to the product σ-algebra $\overline{\mathcal{F}}$, and it preserves the product measure $\overline{P}$.

Defining

$$\kappa_n(t) = \sum_{i=0}^{n-1} \kappa \circ \tau(t)^i \qquad \text{for } n \in \mathbb{N}_0, \tag{5}$$

yields for the iterates of the skew product

$$S^n(t, \omega) = \left(\tau^n(t), \theta_{\kappa_n(t)}\omega\right). \tag{6}$$

The following two properties can be shown easily by induction:

$$\kappa_{n+m}(t) = \kappa_n(t) + \kappa_m\left(\tau^n(t)\right) \qquad \text{for all } n, m \in \mathbb{N}_0, \tag{7}$$

$$\kappa_{jn}(t) = \sum_{i=0}^{j(n-1)} \kappa_n \circ \tau^{ij}(t) \qquad \text{for all } j, n \in \mathbb{N}_0. \tag{8}$$

In Example 2.1 we obtain

$$\theta_{\kappa_n(t)} = \sigma_1^{\sum_{i=0}^{n-1} \kappa^{(1)}\left(\tau^i(t)\right)} \circ \sigma_2^{\sum_{i=0}^{n-1} \kappa^{(2)}\left(\tau^i(t)\right)}. \tag{9}$$

The rest of this section is devoted to of ergodicity and mixing properties of the skew product. First, we recall the definitions of ergodicity and mixing properties which will be used (for instance, cf. [24]). A measure-preserving transformation σ on $(\Omega, \mathcal{F}, P)$ is called *ergodic* with respect to P if P is trivial on the σ-algebra

$$\mathcal{J} := \left\{A \in \mathcal{F} \,\middle|\, \sigma^{-1}A = A\right\} \tag{10}$$

of σ-invariant sets. In this case, we also say that P is ergodic with respect to σ. If P is the only invariant measure with respect to σ it is called *uniquely ergodic.* Ergodicity is equivalent to the mixing property

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} P\left(A \cap \sigma^{-i}B\right) - P\left(A\right) P\left(B\right) \right| \xrightarrow{n \to \infty} 0 \qquad \text{for all } A, B \in \mathcal{F}. \tag{11}$$

A classical example is the translation on the torus by an irrational number.

Example 2.2. Let $\mathbb{T} := \mathbb{R}/\mathbb{Z}$ be the one-dimensional torus equipped with the Borel σ-algebra $\mathcal{B}$ and the Lebesgue measure μ. Then the translation by λ modulo 1,

$$\begin{aligned} \tau_\lambda : \quad \mathbb{T} \quad &\longrightarrow \mathbb{T} \\ t \quad &\longmapsto t + \lambda \mod 1, \end{aligned}$$

defines a μ-invariant transformation on $\mathbb{T}$, which is uniquely ergodic, if λ is irrational and which is periodic if λ is rational. We will often use the notation $\{x\}$ for the fractional part of x; in particular, $\tau_\lambda(t) = \{t + \lambda\}$.

Under which conditions can we prove ergodicity of S with respect to $\overline{P}$? Note that, by a simple projection argument, the ergodicity of τ is necessary. Remember that the uncoupled product of two ergodic transformations need not be ergodic. However, it can be shown that the product is ergodic whenever one of the transformations is ergodic and the other one is weakly mixing (cf. [19]). Recall some definitions for mixing properties for (semi-)groups of transformations. $\| \cdot \|$ denotes the maximum norm on $K = \mathbb{N}_0^d$ or $K = \mathbb{Z}^d$, respectively.

Definition 2.3. *A (semi-)group of measure-preserving transformations $(\theta_k)_{k \in K}$ on $(\Omega, \mathcal{F}, P)$ is called* weakly mixing *with respect to P, if*

$$\frac{1}{n} \sum_{i=0}^{n-1} \left| P\left(A \cap \theta_k^{-1} B\right) - P\left(A\right) P\left(B\right) \right| \xrightarrow{\|k\| \to \infty} 0 \qquad \textit{for all } A, B \in \mathcal{F}. \tag{12}$$

It is called strongly mixing *with respect to P, if*

$$P\left(A \cap \theta_k^{-1} B\right) - P\left(A\right) P\left(B\right) \xrightarrow{\|k\| \to \infty} 0 \qquad \textit{for all } A, B \in \mathcal{F}. \tag{13}$$

To make the skew product an ergodic transformation we have to think about assumptions which bring into play the function κ. In [11], N. Friedman introduced the notation of *weakly mixing along a sequence* for transformations, which we translate here to the d-parameter (semi-)group.

Definition 2.4. *A (semi-)group $(\theta_k)_{k \in K}$ of measure-preserving transformations on $(\Omega, \mathcal{F}, P)$, is called* weakly mixing along the K-valued sequence $(k_n)_{n \in \mathbb{N}}$ *with respect to P, if*

$$\frac{1}{n} \sum_{i=0}^{n-1} \left| P\left(A \cap \theta_{k_i}^{-1} B\right) - P\left(A\right) P\left(B\right) \right| \xrightarrow{n \to \infty} 0 \qquad \textit{for all } A, B \in \mathcal{F}. \tag{14}$$

J. Aaronson suggested to extend the question about ergodicity to further mixing properties of the skew product, and the results are summarized in the next theorem. We will make use of two conditions:

(C1) $(\theta_k)_{k \in K}$ is weakly mixing along the sequence $(\kappa_n(t))_{n \in \mathbb{N}}$ for μ-almost all $t \in M$.

(C2) $(\theta_k)_{k \in K}$ is strongly mixing and $(\kappa_n(t))_{n \in \mathbb{N}}$ goes to infinity for μ-almost all $t \in M$.

Obviously, (C2) implies (C1).

Theorem 2.5. *Assume condition (C1) and that τ is ergodic with respect to μ then S is ergodic with respect to $\overline{P}$.*

Assume condition (C1) and that τ is weakly mixing with respect to μ then S is weakly mixing with respect to $\overline{P}$.

Assume condition (C2) and that τ is strongly mixing with respect to μ then S is weakly mixing with respect to $\overline{P}$.

Proof. We begin with the ergodicity. Assume condition (C1) and let τ be ergodic with respect to μ. We will show that for all bounded $\overline{\mathcal{F}}$-measurable functions F and G

$$\frac{1}{n} \sum_{i=1}^{n} \int_{M \times \Omega} F \circ S^i(t,\omega) \cdot G(t,\omega) \, d\overline{P} \xrightarrow{n \to \infty} \widetilde{F} \cdot \widetilde{G},$$

where

$$\widetilde{F} = \int_{M \times \Omega} F(t,\omega) \, d\overline{P} \qquad \text{and} \qquad \widetilde{G} = \int_{M \times \Omega} G(t,\omega) \, d\overline{P}.$$

By (11), this implies ergodicity. It is sufficient to show this for functions which are products of functions on the factors, i.e.,

$$F(t,\omega) = f(t)\Phi(\omega) \qquad \text{and} \qquad G(t,\omega) = g(t)\Psi(\omega), \tag{15}$$

where f and g are bounded $\mathcal{B}$-measurable functions on M and Φ and Ψ are bounded $\mathcal{F}$-measurable functions on Ω. The general case follows by approximation. For the product functions we have,

$$\overline{F} = \overline{f} \cdot \overline{\Phi} \qquad \text{and} \qquad \overline{G} = \overline{g} \cdot \overline{\Psi},$$

with

$$\widetilde{f} = \int_M f(t) \, d\mu, \quad \widetilde{\Phi} = \int_\Omega \Phi(\omega) \, dP, \quad \widetilde{g} = \int_M g(t) \, d\mu, \quad \text{and} \quad \widetilde{\Psi} = \int_\Omega \Psi(\omega) \, dP.$$

We obtain

$$\left| \frac{1}{n} \sum_{i=1}^{n} \int_{M \times \Omega} F \circ S^i(t,\omega) \cdot G(\omega) \, d\overline{P} - \overline{F} \cdot \overline{G} \right|$$

$$= \left| \frac{1}{n} \sum_{i=1}^{n} \int_{M \times \Omega} f(\tau^i(t))\Phi(\theta_{\kappa_i(t)}(\omega))g(t)\Psi(\omega) \, d\overline{P} - \widetilde{f} \cdot \widetilde{\Phi} \cdot \widetilde{g} \cdot \widetilde{\Psi} \right|$$

$$\leq \left| \frac{1}{n} \sum_{i=1}^{n} \int_{M \times \Omega} f(\tau^i(t))g(t)\Phi(\theta_{\kappa_i(t)}(\omega))\Psi(\omega) - f(\tau^i(t))g(t)\widetilde{\Phi}\,\widetilde{\Psi} \, d\overline{P} \right| \tag{16}$$

$$+ \left| \frac{1}{n} \sum_{i=1}^{n} \left(\int_{M \times \Omega} f(\tau^i(t))g(t) \, d\overline{P} - \widetilde{f} \cdot \widetilde{g} \right) \cdot \widetilde{\Phi} \cdot \widetilde{\Psi} \right|. \tag{17}$$

The term (16) can be bounded by

$$\frac{1}{n} \sum_{i=1}^{n} \left| \int_M f(\tau^i(t))g(t) \left(\int_\Omega \Phi(\theta_{\kappa_i(t)}(\omega))\Psi(\omega) \, dP - \widetilde{\Phi} \cdot \widetilde{\Psi} \right) d\mu \right|,$$

and further, using $\| \cdot \|_\infty$ for the supremum norm, by

$$\|f\|_\infty \|g\|_\infty \frac{1}{n} \sum_{i=1}^{n} \left| \int_\Omega \Phi(\theta_{\kappa_i(t)}(\omega))\Psi(\omega) \, dP - \widetilde{\Phi} \cdot \widetilde{\Psi} \right|. \tag{18}$$

This goes to 0 by the mixing condition (C1). The expression (17) is smaller than

$$\left| \frac{1}{n} \sum_{i=1}^{n} \int_M f(\tau^i(t))g(t) \, d\mu - \widetilde{f} \cdot \widetilde{g} \right| |\widetilde{\Phi}| \, |\widetilde{\Psi}|,$$

which goes to 0 by the ergodicity of τ.

Now assume that τ is weakly mixing. Restricting again to the class of product functions (15), we have to show that

$$\frac{1}{n}\sum_{i=1}^{n}\left|\int_{M}\int_{\Omega}f(\tau^{i}(t))\Phi(\theta_{\kappa_i(t)}(\omega))g(t)\Psi(\omega)\,dP\,d\mu\,-\,\widetilde{f}\cdot\widetilde{\Phi}\cdot\widetilde{g}\cdot\widetilde{\Psi}\right|$$

goes to 0 as n goes to infinity. Similar to the estimates in (16), (17) and (18), we obtain the bound

$$\|f\|_{\infty}\|g\|_{\infty}\frac{1}{n}\sum_{i=1}^{n}\left|\int_{\Omega}\Phi(\theta_{\kappa_i(t)}(\omega))\Psi(\omega)\,dP\,-\,\widetilde{\Phi}\cdot\widetilde{\Psi}\right|$$

$$+\frac{1}{n}\sum_{i=1}^{n}\left|\int_{M}f(\tau^{i}(t))g(t)\,d\mu\,-\,\widetilde{f}\cdot\widetilde{g}\right||\widetilde{\Phi}|\,|\widetilde{\Psi}|.$$

By the mixing condition (C1) and the fact that τ is weakly mixing, this goes to 0.

Finally, for the proof of strong mixing, assume condition (C2) and that τ is strongly mixing. We proceed as in the proof of weak mixing, but do not average. $\square$

This section concludes with a closer look at the case when the transformations are induced by shifts. Let Υ be a finite set, and $\Omega = \Upsilon^{\mathbb{Z}^d}$. For any subset J of $\mathbb{Z}^d$ let $\mathcal{F}_J$ denote the σ-algebra generated by all projections $\omega \mapsto \omega(j)$ with $j \in J$, and let $\mathcal{F} := \mathcal{F}_{\mathbb{Z}^d}$. Consider the *shift transformations* $(\vartheta_v)_{v\in\mathbb{Z}^d}$ on Ω, i.e.,

$$\vartheta_v(\omega)(j) := \omega(j+v) \qquad (j \in \mathbb{Z}^d). \tag{19}$$

Let P be a *random field*, i.e., a measure on $(\Omega, \mathcal{F})$ which is invariant with respect to ϑ_v for all $v \in \mathbb{Z}^d$. The *tail field* is the σ-algebra

$$\mathcal{T} := \bigcap_{V\subset\mathbb{Z}^d\text{ finite}}\mathcal{F}_{\mathbb{Z}^d\backslash V},$$

P is called *tail-trivial* if it fulfills a 0-1 law on $\mathcal{T}$. For $v_1, v_2 \in \mathbb{Z}^d$, $\theta_k := \vartheta_{v_1}^{k^{(1)}}\circ\vartheta_{v_2}^{k^{(2)}}$ $(k \in \mathbb{Z}^d)$ defines a measure preserving group of transformations as explained in Example 2.1.

Corollary 2.6. *Let v_1 and v_2 be linear independent vectors in $\mathbb{Z}^d$. Assume that P is tail-trivial and that the sequence $\left(\|\kappa_n(t)\|\right)_{n\in\mathbb{N}}$ goes to infinity for μ-almost all $t \in M$.*
Then, when τ is ergodic, weakly mixing or strongly mixing with respect to μ, S is ergodic, weakly mixing or strongly mixing with repect to $\overline{P}$, respectively.

Proof. We are going to show condition (C2). Define the boxes $\left\{v \in \mathbb{Z}^d \,\middle|\, \|v\| \leq n\right\}$ $(n \in \mathbb{N})$, and let $B \in \mathcal{F}_J$, for some finite subset J of $\mathbb{Z}^d$. Then there is an $m \in \mathbb{N}$ such that $J \subseteq V_m$. Setting $m(n) := \kappa_n^{(1)}(t)\,v_1 + \kappa_n^{(2)}(t)\,v_2$ we observe that the translated sets $B - m(n)$ are contained in $V_{\widetilde{m}(n)}^c$, where $\widetilde{m}(n) = (m(n) - 2m)\vee 0$. For any $A \in \mathcal{F}$ we obtain

$$\left|P\left(A\cap\theta_{\kappa_n(t)}^{-1}B\right) - P(A)\,P(B)\right|$$

$$= \left|P\left(A\cap\vartheta_{v_1}^{-\kappa_n^{(1)}(t)}\circ\vartheta_{v_2}^{-\kappa_n^{(2)}(t)}B\right) - P(A)\,P(B)\right|$$

$$= \left|P\left(A\cap\vartheta_{\kappa_n^{(1)}(t)\,v_1+\kappa_n^{(2)}(t)\,v_2}^{-1}B\right) - P(A)\,P(B)\right|$$

$$\leq \sup_{C\in\mathcal{F}_{\mathbb{Z}^d\backslash V_{m(n)}}}\left|P(A\cap C) - P(A)P(C)\right|. \tag{20}$$

By the assumptions on v_1, v_2 and κ, $\|\widetilde{m}(n)\|$ goes to infinity. By Proposition 7.9 in [13] tail-triviality is equivalent to *short-range correlations,* i.e.,

$$\sup_{C \in \mathcal{F}_{\mathbb{Z}^d \setminus V_n}} \left| P(A \cap C) - P(A)P(C) \right| \xrightarrow{n \to \infty} 0. \tag{21}$$

Applying this to (20) concludes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3   Ergodic theorems with skew products

Applying Birkhoff's ergodic theorems to the skew-product transformation S yields, for any function $F \in \mathcal{L}^1(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$,

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)}\omega\right) \xrightarrow{n \to \infty} \overline{E}[F|\mathcal{J}] \tag{22}$$

$\overline{P}$-almost surely and in $\mathcal{L}^1(\overline{P})$, where $\mathcal{J}$ is the σ-algebra of all S-invariant sets in $\overline{\mathcal{F}}$. We study this limit more closely. Two different cases will be discussed: when the transformation τ is periodic and when it is ergodic.

Begin with the latter case. Summarizing (22) and Theorem 2.5 leads to

Theorem 3.1. *Assume that τ is ergodic with respect to μ and that the condition (C1) is fullfilled. Then for any function $F \in \mathcal{L}^1(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$,*

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)}\omega\right) \xrightarrow{n \to \infty} \overline{E}[F]$$

$\overline{P}$-almost surely and in $\mathcal{L}^1(\overline{P})$.

We shall illustrate the theorem with two examples.

Example 3.2. *(Shifts in the fibre)*
Consider the situation of Corollary 2.6, and suppose that τ is ergodic with respect to μ. Then by Theorem 3.1, for any function $F \in \mathcal{L}^1(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$,

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \vartheta_{\kappa_n^{(1)}(t)\, v_1 + \kappa_n^{(2)}(t)\, v_2}\right) \xrightarrow{n \to \infty} \overline{E}[F]$$

$\overline{P}$-almost surely and in $\mathcal{L}^1(\overline{P})$.

Example 3.3. *(Irrational translation on the torus in the base)*
Let τ_λ be the translation on the torus defined in Example 2.2. Suppose that λ is irrational. Then τ_λ is ergodic. Choose $(\theta_k)_{k \in K}$ and κ for which the condition (C1) from the last section holds. Then by Theorem 2.5, S is ergodic as well, and Theorem 3.1 tells us that for any integrable function F on $(\mathbb{T} \times \Omega, \mathcal{B} \otimes \mathcal{F}, \mu \otimes P)$,

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(t + i\lambda \mod 1, \theta_{\kappa_i(t)}\omega\right) \xrightarrow{n \to \infty} \int_0^1 E[F(t, \cdot)]\, dt \tag{23}$$

for $\mu \otimes P$-almost all $(t, \omega) \in \mathbb{T} \times \Omega$ and in $\mathcal{L}^1(\mathbb{T} \times \Omega, \mathcal{B} \otimes \mathcal{F}, \mu \otimes P)$.

Let us now consider the case of τ periodic, i.e., where $\tau^q = \mathrm{Id}$ for some $q \in \mathbb{N}$. To see how this differs from the ergodic case, we calculate the iterates of the skew product explicitly.

Lemma 3.4. *Assume that τ is periodic with $q \in \mathbb{N}$. Then for all $j \in \mathbb{Z}$ and all $\nu \in \{0, 1, ..., q-1\}$,*

(i) $\kappa_{jq+\nu}(t) = j\kappa_q(t) + \kappa_\nu(t)$ for all $t \in M$,

(ii) $\theta_{\kappa_{jq+\nu}(t)} = \left(\theta_{\kappa_q(t)}\right)^j \circ \theta_{\kappa_\nu(t)}$ for all $\omega \in \Omega$,

(iii) $S^{jq+\nu}(t, \omega) = \left(\tau^\nu(t), \theta_{\kappa_\nu(t)} \circ (\theta_{\kappa_q(t)})^j \omega\right)$ for all $t \in M$ and all $\omega \in \Omega$.

Proof. By (8) and the periodicity of τ, $\kappa_{jq} = j\kappa_q$. Applying (7) and the periodicity again yields (i). The second statement (ii) is an immediate consequence of (i) and (3). To prove (iii), obtain by (6) that $S^{jq}(t, \omega) = \left(\tau^{jq+\nu}, \theta_{\kappa_{jq+\nu}(t)}\omega\right)$, and apply (i), (ii), and the periodicity of τ. $\qquad\square$

The ergodic theorem for our skew product in the case when τ is periodic has the following form.

Theorem 3.5. *Assume τ is periodic with period $q \in \mathbb{N}$, and $F \in \mathcal{L}^1(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$. Denote by $\mathcal{J}_t$ the σ-algebra of $\theta_{\kappa_q(t)}$-invariant sets in $\mathcal{F}$. Then*

$$\frac{1}{n}\sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)}\omega\right) \xrightarrow{\ n\to\infty\ } \frac{1}{q}\sum_{\nu=0}^{q-1} E\left[F\left(\tau^\nu(t), \theta_{\kappa_\nu(t)}(\cdot)\right)\big|\, \mathcal{J}_t\right],$$

for μ-almost all $t \in M$, and for P-almost all $\omega \in \Omega$ and in $\mathcal{L}^1(P)$. If $\theta_{\kappa_q(t)}$ is ergodic with respect to P for μ-almost all $t \in M$, then the limit simplifies to

$$\frac{1}{q}\sum_{\nu=0}^{q-1} E\left[F\left(\tau^\nu(t), \cdot\right)\right].$$

Proof. Due to (6) we have

$$A_n F := \frac{1}{n}\sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)}\omega\right) = \frac{1}{n}\sum_{i=0}^{n-1} F \circ S^i \quad (n \in \mathbb{N}). \tag{24}$$

The first step in proving the theorem is to show that we can restrict ourselves to a subsequence of the form $(mq)_{m\in\mathbb{N}}$. Any $n \in \mathbb{N}$ can be represented as $n = mq+\nu$, with $m \in \mathbb{N}$ and $\nu \in \{0, 1, ..., q-1\}$, and we may break down $A_n F$ to

$$A_n F = \frac{mq}{mq+\nu}\left(\frac{1}{mq}\sum_{i=0}^{mq-1} F \circ S^i + \frac{1}{mq}\sum_{i=mq}^{mq+\nu-1} F \circ S^i\right).$$

Since

$$\frac{mq}{mq+\nu} \longrightarrow 1 \quad \text{and} \quad \frac{1}{mq}\sum_{i=mq}^{mq+\nu-1} F \circ S^i \xrightarrow{\ m\to\infty\ } 0$$

our question of the limit behavior of (24) reduces to the study of the limit along the subsequence.

For mq instead of n in (24) we get for the ergodic averages

$$A_{mq} F = \frac{1}{q}\sum_{\nu=0}^{q-1}\frac{1}{m}\sum_{j=0}^{m-1} F \circ S^{jq+\nu}.$$

Thus we need only to show that for every $\nu \in \{0, 1, ..., q-1\}$,

$$\frac{1}{m} \sum_{j=0}^{m-1} F \circ S^{jq+\nu} \xrightarrow{m \to \infty} E\left[F\left(\tau^\nu(t), \theta_{\kappa_\nu(t)}(\cdot)\right) \big| \mathcal{J}_t\right].$$

For $\nu \in \{0, 1, ..., q-1\}$ fixed define

$$A_m^{(\nu)} F := \frac{1}{m} \sum_{j=0}^{m-1} F \circ S^{jq+\nu}.$$

By Lemma 3.4, the identity

$$A_m^{(\nu)} F(t, \omega) = \frac{1}{m} \sum_{j=0}^{m-1} F\left(\tau^\nu(t), \theta^{\kappa_\nu(t)} \circ (\theta^{\kappa_q(t)})^j \omega\right).$$

holds for all $t \in M$ and $\omega \in \Omega$. For μ-almost all $t \in M$ fixed, we can define an $\mathcal{F}$-measurable transformation on $(\Omega, \mathcal{F}, P)$ by $\vartheta_t := \theta_{\kappa_q(t)}$, and an $\mathcal{L}^1(\Omega, \mathcal{F}, P)-$ function by $f_t^{(\nu)}(\omega) := F\left(\tau^\nu(t), \theta_{\kappa_\nu(t)}\omega\right)$ $(\omega \in \Omega)$. Now the ergodic averages $A_m^{(\nu)} F(t, \cdot)$ take the form

$$\frac{1}{m} \sum_{j=0}^{m-1} f_t^{(\nu)} \circ \vartheta_t^j.$$

Applying Birkhoff's ergodic theorem we get

$$\lim_{m \to \infty} A_m^{(\nu)} F(t, \cdot) = E[f_t^{(\nu)} | \mathcal{J}_t] = E\left[F\left(\tau^\nu(t), \theta_{\kappa_\nu(t)}(\cdot)\right) | \mathcal{J}_t\right]$$

P-almost surely and in $\mathcal{L}^1(P)$. Combining this with the preceding considerations yields the desired result. In the ergodic case, $\mathcal{J}_t$ is trivial, and the last expression reduces to $E\left[F\left(\tau^\nu(t), \theta_{\kappa_\nu(t)}(\cdot)\right)\right]$. Then the last statement of the theorem follows from the invariance of P under θ. □

Example 3.6. *(Rational translation on the torus in the base)*
Given the conditions of Example 3.3, but with λ a *rational* number. There is a unique representation $\lambda = \frac{p}{q}$, where $p \in \mathbb{Z}, q \in \mathbb{N}$, and p and q have no common divisor. Then τ is periodic with period q and respects the partition $[0, \frac{1}{q}), [\frac{1}{q}, \frac{2}{q}), ..., [\frac{q-1}{q}, 1)$ of $\mathbb{T}$, i.e., for every $\nu \in \{1, ..., q-1\}$ there is a $\widetilde{\nu} \in \{1, ..., q-1\}$ such that $\tau_\lambda([\frac{\nu-1}{q}, \frac{\nu}{q})) = [\frac{\nu-1}{q}, \frac{\nu}{q})$.

Assume that, for P-almost all $\omega \in \Omega$, $F(\cdot, \omega)$ is a step function with respect to the above partition. Then the limit in Theorem 3.5 can be rewritten as

$$\frac{1}{q} \sum_{\nu=0}^{q-1} E\left[F\left(\tau_\lambda^\nu(t), \cdot\right) \big| \mathcal{J}_t\right] = \int_0^1 E\left[F\left(\tau_\lambda^\nu(t), \cdot\right) \big| \mathcal{J}_t\right] dt = \int_0^1 E[F(t, \cdot) | \mathcal{J}_t] dt,$$

and if P is ergodic with respect to θ_{κ_q},

$$\frac{1}{q} \sum_{\nu=0}^{q-1} E\left[F\left(\tau_\lambda^\nu(t), \cdot\right)\right] = \int_0^1 E\left[F\left(\tau_\lambda^\nu(t), \cdot\right)\right] dt = \int_0^1 E[F(t, \cdot)] dt.$$

4   Uniform convergence

In addition to the assumptions at the beginning of Section 2 we suppose that M is a compact separable metric space endowed with metric d, and $\mathcal{B}$ is the Borel σ-algebra on M for the topology

induced by d. In the last section we investigated ergodic theorems for the skew product $S(t, \omega) = (\tau(t), \theta_{\kappa(t)})$ on the product space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$, and we obtained

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \omega\right) = \overline{E}[F|\mathcal{J}](t, \omega),$$

where the convergence was in $\mathcal{L}^1(\overline{P})$ or $\overline{P}$-almost surely.

This section addresses the question of *sure* convergence with respect to the first parameter. Recall that the convergence of ergodic averages need not be true *everywhere*, even if we are in a compact topological space and both the transformation and the function are continuous. This may be seen in simple examples, such as the Bernoulli shift: Consider the space $\{0, 1\}^{\mathbb{N}}$, with σ-algebra $\mathcal{A} = \otimes_{i=1}^{\infty} \sigma(\{0\})$, and measure $\mu = \otimes_{i=1}^{\infty} \frac{1}{2}(\delta_0 + \delta_1)$. Define a transformation on $\{0, 1\}^{\mathbb{N}}$ by $T(t_1, t_2, t_3, ...) = (t_2, t_3, t_4, ...)$, for $t = (t_1, t_2, t_3, ...) \in \{0, 1\}^{\mathbb{N}}$. The function f given by $f(t_1, t_2, t_3, ...) := t_1$ for $(t_1, t_2, t_3, ...) \in \{0, 1\}^{\mathbb{N}}$ is bounded and measurable with respect to $\mathcal{A}$. By Birkhoff's ergodic theorem we obtain

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} f \circ T^i(t) = \frac{1}{2} \quad \text{ for } \mu\text{-almost all } t \in \{0, 1\}^{\mathbb{N}},$$

while for $t = (0, 0, 0, ...)$ this limit is 0.

Which conditions guarantee sure convergence in the first parameter? We will be asking a little more than this, namely about *uniform* convergence in t. We are interested in results of the type

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \omega\right) \xrightarrow{n \to \infty} \overline{E}[F|\mathcal{J}](t, \omega) \qquad \text{uniformly in } t \in M \tag{25}$$

in $\mathcal{L}^1(P)$. To put it another way, we ask that

$$\lim_{n \to \infty} \sup_{t \in M} \left\| \frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)}(\cdot)\right) - \overline{E}[F|\mathcal{J}](t, \cdot) \right\|_{\mathcal{L}^1(P)} = 0. \tag{26}$$

In addition, we investigate whether (25) may take place P-almost surely, i.e., for P-almost all $\omega \in \Omega$,

$$\lim_{n \to \infty} \sup_{t \in M} \left| \frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \omega\right) - \overline{E}[F|\mathcal{J}](t, \omega) \right| = 0. \tag{27}$$

Again, we consider two different cases: when τ is periodic and when it is ergodic. The first case is simple. Establishing an ergodic theorem along the orbits of the skew product, was just a matter of rearranging the sum according to the periodic structure and then applying a classical ergodic theorem only to the second component (see proof of Theorem 3.5). As to the questions about sure and uniform convergence we obtain

Corollary 4.1. *Let τ be periodic with $q \in \mathbb{N}$. Assume $F \in \mathcal{L}^1(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$ such that $F(t, \cdot), F_{ni}(t, \cdot) \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ for all $t \in M$. Then*

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \cdot\right) \xrightarrow{n \to \infty} \frac{1}{q} \sum_{\nu=0}^{q-1} E\left[F\left(\tau^\nu(t), \theta_{\nu_q(t)} \cdot\right) \big| \mathcal{J}_t\right] \qquad \text{uniformly in } t \in M$$

P-a.s. and in $\mathcal{L}^1(P)$, where $\mathcal{J}_t$ denotes the σ-algebra of $\theta_{\kappa_q(t)}$-invariant sets in $\mathcal{F}$.

If P is ergodic with respect to $\theta_{\kappa_q(t)}$ the limit simplifies to

$$\frac{1}{q}\sum_{\nu=0}^{q-1} E\left[F\left(\tau^\nu(t),\,\cdot\,\right)\right].$$

The ergodic case is more delicate. We begin with a more careful investigation of ergodic theorems on the single spaces $(M,\mathcal{B},\mu)$ and $(\Omega,\mathcal{F},P)$. Later, these results will be used to tackle the ergodic theorems for the skew product. With regard to the first component, where the space is assumed to be compact and metrizable, we recall some results about uniform convergence. In these theorems the transformation and the function are both subject to continuity assumptions. With an eye toward later applications we also ask for uniform convergence within the class of Riemann-integrable functions.

Ignoring the first component of the skew product we end up with an ergodic average along subsequences. We therefore derive a d-parameter group version of Blum and Hanson's $\mathcal{L}^p$-ergodic theorem along subsequences. Coming back to the ergodic averages of the skew product, we first study the $\mathcal{L}^p$-convergence. The result in Theorem 4.13 is a consequence of the one-dimensional case and a condition on the coupling sequence $(\kappa_n(t))_{n\in\mathbb{N}}$, uniformly in t. The last part of this section addresses the question of P-almost sure convergence can take place uniformly in $t \in M$. The result in Theorem 4.15 requires equicontinuity of the sequence of ergodic averages.

The classical example for an ergodic theorem that gives a statement about uniform convergence is a theorem of Weyl. We will recall the one-dimensional version here. The metric space is here the torus $\mathbb{T}$, with metric $d(s,t) := |s-t|$ $(s,t \in \mathbb{T})$, and τ the translation on $\mathbb{T}$ by λ (see Example 2.2).

Theorem 4.2. (Weyl) *Let λ be an irrational number and f a continuous function on $\mathbb{T}$. Then*

$$\frac{1}{n}\sum_{i=0}^{n-1} f \circ \tau_\lambda^i \xrightarrow{n\to\infty} \int_0^1 f(t)\,dt \qquad uniformly. \tag{28}$$

To prove Weyl's theorem, Krengel (cf. Theorem 2.6 in Paragraph 1.2.3 in [19]) uses an Arzela-Ascoli technique which we will make use of at the end of this section.

Theorem 4.3. (Krengel) *Let $\tau : M \to M$ be continuous, and assume that f is a function on M, such that the functions*

$$\frac{1}{n}\sum_{i=0}^{n-1} f \circ \tau^i \quad (n \in \mathbb{N}) \tag{29}$$

are equicontinuous on M. Then

$$\frac{1}{n}\sum_{i=0}^{n-1} f \circ \tau^i \xrightarrow{n\to\infty} E[f|\mathcal{J}] \qquad uniformly,$$

where $\mathcal{J}$ denotes the σ-algebra of τ-invariant sets.

Together with the following Lemma this yields Weyl's theorem.

Lemma 4.4. *Let f be a continous function on M, and $\tau : M \to M$ Lipschitz-continuous with Lipschitz constant $c \leq 1$. Then the functions in (29) are equicontinuous on M.*

Proof. We have to show that for every $\varepsilon > 0$ there is a $\delta > 0$ such that for all $n \in \mathbb{N}$ and all $s, t \in M$ with $d(s, t) < \delta$

$$\frac{1}{n} \left| \sum_{i=0}^{n-1} f(\tau^i(s)) - f(\tau^i(t)) \right| < \varepsilon. \tag{30}$$

Fix $\varepsilon > 0$. We will show that there is a $\delta > 0$, such that $|f(\tau^i(s)) - f(\tau^i(t))| < \varepsilon$ for all $d(s, t) < \delta$. Since M is compact, f must be uniformly continuous, i.e., there is a $\delta > 0$ such that for all $x, y \in M$ with $d(x, y) < \delta$ we have $|f(x) - f(y)| < \varepsilon$. By assumption, $d(\tau(s), \tau(t)) \leq c\, d(s, t)$ for all $s, t \in M$, and therefore, $d(\tau^i(s), \tau^i(t)) \leq c^i\, d(s, t) \leq d(s, t)$ for all $s, t \in M$, and for all $i \in \mathbb{N}$. $\square$

We shall ask whether we could replace the assumption of continuity of the function f in Weyl's theorem by a weaker condition. It is certainly not true for all measurable functions, which can be seen in a simple example: Fix $t_0 \in \mathbb{T}$. Its orbit under τ is the set $\mathcal{O} := \{\tau^n(t_0) | n \in \mathbb{N}_0 \}$. Defining the function $f := 1_{\mathcal{O}}$, we obtain

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} 1_{\mathcal{O}} \circ \tau^i(t) = 1 \quad \text{for all } t \in \mathcal{O},$$

but $\int_0^1 1_{\mathcal{O}}(t)\, dt = 0$, which shows that (28) is wrong for f.

The Lebesgue measure is the only probability measure on $(\mathbb{T}, \mathcal{B})$, which is invariant with respect to τ_λ. This observation yields to the following

Definition 4.5. *A continuous transformation τ of a compact metrizable space is called* uniquely ergodic *if it has only one invariant Borel measure.*

It can be shown that this measure must be ergodic, which implies that the ergodic averages of an integrable function converge almost surely to a constant. An extensive discussion of the connections between unique ergodicity and the *uniform* convergence for continuous functions can be found, for instance, in Chapter 4.1.e. of [18] or Theorem 6.19 in [24]. In particular, there is the following

Theorem 4.6. *(Oxtoby) Let $\tau : M \to M$ be continuous and uniquely ergodic with invariant measure μ. Then for any continuous function f on M,*

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i \xrightarrow{n \to \infty} \int_M f\, d\mu \qquad uniformly.$$

The converse need not be true unless further conditions are imposed, such as topological transitivity of τ or constancy of the limit. Below Theorem 2.7 in Chapter 1 of [19], Krengel mentions that Weyl's theorem is somtimes spelled out for to the class of Riemann-integrable functions. Actually, it was proved by de Bruijn and Post [8] that the function is Riemann-integrable if and only if the convergence is uniform. This also follows from our next proposition. We ask the following question: Considering uniform convergence of the ergodic averages along a continuous transformation on a compact real intervall, can we pass automatically from the class of continuous functions to the class of functions which are integrable in the sense of Riemann?

Proposition 4.7. *Let* $a, b \in \mathbb{R}$, $a < b$, μ *a measure on* $([a, b], \mathcal{B}([a, b]))$ *which is absolutely continuous with respect to Lebesgue measure, with a continuous density. Let* $\tau : [a, b] \to [a, b]$ *be continuous. Assume that for any continuous function* f *on* $[a, b]$

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i \xrightarrow{n \to \infty} \int_a^b f \, d\mu \qquad uniformly. \tag{31}$$

Then the convergence holds as well for any function which is integrable in the sense of Riemann.

Proof. With no loss of generality we can assume that μ is the Lebesgue measure on M, and that $M = [0, 1]$. The first step is to construct a sequence of functions $(h_m)_{m \in \mathbb{N}}$ on $[0, 1]$ such that

$$h_m \text{ is continuous on } [0, 1] \text{ and } h_m \geq f \text{ for all } m \in \mathbb{N}, \tag{32}$$

$$\text{and } \lim_{m \to \infty} \int_0^1 h_m \, dt = \int_0^1 f \, dt. \tag{33}$$

Consider the partition of $[0, 1]$ which is given by $\Delta_i = [\frac{i-1}{m}, \frac{i}{m})$ for $i = 1, ..., m - 1$ and $\Delta_m = [\frac{m-1}{m}, 1]$. Since f is Riemann-integrable on $[0, 1]$, there is a constant $c \geq 0$ and a sequence of step functions on $[0, 1)$,

$$f_m(t) = \sum_{i=1}^m \alpha_m^{(i)} \, 1_{\Delta_i}(t) \quad (m \in \mathbb{N})$$

with $|\alpha_m^{(i)}| \leq c$ for $i \in \{1, ..., m\}, m \in \mathbb{N}$, for which

$$f_m \geq f \text{ for all } m \in \mathbb{N} \text{ and } \lim_{m \to \infty} \int_0^1 f_m \, dt = \int_0^1 f \, dt. \tag{34}$$

A slight modification,

$$h_m(t) = \begin{cases} \alpha_m^{(i-1)} + m^2 \left(\alpha_m^{(i)} - \alpha_m^{(i-1)} \right) \left(t - \frac{i}{m} + \frac{1}{m^2} \right) & \text{if } t \in [\frac{i}{m} - \frac{1}{m^2}, \frac{i}{m}), \ i \in \{1, ...m\} \\ f_m(t) & \text{otherwise}, \end{cases}$$

defines a sequence $(h_m)_{m \in \mathbb{N}}$ that fulfills the two conditions of (32). As for the request (33), we first observe

$$\begin{aligned} \left| \int_0^1 h_m(t) - f_m(t) \, dt \right| &= \sum_{i=1}^m \int_{\frac{i}{m} - \frac{1}{m^2}}^{\frac{i}{m}} |h_m(t) - f_m(t)| \, dt \\ &= \sum_{i=1}^m \int_{\frac{i}{m} - \frac{1}{m^2}}^{\frac{i}{m}} m^2 |\alpha_m^{(i+1)} - \alpha_m^{(i)}| \left(t - \frac{i}{m} + \frac{1}{m^2} \right) dt \\ &\leq \sum_{i=1}^m \int_0^{\frac{1}{m^2}} m^2 \, 2 \, c \, t \, dt = m \, c \, m^2 \, \frac{1}{m^4} = \frac{c}{m}, \end{aligned}$$

In the same way, we construct a sequence of functions $(g_m)_{m \in \mathbb{N}}$ on $[0, 1]$ such that

$$g_m \text{ is continuous on } [0, 1] \text{ and } g_m \leq f \text{ for all } m \in \mathbb{N} \tag{35}$$

$$\text{and } \lim_{m \to \infty} \int_0^1 g_m \, dt = \int_0^1 f \, dt. \tag{36}$$

For all $n \in \mathbb{N}, m \in \mathbb{N}$, and $t \in [0, 1]$

$$\frac{1}{n} \sum_{i=0}^{n-1} g_m \circ \tau^i(t) \leq \frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i(t) \leq \frac{1}{n} \sum_{i=0}^{n-1} h_m \circ \tau^i(t).$$

Fix $\delta > 0$. By (36) and (33) we can find $m \in \mathbb{N}$ large enough such that

$$\int_0^1 g_m\, dt < \int_0^1 f\, dt + \frac{\delta}{2} \qquad \text{and} \qquad \int_0^1 f\, dt - \frac{\delta}{2} < \int_0^1 h_m\, dt.$$

Since g_m and h_m are continuous, assumption (31) tells us that there is an $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and $s \in [0, 1]$,

$$\int_0^1 g_m\, dt - \frac{\delta}{2} \leq \frac{1}{n} \sum_{i=0}^{n-1} g_m \circ \tau^i(s) \leq \int_0^1 g_m\, dt + \frac{\delta}{2}$$

and

$$\int_0^1 h_m\, dt - \frac{\delta}{2} \leq \frac{1}{n} \sum_{i=0}^{n-1} h_m \circ \tau^i(s) \leq \int_0^1 h_m\, dt + \frac{\delta}{2}.$$

Putting all these inequalities together yields

$$\int_0^1 f\, dt - \delta \leq \frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i(s) \leq \int_0^1 f\, dt + \delta$$

for all $n \geq n_0$ and for all $s \in [0, 1]$, and the assertion of the corolary follows by taking δ to 0. $\qquad \square$

Applying the last Proposition to the situation of Theorem 4.6 we obtain

Corollary 4.8. *Let $a, b \in \mathbb{R}$, $a < b$, and $\tau : [a, b] \to [a, b]$ continuous and uniquely ergodic with invariant measure μ. Assume that μ is absolutely continuous with respect to Lebesgue measure, with a continuous density. Then for any function f on M which is integrable in the sense of Riemann,*

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i \xrightarrow{n \to \infty} \int_0^1 f\, d\mu \qquad \textit{uniformly.}$$

Remember that the goal of this section was to understand the different modes of convergence of the ergodic averages of a function F on the product space along the orbits of the skew product transformation. We asked about uniform convergence with respect to the first component and $\mathcal{L}^1$-convergence with respect to the second. The above discussion tells us something about the uniform convergence with respect to the first parameter. Now we turn to the second parameter. Choosing for F a function which is constant in t brings us back again to a one-dimensional situation: Fix $t \in M$. Defining a function on Ω by $f(\omega) := F(t, \omega)$ reduces the ergodic averages to $\frac{1}{n} \sum_{i=0}^{n-1} f\big(\theta_{\kappa_i(t)} \omega\big)$, which we can view as a sort of ergodic average along the subsequence $(\kappa_i(t))_{i \in \mathbb{N}}$. Concerning $\mathcal{L}^p$-convergence of classical ergodic averages along subsequences, there is the following characterization.

Theorem 4.9. (Blum & Hanson) *Let T be a transformation on $(\Omega, \mathcal{F})$. Suppose that T is invertible and that both, T and T^{-1} preserve P. Then P is strongly mixing with respect to T if and only if for all p, $1 \leq p < \infty$, every strictly increasing sequence $(m_i)_{i \in \mathbb{N}}$ of integers, and every function $f \in \mathcal{L}^p(\Omega, \mathcal{F}, P)$,*

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ T^{m_i} \xrightarrow{n \to \infty} E[f] \qquad \textit{in } \mathcal{L}^p(P).$$

The key to the proof is the following

Lemma 4.10. *Under the assumptions of Theorem 4.9 and supposing that P is strongly mixing with respect to T we have for all $A \in \mathcal{F}$*

$$\frac{1}{n} \sum_{i=0}^{n-1} 1_A \circ T^{k_i} \xrightarrow{n \to \infty} P(A) \qquad in \ \mathcal{L}^2(P),$$

for every strictly increasing sequence $(k_i)_{i \in \mathbb{N}}$.

It remains to ask how we can carry over Blum and Hanson's theorem to the case of a d-parameter group of transformations $(\theta)_{k \in \mathbb{Z}^d}$, instead of iterates of a transformations T. What we need is a condition on the $\mathbb{Z}^d$-valued sequence $(k_i)_{i \in \mathbb{N}}$ which replaces the strict monotonicity imposed on $(m_i)_{i \in \mathbb{N}}$. With an eye toward later applications on the product space we generalize the result further by admitting a whole family of sequences and showing that the $\mathcal{L}^2$-convergence takes place uniformly over this family. Recall that $\| \cdot \|$ denotes the maximum norm in $\mathbb{Z}^d$. The analogue of Lemma 4.10 is the following

Lemma 4.11. *Assume that P is strongly mixing with respect to $(\theta_k)_{k \in \mathbb{Z}^d}$, and let $(k_n(t))_{n \in \mathbb{N}}$ $(t \in I)$ be a family of sequences with values in $\mathbb{Z}^d$, for which for all $m \in \mathbb{N}$,*

$$\lim_{n \to \infty} \frac{1}{n^2} \sup_{t \in I} \left| \left\{ 1 \leq i, j \leq n \ \middle| \ \|k_i(t) - k_j(t)\| \leq m \right\} \right| = 0. \tag{37}$$

Then for all $A \in \mathcal{F}$,

$$\sup_{t \in I} \left\| \frac{1}{n} \sum_{i=0}^{n-1} 1_A \circ \theta_{k_i(t)} - P(A) \right\|_{\mathcal{L}^2(P)}^2 \xrightarrow{n \to \infty} 0.$$

Proof. For every $A \in \mathcal{F}$ and $t \in I$ we obtain by simple calculations

$$\left\| \frac{1}{n} \sum_{i=0}^{n-1} 1_A \circ \theta_{k_i(t)} - P(A) \right\|_{\mathcal{L}^2(P)}^2$$

$$= \int_\Omega \frac{1}{n^2} \sum_{i,j=0}^{n-1} \left(1_A \circ \theta_{k_i(t)} - P(A) \right) \left(1_A \circ \theta_{k_j(t)} - P(A) \right) dP$$

$$= \frac{1}{n^2} \sum_{i,j=0}^{n-1} \left[\int_\Omega \left(1_A \circ \theta_{k_i(t)} \, 1_A \circ \theta_{k_j(t)} \right) dP - P(A) \int_\Omega \left(1_A \circ \theta_{k_i(t)} + 1_A \circ \theta_{k_j(t)} \right) dP + P(A)^2 \right]$$

$$= \frac{1}{n^2} \sum_{i,j=0}^{n-1} \left(P\left(\theta_{k_i(t)}^{-1} A \cap \theta_{k_j(t)}^{-1} A \right) - P(A)^2 \right),$$

The last term may be bound by

$$\leq \frac{1}{n^2} \sum_{i,j=0}^{n-1} \left| P\left(\theta_{k_i(t) - k_j(t)}^{-1} A \cap A \right) - P(A)^2 \right|. \tag{38}$$

since, by (3), $P\left(\theta_{k_i(t)}^{-1} A \cap \theta_{k_j(t)}^{-1} A \right) = P\left(\theta_{k_i(t) - k_j(t)}^{-1} A \cap A \right)$. It remains to show that the upper bound 38 converges to 0. Fix $\varepsilon > 0$. Due to the mixing condition (13) there is an $m \in \mathbb{N}$ such that

$$\left| P\left(\theta_{k_i(t) - k_j(t)}^{-1} A \cap A \right) - P(A)^2 \right| < \frac{\varepsilon}{2} \qquad \text{for all } k \in \mathbb{Z}^2 \text{ with } \| k_i(t) - k_j(t) \| > m,$$

and

$$\left| P\left(\theta_k^{-1} A \cap A\right) - P\left(A\right)^2 \right| < \frac{\varepsilon}{2} \qquad \text{for all } k \in \mathbb{Z}^2 \text{ with } \|k\| \geq m,$$

and by assumption (37) there is a $n_0 \in \mathbb{N}$ such that

$$\frac{1}{n^2} \sup_{t \in I} \left| \left\{ 1 \leq i, j \leq n \;\middle|\; \|k_i(t) - k_j(t)\| \leq m \right\} \right| < \frac{\varepsilon}{2} \quad \text{for all } n \geq n_0.$$

Applying the last two inequalities to (38) yields for all $n \geq n_0$

$$\begin{aligned}
\sup_{t \in I} &\left\| \frac{1}{n^2} \sum_{i=0}^{n-1} 1_A \circ \theta_{k_i(t)} - P(A) \right\|_{\mathcal{L}^2(\Omega, \mathcal{F}, P)}^2 \\
&\leq \frac{1}{n^2} \sup_{t \in I} \sum_{0 \leq i,j \leq n-1} 1_{\{\|k_i(t)-k_j(t)\| \leq m\}} \left| P\left(\theta_{k_i(t)-k_j(t)}^{-1} A \cap A\right) - P(A)^2 \right| \\
&\quad + \frac{1}{n^2} \sup_{t \in I} \sum_{0 \leq i,j \leq n-1} 1_{\{\|k_i(t)-k_j(t)\| > m\}} \left| P\left(\theta_{k_i(t)-k_j(t)}^{-1} A \cap A\right) - P(A)^2 \right| \\
&< \frac{1}{n^2} \sup_{t \in I} \left| \left\{ 1 \leq i, j \leq n \;\middle|\; \|k_i(t) - k_j(t)\| \leq m \right\} \right| + \frac{1}{n^2} n^2 \frac{\varepsilon}{2} \\
&< \varepsilon,
\end{aligned}$$

and the assertion of the lemma follows by letting ε to 0. $\qquad\square$

Theorem 4.12. *Assume that P is strongly mixing with respect to $(\theta_k)_{k \in \mathbb{Z}^d}$, and let $(k_n(t))_{n \in \mathbb{N}}$ be a family of sequences with values in $\mathbb{Z}^d$, for which for all $m \in \mathbb{N}$,*

$$\lim_{n \to \infty} \frac{1}{n^2} \sup_{t \in I} \left| \left\{ 1 \leq i, j \leq n \;\middle|\; \|k_i(t) - k_j(t)\| \leq m \right\} \right| = 0.$$

Then for $1 \leq p < \infty$ and for any $f \in \mathcal{L}^p(\Omega, \mathcal{F}, P)$,

$$\frac{1}{n} \sum_{i=0}^{n-1} F \circ \theta_{k_i(t)} \xrightarrow{n \to \infty} P(A) \qquad \text{in } \mathcal{L}^p(P). \tag{39}$$

Proof. As an immediate consequence of the preceeding lemma we get for any simple function g on $(\Omega, \mathcal{F})$

$$\frac{1}{n} \sum_{i=0}^{n-1} g \circ \theta_{k_i(t)} \xrightarrow{n \to \infty} E[g] \qquad \text{in } \mathcal{L}^2(P),$$

as n goes to infinity. By a standard argument (for instance, Lemma 4 in [5]), this convergence holds as well in $\mathcal{L}^p(P)$, for $1 \leq p < \infty$. Finally, for any function f in $\mathcal{L}^p(P)$, decomposition into positive and negative parts, $\mathcal{L}^p(P)$-approximation by simple functions, and monotone convergence yields (39). $\qquad\square$

Having studied the first and the second parameter separately, we now go back to the product space. Recall that $\overline{P}$ is the product measure and $\kappa_n = \sum_{i=0}^{n} \kappa \circ \tau^i$.

Theorem 4.13. *Let $\tau : M \longrightarrow M$ be continuous and uniquely ergodic, and suppose that P is strongly mixing with respect to $(\theta_k)_{k \in \mathbb{Z}^d}$. Let $\kappa : M \longrightarrow \mathbb{Z}^d$ be $\mathcal{B}$-measurable such that,*

$$\lim_{n \to \infty} \frac{1}{n^2} \sup_{t \in M} \left| \left\{ 1 \leq i, j \leq n \;\middle|\; \|\kappa_i(t) - \kappa_j(t)\| \leq m \right\} \right| = 0 \tag{40}$$

for all $m \in \mathbb{N}$. Let be $1 \leq p < \infty$. Then for every $\mathcal{F}$-measurable function F on $\overline{\Omega}$ such that $\sup_{t \in M} F(t, \cdot)$ *is in $\mathcal{L}^p(\Omega, \mathcal{F}, P)$ and $F(\cdot, \omega)$ is continuous on M for P-almost every ω,*

$$\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \cdot\right) \xrightarrow{n \to \infty} \overline{E}[F] \qquad in \; \mathcal{L}^p(P), \tag{41}$$

uniformly in $t \in M$.

Proof. We first prove the theorem in the case when F is the indicator of a set of the form $U \times A$, where U is the intersection of finitely many metric balls in M or their complements, and $A \in \mathcal{F}$. By (6), the expression

$$\left\| \frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \cdot\right) - \overline{E}[F] \right\|_{\mathcal{L}^2(P)} \tag{42}$$

then transforms to

$$\frac{1}{n^2} \sum_{i,j=0}^{n-1} \int_{\Omega} 1_U\left(\tau^i(t)\right) 1_U\left(\tau^j(t)\right) 1_A\left(\theta_{\kappa_i(t)}\omega\right) 1_A\left(\theta_{\kappa_j(t)}\omega\right) P(d\omega)$$

$$- \frac{1}{n^2} \sum_{i,j=0}^{n-1} \mu(U)P(A) \left[\int_{\Omega} 1_U\left(\tau^i(t)\right) 1_A\left(\theta_{\kappa_i(t)}\omega\right) P(d\omega) \right. \tag{43}$$

$$\left. + \int_{\Omega} 1_U\left(\tau^j(t)\right) 1_A\left(\theta_{\kappa_j(t)}\omega\right) P(d\omega) \right]$$

$$+ \mu(U)^2 P(A)^2.$$

By $P\left(\theta_{\kappa_i(t)}^{-1} A \cap \theta_{\kappa_j(t)}^{-1} A\right) = P\left(\theta_{\kappa_i(t)-\kappa_j(t)}^{-1} A \cap A\right)$, the first addend equals

$$\frac{1}{n^2} \sum_{i,j=0}^{n-1} 1_U\left(\tau^i(t)\right) 1_U\left(\tau^j(t)\right) P\left(\theta_{\kappa_i(t)-\kappa_j(t)}^{-1} A \cap A\right).$$

It may be replaced by

$$\frac{1}{n^2} \sum_{i,j=0}^{n-1} 1_U\left(\tau^i(t)\right) 1_U\left(\tau^j(t)\right) P(A)^2, \tag{44}$$

without affecting the asymptotic behavior of (43) for n going to infinity: We may bound

$$\left| \frac{1}{n^2} \sum_{i,j=0}^{n-1} 1_U\left(\tau^i(t)\right) 1_U\left(\tau^j(t)\right) \left(P\left(\theta_{\kappa_i(t)-\kappa_j(t)}^{-1} A \cap A\right) - P(A)^2 \right) \right|$$

$$\leq \frac{1}{n^2} \sum_{i,j=0}^{n-1} \left| P\left(\theta_{\kappa_i(t)-\kappa_j(t)}^{-1} A \cap A\right) - P(A)^2 \right|. \tag{45}$$

Now, we argue as in the second part of the proof of Lemma 4.11, replacing the sequence $(k_i)_{i \in \mathbb{N}}$ by $(\kappa_i(t))_{i \in \mathbb{N}}$, and using assumption (40) instead of (37). This we prove that the difference created by the change (44) converges to 0 uniformly with respect to t.

Since the term in the rectangular brackets in the second addend in (43) equals $1_U\left(\tau^i(t)\right) P(A) + 1_U\left(\tau^j(t)\right) P(A)$, the whole expression simplifies to

$$\frac{1}{n^2} \sum_{i,j=0}^{n-1} \left(1_U\left(\tau^i(t)\right) 1_U\left(\tau^j(t)\right) - \mu(U)\left(1_U\left(\tau^i(t)\right) + 1_U\left(\tau^j(t)\right)\right) + \mu(U)^2 \right) P(A)^2,$$

which can be further reduced to

$$\left(\frac{1}{n} \sum_{i=0}^{n-1} 1_U \left(\tau^i(t) \right) - \mu(U) \right)^2 P(A)^2.$$

Since τ is uniquely ergodic and $\mu(\partial U) = 0$, Corollary 4.1.14 in [18] tells us that

$$\frac{1}{n} \sum_{i=0}^{n-1} 1_U \left(\tau^i(t) \right) \xrightarrow{n \to \infty} \mu(U) \qquad \text{uniformly in } t,$$

which concludes the first part of the proof. To pass from $\mathcal{L}^2$-convergence to general $\mathcal{L}^p$, use again a standard argument (for instance, Lemma 4 in [5]).

Now we let F be a general function, satisfying the conditions of the theorem. We need to find for every positive ϵ a sequence of metric balls $U_i \in M$ and $A_i \in \mathcal{F}$, with real numbers a_i such that, for all $t \in M$, $\| F(t, \cdot) - I(t, \cdot) \|_p < \varepsilon$, where

$$I = \sum_{i=1}^{n} a_i 1_{U_i \times A_i}. \tag{46}$$

It will then follow that

$$\left\| \frac{1}{n} \sum_{i=0}^{n-1} F \left(\tau_i(t), \theta_{\kappa_i(t)} \cdot \right) - \frac{1}{n} \sum_{i=0}^{n-1} I \left(\tau_i(t), \theta_{\kappa_i(t)} \cdot \right) \right\|_{\mathcal{L}^p(P)}$$

$$\leq \frac{1}{n} \sum_{i=0}^{n-1} \left\| F \left(\tau_i(t), \theta_{\kappa_i(t)} \cdot \right) - I \left(\tau_i(t), \theta_{\kappa_i(t)} \cdot \right) \right\|_{\mathcal{L}^p(P)}$$

$$= \frac{1}{n} \sum_{i=0}^{n-1} \left\| F \left(\tau_i(t), \cdot \right) - I \left(\tau_i(t), \cdot \right) \right\|_{\mathcal{L}^p(P)}$$

$$< \varepsilon.$$

For $\omega \in \Omega$ and $c > 0$, let $\delta(c, \omega)$ be the modulus of continuity for the function $F(\cdot, \omega)$. Define the sets

$$M_k = \sup_{t \in M} \left\{ \omega \, | \, |F(t, \omega)| \leq k \right\} \quad \text{and} \quad D_k(\varepsilon) = \left\{ \omega \, | \, \delta(1/k, \omega) \leq \varepsilon \right\} \quad (k \in \mathbb{N}).$$

Then the sequence of functions

$$F_k(\omega) := \sup_{t \in M} |F|^p (t, \omega) 1_{D_k(\varepsilon/4)^c \cup M_k^c} \qquad (k \in \mathbb{N}).$$

is bounded by $\sup_{t \in M} |F|^p (t, \omega)$, which is integrable, and converges to 0 for every ω. By the bounded convergence theorem, the integral of F_k converges to 0 as k goes to infinity. Choose a k such that $\int F_k P(d\omega) \leq (\varepsilon/2)^p$. Since M is compact, we may find a finite sequence $t_1, \ldots, t_r \in M$ such that the balls of radius $1/k$ around these centers cover M. We also define a sequence of real numbers $-k - 1 = s_0 < \cdots < s_{r'} = k$ such that the difference between any two successive elements is less than $\varepsilon/8$. Now we define a collection of sets $U_{i,j}$ and $A_{i,j}$ indexed by $r \times r'$. We start with $U_{i,j}$ as the ball of radius $1/k$ around t_i, and then remove the intersections, so that the $U_{i,j}$ is the same for all j, and running through $1 \leq i \leq r$ yields a disjoint cover of M. The sets $A_{i,j}$ are defined by

$$A_{i,j} = \left\{ \omega \, | \, s_{j-1} < F(t_i, \omega) \leq s_j \right\} \cap D_k(\varepsilon/8) \cap M_k.$$

Let $a_{i,j} = s_j$. We throw in one additional product set, $U_0 = M$ and $A_0 = D_k(\varepsilon/8)^c \cup M_k^c$ with $a_0 = 0$, and define the simple function $I(t, \omega)$ as indicated in (46). Then for any $t \in M$,

$$\left\| F(t, \cdot) - I(t, \cdot) \right\|_{\mathcal{L}^p(P)} = \left(\int \left| F(t, \omega) - I(t, \omega) \right|^p 1_{D_k(\varepsilon/8)} 1_{M_k} P(d\omega) \right)^{1/p}$$
$$+ \left(\int F_k(\omega) P(d\omega) \right)^{1/p}.$$

We already assumed (in defining k) that the second term is smaller than $\varepsilon/2$. For every t, there is a unique pair i, j such that $t \in U_{i,j}$ and $\omega \in A_{i,j}$. By construction, $I(t, \omega) = s_j$, so the integrand in the first term is bounded by

$$2^p \left| F(t_i, \omega) - F(t, \omega) \right|^p + 2^p \left| F(t_i, \omega) - s_j \right|^p.$$

This in turn is bounded by $2^p (\varepsilon/4)^p < (\varepsilon/2)^p$, since ω is not in A_0 and $d(t_i, t) < 1/k$, completing the proof. $\square$

Proposition 4.7 yields the following version for functions which are integrable in the sense of Riemann.

Corollary 4.14. *Let $a, b \in \mathbb{R}$, $a < b$, and $\tau : [a, b] \to [a, b]$ be continuous and uniquely ergodic with invariant measure μ, and assume that μ is absolutely continuous with respect to Lebesgue measure, with continuous density. For P and κ assume the same as in Theorem 4.13. Let $F \in L^p([a, b] \times \Omega, \mathcal{B} \otimes \mathcal{F}, \overline{P})$ be Riemann-integrable with respect to the first variable. Then we have*

$$\lim_{n \to \infty} \sup_{t \in M} \left\| \frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \cdot \right) - \overline{E}[F] \right\|_{\mathcal{L}^p(P)} = 0.$$

The last part of this section addresses the question of whether P-almost sure convergence of the ergodic averages of a function F on the product space $M \times \Omega$ may take place uniformly with respect to the first variable. In contrast to the previously discussed case of $\mathcal{L}^1(P)$-convergence, further conditions on F are needed. P-almost sure convergence of ergodic theorems is a very subtle question. Bellow and Losert's article [4] gives an overview of the results and open questions in this field. To get a first impression, choose a function F which is constant in ω. This brings us back to the beginning of this section, and Theorem 4.3 suggests that we would need an equicontinuity assumption in t. Further, we need an additional assumption on the measure.

Theorem 4.15. *Let μ be a τ-invariant measure on $(M, \mathcal{B})$, such that any non-empty open subset U of M has $\mu(U) > 0$. Let $\tau : M \to M$ be continuous and F a function on $M \times \Omega$, for which $F(t, \cdot) \in \mathcal{L}^1(P)$ for all $t \in M$, and the sequence of functions*

$$\left(\frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(\cdot), \theta_{\kappa_i(\cdot)} \omega \right) \right)_{n \in \mathbb{N}}$$

is equicontinuous on M, for all $\omega \in \Omega$. Then

$$\sup_{t \in M} \left| \frac{1}{n} \sum_{i=0}^{n-1} F\left(\tau^i(t), \theta_{\kappa_i(t)} \omega \right) - \overline{E}[F | \mathcal{J}](t, \omega) \right| \xrightarrow{n \to \infty} 0 \tag{47}$$

for P-almost all $\omega \in \Omega$.

Proof. We may assume without loss of generality that $E[F|\mathcal{J}] = 0$. The general case can be reduced to this by subtracting $E[F|\mathcal{J}]$ on both sides and making use of the invariance of $E[F|\mathcal{J}]$ under S.

The first step is to construct a countable dense set $M_1 \subset M$ and a set $N_1 \subset \Omega$ with $P(N_1) = 0$ such that

$$\frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(t, \omega) \longrightarrow 0 \qquad \text{for all } t \in M_1 \text{ and all } \omega \in \Omega \setminus N_1. \tag{48}$$

Since M is compact, the conditions on F assure that $F \in \mathcal{L}^1(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$, and therefore by (22) there is a set $\widetilde{M}_1 \subset M$ with $\mu(\widetilde{M}_1) = 1$ such that for any $t \in \widetilde{M}_1$ there is a set $N(t) \subset \Omega$ with $P(N(t)) = 0$ and

$$\frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(t, \omega) \longrightarrow 0 \qquad \text{for all } \omega \in \Omega \setminus N(t).$$

$\widetilde{M}_1$ is dense in M because its complement has measure zero with respect to μ and therefore, by assumption, contains no nonempty open subsets. Since M is separable we can find a countable dense subset $C \subset M$, and because $\widetilde{M}_1$ is dense in M, we can approximate any $x \in C$ by a sequence $(a_j(x))_{j \in \mathbb{N}}$ with $a_j(x) \in \widetilde{M}_1$ for all $j \in \mathbb{N}$. Then

$$M_1 := \bigcup_{x \in C} \bigcup_{j \in \mathbb{N}} a_j(x) \qquad \text{and} \qquad N_1 := \bigcup_{t \in M_1} N(t)$$

define a countable dense subset M_1 of M and a subset N_1 of Ω, which fulfills (48). This accomplishes the first step.

For the next step, chose $s \in M$ and fix $\varepsilon > 0$. By equicontinuity, there is a set $N_0 \subset \Omega$ with $P(N_0) = 0$ and a $\delta > 0$ such that for all $r, t \in M$ with $d(r, t) < \delta$

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(r, \omega) - F \circ S^i(t, \omega) \right| < \frac{\varepsilon}{2} \quad \text{for all } n \in \mathbb{N} \text{ and all } \omega \in \Omega \setminus N_0. \tag{49}$$

Define $N := N_0 \cup N_1$ and fix $\omega \in \Omega \setminus N$. Since M_1 is dense in M we can find a $t \in M_1$ with $d(s, t) < \delta$, and by (48) there is an $n_1 \in \mathbb{N}$ such that

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(s, \omega) \right| < \frac{\varepsilon}{2} \qquad \text{for all } n \geq n_1.$$

Combining the last two inequalities leads to

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(t, \omega) \right| < \varepsilon \qquad \text{for all } n \geq n_1 \text{ and all } \omega \in \Omega \setminus N,$$

and letting ε go to 0 yields

$$\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(s, \omega) = 0 \qquad \text{for } P\text{-almost all } \omega \in \Omega. \tag{50}$$

To obtain uniform convergence with respect to the first variable, we use a standard compactness argument. Since M is compact, it can be covered by a finite number m of δ-neighborhoods in M,

which centers are denoted by $s_1, ..., s_m$. Applying the reasoning of the last step to each of the $s_1, ..., s_m$ we can find $n_0 \in \mathbb{N}$ such that

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(s_k, \omega) \right| < \frac{\varepsilon}{2} \qquad \text{for all } n \geq n_0, k \in \{1, ..., K\}, \text{ and } \omega \in \Omega \setminus N.$$

For an arbitrary $s \in M$ there exist $k \in \{1, ..., K\}$ such that $d(s, s_k) < \delta$, and by (49) we obtain

$$\frac{1}{n} \left| \sum_{i=0}^{n-1} F \circ S^i(s, \omega) - F \circ S^i(s_k, \omega) \right| < \frac{\varepsilon}{2} \text{ for all } n \in \mathbb{N} \quad \text{and all } \omega \in \Omega \setminus N.$$

Finally, the desired convergence (47) follows by the last two inequalities, and by letting ε go to 0. $\qquad \square$

We conclude the paper with the results to the example mentioned in the introduction.

Example 4.16. *(Ergodic averages of a random field along the lattice approximation of a line)* Let $L_{\lambda, t}(z) = (z, [\lambda z + t])$ $(z \in \mathbb{Z})$ be the approximation of the line with slope λ and y-intercept t by the lattice $\mathbb{Z}^2$. In (2) we were wondering about convergence of the ergodic averages of a function of a $\mathbb{Z}^2$-indexed random field along this lattice approximation. The problem can be treated by the theory developed in this paper.

Consider the situation of Corollary 2.6, and let τ_λ be the translation on the torus defined in Example 2.2. Set $\kappa(t) := (1, [t + \lambda])$. The following lemma shows that the corresponding skew product captures the lattice approximation of the line.

Lemma 4.17. *The iterates of the transformation* $S_\lambda(a, \omega) = (\tau_\lambda(a), \theta_{(1,[a+\lambda])}\omega)$ $(a \in M, \omega \in \Omega)$ *are given by* $S_\lambda^n(a, \omega) = (\{a + n\lambda\}, \theta_{L_a(n)}\omega)$ *for all* $n \in \mathbb{N}_0$.

Proof. We have $S_\lambda^n(a, \omega) = (\tau_\lambda^n(a), \theta_{\kappa_n(a)}\omega)$, where $\kappa_n = \sum_{i=0}^{n-1} \kappa \circ \tau_\lambda^i$. It is easy to see that $\tau_\lambda^n(a) = \{a + n\lambda\}$. It remains to show that $\kappa_n(a) = (n, L_{\lambda, a}(n))$ for all $a \in \mathbb{T}$. For the first component this is obvious. For the second component it follows by induction: It is trivial for $n = 0$. For the step from n to $n + 1$ we obtain $\kappa_{n+1}^{(2)}(a) = \kappa_n^{(2)}(a) + \kappa^{(2)}(\tau_\lambda^n(a)) = L_{\lambda, a}(n) + [\tau_\lambda^n(a) + \lambda]$, and the claim follows from $[\tau_\lambda^n(a) + \lambda] = [a + n\lambda - [a + n\lambda] + \lambda] = -[a + n\lambda] + [a + (n+1)\lambda] = -L_{\lambda, a}(n) + L_{\lambda, a}(n + 1)$. $\quad \square$

Let $f \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$. We want to apply the ergodic theorems on the product space to the function $F(t, \omega) := f(\omega)$ $(\omega \in \Omega)$, for some $t \in M$.

Let λ be irrational. Then τ_λ is ergodic. Since $\|\kappa_n(t)\| \geq n$, the sequence tends to infinity as n goes to infinity. As a special case of the strong 0-1 law, P fulfills a 0-1 law on the tail field. Corollary 2.6 with $v_1 = (1, 0)$ and $v_2 = (0, 1)$ implies the ergodicity of the skew product S_λ. It remains to check condition (40). We have $\|\kappa_i(t) - \kappa_j(t)\| \geq \|i - j\|$, and

$$\lim_{n \to \infty} \frac{1}{n^2} \left| \left\{ 1 \leq i, j \leq n \,\middle|\, \|i - j\| \leq m \right\} \right| = 0$$

for all $m \in \mathbb{N}$. This implies condition (40), and by Theorem 4.13 we have the following result:

$$\frac{1}{n} \sum_{i=0}^{n-1} f(\omega_{L_{\lambda, t}}) \xrightarrow{n \to \infty} \int_0^1 E[f] \, dt$$

in $\mathcal{L}^1(P)$, uniformly in all $t \in M$. Note that the limit does not depend on t, the y-intersect of the approximated line.

In the case when λ is rational, τ_λ is periodic, and a corresponding result follows immediately from Theorem 3.5 (cf. Example 3.6).

Acknowledgements

References

[1] R. Adler and P. Shields. Skew products of Bernoulli shifts with rotations. *Isr. J. of Math.*, 12:215–222, 1972.

[2] R. Adler and P. Shields. Skew products of Bernoulli shifts with rotations II. *Isr. J. of Math.*, 19:228–236, 1974.

[3] H. Anzai. Ergodic skew product transformations on the torus. *Osaka Math. J.*, 3(1), March 1951.

[4] A. Bellow and V. Losert. The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences. *Trans. Am. Math. Soc.*, 288:307–345, 1985.

[5] J.R. Blum and D.L. Hanson. On the mean ergodic theorem for subsequences. *Bull. Am. Math. Soc.*, 66:308–311, 1960.

[6] J. Brettschneider. *Shannon-MacMillan theorems for random fields along curves and lower bounds for surface-order large deviations.* PhD thesis, Humboldt University Berlin, 2000.

[7] J. Brettschneider. Shannon-MacMillan theorems for random fields along curves and lower bounds for surface-order large deviations. EURANDOM *Technical Report*, 2001.

[8] N. G. de Bruijn and K. A. Post. A remark on uniformly distributed sequences and Riemann integrability. *Indag. Math.*, 30:149–50, 1968.

[9] F. den Hollander. Mixing properties for random walk in random scenery. *Ann. Probab.*, 16:1788–1802, 1988.

[10] F. den Hollander and M. Keane. Ergodic properties of color records. *Physica*, 138A:183–193, 1986.

[11] N. Friedman. Mixing on sequences. *Can. J. Math.*, 35:339–352, 1983.

[12] H. Furstenberg. Strict ergodicity and transformation of the torus. *Amer. J. Math.*, 83:573–601, 1961.

[13] H.-O. Georgii. *Gibbs measures and phase transitions.* W. de Gruyter, Berlin, 1988.

[14] H.-O. Georgii. Mixing properties of induced random transformations. *Ergod. Th. and Dynam. Systems*, 17:839–847, 1997.

[15] P. Hellekalek and G. Larcher. On the ergodicity of a class of skew products. *Israel J. Math.*, 54(3):301–306, 1986.

[16] P. Hellekalek and G. Larcher. On Weyl sums and skew products over irrational rotations. *Theoret. Comput. Sci.*, 65(2):189–196, 1989.

[17] S. Kakutani. Random ergodic theorems and Markov processes with a stable distribution. In J. Neyman, editor, *Proceedings of the Second Berkeley Symposium on Statistics and Probability*, pages 247–261, Berkeley, 1951. University of California Press.

[18] A. Katok and B. Hasselblatt. *Introduction to the modern theory of dynamical systems.* Cambridge Univ. Press, Cambridge, 1995.

[19] U. Krengel. *Ergodic theorems.* W. de Gruyter, Berlin, 1985.

[20] I. Meilijson. Mixing properties of a class of skew-products. *Isr. J. of Math.*, 19:266–270, 1974.

[21] J. Milnor. On the entropy geometry of cellular automata. *Complex Syst.*, 2(3):357–385, 1988.

[22] Ishai Oren. Ergodicity of cylinder flows arising from irregularities of distribution. *Israel J. Math.*, 44(2):127–138, 1983.

[23] D. A. Pask. Skew products over the irrational rotation. *Israel J. Math.*, 69(1):65–74, 1990.

[24] P. Walters. *An introduction to ergodic theory.* Springer-Verlag, New York, Heidelberg, Berlin, 1982.

[25] Q. Zhang. On skew products of irrational rotations with tori. In *Convergence in ergodic theory and probability*, volume 5, pages 435–445, 1996.