ON UNIFORM CONVERGENCE IN ERGODIC THEOREMS FOR A CLASS OF SKEW PRODUCT TRANSFORMATIONS

JULIA BRETTSCHEIDER
Department of Statistics
University of Warwick
Coventry, CV4 7AL, UK
and
Department of Community Health & Epidemiology
Cancer Research Institute Division of Cancer Care & Epidemiology
Queen’s University
Ontario, K7L 3N6, Canada

(Communicated by Yuri Kifer)

Abstract. Consider a class of skew product transformations consisting of an ergodic or a periodic transformation on a probability space \((M, \mathcal{B}, \mu)\) in the base and a semigroup of transformations on another probability space \((\Omega, \mathcal{F}, P)\) in the fibre. Under suitable mixing conditions for the fibre transformation, we show that the properties ergodicity, weakly mixing, and strongly mixing are passed on from the base transformation to the skew product (with respect to the product measure). We derive ergodic theorems with respect to the skew product on the product space.

The main aim of this paper is to establish uniform convergence with respect to the base variable for the series of ergodic averages of a function \(F\) on \(M \times \Omega\) along the orbits of such a skew product. Assuming a certain growth condition for the coupling function, a strong mixing condition on the fibre transformation, and continuity and integrability conditions for \(F\), we prove uniform convergence in the base and \(L^p(P)\)-convergence in the fibre. Under an equicontinuity assumption on \(F\) we further show \(P\)-almost sure convergence in the fibre. Our work has an application in information theory: It implies convergence of the averages of functions on random fields restricted to parts of stair climbing patterns defined by a direction.

1. Introduction. The approximation of a line by a planar lattice yields a stair climbing pattern. Consider the averages of a function of a random field along a finite window moving up the stair climbing pattern. Under which conditions does this sequence converge, and what are explicit formulae for the limit?

More formally, let \(L_{\lambda,t}(z) := (z, [\lambda z + t])\) \((z \in \mathbb{Z})\) be a lattice approximation of the line with slope \(\lambda\) and \(y\)-intercept \(t\). Let \(m \in \mathbb{N}\) be a fixed window size and \(L_{\lambda,t}(z, \ldots, z + m - 1) := (L_{\lambda,t}(z), \ldots, L_{\lambda,t}(z + m - 1))\) the vector of lattice point forming the stair climbing pattern visible in the window between the \(z\)th and the \((z + m - 1)\)th step. Let \(P\) be a \(\mathbb{Z}^2\)-indexed random field with values in a set \(\Upsilon\), i.e., a stationary probability measure on \(\Omega := \Upsilon^{\mathbb{Z}^2}\), and let \(\mathcal{F}\) be the

2000 Mathematics Subject Classification. Primary: 37A30; Secondary: 37A50.

Key words and phrases. Skew product transformation, ergodicity, mixing properties, ergodic theorems, uniform convergence.
canonical \(\sigma\)-algebra. For a realization \(\omega \in \Omega\) let \(\omega(L_{\lambda,t}(z,\ldots,z+m-1)) := (\omega(L_{\lambda,t}(z),\ldots,\omega(L_{\lambda,t}(z+m-1))))\) be the averages of the random field along the stair climbing pattern in the window. Consider the averages
\[
\frac{1}{n} \sum_{i=0}^{n-1} f(\omega(L_{\lambda,t}(i,\ldots,i+m-1))) \quad (n \in \mathbb{N})
\]
of a function \(f\) on \(\mathbb{T}^m\). What can we say about \(P\)-almost sure or \(L^1(P)\)-convergence of this sequence? Averages of this type are similar to the ones that occur in the context of directional Shannon-MacMillan theorems for lattice random fields (cf. [7]).

The situation described above can be represented as a special case of a more general set up involving skew product transformations. In the independent component, or base, we have a measure-preserving transformation \(\tau\) on a probability space \((M,B,\mu)\). In the dependent component, or fibre, we have a mixing semigroup of measure-preserving transformations \((\theta_k)_{k \in K}\) on a probability space \((\Omega,F,P)\), which is linked to the base by a \(K\)-valued \(B\)-measurable function \(\kappa\) on \(M\). The class of skew products considered in this paper is given by
\[
S(t,\omega) = (\tau(t),\theta_{\kappa(t)}\omega) \quad (t \in M, \omega \in \Omega).
\]

This paper deals with the following two questions:

(A) Is \(S\) ergodic or mixing with respect to the product measure?

(B) Do the ergodic averages along the skew product converge uniformly with respect to the base?

Under suitable assumptions, we will answer both questions positively.

Ergodicity, and other mixing properties of various classes of skew products have been studied by a number of authors, but the above situation does not fit into any of the settings covered by existing literature. Kakutani [20] introduced a skew product with a Bernoulli-shift in the base and an ergodic transformation in the fibre. 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 [25], den Hollander and Keane [11], and Georgii [15]. Adler and Shields (cf. [1] and [2]) considered a circle rotation for the fibre. Anzai [3] introduced skew products of two rotations on the circle, and derived a criterion for ergodicity. Furstenberg [13] studied unique ergodicity. Zhang [32] investigated this for a translation on a torus in the fibre. A torus translation 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 [27], Hellekalek and Larcher [17], [18], and Pask [28]. Conditions for ergodic properties of skew products were also studied by [23] in the context of group Rokhlin cocycle extensions.

Our answer to question (A) is summarized in Theorem 2.2. We prove that, under suitable mixing conditions for the transformations in the fibre, the properties ergodic, weakly mixing, and 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, the conditions on the fibre transformation can be insured by assuming tail-triviality for \(P\) and a growth condition for ergodic sums of \(\kappa\) along \(\tau\) (cf. Corollary 1).

Our answer to question (B) is given in Theorem 4.5. The proof combines two approaches. The first approach explores uniform convergence theorems in the spirit
of Weyl’s classical result about ergodic averages of a continuous function along the orbit of circle rotations. Oxtoby extended this result to uniquely ergodic transformations of compact metric spaces. Weyl actually proved his classical result for the class of Riemann-integrable functions and we will do the same in the context of Oxtoby’s theorem (cf. Corollary 5). The second approach are techniques developed for ergodic theorems along subsequences. We extend Blum and Hanson’s theorem to the $d$-parameter case replacing the strict monotonicity condition on the sequence by a growth condition on the coupling function $\kappa$, uniformly in $t$. Combining the two approaches we obtain uniform convergence in the base and $L^p$-convergence in the fibre for functions that are continuous with respect to the base and fulfill an integrability condition in the fibre (cf. Theorem 4.5). We further derive a result (cf. Theorem 4.7) about uniform convergence in the base and $P$-almost sure convergence in the fibre, provided the iterates of the function fulfill an equicontinuity condition.

We conclude the paper by returning to our initial questions about the asymptotics of $(1)$. We identify the unit circle $\mathbb{T}$ with the interval $[0, 1)$ equipped it with the Borel $\sigma$-algebra $\mathcal{B}$ and the Lebesgue measure $\mu$. For $\lambda \in \mathbb{R}$ define a rotation of $\mathbb{T}$ by $\tau_\lambda(t) := t + \lambda \mod 1$. For $x \in \mathbb{R}$ let $[x]$ be the integer part of $x$. Define a skew product on the product space $\mathbb{T} \times \Omega$ by

$$S(t, \omega) := (\tau_\lambda(t), \theta_{([1,\lambda+t])\omega}) \quad (t \in \mathbb{T}, \omega \in \Omega).$$

We will see that the iterates of $S$ follow the stair climbing pattern $L_{\lambda,t}$. This allows to rewrite the sequence $(1)$ as an average of $f$ along the orbit of $S$. The convergence of this sequence, for all starting levels $t$ of the stair climbing pattern, will be obtained as a consequence of the uniform ergodic theorems derived in in Section 3 (cf. Corollary 7).

Outline of the paper: In the first section we define mixing properties of semigroups of transformations along sequences and we introduce a class of skew products considered in this paper. In Theorem 2.2, we give the result on ergodicity and mixing properties of these skew products. Finally, we have a closer look at the case when the fibre transformation is a shift operator for a random field. In Section 3 we discuss ergodic theorems for skew products with ergodic base transformation (cf. Corollary 2) and with periodic base transformations (cf. Corollary 3). We illustrate the results with two examples related to sequence $(1)$. In the last section we focus on the main aim of this paper, the uniform convergence with respect to the base. Depending, among other things, on the regularity of the function with respect to the base variable, we obtain different kinds of convergence in the fibre. Theorem 4.5 states $L^p(P)$-convergence provided the function is continuous with respect to the base variable. Corollary 6 is a version of this for Riemann-integrable functions. Theorem 4.7 states $P$-almost sure convergence, provided the functions fulfill a certain equicontinuity condition. Corollary 7 brings us back to the original motivation for this paper, a statement about convergence of sequence $(1)$.

2. Mixing properties of a class of skew products. Let $(\Omega, \mathcal{F}, P)$ be a probability space, and let $(\theta_k)_{k \in \mathbb{N}_0^d}$ be a $d$-parameter semigroup of measure-preserving transformations on $(\Omega, \mathcal{F}, P)$, i.e., each of the transformations preserves the measure $P$, and $\theta_0 = \text{Id}$, and $\theta_k \circ \theta_l = \theta_{k+l}$ for all $k, l \in \mathbb{N}_0^d$. $(\theta_k)_{k \in \mathbb{Z}_0^d}$ is a $d$-parameter
group if \((\theta_k)_{k\in\mathbb{N}^d}\) is a semigroup and \(\theta_{-k} = \theta_k^{-1}\) for all \(k \in \mathbb{N}^d\). The following example will be used frequently in our settings. Let \(\sigma_1\) and \(\sigma_2\) be two commuting measure-preserving transformations on \((\Omega, \mathcal{F}, P)\). Then
\[
\theta_k := \sigma_1^{k(1)} \circ \sigma_2^{k(2)} \quad \text{for} \quad k = (k(1), k(2)) \in \mathbb{N}_0^d
\]
defines a two-parameter semigroup \((\theta_k)_{k\in\mathbb{N}^d}\) of measure-preserving transformations on \((\Omega, \mathcal{F}, P)\). If \(\sigma_1\) and \(\sigma_2\) are invertible it extends to a two-parameter group \((\theta_k)_{k\in\mathbb{Z}^2}\). The constructions extends to \(d\)-parameters in an obvious way.

Let \(K = \mathbb{N}_0^d\) or \(K = \mathbb{Z}^d\). Let \(\tau\) be a measure-preserving transformation of a probability space \((M, \mathcal{B}, \mu)\), and assume that \(\kappa\) is a \(\mathcal{B}\)-measurable function on \(M\) with values in \(K\). Then
\[
S(t, \omega) = (\tau(t), \theta_{\kappa(t)}(\omega)) \quad (t \in M, \omega \in \Omega)
\]
defines a skew product on the product space \(\Omega := M \times \Omega\). In particular, choosing \(\kappa \equiv k_0\) for a constant \(k_0 \in K\) yields the uncoupled product of \(\tau\) and \(\theta_{k_0}\). Obviously, \(S\) is measurable with respect to the product \(\sigma\)-algebra \(\sigma\)-algebra \(\mathcal{F} := \mathcal{B} \otimes \mathcal{F}\), and it preserves the product measure \(P := \mu \otimes P\). It is easy to see that for all \(n, m \in \mathbb{N}_0\), for all \(t \in M\), and for all \(\omega \in \Omega\),
\[
S^n(t, \omega) = (\tau^n(t), \theta_{\kappa_n(t)}(\omega)), \quad \kappa_n = \sum_{i=0}^{n-1} \kappa \circ \tau^i \quad \text{and} \quad \kappa_{n+m} = \kappa_n + \kappa_m \circ \tau^n.
\]

Let us now ask under which conditions the skew product is ergodic. Following a suggestion by J. Aaronson, we actually broaden the question to other mixing properties. Before answering these questions in the next lemma we need to develop some conditions that allow us to derive mixing properties of the skew product.

The ergodicity of \(\tau\) is necessary for the ergodicity of \(S\). This can be shown by a simple projection argument. Would ergodicity of both factors actually imply the ergodicity of the skew product? For uncoupled products the answer is negative, but if one of the factors is weakly mixing, the product is ergodic (cf. [22]). This motivates the introduction of a mixing condition tailored to the dependent factor in the skew products. In other words, we are looking for a mixing condition on \(\theta\) that invokes the function \(\kappa\). N. Friedman (cf. [12]) introduced the concept of weakly mixing along a sequence for a transformation. We carry this over to an analogue notion for \(d\)-parameter (semi-)groups. Note that for the rest of this section properties the transformation \(\tau\) (e.g. measure preserving, ergodic or mixing) are always with respect to the measure \(\mu\), properties of \(\theta\) are always with respect to \(P\) and properties of \(S\) are always with respect to \(P\).

**Definition 2.1.** Let \((k_n)_{n\in\mathbb{N}}\) be a \(K\)-valued sequence. A (semi-)group \((\theta_k)_{k\in K}\) of measure-preserving transformations on \((\Omega, \mathcal{F}, P)\), is called weakly mixing along \((k_n)_{n\in\mathbb{N}}\), if
\[
\frac{1}{n} \sum_{i=0}^{n-1} |P(A \cap \theta_{-k_i}^{-1} B) - P(A) P(B)| \xrightarrow{n \to \infty} 0 \quad \text{for all} \ A, B \in \mathcal{F}.
\]

We will make use of two conditions:

1. **(C1)** \((\theta_k)_{k\in K}\) is weakly mixing along \((\kappa_n(t))_{n\in\mathbb{N}}\) for \(\mu\)-almost all \(t \in M\).
2. **(C2)** \((\theta_k)_{k\in K}\) is strongly mixing and \((\kappa_n(t))_{n\in\mathbb{N}}\) goes to infinity for \(\mu\)-almost all \(t \in M\).
Note that (C2) implies (C1). Condition (C2) can be easily verified for lattice approximations of a line, as discussed in the context of Corollary 7.

**Lemma 2.2.** (i) Assume condition (C1). If $\tau$ is ergodic then $S$ is ergodic.

(ii) Assume condition (C1). If $\tau$ is weakly mixing then $S$ is weakly mixing.

(iii) Assume condition (C2). If $\tau$ is strongly mixing then $S$ is strongly mixing.

**Proof.** We give the proof of the first statement here; the remaining proofs are conducted in a similar fashion. Assume condition (C1) and let $\tau$ denote the $\omega$-almost all $\mu$-τ.

It is sufficient to show this for functions which are products of functions on the $F$ functions $\mu$. To prove the ergodicity of $\mathcal{F}$ on $\omega$ we give the proof of the first statement here; the remaining proofs are conducted in a similar manner. Assume condition (C1) and let $\tau$ be an ergodic random field with values in a set $\Sigma$. For any $J \subseteq \mathbb{Z}^d$ let $\mathcal{F}_J$ denote the $\sigma$-algebra generated by all projections $\omega \mapsto \omega(j)$ with $j \in J$, and let $\mathcal{F} := \mathcal{F}_{2d}$. Consider the shift transformations $(\nu_v)_{v \in \mathbb{Z}^d}$ on $\Omega$, i.e., $\nu_v(\omega)(j) := \omega(j + v)$ ($j \in \mathbb{Z}^d$). Let $P$ be a random field, i.e., a measure on $(\Omega, \mathcal{F})$ which is invariant with respect to $\nu_v$ for all $v \in \mathbb{Z}^d$. The tail field is the $\sigma$-algebra $\mathcal{T} := \bigcap_{V \subseteq \mathbb{Z}^d} \text{finite} \mathcal{F}_V \setminus V$, and $P$ is called tail-trivial if it fulfills a 0-1 law on $\mathcal{T}$. Let $v_1, v_2 \in \mathbb{Z}^d$. As in (4),
\( \theta_k := \vartheta_{v_1}^{(1)} \circ \vartheta_{v_2}^{(2)} \ (k \in \mathbb{Z}^d) \) defines a 2-parameter group of measure-preserving transformations. We have
\[
\theta_{\kappa_n(t)} = \vartheta_{v_1}^{(1)}(t) \circ \vartheta_{v_2}^{(2)}(t) = \theta_{\kappa_n(t)} v_1 + \kappa_n(t) v_2.
\]
Denote the coordinates of elements in \( \mathbb{Z}^d \) by upper indices, and let \( \| \cdot \| \) be the maximum norm. In this situation we have the following

**Corollary 1.** Let \( v_1 \) and \( v_2 \) be linear independent vectors in \( \mathbb{Z}^d \). Assume that \( P \) is tail-trivial and that the sequence \( (\|\kappa_n(t)\|)_{n \in \mathbb{N}} \) goes to infinity for \( \mu \)-almost all \( t \in M \). If \( \tau \) is ergodic, weakly mixing or strongly mixing with respect to \( \mu \), then \( S \) is ergodic, weakly mixing or strongly mixing with respect to \( \overline{P} \), respectively.

**Proof.** We are going to show condition (C2). Define boxes \( V_n = \{ v \in \mathbb{Z}^d \mid \|v\| \leq n \} \) \((n \in \mathbb{N})\), 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 \( J - m(n) \) are contained in \( V_m^{\kappa_n(t)} \), where \( m(n) = (m(n) - 2m) \lor 0 \). For any \( A \in \mathcal{F}, \) we obtain
\[
|P(A \cap \theta_{\kappa_n(t)}^{-1} B) - P(A) P(B)| = |P(A \cap \vartheta_{v_1}^{(1)}(t) \vartheta_{v_2}^{(2)}(t) B) - P(A) P(B)| \leq \sup_{C \in \mathcal{F}_d \setminus A} |P(A \cap C) - P(A) P(C)|.
\]
By the assumptions on \( v_1, v_2 \) and \( \kappa, \|m(n)\| \) goes to infinity. By Proposition 7.9 in [14], tail-triviality is equivalent to short-range correlations, which means that \( \sup_{C \in \mathcal{F}_d \setminus A} |P(A \cap C) - P(A) P(C)| \) converges to 0 as \( n \) goes to infinity. \( \square \)

3. **Ergodic theorems with skew products.** Applying Birkhoff’s ergodic theorem to the skew product \( S \) yields, for any \( F \in \mathcal{L}^1(\Omega, \mathcal{F}, \overline{P}) \),
\[
\frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_n(t)} \omega) \xrightarrow{n \to \infty} \mathbb{E}[F] \quad \mathcal{P}\text{-a.s. and in } \mathcal{L}^1(\overline{P}),
\]
where \( \mathcal{J} \) is the \( \sigma \)-algebra of all \( S \)-invariant sets in \( \mathcal{F} \). We study this limit more closely for two different cases: when the transformation \( \tau \) is ergodic and when it is periodic. In the ergodic case, combining (9) and Lemma 2.2 immediately yields the following ergodic theorem for the skew product.

**Corollary 2.** Assume that \( \tau \) is ergodic with respect to \( \mu \) and that the condition (C1) is fulfilled. Then for any function \( F \in \mathcal{L}^1(\Omega, \mathcal{F}, \overline{P}) \),
\[
\frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_n(t)} \omega) \xrightarrow{n \to \infty} \mathbb{E}[F] \quad \overline{P}\text{-almost surely and in } \mathcal{L}^1(\overline{P}).
\]

Now consider the case that \( \tau \) periodic. We calculate the iterates of the skew product and derive an ergodic theorem with an explicit expression for the limit.

**Lemma 3.1.** Assume that \( \tau \) is periodic with \( q \in \mathbb{N} \). Then for all \( j \in \mathbb{Z} \) and all \( \nu \in \{0, 1, \ldots, q - 1\} \),
(i) \( \kappa_{jq + \nu} = j \kappa_q + \kappa_\nu \),
(ii) \( \theta_{\kappa_{jq + \nu}(t)} = (\theta_{\kappa_q(t)})^j \circ \theta_{\kappa_\nu(t)} \) for all \( t \in M \),
(iii) \( S_{jq + \nu}(t, \omega) = (\tau^\nu(t), \theta_{\kappa_\nu(t)} \circ (\theta_{\kappa_q(t)})^j \omega) \) for all \( t \in M \) and all \( \omega \in \Omega \).
Proof. (i) follows from the definition of $\kappa$ using the periodicity of $\tau$. (ii) is an immediate consequence of (i) and the semigroup property of $\theta$. (iii) follows from (5), the periodicity of $\tau$ and because of (ii).

**Corollary 3.** Assume $\tau$ is periodic with period $q \in \mathbb{N}$. Denote by $\mathcal{J}_t$ the $\sigma$-algebra of $\theta_{\kappa_s(t)}$-invariant sets in $\mathcal{F}$. Then, for any $F \in \mathcal{L}^1(\Omega, \mathcal{F}, \nu)$,
\[
\frac{1}{n} \sum_{i=0}^{n-1} F^t(t, \theta_{\kappa_s(t)} \omega) \underset{n \to \infty}{\longrightarrow} \frac{1}{q} \sum_{\nu=0}^{q-1} E[F^t(t, \theta_{\kappa_s(t)}(\cdot))]_{\mathcal{J}_t},
\]
for $\mu$-almost all $t \in M$, and for $P$-almost all $\omega \in \Omega$ and in $\mathcal{L}^1(P)$. If $\theta_{\kappa_s(t)}$ is ergodic with respect to $P$ for $\mu$-almost all $t \in M$, then the limit simplifies to the constant
\[
\frac{1}{q} \sum_{\nu=0}^{q-1} E[F^t(t, \cdot)]_{\mathcal{J}_t}.
\]

Proof. 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 the ergodic averages to
\[
A_n F := \frac{1}{n} \sum_{i=0}^{n-1} F^t \circ S^i = \frac{mq}{mq + \nu} \left( \frac{1}{mq} \sum_{i=0}^{mq-1} F \circ S^i + \frac{1}{mq} \sum_{i=mq}^{mq+\nu-1} F \circ S^i \right).
\]
Since the first factor converges to 1, and the second addend within the brackets converges to 0, our question reduces to the study of ergodic limits along the subsequence $(mq)_{m\in\mathbb{N}}$. They take the form $A_{mq} F = 1/q \sum_{\nu=0}^{q-1} A_{n}^{(\nu)} F$, where
\[
A_{mq}^{(\nu)} F(t, \omega) := \frac{1}{m} \sum_{j=0}^{m-1} F \circ S^{jq+\nu}(t, \omega) = \frac{1}{m} \sum_{j=0}^{m-1} F^t(t, \theta_{\kappa_s(t)}(\cdot))_{\mathcal{J}_t}(\cdot).
\]
The last equality can be seen by applying Lemma 3.1. For $\mu$-almost all $t \in M$, the function $f_{mq}^{(\nu)}(\omega) := F^t(t, \theta_{\kappa_s(t)}(\omega))$ ($\omega \in \Omega$) is integrable, and applying Birkhoff’s ergodic theorem yields
\[
\lim_{m \to \infty} A_{mq}^{(\nu)} F(t, \cdot) = E[f_{mq}^{(\nu)} | \mathcal{J}_t] = E[F^t(t, \theta_{\kappa_s(t)}(\cdot))]_{\mathcal{J}_t} \text{ P.a.s. and in } \mathcal{L}^1(P).
\]
This implies the first statement of the corollary. The second statement is about the ergodic case. Then $\mathcal{J}_t$ is trivial and, using the invariance of $P$ under $\theta$, the addends in the last expression reduce to $E[F^t(t, \cdot)]_{\mathcal{J}_t}$. 

We end this section with an illustration of our results by two special cases relevant to the skew product tracing a stair climbing pattern introduced in (3).

**Example 1.** Circle rotations as base transformations. Identify the unit circle with $\mathbb{S}$ and equip it with the skew product tracing a stair climbing pattern introduced in (3).

(i) Let $\lambda$ be irrational. Choose $(\theta_k)_{k \in \mathbb{F}}$ and $\kappa$ fulfilling condition (C1). By Lemma 2.2, $S$ is ergodic. By Corollary 2, for any integrable function $F$ on $(\mathbb{T} \times \Omega, \mathcal{B} \otimes \mathcal{F}, \mu \otimes P),$
\[
\frac{1}{n} \sum_{i=0}^{n-1} F(t + i\lambda \mod 1, \theta_{\lambda_i(t)} \omega) \underset{n \to \infty}{\longrightarrow} \int_0^1 E[F(t, \cdot)] dt,
\]
for $\mu \otimes P$-almost all $(t, \omega) \in \mathbb{T} \times \Omega$ and in $L^1(\mathbb{T} \times \Omega, B \otimes F, \mu \otimes P)$.

(ii) Let $\lambda$ be rational. There is a unique representation $\lambda = p/q$, where $p \in \mathbb{Z}$, $q \in \mathbb{N}$, $p$ and $q$ have no common divisor. $\tau_\lambda$ is periodic with period $q$. Furthermore, $\tau_\lambda$ respects the partition $\{0, \frac{1}{q}, \frac{2}{q}, \ldots, \frac{q-1}{q}, 1\}$ of $\mathbb{T}$, i.e., for every $\nu \in \{1, \ldots, q-1\}$ there is a $\tilde{\nu} \in \{1, \ldots, q-1\}$ such that $\tau_\lambda([\frac{\nu-1}{q}, \frac{\nu}{q}]) = [\frac{\tilde{\nu}-1}{q}, \frac{\tilde{\nu}}{q}]$. The limit in Corollary 3 is of the form $1/q \sum_{\nu=0}^{q-1} E[F(t + \nu \lambda \mod 1, \cdot)|J_t]$. If $P$ is ergodic with respect to $\theta_{\kappa_i(t)}$ then limit simplifies to $1/q \sum_{\nu=0}^{q-1} E[F(t + \nu \lambda \mod 1, \cdot)]$.

Example 2. Shifts as fibre transformations. Consider the 2-parameter group defined above (8) and the sequence

$$\frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_1^{(1)}(t) \kappa_2^{(2)}(t) \omega} \omega) \quad (n \in \mathbb{N}). \quad (10)$$

(i) If $\tau$ is ergodic then, for any function $F \in L^1(\Omega, F, \mathcal{P})$, the sequence in (10) converges to $E[F|\mathcal{P}]$, $\mathcal{P}$-almost surely and in $L^1(\mathcal{P})$.

(ii) If $\tau$ is periodic with period $q$ then, for any function $F \in L^1(\Omega, F, \mathcal{P})$, the sequence in (10) converges to

$$\frac{1}{q} \sum_{\nu=0}^{q-1} E[F(\tau^\nu(t), \theta_{\kappa_1^{(1)}(t) \kappa_2^{(2)}(t) \omega} \cdot) | J_t(\omega),$$

for $\mu$-almost all $t \in M$, and for $P$-almost all $\omega \in \Omega$ and in $L^1(P)$. If $\theta_{\kappa_1^{(1)}(t) \kappa_2^{(2)}(t) \omega}$ is ergodic with respect to $P$, for $\mu$-almost all $t \in M$, then the limit simplifies to

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

4. Uniform convergence. This section addresses the question of sure convergence with respect to the first parameter. In addition to the assumptions at the beginning of Section 2 we suppose that $M$ is a compact separable space endowed with metric $d$, and $F$ is the Borel $\sigma$-algebra on $M$ for the topology induced by $d$. Assume that $\mu$ is the completion of a probability measure on $(M, \mathcal{B})$. There exists a largest open subset of $M$ with measure zero; its complement is called support of $\mu$. We assume that $\mu$ is globally supported, i.e., the support of $\mu$ equals $M$. A set $U \in \mathcal{B}$ is called $\mu$-boundaryless if its boundary $\partial U$ is a null set with respect to $\mu$. As before, $\tau$ is a measure-preserving transformation of $(M, \mathcal{B}, \mu)$.

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. Which conditions guarantee sure convergence in the first parameter? We will be asking a little more than this, namely about uniform convergence with respect to the first parameter, i.e.,

$$\frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_i(t)} \omega) \xrightarrow{n \to \infty} E[F(\mathcal{J})(t, \omega) \quad \text{uniformly in } M \text{ and in } L^1(P). \quad (11)$$

We further investigate when the convergence with respect to the second parameter takes place $P$-almost surely. In other words, for $P$-almost all $\omega \in \Omega$,

$$\lim_{n \to \infty} \sup_{t \in M} \left| \frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_i(t)} \omega) - E[F(\mathcal{J})(t, \omega) \quad = 0. \quad (12)$$
Again, we consider two different cases: when \( \tau \) is ergodic and when it is periodic. In the periodic case we proceed as in the proof of Corollary 3 and obtain the following uniform version.

**Corollary 4.** Let \( \tau \) be periodic with \( q \in \mathbb{N} \). Assume \( F \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathcal{P}) \) with \( F(t, \cdot) \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathcal{P}) \) for all \( t \in M \). Then, for \( \mathcal{P} \)-almost all \( \omega \) and in \( \mathcal{L}^1(\mathcal{P}) \),

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

where \( \mathcal{J}_t \) denotes the \( \sigma \)-algebra of \( \theta_{\kappa_\nu(t)} \)-invariant sets in \( \mathcal{F} \). If \( \mathcal{P} \) is ergodic with respect to \( \theta_{\kappa_\nu(t)} \), for all \( t \in M \), the limit equals \( \frac{1}{q} \sum_{\nu=0}^{q-1} \mathbb{E}[F(\tau^\nu(t), \cdot)] \).

The ergodic case is more delicate. In particularly, suitable assumptions on the transformation \( \theta \) in the fibre and on the function \( \kappa \) need to be worked out. We begin with separate investigations of suitable ergodic theorems on the base space and on the fibre spaces, which will later be combined to derive a result on the product space.

In the base we are dealing with a transformation on a compact and metrizable space. We begin by recalling and refining some of the existing results about uniform convergence in this situation. Motivated by applications in information theory, we put particular emphasis on extending uniform convergence results to the class of Riemann-integrable functions. An example for a function that has countably many points of discontinuity but is still Riemann-integrable occurs in the proof of a directional Shannon-MacMillan theorem for random fields (cf. [7]).

The classical example for a statement about the uniform convergence of ergodic averages goes back to Weyl. In the ergodic theory literature, the result is usually stated as follows. The averages of a continuous function along the orbit of an irrational rotation of the circle converge uniformly to the integral of the function. (A higher dimensional version dealing with suitable translations on tori is also due to Weyl.)

The question of uniform convergence has a close connection with unique ergodicity (cf. Chapter 4.1.e. of [21] or Theorem 6.19 in [30]). A continuous transformation \( \tau \) 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. The Lebesgue measure is the only probability measure on the circle \( \mathbb{T} \) equipped with the Borel \( \sigma \)-algebra \( \mathcal{B} \) that is invariant with respect to rotations, so it is uniquely ergodic. Oxtoby extended Weyl’s theorem to the situation where \( \tau \) is a uniquely ergodic transformation on a compact metric space. It states uniform convergence of the ergodic averages of a continuous function. Note that, conversely, uniform convergence does not imply the continuity of the function. (Further conditions for this would be needed, such as topological transitivity of \( \tau \) or constancy of the limit.)

In his original paper, Weyl [31] actually spelled out his theorem more generally than in the form quoted above. First, he states that the averages of a the function \( f \) along every uniformly distributed sequence converge uniformly to the integral of \( f \). In the version of Weyl’s theorem cited above, uniform sequences are replaced by orbits of irrational cycle rotations. Second, he assumes \textit{Riemann integrability} for \( f \) rather than continuity. Note that the statement can \textit{not} be further generalised to the set of all measurable functions. (For example, let \( f \) be the indicator function of the orbit \( O \).)
of a point $t_0$ under an irrational rotation and note that the ergodic averages converge to 1 for any point in $\mathcal{O}$, but $\int_0^1 1_{\mathcal{O}}(t) \, dt = 0$.) In fact, the uniform convergence of the ergodic averages of $f$ along all uniformly distributed sequences was later shown to be sufficient for Riemann integrability of $f$ (cf. [9]). Results of this type, but in the context of Riemann-integrability with respect to a probability measure $\lambda$, are established by [8] introducing the notion of $\lambda$-equidistributed sequences of partitions.

The concept of Riemann integration has been extended to general classes of topological spaces by Hauptmann and Pauc [16] based on the theory of Loomis [24]; see Bauer [4] for a summary of this work. For a function $f$ that is Riemann-integrable with respect to $\mu$ one can find continuous functions $g \leq f \leq h$, such that $\int_M h - g \, d\mu$ gets arbitrarily small. Jacobs uses this approach to define Riemann integration on measure spaces and calls it squeezing-in procedure (cf. Section 9 of Chapter I in [19]). On locally compact topological spaces, it can be shown that a function is Riemann integrable with respect to $\mu$ if and only if it is bounded with compact support and continuous $\mu$-almost everywhere (cf. Section 3.4 in [4] and [29]). Such a characterisation of Riemann-integrable functions can also be found in the short introduction to Riemann integration on compact metric spaces in Doob’s book (cf. Section 20 of Chapter VI in [10]).

The next result will help us to extend uniform ergodic theorems from the class of continuous functions to the class of Riemann-integrable functions.

**Proposition 1.** Assume that $\tau$ is continuous. If for any continuous function $f$ on $M$

$$\frac{1}{n} \sum_{i=0}^{n-1} f \circ \tau^i \xrightarrow{n \to \infty} \int_M f \, d\mu \quad \text{uniformly}, \quad (13)$$

then this convergence holds as well for any Riemann-integrable function $f$ on $M$.

**Proof.** Use the notation $A_n(\phi) := \frac{1}{n} \sum_{i=0}^{n-1} \phi \circ \tau^i$. Let $\varepsilon > 0$. Assume that $f$ is a Riemann-integrable function on $M$. Then there are continuous functions $g$ and $h$ on $M$ with $g \leq f \leq h$,

$$\int_M g \, d\mu - \frac{\varepsilon}{2} \leq \int_M f \, d\mu \quad \text{and} \quad \int_M h \, d\mu \leq \int_M f \, d\mu + \frac{\varepsilon}{2}. \quad (14)$$

Because of (13), there is an $n_0$ such that for all $n \geq n_0$,

$$\int_M g \, d\mu - \frac{\varepsilon}{2} \leq A_n(f) \leq \int_M h \, d\mu + \frac{\varepsilon}{2}$$

Using (14) and $A_n(g) \leq A_n(f) \leq A_n(h)$ yields for all $n \geq n_0$,

$$\int_M f \, d\mu - \varepsilon \leq A_n(f) \leq \int_M f \, d\mu + \varepsilon$$

Letting $\varepsilon$ go to 0 concludes the proof. \qed

**Remark about the proof:** The proof based on the construction of the Riemann integral by the squeezing-in technique with continuous functions as presented, for example, in [4] and [19]. Alternatively, the Riemann integral can be derived using an explicit approximation with Darboux sums (e.g. as in [10]). In that framework, Corollary 5 can by shown by first demonstrating Oxtoby’s theorem for step functions by sandwiching them between continuous functions (as done in the proof of Corollary 4.1.14 in [21]) and then deriving an analogue to Lemma 1 to pass from step functions to Riemann-integrable functions.
Corollary 5. Assume that $\tau$ is continuous and uniquely ergodic with invariant measure $\mu$. Then, for any Riemann-integrable function $f$ on $M$,

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

In particular, this Corollary states uniform convergence for indicator functions of $\mu$-boundaryless sets in $B$, because they are $\mu$-a.e. continuous and therefore integrable in the sense of Riemann.

Now we focus on studying the convergence in the fibre. Fix $t \in M$, and define a function on $\Omega$ by $f(\omega) := F(t, \omega)$. This reduces the ergodic averages of the skew product to $\frac{1}{n} \sum_{i=0}^{n-1} f(\theta_{k_i(t)} \omega)$, which we can view as a sort of ergodic average along the subsequence $(k_i(t))_{i \in \mathbb{N}}$. Recall a classical result about $L^p$-convergence of ergodic averages along subsequences from [6].

**Theorem 4.1.** (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 $f \in L^p(\Omega, \mathcal{F}, P)$,

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

The key to the proof is the following

**Lemma 4.2.** Under the assumptions of Theorem 4.1 and supposing that $P$ is strongly mixing with respect to $(\theta_k)_{k \in \mathbb{Z}^d}$, and let $(k_i(t))_{i \in \mathbb{N}} (t \in I)$ be a family of sequences with values in $\mathbb{Z}^d$ that fulfill, for all $m \in \mathbb{N}$,

$$
\lim_{n \to \infty} \frac{1}{n} \sup_{t \in I} \left| \left\{ 1 \leq i, j \leq n \mid \|k_i(t) - k_j(t)\| \leq m \right\} \right| = 0. \quad (15)
$$

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

$$
\sup_{t \in I} \left\| \frac{1}{n} \sum_{i=0}^{n-1} 1_A \circ \theta_{k_i(t)} - P(A) \right\|_{L^2(P)} \xrightarrow{n \to \infty} 0.
$$
Proof. For every $A \in \mathcal{F}$ and $t \in I$ we obtain by simple calculations,

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

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

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

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

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

(16)

Fix $\varepsilon > 0$. Due to the mixing condition there is an $m \in \mathbb{N}$ such that

\[
\left| P(\theta_{k_i(t)}^{-1} A \cap \theta_{k_j(t)}^{-1} A) - P(A)^2 \right| < \frac{\varepsilon}{2}
\]

for all $i, j \in \mathbb{N}$ with $\|k_i(t) - k_j(t)\| > m$, and

\[
\left| P(\theta_{k}^{-1} A \cap \theta_{k}^{-1} A) - P(A)^2 \right| < \frac{\varepsilon}{2}
\]

for all $k \in \mathbb{Z}^2$ with $\|k\| \geq m$.

By assumption (15) there is a $n_0 \in \mathbb{N}$ such that

\[
\frac{1}{n^2} \sup_{t \in I} \left| \{1 \leq i, j \leq n \mid \|k_i(t) - k_j(t)\| \leq m\} \right| < \frac{\varepsilon}{2}
\]

for all $n \geq n_0$.

Applying the last two inequalities to (16) yields for all $n \geq n_0$

\[
\sup_{t \in I} \left\| \frac{1}{n^2} \sum_{i=0}^{n-1} 1_A \circ \theta_{k_i(t)} - P(A) \right\|_{L^2(\Omega, F, P)}^2 < \varepsilon,
\]

and the assertion of the lemma follows by letting $\varepsilon$ to 0. \hfill \Box

**Theorem 4.4.** Assume that $P$ is strongly mixing with respect to $(\theta_k)_{k \in \mathbb{Z}^d}$. Let $(k_n(t))_{n \in \mathbb{N}}$ (for $t \in I$) be a family of sequences with values in $\mathbb{Z}^d$, for which for all $m \in \mathbb{N}$,

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

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

\[
\frac{1}{n} \sum_{i=0}^{n-1} f \circ \theta_{k_i(t)} \xrightarrow{n \to \infty} E[f] \quad \text{in } L^p(P), \quad \text{uniformly in } t \in I.
\]

(17)

**Proof.** As an immediate consequence of the preceding lemma, (17) is true for $p = 2$ for any simple function $g$ on $(\Omega, F)$. By a standard argument (cf., e.g., Lemma 4 in [6]), this convergence holds as well in $L^p(P)$, for $1 \leq p < \infty$. Finally, for any function $f$ in $L^p(P)$, decomposition into positive and negative parts, $L^p(P)$-approximation by simple functions, and monotone convergence yields (17). \hfill \Box
We will now combine the approaches developed separately for the base and the fibre transformation to derive a result about uniform convergence in the base and \( \mathcal{L}^p \)-convergence in the fibre. A crucial ingredient is a condition that regulates the effects of the coupling sequence \((\kappa_n(t))_{n \in \mathbb{N}}\).

**Theorem 4.5.** Assume that \( \tau \) is continuous and uniquely ergodic, and suppose that \( P \) is strongly mixing with respect to the group of transformations \((\theta_k)_{k \in \mathbb{Z}^d}\). Let \( \kappa : M \to \mathbb{Z}^d \) be \( \mathcal{B} \)-measurable such that,

\[
\lim_{n \to \infty} \frac{1}{n^2} \sup_{t \in M} \{ 1 \leq i, j \leq n \mid \| \kappa_i(t) - \kappa_j(t) \| \leq m \} = 0 \tag{18}
\]

for all \( m \in \mathbb{N} \). Let be \( 1 \leq p < \infty \). Then for every \( F \)-measurable function \( F \) on \( \Omega \) 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 \( \omega \),

\[
\frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_i(t)} \omega) \xrightarrow{n \to \infty} E[F] \quad \text{in } \mathcal{L}^p(P) \text{ and uniformly in } t \in M. \tag{19}
\]

**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 \( \mu \)-boundaryless metric balls in \( M \) or their complements, and \( A \in \mathcal{F} \). By (5), the expression

\[
\left\| \frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_i(t)} \omega) - E[F] \right\|_{\mathcal{L}^2(P)} \tag{20}
\]

then transforms to

\[
(*) = \frac{1}{n^2} \sum_{i,j=0}^{n-1} \int_{\Omega} 1_U(\tau^i(t))1_U(\tau^j(t))1_A(\theta_{\kappa_i(t)} \omega)1_A(\theta_{\kappa_j(t)} \omega) P(d\omega) + \mu(U)^2 P(A)^2
\]

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

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

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

It may be replaced by

\[
\frac{1}{n^2} \sum_{i,j=0}^{n-1} 1_U(\tau^i(t))1_U(\tau^j(t))P(A)^2 \tag{21}
\]

without affecting the asymptotic behavior as can be seen as follows. We may bound

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

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

Now, we argue as in the second part of the proof of Lemma 4.3, replacing the sequence \((\kappa_i)_{i \in \mathbb{N}}\) by \((\kappa_i(t))_{i \in \mathbb{N}}\), and using assumption (18) instead of (15). This
proves that the difference created by the change (21) converges to 0 uniformly with respect to t.

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

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

which can be further reduced to \( \left( 1/n \sum_{i=0}^{n-1} 1_U(\tau^i(t)) - \mu(U) \right)^2 P(A)^2 \). Since \( \tau \) is uniquely ergodic and \( U \) is \( \mu \)-boundaryless, Lemma 5 tells us that the first factor converges to 0 uniformly in t, which concludes the first part of the proof. To pass from \( L^2 \)-convergence to general \( L^p \), use again a standard argument (for instance, Lemma 4 in [6]).

Now we let \( F \) be a general function, satisfying the conditions of the theorem. We need to find, for every positive \( \varepsilon \), a sequence of \( \mu \)-boundaryless metric balls \( U_i \) in \( M \) and \( A_t \in \mathcal{F} \), with non-zero boundaries under \( M \), which can be further reduced to \( 1_U(\tau^i(t)) - \mu(U) \) for every \( t \) and \( i \). We also define a sequence of real numbers \( a_i \), such that, for all \( t \in M \), \( \| F(t, \cdot) - I(t, \cdot) \|_p < \varepsilon \), where \( I = \sum_{i=1}^n a_i 1_{U_i \times A_t} \). It will then follow that

$$
\left\| \frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\kappa_i(t)} \cdot) - \frac{1}{n} \sum_{i=0}^{n-1} I(\tau^i(t), \theta_{\kappa_i(t)} \cdot) \right\|_p \\
\leq \frac{1}{n} \sum_{i=0}^{n-1} \left\| F(\tau^i(t), \cdot) - I(\tau^i(t), \cdot) \right\|_p < \varepsilon.
$$

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

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

Then the sequence of functions on \( \Omega \),

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

is bounded by \( \sup_{t \in M} \|F\|^p(t, \omega) \), which is integrable, and converges to 0 for every \( \omega \). 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 \( t_i \) \( (i = 1, \ldots, r) \) in \( M \) and metric balls with radius \( \rho_i \leq 1/k \) around these centers cover \( M \). The balls can be chosen to be \( \mu \)-boundaryless, because each point in \( M \) is the center of at most countably many metric balls with non-zero boundaries under \( \mu \). 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 \( \rho_i \) around \( t_i \), and then remove the intersections, so that the \( U_{i,j} \) is the same for all \( j \), and running \( i \) through 1, \ldots, \( r \) yields a disjoint cover of \( M \) consisting of \( \mu \)-boundaryless metric balls and their complements. The sets \( A_{i,j} \) are defined by

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

Let \( a_{i,j} = s_j \). We throw in one additional product set, \( U_0 = M \) and \( A_0 = D_k(\varepsilon/8) \cap M_k \) with \( a_0 = s_0 \), and define the simple function \( I(t, \omega) \) as indicated above. Then
for any $t \in M$, $\|F(t, \cdot) - I(t, \cdot)\|_p$ equals

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

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 $2^p |F(t_i, \omega) - F(t, \omega)|^p + 2^p |F(t_i, \omega) - s_j|^p$. This in turn is bounded by $2^p (\varepsilon/4)^p < (\varepsilon/2)^p$, since $\omega$ is not in $A_0$ and $d(t_i, t) < 1/k$, completing the proof. \qed

Proposition 1 yields the following version for Riemann-integrable functions.

**Corollary 6.** Assume that $\tau$ is continuous and uniquely ergodic with invariant measure $\mu$. For $P$ and $\kappa$ assume the same as in Theorem 4.5. Let $F \in L^p(\Omega, F, P)$ be Riemann-integrable with respect to the first variable. Then we have

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

Our next goal is to derive a statement about $P$-almost sure convergence rather than $L^1(P)$-convergence in the fibre. Further conditions on $F$ are needed. $P$-almost sure convergence of ergodic theorems involving weights or subsequences is a very subtle question (cf., e.g., [5]). To motivate equicontinuity assumptions we are going to make on $F$, we recall that Krengel (cf. Theorem 2.6 in Paragraph 1.2.3 in [22]) uses an Arzela-Ascoli technique to proof Weyl’s theorem. Under the assumption that $\tau : M \to M$ is continuous and that $f$ is a function on $M$, such that the functions $F_n := 1/n \sum_{i=1}^{n-1} 1_{F \circ \tau^i}$ $(n \in \mathbb{N})$ are equicontinuous on $M$, Krengel’s theorem states that the convergence in Birkhoff’s ergodic theorem is uniform in $t$. Together with the following Lemma, this yields Weyl’s theorem.

**Lemma 4.6.** Let $f$ be a continuous function on $M$, and $\tau : M \to M$ Lipschitz-continuous with Lipschitz constant $c \leq 1$. Then the functions $F_n$ $(n \in \mathbb{N})$ defined above are equicontinuous.

**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 $d(s, t) < \delta$, $1/n \left| \sum_{i=0}^{n-1} f(\tau^i(s)) - f(\tau^i(t)) \right| < \varepsilon$. 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 $i \in \mathbb{N}$. \qed

Let us go back to the product space setting. Krengel’s approach suggests that we need an equicontinuity assumption in the first argument of $F$. (For example, choose a function $F$ which is constant in $\omega$.)

**Theorem 4.7.** Let $\mu$ be a $\tau$-invariant measure on $(M, \mathcal{B})$ with $\mu(U) > 0$ for all non-empty open $U \subset M$. Let $\tau : M \to M$ be continuous and $F$ a function on $M \times \Omega$, for which $F(t, \cdot) \in L^1(P)$ for all $t \in M$, and the sequence of functions

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

uniformly converges to $\mathbb{E}[F|\mathcal{F}]$. Then the conditional expectation $\mathbb{E}[F|\mathcal{F}]$ is well-defined and continuous in $\omega$. Further, $\mathbb{E}[F|\mathcal{F}] \in L^p(M, \mathcal{B}, \mu)$ for all $p < \infty$.
is equicontinuous on \( M \), for all \( \omega \in \Omega \). Then, for \( P \)-almost all \( \omega \in \Omega \),

\[
\sup_{t \in M} \left| \frac{1}{n} \sum_{i=0}^{n-1} F(\tau^i(t), \theta_{\mu_i}(t)\omega) - \mathbb{E}[F[J](t, \omega)] \right| \xrightarrow{n \to \infty} 0. \tag{22}
\]

**Proof.** We may assume without loss of generality that \( E[F[J]] = 0 \). The general case can be reduced to this by subtracting \( E[F[J]] \) on both sides and making use of the invariance of \( E[F[J]] \) 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

\[
\frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(t, \omega) \to 0 \quad \text{for all } t \in M_1 \text{ and all } \omega \in \Omega \setminus N_1. \tag{23}
\]

Since \( M \) is compact, the conditions on \( F \) assure that \( F \in \mathcal{L}^1(\Omega, \mathcal{F}, \mathbb{P}) \), and therefore by (9) there is a set \( \tilde{M}_1 \subset M \) with \( \mu(\tilde{M}_1) = 1 \) such that for any \( t \in \tilde{M}_1 \) there is a set \( N(t) \subset \Omega \) with \( P(N(t)) = 0 \) and (23) holds for all \( \omega \in \Omega \setminus N(t) \). The complement of \( \tilde{M}_1 \) in \( M \) has \( \mu \)-measure zero. As the support of \( \mu \) equals \( M \), it cannot contain any non-empty open subsets. So, \( \tilde{M}_1 \) is dense in \( M \). Since \( M \) is separable we can find a countable dense subset \( C \subset M \), and because \( \tilde{M}_1 \) is dense in \( M \), we can approximate any \( x \in C \) by a sequence \((a_j(x))_{j \in \mathbb{N}} \) with \( a_j(x) \in \tilde{M}_1 \) for all \( j \in \mathbb{N} \). \( M_1 := \bigcup_{c \in C} \bigcup_{j \in \mathbb{N}} a_j(x) \) defines a countable dense subset of \( M \), and \( N_1 := \bigcup_{t \in M_1} N(t) \) defines a subset of \( \Omega \), which fulfills (23). This completes the first step.

For the next step, choose \( s \in M \) and fix \( \epsilon > 0 \). 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 \( d(r, t) < \delta \)

\[
\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(r, \omega) - F \circ S^i(t, \omega) \right| < \frac{\epsilon}{2} \quad \text{for all } n \in \mathbb{N} \text{ and all } \omega \in \Omega \setminus N_0. \tag{24}
\]

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 (23) there is an \( n_1 \in \mathbb{N} \) such that

\[
\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(t, \omega) \right| < \frac{\epsilon}{2} \quad \text{for all } n \geq n_1.
\]

Combining the last two inequalities leads the desired

\[
\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(s, \omega) \right| < \epsilon \quad \text{for all } n \geq n_1 \text{ and all } \omega \in \Omega \setminus N.
\]

For uniform convergence with respect to the first variable, we use a standard compactness argument. \( M \) can be covered by a finite number \( m \) of \( \delta \)-neighborhoods in \( M \), which centers are denoted by \( s_1, \ldots, s_m \). Applying the reasoning of the last step to each of the \( s_1, \ldots, s_m \) we find \( n_0 \in \mathbb{N} \) such that

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

For an arbitrary \( s \in M \) there exists \( k \in \{1, \ldots, K\} \) such that \( d(s, s_k) < \delta \), and by (24) we obtain

\[
\left| \frac{1}{n} \sum_{i=0}^{n-1} F \circ S^i(s, \omega) - F \circ S^i(s_k, \omega) \right| < \frac{\epsilon}{2} \quad \text{for all } n \in \mathbb{N} \text{ and all } \omega \in \Omega \setminus N.
\]
Finally, the convergence (22) follows by the last two inequalities.

We conclude the paper with an application of our results to the question that originally motivated this work. Consider the ergodic averages of a function of a random field restricted to the staircase pattern defined by the approximation of a line by a lattice as in (1). We actually allow the function to be multivariate to accommodate that, in the staircase pattern context, it may depend on a finite number of steps.

Let \( P \) be a two-dimensional random field, that is, a probability measure on \( \Omega = \mathcal{Y}^2 \) invariant with respect to the group of shift transformations \( (\vartheta_v)_{v \in \mathbb{Z}^2} \) (see just above Corollary 1 for details). For a set \( U \subseteq \mathbb{Z}^2 \) let \( \pi_U(\omega) := \omega(U) \). Then \( P_U := P \circ \pi_U^{-1} \) defines a probability distribution on \( \mathcal{Y}^U \). \([x] \) denotes the integer part of a real number \( x \). Recall the notation \( L_{\lambda,t} \) used in the introduction, to denote the approximation of the line with slope \( \lambda \) and \( y \)-intercept \( t \) by elements of the lattice \( \mathbb{Z}^2 \). We will also use the short form \( P_{\lambda,t,m} \) for the probability distribution \( P_{\lambda,t,m}(0,\ldots,m-1) \) on \( \mathcal{T}^m \).

**Corollary 7.** Let \( m \in \mathbb{N} \) and \( \lambda \in [0,1] \). Let \( f \) be a function on \( \mathcal{Y}^m \) which is integrable with respect to \( P_{\lambda,t,m} \) for all \( t \in [0,1] \).

(i) Let \( \lambda \) be rational. If \( P \) is ergodic with respect to the shift transformations then

\[
\frac{1}{n} \sum_{i=0}^{n-1} f(L_{\lambda,t}(i,\ldots,i+m-1)) \xrightarrow{n \to \infty} \frac{1}{q} \sum_{\nu=0}^{q-1} \int_{\mathcal{T}^m} f(y) P_{\lambda,\tau^\nu(t),m}(dy)
\]

in \( \mathcal{L}^1(P) \), for \( P \)-almost all \( \omega \in \Omega \), and uniformly in \( t \in [0,1] \).

(ii) Let \( \lambda \) be irrational. Assume further that \( f(\omega(L_{\lambda,t}(0,\ldots,m-1))) \) is Riemann-integrable with respect to \( t \in [0,1] \), for \( P \)-almost all \( \omega \in \Omega \). If \( P \) is strongly mixing with respect to the shift transformations then

\[
\frac{1}{n} \sum_{i=0}^{n-1} f(L_{\lambda,t}(i,\ldots,i+m-1)) \xrightarrow{n \to \infty} \int_0^1 \int_{\mathcal{T}^m} f(y) P_{\lambda,t,m}(dy) \, dt
\]

in \( \mathcal{L}^1(P) \) and uniformly in \( t \in [0,1] \).

**Proof.** Let \( \tau \lambda \) be the rotation on the circle defined in Example 1, and \( S_\lambda \) the skew product defined in (2). Use \( \kappa(t) := (1, [t + \lambda]) \) and \( \kappa_n \) as defined in (5). It is easy to show that \( \kappa_n(t) = L_{\lambda,t}(n) \) for all \( n \in \mathbb{N} \) : For \( n = 1 \), it follows immediately from plugging in the definition, and it remains to induce the statement from \( n \) to \( n + 1 \).

It is obvious for the first coordinate. The second coordinate of \( \kappa \) can be written as \( \kappa_{n+1}^{(2)}(t) = \kappa_n^{(2)}(t) + \kappa \circ \tau_\lambda^n(t) = [n\lambda + t] + \tau_\lambda^n(t) + \lambda \). The claim now follows from \([\tau_\lambda^n(t) + \lambda] = [t + n\lambda - [t + n\lambda] + \lambda] = -[t + n\lambda] + [t + (n + 1)\lambda] \).

This implies that the iterates of the skew product are of the form \( S_\lambda^n(t,\omega) = (\tau_\lambda^n(t), \theta_{L_{\lambda,t}(n)}(\omega)) \) (\( n \in \mathbb{N} \)) capturing the lattice approximation of the line. Using the second equality in (5), it also follows that \( L_{\lambda,t}(n+u) = L_{\lambda,t}(n) + L_{\lambda,\tau^n(t)}(u) \) for all \( n, u \in \mathbb{N}_0 \). We thus get \( L_{\lambda,t}(i,\ldots,i+m-1) = L_{\lambda,t}(i) + L_{\lambda,\tau^n(t)}(0,\ldots,m-1) \).

Introducing the function \( F_\lambda(t,\omega) := f(L_{\lambda,t}(0,\ldots,m-1)) \) we obtain

\[
\frac{1}{n} \sum_{i=0}^{n-1} f(L_{\lambda,t}(i,\ldots,i+m-1))
\]

\[
= \frac{1}{n} \sum_{i=0}^{n-1} f \circ \theta_{L_{\lambda,t}(i)} \left( \omega(L_{\lambda,\tau^n(t)}(0,\ldots,m-1)) \right) = \frac{1}{n} \sum_{i=0}^{n-1} F_\lambda \circ S_\lambda(t,\omega). \quad (25)
\]
Consider case (i). There exists a unique representation $\lambda = p/q$ for $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ with no common divisor. $\tau_\lambda$ is periodic with length $q$. By Corollary 4 and (25), the averages on the left side in (i) converge uniformly in $t$, for $P$-almost all $\omega$ and in $L^1(P)$. Since $P$ is ergodic with respect to $\theta_{\kappa_n(t)}$ for all $t \in M$, the limit equals $1/q \sum_{n=0}^{q-1} \int_\Omega F_t(\tau^n(t), \cdot) P(d\omega)$ which simplifies to the expression on the right hand side of (i).

Consider case (ii). For an irrational $\lambda$, $\tau_\lambda$ is uniquely ergodic. Since $\|\kappa_n(t)\|$ (with $\| \cdot \|$ for the maximum norm) is bounded from below by $n$, the sequence tends to infinity as $n$ goes to infinity. Corollary 1 with $v_1 = (1,0)$ and $v_2 = (0,1)$ implies the ergodicity of $S_\lambda$. It remains to verify condition (18). The latter easily follows from $|\kappa_i(t) - \kappa_j(t)| \geq |i - j|$, and since $1/n^2 \{ 1 \leq i, j \leq n \, | i-j | \leq m \}$ converges to 0 for all $m \in \mathbb{N}$. Now, Corollary 6 applied to (25) implies that the averages in (ii) converge uniformly in $t$, and in $L^1(P)$. The limit equals $\int_0^1 \int_\Omega F_t(\cdot) P(d\omega) \, dt$ which simplify to the expression on the right hand side of (ii).

Note that $P$-almost everywhere convergence in (ii) can be derived as well, but requires additional conditions of the form stated in Theorem 4.7. For $m = 1$ the limit actually simplifies to the integral of $f$ with respect to the marginal distribution of $P$ in the origin. In particular, it is independent of $t$, and it is the same in (i) and (ii). The simplest interesting case is $m = 2$. Let $P_{\text{flat}} := P \circ \pi_{\{(0,0),(1,0)\}}^{-1}$ denote the marginal distribution of $P$ on the subset $\{(0,0),(1,0)\}$. This corresponds to the case where $L_{\lambda,t}$ does not have a step in $z = 1$. Let $P_{\text{step}} := P \circ \pi_{\{(0,0),(1,1)\}}^{-1}$ denote the marginal distribution of $P$ on the subset $\{(0,0),(1,1)\}$. This corresponds to the case where $L_{\lambda,t}$ has a step in $z = 1$. Then, the limits in both (i) and (ii) of the above corollary are of the form

$$
\lambda \int_{\mathbb{T}^2} f(y) P_{\text{step}}(dy) + (1 - \lambda) \int_{\mathbb{T}^2} f(y) P_{\text{flat}}(dy).
$$

Acknowledgments. This work was motivated by a special case of Corollary 7 which originally occurred during my thesis work. I am grateful to my supervisor Hans Föllmer for the rich education he gave me in both ergodic theory and probability theory. It is a pleasure for me to thank Jon Aaronson for guiding me toward a broader perspective for Lemma 2.2. David Steinsaltz assisted with an approximation argument I used in the proof of Theorem 4.5, and I thank him for this hint. Benjamin Weiss and Yuval Peres helped me situate my results within the field of ergodic theory; I am grateful to them for the interest they showed in my work. I would like to express my gratitude to Bill Mackillop for supporting my work during the time I spent at Queen’s University. I further thank Hans Crauel, Manfred Denker, Didier Piau, Michael Scheutzow and Masha Saprykina for helpful discussions on different aspects of ergodic theory.

REFERENCES

[1] R. Adler and P. Shields, *Skew products of Bernoulli shifts with rotations*, Israel J. Math., 12 (1972), 215–222.
[2] R. Adler and P. Shields, *Skew products of Bernoulli shifts with rotations. II*, Israel J. Math., 19 (1974), 228–236.
[3] H. Anzai, *Ergodic skew product transformations on the torus*, Osaka Math. J., 3 (1951), 83–99.
[4] H. Bauer, *Über die Beziehungen einer abstrakten Theorie des Riemann-Integrals zur Theorie Radonischer Maße*, Math. Z., 65 (1956), 448–482.
UNIFORM CONVERGENCE WITH SKEW PRODUCTS

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

[6] J. R. Blum and D. L. Hanson, On the mean ergodic theorem for subsequences, Bull. Am. Math. Soc., 66 (1969), 308–311.

[7] J. Brettschneider, Shannon-MacMillan theorems for random fields along curves and lower bounds for surface-order large deviations, Prob. Th. Rel. Fields, 142 (2007), 443–473.

[8] F. Chersi and A. Volčič, $\lambda$-Equidistributed sequences of partitions and a theorem of the De Bruijn-Post type, Annali di Matematica Pura ed Applicata, 162 (1992), 23–32.

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

[10] J. L. Doob, “Measure Theory,” Springer-Verlag, New York, 1993.

[11] F. den Hollander and M. Keane, Ergodic properties of color records, Physica A, 138 (1986), 183–193.

[12] N. Friedman, Mixing on sequences, Can. J. Math., 35 (1983), 339–352.

[13] H. Furstenberg, Strict ergodicity and transformation of the torus, Amer. J. Math., 83 (1961), 573–601.

[14] H.-O. Georgii, “Gibbs Measures and Phase Transitions,” W. de Gruyter, Berlin, 1988.

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

[16] O. Hauptmann and C. Pauc, “Differential - und Integralrechnung, Band III,” 2. Auflage, Göschen Lehrbücherei, 26, Berlin, 1955.

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

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

[19] K. Jacobs, “Measure and Integral,” Academic Press, New York, 1978.

[20] S. Kakutani, Random ergodic theorems and Markov processes with a stable distribution, in “Proceedings of the Second Berkeley Symposium on Statistics and Probability” (ed. J. Neyman), Univ. of California Press, Berkeley, (1951), 247–261.

[21] A. Katok and B. Hasselblatt, “Introduction to the Modern Theory of Dynamical Systems,” Cambridge Univ. Press, Cambridge, 1995.

[22] U. Krengel, “Ergodic theorems,” W. de Gruyter, Berlin, 1985.

[23] M. Lemańczyk and E. Lesigne, Ergodicity of Rokhlin cocycles, J. Anal. Math., 85 (2001), 43–86.

[24] L. H. Loomis, Linear functional and content, Amer. J. Math., 76 (1954), 68–82.

[25] I. Meilijson, Mixing properties of a class of skew-products, Israel J. Math., 19 (1974), 266–270.

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

[27] I. Oren, Ergodicity of cylinder flows arising from irregularities of distribution, Israel J. Math., 44 (1983), 127–138.

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

[29] C. Pau, Intégrale de partition et intégrale topologique. Familles dérivantes topologiques, C. r. Acad. Sci. Paris, 230 (1950), 810–811.

[30] P. Walters, “An Introduction to Ergodic Theory,” Springer-Verlag, New York, Heidelberg, Berlin, 1982.

[31] H. Weyl, Über die Gleichverteilung von Zahlen mod Eins, Math. Ann., 77 (1915/16), 313–352.

[32] Q. Zhang, On skew products of irrational rotations with tori, in “Convergence in Ergodic Theory and Probability” (eds. V. Bergelson, P. March, J. Rosenblatt), Ohio State University Mathematical Research Institute Publications 5, (1996), 435–445.

Received October 2009; revised August 2010.

E-mail address: julia.brettschneider@warwick.ac.uk