Ehrhard Behrends; Jörg Schmeling

Strongly mixing sequences of measure preserving transformations

*Czechoslovak Mathematical Journal*, Vol. 51 (2001), No. 2, 377–385

Persistent URL: http://dml.cz/dmlcz/127654

# STRONGLY MIXING SEQUENCES OF MEASURE PRESERVING TRANSFORMATIONS

Ehrhard Behrends, Jörg Schmeling, Berlin

(Received August 6, 1998)

*Abstract.* We call a sequence $(T_n)$ of measure preserving transformations *strongly mixing* if $P(T_n^{-1}A\cap B)$ tends to $P(A)P(B)$ for arbitrary measurable $A$, $B$. We investigate whether one can pass to a suitable subsequence $(T_{n_k})$ such that $\frac{1}{K}\sum\limits_{k=1}^{K} f(T_{n_k}) \longrightarrow \int f\,\mathrm{d}P$ almost surely for all (or "many") integrable $f$.

*Keywords*: ergodic transformation, strongly mixing, Birkhoff ergodic theorem, Komlós theorem

*MSC 2000*: 28D05

## 1. Introduction

Let $(\Omega,\mathcal{A},P)$ be a Polish probability space and $(T_n)$ a sequence of measure preserving transformations. This sequence will be called *strongly mixing* if

$$\lim_{n\to\infty} P({T_n}^{-1}A\cap B)=P(A)P(B)$$

for arbitrary measurable sets $A$, $B$.

If $T_n$ are of the special form $T^n$ for a fixed transformation $T$ then classical ergodic theory may be applied in order to investigate various mixing properties (see [6] for a good account of the relevant results). Here we have in mind working in the more general setting of arbitrary sequences of transformations, in particular we are interested in the density of orbits and the counterparts of ergodic theorems. One major point of our investigations is to what extent the classical Birkhoff theorem holds. It will turn out that by choosing appropriate subsequences of transformations we can derive an individual ergodic theorem for some class of $L^1$-functions. It is not to be expected that this kind of results holds for general $L^1$-functions (see Section 5 for discussions). Moreover, even if all the transformations are powers of a given transformation there are general examples where the Birkhoff theorem fails in $L^1$ for arbitrary subsequences (see [4]). We discuss in Section 5 how large the class of functions can be in order to get an individual ergodic theorem for subsequences.

Condition (1.1) is not the weakest mixing property. Usually one starts with ergodicity which means that the right-hand-side limit in (1.1) is attained only in the Cesàro mean: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{\infty} ({T_k}^{-1} A \cap B) = P(A)P(B)$. We call such sequences of transformations *ergodic*. Although some of the following results can be derived under the latter condition we prefer (1.1). The reason is that the ergodicity condition is not invariant under taking subsequences while (1.1) is. In what follows we will sometimes use the equivalent condition:

$$\lim_{n\to\infty} \int_\Omega f(T_n x) g(x) P(\mathrm{d}x) = \int_\Omega f(x) P(\mathrm{d}x) \int_\Omega g(x) P(\mathrm{d}x) \tag{1.2}$$

for all $L^1$-functions $f$ and $g$ such that $g \cdot (f \circ T_n)$ is in $L^1$ for sufficiently large $n$. The last condition is fulfilled whenever $f \in L^1$, $g \in L^\infty$ or both $f, g \in L^2$.

## 2. Examples of strongly mixing sequences

The first examples are strongly mixing dynamical systems, i.e. $T_n = T^n$ for $T\colon \Omega \to \Omega$ is a strongly mixing map. In this case all classical ergodic theorems hold.

More interesting are subsequences of $(T^n)_n$. These are treated in [4].

Let $\varphi\colon \mathbb{N} \to \mathbb{N}$ be an unbounded monotone increasing function. Then the sequence $T_n = S^{\varphi(n)}$ is mixing provided $S$ is so. We note that for strictly monotone increasing sequences we are in the case of subsequences. We also can allow $\varphi(n)$ to be eventually monotone increasing or even to have the property that the preimage of any compact subset of $\mathbb{N}$ is compact. This leads to special rearrangements of the previous sequences.

It is not hard to check that sequences of the form $T_n\colon [0,1) \to [0,1)$, defined by $T_n = a_n x \bmod(1)$ where $(a_n)$ is an unbounded monotone increasing subsequence of the natural (or even real) numbers is strongly mixing.

Let $\Psi_t\colon M \to M$, $t \in \mathbb{N}$ be a strongly mixing flow with respect to some invariant measure. Then for any unbounded increasing sequence of real numbers $t_1, \dots, t_n, \dots$ the sequence $t_n$ of time maps $T_n := \Psi_{t_n}\colon M \to M$ is strongly mixing.

**Lemma 3.1.** *Let $(T_n)$ be strongly mixing. If $A$ is measurable such that $P(A) > 0$ then $P(\bigcup_{n\geqslant n_0} T_n^{-1}(A)) = 1$ for every $n_0$.*

P r o o f. Denote by $B$ the complement of $\bigcup_{n\geqslant n_0} T_n^{-1}(A)$. Clearly we have that $P(T_n^{-1}(A) \cap B) = 0$ for all $n \geqslant n_0$. But $P(T_n^{-1}(A) \cap B)$ tend to $P(B)P(A)$, hence $P(B) = 0$. □

This lemma shows that mixing transformations are sweeping out in the measure sense. It copies the classical lemma for ergodic systems (see [6], Theorem 1.5) and shows that the assertion does not depend on the fact that it is a classical ergodic system, i.e. the consecutive powers of a single transformation. The next proposition shows that there is also a topological variant of this result in the general case. We want to remark that for both results the weaker notion of ergodicity is sufficient.

**Proposition 3.2.** *Let $(M, d)$ be a compact metric space, and $P$ a Borel measure such that $P(U) > 0$ whenever $U$ is an open nonvoid subset. Then, for every strongly mixing sequence $(T_n)$, the orbit $\{T_n(x)\colon\ n = 1, \ldots\}$ is dense for almost every $x$.*

P r o o f. Let $U_1, U_2, \ldots$ be a basis of the topology. Then, as is easy to see, an orbit $\{T_n(x)\colon\ n = 1, \ldots\}$ is dense iff $x$ lies in $\bigcap_{r,n_0} \bigcup_{n\geqslant n_0} T_n^{-1}(U_r)$. By the preceding lemma $\bigcup_{n\geqslant n_0} T_n^{-1}(U_r)$ have full measure for all $r, n_0$ so that we are done. □

## 4. Mean values on the orbit approximate the integral

Let $(T_n)$ be a strongly mixing sequence. We will use the following notation: For any scalar-valued measurable function $f$ and any $x$ the arithmetic mean $(f(T_1(x)) + \ldots + f(T_n(x)))/n$ will be abbreviated by $S_n(f, x)$. (Sometimes we will pass to a subsequence $(T_{n_k})$ of $(T_n)$. Then $S_n(f, x)$ are meant to be defined with respect to this subsequence.)

In this section we will be concerned with the problem to what extent $S_n(f, x)$ tend to $\int f\,\mathrm{d}P$.

**Lemma 4.1.** *For all $f \in L^1$ and almost all $x$ one has* $\liminf S_n(f, x) \leqslant \int f\,\mathrm{d}P \leqslant \limsup S_n(f, x)$.

P r o o f. For integers $n_0$, $k$ let $A_{n_0,k}$ be the set

$$\left\{x\colon\ S_n(f, x) \leqslant \int f\,\mathrm{d}P - 1/k \quad \text{for all} \quad n \geqslant n_0\right\};$$

denote by $g$ the characteristic function of this set. Then, whenever $n \geqslant n_0$, one has

$$\int g(x) S_n(f,x) P(\mathrm{d}x) = \int_{A_{n_0,k}} S_n(f,x) P(\mathrm{d}x) \leqslant \left( \int f \, \mathrm{d}P - 1/k \right) P(A_{n_0,k}).$$

On the other hand, as $n$ goes to infinity the integrals $\int g(x) S_n(f,x) P(\mathrm{d}x)$ tend to $\int g \, \mathrm{d}P \int f \, \mathrm{d}P = P(A_{n_0,k}) \int f \, \mathrm{d}P$ by formula (1.2) so that necessarily $P(A_{n_0,k}) = 0$. Thus

$$P\left( \left\{ x \colon \limsup S_n(f,x) < \int f \, \mathrm{d}P \right\} \right) = P\left( \bigcap_{n_0,k} A_{n_0,k} \right) = 0.$$

Similarly it is shown that $\liminf S_n(f,x) \leqslant \int f \, \mathrm{d}P$ for almost all $x$. □

In general it is not to be expected that the preceding lemma could be sharpened in that $S_n(f,x)$ would converge to $\int f \, \mathrm{d}P$.

Consider *any* mixing sequence $\tilde{T}_n$ on any nontrivial $\Omega$ and consider a new sequence $T_1, T_2, \ldots$ defined by $\tilde{T}_1, \tilde{T}_2, \tilde{T}_2, \tilde{T}_2, \tilde{T}_3, \tilde{T}_3, \ldots$, (1 time $\tilde{T}_1$, 3 times $\tilde{T}_2$, 12 times $\tilde{T}_3, \ldots$ ($\tilde{T}_k$ occurs $3r_k$ times, where $r_k$ are defined by $r_1 := 1, r_{k+1} := 3(r_1 + \ldots + r_k)$). The new sequence surely is also strongly mixing, but the construction ensures that the sequence $(S_n(f,x))$ oscillates; for example, if $f$ is the characteristic function of a set $A$ with $P(A)$, $P(\Omega \setminus A) > 0$, then $S_n(f,x)$ will be $\geqslant 2/3$ and $\leqslant 1/3$ infinitely often for almost all $x$.

Let $(b_n)$ be a bounded sequence of real numbers. Then it is easy to see that for any $b$ lying between the *lim inf* and the *lim sup* of the Cesàro means there is a subsequence which is Cesàro convergent to $b$. One might suspect that this elementary argument would yield for any bounded $f$ a subsequence of the original $(T_n)$ such that $S_n(f,x)$ for this subsequence converge almost everywhere to $\int f \, \mathrm{d}P$. However, since one might have to take into account uncountably many $x$ it is not obvious whether one may pass from numbers to functions.

Nevertheless, it is possible to prove a variant of Birkhoff's theorem:

**Theorem 4.2.** *Let $(T_n)$ be strongly mixing and $\mathcal{F} = \{f_1, f_2, \ldots\}$ a countable subset of $L^1$. Then there is a subsequence $(T_{n_k})$ such that the associated $S_n(f,.)$ tend to $\int f \, \mathrm{d}P$ almost everywhere for all $f \in \mathcal{F}$.*

Proof. This follows from the *Komlós theorem* ([3]): Whenever $(g_m)$ is a uniformly bounded sequence in $L^1$, there are an $L^1$-function $g$ and a subsequence $(g_{m_k})$ such that every subsequence of this subsequence is pointwise Cesàro convergent almost everywhere to $g$.

To derive our result apply the Komlós theorem to the sequence $(f_1 \circ T_n)_n$; a suitable subsequence $(f_2 \circ T_{n_k})$ will have the property that all subsequences are

Cesàro convergent almost everywhere. Now consider $(f_2 \circ T_{n_k})_k$ and choose a suitable subsequence again. It should be clear now how to proceed and that a standard diagonal construction leads to a subsequence of $(T_n)_n$—which we will denote by $(T_{n_k})_k$ again—such that $(f_r \circ T_{n_k})$ are Cesàro convergent almost everywhere for all $f_r$. It now only remains to note that the respective limit necessarily is $\int f_r \, \mathrm{d}P$ by the preceding lemma.

Under additional conditions on $f_1, \ldots$ one can give a proof which does not depend on the Komlós theorem. To indicate the idea let us assume that there is only one function $f$ in $\mathcal{F}$ and that this function is bounded. Then it is possible to mimic the idea of the proof of the strong law of large numbers as it is presented in [1]. One chooses a subsequence $(T_{n_k})$ such that $(f \circ T_{n_k})$ is "nearly" an iid sequence of random variables. Using $\int g(f \circ T_n) \, \mathrm{d}P \longrightarrow \int fg \, \mathrm{d}P$ for bounded $f, g$ it can be achieved that $\int |f \circ T_{n_1} + \ldots f \circ T_{n_k}|^4 \, \mathrm{d}P \leqslant Ck^2$ for a suitable subsequence, and then $S_n(f, x) \longrightarrow \int f \, \mathrm{d}P$ follows by a Borel-Cantelli argument.

It is standard to pass from one $f$ to countably many $f$, however, it seems to be no way to treat the case of arbitrary $L^1$-functions in this way. □

**Corollary 4.3.** *Let $P$ be a Borel probability measure on a compact metric space $(M, d)$ and $(T_n)$ a strongly mixing sequence. Then there is a subsequence $(T_{n_k})_k$ such that $(S_n(f, x))$ tends to $\int f \, \mathrm{d}P$ for almost all $x$ and every continuous function $f$.*

Proof. One only has to combine the preceding theorem with the remark that $CM$ is separable with respect to the supremum norm and the observation that uniform limits of functions $f$ with "$(S_n(f, x)$ tends to $\int f \, \mathrm{d}P$ for almost all $x$" also have this property. □

## 5. Further discussions

In the previous section we have seen that by passing to a subsequence the Birkhoff theorem holds for continuous functions. A natural question is whether this can be generalized to all of $L^1$ as in the classical ergodic theorem. Unfortunately we cannot fill up $L^1$ with the closure of a countable set in the supremum norm and push our arguments through. Moreover, in a series of papers Bourgain and others have shown that if $\inf m_{n+1}/m_n > 1$ then for any ergodic transformation (and in particular for any mixing transformation) $S$ there is a function $f \in L^1 \cap L^\infty$ such that for $T_n = S^{m_n}$ the sequence $(S_n(f, x))_n$ does *not* converge almost everywhere; see [2] or [4]. So the question arises to what extent the class of continuous functions can be enlarged that the individual ergodic theorem holds within this class.

First we want to improve the statement of Lemma 4.1 for special functions.

**Proposition 5.1.** *Let $(\Omega, \mathcal{A}, P)$ be a Polish space and $(T_n)_n$ a mixing sequence of transformations. Let $U$ be an open set, $\chi_U$ its indicator function and $(T_{n_k})_k$ the subsequence from Corollary 4.3. Then*

$$\liminf \frac{1}{K} \sum_{k=1}^{K} \chi_U(T_{n_k} x) = P(U)$$

*for $P$, a.e. $x \in \Omega$.*

Proof. For given $\varepsilon > 0$ we always can find a continuous function $f$ with $f\big|_{\Omega\setminus U} = 0$, $0 \leqslant f \leqslant \chi_U$ and $P(U) - \int f \,\mathrm{d}P < \varepsilon$. In view of Corollary 4.3 we have $\lim \frac{1}{K} \sum\limits_{k=1}^{K} f(T_{n_k} x) = \int f \,\mathrm{d}P$. Now Lemma 4.1 implies the statement since $\varepsilon$ was arbitrary. □

Similarly we have

**Proposition 5.2.** *Let $(\Omega, \mathcal{A}, P)$ be a Polish space and $(T_n)_n$ a mixing sequence of transformations. Let $F$ be a closed set, $\chi_F$ its indicator function and $(n_k)_k$ the subsequence from Corollary 4.3. Then for $P$, a.e. $x \in \Omega$*

$$\limsup \frac{1}{K} \sum_{k=1}^{K} \chi_F(T_{n_k} x) = P(F).$$

In the remainder of this section we want to stress that one can enlarge the set of functions for which there is a universal subsequence such that the Cesàro means along this subsequence converge for all functions in this set to the integral. However, this subsequence becomes sparser as we extend the set of functions.

Let $\mathcal{C} = \{C_1, C_2, \ldots\}$ be a countable collection of measurable sets and $(p_k)$ an increasing sequence of natural numbers. We denote by $\mathcal{V} = \mathcal{V}(\mathcal{C}, (p_k))$ the following class of sets. $V \in \mathcal{V}$ iff there is a sequence of sets $A_1, A_2, \ldots$ with the properties

- $A_k \in \{C_1, C_2, \ldots, C_{p_k}\} := \mathcal{C}_k$,
- $A_k \subset A_{k-1}$,
- $V = \bigcap A_k$,
- $P(A_k) < P(V)(1 + 1/k^2)$.

A set $V \in \mathcal{V}$ is called $(\mathcal{C}, (p_k))$-approximable.

We can prove the following theorem.

**Theorem 5.3.** *Let $\mathcal{C}$ and $(p_k)$ be given as in the previous definition. Let $(T_n)_n$ be strongly mixing. Then there is a subsequence $(n_k)_k$ of natural numbers such that for $V \in \mathcal{V}(\mathcal{C},(p_k))$ we have*

$$\lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \chi_V(T_{n_k}x) = P(V)$$

*for almost all $x \in \Omega$.*

Before proving this theorem we recall the following result of Philipp (see [5], p. 66):

**Theorem 5.4** (Philipp, Schweiger)**.** *Let $(\Omega, \mathcal{A}, P)$ be a probability space and $(E_n)_n$ a sequence of measurable sets. For $x \in \Omega$ we define $A(N,x) := \sum_{n\leqslant N} \chi_{E_n}(x)$ and $\varphi(N) := \sum_{n\leqslant N} P(E_n)$. Suppose there exists a convergent series $\sum \alpha_k$ with $\alpha_k \geqslant 0$ such that*

$$P(E_n \cap E_{n+m}) \leqslant P(E_n)P(E_{n+m}) + \big(P(E_n) + P(E_{n+m})\big)\alpha_m + P(E_{n+m})\alpha_n.$$

*Then for $\varepsilon > 0$ and for almost all $x$*

$$A(N;x) = \varphi(N) + O\big(\varphi^{\frac{1}{2}}(N)\log^{\frac{3}{2}+\varepsilon}\varphi(N)\big)$$

*holds true.*

This theorem has the following immediate corollary:

**Corollary 5.5.** *Let $(T_n)$ be a sequence of measure-preserving transformations and $A$ a measurable set. If there is a summable sequence $\alpha_m$ with*

$$P(T_n^{-1}A \cap T_{n+m}^{-1}A) \leqslant P(A)^2 + \alpha_m + \alpha_n$$

*then*

$$\lim_{K\to\infty} \sum_{k=1}^{K} \chi_A(T_{n_k}x) = P(A)$$

*for almost all $x \in \Omega$.*

P r o o f of Theorem 5.3. Let $\mathcal{C}$, $(p_k)$ be given. We construct the subsequence $(n_k)$ inductively. First we are going to construct a subsequence $\big(n_k^{(1)}\big)$ such that the assumptions of Corollary 5.5 hold for the sequence $(T_{n_k^{(1)}})_k$ and all sets $C_1, \ldots, C_{p_1}$.

Let $0 < \gamma < \frac{1}{2}$. In view of the strong mixing property we first choose a subsequence $\big(m_k^{(1)}\big)$ of naturals such that $P\big(A \cap T^{-1}_{m_k^{(1)}} A\big) \leqslant P(A)^2 + \gamma^k$ for all $A \in \mathcal{C}_1$. Now we choose a subsequence $\big(m_k^{(2)}\big) \subset \big(m_k^{(1)}\big)$ with

$$P\Big(T^{-1}_{m_1^{(1)}} A \cap T^{-1}_{m_k^{(2)}} A\Big) \leqslant P(A)^2 + \gamma^{k+1}$$

for all $A \in \mathcal{C}_1$. We continue this process by induction. Finally we set $n_k^{(1)} = m_k^{(k)}$. By the properties of the construction we have

$$P\Big(T^{-1}_{n_k^{(1)}} A \cap T^{-1}_{n_{k+m}^{(1)}} A\Big) \leqslant P(A)^2 + \gamma^{k+m} \tag{5.1}$$

for all $A \in \mathcal{C}_1$ and $m > 0$.

Let us inductively choose subsequences $\big(n_k^{(l)}\big) \subset \big(n_k^{(l-1)}\big)$ such that (5.1) holds for the sequence $\big(n_k^{(l)}\big)$ and all sets $A \in \mathcal{C}_l$. Again applying the diagonalization procedure we produce a sequence $(n_k) := \big(n_k^{(k)}\big)$ with the property that

$$P(T^{-1}_{n_k} A \cap T^{-1}_{n_{k+m}} A) \leqslant P(A)^2 + \gamma^{k+m}$$

provided $A \in \mathcal{C}_l$, $k \geqslant l$, $m > 0$.

Let $V \in \mathcal{V}(\mathcal{C}, (p_k))$. Choose $A_k \in \mathcal{C}_k$ such that $P(A_k) \leqslant (1 + 1/k^2)P(V)$ and $V \subset A_k$. Hence,

$$P(T^{-1}_{n_k} V \cap T^{-1}_{n_{k+m}} V) \leqslant P(T^{-1}_{n_k} A_k \cap T^{-1}_{n_{k+m}} A_k) \tag{5.3}$$

$$\leqslant P(A_k)^2 + \gamma^{k+m} \tag{5.4}$$

$$\leqslant \Big(1 + \frac{1}{k^2}\Big)^2 P(V)^2 + \gamma^{k+m} \tag{5.5}$$

for $m > 0$. If we set $\alpha_m = 3/m^2 > \gamma^{k+m}$ we have $\sum \alpha_m < \infty$ and the assumptions of Corollary 5.5 are fulfilled. Hence, the assertions follow immediately. □

**Corollary 5.6.** *Let $\mathcal{V}$, $(p_k)$, $T_{n_k}$ be as in Theorem* 5.3 *and let $f$ be a finite linear combination of characteristic functions of sets from $\mathcal{V}$ (or even a uniform limit of such functions). Then the sums $1/K \sum_{k=1}^{K} f(T_{n_k})$ tend almost everywhere to $\int f \, \mathrm{d}P$.*

Here we have an interesting phenomenon: It is not to be expected that the result holds for *every* $L^1$-function; this fails already for arbitrary lacunary subsequences (see [4]). On the other hand, the hierarchy of sets $\mathcal{V}$ grows with the growth rate of $(p_n)$. The above theorem tells us that we can find sparser and sparser subsequences to have an ergodic theorem for the classes $\mathcal{V}$. It seems that sparse sequences are at the same time contraproductive and helpful.

## 6. Other subsequence averaging

We can also raise the question what happens if we take the entire sequence $(T_n)$ but evaluate the Cesàro mean at special places only. There is no way of getting an almost sure convergence result in general. Here we have the following example which is a modification of the example above.

**Example.** Let $\Sigma_2 := \{0,1\}^{\mathbb{N}}$ be the space of all one-sided infinite sequences of 0's and 1's endowed with the Tichonov product topology generating the Borel sets. Let $\sigma$ be the shift to the left and $P$ the Bernoulli measure assigning to the symbols 0 and 1 equal probability 1/2. We set, for $m \in \mathbb{N}$, $r(m) := \sup\{r\colon r! < m\}$. The strongly mixing sequence of transformations is defined as $T_n := \sigma^{r(n)}$ (note that $\sigma$ itself is mixing). For $\underline{x} = x_1x_2\ldots \in \Sigma_2$ we define $f\colon \Sigma_2 \to \mathbb{R}$ via $f(\underline{x}) = x_1$. We claim that no subsequence of $(S_n(f,\underline{x}))$ converges almost everywhere. Let us assume the contrary, i.e. there is a sequence $N_k$ such that $(S_{N_k}(f,\underline{x}))_k$ converges almost everywhere to a function $g$. Let $\frac{1}{32} > \varepsilon > 0$ be fixed. Then we can find a set $\Gamma \in \Sigma_2$ with $P(\Gamma) > 1-\varepsilon$ and a number $k_0 = k_0(\varepsilon)$ such that $|S_{N_k}(f,\underline{x}) - S_{N_l}(f,\underline{x})| < \varepsilon$ for all $k,l > k_0$. By the definition of $(T_n)$ we have for all sufficiently large $n$ that $S_{n!}(f,\underline{x}) < \varepsilon$ provided $x_n = 0$ or $S_{n!}(f,\underline{x}) > 1-\varepsilon$ provided $x_n = 1$, respectively. Hence for sufficiently large $k$ we have $S_{N_k}(f,\underline{x}) < \varepsilon$ whenever $x_{r(N_k)} = x_{r(N_k)+1} = 0$ or $S_{N_k}(f,\underline{x}) > 1-\varepsilon$ whenever $x_{r(N_k)} = x_{r(N_k)+1} = 1$, respectively. We consider sufficiently large numbers $k$, $l$ with $k_0 < k < l$ such that $r(N_k)+1 < r(N_l)$. Then the set $B := \{\underline{x}\colon x_{r(N_k)} = x_{r(N_k)+1} = 0 \text{ and } x_{r(N_l)} = x_{r(N_l)+1} = 1\}$ consists of points $\underline{x}$ with $|S_{N_k}(f,\underline{x}) - S_{N_l}(f,\underline{x})| > 1-2\varepsilon > \varepsilon$ implying $B \cap \Gamma = \emptyset$. On the other hand, $P(B) = \frac{1}{16}$ and hence $P(B\cap\Gamma) > \frac{1}{16} - \varepsilon > 0$, a contradiction.

### *References*


[1] *P. Billingsley*: Probability and Measure. John Wiley & Sons, New York, 1995.

[2] *J. Bourgain*: Almost sure convergence and bounded entropy. Israel J. Math. *63* (1988), 79–97.

[3] *J. Komlós*: A generalization of a problem of Steinhaus. Acta Math. Acad. Sci. Hungar *18* (1967), 217–229.

[4] *J. M. Rosenblatt and M. Wierdl*: Pointwise ergodic theorems via harmonic analysis. Ergodic theory and its connections with harmonic analysis (K. M. Petersen and I. A. Salama, eds.). London Math. Soc. Lecture Note Series 205, Cambridge Univ. Press, 1995.

[5] *F. Schweiger*: Ergodic theory of fibred systems and metric number theory. Oxford Science Publications, 1995.

[6] *P. Walters*: An Introduction to Ergodic Theory. Springer, 1982.



*Author's address*: I. Mathematisches Institut, Freie Universität Berlin, Arnimallee 2–6, D-14 195 Berlin, Germany, e-mails: `behrends@math.fu-berlin.de; shmeling@math.fu-berlin.de`.