C O L L O Q U I U M   M A T H E M A T I C U M

VOL. 84/85 | 2000 | PART 2

REMARKS ON THE TIGHTNESS OF COCYCLES

BY

JON AARONSON (TEL AVIV)
AND BENJAMIN WEISS (JERUSALEM)

Dedicated to the memory of Anzelm Iwanik

Abstract. We prove a generalised tightness theorem for cocycles over an ergodic probability preserving transformation with values in Polish topological groups. We also show that subsequence tightness of cocycles over a mixing probability preserving transformation implies tightness. An example shows that this latter result may fail for cocycles over a mildly mixing probability preserving transformation.

Let $(\Omega,\mathcal{B},m)$ be a probability space, let $T:\Omega\to\Omega$ be an ergodic probability preserving transformation, let G be a Polish topological group and let $\phi:\Omega\to G$ be measurable.

We consider S_n, the *random walk* or *cocycle* on G defined by

$$S_0(\omega)=e,\qquad S_{n+1}(\omega):=\phi(T^n\omega)S_n(\omega).$$

This random walk is generated by the *skew product* transformation $T_\phi: X\times G\to X\times G$ where $T_\phi^n(\omega,y)=(T^n\omega,S_n(\omega)y)$. In case G is a locally compact topological group, T_ϕ preserves the measure $m\times m_G$ where m_G is a left Haar measure on G.

1. Tightness theorem. We consider the situation where $\{m\text{-dist.}(S_n): n\ge 1\}$ is *tight* in the sense that for every $\varepsilon>0$, there is a compact $C\subset G$ such that $\sup_{n\ge1} m(S_n\notin C)<\varepsilon$ (equivalently, tightness is precompactness in the space $\mathcal{P}(G)$ of probability measures on G). One way this can happen is when ϕ is cohomologous to a compact-group-valued function, i.e. there is a compact subgroup $K\subseteq G$ and measurable $\psi:\Omega\to K$, $g:\Omega\to G$ such that $\phi(\omega)=g(T\omega)^{-1}\psi(\omega)g(\omega)$; then $S_n(\omega)=g(T^n\omega)^{-1}k_n(\omega)g(\omega)$ where $k_n(\omega):=\psi(T^{n-1}\omega)\psi(T^{n-2}\omega)\ldots\psi(\omega)\in K$.

TIGHTNESS THEOREM. *The distributions* $\{m\text{-dist.}(S_n): n\ge 1\}$ *are tight in* $\mathcal{P}(G)\Leftrightarrow\phi$ *is cohomologous to a compact-group-valued function.*

2000 *Mathematics Subject Classification*: Primary 28D05.

Remarks about $\Leftarrow$. 1) The $\Leftarrow$ of the tightness theorem is an easy consequence of the tightness of a single probability on a Polish space (Prokhorov's theorem, see [Par]) and the probability preserving property of T.

2) If m is not absolutely continuous with respect to some T-invariant probability on $(\Omega, \mathcal{B})$ then $\Leftarrow$ may fail.

In this case, there is a set $W \in \mathcal{B}$ with $m(W) > 0$ and a sequence $n_k \to \infty$ such that $\{T^{-n_k}W : k \geq 1\}$ are disjoint (such a set is called *weakly wandering*). Given a noncompact Polish space G, we choose $x_0 \in G$ and a sequence $y_k \in G$, $y_k \to \infty$ (i.e. for each compact $C \subset G$, $y_k \notin C$ eventually) and define $f : \Omega \to G$ by

$$f(x) = \begin{cases} y_k, & x \in T^{-n_k}W \ (k \geq 1), \\ x_0, & x \in \Omega \setminus \bigcup_{k=1}^{\infty} T^{-n_k}W. \end{cases}$$

It follows that $\{m\text{-dist.}(f \circ T^n) : n \geq 1\}$ cannot be tight in $\mathcal{P}(G)$ since $m([f \circ T^{n_k} = y_k]) \geq m(W) \nrightarrow 0$.

If G is a noncompact Polish topological group, we set $\phi = f^{-1} f \circ T$ and obtain a coboundary for which the distributions $\{m\text{-dist.}(S_n) : n \geq 1\}$ are not tight in $\mathcal{P}(G)$.

In case G has no nontrivial compact subgroups, the tightness theorem boils down to the so-called *coboundary theorem*:

The distributions $\{m\text{-dist.}(S_n) : n \geq 1\}$ *are tight in* $\mathcal{P}(G) \Leftrightarrow \phi$ *is a coboundary.*

The first version of the coboundary theorem seems to be:

L^2 Coboundary Theorem [Leo]. *If* $\{Z_n : n \geq 1\}$ *is a wide sense stationary process, then there exists a wide sense stationary process* $\{Y_n : n \geq 1\}$ *such that* $Z_n = Y_n - Y_{n+1}$ *iff* $\sup_{n \geq 1} \mathbb{E}(|\sum_{k=1}^{n} Z_k|^2) < \infty$.

Proof. If there is $\{Y_n : n \geq 1\}$ wide sense stationary such that $Z_n = Y_n - Y_{n+1}$, then $\sum_{k=1}^{n} Z_k = Y_1 - Y_{n+1}$ and $\|\sum_{k=1}^{n} Z_k\|_2 \leq 2\|Y_1\|_2$ for all $n \geq 1$.

Conversely, if $\|\sum_{k=1}^{n} Z_k\|_2 \leq M$ for all $n \geq 1$, then by weak* sequential compactness of norm bounded sets, there are $N_a \to \infty$ and a r.v. $Y = Y(Z_1, Z_2, \ldots)$ such that

$$\frac{1}{N_a} \sum_{n=1}^{N_a} \sum_{k=1}^{n} Z_k \rightharpoonup Y$$

where $\rightharpoonup$ denotes weak convergence in L^2. Write $Y_n := Y(Z_n, Z_{n+1}, \ldots)$. Then $\{Y_n : n \geq 1\}$ is a wide sense stationary process and

$$\frac{1}{N_a} \sum_{n=1}^{N_a} \sum_{k=1}^{n} Z_{k+\nu-1} \rightharpoonup Y_\nu \quad \forall \nu \geq 1.$$

It follows that

$$\begin{aligned} Y_{\nu+1} &\leftharpoonup \frac{1}{N_a}\sum_{n=1}^{N_a}\sum_{k=\nu+1}^{n+\nu} Z_k = \frac{1}{N_a}\sum_{n=1}^{N_a}\Big(\sum_{k=\nu}^{n+\nu-1} Z_k + Z_{n+\nu} - Z_\nu\Big) \\ &= \frac{1}{N_a}\sum_{n=1}^{N_a}\sum_{k=1}^{n} Z_{k+\nu-1} + \frac{1}{N_a}\sum_{n=1}^{N_a} Z_{n+\nu} - Z_\nu \rightharpoonup Y_\nu - Z_\nu \end{aligned}$$

because $\|\sum_{n=1}^{N_a} Z_{n+\nu}\|$ is uniformly bounded. ■

Leonov's theorem has the L^p analogues:

L^p Coboundary Theorem. *Let $(X,\mathcal{B},m,T)$ be a probability preserving transformation, let $1 \le p < \infty$ and let $f : X \to \mathbb{R}$ be measurable. There exists $g \in L^1(m)$ such that $f = g - g\circ T$ iff $\sup_{n\ge1}\|\sum_{k=1}^n f\circ T^k\|_p < \infty$.*

The proof of the L^p coboundary theorem is the same as that of Leonov with Komlos type convergence replacing weak convergence when $p = 1$.

The coboundary theorem is established in [Sch1] for the case $G = \mathbb{R}$, and in [Mo-Sch] for G locally compact, second countable, Abelian without compact subgroups.

The tightness theorem for locally compact, second countable groups was established in [Sch2]; related partial results are given in [Co] and [Zim].

Bradley has proved $\Rightarrow$ of the coboundary theorem assuming only that T is measurable: in [Br1] for $G = \mathbb{R}$, in [Br2] for G a Banach space and in [Br3] for G a group of upper triangular matrices.

The present methods can be stretched to prove the $\Rightarrow$ of the tightness theorem assuming only that T is measurable and invertible.

Basic Lemma. *If the family $\{P\text{-dist.}(S_n) : n \ge 1\}$ is tight in $\mathcal{P}(G)$, then there is a measurable $P : \Omega \to \mathcal{P}(G)$ such that*

$$P_{T\omega}(A) = P_\omega(\phi(\omega)^{-1}A) \quad (A \in \mathcal{B}(G)).$$

This basic lemma is implicit in [Br1] for $G = \mathbb{R}$. The general proof is essentially as in [Br1] (see below).

The coboundary theorem for $\mathbb{R}$ is easily established using it ([Br1]). Indeed if for $\omega \in \Omega$, $\mu(\omega)$ is defined as the minimal number satisfying

$$P_\omega((-\infty,\mu(\omega)]), P_\omega([\mu(\omega),\infty)) \ge 1/2,$$

then $\mu : \Omega \to \mathbb{R}$ is measurable and (since $P_{T\omega}(A) = P_\omega(A - \phi(\omega))$) we have $\mu(T\omega) = \mu(\omega) - \phi(\omega)$.

The proof of the tightness theorem given the basic lemma uses a generalisation of the characterisation of invariant measures for group extensions in [Key-New]. The proof is an adaptation of Lemańczyk's proof of [Key-New] in [Lem]. See also the proof of Theorem 8.3.2 in [A].

Proof of the basic lemma. Choose first $K_\nu \subset K_{\nu+1} \subset \ldots \subset G$, a sequence of compact sets in G with the property (ensured by tightness) that

$$m([S_n \in K_\nu^{\mathrm{c}}]) \le 1/4^\nu \quad \forall n, \nu \ge 1. \tag{1}$$

Consider the random measures $W_n : \Omega \to \mathcal{P}(G)$ defined by

$$W_n(A) := \frac{1}{n}\sum_{j=1}^{n} 1_A(S_j).$$

Next, for $\nu \ge 1$ let $\mathcal{A}_\nu \subset C(K_\nu)$ be a countable family, dense in $C(K_\nu)$; and let $\mathcal{A} = \bigcup_{\nu=1}^\infty \mathcal{A}_\nu$.

We now claim that there are $n_k \to \infty$ and $L : \mathcal{A} \to L^\infty(\Omega)$ such that

$$\int_G f\, dW_{n_k} \to L(f) \quad \text{weak* in } L^\infty(\Omega)\ \forall f \in \mathcal{A}. \tag{2}$$

This is shown using weak* precompactness of $L^\infty(\Omega)$-bounded sets, and a diagonalisation.

By possibly passing to a subsequence, we can ensure that for each $f \in \mathcal{A}$, there is N_f such that

$$\Big|\int_X \Big(\Big(\int_G f\, dW_{n_k} - L(f)\Big)\Big(\int_G f\, dW_{n_j} - L(f)\Big)\Big)\, dm\Big| < \frac{1}{2^k} \quad \forall k \ge N_f,\ j < k,$$

whence ([Rev])

$$\frac{1}{N}\sum_{k=1}^{N}\int_G f\, dW_{n_k} \to L(f) \quad \text{a.e. } \forall f \in \mathcal{A} \tag{3}$$

and hence (by density) for all $f \in \bigcup_{\nu=1}^\infty C(K_\nu)$.

By the Chebyshev–Markov inequality,

$$\begin{aligned} m(L(1_{K_\nu^{\mathrm{c}}}) > 1/2^\nu) \leftarrow m(W_{n_k}(K_\nu^{\mathrm{c}}) > 1/2^\nu) &< 2^\nu \int_X W_{n_k}(K_\nu^{\mathrm{c}}) dm \\ &< 1/2^\nu \quad \forall \nu \ge 1 \end{aligned}$$

and so by the Borel–Cantelli lemma, $L(1_{K_\nu^{\mathrm{c}}}) \le 1/2^\nu$ a.e. for ν large.

It follows that there is a measurable $P : \Omega \to \mathcal{P}(G)$ such that $L(f)(\omega) = \int_G f\, dP_\omega$ for all $f \in \mathcal{A}$.

To see that $P_{T\omega} = P_\omega \circ R_{\phi(\omega)}$ ($R_g(y) := yg$), note that

$$\begin{aligned} \int_G f\, dW_n(T\omega) &= \frac{1}{n}\sum_{j=1}^{n} f(S_j(T\omega)) \\ &= \frac{1}{n}\sum_{j=1}^{n} f(S_{j+1}(\omega)\phi(\omega)^{-1}) = \frac{1}{n}\sum_{j=2}^{n+1} f \circ R_{\phi(\omega)^{-1}}(S_j(\omega)) \end{aligned}$$

$$= \int_G f \circ R_{\phi(\omega)^{-1}}\, dW_n(\omega) \pm \frac{2\|f\|_\infty}{n}$$

$$= \int_G f\, dW_n(\omega) \circ R_{\phi(\omega)} \pm \frac{2\|f\|_\infty}{n}. \ \blacksquare$$

Proof of $\Rightarrow$ *in the tightness theorem.* Given probabilities $\omega \mapsto p_\omega$ on G satisfying

$$p_{T\omega} = p_\omega \circ L_{\phi(\omega)^{-1}},$$

define a probability $\mu \in \mathcal{P}(\Omega \times G)$ by

$$\mu(A \times B) := \int_A p_\omega(B)\, dm(\omega).$$

We first note that this probability is T_ϕ-invariant:

$$\begin{aligned}\int_{X\times G} (u \otimes v) \circ T_\phi\, d\mu &= \int_X u(Tx) \int_G v(\phi(x)y)\, dp_x(y)\, dm(x)\\ &= \int_X u(Tx) \int_G v(y)\, dp_{Tx}(y)\, dm(x)\\ &= \int_X u(x) \int_G v(y)\, dp_x(y)\, dm(x) = \int_{X\times G} u \otimes v\, d\mu.\end{aligned}$$

Almost every ergodic component P of μ has a disintegration over m of the form

$$P(A \times B) := \int_A \widetilde{p}_\omega(B)\, dm(\omega)$$

where $\omega \mapsto \widetilde{p}_\omega \in \mathcal{P}(G)$ is measurable, and $\widetilde{p}_{T\omega} = \widetilde{p}_\omega \circ R_{\phi(\omega)}$. Fix one such P.

Define $p \in \mathcal{P}(G)$ by $p(B) := P(\Omega \times B)$. There are compact sets $C_1 \subset C_2 \subset \ldots$ such that $\bigcup_{n=1}^\infty C_n = G \bmod p$. Define compact subsets $\{K_n : n \geq 0\}$ by

$$K_0 := \{e\}, \qquad K_{n+1} = (K_n \cup C_n)(K_n \cup C_n)^{-1}(K_n \cup C_n)(K_n \cup C_n)^{-1}.$$

Evidently, $G_0 := \bigcup_{n=1}^\infty K_n$ is a subgroup of G and $p(G \setminus G_0) = 0$, whence $\widetilde{p}_\omega(G \setminus G_0) = 0$ for m-a.e. $\omega \in \Omega$.

Next, consider the space $C_{\mathrm{B}}(G_0)$ of bounded, continuous, $\mathbb{R}$-valued functions on G_0 (equipped with the supremum norm) and set

$$\mathcal{C} := \{f \in C_{\mathrm{B}}(G_0) : \sup_{y \in K_n^{\mathrm{c}}} |f(y)| \underset{n\to\infty}{\longrightarrow} 0\}.$$

Evidently $\mathcal{C} = \overline{\bigcup_{n=1}^\infty C_{\mathrm{B}}(K_n)}$ is separable, and $f \in \mathcal{C} \Rightarrow f \circ R_g \in \mathcal{C}$ for all $g \in G_0$ (since if $g \in K_i$, then $x \notin K_{n+i} \Rightarrow\ xg \notin K_n$).

For each $a \in G$, $P \circ Q_a$ (where $Q_a(\omega, y) := (\omega, ya)$) is also an ergodic T_ϕ-invariant probability (since $T_\phi \circ Q_a = Q_a \circ T_\phi$), and therefore either

$P \circ Q_a = P$ or $P \circ Q_a \perp P$. Define $H := \{a \in G_0 : P \circ Q_a = P\}$, a closed subgroup of G_0. For a.e. $\omega \in \Omega$, $p_\omega(Aa) = p_\omega(A)$ $(a \in H,\ A \in \mathcal{B}(G))$.

Consider the Banach space $\mathcal{M}(\Omega \times G_0)$ of bounded measurable functions $\Omega \times G_0 \to \mathbb{R}$ equipped with the supremum norm. We need a separable subspace $\mathcal{A} \subset \mathcal{M}(\Omega \times G_0)$ which separates the points of $\Omega \times G_0$ such that $f \in \mathcal{A} \Rightarrow f \circ Q_a \in \mathcal{A}$ for all $a \in G_0$. In particular,

$$a, b \in G_0, \quad \int_{\Omega \times G} f\, dP \circ Q_a = \int_{\Omega \times G} f\, dP \circ Q_b \quad \forall f \in \mathcal{A} \Rightarrow P \circ Q_a = P \circ Q_b.$$

To obtain such a subspace, fix a compact metric topology on Ω generating $\mathcal{B}$; then $\mathcal{A} = C(\Omega) \otimes \mathcal{C}$ is as needed.

By Birkhoff's ergodic theorem,

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T_\phi^k(\omega, y) \to \int_{\Omega \times G} f\, dP \qquad \text{a.e. } \forall f \in L^1(P).$$

Set

$$Y := \left\{(\omega, y) \in \Omega \times G_0 : \frac{1}{n} \sum_{k=0}^{n-1} f \circ T_\phi^k(\omega, y) \to \int_{\Omega \times G} f\, dP \ \forall f \in \mathcal{A}\right\}.$$

Since $\mathcal{A}$ is a separable subspace of $\mathcal{M}(\Omega \times G_0)$, the set Y is determined by a countable subcollection of $\mathcal{A}$, whence $Y \in \mathcal{B}(\Omega \times G_0)$, and by Birkhoff's ergodic theorem $P(Y) = 1$. For $\omega \in \Omega$, set $Y_\omega = \{y \in G_0 : (\omega, y) \in Y\}$. We claim that Y_ω is a coset of H whenever it is nonempty.

To see this, suppose that $a \in G$. Then for all $f \in \mathcal{A}$ and for a.e. $(x, y) \in Y$,

$$\frac{1}{n} \sum_{k=0}^{n-1} f \circ T_\phi^k(\omega, ya) \to \int_{\Omega \times G} f \circ Q_a\, dP = \int_{\Omega \times G} f\, dP \circ Q_a^{-1}.$$

Thus, $(\omega, ya) \in Y$ iff $P \circ Q_a^{-1} = P$, equivalently $a \in H$; and Y_ω is indeed a coset of H whenever it is nonempty (i.e. a.e.).

By the analytic section theorem, there is a measurable $h : \Omega \to G$ such that $h(\omega) \in Y_\omega$ for a.e. $\omega \in \Omega$, whence $Y_\omega = h(\omega)H$.

Now let $P'_\omega \in \mathcal{P}(G)$ be defined by $P'_\omega(A) := p_\omega(h(\omega)^{-1}A)$. Clearly $P'_\omega(H) = 1$ and $P'_\omega(Aa) = P'_\omega(A)$ $(a \in H,\ A \in \mathcal{B}(G))$. Thus by [Weil], H is compact and $P'_\omega = m_H$, Haar measure on H.

Defining $\Psi : \Omega \times G \to \Omega \times G$ by $\Psi(\omega, y) := (\omega, h(\omega)y)$, we have $P \circ \Psi^{-1} = m \times m_H$. If $V := \Psi \circ T_\phi \circ \Psi^{-1}$ then $m \times m_H \circ V = m \times m_H$ and $V = T_\psi$ where $\psi(\omega) := h(\omega)\phi(\omega)h(\omega)^{-1}$.

Since $(\Omega \times G, \mathcal{B}(\Omega \times G), m \times m_H, V)$ is a probability preserving transformation, we see that $\psi : \Omega \to H$. ∎

2. Subsequence tightness. Let $(X,\mathcal{B},m,T)$ be a mixing probability preserving transformation and let $\phi : X \to \mathbb{R}$ be measurable. Bradley [Br4] showed that if the stochastic process $\{\phi\circ T^n : n \ge 1\}$ is strongly Rosenblatt mixing, then either

1) $\sup_{r\in\mathbb{R}} m([|S_n - r| \le C]) \to 0$ for every $0 < C < \infty$, or

2) there are constants a_n such that $\{m\text{-dist.}(S_n - a_n) : n \ge 1\}$ is tight (whence ϕ is cohomologous to a constant).

A weaker version of this generalises to an arbitrary stationary stochastic process driven by a mixing probability preserving transformation.

THEOREM 2. *Suppose that $(X,\mathcal{B},m,T)$ is a mixing probability preserving transformation and that $\phi : X \to \mathbb{R}$ is measurable. If there are $n_k \to \infty$ and $d_k \in \mathbb{R}$ such that $\{m\text{-dist.}(S_{n_k} - d_k) : k \ge 1\}$ is tight, then there are $a \in \mathbb{R}$ and $g : \Omega \to \mathbb{R}$ measurable such that $\phi(\omega) = a + g(T\omega) - g(\omega)$. If $\sup_k |d_k| < \infty$, then $a = 0$.*

Proof. Consider $(X\times X, \mathcal{B}\otimes\mathcal{B}, m\times m, T\times T)$, and $\phi,\phi' : X\times X \to \mathbb{R}$ defined by $\phi(x,y) := \phi(x)$, $\phi'(x,y) := \phi(y)$.

• We first show that $\{m\times m\text{-dist.}(S_n - S'_n) : n \ge 1\}$ is tight. Let $\varepsilon > 0$ and choose $M > 0$ such that $m([|S_{n_k} - d_k| > M/2]) < \varepsilon/2$ for all $k \ge 1$. By mixing of T, for all $n \ge 1$,

$$m([|S_n - S_n\circ T^{n_k}| > M]) \to m\times m([|S_n - S'_n| > M])$$

as $k \to \infty$. Now

$$S_n - S_n \circ T^{n_k} = S_n - S_{n+n_k} + S_{n_k} = S_{n_k} - S_{n_k}\circ T^n,$$

whence

$$\begin{aligned} m([|S_n - S_n\circ T^{n_k}| > M]) &= m([|S_{n_k} - S_{n_k}\circ T^n| > M]) \\ &\le 2m([|S_{n_k} - d_k| > M/2]) < \varepsilon. \end{aligned}$$

• Next, as in [Br4], there are $a_n \in \mathbb{R}$ such that $\{m\text{-dist.}(S_n - a_n) : n \ge 1\}$ is tight. To see this, given $\varepsilon > 0$, let $M(\varepsilon) > 0$ be such that

$$m\times m([|S_n - S'_n| > M(\varepsilon)]) < \varepsilon^2 \quad \forall n \ge 1.$$

It follows that

$$\begin{aligned} m(\{x\in X : m([|S_n - S_n(x)| > M(\varepsilon)]) > \varepsilon\}) \\ &\le \frac{1}{\varepsilon}\int_X m([|S_n - S_n(x)| > M(\varepsilon)])\,dm(x) \\ &= \frac{1}{\varepsilon} m\times m([|S_n - S'_n| > M(\varepsilon)]) < \varepsilon \quad \forall n \ge 1, \end{aligned}$$

whence there are $a_n(\varepsilon) \in \mathbb{R}$ such that

$$m([|S_n - a_n(\varepsilon)| > M(\varepsilon)]) \le \varepsilon \quad \forall n \ge 1.$$

Set $a_n = a_n(1/3)$. For each $0 < \varepsilon < 1/2$, $n \ge 1$, we have

$$m([|S_n - a_n(\varepsilon)| < M(\varepsilon)] \cap [|S_n - a_n| < M(1/3)]) > 0,$$

whence $|a_n - a_n(\varepsilon)| < M(1/3) + M(\varepsilon)$ and

$$m([|S_n - a_n| > 2M(\varepsilon) + M(1/3)]) < \varepsilon \quad \forall n \ge 1.$$

• We show that there is an $a \in \mathbb{R}$ such that $\sup_{n\ge1} |a_n - na| < \infty$. To this end, note that there is an $M > 0$ such that

$$(\ddagger) \qquad |a_{k+l} - a_k - a_l| < M \quad \forall k, l \ge 1.$$

Indeed, if $m([|S_n - a_n| > K]) < 1/8$ for all $n \ge 1$, then (since $S_{k+l} = S_k + S_l \circ T^k$)

$$\begin{aligned} &m([|S_{k+l} - a_k - a_l| > 2K]) \\ &\qquad \le m([|S_k - a_k| > K] \cup [|S_l \circ T^k - a_l| > K]) < 1/4, \end{aligned}$$

whence

$$m([|S_{k+l} - a_k - a_l| \le 2K] \cap [|S_{k+l} - a_{k+l}| \le K]) > 0$$

and $|a_{k+l} - a_k - a_l| \le 3K$ for $k, l \ge 1$.

By (‡), there are $N_k \to \infty$ and $b_\nu \in \mathbb{R}$ ($\nu \ge 1$) such that

$$\frac{1}{N_k}\sum_{j=1}^{N_k}(a_{j+\nu} - a_j) \to b_\nu \quad \text{as } k \to \infty\ \forall \nu \ge 1.$$

It follows from (‡) that

$$|b_\nu - a_\nu| = \lim_{k\to\infty}\left|\frac{1}{N_k}\sum_{j=1}^{N_k}(a_{j+\nu} - a_j - a_\nu)\right| \le M$$

and that

$$\begin{aligned} b_{\nu+\mu} &\leftarrow \frac{1}{N_k}\sum_{j=1}^{N_k}(a_{j+\mu+\nu} - a_j) \\ &= \frac{1}{N_k}\sum_{j=1}^{N_k}(a_{j+\mu} - a_j) + \frac{1}{N_k}\sum_{j=\mu+1}^{N_k+\mu}(a_{j+\nu} - a_j) \\ &= \frac{1}{N_k}\sum_{j=1}^{N_k}(a_{j+\mu} - a_j) + \frac{1}{N_k}\sum_{j=1}^{N_k}(a_{j+\mu} - a_j) \pm \frac{M + |a_\mu|}{N_k} \\ &\to b_\mu + b_\nu. \end{aligned}$$

Thus $b_\nu = \nu a$ and $|a_\nu - \nu a| \le M$ where $a = b_1 = \lim_{n\to\infty} a_n/n$.

In case $\sup_k |d_k| < \infty$, because of the tightness of $\{m\text{-dist.}(S_{n_k}) : k \geq 1\}$ we have $\sup_{k\geq 1} |a_{n_k}| < \infty$, whence $a = 0$.

• It now follows from the coboundary theorem that ϕ is cohomologous to a. ∎

3. An example. We show that there is a probability preserving transformation $(X, \mathcal{B}, m, T)$ which is *mildly mixing* in the sense that there is no $A \in \mathcal{B}$ with $0 < m(A) < 1$ such that $\liminf_{n\to\infty} m(A \triangle T^n A) = 0$ (see §2.7 of [A]), but there is a measurable function $\phi : X \to \mathbb{R}$ such that T_ϕ is ergodic and for some $n_k \to \infty$, $\limsup_{k\to\infty} |S_{n_k}| < \infty$ m-almost everywhere.

Chacon's transformation [Cha]. This transformation $(X, \mathcal{B}, m, T)$ is defined inductively on $X := \bigcup_{n=1}^\infty C_n \subset \mathbb{R}$ where $m =$ Lebesgue measure.

Here $C_n = \bigcup_{k=0}^{l_n-1} T^k J_n$ where

• $l_1 = 1,\ l_{n+1} = 3l_n + 1$ ($\Rightarrow l_n = (3^n - 1)/2$);

• $\{T^k J_n : 0 \leq k \leq l_n - 1\}$ are disjoint intervals of length $1/3^{n-1}$ and $T : T^k J_n \to T^{k+1} J_n$ is a translation;

• C_{n+1} is obtained by writing $J_n = \bigcup_{i=0}^2 J_{n,i}$ where the $J_{n,i}$ $(i = 0, 1, 2)$ are disjoint intervals of length $1/3^n$ and setting $J_{n+1} := J_{n,0}$ and

$$T^k J_{n+1} := \begin{cases} T^k J_{n,0}, & 0 \leq k \leq l_n - 1, \\ T^{k-l_n} J_{n,1}, & l_n \leq k \leq 2l_n - 1, \\ \mathcal{S}_{n+1}, & k = 2l_n, \\ T^{k-2l_n-1} J_{n,2}, & 2l_n + 1 \leq k \leq 3l_n = l_{n+1} - 1 \end{cases}$$

where $\mathcal{S}_{n+1}$ is an interval of length $1/3^n$, disjoint from C_n (called the *spacer*).

The set X has finite measure which can be normalized to equal one but we keep the standard Lebesgue measure in order to simplify the later formulae. We first give a proof of the ergodicity based on a careful analysis of how the intervals $T^k J_n$ approximate arbitrary measurable sets. This analysis will also be the base for our proof of the mild mixing property.

Define

$$\mathcal{C}_n := \Big\{U_n(K) := \bigcup_{k\in K} T^k J_n : K \subset \{0, 1, \ldots, l_n - 1\}\Big\}.$$

For $A \in \mathcal{B}$, $\varepsilon > 0$ and $n \geq 1$ define

$$K_{A,\varepsilon}^{(n)} := \{0 \leq k \leq l_n - 1 : m(T^k J_n \cap A) < \varepsilon m(J_n)\} \subset \{0, 1, \ldots, l_n - 1\}.$$

Evidently, for $A, B \in \mathcal{B}$ disjoint and $0 < \varepsilon < 1/2$, $K_{A,\varepsilon}^{(n)}$ and $K_{B,\varepsilon}^{(n)}$ are disjoint.

It is standard that for all $A \in \mathcal{B}$ and $\varepsilon > 0$, there is $N_{A,\varepsilon}$ such that

$$|E_A^{(n)}| < \varepsilon l_n \quad \forall n \geq N_{A,\varepsilon}$$

where

$$E_A^{(n)} := \{0, 1, \ldots, l_n - 1\} \setminus (K_{A,\varepsilon}^{(n)} \cup K_{A^c,\varepsilon}^{(n)}),$$

whence (for such n)

$$m(U_n(K^{(n)}_{A,\varepsilon}) \setminus A) = \sum_{k \in K^{(n)}_{A,\varepsilon}} m(T^k J_n \setminus A) < \varepsilon m(C_n)$$

and

$$\begin{aligned} m(A \setminus U_n(K^{(n)}_{A,\varepsilon})) &= m(A \cap U_n(K^{(n)}_{A^c,\varepsilon})) + m(A \cap U_n(E^{(n)}_A)) \\ &\le \sum_{k \in K^{(n)}_{A^c,\varepsilon}} m(T^k J_n \setminus A) + \varepsilon m(C_n) < 2\varepsilon m(C_n) \end{aligned}$$

and $m(A \triangle U_n(K^{(n)}_{A,\varepsilon})) < 3\varepsilon m(C_n)$. Henceforth, we let $n_{A,\varepsilon}$ be the minimal N with $|E^{(n)}_A| < \varepsilon l_n$ for all $n \ge N$.

Conversely, suppose that $A \in \mathcal{B}$ and $U = U_n(K) \in \mathcal{C}_n$ satisfy $m(A \triangle U) < \varepsilon m(U)$. Then

$$\begin{aligned} &\sum_{k \in K,\, m(T^k J_n \setminus A) \ge \sqrt{\varepsilon}\, m(J_n)} m(T^k J_n) \\ &\qquad \le \frac{1}{\sqrt{\varepsilon}} \sum_{k \in K,\, m(T^k J_n \setminus A) \ge \sqrt{\varepsilon}\, m(J_n)} m(T^k J_n \setminus A) \\ &\qquad \le \frac{1}{\sqrt{\varepsilon}} m(U \setminus A) < \sqrt{\varepsilon} \end{aligned}$$

and

$$\begin{aligned} &\sum_{k \in K^c,\, m(T^k J_n \setminus A^c) \ge \sqrt{\varepsilon} m(J_n)} m(T^k J_n) \\ &\qquad \le \frac{1}{\sqrt{\varepsilon}} \sum_{k \in K^c,\, m(T^k J_n \setminus A^c) \ge \sqrt{\varepsilon}\, m(J_n)} m(T^k J_n \setminus A^c) \\ &\qquad \le \frac{1}{\sqrt{\varepsilon}} m(A \setminus U) < \sqrt{\varepsilon} \end{aligned}$$

whence

$$|K \setminus K^{(n)}_{A,\varepsilon}|, |K^c \setminus K^{(n)}_{A^c,\varepsilon}| \le \sqrt{\varepsilon}\, l_n$$

and $n \ge n_{A,2\sqrt{\varepsilon}}$.

To see (the well known fact [Fr]) that $(X, \mathcal{B}, m, T)$ is an ergodic measure preserving transformation, let $A \in \mathcal{B}$ with $m(A) > 0$ satisfy $TA = A$. Evidently, $K^{(n)}_A \ne \emptyset \Rightarrow K^{(n)}_A = \{0, 1, \ldots, l_n - 1\}$, whence $U_n(K^{(n)}_{A,\varepsilon}) = C_n$.

It follows that $m(A) > m(C_n)(1 - 3\varepsilon)$ for all $\varepsilon > 0$ and $n \ge n_{A,\varepsilon}$, whence $A = X \bmod m$.

In [Cha] it was shown that Chacon's transformation $(X, \mathcal{B}, m, T)$ is weakly mixing and not strongly mixing. We next claim that it is *mildly mixing*. For a related result, see [F-K].

To see this, we first need some notation to record how sets in $\mathcal{C}_n$ appear in C_{n+2}. Define e_j $(0 \le j \le 7)$ by

$$e_j := \begin{cases} 0, & j = 0, 2, 3, 6, \\ 1, & j = 1, 4, 5, 7, \end{cases}$$

$\kappa_j = \kappa_{j,n}$ by

$$\kappa_0 = 0, \qquad \kappa_{j+1} := \kappa_j + l_n + e_j$$

and

$$X_j = X_{j,n} := \bigcup_{i=0}^{l_n - 1} T^{i+\kappa_{j,n}} J_{n+2} \qquad (0 \le j \le 8).$$

Then given $n \ge 1$, $K \subset \{0, 1, \ldots, l_n - 1\}$ and $U = U_n(K) \in \mathcal{C}_n$, we have

$$T^{\kappa_{j,n}}(U \cap X_0) = \bigcup_{i \in K} T^{i+\kappa_{j,n}} J_{n+2} = U \cap X_j \qquad (0 \le j \le 7)$$

and

$$T^{l_n + e_j}(U \cap X_j) = U \cap X_{j+1}.$$

Next suppose that $A \in \mathcal{B}$, $\varepsilon > 0$ and $n \ge n_{A,\varepsilon}$. Then

$$m(T^{i+\kappa_{j,n}} J_{n+2} \cap A) < 9\varepsilon m(J_{n+2}) \qquad \forall i \in K^{(n)}_{A^c,\varepsilon},\ 0 \le j \le 8$$

and

$$m(T^{i+\kappa_{j,n}} J_{n+2} \setminus A) < 9\varepsilon m(J_{n+2}) \qquad \forall i \in K^{(n)}_{A,\varepsilon},\ 0 \le j \le 8;$$

whence

$$m(T^{\kappa_{j,n}}(A \cap X_0) \,\triangle\, (A \cap X_j)) < 36\varepsilon.$$

Now suppose $A \in \mathcal{B}$ with $m(A) > 0$ satisfies $\liminf_{n\to\infty} m(A \triangle T^n A) = 0$. We claim that $A = T^{-1}A$.

To see this, fix $\varepsilon > 0$. Then there are $n \ge n_{A,\varepsilon}$ and $N \in [l_n, l_{n+1} - 1]$ such that $m(A \,\triangle\, T^N A) < \varepsilon$, whence there is $B \in \mathcal{C}_n$ such that $m(B \,\triangle\, T^N B) < 3\varepsilon$. Write $N = a l_n + b$ where $a = 1, 2$ and $0 \le b \le l_n$. For $0 \le j \le 6 - a$ we have

$$T^N X_j = T^{a l_n + b} X_j = T^{b - e_{j,a}} X_{j+a}$$

where $e_{j,1} = e_j$ and $e_{j,2} = e_j + e_{j+1}$. Thus, on the one hand

$$\begin{aligned} T^N(B \cap X_j) &= T^N B \cap T^N X_j \\ &\approx^{3\varepsilon} B \cap T^N X_j = B \cap T^{b - e_{j,a}} X_{j+a} \qquad (0 \le j \le 7) \end{aligned}$$

(where $C \approx^\eta D$ means $m(C \,\triangle\, D) < \eta$) and on the other hand

$$T^N(B \cap X_j) = T^{b - e_{j,a}}(B \cap X_{j+a}) \qquad (0 \le j \le 6 - a)$$

whence

$$\begin{aligned} B \cap X_{j+a} &\approx^{3\varepsilon} T^{-b + e_{j,a}} B \cap X_{j+a} \qquad && \forall 0 \le j \le 6 - a, \\ B &\approx^{27\varepsilon} T^{-b + e_{j,a}} B && \forall 0 \le j \le 6 - a, \end{aligned}$$

whence (choosing j, j' with $e_{j,a} - e_{j',a} = 1$)

$$B \approx^{54\varepsilon} TB \Rightarrow A \approx^{56\varepsilon} TA.$$

The cocycle. This cocycle $\phi : X \to \mathbb{Z}$ will be defined successively as a sum of coboundaries. Define $g^{(n)} : C_{n+2} \to \mathbb{Z}$ by

$$g^{(n)}(x) = \begin{cases} 1, & x \in \mathcal{S}_{n+1}, \\ -3, & x \in \mathcal{S}_{n+2}, \\ 0, & \text{else}. \end{cases}$$

Note that

$$(\ddagger) \qquad \forall n \geq 1,\ k \geq n+2, \qquad T^N X_{i,k} = X_{i+j,k} \Rightarrow g_N^{(n)} \equiv 0 \text{ on } X_{i,k}$$

(this is because $g_N^{(n)}|_{X_{i,k}} = j g_{l_k}^{(n)}|_{J_k} = 0$); whereas for all $U \in \mathcal{C}_n$,

$$U \cap T^{-(2l_n+1)} U \cap [g_{2l_n+1}^{(n)} = 1] \supset U \cap \bigcup_{k=0,1,3,7} X_{k,n} =: U \cap Y_n,$$

whence

$$m(U \cap T^{-(2l_n+1)} U \cap [g_{2l_n+1}^{(n)} = 1]) \geq \tfrac{4}{9} m(U).$$

Now fix a sequence $n_k \nearrow \infty$ such that

- $n_{k+1} > n_k + 2$,
- $\sum_{j \geq k+1} m(\mathcal{S}_{n_j}) < m(J_{n_k})/(45(2l_{n_k}+1))$

and define $\phi := \sum_{k=1}^{\infty} g^{(n_k)}$.

Ergodicity of T_ϕ. We see by (‡) that for all $k \geq 1$,

$$\phi_{2l_{n_k}+1} = \sum_{j \geq k} g_{2l_{n_k}+1}^{(n_j)} \qquad \text{on } Y_{n_k},$$

whence

$$\begin{aligned} m(Y_{n_k} \cap [\phi_{2l_{n_k}+1} \neq g_{2l_{n_k}+1}^{(n_k)}]) &\leq \sum_{j \geq k+1} m([g_{2l_{n_k}+1}^{(n_j)} \neq 0]) \\ &\leq (2l_{n_k}+1) \sum_{j \geq k+1} m(\mathcal{S}_{n_j}) \leq \frac{m(J_{n_k})}{45} \end{aligned}$$

and for $U \in \mathcal{C}_{n_k},\ U \neq \emptyset$, we have

$$\begin{aligned} &m(U \cap T^{-(2l_{n_k}+1)} U \cap [\phi_{2l_{n_k}+1} = 1]) \\ &\quad \geq m(U \cap T^{-(2l_{n_k}+1)} U \cap [g_{2l_{n_k}+1}^{(n_k)} = 1]) - m([\phi_{2l_{n_k}+1} \neq g_{2l_{n_k}+1}^{(n_k)}]) \\ &\quad \geq \frac{4}{9} m(U) - \frac{m(J_{n_k})}{45} \geq \frac{19 m(U)}{45}. \end{aligned}$$

To show that $T_\phi : X \times \mathbb{Z} \to X \times \mathbb{Z}$ is ergodic, it suffices by [Sch1] to show that if $A \in \mathcal{B}$, $m(A) > 0$ and $k \geq 1$ is large enough, then

$$m(A \cap T^{-(2l_{n_k}+1)}A \cap [\phi_{2l_{n_k}+1} = 1]) > 0.$$

To see this, note that for $k \geq 1$ large enough, there exists $U \in \mathcal{C}_n$ with $m(A \triangle U) < 2m(U)/45$, whence

$$\begin{aligned} m(A \cap T^{-(2l_{n_k}+1)}A \cap [\phi_{2l_{n_k}+1} = 1]) & \\ &\geq m(U \cap T^{-(2l_{n_k}+1)}U \cap [\phi_{2l_{n_k}+1} = 1]) - 2m(A \triangle U) \\ &\geq m(U)/3 > 0. \end{aligned}$$

Tightness of $\{m\text{-dist.}(S_{l_{n_k}}) : k \geq 1\}$. We first claim that

$$(\diamond) \qquad \Big|\Big(\sum_{k=1}^{K} g^{(n_k)}\Big)_{l_N}\Big| \leq 3 \quad \forall K \geq 1,\ N \geq n_K + 2.$$

To see this, we consider the tower C_{N+2} which consists of C_N-blocks, and the spacers $\mathcal{S}_{N+1} \cup \mathcal{S}_{N+2}$ on which $\sum_{k=1}^{K} g^{(n_k)} \equiv 0$. The cocycle sum over a C_N-block is zero by construction.

An arbitrary cocycle sum of length l_N in C_{N+2} begins in the middle of a C_N-block, either passes over a spacer interval (in $\mathcal{S}_{N+1} \cup \mathcal{S}_{N+2}$) or not, and continues to the middle of the next C_N-block. In the second case, the cocycle sum will be as over a C_N-block, and will be zero. In the first case, it will be as over a C_N-block less one interval (the one before the starting place) and

$$\Big(\sum_{k=1}^{K} g^{(n_k)}\Big)_{l_N} = -\sum_{k=1}^{K} g^{(n_k)}(x_0).$$

The claim $(\diamond)$ follows since $\sum_{k=1}^{K} g^{(n_k)} = 0, 1, -3$.

To prove our tightness claim, we prove that $m([|S_{l_{n_K}}| \geq 4]) \to 0$ as $K \to \infty$. Indeed, by $(\diamond)$,

$$\begin{aligned} m([|S_{l_{n_K}}| \geq 4]) &\leq m\Big(\Big[S_{l_{n_K}} \neq \Big(\sum_{k=1}^{K} g^{(n_k)}\Big)_{l_{n_K}}\Big]\Big) \\ &= m\Big(\Big[\Big(\sum_{k=K+1}^{\infty} g^{(n_k)}\Big)_{l_{n_K}} \neq 0\Big]\Big) \\ &\leq l_{n_K} m\Big(\Big[\sum_{k=K+1}^{\infty} g^{(n_k)} \neq 0\Big]\Big) \\ &\leq l_{n_K} \sum_{k=K+1}^{\infty} m(\mathcal{S}_{n_k}) \leq \frac{m(J_{n_K})}{90}. \end{aligned}$$

REFERENCES

[A] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Math. Surveys Monogr. 50, Amer. Math. Soc., Providence, RI, 1997.

[Br1] R. C. Bradley, *On a theorem of K. Schmidt*, Statist. Probab. Lett. 23 (1995), 9–12.

[Br2] —, *A "coboundary" theorem for sums of random variables taking values in a Banach space*, Pacific J. Math. 178 (1997), 201–224.

[Br3] —, *A "multiplicative coboundary" theorem for some sequences of random matrices*, J. Theoret. Probab. 9 (1996), 659–678.

[Br4] —, *On the dissipation of the partial sums of a stationary strongly mixing sequence*, Stochastic Process. Appl. 54 (1994), 281–290.

[Cha] R. V. Chacon, *Weakly mixing transformations which are not strongly mixing*, Proc. Amer. Math. Soc. 22 (1969), 559–562.

[Co] J.-P. Conze, *Transformations cylindriques et mesures finies invariantes*, Ann. Sci. Univ. Clermont Math. 17 (1979), 25–31.

[Fr] N. Friedman, *Introduction to Ergodic Theory*, van Nostrand, New York, 1970.

[F-K] N. Friedman and J. King, *Rank one lightly mixing*, Israel J. Math. 73 (1991), 281–288.

[Key-New] H. Keynes and D. Newton, *The structure of ergodic measures for compact group extensions*, Israel J. Math. 18 (1974), 363–389.

[Lem] M. Lemańczyk, *Ergodic compact Abelian group extensions*, habilitation thesis, Nicholas Copernicus Univ., Toruń, 1990.

[Leo] V. P. Leonov, *On the dispersion of time averages of a stationary random process*, Theory Probab. Appl. 6 (1961), 93–101.

[Mo-Sch] C. C. Moore and K. Schmidt, *Coboundaries and homomorphisms for non-singular actions and a problem of H. Helson*, Proc. London Math. Soc. 40 (1980), 443–475.

[Par] K. R. Parthasarathy, *Introduction to Probability and Measure*, Springer, New York, 1978.

[Rev] P. Revesz, *On a problem of Steinhaus*, Acta Math. Acad. Sci. Hungar. 16 (1965), 311–318.

[Sch1] K. Schmidt, *Cocycles of Ergodic Transformation Groups*, Lecture Notes in Math. 1, MacMillan of India, 1977.

[Sch2] —, *Amenability*, *Kazhdan's property* T, *strong ergodicity and invariant means for ergodic group-actions*, Ergodic Theory Dynam. Systems 1 (1981), 223–236.

[Weil] A. Weil, *L'intégration dans les groupes topologiques et ses applications*, Actualités Sci. Indust. 869, Hermann, Paris, 1940.

[Zim] R. J. Zimmer, *On the cohomology of ergodic group actions*, Israel J. Math. 35 (1980), 289–300.

School of Mathematical Sciences
Tel Aviv University
69978 Tel Aviv, Israel
E-mail: aaro@math.tau.ac.il

Mathematical Institute
Hebrew University
Jerusalem, Israel
E-mail: weiss@math.huji.ac.il

Received 13 August 1999;
revised 4 December 1999

(3816)