ISRAEL JOURNAL OF MATHEMATICS **130** (2002), 285–321

RANDOM ERGODIC THEOREMS AND REAL COCYCLES

BY

MARIUSZ LEMAŃCZYK

Department of Mathematics and Computer Science, Nicholas Copernicus University
ul. Chopina 12/18, 87-100 Toruń, Poland
e-mail: mlem@mat.uni.torun.pl

AND

EMMANUEL LESIGNE

Laboratoire de Mathématiques et Physique Théorique, UMR 6083 CNRS
Université François Rabelais, Parc de Grandmont, 37200 Tours, France
e-mail: lesigne@univ-tours.fr

AND

FRANÇOIS PARREAU

Laboratoire d'Analyse, Géométrie et Applications, UMR 7539, Institut Galilée
Université Paris 13, 99, av. J.B. Clément, 93430 Villetaneuse, France
e-mail: parreau@math.univ-paris13.fr

AND

DALIBOR VOLNÝ

Laboratoire Raphaël Salem, UMR 6085 CNRS, Université de Rouen
76821 Mont Saint-Aignan, France
e-mail: dalibor.volny@univ-rouen.fr

AND

MATÉ WIERDL

Department of Mathematical Sciences, University of Memphis
Memphis, TN 38152, USA
e-mail: mw@moni.msci.memphis.edu

Received July 23, 1998 and in revised form September 27, 2000

ABSTRACT

We study mean convergence of ergodic averages $\frac{1}{N}\sum_{n=0}^{N-1} f \circ \tau^{k_n(\omega)}$ (*) associated to a measure-preserving transformation or flow τ along the random sequence of times $k_n(\omega) = \sum_{j=0}^{n-1} F(T^j\omega)$ given by the Birkhoff sums of a measurable function F for an ergodic measure-preserving transformation T.

We prove that the sequence $(k_n(\omega))$ is almost surely universally good for the mean ergodic theorem, i.e., that, for almost every ω, the averages (*) converge for every choice of τ, if and only if the "cocycle" F satisfies a cohomological condition, equivalent to saying that the eigenvalue group of the "associated flow" of F is countable. We show that this condition holds in many natural situations.

When no assumption is made on F, the random sequence $(k_n(\omega))$ is almost surely universally good for the mean ergodic theorem on the class of mildly mixing transformations τ. However, for any aperiodic transformation T, we are able to construct an integrable function F for which the sequence $(k_n(\omega))$ is not almost surely universally good for the class of weakly mixing transformations.

1. Introduction

Let $(\Omega, \mathcal{T}, \mu)$ be a standard probability space, T a measure-preserving transformation of this space which we throughout assume to be ergodic, and F a real measurable function on it. For x real, we denote $e(x) = \exp(2i\pi x)$. We study the convergence of the averages

$$(1.1) \qquad \left(\frac{1}{N}\sum_{n=0}^{N-1} e\left(\theta F^{(n)}(\omega)\right)\right)_{N>0},$$

where $\theta \in \mathbb{R}$, $\omega \in \Omega$ and $F^{(n)}$ denotes the Birkhoff sum $F^{(n)} = \sum_{k=0}^{n-1} F \circ T^k$ of F.

The behavior of such sums is related to the theory of real cocycles over measure-preserving systems and their associated flows. As usual, the function F will also be called a **cocycle**.

Denoting by $\mathcal{B}$ the Borel σ-algebra of the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and by λ the Lebesgue measure on $(\mathbb{T}, \mathcal{B})$, we consider for each real θ the skew product transformation $T_{\theta,F}$ on the product space $(\Omega \times \mathbb{T}, \mathcal{T} \otimes \mathcal{B}, \mu \otimes \lambda)$ defined by

$$T_{\theta,F}(\omega, x) = (T\omega, x + \theta F(\omega)).$$

This is a measure-preserving transformation and it satisfies

$$T^n_{\theta,F}(\omega,x) = (T^n\omega, x+\theta F^{(n)}(\omega)).$$

A direct application of the pointwise ergodic theorem to the function $(\omega,x)\mapsto e(x)$ gives: for every θ, the sequence (1.1) converges μ-almost everywhere and in any $L^p(\mu)$, $1\le p<\infty$. Furthermore, if the skew product is ergodic, the limit is zero.

We are interested in the following question: can the set of full measure on which the sequence (1.1) converges be chosen independently of θ?

Definition: We call F a **good averaging cocycle** if there exists a subset Ω' of Ω of full measure such that for all $\omega\in\Omega'$ and for all $\theta\in\mathbb{R}$ the sequence (1.1) converges.

This "good averaging" property is related to the mean ergodic theorem along random sequences since, in case F takes integer values, there is equivalence between:

- F is a good averaging cocycle;
- for μ-almost every ω, the sequence $(F^{(n)}(\omega))$ is a good sequence for the mean ergodic theorem, that is to say, if $(X,\mathcal{A},\nu,\tau)$ is any invertible probability measure-preserving system, if $1\le p<\infty$ and if $f\in L^p(\nu)$, then the sequence

$$\left(\frac{1}{N}\sum_{n=0}^{N-1} f\circ\tau^{F^{(n)}(\omega)}\right) \tag{1.2}$$

  converges in $L^p(\nu)$.

When F takes real values there is a similar equivalence for measure-preserving flows (X,τ) (see Section 3).

In this paper we study the good averaging property, we give a variety of examples, and we investigate this random mean ergodic theorem. Similar questions have already been studied in [23] and [14]. A general reference for ergodic theorems along subsequences is [33].

Here is a summary of our results.

The good averaging property is not satisfied in general but it is satisfied in some interesting situations; in particular, in each of the following cases F is a good averaging cocycle:

- F is a regular real cocycle;
- F is integrable with non-zero mean, and more generally when its associated flow preserves a probability measure;

- the process $(F \circ T^n)$ is independent;
- the process $(F \circ T^n)$ is Gaussian;
- the dynamical system $(\Omega, \mathcal{T}, \mu, T)$ possesses a hyperbolic character and F is sufficiently regular.

Furthermore, the limit of (1.1) is zero for all $\theta \neq 0$, and the limit of (1.2) is the limit of the usual ergodic averages, in some of these cases, namely when:

- F is an ergodic real cocycle;
- the process $(F \circ T^n)$ is independent and F does not take its values in a subgroup $a\mathbb{Z}$;
- the process $(F \circ T^n)$ is Gaussian and F is not a coboundary;
- the associated flow of the cocycle F is weakly mixing.

We obtain a complete characterization of the good averaging property: F is a good averaging cocycle if and only if either F is a real coboundary, or the set of all θ such that θF is a coboundary modulo 1 is countable, that is the associated flow of F has countably many eigenvalues (Theorem 4.1).

Since not every function F is a good averaging cocycle, another question can be asked: under which assumption on the dynamical system $(X, \mathcal{A}, \nu, \tau)$ does the sequence (1.2) converge, whatever be F and f? We prove that given any F, for μ-almost every ω, we have: for any mildly mixing system $(X, \mathcal{A}, \nu, \tau)$ and any $f \in L^p(\nu)$, the sequence (1.2) converges in $L^p(\nu)$. On the other hand, for any aperiodic ergodic system $(\Omega, \mathcal{T}, \mu, T)$, there exists an integrable function F such that, for μ-almost all ω, there exists a weakly mixing system $(X, \mathcal{A}, \nu, \tau)$ and $f \in L^2(\nu)$ such that the sequence (1.2) does not converge. This means that there always exists a universal set of measure one for the random mean ergodic theorem on the class of mildly mixing systems, and there can be no such set for the class of weakly mixing systems.

We should add that the problem of mean convergence of averages (1.2) is quite different from the problem of pointwise convergence. For example, it is proved in [23] that if the random process $(F \circ T^n)$ is independent, centered and square integrable, then for almost all ω and for any choice of $(X, \mathcal{A}, \nu, \tau)$ there exists a bounded measurable function f on X such that the averages (1.2) do not converge almost everywhere. But these averages do converge in the mean (see [23] Section 7, or the present paper, Section 6). Other examples enlightening the great difference between problems of mean and pointwise convergence are given in Section 7.

The case of non-negative integrable functions and the case of integer-valued non-centered functions F have already been studied. In these cases, the strongest result of ν-almost everywhere convergence of the averages (1.2) is proved in [23]. Just a word on the method used in [23]: it is based on the return-times ergodic theorem, in the discrete and in the continuous time cases, and some subtle constructions. We are studying here only mean convergence and in this setting the return-times theorem is nothing else than the classical Wiener–Wintner ergodic theorem. The approach proposed in this paper will be quite different.

Let us notice finally that the only assumptions on the system $(X, \mathcal{A}, \nu, \tau)$ that we will make are on the spectral level, and therefore our results on the random mean ergodic theorem extend to any representation of $\mathbb{Z}$ (or $\mathbb{R}$) in the unitary group of a Hilbert space.

2. Sets of critical values associated to the function

For each $\theta \in \mathbb{R}$, let

$$\ell_\theta(\omega) := \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e(\theta F^{(n)}(\omega)),$$

which is well defined a.e. and in $L^2(\mu)$, as we already noticed. This limit function satisfies the **transfer equation**

$$\ell_\theta(T\omega) = e(-\theta F(\omega))\ell_\theta(\omega).$$

Recall that we assume that T is ergodic. Thus $|\ell_\theta|$ is a.s. constant and if it is not zero then F is a coboundary modulo 1, i.e., there exists a real measurable function G on Ω with $\theta F(\omega) = G(T\omega) - G(\omega)$ mod 1 a.e., and such a function G is called a **transfer function**.

We moreover have, almost everywhere, $\ell_\theta(T^n\omega) = e(-\theta F^{(n)}(\omega))\ell_\theta(\omega)$ for every $n > 0$, so

$$\frac{1}{N}\sum_{n=0}^{N-1} \ell_\theta(T^n\omega) = \ell_\theta(\omega)\frac{1}{N}\sum_{n=0}^{N-1} e(-\theta F^{(n)}(\omega)),$$

which implies, by the ergodic theorem,

$$\mathbb{E}[\ell_\theta] = |\ell_\theta|^2 \quad \text{a.e.} \tag{2.1}$$

Definitions:

- We denote by Θ_F the set of real numbers θ such that $\ell_\theta \neq 0$ a.e. or equivalently $\mathbb{E}[\ell_\theta] \neq 0$.
- We denote by Λ_F the set of real numbers θ such that θF is a coboundary modulo 1.

Notice that $\Theta_F \subset \Lambda_F$. We shall need basic properties of Λ_F, which come back to [6], [17], [25]. Namely, Λ_F is a Borel subgroup of $\mathbb{R}$ and it can be endowed with a natural Polish topology, which we describe at the end of this section; it is also equal to the eigenvalue group of the so-called associated flow (or "Mackey range") of F (see Section 5). If F is a real coboundary, that is if there exists a measurable real function G on Ω such that $F(\omega) = G(T\omega) - G(\omega)$ a.e., then $\Lambda_F = \mathbb{R}$. The converse is true (see Section 3).

The next results give a first link between these sets and the good averaging property.

THEOREM 2.1: *For μ-almost every ω,*

$$\text{(2.2)} \qquad \text{for all } \theta \notin \Theta_F, \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e\left(\theta F^{(n)}(\omega)\right) = 0.$$

The proof of this theorem is based on the Van der Corput inequality and the following proposition.

PROPOSITION 2.2: *Let G be a real measurable function on Ω. For μ-almost every ω,*

$$\text{(2.3)} \qquad \text{for all } \theta \in \mathbb{R}, \quad \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} e(\theta G(T^n\omega)) = \mathbb{E}[e(\theta G)].$$

In particular, if $F(\omega) = G(T\omega) - G(\omega)$ a.e., then for almost every ω we have $F^{(n)}(\omega) = G(T^n\omega) - G(\omega)$ for every n, so we obtain the following immediate corollary.

COROLLARY 2.3: *If F is a real coboundary, then it is a good averaging cocycle.*

Proof of Proposition 2.2: By the pointwise ergodic theorem, there exists a subset Ω' of Ω, with $\mu(\Omega') = 1$, such that for any rational number θ and any ω in Ω' convergence (2.3) holds. By density we just need to prove that for almost every ω the sequence $M_N^\omega \colon \mathbb{R} \to \mathbb{C}$, $N = 1, 2, \ldots$ defined by

$$M_N^\omega(\theta) = \frac{1}{N} \sum_{n=0}^{N-1} e(\theta G(T^n\omega))$$

is equicontinuous.

For each positive integer K, let η_K be the characteristic function of the set $[|G| > K]$, and let $\chi_K = 1 - \eta_K$. Since T is ergodic, for each K we have for almost all ω

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\eta_K(T^n\omega)=\mathbb{E}[\eta_K]. \tag{2.4}$$

Let us fix ω for which (2.4) holds for every K. We want to show that for every $\epsilon>0$ there is a $\delta>0$ so that if $|\theta-\theta'|<\delta$ then for every N

$$|M_N^\omega(\theta)-M_N^\omega(\theta')|<\epsilon.$$

Since for each fixed N, M_N^ω is continuous, we just need to find a δ which works for every large N. Let us write

$$\begin{aligned}M_N^\omega(\theta)-M_N^\omega(\theta') =&\frac{1}{N}\sum_{n=0}^{N-1}\chi_K(T^n\omega)(e(\theta G(T^n\omega))-e(\theta' G(T^n\omega)))\\ &+\quad\frac{1}{N}\sum_{n=0}^{N-1}\eta_K(T^n\omega)(e(\theta G(T^n\omega))-e(\theta' G(T^n\omega))).\end{aligned}$$

We can choose K so that $\mathbb{E}(\eta_K)=\mu([|G|>K])<\epsilon/4$ and then, because of (2.4), we can choose N_0 large enough so that if $N>N_0$ the absolute value of the second term is less than $\epsilon/2$. The first term can be made less than $\epsilon/2$ by choosing δ small enough because of the uniform (in n) estimate

$$\chi_K(T^n\omega)|e(\theta G(T^n\omega))-e(\theta' G(T^n\omega))|\leq 2\pi\cdot|\theta-\theta'|\cdot K. \qquad \blacksquare$$

Proof of Theorem 2.1: The proof is based on the classical Van der Corput inequality (see, for example, [22, A71.3]), which implies that if $(u_n)_{n\geq 0}$ is a sequence of complex numbers such that, for any integer $h\geq 0$,

$$\gamma_h:=\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}u_{n+h}\overline{u_n}$$

exists and satisfies

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}\gamma_h=0$$

then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}u_n=0.$$

By Proposition 2.2, we know that, for almost all ω, for any $h\geq 0$ and any θ,

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}e\left(\theta F^{(h)}(T^n\omega)\right)=\mathbb{E}\left[e\left(\theta F^{(h)}\right)\right].$$

Let us apply Van der Corput's argument to the sequence $u_n := e\left(\theta F^{(n)}(\omega)\right)$. We have

$$\begin{aligned}\gamma_h :&= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} e\left(\theta\left(F^{(n+h)}(\omega)-F^{(n)}(\omega)\right)\right)\\ &= \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} e\left(\theta F^{(h)}(T^n\omega)\right)\\ &= \mathbb{E}\left[e\left(\theta F^{(h)}\right)\right]\end{aligned}$$

and

$$\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}\gamma_h = \mathbb{E}\left[\lim_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1} e(\theta F^{(h)})\right]$$

which is equal to zero if $\theta\notin\Theta_F$. ∎

Corollary 2.4: *If the set Θ_F is countable, then F is a good averaging cocycle.*

Proof of Corollary 2.4: We know that for each $\theta\in\Theta_F$ there exists a subset Ω_θ of Ω of full measure such that for any $\omega\in\Omega_\theta$ the sequence (1.1) converges. ∎

We finish this section by the description of the natural topology of Λ_F. This topology is inherited from the ordinary topology on $\mathbb{R}$ and from the L^2-topology on the transfer functions. More precisely, to each $\theta\in\Lambda_F$ we choose a measurable transfer function G_θ from Ω into the torus $\mathbb{T}$ corresponding to $\theta F(\omega)$ (mod 1). Then, we define the metric d on Λ_F by

$$\begin{aligned}d(\theta,\theta') :&= |\theta-\theta'| + \inf_{x\in\mathbb{R}}\left(\int_\Omega |e(x+G_\theta)-e(G_{\theta'})|^2\,d\mu\right)^{\frac12}\\ &= |\theta-\theta'| + \left(2-2\left|\int_\Omega e(G_\theta-G_{\theta'})d\mu\right|\right)^{\frac12}\end{aligned}$$

(note that this expression does not depend on the choice of G_θ, $G_{\theta'}$ since $e(G_\theta)$ is determined a.e. up to a constant).

This metric gives to Λ_F the structure of a separable complete topological group, in other words Λ_F is a Polish group.

Now, Θ_F is an open neighborhood of 0 in Λ_F. Indeed, for every $\theta\in\Lambda_F$, we have

$$\ell_\theta(\omega) = \lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1} e\left(G_\theta(T^n\omega)-G_\theta(\omega)\right) = e\left(-G_\theta(\omega)\right)\int e(G_\theta)d\mu \text{ a.e.},$$

whence

$$|\ell_\theta| = \left| \int e(G_\theta)\, d\mu \right|, \tag{2.5}$$

and $\theta \to |\ell_\theta|$ is continuous in the topology of Λ_F.

In particular Λ_F is a countable union of translates of Θ_F. As Θ_F is clearly a Borel set, it follows that Λ_F is a Borel subgroup of $\mathbb{R}$. It also follows that Θ_F is countable iff Λ_F is countable.

3. The random mean ergodic theorem for mildly mixing systems

To begin, we recall in a general setting the link between convergence of trigonometric averages and the mean ergodic theorem along subsequences. A sequence of real numbers $(a_n)_{n\geq 0}$ is called **a good averaging sequence** if for every real number θ the sequence of the averages

$$\frac{1}{N} \sum_{n=0}^{N-1} e\left(\theta a_n\right) \tag{3.1}$$

converges.

PROPOSITION 3.1: *Let (a_n) be a good averaging sequence and $p \in [1, +\infty)$.*

If (a_n) is a sequence of integers then, given any invertible probability measure-preserving system $(X, \mathcal{A}, \nu, \tau)$ and $f \in L^p(\nu)$, the sequence

$$\left(\frac{1}{N} \sum_{n=0}^{N-1} f \circ \tau^{a_n} \right) \tag{3.2}$$

converges in $L^p(\nu)$.

Given any probability measure-preserving flow $(X, \mathcal{A}, \nu, (\tau_t)_{t\in\mathbb{R}})$ and $f \in L^p(\nu)$, the sequence

$$\left(\frac{1}{N} \sum_{n=0}^{N-1} f \circ \tau_{a_n} \right) \tag{3.3}$$

converges in $L^p(\nu)$.

These results are well known. In order to be complete, and because it was an important part of our initial motivation for the study of (1.1), we recall the proof in the case of an integer-valued sequence and a measure-preserving transformation. The argument can be repeated word for word in the case of a real sequence (a_n) and a measure-preserving $\mathbb{R}$-action (then the spectral measure in the argument below is defined on $\mathbb{R}$).

Proof of Proposition 3.1: If $f \in L^2(\nu)$ and σ_f denotes the spectral measure of f under τ, then

$$\left\| \frac{1}{N} \sum_{n=0}^{N-1} f \circ \tau^{a_n} - \frac{1}{K} \sum_{k=0}^{K-1} f \circ \tau^{a_k} \right\|_2^2 = \int_{\mathbb{T}} \left| \frac{1}{N} \sum_{n=0}^{N-1} e(\theta a_n) - \frac{1}{K} \sum_{k=0}^{K-1} e(\theta a_k) \right|^2 d\sigma_f(\theta).$$

Since (a_n) is a good averaging sequence, by the dominated convergence theorem and the Cauchy criterion, we obtain the convergence of (3.2) in $L^2(\nu)$. For bounded functions f, the sequence (3.2) converges in probability and is uniformly bounded, hence it converges in $L^p(\nu)$. The result follows by density of L^∞ in L^p. ∎

Remark 3.2: In case when the limit of the averages (3.1) is zero for any non-zero real number θ, i.e., when the sequence $(a_n\theta)$ is uniformly distributed mod 1 for any non-zero real number θ, the limit of the averages (3.3) is the limit of usual ergodic averages and the sequence (a_n) is a Poincaré recurrence sequence ([12]).

Following Furstenberg's classical argument, this property has consequences in combinatorial number theory. This remark, which is very well known in its statement for integer-valued sequences (a_n), is developed in a short appendix (Section 9) for the real case.

The proof of Proposition 3.1 shows that, given $(X, \mathcal{A}, \nu, \tau)$ and $f \in L^p(\nu)$, the conclusion holds if the averages (3.1) converge σ_f-a.e. Therefore, from Theorem 2.1 we obtain a spectral condition on $(X, \mathcal{A}, \nu, \tau)$ for the validity of the random mean ergodic theorem.

PROPOSITION 3.3: *Given any integer-valued cocycle F, there exists a subset Ω' of Ω of full measure such that, for any $\omega \in \Omega'$, we have: if σ denotes the maximal spectral type of the system $(X, \mathcal{A}, \nu, \tau)$ in the orthocomplement of the constant functions in $L^2(\nu)$ and $\sigma(\Theta_F) = 0$, then for each $f \in L^p(\nu)$, the sequence (1.2) converges in $L^p(\nu)$ to the integral of f.*

The similar statement holds for real-valued F and measure-preserving flows.

Remark: If we are looking for explicit constructions of "good sequences of times" for the mean ergodic theorem, we can remark that, in the case when T is a uniquely ergodic transformation of the regular measure space $(\Omega, \mathcal{T}, \mu)$ and F is continuous, we can choose $\Omega' = \Omega$ in the statement of Proposition 3.3. This is so because, in that particular case, the universal set Ω' for $\theta \notin \Theta_F$ is the set where the ergodic averages of the continuous functions $e(\theta F^{(h)}(\cdot))$ converge (proof of Theorem 2.1).

The next results give us precise information on the possible size of the Borel subgroup Λ_F. In [17] and [6] it is proved that

(3.4) *either the set Λ_F has Lebesgue measure zero, or F is a real coboundary.*

In fact, much more is known. Moore and Schmidt in [25] gave a considerable strengthening of this result. To formulate their result let us recall the definition of a full measure on $\mathbb{R}$, which appears in [25]: a Borel probability measure σ on $\mathbb{R}$ is **full** if $\limsup_{t\to\infty} |\widehat{\sigma}(t)| < 1$.

THEOREM 3.4 ([Moore and Schmidt]): *If a full measure is concentrated on Λ_F then F is a real coboundary and $\Lambda_F = \mathbb{R}$.*

In other words, if F is not a real coboundary, then Λ_F is a **weak Dirichlet subgroup** (see [19], where the more precise result that Λ_F is "saturated" is shown). Although Theorem 3.4 is not original, we propose a new short proof, whose first argument will be used later.

Proof of Theorem 3.4: We first notice that there exists a real measurable function G on $\Lambda_F \times \Omega$ such that, for each $\theta \in \Lambda_F$, for μ-almost all ω,

$$\theta F(\omega) = G(\theta, T\omega) - G(\theta, \omega) \bmod 1. \tag{3.5}$$

Indeed, as the averages in (1.1) are measurable functions of (θ, ω), we can choose the measurable function G on $\Theta_F \times \Omega$ such that $e(G(\theta,\omega)) = \ell_\theta(\omega)/|\ell_\theta(\omega)|$ whenever the limit $\ell_\theta(\omega)$ exists and is non-zero (and, e.g., we let $G(\theta,\omega) = 0$) otherwise). So (3.5) holds μ-a.e. for each $\theta \in \Theta_F$. Consider then a countable Borel partition (Λ_n) of Λ_F, where $\Lambda_0 = \Theta_F$ and each Λ_n, $n \neq 0$ is the translate of a part of Θ_F by some $\theta_n \in \Lambda_F$ (which is possible because Θ_F is open in Λ_F). For each $n \neq 0$, we choose a transfer function G_{θ_n} for $\theta_n F$, and we define G on $\Lambda_n \times \Omega$ by

$$G(\theta, \cdot) := G_{\theta_n}(\cdot) + G(\theta - \theta_n, \cdot).$$

Now, suppose that a full probability measure σ is concentrated on Λ_F and fix $\epsilon > 0$ such that

$$\limsup_{t\to\infty} |\widehat{\sigma}(t)| < 1 - 3\epsilon. \tag{3.6}$$

Then the map $\omega \mapsto g_\omega := e(-G(\cdot,\omega))$ is measurable from Ω into $L^1(\sigma)$ and there exists a measurable subset A of positive measure in Ω such that

$$\text{for all } \omega,\, \omega' \in A, \quad \int |g_\omega - g_{\omega'}| d\sigma < \epsilon. \tag{3.7}$$

We fix $\omega_0 \in A$ and we associate to every ω the Fourier transform ϕ_ω defined on $\mathbb{R}$ by

$$\phi_\omega(t) = \int g_\omega(\theta)\overline{g_{\omega_0}(\theta)} e(-t\theta) d\sigma(\theta).$$

The set of (θ,ω) satisfying (3.5) is measurable. So, by the Fubini theorem, we have, for μ-almost all ω, for σ-almost all θ,

$$e(\theta F(\omega)) \cdot g_{T\omega}(\theta) = g_\omega(\theta),$$

and, after Fourier transform, this gives

$$\phi_{T\omega}(t) = \phi_\omega\left(t + F(\omega)\right). \tag{3.8}$$

On the other hand, by (3.7), we have

$$\text{for all } \omega \in A, \quad |\widehat{\sigma}(t) - \phi_\omega(t)| < \epsilon. \tag{3.9}$$

The Fourier transform $g \mapsto [t \mapsto \int g(\theta) e(-t\theta) d\sigma(\theta)]$ is continuous from $L^1(\sigma)$ into the space of bounded continuous functions on $\mathbb{R}$ equipped with the uniform norm. Hence the map

$$M(\omega) := \sup_{t\in\mathbb{R}} |\phi_\omega(t)|$$

is measurable on Ω. But, by (3.8), the map M is T-invariant, hence it is a.e. constant, and by (3.9) we have

$$M \geq \sup_{t\in\mathbb{R}} |\widehat{\sigma}(t)| - \epsilon = 1 - \epsilon.$$

Similarly, by (3.8), (3.6) and (3.9),

$$\limsup_{t\to\pm\infty} |\phi_\omega(t)| \text{ is a.e. constant and } < 1 - 2\epsilon.$$

This implies that the real function

$$G(\omega) := \sup\{t \in \mathbb{R} : |\phi_\omega(t)| \geq 1 - \epsilon\}$$

is almost everywhere defined, and it is measurable, again by the continuity of the Fourier transform. By (3.8), we have

$$G(T\omega) = G(\omega) - F(\omega)$$

and the proof is finished. ∎

We give now an application of Theorem 3.4 to the random mean ergodic theorem. For simplicity, we discuss here the case of an integer-valued function F, but all can be extended to the real case.

Recall that a measure-preserving system $(X, \mathcal{A}, \nu, \tau)$ is **mildly mixing** ([13], [12], [1], [31]) if it has no rigid factor. Equivalently, if σ denotes the maximal spectral type of τ in the orthocomplement of the constant functions in $L^2(\nu)$, then each probability measure absolutely continuous with respect to σ is full. A strongly mixing system is mildly mixing and mild mixing implies weak mixing.

THEOREM 3.5: *For μ-almost ω, we have: for each mildly mixing measure-preserving system $(X, \mathcal{A}, \nu, \tau)$, for each $1 \leq p < +\infty$ and $f \in L^p(\nu)$ the sequence*

$$\left(\frac{1}{N}\sum_{n=0}^{N-1} f \circ \tau^{F^{(n)}(\omega)}\right)$$

converges in $L^p(\nu)$. Moreover, if F is not a real coboundary then the limit of this sequence is $\int_X f\, d\nu$.

Proof of Theorem 3.5: If F is a real coboundary, the result is given by Corollary 2.3 and Proposition 3.1. If F is not a real coboundary, Theorem 3.4 implies that there does not exist any measure absolutely continuous with respect to the maximal spectral type σ of τ concentrated on Λ_F; so $\sigma(\Lambda_F) = 0$, hence $\sigma(\Theta_F) = 0$ and Proposition 3.3 applies. ∎

Remark: We will show later that the class of τ for which this random mean ergodic theorem holds universally does not contain all weakly mixing transformations.

The previous result holds in particular for all strongly mixing dynamical systems $(X, \mathcal{A}, \nu, \tau)$. This can be seen as an application of Rosenblatt's study of norm convergence in [32]. Indeed, it is possible to prove that, for almost every, ω the sequence of probability measures given on $\mathbb{Z}$ by

$$\frac{1}{N}\sum_{n=0}^{N-1} \delta(F^{(n)}(\omega))$$

(where $\delta(t)$ is the Dirac mass at point t) is uniformly dissipative and hence the norm convergence of (1.2) when τ is strongly mixing is the generalization of the Blum–Hanson theorem given in [32].

4. A complete characterization

THEOREM 4.1: *F is a good averaging cocycle if and only if either it is a real coboundary or Λ_F is countable.*

We just have to show the necessity of the condition. Indeed, the fact that it is sufficient follows from Corollaries 2.3 and 2.4. We will proceed in several steps that we can summarize in the following way: under the good averaging hypothesis, we prove the existence of a measurable selector of the transfer functions for the mod 1 coboundaries θF, $\theta \in \Lambda_F$, which is continuous and homomorphic with respect to θ; then, we show that convergence in Λ_F implies uniform convergence on a set of positive measure and, using a theorem of Rosenthal, we deduce that Λ_F is locally compact, and we conclude.

From now on we suppose that F is a good averaging cocycle. The function ℓ_θ is defined as in Section 2. We will denote $\ell(\theta,\omega) := \ell_\theta(\omega)$. For almost all ω, the function $\ell(\cdot,\omega)$ is well defined everywhere and $\ell(0,\omega) = 1$.

LEMMA 4.2: *For almost all ω, the function $\ell(\cdot,\omega)$ is continuous on Λ_F.*

Proof of Lemma 4.3: By a standard representation theorem, considering for example the sample space of the process $(F \circ T^n)$, we can always assume that Ω is a Polish space, equipped with its Borel σ-algebra and a regular probability measure μ, and that the transformation T and the function F are continuous. For $N > 0$ and $(\theta,\omega) \in \mathbb{R} \times \Omega$ we define

$$\ell_N(\theta,\omega) := \frac{1}{N}\sum_{n=0}^{N-1} e\left(\theta F^{(n)}(\omega)\right).$$

Thus we have a sequence (ℓ_N) of continuous functions on $\mathbb{R} \times \Omega$ and hence the more so on $\Lambda_F \times \Omega$. There exists $\Omega' \subset \Omega$ with $\mu(\Omega') = 1$ such that for all $(\theta,\omega) \in \mathbb{R} \times \Omega'$ the limit

$$\ell(\theta,\omega) := \lim_{N\to\infty} \ell_N(\theta,\omega)$$

exists. We can suppose that Ω' is T-invariant and we have

$$\ell(\theta,T\omega) = e(-\theta F(\omega))\ell(\theta,\omega) \quad \text{on } \mathbb{R} \times \Omega'. \tag{4.1}$$

Furthermore, by (2.5), there exists an open neighborhood U_0 of 0 in Λ_F on which $|\ell_\theta| > \frac{1}{2}$, that is $\ell(\theta,\cdot)$ has a (a.e.) constant modulus $> \frac{1}{2}$ if $\theta \in U_0$.

Let K be a compact subset of Ω' with $\mu(K) > 0$ and such that K is equal to the topological support of the restriction of μ to K. The product space $U_0 \times K$

equipped with the $\Lambda_F \times \Omega$-topology is a Baire space, since it is open in the complete metric space $\Lambda_F \times K$. On this space, the function ℓ is the limit of a sequence of continuous functions.

Let $0 < \epsilon < \frac{1}{4}$. By the Baire Theorem (e.g., [28], Theorem 7.3), there exist non-empty open subsets $U' \subset U_0$ and $A \subset K$ such that, for any $\theta, \theta' \in U'$ and any ω, $\omega' \in A$,

$$|\ell(\theta, \omega) - \ell(\theta', \omega')| < \epsilon. \tag{4.2}$$

We fix $\theta_0 \in U'$ and find $\omega_0 \in A$ such that $|\ell(\theta_0, \omega_0)| > \frac{1}{2}$. By (4.2), for all $\theta \in U'$ and all $\omega \in A$ we have $|\ell(\theta, \omega)| > \frac{1}{4}$ and, by (4.1), this implies that, for all $\theta \in U'$ and all $\omega \in \Omega'' := \bigcup_{n \geq 0} T^{-n} A$,

$$|\ell(\theta, \omega)| > 1/4. \tag{4.3}$$

Thanks to our assumption on the support of $\mu_{|K}$ we have $\mu(A) > 0$, hence Ω'' is of full measure. We put $U := U' - \theta_0$ which is a neighborhood of zero in Λ_F, and, for any $(\theta, \omega) \in U \times \Omega''$, we define $G_\epsilon(\theta, \omega) \in \mathbb{T}$ by

$$e(-G_\epsilon(\theta, \omega)) = \frac{\ell(\theta + \theta_0, \omega)}{|\ell(\theta + \theta_0, \omega)|} \cdot \frac{|\ell(\theta_0, \omega)|}{\ell(\theta_0, \omega)}.$$

Then for any $\omega \in \Omega''$ and any $\theta \in U$,

$$G_\epsilon(\theta, T\omega) = \theta F(\omega) + G_\epsilon(\theta, \omega) \bmod 1,$$

and, by (4.2) and (4.3), for any $\omega \in A$ and any $\theta \in U$,

$$|1 - e(G_\epsilon(\theta, \omega))| < 8\epsilon.$$

Now, we want to increase the measure of the set A. Let m be a positive integer such that $\mu(\bigcup_{k=0}^{m-1} T^{-k} A) > 1 - \epsilon$ and let $M > 0$ be such that $\mu([|F| > M]) < \epsilon/m$. Define

$$A_\epsilon := \Big(\bigcap_{k=0}^{m-1} T^{-k}[|F| \leq M]\Big) \bigcap \Big(\bigcup_{k=0}^{m-1} T^{-k} A\Big)$$

and

$$U_\epsilon := U \cap \Big(-\frac{\epsilon}{mM}, \frac{\epsilon}{mM}\Big).$$

U_ϵ is a neighborhood of zero in Λ_F and, for any $\omega \in A_\epsilon$, for any $\theta \in U_\epsilon$, we can choose k $(0 \leq k < m)$ such that $T^k\omega \in A$, whence

$$|1 - e(G_\epsilon(\theta, \omega))| \leq |1 - e(G_\epsilon(\theta, T^k\omega))| + |1 - e(\theta F^{(k)}(\omega))| < 8\epsilon + kM|\theta| < 9\epsilon,$$

and we moreover have $\mu(A_\epsilon) > 1 - 2\epsilon$.

We now consider a sequence (ϵ_k) of positive numbers such that $\sum_{k\geq 1} \epsilon_k < \frac{1}{2}$ and we put $U_k := U_{\epsilon_k}$, $A_k := A_{\epsilon_k}$, $A_0 := \bigcap_{k\geq 1} A_k$ and $G_k := G_{\epsilon_k}$. We denote by A_0' the set of $\omega \in A_0$ such that, for any $k \geq 1$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_{A_k}(T^n\omega) = \mu(A_k) > 1 - 2\epsilon_k.$$

Then $\mu(A_0') = \mu(A_0) > 0$. We are going to show that, for every $\omega \in A_0'$, the sequence of functions $\ell_N(\cdot, \omega)$ is equicontinuous on Λ_F.

Fix $\omega \in A_0'$. For any given $\delta > 0$, there exist $k \geq 1$ such that $\epsilon_k < \delta$ and $N_k > 0$ such that, for $N \geq N_k$,

$$\frac{1}{N} \sum_{n=0}^{N-1} \chi_{A_k}(T^n\omega) > 1 - 2\delta.$$

Then, for every $N \geq N_k$ and every $\theta, \theta' \in \Lambda_F$ with $\theta - \theta' \in U_k$,

$$\begin{aligned} |\ell_N(\theta,\omega) - \ell_N(\theta',\omega)| \leq & \frac{1}{N} \sum_{n=0}^{N-1} |e((\theta - \theta')F^{(n)}(\omega)) - 1| \\ = & \frac{1}{N} \sum_{n=0}^{N-1} |e(G_k(\theta - \theta', T^n\omega)) - e(G_k(\theta - \theta', \omega))| \\ \leq & \frac{2}{N} \sum_{n=0}^{N-1} (1 - \chi_{A_k}(T^n\omega)) + 36\pi\epsilon_k \leq 118\delta. \end{aligned}$$

This proves the equicontinuity property, and it implies that, for every $\omega \in A_0'$, $\ell(\theta, \omega)$ is a continuous function of θ on Λ_F. But, by (4.1), the set of ω's satisfying this property is invariant under T. Hence this set is of full measure and Lemma 4.2 is proved. ∎

LEMMA 4.3: *There exist a T-invariant set Ω' of full measure and a measurable $G: \Lambda_F \times \Omega' \to \mathbb{T}$ such that for every $\omega \in \Omega'$, θ, $\theta' \in \Lambda_F$*

$$(4.4) \quad G(\theta, T\omega) = \theta F(\omega) + G(\theta, \omega), \quad G(\theta + \theta', \omega) = G(\theta, \omega) + G(\theta', \omega) \bmod 1$$

and, moreover, for every $\omega \in \Omega'$, the function $\theta \mapsto G(\theta, \omega)$ is continuous on Λ_F.

Proof of Lemma 4.3: We start from the conclusion of Lemma 4.2. For $\theta \in \Theta_F$, the function $|\ell(\theta, \cdot)|$ is a.e. equal to a positive constant. Using the separability of Θ_F and the continuity of $\theta \mapsto \ell(\theta,\omega)$, we obtain that there exists a set of full measure Ω'' such that, for every $\omega \in \Omega''$ and every $\theta \in \Theta_F$, $\ell(\theta,\omega) \neq 0$. We may also assume that the transfer equation holds everywhere on $\Theta_F \times \Omega''$.

We construct a measurable selector G for the transfer functions on $\Lambda_F \times \Omega''$ as in the proof of Theorem 3.4, starting with $e(-G(\theta,\omega)) = \ell(\theta,\omega)/|\ell(\theta,\omega)|$ on $\Theta_F \times \Omega''$. We consider again a countable partition (Λ_n) of Λ_F, such that $\Lambda_0 = \Theta_F$ and for each $n \neq 0$ there exists $\theta_n \in \Lambda_F$ with $\Lambda_n - \theta_n \subset \Theta_F$, and we let

$$G(\theta,\cdot) := G_{\theta_n}(\cdot) + G(\theta-\theta_n,\cdot) \quad \text{if } \theta \in \Lambda_n,$$

where G_{θ_n} is a transfer function (mod 1) for $\theta_n F$.

Now, there is a subset of Ω'' with full measure on which $G(\cdot,\omega)$ is continuous on each Λ_n and the transfer equation $G(\theta,T\omega) = \theta F(\omega) + G(\theta,\omega)$ mod 1 holds for every $\theta \in \Lambda_F$. This equation implies that, for all $\theta,\theta' \in \Lambda_F$,

$$G(\theta+\theta',\omega) - G(\theta,\omega) - G(\theta',\omega)$$

is μ-almost everywhere equal to a constant $c(\theta,\theta')$. For any choice of n, n', n'', let us denote $C_{n,n',n''} := \{(\theta,\theta') \in \Lambda_n \times \Lambda_{n'}\colon \theta+\theta' \in \Lambda_{n''}\}$. In each non-empty $C_{n,n',n''}$, we choose a countable dense subset, and we denote by D the union of all these subsets. There is a further subset Ω' of full measure of Ω'' such that, for all $\omega \in \Omega'$, for all $(\theta,\theta') \in D$,

$$G(\theta+\theta',\omega) - G(\theta,\omega) - G(\theta',\omega) = c(\theta,\theta'). \tag{4.5}$$

Since, for all $\omega \in \Omega'$, $G(\cdot,\omega)$ is continuous on each Λ_n and thanks to the choice of D, this implies that, for all $\omega \in \Omega'$, the identity (4.5) is true for all $\theta,\theta' \in \Lambda_F$. Now we can pick any ω_0 in Ω' and replace $G(\theta,\omega)$ by $G(\theta,\omega) - G(\theta,\omega_0)$. This transfer function satisfies (4.4).

Moreover, for every $\omega \in \Omega'$, $G(\cdot,\omega)$ is continuous on Θ_F, which is a neighborhood of 0. Hence this homomorphism is continuous on all Λ_F. ∎

Now, we are able to show that the convergence in Λ_F implies the uniform convergence of the transfer functions $G(\theta,\cdot)$ on a set of positive measure. In order to obtain a topological equivalence, we moreover have to take in account the usual $\mathbb{R}$-convergence. We shall do that by using the correspondence between transfer functions and eigenfunctions of the associated flow described in the next section, but note that the functions we construct here are everywhere defined.

To each $\theta \in \Lambda_F$, we associate the modulus one function g_θ defined on $\Omega' \times [0,1]$ by

$$g_\theta(\omega,t) = e\left(G(\theta,\omega) + \theta t\right).$$

We denote by Λ the multiplicative group

$$\Lambda = \{g_\theta : \theta \in \Lambda_F\}$$

and by λ the Lebesgue measure on $[0,1]$.

LEMMA 4.4: *There exists a measurable subset A of Ω with positive measure such that the topology of $L^2(A\times[0,1], \mu|_A \otimes \lambda)$ and the topology of uniform convergence on $A\times[0,1]$ coincide on Λ, and such that, if Λ is equipped with this topology, then the map $\theta \mapsto g_\theta$ is a topological group isomorphism.*

Proof of Lemma 4.4: For $n>0$, let B_n be the closed ball of radius $1/n$ centered at 0 in Λ_F. Given $\epsilon>0$, since B_n is separable and each $e(G(\cdot,\omega))$ with $\omega\in\Omega'$ is a continuous character of Λ_F, the set $A_{n,\epsilon}$ of all $\omega\in\Omega'$ such that $|1-e\left(G(\theta,\omega)\right)|\le\epsilon$ for every $\theta\in B_n$ is measurable, and moreover

$$\bigcup_{n>0} A_{n,\epsilon} = \Omega'.$$

Let (n_k) $(k>0)$ be a sequence of positive integers such that

$$\sum_{k>0} \mu\big(\Omega' \setminus A_{n_k,1/k}\big) < 1$$

and let $A = \bigcap A_{n_k,1/k}$. Then $\mu(A)>0$ and, when $d(\theta,0)\le 1/n_k$, we have $|1-e(G(\theta,\omega)|\le 1/k$ uniformly on A.

We also know that if $\theta_n\to 0$ in Λ_F then $\theta_n\to 0$ in $\mathbb{R}$. Hence, if $\theta_n\to 0$ in Λ_F, then $g_{\theta_n}\to 1$ uniformly on $A\times[0,1]$.

Conversely, let us suppose that $g_{\theta_n}\to 1$ in $L^2(\mu|_A\otimes\lambda)$. We have $\theta_n\to 0$ in $\mathbb{R}$ and $G(\theta_n,\cdot)\to 0$ in $L^2(\mu|_A)$. Indeed, denoting by μ_A the conditional probability with respect to A, it is easy to verify that

$$\int_0^1\int_A |1-g_{\theta_n}(\omega,t)|^2 d\mu_A(\omega)dt \ge \max(\theta_n^2, \int_A |1-e(G(\theta_n,\omega))|^2 d\mu_A(\omega)).$$

Using the transfer equation, we deduce that, for any $k\ge 0$, we have $G(\theta_n,\cdot)\to 0$ in $L^2(\mu|_{T^{-k}A})$. By ergodicity of T we conclude that $G(\theta_n,\cdot)\to 0$ in $L^2(\mu)$. The sequence (θ_n) goes to zero in Λ_F. ∎

The next step of the proof is a general theorem of independent interest.

THEOREM 4.5: *Let $(A,\mathcal{A},\nu)$ be a probability space and Λ either a multiplicative subgroup of measurable functions of modulus one on A or an additive subgroup* *of real bounded measurable functions on A. If the L^2-topology and the topology of uniform convergence coincide on Λ, and Λ is closed in these topologies, then Λ is locally compact.*

We give the proof of the theorem in the multiplicative case, which is more difficult and is the one we need. For the additive case, we need only make obvious simplifications.

Proof of Theorem 4.5: We will make use of a result of Haskell Rosenthal in functional analysis ([34]). In a real Banach space, a sequence (f_k) is said to be a **Sidon sequence** if there exists $\delta > 0$ such that, for every finite sequence of real coefficients c_k, we have

$$\left\|\sum c_k f_k\right\| \geq \delta \sum |c_k|. \tag{4.6}$$

If (f_k) is a Sidon sequence then its span is isomorphic to ℓ^1. Rosenthal's theorem asserts:

In a real Banach space, every bounded sequence contains either a weak Cauchy subsequence or a Sidon subsequence.

We will apply this theorem in the Banach space of bounded real functions defined on A, equipped with the uniform norm. Note that, in this space, a weak Cauchy sequence converges pointwise and thus, by the dominated convergence theorem, it converges in the $L^2(\nu)$-norm.

Denote by Λ' the image of Λ by the lifting $e(\theta) \mapsto \theta \in [-\frac{1}{2}, \frac{1}{2})$ and denote by $B(r)$ the closed ball or radius r centered at 0 in Λ' for the uniform norm. From our hypothesis we deduce that, in the ball $B(1/4)$, the L^2-convergence implies the uniform convergence.

Let $\epsilon \in (0, 1/2)$. If m is a positive integer and if $f/m \in B(\epsilon/m)$, then $f \in B(\epsilon)$. We claim that if the ball $B(\epsilon)$ can be covered by a finite number of subsets of diameter less than $\epsilon/2$ then the ball $B(\epsilon)$ is precompact, i.e., $B(\epsilon)$ can be covered by a finite number of balls of arbitrary small radius. Indeed, if the ball $B(\epsilon)$ can be covered by a finite number of subsets of diameter less than $\epsilon/2^k$, then it can be covered by a finite number of balls of radius $\epsilon/2^k$ and the ball $B(\epsilon/2^k)$ can be covered by a finite number of subsets of diameter less than $\epsilon/4^k$ (take the preimage of the covering of $B(\epsilon)$ under the map $f \mapsto 2^k f$), hence $B(\epsilon)$ can be covered by a finite number of subsets of diameter less than $\epsilon/4^k$. Our claim follows by induction.

Now, assume that none of the $B(2^{-k})$ is precompact. We construct inductively an infinite sequence (f_k) in Λ' with $2^{-k} f_k \in \Lambda'$ for every k and $\|f_i - f_j\|_\infty \geq 1/8$

for every $i \neq j$. Pick any f_0 in Λ'. Suppose f_0, ..., f_{k-1} are chosen. By the previous claim, we can find g_0, ..., g_k in $B(2^{-(k+1)})$ such that $\|g_i - g_j\|_\infty > 2^{-(k+2)}$ whenever $i \neq j$. Then no pair among the $2^k g_j$'s can be in the same subset of diameter 1/4, so we can choose j such that the distance of $f_k := 2^k g_j$ to each f_i, $i < k$, is $\geq 1/8$.

By construction, the sequence (f_k) has no uniformly convergent subsequence; thanks to our preliminary remarks, we know that this sequence has no weak Cauchy subsequence. By Rosenthal's theorem, the sequence $(f_k)_{k\geq 2}$ has a Sidon subsequence, which we still denote by $(f_k)_{k\geq 2}$. We fix $\delta > 0$ such that (4.6) holds. Note that, for every k, we have $2^{-k} f_k \in \Lambda'$.

Consider a finite random sum $\sum_{j\in J} \epsilon_j f_j$, where (ϵ_j) is an independent random sequence of ± 1 with equal probabilities. We have

$$E\left(\left\|\sum_{j\in J}\epsilon_j f_j\right\|_2^2\right) = E\left(\sum_{i,j\in J}\epsilon_j\epsilon_i\langle f_j, f_i\rangle\right) = \sum_{j\in J}\|f_j\|_2^2 \leq \#J.$$

It follows that there exists a deterministic sequence $(\epsilon_k)_{k<j\leq k+2^k}$ such that

$$\left\|\sum_{j=k+1}^{k+2^k}\epsilon_j f_j\right\|_2 \leq 2^{k/2},$$

and we define

$$g_k := 2^{-k-1}\sum_{j=k+1}^{k+2^k}\epsilon_j f_j.$$

Since $2^{-k-1} f_j \in \Lambda'$ for $j > k$, we have $g_k \in B(1/4)$. Moreover the sequence (g_k) goes to zero in L^2-norm, but, by the Sidon property,

$$\|g_k\|_\infty \geq \delta/2.$$

This contradicts our hypothesis. We conclude that some closed ball $B(2^{-k})$ is precompact. In the complete metric group Λ there exists a closed precompact, hence compact, neighborhood of zero. This proves that Λ is locally compact. ∎

Proof of Theorem 4.1: Assume that F is a good averaging cocycle. By Lemma 4.4 and Theorem 4.5, Λ_F is locally compact. Now, a locally compact group which is continuously embedded in $\mathbb{R}$ is either discrete or equal to $\mathbb{R}$ (e.g., [35], Theorem 3.2.2). In the former case, it is countable, since it is separable, and in the latter case we know that F is a coboundary. ∎

Thanks to Theorem 4.1 we have explicit examples of functions which are not good averaging cocycles. One now classical construction has been given by Osikawa ([27]) of a cocycle F such that Λ_F is uncountable and which is not a coboundary. It is described in Aaronson's book ([1, §2.6]). In this example the dynamical system $(\Omega, \mathcal{T}, \mu, T)$ is an odometer and the cocycle is positive.

5. Some links with the theory of real cocycles

In this section, we describe some relations between the concept of regular cocycle, the concept of associated flow and the good averaging property.

Let the space $\Omega\times\mathbb{R}$ be equipped with the infinite measure $\mu\otimes\lambda$ where λ denotes the Lebesgue measure on $\mathbb{R}$. Given a real cocycle F on Ω, the corresponding cylindrical flow T_F is the skew product defined on $\Omega \times \mathbb{R}$ by

$$T_F(\omega, x) = (T\omega, x + F(\omega)),$$

and F is said to be ergodic if T_F is ergodic.

Following [36], a real number α is called an **essential value** of a cocycle F if, for all $\epsilon > 0$, for all $B \in \mathcal{T}$ with $\mu(B) > 0$, there exists an integer $n > 0$ such that

$$\mu\left(B \cap T^{-n}B \cap \left\{\omega \in \Omega : |F^{(n)}(\omega) - \alpha| < \epsilon\right\}\right) > 0.$$

The set $E(F)$ of essential values of F is a closed subgroup of $\mathbb{R}$ and it is known that F is ergodic if and only if $E(F) = \mathbb{R}$. The cocycle F is called **regular** if it is cohomologous to a cocycle with values in the set of its essential values, that is if there exists a real measurable function G on Ω such that, for μ-almost all ω, $F(\omega) + G(\omega) - G(T\omega) \in E(F)$. In that case, this new cocycle defines an ergodic extension of $(\Omega, \mathcal{T}, \mu, T)$ by $E(F)$.

PROPOSITION 5.1: *Any regular real cocycle is a good averaging cocycle.*

Proof of Proposition 5.1: If $\theta \in \Lambda_F$, there exists a real measurable function G_θ such that $G_\theta(T\omega) = \theta F(\omega) + G_\theta(\omega) \bmod 1$; the function $(\omega, x) \mapsto e(G_\theta(\omega) - \theta x)$ is T_F-invariant.

If T_F is ergodic, that is if $E(F) = \mathbb{R}$, this implies that $\Lambda_F = \{0\}$.

If $E(F) = a\mathbb{Z}$, with $a \in \mathbb{R}$ and $a \neq 0$, then F is cohomologous to a cocycle F', with values in $a\mathbb{Z}$, which defines an ergodic extension of (Ω, T) by $a\mathbb{Z}$. As in the above case, we conclude that $\Lambda_F = \Lambda_{F'} = a^{-1}\mathbb{Z}$.

In the case where $E(F) = \{0\}$, the fact that the cocycle is regular means that it is a real coboundary, and the result follows from Corollary 2.3. ∎

Remark: If F is ergodic, since $\Lambda_F = \{0\}$, we have moreover that the sequence (1.1) converges to zero for every $\theta \neq 0$, for almost every ω, hence the limit in the mean ergodic theorem along the random sequence $(F^{(n)}(\omega))$ is a.s. the same as the limit in the usual ergodic theorem.

Ergodicity of cocycles, in particular over irrational rotations, is a widely studied subject (e.g., [3], [4], [7], [10], [18], [20], [24], [26], [29, 30]).

Non-ergodic real cocycles (for example, positive ones) have non-trivial associated flows, and the eigenvalues of this flow are connected with our previous discussion. Let us recall the definition of the associated flow, as it appears, for example, in [17] or [2]. On $\Omega \times \mathbb{R}$, the cylindrical flow T_F commutes with the flow $(\tau_t)_{t\in\mathbb{R}}$ defined by translation on the second coordinate: $\tau_t(\omega, x) = (\omega, x+t)$. We consider the σ-algebra $\mathcal{I}$ of T_F-invariant measurable subsets of $\Omega \times \mathbb{R}$ and we consider on $\mathbb{R}$ a probability measure ν equivalent to the Lebesgue measure. The non-singular action of $\mathbb{R}$ given by the flow (τ_t) on the measure space $(\Omega \times \mathbb{R}, \mathcal{I}, \mu\otimes\nu)$ is the flow associated to the cocycle F (this flow is also called "Mackey range").

If the cocycle F is non-negative, the associated flow is isomorphic to the special flow built over the base $(\Omega, \mathcal{T}, \mu, T)$ under the function F. If moreover F is integrable, this flow preserves a finite measure.

The following result is well-known ([17]) but we will give here a short proof because it is central in our analysis.

PROPOSITION 5.2: *The set $2\pi\Lambda_F$ is the set of L^∞-eigenvalues of the associated flow.*

Proof of Proposition 5.2: Let $\theta \in \Lambda_F$. There exists a real measurable function G on Ω such that $\theta F = G - G\circ T \bmod 1$. The function h defined on $\Omega \times \mathbb{R}$ by $h(\omega, x) = e(G(\omega) + \theta x)$ is T_F-invariant, thus it is $\mathcal{I}$-measurable. But we have $h \circ \tau_t = e(\theta t)h$, which implies that $2\pi\theta$ is an eigenvalue of the associated flow.

Conversely, if $2\pi\theta$ is an eigenvalue of the associated flow, then there exists a non-zero $\mathcal{I}$-measurable function h on $\Omega \times \mathbb{R}$ such that $h(\omega, x+t) = e(\theta t)h(\omega, x)$ a.e. The $\mathcal{I}$-measurability means $h(T\omega, x + F(\omega)) = h(\omega, x)$ a.e. These two equations yield $h(\omega, x) = e(\theta F(\omega))h(T\omega, x)$ a.e. For almost all x, this is true for μ-almost all ω, and we can choose x such that it holds μ-a.e. and the function $h(\cdot, x)$ is not μ-almost everywhere zero. It follows $\theta \in \Lambda_F$. ∎

Since the set of eigenvalues of a finite measure-preserving system is at most countable, we deduce:

COROLLARY 5.3: *If the associated flow preserves a finite measure equivalent to* $\mu \otimes \nu$ *on* $(\Omega \times \mathbb{R}, \mathcal{I})$, *then* F *is a good averaging cocycle.*

As a particular case, we obtain that if F is non-negative and integrable then it is a good averaging cocycle. Besides, if the cocycle F is integrable with non-zero mean then it is cohomologous to an integrable cocycle of constant sign (it is shown by Kočergin in [21] among deeper results; see also [9] for a detailed proof). Since two cohomologous cocycles have the same group Λ_F, we have:

COROLLARY 5.4: *Any integrable function with non-zero mean is a good averaging cocycle.*

Note that two cohomologous cocycles also have isomorphic associated flows (and, in particular, if F is integrable with non-zero mean then the associated flow is isomorphic to a special flow preserving a finite measure). More generally, here is another direct consequence of Proposition 5.2.

COROLLARY 5.5: *If two cocycles have isomorphic associated flows and if one of them is a good averaging cocycle, then so is the other one.*

This corollary implies that the good averaging property is preserved by orbit equivalence. It is also preserved by induction: for a measurable subset A of positive measure in Ω, denoting by $n_A(\omega)$ the first return time of $\omega \in A$ in A and by T_A the induced transformation $\omega \mapsto T^{n_A}(\omega)$, we have that the cocycles F on (Ω, T) and $F_A\colon \omega \mapsto F^{(n_A)}(\omega)$ on (A, T_A) have isomorphic associated flows (it can also be directly checked that $\Lambda_{F_A} = \Lambda_F$).

It is proved in [15], [16] (see also [2]) that, up to an isomorphism, every non-singular, conservative, ergodic and free $\mathbb{R}$-action is the associated flow of a recurrent cocycle over an arbitrary ergodic aperiodic measure-preserving dynamical system. So a variety of examples of possible sets Λ_F is known.

Small modifications of the proof of Theorem 4.1 yield the following result which seems to be of general interest in the spectral theory of non-singular transformations. The Polish topology of the eigenvalue group is defined in the same way as the topology of Λ_F from the L^2-topology on eigenfunctions, and when the transformation is conservative, the eigenvalue group cannot carry a full measure (see, e.g., [19]).

THEOREM 5.6: *Let* S *be an ergodic conservative non-singular automorphism or flow of a standard space* $(Y, \mathcal{B}, \nu)$, *and let* Λ *be the eigenvalue group of* S, *equipped with its Polish topology.*

Assume that there is a set of full measure Y' in Y, and an eigenfunction selector $\theta \mapsto g_\theta = g(\theta, \cdot)$ on Λ, where g is a measurable function of modulus one on $\Lambda \times Y'$ such that, for any $x \in Y'$, the function $\theta \mapsto g(\theta, x)$ is continuous on Λ.

Then Λ is countable.

6. Independent identically distributed processes

THEOREM 6.1: *If the process $(F \circ T^n)$, defined on the probability space Ω, is independent, then either*

$F = 0$ a.e. or

there exists $a \neq 0$ such that F takes (a.s.) its values in a discrete subgroup $a\mathbb{Z}$ of the real line and $\Theta_F = a^{-1}\mathbb{Z}$ or

F does not take its values in a discrete subgroup of $\mathbb{R}$ and $\Theta_F = \{0\}$.

COROLLARY 6.2: *If the process $(F \circ T^n)$ is independent, then F is a good averaging cocycle.*

Proof of Theorem 6.1: Fix $\theta \in \Theta_F$. We have

$$\lim_{N\to\infty} \mathbb{E}\left[\frac{1}{N}\sum_{n=0}^{N-1} e(\theta F^{(n)})\right] \neq 0.$$

But

$$\mathbb{E}[e(\theta F^{(n)})] = \mathbb{E}[e(\theta F)]^n$$

so we have $\mathbb{E}[e(\theta F)] = 1$ and thus $\theta F \in \mathbb{Z}$ almost surely. ∎

Remarks:

- If the function F is integrable centered and if the stationary process associated is independent, the random walk $(F^{(n)})$ is recurrent and the cocycle F is regular. (See [37] for a description of the relationships between recurrence of random walks and ergodicity of cocycles.) In this case Proposition 5.1 applies[1].
- The good averaging property can also be proved when the process $(F \circ T^n)$ has some kind of asymptotic independence property. By way of example, using the spectral property of Perron–Frobenius operators, we can show that the group Λ_F is countable in the following cases:
  - (Ω, T) is a subshift of finite type, μ is a Gibbs measure and F is Lipschitz.

1 This remark has been communicated to us by J. Aaronson.

- Ω is an interval, μ is absolutely continuous, T is expanding and F has bounded variation.

7. Gaussian processes

THEOREM 7.1: *If the process $(F \circ T^n)$, defined on the probability space Ω is ergodic, Gaussian and centered, then either the set Θ_F reduces to $\{0\}$ or it is all the real line.*

COROLLARY 7.2: *If the process $(F \circ T^n)$ is Gaussian and ergodic, then F is a good averaging cocycle.*

Proof of Theorem 7.1: Recall that an ergodic Gaussian process is necessary weak-mixing ([8, chap. 14, §2]). This implies that there exists a subset E of $\mathbb{N}$, of asymptotic density one (i.e., $\liminf_{N\to\infty} \frac{1}{N}\#(E \cap [0,N)) = 1$) such that the sequence $(F \circ T^n)_{n\in E}$ goes weakly (in $L^2(\mu)$) to zero.

Let us suppose that $1 \in \Theta_F$, and show that F is a real coboundary (for the general case, replace F by θF). We have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mathbb{E}[e(F^{(n)})] = \mathbb{E}\Big[\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} e(F^{(n)})\Big] \neq 0.$$

Since the random variable $F^{(n)}$ has a Gaussian distribution, we have

$$\mathbb{E}[e(F^{(n)})] = \exp(-2\pi^2\|F^{(n)}\|_2^2).$$

The Cesaro means of the sequence of positive numbers

$$(\exp(-2\pi^2\|F^{(n)}\|_2^2))$$

do not go to zero; hence there exist $\epsilon > 0$ and a subset E' of $\mathbb{N}$, of positive asymptotic density such that, for any $n \in E'$,

$$\exp(-2\pi^2\|F^{(n)}\|_2^2) > \epsilon.$$

The sequence $\big(F^{(n)}\big)_{n\in E'}$ is bounded in $L^2(\mu)$. Now, $E \cap E'$ has positive asymptotic density, hence it is infinite. The sequence $\big(F^{(n)}\big)_{n\in E\cap E'}$ is bounded in $L^2(\mu)$, and we can extract from it a subsequence $(F^{(n)})_{n\in E''}$ which converges weakly to a limit G in L^2. We write

$$F = F^{(n)} - F^{(n)} \circ T + F \circ T^n$$

and we recall that, along the sequence E'', weak-lim $F \circ T^n = 0$. We conclude that $F = G - G \circ T$. ∎

Remark: It is interesting to notice that the positive results obtained in the i.i.d. case and in the Gaussian case for the random mean ergodic theorem do not at all extend to the pointwise convergence. Indeed it is proved in [5] that if (a_n) is a sequence of real numbers linearly independent over $\mathbb{Q}$, then it is universally bad for pointwise convergence, i.e., for any non-trivial probability measure-preserving flow $(X, \mathcal{A}, \nu, (\tau_t)_{t\in\mathbb{R}})$ there exists a bounded measurable function f on X such that (3.3) is not an almost everywhere convergent sequence. In our context this implies that if F has a continuous distribution and if the process $(F \circ T^n)$ is i.i.d. or Gaussian, then almost surely the sequence $(F^{(n)})$ is universally bad for pointwise convergence.

8. The random mean ergodic theorem for weakly mixing systems

In this section we describe a construction of an integrable cocycle F which is not a good averaging cocycle, and which can be chosen with integer values. This construction is universal, in the sense that it takes place in an arbitrary dynamical system, and it allows us to show that the random mean ergodic theorem for mildly mixing transformations (Theorem 3.5) is not true for the class of weakly mixing transformations.

THEOREM 8.1: *For any aperiodic and ergodic probability measure-preserving system $(\Omega, \mathcal{T}, \mu, T)$ there exists an integrable function F such that, for almost every $\omega \in \Omega$, there exists a number θ such that the sequence*

$$\frac{1}{N}\sum_{n=0}^{N-1} e(\theta F^{(n)}(\omega))$$

does not converge.

THEOREM 8.2: *For any aperiodic and ergodic probability measure-preserving system $(\Omega, \mathcal{T}, \mu, T)$ there exists an integrable function F such that, for almost every $\omega \in \Omega$, there exists a weakly mixing probability measure-preserving system $(X, \mathcal{A}, \nu, \tau)$ and $f \in L^2(\mu)$ such that the sequence*

$$\left(\frac{1}{N}\sum_{n=0}^{N-1} f \circ \tau^{F^{(n)}(\omega)}\right)$$

does not converge in $L^2(\mu)$.

In the proofs of these theorems we will use a lemma, which says that if a sequence (L_n) of numbers grows sufficiently fast, then the random variables

$\theta \mapsto L_n\theta$, defined on the unit interval with uniform probability, are almost independent. A proof of this lemma is postponed to the end of the section. For any real number x, we denote $\|x\| = \frac{1}{2\pi} d(x,\mathbb{Z})$ where $d(x,\mathbb{Z})$ is the usual distance from x to $\mathbb{Z}$.

LEMMA 8.3: *For any sequence $(\epsilon_n)_{n\geq 0}$ of positive real numbers, there exists a sequence $(L_n)_{n\geq 0}$ of integers such that, for any sequence $(\alpha_n)_{n\geq 0}$ of real numbers, the set of real numbers θ satisfying*

$$\|L_n\theta - \alpha_n\| \leq \epsilon_n \quad (n \geq 0),$$

contains a Cantor set, hence an uncountable compact subset of $\mathbb{R}$.

Proof of Theorem 8.1: We consider a triangular array $(L_{n,i}), n \geq 1, 0 \leq i < 2^n$, of positive integers such that for any $E \subset \{(n,i) \in \mathbb{N}^2 : n \geq 1, 0 \leq i < 2^n\}$, there exists $\theta \in \mathbb{R}$ satisfying

$$(n,i) \notin E \implies \|L_{n,i}\theta\| < 2^{-n},$$
$$(n,i) \in E \implies \|L_{n,i}\theta - \frac{1}{4}\| < 2^{-n}.$$

The existence of such an array is ensured by Lemma 8.3.

For each $n \geq 1$, we consider a Rokhlin tower of height $4^n h_n$ which covers all the space Ω except a part of measure less than 2^{-n}. The base of this tower is a set $A_n \in \mathcal{T}$ such that $T^m A_n, 0 \leq m < 4^n h_n$, are pairwise disjoint and $\mu\left(\bigcup_{0\leq m<4^n h_n} T^m A_n\right) > 1 - 2^{-n}$. The sequence of positive integers (h_n) is assumed to satisfy the following conditions:

C1: $h_n \geq 2^n$;

C2: $12\sum_{n\geq 1}\sum_{\ell>n}(h_n/h_\ell) < 1$;

C3: $\sum_{n\geq 1}(1/h_n)\max_{0\leq i<2^n} L_{n,i} < +\infty$;

C4: there exists $\Omega_n \in \mathcal{T}$ such that $\mu(\Omega_n) > 1 - 2^{-n}$ and, for all $\omega \in \Omega_n$, for all $\ell = 1, \ldots, n-1$, for each set A which is a union of levels of the tower $\{T^m A_\ell : 0 \leq m < 4^\ell h_\ell\}$ and for all $N > h_n$,

$$\left|\frac{1}{N}\sum_{t=0}^{N-1} \chi_A(T^t\omega) - \mu(A)\right| < 2^{-n}.$$

The existence of the sequence of heights (h_n) and corresponding towers is ensured by an inductive construction (to have C4 we use the pointwise ergodic theorem and the Egorov theorem).

We denote $\Omega^1 := \bigcap_{n\geq 7}\Omega_n$ and we have $\mu(\Omega^1) > \frac{63}{64}$.

We introduce a new numbering of the tower's levels:

$$B_n(i,j,k) := T^m A_n$$

if

$$m = 2^n h_n i + h_n j + k,$$

where $0 \le i, j < 2^n$ and $0 \le k < h_n$.

We define a function F_n on Ω by

$$\begin{aligned} F_n &= L_{n,i} && \text{on } B_n(i,2j,0) \text{ and } B_n(i,2j+1,h_n-1) \\ & && \text{for } 0 \le i < 2^n,\ 0 \le j < 2^{n-1}, \\ F_n &= -L_{n,i} && \text{on } B_n(i,2j+1,0) \text{ and } B_n(i,2j,h_n-1) \\ & && \text{for } 0 \le i < 2^n,\ 0 \le j < 2^{n-1}, \\ F_n &= 0 && \text{elsewhere.} \end{aligned}$$

This function F_n is an additive coboundary; indeed,

$$F_n = G_n \circ T - G_n,$$

where G_n is defined by

$$\begin{aligned} G_n &= L_{n,i} && \text{on } B_n(i,2j,k) && \text{for } 0 \le i < 2^n,\ 0 \le j < 2^{n-1},\ 0 < k < h_n, \\ G_n &= -L_{n,i} && \text{on } B_n(i,2j+1,k) && \text{for } 0 \le i < 2^n,\ 0 \le j < 2^{n-1},\ 0 < k < h_n, \\ G_n &= 0 && \text{elsewhere.} \end{aligned}$$

We have

$$\mathbb{E}[|F_n|] = \sum_{i<2^n} 2^{n+1} L_{n,i}\mu(A_n) \le \sum_{i<2^n} 2^{n+1} L_{n,i}\frac{1}{4^n h_n}$$

and C3 implies that

$$\sum_{n\ge1} \mathbb{E}[|F_n|] < +\infty.$$

Hence the series $\sum_{n\ge7} F_n$ converges in L^1 and also almost everywhere. Denote its sum by F.

We denote by Ω_n' a part of the tower $\bigcup_m T^m A_n$ (we will remove some levels). To be precise, we define

$$E_n := \{(i,j,k) : 0 \le i < 2^n, 0 \le j < 2^n - 2, 0 < k < h_n\}$$

and

$$\Omega_n' := \bigcup_{(i,j,k)\in E_n} B_n(i,j,k).$$

We have $\#E_n^c \le 2^{n+1}h_n + 4^n$, so using C1,

$$\mu(\Omega_n') \ge 1 - 2^{-n} - \#E_n^c \cdot \mu(A_n) \ge 1 - 2^{-n+2}.$$

We define $\Omega^2 := \bigcap_{n\geq 7}\Omega'_n$ and we have $\mu(\Omega^2) > \frac{15}{16}$. Let $\omega \in \Omega^2$. For each $n \geq 7$ there exists a unique $(i_\omega, j_\omega, k_\omega) = (i_{\omega,n}, j_{\omega,n}, k_{\omega,n})$ such that

$$\omega \in B_n(i_\omega, j_\omega, k_\omega).$$

Putting $E := \{(n, i_{\omega,n}) : n \in \mathbb{N}\}$ we obtain that there exists $\theta = \theta(\omega) \in \mathbb{R}$ such that, for each $n \geq 7$,

$$\begin{cases} i \neq i_\omega \Longrightarrow & \|L_{n,i}\theta\| < 2^{-n} \\ & \|L_{n,i_\omega}\theta - \frac{1}{4}\| < 2^{-n}. \end{cases} \tag{8.1}$$

Let us suppose that j_ω is even. Since $\omega \in \Omega^2$, $k_\omega > 0$ and hence $G_n(\omega) = L_{n,i_\omega}$. There exists an integer $c = c(n,\omega)$ such that $0 < c < h_n$ and

$$\begin{array}{lll} G_n(T^t\omega) & = L_{n,i_\omega} & \text{for } 0 \leq t < c, \\ G_n(T^t\omega) & = -L_{n,i_\omega} & \text{for } c < t < c + h_n, \\ G_n(T^t\omega) & = L_{n,i_\omega} & \text{for } c + h_n < t < c + 2h_n. \end{array}$$

(Here we used the fact that $\omega \in \Omega^2$ implies that $j_\omega \leq 2^n - 3$.)

If j_ω is odd, we have the same property providing that we replace $L_{n,i}$ by -$L_{n,i}$.

Note that if $G_n(\omega') = \pm L_{n,i_\omega}$ then

$$|e(\theta G_n(\omega')) - e(\pm\frac{1}{4})| < 2^{-n},$$

and if $G_n(\omega') \neq \pm L_{n,i_\omega}$ then

$$|e(\theta G_n(\omega')) - 1| < 2^{-n}.$$

We have

$$\begin{aligned} \left|\sum_{t=0}^{c+h_n-1} e(\theta G_n(T^t\omega)) - e\Big(-\frac{1}{4}\Big)\right| &\leq \sum_{t=0}^{c+h_n-1} \left|e(\theta G_n(T^t\omega)) - e\Big(-\frac{1}{4}\Big)\right| \\ &\leq \Big(\sum_{0\leq t<c} 2\Big) + 1 + \sum_{t=c+1}^{c+h_n-1} 2^{-n} \\ &\leq 2c + 1 + (h_n - 1)2^{-n} \leq 2c + (c + h_n)2^{1-n} \end{aligned}$$

and, by similar arguments,

$$\begin{aligned} \left|\sum_{t=0}^{c+2h_n-1} e(\theta G_n(T^t\omega)) - e\Big(\frac{1}{4}\Big)\right| &\leq c2^{-n} + 1 + (h_n - 1)2 + 1 + (h_n - 1)2^{-n} \\ &\leq 2h_n + (c + 2h_n)2^{-n}. \end{aligned}$$

These inequalities imply that

$$\left|\frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1} e(\theta G_n(T^t\omega)) - \frac{1}{c+2h_n}\sum_{t=0}^{c+2h_n-1} e(\theta G_n(T^t\omega))\right| \geq$$

$$2-\frac{2c}{c+h_n}-2^{1-n}-\frac{2h_n}{c+2h_n}-2^{-n}=\frac{2h_n^2}{(c+h_n)(c+2h_n)}-2^{1-n}-2^{-n}$$

$$>\frac{1}{3}-2^{1-n}-2^{-n},$$

because $c < h_n$. Hence (remembering that $n \geq 7$) we obtain

$$\left|\frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1} e(\theta G_n(T^t\omega)) - \frac{1}{c+2h_n}\sum_{t=0}^{c+2h_n-1} e(\theta G_n(T^t\omega))\right| \geq \frac{1}{4}. \tag{8.2}$$

We have

$$\mu(\Omega^1 \cap \Omega^2) > \frac{7}{8} \tag{8.3}$$

and let us suppose now that $\omega \in \Omega^1 \cap \Omega^2$. For positive integers ℓ we define

$$V_{\ell,\omega} := \{\omega' \in \Omega : G_\ell(\omega') \neq L_{\ell, \imath_{\omega,\ell}}\}.$$

Since $\omega \in \Omega^1$, for any $N > h_n$ and any $\ell < n$,

$$\frac{1}{N}\sum_{t=0}^{N-1} \chi_{V_{\ell,\omega}}(T^t\omega) \leq \mu(V_{\ell,\omega}) + 2^{-\ell} \leq 2^{1-\ell}.$$

Indeed, $\Omega\backslash V_{\ell,\omega}$ is a union of levels of the ℓ's tower and we apply C4. Furthermore, for any $\omega' \notin V_{\ell,\omega}$, we have $\|\theta G_\ell(\omega')\| < 2^{-\ell}$. This will allow us to replace G_n by $\sum_{\ell\leq n} G_\ell$ in (8.2). Indeed,

$$\left|\frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1} e\Big(\theta\sum_{\ell=7}^{n} G_\ell(T^t\omega)\Big) - \frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1} e(\theta G_n(T^t\omega))\right|$$

$$\leq \frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1}\sum_{\ell=7}^{n-1} |e(\theta G_\ell(T^t\omega)) - 1|$$

$$\leq \frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1}\sum_{\ell=7}^{n-1} (2\chi_{V_{\ell,\omega}}(T^t\omega) + 2^{-\ell})$$

$$\leq \sum_{\ell=7}^{n-1}(2^{2-\ell}+2^{-\ell}) \leq \frac{1}{16}+\frac{1}{64}.$$

The same estimate holds if we replace $c+h_n$ by $c+2h_n$. Therefore, in view of (8.2),

$$\left|\frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1} e\Big(\theta\sum_{\ell=7}^{n} G_\ell(T^t\omega)\Big) - \frac{1}{c+2h_n}\sum_{t=0}^{c+2h_n-1} e\Big(\theta\sum_{\ell=7}^{n} G_\ell(T^t\omega)\Big)\right| \geq \frac{1}{16}.$$

Let us show now that we can choose ω such that, for all $t < c+2h_n$ and all $\ell > n$, $F_\ell(\omega) = 0$. We have

$$\mu(F_\ell \neq 0) \leq \frac{2}{h_\ell}$$

and, for any $n \geq 1$,

$$\mu\Big(\bigcup_{\ell>n}\{F_\ell \neq 0\}\Big) \leq \sum_{\ell>n}\frac{2}{h_\ell}.$$

We denote by Ω^3 the set of ω's such that, for any $n \geq 1$, for $\ell > n$ and $0 \leq t < 3h_n$, $F_\ell(T^t\omega) = 0$. We have

$$\mu(\Omega^3) \geq 1 - \sum_{n\geq 1} 3h_n \sum_{\ell>n}\frac{2}{h_\ell}.$$

It follows from C2 and (8.3) that

$$\mu(\Omega^1 \cap \Omega^2 \cap \Omega^3) > 0.$$

If $\omega \in \Omega^3$, $n \geq 7$ and $N > 3h_n$, then

$$\begin{aligned}\frac{1}{N}\sum_{t=0}^{N-1} e(\theta F^{(t)}(\omega)) &= \frac{1}{N}\sum_{t=0}^{N-1} e\Big(\theta\sum_{\ell=7}^{n} F_\ell^{(t)}(\omega)\Big)\\ &= e\Big(-\theta\sum_{\ell=7}^{n} G_\ell(\omega)\Big)\frac{1}{N}\sum_{t=0}^{N-1} e\Big(\theta\sum_{\ell=7}^{n} G_\ell(T^t\omega)\Big).\end{aligned}$$

In view of (8.2) this implies that for any $\omega \in \Omega^1 \cap \Omega^2 \cap \Omega^3$ and any $n \geq 7$,

$$\left|\frac{1}{c+h_n}\sum_{t=0}^{c+h_n-1} e(\theta F^{(t)}(\omega)) - \frac{1}{c+2h_n}\sum_{t=0}^{c+2h_n-1} e(\theta F^{(t)}(\omega))\right| \geq \frac{1}{16}. \tag{8.5}$$

Consequently, for any $\omega \in \Omega^1 \cap \Omega^2 \cap \Omega^3$ the sequence

$$\frac{1}{N}\sum_{t=0}^{N-1} e(\theta F^{(t)}(\omega)) \tag{$*$}$$

does not converge. But it is clear that the set of ω's such that $(*)$ does not converge is T-invariant. By ergodicity of T, the result follows. ∎

Proof of Theorem 8.2: This proof requires only a small modification of the proof of Theorem 8.1. By Lemma 8.3 the family $(L_{n,i}), n \geq 1, 0 \leq i < 2^n$ can be chosen so that, for each $\omega \in \Omega^2$, conditions (8.1) are satisfied by all the numbers θ in an uncountable compact subset K_ω of $\mathbb{T}$. We fix $\omega \in \Omega^1 \cap \Omega^2 \cap \Omega^3$. For any $n \geq 7$ and any $\theta \in K_\omega$ the condition (8.5) is fulfilled. Let us denote $K'_\omega := K_\omega \cup (-K_\omega)$. For any $n \geq 7$ and any $\theta \in K'_\omega$ the condition (8.5) is fulfilled. There exists a continuous Borel probability measure σ_ω whose support is contained in K'_ω and such that for any measurable subset A of $\mathbb{T}$, $\sigma_\omega(A) = \sigma_\omega(-A)$. We consider now the Gaussian dynamical system $(X, \mathcal{A}, \nu, \tau)$ of spectral measure σ_ω (see [8]). The spectral measure being continuous this Gaussian dynamical system is weakly mixing, and in the Gaussian subspace of $L^2(\nu)$ we can find f whose spectral measure σ_f equals σ_ω. By (8.5) and the spectral theorem we have, for any $n \geq 7$,

$$\begin{aligned}
&\left\| \frac{1}{c+h_n} \sum_{t=0}^{c+h_n-1} f \circ T^{F^{(t)}(\omega)} - \frac{1}{c+2h_n} \sum_{t=0}^{c+2h_n-1} f \circ T^{F^{(t)}(\omega)} \right\|_2^2 \\
&= \int_{\mathbb{T}} \left| \frac{1}{c+h_n} \sum_{t=0}^{c+h_n-1} e(\theta F^{(t)}(\omega)) - \frac{1}{c+2h_n} \sum_{t=0}^{c+2h_n-1} e(\theta F^{(t)}(\omega)) \right|^2 d\sigma_\omega(\theta) \\
&\geq \inf_{\theta \in K_\omega} \left| \frac{1}{c+h_n} \sum_{t=0}^{c+h_n-1} e(\theta F^{(t)}(\omega)) - \frac{1}{c+2h_n} \sum_{t=0}^{c+2h_n-1} e(\theta F^{(t)}(\omega)) \right|^2 \geq \frac{1}{16}.
\end{aligned}$$

Consequently, the sequence

$$\frac{1}{N} \sum_{t=0}^{N-1} f \circ T^{F^{(t)}(\omega)}$$

does not converge in $L^2(\nu)$. ∎

Proof of Lemma 8.3: We suppose that for any $n \geq 0$, $\epsilon_n < \frac{1}{2}$ and we fix a sequence (L_n) such that

$$L_0 = 1 \quad \text{and} \quad L_n \geq \frac{3L_{n-1}}{2\epsilon_{n-1}} \quad (n \geq 1).$$

Let us prove by induction that there exists a decreasing sequence of compact subsets $(K_n)_{n\geq 0}$ in $\mathbb{R}$ such that each K_n is a union of 2^n intervals pairwise disjoint of length $2\epsilon_n / L_n$ and satisfies:

$$\theta \in K_n \Longrightarrow \|L_k\theta - \alpha_k\| \leq \epsilon_k, \quad k = 0, 1, \ldots, n.$$

We define $K_0 := [\alpha_0 - \epsilon_0, \alpha_0 + \epsilon_0]$. Now we fix $n \geq 0$ and we suppose that $K_0 \supset K_1 \supset \cdots K_n$ are given and satisfy the announced property. Let I be an interval of length $2\epsilon_n/L_n$. Since $L_{n+1} \geq 3L_n/2\epsilon_n$, $\{L_{n+1}\theta : \theta \in I\}$ contains an interval of length 3 and $\{\theta \in I : \|L_{n+1}\theta - \alpha_{n+1}\| \leq \epsilon_{n+1}\}$ contains two disjoint closed intervals of length $2\epsilon_{n+1}/L_{n+1}$. This construction can be done in each of the 2^n intervals whose union is K_n, and we obtain 2^{n+1} disjoint closed intervals of length $2\epsilon_{n+1}/L_{n+1}$ whose union is denote K_{n+1}. If $\theta \in K_{n+1}$ and $k \leq n+1$ then $\|L_k\theta - \alpha_k\| \leq \epsilon_k$. This completes the induction construction.

The set $K := \bigcap_n K_n$ is a Cantor set and, if $\theta \in K$, then for any $n \geq 0$, $\|L_n\theta - \alpha_n\| \leq \epsilon_n$. ∎

9. Appendix

The aim of this appendix is the justification of Remark 3.2. A real sequence $(a_n)_{n\geq 0}$ is called **ergodic** if it is a good averaging sequence and if the limit of (3.1) is zero for any non-zero number θ; so, by Weyl's criterion, the real sequence (a_n) is ergodic if and only if its normal set is $\mathbb{R}\backslash\{0\}$, that is to say, for any non-zero number θ, the sequence $(a_n\theta)$ is uniformly distributed mod 1. The sequence $a_n = \sqrt{n}$ is an example of an ergodic sequence.

The arguments of the proof of Proposition 3.1 lead directly to the following result.

Proposition 9.1: *Let (a_n) be an ergodic sequence, $(X, \mathcal{A}, \nu, (\tau_t)_{t\in\mathbb{R}})$ a measure preserving flow and $p \in [1, +\infty)$. Denote by $\mathcal{I}$ the σ-algebra of (τ_t)-invariant elements of $\mathcal{A}$. Then:*

(i) *for any $f \in L^p(\nu)$, the limit of sequence (3.3) equals the conditional expectation $\mathbb{E}_\nu[f|\mathcal{I}]$,*

(ii) *for any $A \in \mathcal{A}$, $\nu(A) > 0$, there exists $n \in \mathbb{N}$ such that $\nu(A \cap \tau_{-a_n}A) > 0$.*

Following ideas that Furstenberg developed in the integer-valued sequence case we show now that the recurrence property stated in the last Proposition gives rise to a result in combinatorial number theory. Let us recall the definition of Banach density. If E is a subset of $\mathbb{R}$, its **upper Banach density** is defined as

$$\limsup_{N\to\infty} \sup_M \frac{1}{N}|E \cap [M, M+N)|$$

where $|\cdot|$ denotes the inner Lebesgue measure.

Proposition 9.2: *Let R be a Poincaré recurrence set of real numbers, i.e. a subset of $\mathbb{R}$ such that for any probability measure-preserving flow $(X, \mathcal{A}, \nu, (\tau_t)_{t\in\mathbb{R}})$*

and any $A \in \mathcal{A}$ with $\nu(A) > 0$, there exists $r \in R$ such that $\nu(A \cap \tau_{-r}A) > 0$. Then for any $E \subset \mathbb{R}$ of non-zero upper Banach density, we have

$$\overline{E - E} \cap R \neq \emptyset.$$

Proof of Proposition 9.2: In the definition of Banach density we have used the inner measure, so it is easy to see that if E has non-zero upper Banach density, then E contains a closed set of non-zero upper Banach density. Therefore we can suppose, without loss of generality, that E is closed. We define $f\colon \mathbb{R} \longrightarrow [0,1]$ by

$$f(x) = (1 - d(x, E))^+$$

(where $d(x, E)$ denotes the usual distance between x and E).

On the space $\mathcal{C}$ of continuous functions from $\mathbb{R}$ into $[0,1]$, equipped with the topology of uniform convergence on compact subsets, we denote by τ the action of $\mathbb{R}$ by translation: if $g \in \mathcal{C}$ then $(\tau_t g)(x) = g(x+t)$.

We denote by X the closure of the orbit of f, i.e.,

$$X := \overline{\{\tau_t(f) : t \in \mathbb{R}\}} = \overline{\{f(\cdot + t) : t \in \mathbb{R}\}}.$$

Since f is uniformly continuous, by the Arzela–Ascoli theorem, X is a compact metric space.

By the standing assumption of non-zero density, we know that there exist $\delta > 0$ and two real sequences (α_n), (β_n) such that $\lim_{n\to\infty}(\beta_n - \alpha_n) = +\infty$ and, for $n \geq 1$,

$$\frac{1}{\beta_n - \alpha_n} |E \cap [\alpha_n, \beta_n]| > \delta.$$

We define a sequence (ν_n) of probability measures on X by

$$\nu_n(\Phi) := \frac{1}{\beta_n - \alpha_n} \int_{\alpha_n}^{\beta_n} \Phi(\tau_t(f))dt$$

for any continuous function Φ on X. We extract from (ν_n) a weakly convergent subsequence (ν_{n_k}) and we denote by ν its limit. From the fact that $\beta_n - \alpha_n \to +\infty$, we deduce that ν is τ-invariant. We want now to apply the recurrence property along R to the probability measure-preserving flow (X, ν, τ).

Fix $\epsilon > 0$. Sets $A := \{g \in X : g(0) = 1\}$ and $\{g \in X : g(0) \leq 1 - \epsilon\}$ are closed disjoint subsets of X. Hence there exists a continuous function Φ from X into $[0,1]$ such that

$$\Phi(g) = \begin{cases} 1 & \text{if } g(0) = 1; \\ 0 & \text{if } g(0) \leq 1 - \epsilon. \end{cases}$$

We have $\nu(\Phi) = \lim_{n\to\infty} \nu_{n_k}(\Phi)$ and

$$\begin{aligned}\nu_n(\Phi) &= \frac{1}{\beta_n - \alpha_n}\int_{\alpha_n}^{\beta_n} \Phi(f(\cdot + t))dt \\ &\geq \frac{1}{\beta_n - \alpha_n}\int_{\alpha_n}^{\beta_n} \mathbf{1}_E(t)dt,\end{aligned}$$

because, if $t \in E$, then $f(0+t) = 1$ and $\Phi(f(\cdot + t)) = 1$. Therefore, for each n, $\nu_n(\Phi) > \delta$ so $\nu(\Phi) \geq \delta$. This implies that $\nu(\{g \in X : g(0) \geq 1 - \epsilon\}) \geq \delta$ and, letting ϵ go to zero, we obtain $\nu(A) \geq \delta$. Now we can use the recurrence property in the dynamical system (X, ν, τ), and we have: there exists $r \in R$ such that $\nu(A \cap \tau_{-r}(A)) > 0$, hence there exists $r \in R$ and $g \in X$ such that $g(0) = 1$ and $g(r) = 1$. By definition of X such a function g is a limit of a sequence $(f(\cdot + t_n))$. Thus we have proved the existence of $r \in R$ and of a sequence (t_n) of real numbers such that

$$\lim_{n\to\infty} f(t_n) = \lim_{n\to\infty} f(r + t_n) = 1.$$

By the definition of f this means exactly that $r \in \overline{E - E}$. ∎

References

[1] J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Mathematical Surveys and Monographs **50**, American Mathematical Society, Providence, RI, 1997.

[2] J. Aaronson, T. Hamachi and K. Schmidt, *Associated actions and uniqueness of cocycles*, in *Algorithms, Fractals and Dynamics* (Y. Takahashi, ed.), Plenum Press, New York, 1995, pp. 1–25.

[3] J. Aaronson, M. Lemańczyk, C. Mauduit and H. Nakada, *Koksma inequality and group extensions of Kronecker transformations*, in *Algorithms, Fractals and Dynamics* (Y. Takahashi, ed.), Plenum Press, New York, 1995, pp. 27–50.

[4] L. Bagget and K. Merrill, *Smooth cocycles for an irrational rotation*, Israel Journal of Mathematics **79** (1992), 281–288.

[5] V. Bergelson, M. Boshernitzan and J. Bourgain, *Some results on nonlinear recurrence*, Journal d'Analyse Mathématique **62** (1994), 29–46.

[6] J. P. Conze, *Remarques sur les transformations cylindriques et les équations fonctionnelles*, Publications du Séminaire de Probabilités, Université Rennes, 1976.

[7] J. P. Conze, *Ergodicité d'un flot cylindrique*, Bulletin de la Société Mathématique de France **108** (1980), 441–456.

[8] I. Cornfeld, S. V. Fomin and Ya. G. Sinai, *Ergodic Theory*, Springer, New York, 1982.

[9] J. M. Derrien, *Sur l'existence de cohomologues réguliers pour les cocycles intégrables*, Séminaires de Probabilités de Rennes, 1995.

[10] K. Frączek, *On ergodicity of some cylinder flows*, Fundamenta Mathematicae **163** (2000), 117–130.

[11] H. Furstenberg, *Strict ergodicity and transformations of the torus*, American Journal of Mathematics **83** (1961), 573–601.

[12] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[13] H. Furstenberg and B. Weiss, *The finite multipliers of infinite ergodic transformations*, Lecture Notes in Mathematics **668**, Springer-Verlag, Berlin, 1978, pp. 127–132.

[14] C. Gamet and D. Schneider, *Théorèmes ergodiques multidimensionnels et suites aléatoires universellement représentatives en moyenne*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **33** (1997), 269–282.

[15] V. I. Golodets and S. D. Sinel'shchikov, *Existence and uniqueness of cocycles of an ergodic automorphism with dense ranges in amenable groups*, Preprint, Institute of Low Temperature Physics and Engineerings, **19–83**, Kharkhov, 1983.

[16] V. I. Golodets and S. D. Sinel'shchikov, *Amenable ergodic actions of groups and images of cocycles*, Soviet Mathematics Doklady **41** (1990), 523–525.

[17] T. Hamachi and M. Osikawa, *Ergodic groups of automorphisms and Krieger's Theorems*, Seminar on Mathematical Sciences, Keio University **3**, 1981.

[18] P. Hellekalek and G. Larcher, *On ergodicity of a class of skew products*, Israel Journal of Mathematics **54** (1986), 301–306.

[19] B. Host, J. F. Méla and F. Parreau, *Non-singular transformations and spectral analysis of measures*, Bulletin de la Société Mathématique de France **119** (1991), 33–90.

[20] A. Iwanik, *Ergodicity of piecewise smooth cocycles over toral rotations*, Fundamenta Mathematicae **157** (1998), 235–244.

[21] A. V. Kočergin, *On the homology of functions over dynamical systems*, Doklady Akademii Nauk SSSR **231** (1976); transl.: Soviet Mathematics Doklady **17** (1976), 1637–1641.

[22] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974.

[23] M. Lacey, K. Petersen, D. Rudolph and M. Wierdl, *Random ergodic theorems with universally representative sequences*, Annales de l'Institut Henri Poincaré, Probabilités et Statistiques **30** (1994), 353–395.

[24] M. Lemańczyk, F. Parreau and D. Volný, *Ergodic properties of real cocycles and pseudo-homogeneous Banach spaces*, Transactions of the American Mathematical Society **348** (1996), 4919–4938.

[25] C. Moore and K. Schmidt, *Coboundaries and homomorphisms for non-singular actions and a problem of H. Helson*, Proceedings of the London Mathematical Society **40** (1980), 443–475.

[26] I. Oren, *Ergodicity of cylinder flows arising from irregularities of distribution*, Israel Journal of Mathematics **44** (1983), 127–138.

[27] M. Osikawa, *Point spectra of non-singular flows*, Publications of the Research Institute for Mathematical Sciences of Kyoto University **13** (1977), 167–172.

[28] J. C. Oxtoby, *Measure and Category*, Graduate Texts in Mathematics **2**, Springer-Verlag, New York, 1971.

[29] D. Pask, *Skew products over the irrational rotation*, Israel Journal of Mathematics **69** (1990), 65–74.

[30] D. Pask, *Ergodicity of certain cylinder flows*, Israel Journal of Mathematics **76** (1991), 129–152.

[31] M. Queffelec, *Substitution dynamical systems, Spectral analysis*, Lecture Notes in Mathematics **1294**, Springer, Berlin, 1987.

[32] J. Rosenblatt, *Norm convergence in ergodic theory and the behaviour of Fourier transforms*, Canadian Journal of Mathematics **46** (1994), 184–199.

[33] J. Rosenblatt and M. Wierdl, *Pointwise ergodic theorems via harmonic analysis*, in *Ergodic Theory and its Connections with Harmonic Analysis* (K. Petersen and I. Salama, eds.), London Mathematical Society Lecture Note Series **205**, Cambridge University Press, 1995.

[34] H. P. Rosenthal, *A characterization of Banach spaces containing ℓ^1*, Proceedings of the National Academy of Sciences of the United States of America **71** (1974), 2411–2413.

[35] W. Rudin, *Fourier Analysis on Groups*, Interscience Publishers, Wiley, New York, 1962.

[36] K. Schmidt, *Cocycles of Ergodic Transformation Groups*, Lecture Notes in Mathematics **1**, MacMillan Co. of India, 1977.

[37] R. J. Zimmer, *Random walks on compact groups and the existence of cocycles*, Israel Journal of Mathematics **26** (1977), 214–220.