

The return times theorem, auto-correlation and sequences with an empty Fourier–Bohr spectrum

by

MATAN TAL

Abstract. This paper explores the proof by J. Bourgain, H. Furstenberg, Y. Katznelson and D. S. Ornstein (1989) of their return times theorem and lights a corner in it regarding the role of auto-correlation. As for pointwise convergence, this was already observed by E. Lesigne, C. Mauduit and B. Mosse (1994), and here we write down the proof. This yields a more intrinsic characterization of the sequences satisfying the pointwise theorem. Then we obtain a characterization linked to auto-correlation also for sequences satisfying the mean theorem – by that theorem those were already known to be exactly the sequences with an empty Fourier–Bohr spectrum. Some examples are provided regarding generic sequences satisfying the pointwise theorem for which the measure on the circle that the auto-correlation function represents (by Fourier transform) is not atomless, and also regarding the existence of sequences that satisfy the mean theorem but not the pointwise one.

1. Introduction. Throughout, we work in $\mathbb{C}^{\mathbb{N}}$. We denote its shift map by σ , and the function that gives the value at the first index by $\pi_1 : \mathbb{C}^{\mathbb{N}} \rightarrow \mathbb{C}$.

Our object of interest is sequences $x = (x_n)_{n \in \mathbb{N}} \in \mathbb{C}^{\mathbb{N}}$ that satisfy

$$\frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \xrightarrow{N \rightarrow \infty} 0$$

almost surely for any complex-valued stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$ of finite variance – we shall call them *pointwise cancellation sequences* – and also sequences for which

$$\left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0$$

for every such process, which we shall call *mean cancellation sequences*.

2020 *Mathematics Subject Classification*: Primary 37A46; Secondary 60G10.

Key words and phrases: return times, correlation, auto-correlation, cancellation, Fourier–Bohr.

Received 17 January 2025; revised 18 December 2025.

Published online 6 May 2026.

By a short argument (in the proof of Theorem 2) we shall see that x is a mean cancellation sequence if and only if $\left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0$ for every *bounded* complex-valued stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$. So for x satisfying $\sup_{N \in \mathbb{N}} \frac{|x_1| + \dots + |x_N|}{N} < \infty$ (we shall anyway be concerned only with such x), pointwise cancellation implies mean cancellation by Lebesgue's dominated convergence theorem. In Section 5 we shall see that there exist bounded sequences that are mean cancellation sequences but not pointwise ones.

In the opening paragraph of [2] the following result is proved by a direct application of the spectral theorem to the unitary shift map $(Y_n \mapsto Y_{n+1})$ and the vector Y_1 .

THEOREM 0. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence satisfying $\sup_{N \in \mathbb{N}} \frac{|x_1| + \dots + |x_N|}{N} < \infty$. The following conditions are equivalent:*

- (i) $\frac{1}{T} \sum_{n=1}^{T-1} x_n z^n \xrightarrow{T \rightarrow \infty} 0$ for all $z \in \mathbb{C}$ on the unit circle.
- (ii) $x = (x_n)_{n \in \mathbb{N}}$ is a mean cancellation sequence.

(The implication (ii) \Rightarrow (i) is not mentioned in [2] but its proof is an immediate conclusion from the fact that for every z_0 on the unit circle there exist stationary processes $(Y_n(\omega))_{n \in \mathbb{Z}}$ for which z_0 is in the point spectrum of Y_1 relative to the unitary shift map – for example, ones originating from circle rotations.)

A sequence satisfying (i) is customarily said to have *empty Fourier–Bohr spectrum*. This is a special case of sequences x for which $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{n=1}^{T-1} x_n z^n$ exists for all $z \in \mathbb{C}$ on the unit circle, which are customarily called *Hartman almost-periodic sequences* (for instance see [3]). In [5] J.-P. Kahane proved that for this last type of sequences there is at most a countable set of z for which the limit is not zero. Theorem 0 gives another characterization of sequences with empty Fourier–Bohr spectrum.

The main result in [2] is the following theorem.

THEOREM 1'. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence generic relative to a σ -invariant Borel probability measure μ on $\mathbb{C}^{\mathbb{N}}$ with $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$. The following conditions are equivalent:*

- (i) $\frac{1}{T} \sum_{n=1}^{T-1} x_n \bar{\xi}_n \xrightarrow{T \rightarrow \infty} 0$ for μ -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$.
- (ii) $x = (x_n)_{n \in \mathbb{N}}$ is a pointwise cancellation sequence.

REMARK. This was actually formulated in [2] for x bounded, and the cancellation property referred there to processes of finite expectation. The formulation given above fits better the character of our paper and requires only a straightforward modification of the original proof. In fact, also in [2] the proof that (i) \Rightarrow (ii) is written explicitly only for processes $(Y_n(\omega))_{n \in \mathbb{Z}}$ of

finite variance. The deduction from that to finite expectation requires a step which is implicit in that proof, and its details can be found in [7, Lemmas 5 and 8].

Unlike in Theorem 0, condition (i) of Theorem 1' refers to the μ for which x is generic. It is observed in [6] (cf. also [1, p. 15]) that the proof of Theorem 1' can actually yield an equivalent condition to (ii) that, unlike (i), does not refer to μ and hence is a more intrinsic characterization of such x . It involves an expression linked to the auto-correlation of x . This is Theorem 1 in the present paper.

THEOREM 1. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence generic relative to a σ -invariant Borel probability measure μ on $\mathbb{C}^{\mathbb{N}}$ with $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$. The following conditions are equivalent:*

- (i) $\frac{1}{T} \sum_{n=1}^{T-1} x_n \overline{x_{n+\tau}} \xrightarrow{T \rightarrow \infty} 0$ for μ -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$.
- (ii) For any $\varepsilon > 0$ there exists $N' > 0$ such that for any $N'' > N'$ there exists $T' > 0$ for which

$$\left| \frac{\{0 < \tau \leq T : \exists N' \leq N \leq N'' \mid \frac{1}{N} \sum_{n=1}^N x_{n+\tau} \overline{x_n} \geq \varepsilon\}}{T} \right| < \varepsilon \quad \text{for all } T > T'.$$

- (iii) $x = (x_n)_{n \in \mathbb{N}}$ is a pointwise cancellation sequence.

REMARK. The proof of the implication (ii) \Rightarrow (iii) (Proposition 2 below) does not make use of the fact that x is generic relative to some measure. Therefore, (ii) is a sufficient condition for (iii) also without that assumption. We do not know whether, in this more general situation, it is also a necessary condition.

In the next section, we write down the full proof of Theorem 1 (since in [6] the theorem was merely formulated).

Based on a proof similar in its core to the proof of Theorem 1, we prove the following theorem reinforcing Theorem 0 when $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$.

THEOREM 2. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence satisfying $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$. The following conditions are equivalent:*

- (i) $\frac{1}{T} \sum_{n=1}^{T-1} x_n z^n \xrightarrow{T \rightarrow \infty} 0$ for all $z \in \mathbb{C}$ on the unit circle.
- (ii) For any increasing sequence of integers T_k there exists a subsequence T_{k_l} with the property that for any $\varepsilon > 0$ there exists $r' > 0$ such that for any $r'' > r'$ there exists $l' > 0$ for which

$$\left| \frac{\{0 < \tau \leq T_{k_l} : \exists r' \leq r \leq r'' \mid \frac{1}{T_{k_r}} \sum_{n=1}^{T_{k_r}} x_{n+\tau} \overline{x_n} \geq \varepsilon\}}{T_{k_l}} \right| < \varepsilon \quad \text{for all } l > l'.$$

- (iii) $\left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0$ for any complex-valued bounded stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$.
- (iv) $x = (x_n)_{n \in \mathbb{N}}$ is a mean cancellation sequence.

So Theorem 2 gives another characterization for a sequence to have an empty Fourier–Bohr spectrum, this time applying an expression linked to its auto-correlation. It is proved in Section 3. To make this paper self-contained we include in the proof also the part that has already appeared in [2].

Turning to another line of inquiry, it was proved in [2] that for all ergodic stationary processes $(Y_n(\omega))_{n \in \mathbb{Z}}$ of finite variance with Y_1 having a continuous spectrum relative to the unitary shift map, the event of being generic and also a pointwise cancellation sequence is of probability 1. A generic x represents a stationary process with such a continuous spectrum if and only if

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{\tau=0}^{T-1} \lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N x_{n+\tau} \overline{x_n} \right| = 0.$$

Section 4 is devoted to the possibility of a pointwise cancellation sequence that is generic relative to a stationary process for which the above mentioned spectrum is not continuous. A posteriori, after establishing Theorem 1, this is more interesting to contemplate. We discuss this issue and provide examples that constitute our present understanding of it.

2. The proofs of Theorems 1 and 2. Theorem 1 follows right away from Propositions 3 and 4: Proposition 3 is (ii) \Rightarrow (iii) and Proposition 4 is (i) \Rightarrow (ii), while (iii) \Rightarrow (i) is trivial as (i) is a special case of (iii) (by taking the stationary process with values which are the complex conjugates of the values of the stationary process defined by μ).

PROPOSITION 3. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence satisfying $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$. Assume that for any $\varepsilon > 0$ there exists $N' > 0$ such that for any $N'' > N'$ there exists $T' > 0$ for which*

$$(*) \quad \frac{|\{0 < \tau \leq T : \exists N' \leq N \leq N'' \mid \frac{1}{N} \sum_{n=1}^N x_{n+\tau} \overline{x_n} \geq \varepsilon\}|}{T} < \varepsilon \text{ for all } T > T'.$$

Then x is a pointwise cancellation sequence.

Proof. Assuming the theorem is false, there is a σ -invariant Borel probability measure on $\mathbb{C}^{\mathbb{N}}$ for which $\limsup_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N x_n y_n \right| > 0$ on a set of positive measure. By considering its ergodic decomposition, there exists an ergodic σ -invariant Borel probability measure ν on $\mathbb{C}^{\mathbb{N}}$, a positive ν -measure set $B' \subseteq \mathbb{C}^{\mathbb{N}}$ and some $a > 0$ such that $\limsup_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N x_n y_n \right| > a$ for all $y \in B'$. We may also assume that $\|\pi_1\|_2 = 1$ in $L^2_{\nu}(\mathbb{C}^{\mathbb{N}})$ (by performing dilation).

There is a positive ν -measure set $B \subseteq B'$ and a sequence of intervals (L_j, M_j) satisfying

$$0 < L_1 < M_1 < L_2 < M_2 < \dots$$

such that for every $y \in B$ and $j \in \mathbb{N}$ there exists $n_j(y) \in (L_j, M_j)$ for which

$$\left| \frac{1}{n_j(y)} \sum_{n=1}^{n_j(y)} x_n y_n \right| > a.$$

We now fix $\delta = 1/3$. There exists, by ergodicity, some $K > 0$ such that

$$\nu \left(\bigcup_{k=0}^{K-1} \sigma^{-k}(B) \right) > 1 - \delta/2.$$

We denote by G the set of all $y \in \mathbb{C}^{\mathbb{N}}$ satisfying

$$\frac{1}{N} \sum_{n=0}^{N-1} \mathbb{1}_{\bigcup_{k=0}^{K-1} \sigma^{-k}(B)}(\sigma^n(y)) \xrightarrow{N \rightarrow \infty} \nu \left(\bigcup_{k=0}^{K-1} \sigma^{-k}(B) \right)$$

and $\frac{1}{N} \sum_{n=1}^N |y_n|^2 \xrightarrow{N \rightarrow \infty} \|\pi_1\|_2^2 = 1$. We have $\nu(G) = 1$ so in particular G is non-empty and this is all we shall need.

Given any $y \in G$ and $J \in \mathbb{N}$, for a large enough D (depending on y and J) we aim to construct a sequence

$$c^{(1)}(y), \dots, c^{(J)}(y) \in \mathbb{C}^D$$

satisfying the following (when considering these vectors with the inner product of the normalized uniform measure on $\{1, \dots, D\}$):

- $\|c^{(j)}(y)\|_2^2 \leq \sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N}$ for every j .
- $\langle c^{(j_1)}(y), c^{(j_2)}(y) \rangle \leq \delta/J$ for all $j_1 \neq j_2$.
- $\langle c^{(j)}(y), \overline{(y_n)_{n=1}^D} \rangle > (1 - \delta)a$ for every j .

This implies that $c(y) = c^{(1)}(y) + \dots + c^{(J)}(y)$ satisfies on the one hand

$$\|c(y)\|_2^2 = \|c^{(1)}(y) + \dots + c^{(J)}(y)\|_2^2 \leq \sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} J + \delta(J - 1).$$

On the other hand, for D large enough,

$$\begin{aligned} (1 - \delta)aJ &< \langle c^{(1)}(y) + \dots + c^{(J)}(y), \overline{(y_n)_{n=1}^D} \rangle \leq \|c(y)\|_2 \sqrt{\frac{1}{D} \sum_{n=1}^D |y_n|^2} \\ &< (1 + \delta)\|\pi_1\|_2 \|c(y)\|_2 = (1 + \delta)\|c(y)\|_2. \end{aligned}$$

Hence

$$(1 - \delta)^2 a^2 J^2 < (1 + \delta)^2 \left(\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} J + \delta(J - 1) \right)$$

for all $J \in \mathbb{N}$, and this leads to the contradiction $a = 0$.

Constructing the sequence $c^{(1)}(y), \dots, c^{(J)}(y) \in \mathbb{C}^D$. Fixing some $y \in G$ and $J \in \mathbb{N}$, we construct these vectors for a large enough D . By deleting some of the intervals (L_j, M_j) we may assume without loss of generality that $\frac{K}{L_1} < \frac{\delta}{6J}$, $\frac{K+M_j}{L_{j+1}} < \frac{\delta}{6J}$, and also, placing $\varepsilon = \frac{\delta}{6J^2K}$ in (*), and using the notation of (*) the following requirements hold:

(Δ) $L_1 \geq N'$, and for $1 \leq j < J$, we have $N'' = M_j$ and $L_{j+1} \geq M_j + T'$ (N', N'' and T' are as in (*); notice that N'' and T' depend on j).

We define $c^{(j)}(y)$ recursively on j from J downwards to 1. As the base case, $c^{(J)}(y)$ is defined as follows: we denote by $l_1^{(J)}(y)$ the minimal $0 \leq l \leq D - M_J$ for which $\sigma^l(y) \in B$ and define

$$c_{l_1^{(J)}(y)}^{(J)}(y), \dots, c_{l_1^{(J)}(y) + n_J(\sigma^{l_1^{(J)}(y)}(y)) - 1}^{(J)}(y)$$

to be equal to

$$\omega_1^{(J)} x_1, \dots, \omega_1^{(J)} x_{n_J(\sigma^{l_1^{(J)}(y)}(y))}$$

for $\omega_1^{(J)}$ with $|\omega_1^{(J)}| = 1$ that has the property that

$$\frac{\omega_1^{(J)}}{n_J(\sigma^{l_1^{(J)}(y)}(y))} \sum_{n=1}^{n_J(\sigma^{l_1^{(J)}(y)}(y))} x_n \cdot (\sigma^{l_1^{(J)}(y)}(y))_n$$

is a positive number (greater than a). Then denote by $l_2^{(J)}(y)$ the minimal $l_1^{(J)}(y) + n_J(\sigma^{l_1^{(J)}(y)}(y)) \leq l \leq D - M_J$ for which $\sigma^l(y) \in B$, and define

$$c_{l_2^{(J)}(y)}^{(J)}(y), \dots, c_{l_2^{(J)}(y) + n_J(\sigma^{l_2^{(J)}(y)}(y)) - 1}^{(J)}(y)$$

to be equal to

$$\omega_2^{(J)} x_1, \dots, \omega_2^{(J)} x_{n_J(\sigma^{l_2^{(J)}(y)}(y))}$$

for $\omega_2^{(J)}$ with $|\omega_2^{(J)}| = 1$ such that

$$\frac{\omega_2^{(J)}}{n_J(\sigma^{l_2^{(J)}(y)}(y))} \sum_{n=1}^{n_J(\sigma^{l_2^{(J)}(y)}(y))} x_n \cdot (\sigma^{l_2^{(J)}(y)}(y))_n$$

is a positive number (greater than a). We continue until reaching step R_J for which after $l_{R_J}^{(J)}(y) + n_J(\sigma^{l_{R_J}^{(J)}(y)}(y)) - 1$ there is no such l . For all the rest of the indices of $c^{(J)}(y)$ we set the value to be zero.

For indices $1 \leq n \leq D$ that do not belong to the union

$$\bigcup_{i=1}^{R_J} \{l_i^{(J)}(y) - K + 1, \dots, l_i^{(J)}(y), \dots, l_i^{(J)}(y) + n_J(\sigma^{l_i^{(J)}(y)}(y)) - 1\},$$

necessarily $\sigma^n(y) \notin \bigcup_{k=0}^{K-1} \sigma^{-k}(B)$ or $n \in \{D - M_J - K + 1, \dots, D\}$. Thus, for a large D , there are more than $(1 - \frac{\delta}{2})D$ indices $1 \leq n \leq D$ that belong to that union, and thus there are more than

$$\begin{aligned} \left(1 - \frac{K}{L_1}\right) \left(1 - \frac{\delta}{2}\right) D &> \left(1 - \frac{K}{L_1} - \frac{\delta}{2}\right) D > \left(1 - \frac{\delta}{6J} - \frac{\delta}{2}\right) D \\ &> \left(1 - \frac{\delta}{2J} - \frac{\delta}{2}\right) D \end{aligned}$$

indices $1 \leq n \leq D$ that are in $\bigcup_{i=1}^{R_J} \{l_i^{(J)}(y), \dots, l_i^{(J)}(y) + n_J(\sigma^{l_i^{(J)}}(y)(y)) - 1\}$ and also satisfy $\sigma^n(y) \in \bigcup_{k=0}^{K-1} \sigma^{-k}(B)$. Likewise, in the step corresponding to each $1 \leq j_0 \leq J$ in the recursion we will prove inductively that there are more than $(1 - (J - j_0 + 1)\frac{\delta}{2J} - \frac{\delta}{2})D$ indices $1 \leq n \leq D$ that belong to $\bigcup_{i=1}^{R_{j_0}} \{l_i^{(j_0)}(y), \dots, l_i^{(j_0)}(y) + n_{j_0}(\sigma^{l_i^{(j_0)}}(y)(y)) - 1\}$ and also satisfy $\sigma^n(y) \in \bigcup_{k=0}^{K-1} \sigma^{-k}(B)$.

We turn to the step of recursion for $1 \leq j_0 < J$. The number $c^{(j_0)}(y)$ is defined with n_{j_0} as we did in the base case with $c^{(J)}(y)$ and n_J , with an additional requirement when determining $l_i^{(j_0)}(y)$ that we now describe. We determine $l_i^{(j_0)}(y)$ similarly to how we determined $l_i^{(J)}(y)$, only that we want $1 \leq l \leq D - M_{j_0}$ to satisfy not only $\sigma^l(y) \in B$, but also

$$\begin{aligned} \{l, \dots, l + n_{j_0}(\sigma^l(y)) - 1\} \\ \subseteq \{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0+1}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\} \end{aligned}$$

for some i , and moreover for all $j_0 < j \leq J$,

$$\frac{1}{n_{j_0}(\sigma^l(y))} |x_1 c_l^{(j)}(y) + \dots + x_{n_{j_0}(\sigma^l(y))} c_{l+n_{j_0}(\sigma^l(y))-1}^{(j)}(y)| < \frac{\delta}{2J}.$$

As before, the rest of the indices are set to zero.

If an index $1 \leq n \leq D$ does not belong to either

$$\bigcup_{i=1}^{R_{j_0}} \{l_i^{(j_0)}(y) - K + 1, \dots, l_i^{(j_0)}(y), \dots, l_i^{(j_0)}(y) + n_{j_0}(\sigma^{l_i^{(j_0)}}(y)(y)) - 1\}$$

or $\bigcup_{k=0}^{K-1} \sigma^{-k}(B)$ then it must belong to at least one of the following sets:

$$\begin{aligned} A_1 = \left\{ 1 \leq n \leq D : \sigma^n(y) \notin \bigcup_{k=0}^{K-1} \sigma^{-k}(B) \text{ or} \right. \\ \left. n \notin \bigcup_{i=1}^{R_{j_0+1}} \{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\} \right\}, \end{aligned}$$

$$A_2 = \left\{ 1 \leq n \leq D : n \notin A_1, \forall i \{n, \dots, n + K + M_{j_0} - 1\} \right. \\ \left. \not\subseteq \{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0+1}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\} \right\},$$

$$A_3 = \bigcup_{k=0}^{K-1} B_k$$

where

$$B_k = \left\{ 1 \leq n \leq D : \sigma^{n+k}(y) \in B, \right.$$

$$\exists i \{n + k, \dots, n + k + n_{j_0}(\sigma^{n+k}(y)) - 1\} \\ \subseteq \{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0+1}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\},$$

$$\exists j_0 < j \leq J$$

$$\left. \frac{1}{n_{j_0}(\sigma^l(y))} |x_1 c_{n+k}^{(j)}(y) + \dots + x_{n_{j_0}(\sigma^l(y))} c_{n+k+n_{j_0}(\sigma^{n+k}(y))-1}^{(j)}(y)| \geq \frac{\delta}{2J} \right\}.$$

For a large D , by the induction hypothesis, $|A_1| < ((J - j_0)\frac{\delta}{2J} + \frac{\delta}{2})D$.

As for $|A_2|$, since $n \notin A_1$, there exists some i for which

$$n \in \{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0+1}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\},$$

but

$$\{n, \dots, n + K + M_{j_0} - 1\} \\ \not\subseteq \{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0+1}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\}.$$

Thus n is one of the largest $K + M_{j_0} - 1$ numbers in

$$\{l_i^{(j_0+1)}(y), \dots, l_i^{(j_0+1)}(y) + n_{j_0+1}(\sigma^{l_i^{(j_0+1)}}(y)(y)) - 1\},$$

and $|A_2|$ is bounded from above by $\frac{K+M_{j_0}}{L_{j_0+1}}D$.

Regarding $|A_3|$, let us first define for every $j_0 < j \leq J$ the set $B_{k,j}$ as we defined B_k except without the expression “ $\exists j_0 < j \leq J$ ”. So $B_k = \bigcup_{j_0+1}^J B_{k,j}$. By (Δ) , from those n that satisfy the first two requirements in the definition of $B_{k,j}$ only a fraction of less than ε satisfies the third, thus certainly $|B_{k,j}| < \varepsilon D$. Hence $|B_k| < J\varepsilon D = J \cdot \frac{\delta}{6J^2K} \cdot D = \frac{\delta}{6JK}D$ and thus $|A_3| < \frac{\delta}{6J}D < \frac{\delta}{2J}D$.

Overall, the number of indices $1 \leq n \leq D$ that do not belong to either

$$\bigcup_{i=1}^{R_{j_0}} \{l_i^{(j_0)}(y) - K + 1, \dots, l_i^{(j_0)}(y), \dots, l_i^{(j_0)}(y) + n_{j_0}(\sigma^{l_i^{(j_0)}}(y)(y)) - 1\}$$

or $\bigcup_{k=0}^{K-1} \sigma^{-k}(B)$ is not greater than

$$\begin{aligned} |A_1| + |A_2| + |A_3| &< \left((J - j_0) \frac{\delta}{2J} + \frac{\delta}{2} \right) D + \frac{K + M_{j_0}}{L_{j_0+1}} D + \frac{\delta}{6J} D \\ &< \left((J - j_0) \frac{\delta}{2J} + \frac{\delta}{2} + \frac{\delta}{6J} + \frac{\delta}{6J} \right) D. \end{aligned}$$

And since $\frac{K}{L_{j_0}} < \frac{K}{L_1} < \frac{\delta}{6J}$, the number of indices $1 \leq n \leq D$ that do not belong to either

$$\bigcup_{i=1}^{R_{j_0}} \{l_i^{(j_0)}(y), \dots, l_i^{(j_0)}(y) + n_{j_0}(\sigma^{l_i^{(j_0)}}(y)(y)) - 1\}$$

or $\bigcup_{k=0}^{K-1} \sigma^{-k}(B)$ is smaller than

$$\frac{\delta}{6J} D + \left((J - j_0) \frac{\delta}{2J} + \frac{\delta}{2} + \frac{\delta}{6J} + \frac{\delta}{6J} \right) D = \left((J - j_0 + 1) \frac{\delta}{2J} + \frac{\delta}{2} \right) D.$$

Up to now, we have constructed $c^{(1)}(y), \dots, c^{(J)}(y)$. By the construction it is clear that $\|c^{(j)}(y)\|_2^2 \leq \sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N}$ for every j , and that $\langle c^{(j_1)}(y), c^{(j_2)}(y) \rangle \leq \delta/J$ for every $j_1 \neq j_2$. We have also proved that for every j there are more than $(1 - (J - j_0 + 1) \frac{\delta}{2J} - \frac{\delta}{2}) D$ indices $1 \leq n \leq D$ that belong to $\bigcup_{i=1}^{R_j} \{l_i^{(j)}(y), \dots, l_i^{(j)}(y) + n_j(\sigma^{l_i^{(j)}}(y)(y)) - 1\}$. With the aid of this fact we can infer that $\langle c^{(j)}(y), (y_n)_{n=1}^D \rangle > (1 - \delta)a$:

$$\begin{aligned} &\left| \frac{1}{D} \sum_{n=1}^D c_n^{(j)}(y) y_n \right| \\ &= \left| \frac{n_j(\sigma^{l_1^{(j)}}(y)(y))}{D} \cdot \frac{1}{n_j(\sigma^{l_1^{(j)}}(y)(y))} \sum_{n=1}^{n_j(\sigma^{l_1^{(j)}}(y)(y))} x_n \cdot (\sigma^{l_1^{(j)}}(y)(y))_n \right| + \dots \\ &\quad + \left| \frac{n_j(\sigma^{l_{R_j}^{(j)}}(y))}{D} \cdot \frac{1}{n_j(\sigma^{l_{R_j}^{(j)}}(y))} \sum_{n=1}^{n_j(\sigma^{l_{R_j}^{(j)}}(y))} x_n \cdot (\sigma^{l_{R_j}^{(j)}}(y)(y))_n \right| \\ &> \left(\frac{n_j(\sigma^{l_1^{(j)}}(y)(y))}{D} + \dots + \frac{n_j(\sigma^{l_{R_j}^{(j)}}(y)(y))}{D} \right) a \\ &> \left(1 - (J - j_0 + 1) \frac{\delta}{2J} - \frac{\delta}{2} \right) a \geq (1 - \delta)a. \quad \blacksquare \end{aligned}$$

PROPOSITION 4. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence generic relative to a σ -invariant Borel probability measure μ on $\mathbb{C}^{\mathbb{N}}$ that satisfies $\frac{1}{T} \sum_{n=1}^{T-1} x_n \overline{x_{n+\ell}}$ $\xrightarrow{T \rightarrow \infty} 0$ for μ -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$. Then for any $\varepsilon > 0$ there exists $N' > 0$*

such that for any $N'' > N'$ there exists $T' > 0$ for which

$$\frac{|\{0 < \tau \leq T : \exists N' \leq N \leq N'' \mid \frac{1}{N} \sum_{n=1}^N x_{n+\tau} \overline{x_n} \geq \varepsilon\}|}{T} < \varepsilon \quad \text{for all } T > T'.$$

Proof. Given $\varepsilon > 0$ and $\tilde{N} \in \mathbb{N}$, we denote by $A_\varepsilon^{\tilde{N}} \subseteq \mathbb{C}^{\mathbb{N}}$ the measurable set containing all points $\xi \in \mathbb{C}^{\mathbb{N}}$ for which $|\frac{1}{N} \sum_{n=1}^{N-1} x_n \overline{\xi_n}| < \varepsilon$ for all $N \geq \tilde{N}$. For a large enough \tilde{N} we have $\mu(A_\varepsilon^{\tilde{N}}) > 1 - \varepsilon$. We denote the minimal such \tilde{N} by N' . For every $N'' \geq N'$, the set $B_\varepsilon^{N', N''} = \{\xi \in \mathbb{C}^{\mathbb{N}} : \forall N' \leq N \leq N'' \mid \frac{1}{N} \sum_{n=1}^{N-1} x_n \overline{\xi_n} < \varepsilon\}$ is of μ -measure greater than $1 - \varepsilon$. Its complement is a closed set, and hence, since x is generic, there exists some $T' \in \mathbb{N}$ such that $|\{0 < \tau \leq T : \sigma^\tau(x) \notin B_\varepsilon^{N', N''}\}|/T < \varepsilon$ for every $T > T'$. But

$$\begin{aligned} & \{0 < \tau \leq T : \sigma^\tau(x) \notin B_\varepsilon^{N', N''}\} \\ &= \left\{ 0 < \tau \leq T : \exists N' \leq N \leq N'' \mid \left| \frac{1}{N} \sum_{n=1}^N x_{n+\tau} \overline{x_n} \right| \geq \varepsilon \right\} \end{aligned}$$

for every T , and hence we are done. ■

3. Mean cancellation sequences. By a similar proof to that of Theorem 1, the following proposition can be obtained.

PROPOSITION 5. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence, and let T_k be an increasing sequence of integers for which $\sup_{k \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_{T_k}|^2}{T_k} < \infty$ and $\frac{1}{T_k} \sum_{n=1}^{T_k} \delta_{\sigma^n(x)}$ converges weak* to some probability measure μ as $k \rightarrow \infty$. The following conditions are equivalent:*

- (i) $\frac{1}{T_k} \sum_{n=1}^{T_k-1} x_n \overline{x_n} \xrightarrow{k \rightarrow \infty} 0$ for μ -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$.
- (ii) For any $\varepsilon > 0$ there exists $s' > 0$ such that for any $s'' > s'$ there exists $k' > 0$ for which

$$\frac{|\{0 < \tau \leq T_k : \exists s' \leq s \leq s'' \mid \frac{1}{T_s} \sum_{n=1}^{T_s} x_{n+\tau} \overline{x_n} \geq \varepsilon\}|}{T_k} < \varepsilon \quad \text{for all } k > k'.$$

- (iii) For any complex-valued stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$ of finite variance,

$$\frac{1}{T_k} \sum_{n=1}^{T_k} x_n Y_n(\omega) \xrightarrow{k \rightarrow \infty} 0 \quad \text{almost surely.}$$

Proof. The proof is carried out similarly to that of Theorem 1. In the proof of Proposition 3 the index $n_j(y)$ should be equal to a T_k . ■

The condition $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$ implies $\sup_{N \in \mathbb{N}} \frac{|x_1| + \dots + |x_N|}{N} < \infty$, and the latter condition suffices to ensure that for every $\varepsilon > 0$ there exists $M > 0$ for which the $\frac{1}{T} \sum_{n=1}^T \delta_{\sigma^n(x)}$ -measure of the set $\{z \in \mathbb{C}^{\mathbb{N}} : |z_1| \leq M\}$

is greater than $1 - \varepsilon$ for all $T \in \mathbb{N}$. Thus, in the weak* topology, any partial limit of the sequence $\frac{1}{T} \sum_{n=1}^T \delta_{\sigma^n(x)}$ is a probability measure (there is no escape of mass).

Hence Corollary 6 below is an immediate conclusion from Proposition 5. It does not presuppose any genericity of x .

COROLLARY 6. *Let $x = (x_n)_{n \in \mathbb{N}}$ be a complex sequence satisfying $\sup_{N \in \mathbb{N}} \frac{|x_1|^2 + \dots + |x_N|^2}{N} < \infty$. The following conditions are equivalent:*

- (i) *For any increasing sequence of integers T_k there exists a subsequence T_{k_l} with the property that for any $\varepsilon > 0$ there exists $r' > 0$ such that for any $r'' > r'$ there exists $l' > 0$ for which*

$$\left| \frac{\left\{ 0 < \tau \leq T_{k_l} : \exists r' \leq r \leq r'' \mid \left| \frac{1}{T_{k_r}} \sum_{n=1}^{T_{k_r}} x_{n+\tau} \overline{x_n} \right| \geq \varepsilon \right\}}{T_{k_l}} \right| < \varepsilon \quad \text{for all } l > l'.$$

- (ii) *For any increasing sequence of integers T_k there exists a subsequence T_{k_l} such that for any complex-valued stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$ of finite variance*

$$\frac{1}{T_{k_l}} \sum_{n=1}^{T_{k_l}} x_n Y_n(\omega) \xrightarrow{l \rightarrow \infty} 0 \quad \text{almost surely.}$$

We are ready to prove Theorem 2.

Proof of Theorem 2. Given a stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$ of finite variance, applying the spectral theorem with the vector Y_1 and the unitary shift map, we find that $\left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0$ if and only if $\left\| \frac{1}{T} \sum_{n=1}^T x_n z^n \right\|_2 \xrightarrow{T \rightarrow \infty} 0$ in L^2 on the unit circle equipped with the spectral measure.

(i) \Rightarrow (iv): Assume $\frac{1}{T} \sum_{n=1}^T x_n z^n \xrightarrow{T \rightarrow \infty} 0$ for all $z \in \mathbb{C}$ on the unit circle.

Because

$$\left| \frac{1}{T} \sum_{n=1}^T x_n z^n \right| \leq \left| \frac{1}{T} \sum_{n=1}^T |x_n| \right|,$$

and $\sup_{T \in \mathbb{N}} \left| \frac{1}{T} \sum_{n=1}^T |x_n| \right| < \infty$, we can apply the dominated convergence theorem to deduce the desired result.

(iv) \Rightarrow (iii): A bounded random variable is of finite variance.

(iii) \Rightarrow (i): Assume the contrary, that there exists some z_0 on the unit circle satisfying $\limsup_{T \rightarrow \infty} \left| \frac{1}{T} \sum_{n=1}^T x_n z_0^n \right| > 0$. Considering $Y_n(\omega)$ to be ωz_0^n where the underlying probability space is the unit circle with the uniform measure, the procedure above of applying the spectral theorem yields

a spectral measure possessing an atom at z_0 , so

$$\left\| \frac{1}{T} \sum_{n=1}^T x_n z^n \right\|_2 \xrightarrow{T \rightarrow \infty} 0 \quad \text{and thus} \quad \left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0.$$

Up to now we have proved (i) \Leftrightarrow (iii) \Leftrightarrow (iv), and it is left to link these conditions to condition (ii) which is condition (i) of Corollary 6. We will link them to condition (ii) of Corollary 6 (which is equivalent to the latter by that corollary).

((ii) of Cor. 6) \Rightarrow (iii): Given a stationary process $(Y_n(\omega))_{n \in \mathbb{Z}}$ bounded by $M > 0$, we have

$$\left| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right| \leq M \left| \frac{1}{T} \sum_{n=1}^T |x_n| \right|,$$

and thus, by the dominated convergence theorem, for any increasing sequence of integers T_k there exists a subsequence T_{k_l} satisfying

$$\left\| \frac{1}{T_{k_l}} \sum_{n=1}^{T_{k_l}} x_n Y_n(\omega) \right\|_2 \xrightarrow{l \rightarrow \infty} 0.$$

This implies that $\left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0$.

(iv) \Rightarrow ((ii) of Cor. 6): $\left\| \frac{1}{T} \sum_{n=1}^T x_n Y_n(\omega) \right\|_2 \xrightarrow{T \rightarrow \infty} 0$ implies that $\left\| \frac{1}{T_k} \sum_{n=1}^{T_k} x_n Y_n(\omega) \right\|_2 \xrightarrow{k \rightarrow \infty} 0$ for any increasing sequence of integers T_k . In particular, there exists a subsequence T_{k_l} for which the convergence holds almost surely. ■

4. Generic cancellation sequences and the spectrum of π_1 relative to the Koopman operator. In [2], it is shown that for a σ -invariant Borel probability measure μ on $\mathbb{C}^{\mathbb{N}}$ with π_1 square integrable, the spectrum of π_1 relative to the Koopman operator of σ is continuous if and only if μ -almost every $x \in \mathbb{C}^{\mathbb{N}}$ satisfies $\frac{1}{T} \sum_{n=0}^{T-1} x_n \bar{\xi}_n \xrightarrow{T \rightarrow \infty} 0$ for μ -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$. If μ is ergodic then $x \in \mathbb{C}^{\mathbb{N}}$ is μ -almost surely generic relative to μ , and hence when also these two conditions hold then it is μ -almost surely a pointwise cancellation sequence (by Theorem 1'). However, we now give an example of a pointwise cancellation sequence $x \in \mathbb{C}^{\mathbb{N}}$ which is generic relative to a σ -invariant Borel probability measure μ but the two equivalent conditions above do not hold for that μ .

We make use of the following two standard propositions whose proofs are found in the appendix.

PROPOSITION 7. *For every irrational β and real $\alpha \neq 0$, the sequence $(e^{2\pi n \beta i}, e^{2\pi \sqrt{n} \alpha i})_{n=1}^{\infty}$ equidistributes on the 2-torus.*

PROPOSITION 8. *Let (Y, \mathcal{B}, T, ν) be probability measure preserving system and let α be an irrational number. If h is an integrable function on the circle (relative to the uniform measure) and $f \in L^2_\nu(y)$ is orthogonal in $L^2_\nu(y)$ to all eigenfunctions (relative to the Koopman operator) with eigenvalues of the form $e^{2\pi m\alpha i}$ for $0 \neq m \in \mathbb{Z}$, then, under every joining of (Y, \mathcal{B}, T, ν) and the system of the circle rotation by $e^{2\pi\alpha i}$, the integral of $f(y)h(z)$ equals the product of the integrals of each of these functions.*

EXAMPLE. Let α be an irrational number, and set $\omega = e^{2\pi\alpha i}$. The sequence $x = (e^{2\pi(n+\sqrt{n})\alpha i})_{n=1}^\infty$ equidistributes on the circle (by Proposition 7) and – because also $\sqrt{n+1} - \sqrt{n} \xrightarrow{n \rightarrow \infty} 0$ – considered as a point in $\mathbb{C}^\mathbb{N}$ it is generic relative to the measure which is the push-forward of the uniform measure of the circle through the map $z \mapsto (z\omega^n)_{n=1}^\infty$. The system of this measure together with the shift is measure-theoretically isomorphic to the system of the rotation of the circle by ω , and so it is not weak-mixing.

We now prove that x is a pointwise cancellation sequence. Given an ergodic σ -invariant Borel probability measure ν on $\mathbb{C}^\mathbb{N}$ for which π_1 is square integrable, we can write, in $L^2_\nu(\mathbb{C}^\mathbb{N})$, $\pi_1(y) = f(y) + g(y)$ where g belongs to the closure of the span of the eigenfunctions of eigenvalues ω^m for $0 \neq m \in \mathbb{Z}$ and f is orthogonal to this span. It is straightforward to prove using the Cauchy–Schwarz inequality that the functions $h \in L^2_\nu(\mathbb{C}^\mathbb{N})$ for which $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{n=1}^T x_n h(\sigma^n(y)) = 0$ ν -almost surely form a closed subspace (this is done in [7]), hence to show that $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{n=1}^T x_n g(\sigma^n(y)) = 0$ we may assume without loss of generality that $g = g_m$ where g_m is an eigenfunction of eigenvalue ω^m .

The equality $g_m(\sigma^n(y)) = \omega^{nm} g_m(y)$ holds for ν -almost every $y \in \mathbb{C}^\mathbb{N}$. So for ν -almost every $y \in \mathbb{C}^\mathbb{N}$,

$$\frac{1}{T} \sum_{n=1}^T x_n y_n = \frac{1}{T} \sum_{n=1}^T x_n f(\sigma^n(y)) + \frac{1}{T} \sum_{n=1}^T x_n g_m(\sigma^n(y)) \xrightarrow{T \rightarrow \infty} 0 + 0 = 0.$$

The first limit vanishes for ν -almost every $y \in \mathbb{C}^\mathbb{N}$ by Proposition 8 and the fact that the expectation of the stationary process defined by x is 0. Regarding the second limit, it is the limit in T of the expression

$$\begin{aligned} \frac{1}{T} \sum_{n=1}^T x_n g_m(\sigma^n(y)) &= \frac{1}{T} \sum_{n=1}^T e^{2\pi(n+\sqrt{n})\alpha i} \omega^{mn} g_m(y) \\ &= g_m(y) \frac{1}{T} \sum_{n=1}^T e^{2\pi(n(m+1)+\sqrt{n})\alpha i}, \end{aligned}$$

and $\frac{1}{T} \sum_{n=1}^T e^{2\pi(n(m+1)+\sqrt{n})\alpha i} \xrightarrow{T \rightarrow \infty} 0$ by Proposition 7 (the case of $m = 0$ is in fact the case $m = 0$ in the proof of Proposition 7).

For x of the above example, it is interesting to look into the behaviour of the expression

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N x_{n+\tau} \bar{x}_n &= \frac{1}{N} \sum_{n=1}^N e^{2\pi(n+\tau+\sqrt{n+\tau})\alpha i} e^{-2\pi(n+\sqrt{n})\alpha i} \\ &= \frac{1}{T} \sum_{n=1}^N e^{2\pi(\tau+\sqrt{n+\tau}-\sqrt{n})\alpha i}. \end{aligned}$$

By Theorem 1, for any $\varepsilon > 0$ there exists $N' > 0$ such that for any $N'' > N'$ there exists $T' > 0$ for which

$$\left| \left\{ 0 < \tau \leq T : \exists N' \leq N \leq N'' \left| \frac{1}{N} \sum_{n=1}^N e^{2\pi(\tau+\sqrt{n+\tau}-\sqrt{n})\alpha i} \right| \geq \varepsilon \right\} \right| < \varepsilon$$

for every $T > T'$. And yet,

$$\frac{1}{N} \sum_{n=1}^N e^{2\pi(\tau+\sqrt{n+\tau}-\sqrt{n})\alpha i} \xrightarrow{N \rightarrow \infty} e^{2\pi\tau\alpha i} \quad \text{for all } \tau.$$

An objection to this example can be made claiming that x is not in the support of the measure for which it is generic and maybe if this additional requirement would have been added then we could not find such an x . We will show this is not the case with a new example that makes use of the previous one.

EXAMPLE. Consider any weak-mixing σ -invariant Borel probability measure η on $\mathbb{C}^{\mathbb{N}}$ of full support with π_1 being square integrable (e.g. independent trials of any fully supported distribution on \mathbb{C} of finite variance). We take $u \in \mathbb{C}^{\mathbb{N}}$ to be any of its generic points that also satisfy $\frac{1}{T} \sum_{n=0}^{T-1} u_n \xi_n \xrightarrow{T \rightarrow \infty} 0$ for η -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$, and continue with x from the previous example. The stationary process defined by η is disjoint from the Kronecker stationary process for which x is generic, and hence (x, u) is generic relative to the product measure of the two processes. Thus $x + u$ is generic relative to the ergodic process which is the sum of the two processes taken independently, i.e. the convolution of the two measures (these are measures on the additive group $\mathbb{C}^{\mathbb{N}}$). This new measure is of full support and, according to it, π_1 is square integrable and does not have a continuous spectrum relative to the Koopman operator. We claim that for any complex-valued stationary process $(Y_n(\omega))_{n \in \mathbb{N}}$ of finite variance,

$$\frac{1}{N} \sum_{n=1}^N (x + u)_n Y_n(\omega) = \frac{1}{N} \sum_{n=1}^N x_n Y_n(\omega) + \frac{1}{N} \sum_{n=1}^N u_n Y_n(\omega) \xrightarrow{N \rightarrow \infty} 0 + 0$$

almost surely. The proof of the limit of the first summand was carried out in the previous example, and the second one follows from Theorem 1'.

This last example of $x + u$ fixed the problem of x not belonging to the support of the measure for which it is generic, but, unlike x , $x + u$ is not bounded. We do not have in hand such an example that is also bounded.

5. A sequence with mean but not pointwise cancellation. This section demonstrates the existence of weak-mixing bounded stationary processes of zero expectation that admit a generic point which is not a pointwise cancellation sequence. However, any generic point of a bounded weak-mixing stationary process of zero expectation satisfies condition (i) in Theorem 0 (by the disjointness of weak-mixing and Kronecker processes) and as such forms a mean cancellation sequence.

The proof relies on a yet unpublished example of M. Hochman appearing in the appendix, and in this section we assume the reader is familiar with it. Hochman shows that there exists a weak-mixing system with invariant measure μ which has a generic point that satisfies the following: the points in the fiber belonging to it in the product of the system with itself are μ -almost surely not generic relative to any $\sigma \times \sigma$ -invariant measure. In fact, he proves the existence of a family of such measures which he refers to as weak-mixing measures *admitting tall covers*.

We here just indicate how to modify what is done in Theorem 10 for our current needs.

We choose $\Lambda = \{a, b\}$ (a and b being some symbols) and take any weak-mixing σ -invariant Borel probability measure μ admitting tall covers. Instead of defining x as in Theorem 10, we define it in a simpler manner: $x_i = (a'_n)_i$ for $|a'_{n-1}| < i \leq |a'_n|$. Substitute a and b with any two real numbers $a' < 0 < b'$ for which the corresponding process μ' on $\{a', b'\}^{\mathbb{N}}$ is of zero expectation and denote by x' the point in $\{a', b'\}^{\mathbb{Z}}$ corresponding to x . Then in $(\mathbb{C}^{\mathbb{N}} \times \mathbb{C}^{\mathbb{N}}, \sigma \times \sigma)$ the point (x', ξ) has a sequence of orbital measures that converges to the measure $\Delta_{\mu'}$ (in the notation of the proof of Theorem 10) for μ' -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$. Thus for μ' -almost every $\xi \in \mathbb{C}^{\mathbb{N}}$ the sequence $\frac{1}{N} \sum_{i=1}^N x'_i \bar{\xi}_i$ has a subsequence that converges to $\mathbb{E}_{\mu'}(\pi_1^2) > 0$.

6. Appendix

6.1. Proofs of Propositions 7 and 8. We present here the proofs of Propositions 7 and 8 from the last section. These claims are standard, but are proved here in order to make this paper self-contained.

Proof of Proposition 7. Let β be an irrational number and $\alpha \neq 0$ be a real number. It suffices to prove that for every function $I : S^1 \rightarrow \{0, 1\}$ which is the indicator of an arc of the circle between angles $2\pi a$ and $2\pi b$ for

$0 \leq a < b \leq 1$ and $m \in \mathbb{Z}$, the sequence $(e^{2\pi n\beta i}, e^{2\pi\sqrt{n}\alpha i})_{n=1}^\infty$ satisfies

$$\frac{1}{N} \sum_{n=0}^N I(e^{2\pi\sqrt{n}\alpha i}) e^{2\pi mn\beta i} \xrightarrow{m \rightarrow \infty} \begin{cases} 0, & m \neq 0, \\ b - a, & m = 0, \end{cases}$$

(since step functions uniformly approximate continuous functions).

We first deal with the case $m = 0$. Given an $\varepsilon > 0$, we fix a positive integer d for which the distance between $d\alpha$ and the nearest integer is smaller than $\varepsilon/2$. It suffices to prove that

$$\left| \frac{1}{2Nd + d^2} \sum_{n=N^2}^{(N+d)^2-1} I(e^{2\pi\sqrt{n}\alpha i}) - (b - a) \right| < \varepsilon$$

for every N large enough. And this follows easily from the following considerations: $\sqrt{N^2 + k} = N + \frac{1}{2N}k + R(N, k)$, where $|R(N, k)| < \frac{1}{4N^3}$ for $k \geq 0$ (by Taylor's expansion for the real variable k). So when k ranges over the integers between 0 and $2Nd + d^2 - 1$, the sequence $\sqrt{N^2 + k} \cdot \alpha$ varies from $\sqrt{N^2} \cdot \alpha = N\alpha$ to $\sqrt{(N+d)^2} \cdot \alpha = (N+d)\alpha$ in nearly equal steps of $\frac{1}{2N}\alpha$ (for large N), and $(N+d)\alpha - N\alpha = d\alpha$ is an integer up to a difference of $\varepsilon/2$.

As for $m \neq 0$, suppose $\varepsilon > 0$ is given. It suffices to prove that for large enough N , if $e^{2\pi\sqrt{N}\alpha i}$ is inside the arc that I indicates and $e^{2\pi\sqrt{N-1}\alpha i}$ is not, then $\left| \frac{1}{r(N)} \sum_{n=N}^{N+r(N)-1} e^{2\pi mn\beta i} \right| < \varepsilon$ where $r(N) \in \mathbb{N}$ is the number for which $e^{2\pi\sqrt{N+k}\alpha i}$ stays inside the arc from $k = 0$ precisely up to $k = r(N) - 1$. Evaluating a geometric series

$$\left| \frac{1}{r(N)} \sum_{n=N}^{N+r-1} e^{2\pi mn\beta i} \right| < \frac{1}{r(N)} \cdot \frac{1}{|1 - e^{2\pi m\beta i}|}$$

and observing that $r(N) \rightarrow \infty$ as $N \rightarrow \infty$ completes the proof. ■

Proof of Proposition 8. Let λ be a joining of the two systems. We may assume without loss of generality that $h(z) = z^m$ for some $0 \neq m \in \mathbb{Z}$. We need to prove that $\int_{Y \times S^1} f(y) z^m d\lambda(y, z) = 0$. By the invariance of λ ,

$$\begin{aligned} \int_{Y \times S^1} f(y) z^m d\lambda(y, z) &= \int_{Y \times S^1} \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi mn\alpha i} z^m f(T^n(y)) d\lambda(y, z) \\ &= \int_{Y \times S^1} z^m \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi mn\alpha i} f(T^n(y)) d\lambda(y, z). \end{aligned}$$

On the other hand, $u(y) \mapsto e^{2\pi m\alpha i} f(T(y))$ is an isometry from $L^2_\nu(y)$ to itself, and thus by the mean ergodic theorem $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi mn\alpha i} f(T^n(y))$ exists in $L^2_\nu(y)$ and equals the projection of f on the eigenspace – relative to the Koopman operator of T – of the eigenvalue $e^{-2\pi m\alpha i}$. So by the as-

sumption on f this yields $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi mn\alpha i} f(T^n(y)) = 0$ in $L^2_\nu(y)$. Hence

$$\int_{Y \times S^1} f(y) z^m d\lambda(y, z) = \lim_{N \rightarrow \infty} \int_{Y \times S^1} z^m \frac{1}{N} \sum_{n=0}^{N-1} e^{2\pi mn\alpha i} f(T^n(y)) d\lambda(y, z) = 0. \blacksquare$$

6.2. Hochman's example. Here we describe M. Hochman's example of a weak-mixing system with invariant measure μ and its generic point whose fiber in the product of the system with itself has a peculiar property: the points on it are μ -almost surely not generic relative to any $\sigma \times \sigma$ -invariant measure (the reader interested in this topic can also look into [8, Section 3]). It is actually not a construction of a specific example, but rather proving the existence of a family of such invariant measures.

In the following, Λ is some finite alphabet. As usual, Λ^* will denote the collection of all finite length words in Λ , and if $\alpha \in \Lambda^*$ then $|\alpha|$ will denote its length. Also, $[\alpha]$ is the set of points $y \in \Lambda^{\mathbb{N}}$ satisfying $y_1 \dots y_{|\alpha|} = \alpha$.

DEFINITION. Let μ be a σ -invariant Borel probability measure on $\Lambda^{\mathbb{N}}$. For $\varepsilon > 0$ and $N \in \mathbb{N}$ we shall say that μ admits an (ε, N) -cover if there exist words $a_1, \dots, a_p \in \Lambda^*$ satisfying $|a_1| > N$, $|a_{i+1}| > 3|a_i|$ and $\mu(\bigcup_{i=1}^p [a_i]) > 1 - \varepsilon$. We shall say that μ admits tall covers if it admits an (ε, N) -cover for all (ε, N) .

Although measures admitting tall covers have zero entropy (this follows easily from the definition of entropy on the obvious generating partition), the family of those measures is a dense G_δ -set relative to the weak* topology on the Borel probability measures on $\Lambda^{\mathbb{N}}$. This is because, for every ε and N , the set of measures admitting (ε, N) -covers is open and contains all periodic measures.

The family of Borel probability measures that both admit tall covers and are weak-mixing contains a dense G_δ -set since also the set of all weak-mixing Borel probability measures on $\Lambda^{\mathbb{N}}$ contains a dense G_δ -set (cf. [4, the chapter "Category"]).

DEFINITION. Let μ be a σ -invariant ergodic Borel probability measure on $\Lambda^{\mathbb{N}}$, and let $\alpha \in \Lambda^*$. For $\varepsilon > 0$, we shall say that α is ε -generic if each word β of length not greater than $1/\varepsilon$ appears in it with a frequency closer than ε to its μ -probability, i.e.

$$\left| \mu(\beta) - \frac{1}{|\alpha|} |\{1 \leq i \leq |\alpha| - |\beta| + 1 : \alpha_i \dots \alpha_{i+|\beta|-1} = \beta\}| \right| < \varepsilon.$$

For $M \in \mathbb{N}$, we shall say that α is strongly (ε, M) -generic if

$$\frac{1}{|\alpha|} |\{1 \leq i < |\alpha| - M + 1 : \alpha_i \dots \alpha_{i+M-1} \text{ is } \varepsilon\text{-generic}\}| > 1 - \varepsilon.$$

PROPOSITION 9. *If μ is a σ -invariant ergodic Borel probability measure on $A^{\mathbb{Z}}$ that admits tall covers, then for every $\varepsilon > 0$ and every large enough M one can take for any $N \in \mathbb{N}$ an (ε, N) -cover consisting only of strongly (ε, M) -generic words.*

Proof. By the mean ergodic theorem, if M is large enough, there is a μ -probability greater than $1 - \varepsilon^2$ that a point $y \in A^{\mathbb{N}}$ has $y_1 \dots y_M$ ε -generic. Given an (ε, N) -cover \tilde{E} , we claim that if N is large enough (which we may suppose without loss of generality) then its elements that are not strongly (ε, M) -generic have μ -probability less than 3ε . The mean ergodic theorem implies we can find for every large enough L some $y \in A^{\mathbb{N}}$ for which both $y_1 \dots y_L$ is strongly (ε, M) -generic and each of the elements of \tilde{E} appears in it with a frequency closer than $\varepsilon^2/|\tilde{E}|$ to its μ -probability.

Denote by E the collection of elements of \tilde{E} that are not strongly (ε, M) -generic and assume without loss of generality that none of them forms an initial segment of another. Thus we can parse $y_1 \dots y_L$ to read elements of E in the following way: starting from y_1 , we progress to the right until we meet an element of E , then we keep progressing to the right from the first index after that element ends until we meet another element and so on. An index $1, \dots, L$ not covered by this parsing cannot be a beginning of an element of E in $y_1 \dots y_L$ and thus there are at most $(1 - \mu(E) + \varepsilon^2)L$ such indices ($\mu(E)$ denotes, by abuse of notation, the μ -measure of the set that the elements of E occupy together). Hence this parsing involves at least $(\mu(E) - \varepsilon^2)L$ of the indices $1, \dots, L$. For N much larger than M this means that at least $\frac{1}{2}(\mu(E) - \varepsilon^2)L$ of the M -blocks in $y_1 \dots y_L$ are not ε -generic. So if $\mu(E) \geq 3\varepsilon$ then $\frac{1}{2}(\mu(E) - \varepsilon^2) \geq \frac{3}{2}\varepsilon - \frac{1}{2}\varepsilon^2 > \varepsilon$, which contradicts the fact that $y_1 \dots y_L$ is strongly (ε, M) -generic.

Hence we can omit from the cover the elements of E and remain with a $(4\varepsilon, N)$ -cover consisting only of strongly (ε, M) -generic words. ■

We shall also need an elementary lemma.

LEMMA 10. *If α is strongly (ε, M) -generic then it is 2ε -generic.*

Proof. We parse α as follows. We find the minimal i for which $\alpha_i \dots \alpha_{i+M-1}$ is ε -generic, include this word in the parsing, then progress to the least index greater than $i + M - 1$ satisfying the same property, include that word and continue. All words of length not greater than $1/\varepsilon$ appear with frequency closer than ε to their μ -probability among the indices that are included in the parsing as initial indices. All the indices i that are not included in this parsing are in particular indices for which $\alpha_i \dots \alpha_{i+M-1}$ is not ε -generic, hence there are less than $\varepsilon|\alpha|$ such indices. So they cannot distort the frequency of those words by more than ε . ■

It remains to prove the following theorem.

THEOREM 11. *Let μ be a σ -invariant ergodic Borel probability measure on $\Lambda^{\mathbb{Z}}$ admitting tall covers. Then there exists a generic point $x \in \Lambda^{\mathbb{N}}$ relative to μ such that the points composing $\{x\} \times \Lambda^{\mathbb{N}}$ are μ -almost surely not generic relative to any $\sigma \times \sigma$ -invariant Borel probability measure μ .*

Proof. For $\varepsilon_k = \frac{1}{2^k}$ and a large enough M_k as in Proposition 9, we choose – inductively on $k \in \mathbb{N}$ – (N_k, ε_k) -covers $a_{k,1}, \dots, a_{k,p_k}$ with all elements strongly (ε_k, M_k) -generic and with N_k greater than both $k|a_{k-1,p_{k-1}}|$ (for $k > 1$) and $3 \cdot 2^k M_k$.

For notational convenience we denote the sequence made by concatenating all the elements of these covers together in order $a_{1,1}, \dots, a_{1,p_1}, a_{2,1}, \dots, a_{2,p_2}, a_{3,1}, \dots, a_{3,p_3}, \dots$ by $\{a'_n\}_{n=1}^{\infty}$. And we denote by $k(n)$ the k corresponding to each n .

We are ready to construct x :

$$x_i = \begin{cases} (a'_n)_i, & |a'_{n-1}| < i \leq |a'_{n-1}| + \lceil \frac{1}{3}|a'_n| \rceil, \\ (a'_n)_{i-1}, & |a'_{n-1}| + \lceil \frac{1}{3}|a'_n| \rceil < i \leq |a'_n|. \end{cases}$$

Since a'_n is strongly $(\frac{1}{2^{k(n)}}, M_{k(n)})$ -generic, every subblock of length at least $\frac{1}{3}|a'_n|$ must at least be strongly $(\frac{4}{2^{k(n)}}, M_{k(n)})$ -generic (the percentage of “good” M -blocks in the subblock need not be less than $3/2^{k(n)}$ due to “boundary effects”, but the requirement $M_k < \frac{1}{2^k} \frac{N_k}{3}$ ensures it is less than $\frac{4}{2^{k(n)}}$), so in particular $\frac{8}{2^{k(n)}}$ -generic by Lemma 10. This, together with the fact that $k(n) \xrightarrow{n \rightarrow \infty} \infty$, guarantees that x is generic.

By the Borel–Cantelli lemma ($\sum_{k=1}^{\infty} \frac{1}{2^k} = 1 < \infty$), $y \in \Lambda^{\mathbb{N}}$ μ -almost surely belongs – for k large enough – to $[a_{k,r(y,k)}]$ for some $1 \leq r(y,k) \leq p_k$. Denoting by $n(k,r)$ the n for which $a'_n = a_{k,r}$, the fact that $a_{k,r(y,k)} = a'_{n(k,r(y,k))}$ implies that y agrees with x on the indices

$$|a'_{n(k,r(y,k))-1}| < i \leq |a'_{n(k,r(y,k))-1}| + \lceil \frac{1}{3}|a'_{n(k,r(y,k))}| \rceil,$$

while on

$$|a'_{n(k,r(y,k))-1}| + \lceil \frac{1}{3}|a'_{n(k,r(y,k))}| \rceil < i \leq |a'_{n(k,r(y,k))}|$$

they differ by a shift of one symbol. Thus, recalling that $N_k > k|a_{k-1,p_{k-1}}|$, the sequence in k of $\sigma \times \sigma$ -orbital measures of (x, y) , taken for each k until time $\{|a'_{n(k,r(y,k))-1}| + \lceil \frac{1}{3}|a'_{n(k,r(y,k))}| \rceil\}_{k \in \mathbb{N}}$ converges to the push-forward of μ through the map $z \mapsto (z, z)$, which we denote by Δ_μ , while the similar sequence taken for each k until time $|a'_{n(k,r(y,k))}|$ converges to the sum of $\frac{1}{3}\Delta_\mu$ and $\frac{2}{3}$ times the push-forward of μ through $z \mapsto (z, \sigma(z))$. So the point (x, y) is not generic relative to any $\sigma \times \sigma$ -invariant measure. ■

Acknowledgements. We thank Asaf Katz and Yuval Peres for sketching to the author the proof of Proposition 7, and Hillel Furstenberg for

showing him the proof of Proposition 8 as well as providing other valuable remarks. In addition, we want to thank Mike Hochman for generously sharing with the author his example appearing in the second subsection of the appendix; Section 5 relies on it.

References

- [1] I. Assani and K. Presser, *A survey of the return times theorem*, in: Ergodic Theory and Dynamical Systems, de Gruyter, 2014, 19–58.
- [2] J. Bourgain, H. Furstenberg, Y. Katznelson and D. S. Ornstein, *Return times of dynamical systems*, Publ. Math. Inst. Hautes Études Sci. 69 (1989), 42–45.
- [3] G. Cohen and V. Losert, *On Hartman almost periodic functions*, Studia Math. 173 (2006), 81–101.
- [4] P. R. Halmos, *Lectures on Ergodic Theory*, Dover Publ., 2017.
- [5] J.-P. Kahane, *Sur les coefficients de Fourier–Bohr*, Studia Math. 21 (1961), 103–106.
- [6] E. Lesigne, C. Mauduit et B. Mosse, *Le théorème ergodique le long d’une suite q -multiplicative*, Compos. Math. 93 (1994), 49–79.
- [7] D. J. Rudolph, *A joinings proof of Bourgain’s return time theorem*, Ergodic Theory Dynam. Systems 14 (1994), 197–203.
- [8] D. J. Rudolph, *Fully generic sequences and a multiple-term return-times theorem*, Invent. Math. 131 (1997), 199–228.

Matan Tal
Einstein Institute of Mathematics
Edmond J. Safra Campus
The Hebrew University of Jerusalem
Jerusalem, Israel
E-mail: matan.tal@mail.huji.ac.il