Ergod. Th. & Dynam. Sys., (2024), **44**, 2581–2598
doi:10.1017/etds.2023.75

Some measure rigidity and equidistribution results for β-maps

NEVO FISHBEIN

Independent Scholar, Israel
(e-mail: nevofi@gmail.com)

(*Received* 27 *April* 2023 *and accepted in revised form* 6 *September* 2023)

Abstract. We prove $\times a \times b$ measure rigidity for multiplicatively independent pairs when $a \in \mathbb{N}$ and $b > 1$ is a 'specified' real number (the b-expansion of 1 has a tail or bounded runs of 0s) under a positive entropy condition. This is done by proving a mean decay of the Fourier series of the point masses average along $\times b$ orbits. We also prove a quantitative version of this decay under stronger conditions on the $\times a$ invariant measure. The quantitative version together with the $\times b$ invariance of the limit measure is a step toward a general Host-type pointwise equidistribution theorem in which the equidistribution is for Parry measure instead of Lebesgue. We show that finite memory length measures on the a-shift meet the mentioned conditions for mean convergence. Our main proof relies on techniques of Hochman.

Key words: β-maps, Parry measure, measure rigidity
2020 Mathematics Subject Classification: 28D05 (Primary); 11A63 (Secondary)

1. *Introduction*

1.1. *Definitions and notation.* First, we introduce a couple of frequently used definitions. Given a positive real number b, we define the b-fold map $T_b : [0, 1) \to [0, 1)$ by $T_b(x) = b \cdot x \bmod 1$. We identify $[0, 1)$ with $\mathbb{R}/\mathbb{Z}$ so that T_b is referred interchangeably as a toral map that has at most one discontinuity at 0. For a real pair $(s, t) \in (1, \infty) \times (1, \infty)$, we say that they are *multiplicatively independent* and write $s \nsim t$ if $(\log s/\log t) \notin \mathbb{Q}$. As is customary, we write $[\cdot]$, $\lceil\cdot\rceil$ for the floor and the ceiling functions, respectively, and $\{x\} = x - [x]$ for the fractional part of a non-negative real number x.

For a Polish space X, denote by (X, Σ) the Borel space that is associated to it. Let μ be some probability measure on (X, Σ). A sequence of points $\{x_n\}_{n=1}^{\infty}$ in (X, E) is said to be *equidistribute for* μ if the mean of their point masses weakly-* converges to μ: $(1/N) \sum_{n=1}^{N} \delta_{x_n} \xrightarrow{w-*} \mu$. Let $T : X \to X$ be a function. If for some $x \in X$ we have $(1/N) \sum_{n=1}^{N} \delta_{T^n x} \xrightarrow{w-*} \mu$, then we say that x *equidistributes for* μ *under* T.

For every $m \in \mathbb{Z}$ and $x \in \mathbb{R}/\mathbb{Z}$, we write $e_m(x) = e(mx) = \exp(2\pi imx)$. Let μ be a finite Borel measure on $\mathbb{R}/\mathbb{Z}$ and $M(\mathbb{R}/\mathbb{Z})$ be the set of all such measures. The *mth Fourier coefficient of* μ is defined to be $\hat{\mu}(m) = \int e_m(x)\, d\mu(x)$, the *Fourier transform of* μ is the sequence $\hat{\mu} = (\hat{\mu}(n))_{n\in\mathbb{Z}}$ and the map $\mathcal{F} : M(\mathbb{R}/\mathbb{Z}) \to \ell^\infty(\mathbb{Z})$ defined by $\mathcal{F}(\mu) = \hat{\mu}$ is the *Fourier transform*.

We denote the Lebesgue measure on $\mathbb{R}/\mathbb{Z}$ by m and the integration with respect to Lebesgue measure by dz. In our context, all the absolutely continuous measures are with respect to Lebesgue.

1.2. *Background.* Furstenberg's Diophantine theorems [3] formed the background to Furstenberg's pioneering $\times 2 \times 3$ conjecture about the measure rigidity of T_2, T_3. This conjecture suggests that the only non-atomic and ergodic Borel probability measure on the circle which is also T_2 and T_3 invariant is Lebesgue measure.

The best result so far on this problem was proved by Rudolph and later strengthened further by Johnson, now known as Rudolph–Johnson theorem [7, 10]. It establishes the conjecture for every multiplicatively independent pair of integers $m, n \geq 2$ under the additional assumption of positive entropy.

Later, the following pointwise 'equidistributional' version was proved by Host [6] when $\gcd(a, b) = 1$ and improved to the case $a \not\sim b$ by Hochman and Shmerkin [5]. Let μ be a probability measure on $\mathbb{R}/\mathbb{Z}$ which is invariant, ergodic, and has positive entropy with respect to an endomorphism T_a. Then μ-almost every (a.e.) x equidistributes for Lebesgue measure under every endomorphism T_b with $a \not\sim b$. We write HET (Host's equidistribution theorem) for short when referring to this theorem.

In another direction, the next result is due to Parry [9]. Given a real $b > 1$, there exists a unique T_b-invariant Borel probability measure that is equivalent to Lebesgue measure. Its Radon–Nikodym derivative can be written explicitly as

$$f(x) = \sum_{x<T_b^n(1)} \frac{1}{b^n} \quad \text{and} \quad 1 - \frac{1}{b} \leq f \leq \frac{1}{1-1/b}.$$

Therefore, for an integer $a \geq 2$ and a non-integer $b > 1$, there does not exist a joint T_a and T_b invariant and absolutely continuous probability measure.

This suggests the following problem.

Problem 1.1. Let $a \geq 2$ be an integer and let $b > 1$ be a non-integer such that $a \not\sim b$. Is it true that there are no non-atomic and ergodic Borel probability measures on the circle which are both T_a and T_b invariant?

The HET and the Parry measure may also be related to each other via the next possible generalization of HET.

Problem 1.2. (Generalized HET) Let μ be a probability measure on $\mathbb{R}/\mathbb{Z}$ which is invariant, ergodic, and has positive entropy with respect to an endomorphism T_a. Let $1 < b \in \mathbb{R}$ with $a \not\sim b$. Is it true that μ-a.e. x equidistributes for Parry measure under T_b?

These two problems are the main concern of this paper.

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

1.3. *Results.* We begin by outlining our strategy for tackling Problems 1.1 and 1.2 from a harmonic analysis perspective. We denote the space of bi-infinite sequences both of whose limits are zero by $c_0(\mathbb{Z})$.

In the case of Problem 1.1, a possible strategy is the following.

(1) To show that for every probability measure μ on $\mathbb{R}/\mathbb{Z}$ which is invariant, ergodic, and has positive entropy with respect to an endomorphism T_a, it holds that μ-a.e. x,

$$\left\{\int \limsup_{N\to\infty}\left|\frac{1}{N}\sum_{i=0}^{N-1} e_m(T_b^i(x))\right| d\mu\right\}_{m=-\infty}^{\infty} \in c_0(\mathbb{Z}).$$

(2) Assume by contradiction that μ is a non-atomic and ergodic Borel probability measure on the circle which is T_a- and T_b-invariant as in Problem 1.1. Notice that

$$|\hat{\mu}(m)| \leq \int \limsup_{N\to\infty}\left|\frac{1}{N}\sum_{i=0}^{N-1} e_m(T_b^i(x))\right| d\mu$$

and therefore by the first step,

$$\{|\hat{\mu}(m)|\}_{m=-\infty}^{\infty} \in c_0(\mathbb{Z}).$$

However, the T_a invariance implies that for every $m \in \mathbb{Z}$, we have $\hat{\mu}(m) = \hat{\mu}(a\cdot m)$ and the limit above is possible only when

$$\hat{\mu}(m) = \begin{cases} 1 & \text{for } m = 0, \\ 0 & \text{for } m \neq 0, \end{cases}$$

that is, when μ is Lebesgue measure. This contradicts the assumption that μ is T_b-invariant because Lebesgue measure is not T_b-invariant.

For the case of Problem 1.2, we will need another definition. Let a, b and μ be as in Problem 1.2. Given some $x \in [0, 1)$ and some subsequence $\{N_k\}$ of $\mathbb{N}$, we denote $\lambda_{x,\{N_k\}} = \lim_{k\to\infty}(1/N_k)\sum_{i=0}^{N_k-1}\delta_{T_b^i(x)}$, if it exists. It is not hard to show that the only T_b-invariant measure with $\ell^2(\mathbb{Z})$ Fourier transform is the Parry measure. Therefore, a possible strategy to solve Problem 1.2 is the following.

(1) To show that μ-a.e. x,

$$\left\{\limsup_{N\to\infty}\left|\frac{1}{N}\sum_{i=0}^{N-1} e_m(T_b^i(x))\right|\right\}_{m=-\infty}^{\infty} \in \ell^2(\mathbb{Z}).$$

(2) To show that for every μ-typical x, $\lambda_{x,\{N_k\}}$ is necessarily T_b-invariant (for every $\{N_k\}$ such that the limit exists).

(3) By the uniqueness of Parry measure, we conclude that it is the limit of every proper convergent subsequence of the sequence

$$\left\{\frac{1}{N}\sum_{i=0}^{N-1}\delta_{T_b^i(x)}\right\}_{N=1}^{\infty}$$

and we know that such a convergent subsequence does exist. Thus, the whole sequence also converges to Parry measure.

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

Before stating our first result, we give another definition. Given $1 < b \in \mathbb{R}$, there is a non-surjective measurable embedding of $[0, 1)$ in $\Lambda_b^{\mathbb{N}} = \{0, 1, \ldots, [b] - 1\}^{\mathbb{N}}$ which is given by the *b-expansion*,

$$r \mapsto ([b \cdot r], [b \cdot T_b r], [b \cdot T_b^2 r] \ldots).$$

That is, a subset of $\Lambda_b^{\mathbb{N}}$ can represent $[0, 1)$. We say that a real positive b is a *specified number* or with the *specification property* if the b-expansion of 1 has bounded runs of 0s or has a tail of 0s. The set of specified numbers is uncountable and dense in $(1, \infty)$ (see §2.1). In §3, we prove the following theorem.

THEOREM 1.3. *Let μ be a probability measure on $\mathbb{R}/\mathbb{Z}$ which is invariant, ergodic, and has positive entropy with respect to an endomorphism T_a for some positive integer $a \geq 2$ and let $b > 1$ be a real specified number such that $a \not\sim b$. Then,*

$$\left\{ \int \limsup_{N \to \infty} \left| \frac{1}{N} \sum_{i=0}^{N-1} e_m(T_b^i(x)) \right| d\mu \right\}_{m=-\infty}^{\infty} \in c_0(\mathbb{Z}).$$

According to our strategy, Theorem 1.3 yields a Furstenberg-type measure rigidity for the class of specified b values.

COROLLARY 1.4. *Let $a \geq 2$ be an integer and $b > 1$ be a specified non integer such that $a \not\sim b$. Then there are no jointly T_a and T_b invariant non-atomic and ergodic Borel probability measures with positive entropy under T_a.*

This is an answer to Problem 1.1 for the class of specified b values under an entropy assumption (the same assumption currently needed when $b \in \mathbb{N}$).

To explain our next result, we need a few more definitions. For convenience, we denote $\Lambda_a^{-\mathbb{N}} = \Omega^-$. For an integer $a \geq 2$, let $\mathcal{A}$ denote the a-adic partition of $[0, 1)$: $\{[k/a, (k+1)/a)\}_{k=0}^{a-1}$ and $\mathcal{A}(x) \in \mathcal{A}$ denote the element which contains $x \in X$. Let $\widetilde{\Omega} = \Omega^- \times [0, 1)$ be the natural extension of $(\mathbb{R}/\mathbb{Z}, \mu, T_a)$ together with the map $\widetilde{T}_a(\omega, x) = (\omega\mathcal{A}(x), T_a x)$ and write $\widetilde{\mu}$ for the unique extension of μ to a $\widetilde{T}_a$-invariant measure on $\widetilde{\Omega}$. Let $\mathcal{C} = \bigvee_{i=-\infty}^{0} \widetilde{T}_a^i \mathcal{A}$ denote the σ-algebra in $\widetilde{\Omega}$ generated by projection to the past: $\widetilde{\omega} = (\omega, x) \mapsto \omega$ and let $\{\widetilde{\mu}_\omega^{\mathcal{C}}\}_\omega$ be the corresponding disintegration. Notice that the members of $\{\widetilde{\mu}_\omega^{\mathcal{C}}\}_\omega$ depend only on the Ω^--component and that we can identify them as measures on $[0, 1)$ such that $\mu = \int \mu_\omega \, d\widetilde{\mu}(\omega)$.

In §4, under stronger assumptions, we prove a quantitative mean version of Theorem 1.3.

THEOREM 1.5. *Let a, b, μ be as in Theorem 1.3 and let $\{I_j^n\}_{j=1}^n$ be the uniform partition of the unit interval into n sub-intervals. Assume that there exists some positive α, such that for every n, the following holds:*

$$\operatorname*{ess\,sup}_{\eta, j} \mu_\eta(I_j) \leq O(n^{-\alpha})$$

and such that for some positive β, for every positive integer n, we have

$$\operatorname*{ess\,sup}_{\eta} \iint \chi_{\{x : |x-y| < n^{-1}\}} \, d\mu_\eta(x) \, d\mu_\eta(y) \leq O(n^{-\beta}).$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

Then for every integer m,

$$\boldsymbol{E}_\mu\bigg(\limsup_N \bigg|\frac{1}{N}\sum_{i=0}^{N-1} e_m(T_b^i x)\bigg|\bigg) < O(|m|^{\min_{\gamma,\delta>0}\max\{-\delta\alpha,(-1-\delta(\alpha-1)+\gamma\delta)/2,\delta(1-\gamma\beta)/2\}}).$$

We remark that one can always assume $\alpha \le \beta$ (see §4).

In §6.4, we show that the conditions on μ in Theorem 1.5 are flexible enough to work for every process with memory of finite length such as variable length mixing Markov chains. Nonetheless, it certainly does not always work as we show in §6.3.

Our secondary result from §5 provides the following theorem.

THEOREM 1.6. *For every real $b > 1$ and a μ-typical x, if $\{0\}$ is not an atom of a partial limit $\lambda_{x,\{N_k\}} = \lim_{k\to\infty}(1/N_k)\sum_{i=0}^{N_k-1}\delta_{T_b^i}(x)$, then $\lambda_{x,\{N_k\}}$ must be T_b-invariant.*

In particular, Theorem 1.3 together with Wiener's lemma [2] imply that $\{0\}$ is not an atom of $\lambda_{x,\{N_k\}}$, so Theorem 1.6 proves Step 2 in the second strategy under the assumptions of Theorem 1.3.

Notice that since

$$-0.5 < \min_{\gamma,\delta>0}\max\bigg\{-\delta\alpha, \frac{-1-\delta(\alpha-1)+\gamma\delta}{2}, \frac{\delta(1-\gamma\beta)}{2}\bigg\} < 0,$$

our strategy fails to solve Problem 1.2. Let us briefly examine some other alternatives. We focused on trying to relax the $\ell^2(\mathbb{Z})$ condition as follows.

Let $L^1(m)$ be the set of absolutely continuous measures and let $M_c(\mathbb{R}/\mathbb{Z})$ be the set of continuous measures. We say that a measure μ on $\mathbb{R}/\mathbb{Z}$ is *Rajchman* if $\lim_{|n|\to\infty}\hat{\mu}(n) = 0$ and denote the set of Rajchman measures by $\mathcal{R}$. The Riemman–Lebesgue lemma implies that every absolutely continuous measure is Rajchman. Works by several mathematicians at the beginning of the 20th century (see [8]) established the more extensive result that $L^1(m) \subsetneq \mathcal{R} \subsetneq M_c(\mathbb{R}/\mathbb{Z})$. We can naturally ask the following questions. Is it true that Rajchman T_b-invariant measure must be absolutely continuous? Equivalently, is it true that $L^1(m) \cap \{T_b\text{-invariant}\} = \mathcal{R} \cap \{T_b\text{-invariant}\}$? Some evidence for a positive answer comes from the special case of an integer b. In this case, the Parry measure is just the Lebesgue measure and if μ is T_b-invariant and Rajchman, then, as we already observed for every $n \in \mathbb{Z}$, $\hat{\mu}(T_b n) = \hat{\mu}(b\cdot n)$. Thus, in addition to $\hat{\mu}(0)$, all the other Fourier coefficients must vanish and μ is indeed Lebesgue. However, in §7, we show the following proposition.

PROPOSITION 1.7. *There exists a T_b-invariant Rajchman measure that is not absolutely continuous.*

2. *Preliminaries*

2.1. *β-shifts.* Recall our notation $\Lambda_b = \{0, 1, \ldots, [b]-1\}$. With respect to the product σ-algebra on $\Lambda_b^{\mathbb{N}}$, the *shift transformation* $\sigma : \Lambda_b^{\mathbb{N}} \to \Lambda_b^{\mathbb{N}}$ which is defined by $\sigma((\lambda_i)) = (\lambda_{i+1})$ turns $\Lambda_b^{\mathbb{N}}$ into a dynamical system that we call the *full* $[b]$ *shift*. The restriction of the full shift to the closure of the subset of sequences which encodes b-expansions is a subshift that we call the b-shift and denote by $X_b \subset \Lambda_b^{\mathbb{N}}$.

One says that a real positive b is the following.

- A *simple number* if the b-expansion of 1 has a 0s tail.
- A *simple Parry number* if the b-expansion of 1 is a periodic sequence. In some sources, it is also called a (purely) *periodic number*.
- A *Parry number* if the b-expansion of 1 has a periodic tail. In some sources, it is also called an *eventually periodic number*.

It is immediate to conclude that

$$\{\text{simple \#s}\} \subset \{\text{simple Parry \#s}\} \subset \{\text{Parry \#s}\} \subset \{\text{specified \#s}\}.$$

Parry showed [9] that the simple numbers are everywhere dense in $(1, \infty)$, and so is the set of specified numbers. Schmeling [11] showed that the set of specified numbers also has Hausdorff dimension 1, but it is meager and has Lebesgue measure 0. In particular, it has the cardinality of the continuum.

An important property of specified numbers is the following proposition.

PROPOSITION 2.1. *When b has the specification property, the orbit of* 1 *under T_b (in* $[0, 1)$*) remains bounded away from* 0 *unless it hits it.*

Proof. Let $1 < b \in \mathbb{R}$ be a specified number. We write $b_0 = [b]$, $b_1 = [b\{b\}], \ldots$ and similarly we write $r_0 = T_b^1(1) = \{b\}$, $r_1 = T_b^2(1) = \{b\{b\}\}, \ldots$. We need to prove that $0 < \inf_n\{r_n : r_n > 0\}$. The special case of simple b is trivial. If we assume by contradiction that $0 = \inf_n\{r_n : r_n > 0\}$, then b cannot be simple and there is an upper bound $k \in \mathbb{N}$ on the length of runs of 0s. However, for every k, there exists an $n_0 \in \mathbb{N}$ with $r_{n_0} < b^{-(k+1)}$ and therefore $b_{n_0+i} = 0$ for $1 \leq i \leq k+1$ in contradiction. □

Finally, we present a result by Parry [9] that gives a criterion to determine whether a given sequence $(b_n) \in \{0, \ldots [b]\}^{\mathbb{N}}$ is a b-expansion of some $x \in [0, 1)$, $x = b_0 + b_1/\beta + \cdots$. We emphasize that there might be many representations of x in this form but only one of them corresponds to the b-expansion that we described earlier. This will be useful for constructing the counterexample in §7.

If $(a_0, a_1, \ldots)$, $(b_0, b_1, \ldots)$ are sequences of the same length (finite or infinite) of non-negative integers less than b, we write $(a_0, a_1, \ldots) < (b_0, b_1, \ldots)$ when $a_n < b_n$ for the first $a_n \neq b_n$.

THEOREM 2.2. (Parry's criterion) *Let $b > 1$ be a non-simple number. If the b-expansion of b is $b = a_0 + a_1/b + \cdots$ and $(b_0, b_1, \ldots)$ is a sequence of non-negative integers, a necessary and sufficient condition for the existence of x with b-expansion, $x = b_0 + b_1/b + \cdots$, is that $(b_n, b_{n+1}, \ldots) < (a_0, a_1, \ldots)$ for all $n \geq 1$. In particular, $(a_n, a_{n+1}, \ldots) < (a_0, a_1, \ldots)$ for all $n \geq 1$.*

2.2. *Entropy theory.* Let $(X, \mathcal{B})$ be a standard Borel space and $\mathcal{D} \subset \mathcal{B}$ be a measurable partition. We write $\mathcal{D}(x) \in \mathcal{D}$ for the element which contains $x \in X$. This is also well defined when $\mathcal{D}$ is countably generated σ-algebra. In addition, we denote the joining of two finite partitions $\mathcal{A}, \mathcal{B}$ by $\mathcal{A} \vee \mathcal{B} = \{A \cap B : A \in \mathcal{A}, B \in \mathcal{B}\}$. Let $(X, \mathcal{B}, \mu)$ be a probability space then the *Shannon entropy* of μ with respect to a partition $\mathcal{A}$ of X is the non-negative

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

number $H_\mu(\mathcal{A}) = -\sum_{A\in\mathcal{A}} \mu(A)\log\mu(A)$. The *entropy* $h_\mu(T,\mathcal{A})$ *of a partition* $\mathcal{A}$ of a measure-preserving system $(X,\mathcal{F},\mu,T)$ is the limit $h_\mu(T,\mathcal{A}) = \lim_{n\to\infty}(1/n)H_\mu(\mathcal{A}_n)$, where $\mathcal{A}_n = \bigvee_{i=0}^{n-1} T^{-i}\mathcal{A}$. The *Kolmogorov–Sinai entropy* (or just the *entropy*) of the m.p.t. $(X,\mathcal{F},\mu,T)$ is $h_\mu(T) = \sup_{\mathcal{A}} h_\mu(T,\mathcal{A})$ where the supremum is taken over all the finite partitions $\mathcal{A}$. Equality is achieved if $\mathcal{A}$ is a generating partition, $\mathcal{F} = \bigvee_{n=1}^{\infty}\mathcal{A}_n \mod \mu$.

A landmark result in ergodic theory is the Shannon–McMillan–Breiman theorem [14]. Let $(X,\mathcal{F},\mu,T)$ be an ergodic measure-preserving system and $\mathcal{A}$ a finite partition. Then μ-a.e. x $\lim_{n\to\infty}(1/n)\log\mu(\mathcal{A}_n(x)) = h_\mu(T,\mathcal{A})$. It is not hard to deduce from it that an ergodic and T_a-invariant Borel probability measure μ with positive entropy is non-atomic (this also can be proved directly). We will use this corollary in occasional places.

2.3. *General results on equidistribution (due to Hochman).* This subsection covers three results that we adopt from Hochman [4]. Two of them are presented with a very superficial description of their proofs to help the reader gain some intuition. A more thorough treatment can be found in the original paper.

We denote the real line translation and scaling maps by

$$R_\theta x = x + \theta,$$
$$S_t x = t \cdot x,$$

respectively. Here R_θ is taken mod 1 when acting on $[0, 1) \cong \mathbb{R}/\mathbb{Z}$

Let μ be a probability measure on $\mathbb{R}/\mathbb{Z}$ and $E \in \mathcal{B}$ such that $\mu(E) > 0$. We write $\mu_E = (1/\mu(E)) \cdot \mu|_E$ for the normalized restriction of μ to E.

The next technique relates orbits to the local structure of μ.

THEOREM 2.3. [4] *Let $T : X \to X$ be a continuous map of compact metric space. Let $\mathcal{D}_1, \mathcal{D}_2, \ldots$ be a refining sequence of finite Borel partitions. Let μ be a Borel probability measure on X and assume that* $\sup_{n\in\mathbb{N}}\{\operatorname{diam} T_a^n D : D \in \mathcal{D}_{n+k}, \mu(D) > 0\} \to 0$ *as $k \to \infty$. Then for μ-a.e. x,*

$$\left(\frac{1}{N}\sum_{n=1}^{N}\delta_{T^n x} - \frac{1}{N}\sum_{n=1}^{N} T^n \mu_{\mathcal{D}_n(x)}\right) \xrightarrow[w-*]{N\to\infty} 0.$$

The idea of the proof is to take a countable dense set in $C(X)$ and prove the weak-* convergence with respect to its members. The left average can be replaced with the $\mathcal{D}_{n+k}$-conditional mean by the assumption. The right average is just the $\mathcal{D}_n(x)$-conditional mean. A variant of the ergodic theorem for martingale differences implies that their limits are equal.

The second theorem is about equidistribution along orbits of the form $(n\theta, T^{[\beta n]}x)$, where x is a typical point for μ.

THEOREM 2.4. [4] *Let (X,μ,T) be an ergodic m.p.s. on a compact metric space. Let $\beta > 0$ and $\theta \neq 0$. Then for μ-a.e. x, the sequence $(n\theta, T^{[\beta n]}x)$ equidistributes for a measure ν_x on $[0,1) \times X$ that satisfies $\int \nu_x \, d\mu(x) = \tau \times \mu$, where τ is the invariant measure on $([0,1), R_\theta)$ supported on the orbit closure of* 0.

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

The proof uses an intermediate result which says that for $\tau \times \mu$-a.e. (u, x), the orbit $(n\theta + u, T^{[\beta n]}x)$ equidistributes for a measure $\nu_{u,x}$ on $[0, 1) \times X$ satisfying

$$\int \nu_{u,x}\, dz \times d\mu(u, x) = \tau \times \mu.$$

We will not go into its detail besides mentioning that a suspension of height 1 is used to overcome the integer part issue. Back to the proof of Theorem 2.4, it implies that Lebesgue-a.e. $u \in [0, 1)$ and μ-a.e. x, $\delta_{n\theta+u} \times \delta_{T^{[n\beta]}x} \xrightarrow{w-*} \nu_{u,x}$ and also that $\int \nu_{u,x}\, dz \times \mu(u, x) = m \times \mu$. By translation of the first coordinate, we get that $(n\theta, T^{[\beta n]}x)$ equidistributes for a measure ν_x on $[0, 1) \times X$, where $\nu_x = \nu_{x,0}$, and with Lemma 2.3 from the original paper, we find that $\int \nu_x\, d\mu(x) = \tau \times \mu$.

Lastly, we give a slightly modified version for Hochman's evaluation of the Fourier transform of scaled measures.

LEMMA 2.5. *Let μ be a non-atomic probability measure on $\mathbb{R}$. Then for every $(c, d) \subset \mathbb{R}$ and for every $r > 0$ and $m \neq 0$,*

$$\int_0^1 |\mathcal{F}((S_{b^z}\mu)|_{[c,d]})(m)|^2\, dz \leq \frac{2\mu([c, d])^2}{r \cdot |m|} + \int_c^d \int_c^d \chi_{B_r(y')}(y)\, d\mu(y)\, d\mu(y'),$$

where $B_r(x) = \{y : |x - y| < r\}$.

Proof. (Based on Hochman's proof [4]) Using Fubini,

$$\begin{aligned}\int_0^1 |\mathcal{F}((S_{b^z}\mu)|_{[c,d]})(m)|^2\, dz &= \int_0^1 \left| \int_c^d e(mb^z y)\, d\mu(y)\right|^2 dz \\ &= \int_0^1 \int_c^d \int_c^d e(mb^z y)\overline{e(mb^z y')}\, d\mu(y)\, d\mu(y')\, dz \\ &= \int_0^1 \int_c^d \int_c^d e(mb^z (y - y'))\, d\mu(y)\, d\mu(y')\, dz \\ &= \int_c^d \int_c^d \int_0^1 e(mb^z (y - y'))\, dz\, d\mu(y)\, d\mu(y')\end{aligned}$$

and then changing of variables $t = b^z$,

$$\leq \int_c^d \left(\int_{[c,d]\setminus B_r(y')} \left| \int_1^b \frac{1}{\log(b)t} e(m(y - y')t)\, dt\right| d\mu(y) + \int_{B_r(y')\cap[c,d]} 1\, d\mu(y)\right) d\mu(y').$$

Finally, using integration by parts for the inner integral in the left summand,

$$\leq \frac{2\mu([c, d])^2}{r \cdot |m|} + \int_c^d \int_c^d \chi_{B_r(y')}(y)\, d\mu(y)\, d\mu(y'). \qquad \square$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

3. *Proof of Theorem 1.3*

Let a, b, μ be as in Theorem 1.3 and denote $\alpha = \log b/\log a$. Given a positive integer n, denote $n' = [\alpha n]$ and $z_n = \{\alpha n\} = \alpha n \bmod 1$. That is, $\{z_n\}_{n\in\mathbb{N}}$ is the orbit of $0 \in \mathbb{R}/\mathbb{Z}$ under the irrational rotation by α. Recall that $(\widetilde{\Omega}, \widetilde{\mu}, \widetilde{T}_a)$ is the natural extension of $(\mathbb{R}/\mathbb{Z}, \mu, T_a)$ and $\mu = \int \mu_\omega \, d\widetilde{\mu}(\omega)$ is the disintegration of μ given the past.

Recall that $\mathcal{A}$ denotes the a-adic partition of $[0, 1)$: $\{[k/a, (k+1)/a)\}_{k=0}^{a-1}$ and correspondingly $\mathcal{A}_n = \bigvee_{i=0}^{n-1} T_a^{-i}\mathcal{A}$ is the a-adic partition of generation-n: $\{[k/a^n, (k+1)/a^n)\}_{k=0}^{a^n-1}$. This simple partition is convenient to work with and it can easily be shown that it is a generator for T_a. Naturally, $\mathcal{A}_n(x)$ stands for the nth-generation atom which contains x.

Let f be a non-negative piecewise linear function on $[0, 1)$. We denote the set of its discontinuities. The *minimal jumps oscillation of f* is defined by

$$\mathrm{mjo}(f) = \min_{x\in J} \Big\{ \lim_{x'\in x^+} f(x') - \lim_{x'\in x^-} f(x') \Big\}.$$

Recall the notation $r_0 = T_b^1(1) = \{b\}$, $r_1 = T_b^2(1) = \{b\{b\}\}, \ldots$. Denote $m_b = \inf_n\{r_n : r_n > 0\}$. Since b is specified, we have $0 < m_b \le 1$, as shown in Proposition 2.1. Specifically, $\inf_n \mathrm{mjo}(\{T_b^n\}) = m_b$.

Notice that for every $0 < \theta < 1 - a^{-n'}$, the function $T_b^n \circ R_\theta \circ S_{a^{-n'}}$ is T_b^n composed on the affine map $R_\theta \circ S_{a^{-n'}}(x) = a^{-n'}x + \theta$ of the real line. That is, this composition is T_b^n stretched horizontally by $a^{n'}$ and translated by θ. Hence it is a well-defined piecewise linear map with minimal jumps oscillation which is greater or equal to $\mathrm{mjo}(T_b^n)$. In our notation, it means that mjo $(T_b^n \circ R_\theta \circ S_{a^{-n'}}) \ge m_b$. Notice that $0 < a^{-n'} \cdot b^n$¡a, so $T_b^n \circ R_\theta \circ S_{a^{-n'}}$ also has a uniform slope bounded from above by $0 < a$. These last two properties imply that $T_b^n \circ R_\theta \circ S_{a^{-n'}}$ has at most $\lceil a/m_b \rceil$ discontinuities with a minimal gap of m_b/a between them. Thus, for a sufficiently refined uniform partition of the unit interval, each member of the partition contains at most one discontinuity.

Now we turn to prove Theorem 1.3. Fix a $\widetilde{\mu}$-typical $\omega \in \Omega^-$ and a μ_ω-typical $x \in [0, 1)$. Thus, we want to show asymptotic decay of the $\widetilde{\mu}$-expectation of

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} e_m(T_b^n x) \right|. \tag{3.1}$$

It holds that $T_b^n \mathcal{A}_{n'+l}(x)$ has diameter $O(a^{-l})$ under the metric on $[0, 1) \cong \mathbb{R}/\mathbb{Z}$ and by Theorem 2.3, we have

$$= \limsup_N \left| \frac{1}{N} \sum_{n=1}^{N} \int e_m \, d(T_b^n (\mu_\omega)_{\mathcal{A}_{n'}(x)}) \right|.$$

Recall that $\mathcal{C} = \bigvee_{i=-\infty}^{0} \widetilde{T}_a^i \mathcal{A}$ denotes the σ-algebra in $\widetilde{\Omega}$ generated by projection to the past. Since $\mathcal{C} \vee \mathcal{A}_n = \widetilde{T}_a^n \mathcal{C}$, we have the equivariance relation $T_a^n((\mu_\omega)_{\mathcal{A}_n(x)}) = \mu_{\widetilde{T}_a^n(\omega,x)}$ and also $(\mu_\omega)_{\mathcal{A}_n(x)} = R_{\theta_{\omega,x,n}}(S_{a^{-n}} \mu_{\widetilde{T}_a^n(\omega,x)})$ for some phase $\theta_{\omega,x,n}$. This allows us to write

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

$$= \limsup_{N\to\infty} \left|\frac{1}{N}\sum_{n=1}^{N}\int e_m(T_b^n R_{\theta_{\omega,x,n}} S_{a^{-n'}}(y))\, d\mu_{\tilde{T}_a^{n'}(\omega,x)}\right|$$
$$\le \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\left|\int e_m(T_b^n R_{\theta_{\omega,x,n}} S_{a^{-n'}}(y))\, d\mu_{\tilde{T}_a^{n'}(\omega,x)}\right|. \tag{3.2}$$

If we split the integral above into the sum of integrals on the elements of the uniform partition $\{I_j^k\}_{j=1}^k$ such that $|I_j^k| = 1/k$, we get that for every interval $I_{j_0}^k$ that does not contain a discontinuity, we have

$$\int_{I_{j_0}^k} e_m(T_b^n R_{\theta_{\omega,x,n}} S_{a^{-n'}}(y))\, d\mu_{\tilde{T}_a^{n'}(\omega,x)} = \int_{I_{j_0}^k} e_m(a^{z_n} y + \theta_{j_0,\omega,x,n})\, d\mu_{\tilde{T}_a^{n'}(\omega,x)},$$

where $\theta_{j_0,\omega,x,n}$ is some phase that can be omitted under absolute value. Otherwise, $I_{j_0}^k$ contains a discontinuity and its measure is less than $\sup_j \mu_\omega(I_j^k)$. We take it into account in equation (3.2), denoting $c_{\omega,k} = \lceil a/m_b \rceil \sup_j \mu_\omega(I_j^k)$ and write

$$\le c_{\omega,k} + \limsup_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{k}\left|\int_{I_j^k} e_m(a^{z_n} y + \theta_{j,\omega,x,n})\, d\mu_{\tilde{T}_a^{n'}(\omega,x)}\right|.$$

Now apply Theorem 2.4, after omitting the phases because of the absolute value, to obtain

$$\le c_{\omega,k} + \sum_{j=1}^{k}\int\left|\int_{I_j^k} e_m(a^z y)\, d\mu_\eta\right| d\nu_{\omega,x}(z,\eta). \tag{3.3}$$

If we integrate both sides of the inequality in equations (3.1) and (3.3) with respect to $\widetilde{\mu}$, then by Corollary 2.4, it becomes

$$\int \limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=1}^{m} e_m(T_b^n x)\right| d\widetilde{\mu}(\omega,x)$$
$$\le \int\left(c_{\omega,k} + \sum_{j=1}^{k}\int\left|\int_{I_j^k} e_m(a^z y)\, d\mu_\omega\right|\right) dz\, d\widetilde{\mu}(\omega,x).$$

Next, we apply the Cauchy–Schwartz inequality to get

$$\le \int\left(c_{\omega,k} + \sqrt{k\sum_{j=1}^{k}\int\left|\int_{I_j^k} e_m(a^z y)\, d\mu_\eta\right|^2 dz}\right) d\widetilde{\mu}(\omega,x),$$

where Lemma 2.5 can provide the following evaluation. For any $r > 0$,

$$\le \int\left(c_{\omega,k} + \sqrt{\frac{k\sum_{j=1}^{k}\mu_\eta(I_j^k)^2}{r|m|} + k\iint \chi_{B_r(y)}\, d\mu_\omega(y')\, d\mu_\omega(y)}\right) d\widetilde{\mu}(\omega,x).$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

Notice that the conditional measures μ_ω are continuous because μ was assumed to have positive entropy so by the dominated convergence theorem,

$$\int c_{\omega,k}\, d\widetilde{\mu}(\omega, x) = \left\lceil \frac{a}{m_b} \right\rceil \int \sup_j \mu_\omega(I_j^k)\, d\widetilde{\mu}(\omega, x) \xrightarrow{k\to\infty} 0$$

and similarly,

$$\mathbf{E}_{\mu_\omega}(\mu_\omega(B_r(y))) = \iint \chi_{B_r(y)}(y')\, d\mu_\omega(y')\, d\mu_\omega(y) \xrightarrow{r\to 0} 0.$$

Finally, we can choose $r = |m|^{-1+\epsilon}$ and $k = \min\{|m|^{0.5\epsilon}, \mathbf{E}_{\mu_\omega}(\mu_\omega(B_r(y)))^{-1+\epsilon}\}$ for a small $\epsilon > 0$ in a way that

$$\int c_{\omega,k} + \sqrt{\frac{k\sum_{j=1}^k \mu_\eta(I_j^k)^2}{r|m|} + k\mathbf{E}_{\mu_\omega}(\mu_\omega(B_r(y)))}\, d\widetilde{\mu}(\omega, x) \xrightarrow{m\to\infty} 0. \qquad \square$$

4. *Proof of Theorem 1.5*

The main difference between the previous proof and the quantitative one here is that we assume explicit bounds on $\int c_{\omega,k}\, d\widetilde{\mu}(\omega, x)$ and $\mathbf{E}_{\mu_\omega}(\mu_\omega(B_r(y)))$. In addition, we aim for a pointwise decay instead of mean decay. This may be achieved by relating the decay of a stochastic process to the decay of its mean (as in §4.1).

We begin by repeating our assumptions. Let a, b, μ be as in Theorem 1.3. Denote the uniform partition of the interval $\mathbb{R}/\mathbb{Z}$ into k pieces by $\{I_j^k\}_{j=1}^k$ such that $|I_j^k| = 1/k$ and assume that for some $0 < \alpha$ and for every $k \in \mathbb{N}$,

$$\operatorname*{ess\,sup}_{j,\omega} \mu_\omega(I_j^k) \le O(k^{-\alpha}). \tag{4.1}$$

When this holds, it also imposes a secondary property which is important for us,

$$\operatorname*{ess\,sup}_{\eta\in\Omega^-} \iint \chi_{B_{k^{-1}}(y)}(x)\, d\mu_\eta(x)\, d\mu_\eta(y) \le O(k^{-\alpha}),$$

but it is useful to have here a distinct parameter $\beta \ge \alpha$ such that

$$\operatorname*{ess\,sup}_{\eta\in\Omega^-} \iint \chi_{B_{k^{-1}}(y)}(x)\, d\mu_\eta(x)\, d\mu_\eta(y) \le O(k^{-\beta}). \tag{4.2}$$

The first part of the proof is identical to the previous one but now we think of the limit

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^N e_m(T_b^n x) \right|$$

as a random variable from $\mathbb{R}/\mathbb{Z}$ to $\mathbb{R}$. Again, we fix a $\widetilde{\mu}$-typical $\omega \in \Omega^-$ and a μ_ω-typical $x \in [0, 1)$, and consider

$$\limsup_N \left| \frac{1}{N} \sum_{n=1}^N e_m(T_b^n x) \right|.$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

We repeat the steps in equations (3.1) to (3.3) from the previous proof up to replace $c_{\omega,k}$ with $O(k^{-\alpha})$ according to the condition in equation (4.1). That is,

$$\int \limsup_N \left|\frac{1}{N}\sum_{n=1}^{N} e_m(T_b^n x)\right| d\widetilde{\mu}(\omega, x)$$
$$\leq \int O(k^{-\alpha}) + \sqrt{k\sum_{j=1}^{k}\int\left|\int_{I_j^k} e_m(a^z y)\, d\mu_\eta\right|^2 dz}\, d\widetilde{\mu}(\omega, x).$$

In this case, Lemma 2.5 can provide the following evaluation. For any $r > 0$, using $|I_j^k| = 1/k$ and $\mu_\eta(I_j^k) \leq O(k^{-\alpha})$,

$$\leq \int O(k^{-\alpha}) + \sqrt{\frac{O(k^{1-\alpha})}{r|m|} + k\int \chi_{B_r}(y)\, d\mu_\omega(y')\, d\mu_\omega(y)}\, d\widetilde{\mu}(\omega, x).$$

Finally, denote $r = k^{-\gamma}$ and choose $k = [|m|^\delta]$ for some $0 < \delta$. Then with the condition in equation (4.2),

$$\leq \int O(|m|^{-\delta\alpha}) + \sqrt{O(|m|^{-1-\delta(\alpha-1)+\gamma\delta}) + O(|m|^{\delta(1-\gamma\beta)})}\, d\widetilde{\mu}(\omega, x)$$
$$\leq O(|m|^{\min_{\gamma,\delta>0}\max\{-\delta\alpha,(-1-\delta(\alpha-1)+\gamma\delta)/2,\delta(1-\gamma\beta)/2\}}).$$

Notice that $(-1-\delta(\alpha-1)+\gamma\delta)/2 \geq -0.5$, which means that the referred expression in the exponent is bounded from below by -0.5. An upper bound is achieved under the equality $2\delta\alpha = 1+\delta(\alpha-1)-\gamma\delta = \gamma\delta\beta-\delta$ that leads to $\gamma = (2\alpha+1)/\beta$ and $\delta = \beta/(\beta(1+\alpha)+2\alpha+1)$. Therefore,

$$\leq O(|m|^{-(\alpha\beta/(\beta(1+\alpha)+2\alpha+1))})$$

and the conclusion is that the exponent is strictly bounded between -0.5 and 0. As already mentioned, this decay rate is too slow for applying our second strategy from the introduction.

We remark that both the conditions in equations (4.1) and (4.2) can be relaxed in several ways and mention here one of them as an example. If the conditions in equations (4.1) and (4.2) hold outside a sequence of measurable sets whose measure decays at some known rate, we still promise the mean decay of the Fourier transform at a rate that we are able to bound.

Lastly, although we did not achieve the desired result, we want to make a suggestion of how to progress further in that direction. This will be relevant if someone manages to improve the evaluation of the decay rate from above.

4.1. *Relating the decay rate of a stochastic process to the decay rate of its mean.* Let $(X, \mathcal{B}, \mu)$ be a probability space and $\{A_m\}_{m=1}^\infty$ be a stochastic process. A conventional application of Markov's inequality and the first Borel–Cantelli lemma implies that the decay rate in m of $\{|A_m|\}_{m=1}^\infty$ relates to the decay rate in m of its means $\{\mathbf{E}_\mu(|A_m|)\}_{m=1}^\infty$ in the following way.

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

PROPOSITION 4.1. *If there exists a real* $0 < \alpha$ *such that* $\mathbf{E}_\mu |A_m| < O(m^{-1-\alpha})$*, then for every* $0 < \epsilon < \alpha$*, almost surely* $|A_m| \le o(m^{-\alpha+\epsilon})$.

Proof. Fix $c > 1$. By Markov's inequality $\mathbf{P}_\mu\{|A_m| > (1/c)m^{-\alpha+\epsilon}\} < c\mathbf{E}_\mu|A_m|/m^{-\alpha+\epsilon}$ and from the hypothesis above, there exists some $s > 0$ such that $\mathbf{P}_\mu\{|A_m| > (1/c)m^{-\alpha+\epsilon}\} < (c \cdot s \cdot m^{-1-\alpha})/m^{-\alpha+\epsilon}$. Thus,

$$\mathbf{P}_\mu\bigg\{|A_m| > \frac{1}{c}m^{-\alpha+\epsilon}\bigg\} < \frac{c \cdot s \cdot m^{-1-\alpha}}{m^{-\alpha+\epsilon}} = c \cdot s \cdot m^{-1-\epsilon}$$
$$\Longrightarrow \sum_m \mathbf{P}_\mu\bigg\{|A_m| > \frac{1}{c}m^{-\alpha+\epsilon}\bigg\} < \infty.$$

Finally, by the first Borel–Cantelli lemma, $\mathbf{P}_\mu(\limsup_m\{|A_m| > (1/c)m^{-\alpha+\epsilon}\}) = 0$, which implies that $\{|A_m| > (1/c)m^{-\alpha+\epsilon}\}$ occur finitely often with probability 1. □

5. *Proof of Theorem 1.6*

We begin with providing an equivalent definition of weak-* convergence that will be needed in that section, via the next continuous mapping theorem (see [**1**, Theorem 2.57]).

THEOREM 5.1. $\mu_n \xrightarrow{w-*} \mu$ *if and only if* $\int f\, d\mu_n \xrightarrow{n\to\infty} \int f\, d\mu$ *for every bounded function* $f : X \to \mathbb{R}$ *with* $\mu(\{x : f$ *has a discontinuity at* $x\}) = 0$.

Let x be a μ-typical point. In this part, we assume that x equidistributes along a subsequence $\{N_k\}$ under T_b for some measure where all the hypotheses are as in Problem 1.2 and also that $\{0\}$ is not an atom of this measure. Recall that in our notation, it means that the limit $\lambda_{x,\{N_k\}} = \lim_{k\to\infty}(1/N_k)\sum_{i=0}^{N_k-1}\delta_{T_b^i(x)}$ is well defined. Then we show that $\lambda_{x,\{N_k\}}$ must be T_b-invariant. For convenience, our proof is given as if the equidistribution is along the whole sequence since for equidistribution along a subsequence, the argument is identical. Here we denote the limit measure by λ_∞ for short.

Fix $1 < b \in \mathbb{R}$ and let $x \in \mathbb{R}/\mathbb{Z}$ be a μ-typical point and denote

$$\lambda_N = \frac{1}{N}\sum_{i=0}^{N-1}\delta_{T_b^i x} \xrightarrow{w-*} \lim_{N\to\infty}\frac{1}{N}\sum_{i=0}^{N-1}\delta_{T_b^i x} = \lambda_\infty.$$

We need to show that

$$\int f\, d\lambda_\infty = \int f\, dT_{b*}\lambda_\infty$$

for all $f \in C(\mathbb{R}/\mathbb{Z})$.

We have

$$\lambda_N - T_{b*}\lambda_N = \frac{1}{N}\sum_{i=0}^{N-1}\delta_{T_b^i(x)} - \frac{1}{N}\sum_{i=0}^{N-1}\delta_{T_b^{i+1}(x)} = \frac{\delta_x - \delta_{T_b^N(x)}}{N} \xrightarrow{N\to\infty} 0,$$

so $\lim_{N\to\infty}(\lambda_N - T_{b*}\lambda_N) = 0$.

Additionally, for every $f \in C(\mathbb{R}/\mathbb{Z})$, $f \circ T_b$ has at most one discontinuity which is located at 0 and by assumption, $\lambda_\infty(\{0\}) = 0$. Since $\lambda_N \xrightarrow{w-*} \lambda_\infty$ for any $f \in C(\mathbb{R}/\mathbb{Z})$, then by Theorem 5.1,

$$\mathbf{E}_{T_{b*}\lambda_N}(f) = \mathbf{E}_{\lambda_N}(f \circ T_b) \xrightarrow{N\to\infty} \mathbf{E}_{\lambda_\infty}(f \circ T_b) = \mathbf{E}_{T_{b*}\lambda_\infty}(f).$$

Now we can stitch it all together and get for every $f \in C(\mathbb{R}/\mathbb{Z})$ that

$$\begin{aligned}|\mathbf{E}_{\lambda_\infty}(f) - \mathbf{E}_{T_{b*}\lambda_\infty}(f)| \leq &|\mathbf{E}_{\lambda_\infty}(f) - \mathbf{E}_{\lambda_N}(f)| \\ &+ |\mathbf{E}_{\lambda_N}(f) - \mathbf{E}_{T_{b*}\lambda_N}(f)| \\ &+ |\mathbf{E}_{T_{b*}\lambda_N}(f) - \mathbf{E}_{T_{b*}\lambda_\infty}(f)| \to 0\end{aligned}$$

as N tends to infinity which implies the desired result.

6. *Exploring conditions in equations (4.1) and (4.2)*

In this section, we inquire the validity range of conditions in equations (4.1) and (4.2). We added some brief theoretical background that will be used for that subjective.

6.1. *Stationary coding.*

The content of this subsection is taken from the book of P. Shields on ergodic theory [12].

Let A and B be finite sets. A Borel measurable map $F : A^{\mathbb{Z}} \to B^{\mathbb{Z}}$ is a *stationary coder* if $F(T_A x) = T_B F(x)$ for every $x \in A^{\mathbb{Z}}$, where T_A and T_B denote the shifts on the respective two-sided sequence spaces $A^{\mathbb{Z}}$ and $B^{\mathbb{Z}}$. Stationary coding carries a Borel measure μ on $A^{\mathbb{Z}}$ into the measure $\nu = \mu \circ F^{-1}$ defined for Borel subsets C of $B^{\mathbb{Z}}$ by $\nu(C) = \mu(F^{-1}(C))$. The encoded process ν is said to be a *stationary coding* of μ, with coder F. It is immediate to conclude that a stationary coding of a stationary process is itself a stationary process.

The map $f : A^{\mathbb{Z}} \to B$ defined by the formula $f(x) = F(x)_0$ is the *time-zero coder* associated with F. Notice that stationary coding preserves ergodicity.

Let T be an invertible, ergodic transformation of the probability space (X, Σ, μ). Let S be a set of positive measure. Denote by $\{R_n\}$ the $(T, \mathcal{A}_S)$-process defined by the partition $\mathcal{A}_S = \{S, X \setminus S\}$ with $X \setminus S$ labeled by 0 and S labeled by 1. That is, $R_n(x) = \chi_S(T^n x)$ for every integer $n \in \mathbb{Z}$. Notice that $\{R_n\}$ is a stationary coding of the $(T, \mathcal{A}_S)$-process with time zero coder χ_S and therefore it is ergodic.

Notice that such a non-trivial $\{R_n\}$ process (i.e $0 < \mu(S) < 1$) has positive entropy where (X, Σ, μ) is a product measure on $\mathcal{A}_S^{\mathbb{Z}}$. This is because a non-trivial $(T, \{S, X \setminus S\})$-process is a factor of Bernoulli measure which has completely positive entropy.

6.2. *Reverse Markov's inequality.*

The next reverse version of Markov's inequality is famous and easy to prove. We present it here with a proof for completeness.

THEOREM 6.1. (Reverse Markov's inequality) *Let X be a random variable on a probability space $(\Omega, \boldsymbol{P})$ that satisfies $\boldsymbol{P}(X \leq a) = 1$ for some constant a. Then, for $d < \boldsymbol{E}(X)$,*

$$\boldsymbol{P}(X > d) \geq \frac{\boldsymbol{E}(X) - d}{a - d}.$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

Proof. Define the random variable $V = a - X$ which is almost surely non-negative by assumption. The event $\{X \le d\}$ is equivalent to the event $\{V \ge a - d\}$. Now with Markov's inequality,

$$\mathbf{P}(\{X \le d\}) = \mathbf{P}(\{V \ge a - d\}) \le \frac{\mathbf{E}(V)}{a - d} = \frac{a - \mathbf{E}(X)}{a - d},$$

where the right-hand side numerator and denominator are strictly positive since $d < \mathbf{E}(X) \le a$. Finally,

$$\mathbf{P}(\{X > d\}) = 1 - \mathbf{P}(\{X \le d\}) > \frac{\mathbf{E}(X) - d}{a - d}. \qquad \square$$

6.3. *Not all the μ values meet the condition in equation (4.2).* Here we construct a counterexample which violates the weaker condition in equation (4.2). We anticipate that the condition in equation (4.1) can be violated by simpler examples.

We use the formalism from §§2.2 and 6.1. Let $(X_l, \Sigma, \nu, \sigma)$ be a uniformly distributed full Bernoulli shift in $2 < l$ symbols. We define a measurable set $\mathcal{Y} \in \Sigma$ recursively.

Fix $0 < \epsilon < 0.5$. For $k = 1$, fix $2 \le n_1 \in \mathbb{N}$ such that $\log(n_1)^{-1} < 1 - \epsilon$ and let Y_1 be a set such that $0.5 \log(n_1)^{-2} < \nu(Y_1) < \log(n_1)^{-2}$. Define $\mathcal{Y}_1 = \bigcup_{m=0}^{[\log n_1]} \sigma^{-m} Y_1$ so that $0 < \nu(\mathcal{Y}_1) \le \log(n_1)^{-1}$ and also $\nu(\mathcal{Y}_1) < 1 - \epsilon$.

For $1 < k \in \mathbb{N}$, choose $n_{k-1} < n_k \in \mathbb{N}$ such that $\log(n_k)^{-1} < 1 - \epsilon - \sum_{i=1}^{k-1} \nu(\mathcal{Y}_i)$ and let Y_k be a set such that $0.5 \cdot \log(n_k)^{-2} < \nu(Y_k) < \log(n_k)^{-2}$. Define $\mathcal{Y}_k = \bigcup_{m=0}^{[\log n_k]} \sigma^{-m} Y_k$ so that $0 < \nu(\mathcal{Y}_k) \le 1/\log n_k \le 1 - \epsilon - \sum_{i=1}^{k-1} \nu(\mathcal{Y}_i)$ and therefore $0 < \sum_{i=1}^{k} \nu(\mathcal{Y}_i) < 1 - \epsilon$.

Finally, define $\mathcal{Y} = \bigcup_{i=1}^{\infty} \mathcal{Y}_i$ and define a process on $\{0, 1\}^{\mathbb{Z}}$ by $R_n(x) = \chi_{\mathcal{Y}}(\sigma^{n-1} x)$ with the induced measure $\widetilde{\mu} = \nu \circ R^{-1}$ as in §6.1. By our construction, $0 < \nu(\mathcal{Y}) < 1 - \epsilon$ and $\{n_k\}$ is strictly increasing. Since (X_l, Σ, ν) is ergodic, so is $\{R_n\}$ as a stationary coding with time zero coder (which trivially also preserves the measure). That is, we can interpret $\widetilde{\mu}$ as a probability measure which is ergodic and shift invariant on the binary full shift. It also has positive entropy by §6.1.

In another direction, notice that $[0, 1)$ is the image of the infinite binary sequences $\{0, 1\}^{\mathbb{N}}$ under the natural map $(x_i)_{i=1}^{\infty} \mapsto \sum_{i=1}^{\infty} (x_i / 2^i)$ which is also a bijection on the complement of a countable set, that is, a bijection on the complement of a null set for every continuous measure. Thus, if we denote $X_2 = \{0, 1\}^{\mathbb{Z}}$ (the 2-shift), then we can define up to a null set the bijection $\iota : X_2 \to \{0, 1\}^{-\mathbb{N}_0} \times [0, 1)$. We can also define the pushforward $\widetilde{\rho} = \widetilde{\mu} \circ \iota^{-1}$ (that is, a measure on $\{0, 1\}^{-\mathbb{N}_0} \times [0, 1)$) which preserves all the relevant properties of $\widetilde{\mu}$ (ergodicity is trivial and positive entropy by being a non-trivial factor of the system with completely positive entropy). Now, $\widetilde{\rho} = \int \rho_\eta \, d\widetilde{\rho}(\eta)$, where it is disintegrated with respect to the σ-algebra that is generated by projection to the past (like in §1.3) and the conditional measures ρ_η identified as measures on $[0, 1)$. Similarly, $\widetilde{\mu} = \int d\mu_\eta \, d\widetilde{\mu}(\eta)$, where $\eta \in \{0, 1\}^{-\mathbb{N}_0}$.

For every $k \in \mathbb{N}$, denote the event $\bigcap_{i=0}^{[\log n_k]} \{R_i = 1\}$ by E_k. We begin with

$$\int \int \int \chi_{B_{n_k^{-1}}(y)}(x) \, d\rho_\eta(x) \, d\rho_\eta(y) \, d\widetilde{\rho}(\eta),$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

where $B_{n_k^{-1}}(y) = \{x \in [0,1) : |x-y| < n_k^{-1}\}$. We need to show that this integral decays in n_k with sub polynomial rate. We can pullback this integration to the space X_2 and restrict it to the event $E_k \times E_k$ since every pair $(x', y') \in E_k \times E_k$ corresponds to a pair $(x, y) \in [0,1)^2$ with $|x-y| < 1/n_k$. Thus,

$$\begin{aligned}
&\geq \int\iint \chi_{E_k\times E_k}(x,y)\,d\mu_\eta \times \mu_\eta(x,y)\,d\widetilde{\mu}(\eta)\\
&= \int\int \chi_{E_k}(x)\,d\mu_\eta(x)\int \chi_{E_k}(y)\,d\mu_\eta(y)\,d\widetilde{\mu}(\eta)\\
&= \int\bigg(\int \chi_{E_k}(x)\,d\mu_\eta(x)\bigg)^2\,d\widetilde{\mu}(\eta)\\
&\geq \bigg(\int\int \chi_{E_k}(x)\,d\mu_\eta(x)\,d\widetilde{\mu}(\eta)\bigg)^2\\
&= \bigg(\int \chi_{E_k}(x)\,d\widetilde{\mu}(x)\bigg)^2,
\end{aligned}$$

where in the first equality, we used Fubini and then split the integral into two separate integrals, and the second inequality is Jensen. We can pullback the integration once again to the full shift X_l and restrict it to the event Y_k which included in the preimage of E_k,

$$\begin{aligned}
&\geq \bigg(\int \chi_{Y_k}(x)\,d\nu(x)\bigg)^2\\
&= 0.25\cdot \log(n_k)^{-4}.
\end{aligned}$$

Now, for every $k \in \mathbb{N}$, we can use the reverse Markov inequality (Theorem 6.1) with $a = 2$ and $d = 1/8\log(n_k)^4$ to get

$$\mathbf{P}_{\widetilde{\mu}}\bigg(\bigg\{\int \mu_\eta(B_{n_k^{-1}}(y))\,d\mu_\eta(y) \geq \frac{1}{8\log(n_k)^4}\bigg\}\bigg) \geq \frac{1}{16\log(n_k)^4}.$$

This violates the condition in equation (4.2).

6.4. *Processes with memory of finite length meet the condition in equation (4.1).* In this special case, it will be enough to use the total probability formula to reach the condition in equation (4.1) which also implies that the condition in equation (4.2) holds.

Let A be a finite alphabet $|A| = l$ for some integer $l \geq 2$ and let $(A^{\mathbb{Z}}, \mu, T)$ be a finite memory length symbolic process which is ergodic, invariant, and with positive entropy with respect to T. That is, the prediction of A_0 depends only on a finite portion of the past $A_{-n}, \ldots, A_{-1}$ for some $n < \infty$ or explicitly for every $m \in \mathbb{Z}$ and $x \in A^{\mathbb{Z}}$, $\mu([x_{m+n}^{\infty}]|[x_{-\infty}^{m}]) = \mu([x_{m+n}^{\infty}])$. Denote $s = \max_{a\in A}\mu([a]) < 1$. Assume that $n < m$, then with direct computation,

$$\begin{aligned}
\mu_{[x_{-\infty}^0]}([x_0^m]) &\leq \mu([x_n^m])\\
&= \mu([x_n])\mu([x_{n+1}^m]|[x_n])
\end{aligned}$$

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press

$$
\begin{aligned}
&\leq \mu([x_n])\mu([x_{2n}^m]) \\
&\vdots \\
&= \prod_{i=1}^{[m/n]} \mu([x_{n\cdot i}]) \\
&\leq s^{m/n+1}.
\end{aligned}
$$

If we consider a uniform partition of the unit interval $\{I_j^k\}_{j=1}^k$ such that $|I_j^k| = 1/k$ and where $k = l^m$, then the display above implies that $\operatorname{ess\,sup}_{j,\omega} \mu_\omega(I_j^k) \leq O(k^{\log(s)/(n \log(l))})$.

7. *Rajchman and invariance does not force uniqueness*

7.1. *Self-similarity and classes of algebraic numbers.* Here we set out some of the basic definitions regarding self similarity. A set $\mathcal{F} = \{f_1, \ldots f_k\}$ for $k \geq 2$ of contractions on $\mathbb{R}$, $f_i(x) = r_i x + a_i$, with $0 < r_i < 1$ for each $i \in \{1, \ldots, k\}$, is an *iterated function system* or *IFS* for short. We also call the f_i terms *similarities*.

A key fact in this topic is that there exists a unique non-empty compact set $K \subset \mathbb{R}$ such that $K = \bigcup_{i=1}^k f_i(K)$. We call it the *attractor* or the *self-similar set* of the IFS $\mathcal{F}$.

Given a list of positive numbers $\boldsymbol{p} = (p_1, \ldots, p_k)$ with $\sum_{i=1}^k p_i = 1$, we call it a *positive probability vector* and there is a unique probability measure $\mu_{\boldsymbol{p}}$ with $\mu_{\boldsymbol{p}} = \sum_{i=1}^k p_i \cdot f_i \mu_{\boldsymbol{p}}$. This measure is supported on the attractor of $\mathcal{F}$ and called a *self-similar measure*.

A *Pisot number* is a real algebraic integer greater than 1 all of whose Galois conjugates are less than 1 in absolute value. A *Salem number* is a real algebraic integer greater than 1 whose all conjugate roots have absolute value no greater than 1, and at least one of them has an absolute value which equals 1. There are countably many Pisot and Salem numbers.

We conclude this part with a recent theorem of Varjú and Yu [13] which plays a major role in our proof that T_b-invariant Rajchman measure need not be Parry.

THEOREM 7.1. *Let $k \geq 2$ be an integer. Let $r_1 = r^{l_1}, \ldots, r_k = r^{l_k}$ for some $r \in (0, 1)$ and $l_1, \ldots, l_k \in \mathbb{Z}_{>0}$ with $\gcd(l_1, \ldots, l_k) = 1$. Assume that r^{-1} is not a Pisot or Salem number. Let μ be a non-singleton self-similar measure associated to the IFS $\mathcal{F} = \{f_1, \ldots, f_k\}$ ($f_i(x) = r_i x + a_i$) and a positive probability measure. Then,*

$$|\hat{\mu}(z)| = O(|\log(z)^{-c}|)$$

for some $c > 0$.

We use Theorem 7.1 to prove that a Rajchman T_b-invariant measure must not be unique. Let $b \geq 2$ be a specified non-simple number which is not Salem or Pisot. We can guarantee the existence of such a number by cardinality considerations (recall that Pisot numbers as well as Salem numbers and simple numbers are countable while specified numbers have the cardinality of continuum).

Denote $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$ for short and denote the b-expansion of b by (a_i). That is, $b = a_0 + a_1/b + \cdots$ and in particular $a_0 = [b]$. Every sequence $(x_i) \in \{0, 1\}^{\mathbb{N}_0}$ and $n \in \mathbb{N}$ satisfies $(x_{i+n}) < (a_i)$ since $x_n \in \{0, 1\}$ which is either way less than $[b] = a_0$.

Recall that $X_b \subset \Lambda_b^{\mathbb{N}_0}$ is the subset of sequences that encodes b-expansions, so by Parry's criterion (Theorem 2.2), we get that $\{0, 1\}^{\mathbb{N}_0} \subset X_b$.

Define the set $K = \{\sum_{i=0}^{\infty}(b_i/b^i) : (b_i) \in \{0, 1\}^{\mathbb{N}_0}\}$ which is the image of $\{0, 1\}^{\mathbb{N}}$ under the b-expansion and define the IFS $\mathcal{F} = \{f_0(x) = x/b,\ f_1(x) = x/b + 1/b\}$. If we denote concatenation of symbols by **.**, then clearly $\{0, 1\}^{\mathbb{N}_0} = 0 \,\mathbf{.}\, \{0, 1\}^{\mathbb{N}_0} \cup 1 \,\mathbf{.}\, \{0, 1\}^{\mathbb{N}_0}$, which means under the b-expansion, $K = \bigcup_{i=0}^{1} f_i(K)$. Thus, K is the attractor of $\mathcal{F}$ and for every positive probability vector $\boldsymbol{p} = (p_0, p_1)$, the self-similar measure that satisfies $\mu_{\boldsymbol{p}} = \sum_{i=0}^{1} p_i \cdot f_i\mu_{\boldsymbol{p}}$ is well defined. Pulling it back again under the b-expansion, it translated into a Bernoulli shift $\boldsymbol{p}^{\mathbb{N}_0}$ on $X_b|_{\{0,1\}^{\mathbb{N}_0}} = \{0, 1\}^{\mathbb{N}_0}$, where the former is T_b-invariant and correspondingly the latter is shift invariant.

If we consider Theorem 7.1 where both of the l_i terms are equal to 1 such that $\gcd(l_0, l_1) = 1$, then all the hypotheses hold with respect to b and $\mathcal{F}$ and any self-similar measure $\mu_{\boldsymbol{p}}$ provides the Rajchman property,

$$|\hat{\mu}_{\boldsymbol{p}}(z)| = O(|\log(z)^{-c}|).$$

Since there are infinitely many of them, the uniqueness is violated.

Acknowledgements. First, I would like to thank Prof. B. (Benjy) Weiss for many helpful discussions about dynamics, and especially for his suggestion regarding the example in §6.3. An earlier version of this study was submitted at the Hebrew University as part of an academic degree. It was done under formal guidance of Prof. M. Hochman, for which I am grateful. Lastly, I would also like to thank Prof. Z. Wang for a thorough review of this work and Prof. S. Baker for providing advice on the example in §7.

REFERENCES

[1] P. Billingsley. *Probability and Measure*, 2nd edn. Wiley, New York, 1986.
[2] I. P. Cornfeld, S. V. Fomin and Y. G. Sinai. *Ergodic Theory*. Springer, Berlin, 1982.
[3] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation. *Math. Syst. Theory* **1** (1967), 1–49.
[4] M. Hochman. A short proof of Host's equidistribution theorem. *Israel J. Math.* **251** (2022), 527–539.
[5] M. Hochman and P. Shmerkin. Equidistribution from fractal measures. *Invent. Math.* **202** (2015), 427–479.
[6] B. Host. Nombres normaux, entropie, translations. *Israel J. Math.* **91** (1995), 419–428.
[7] A. S. A. Johnson. Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers. *Israel J. Math.* **77** (1992), 211–240.
[8] R. Lyons. Seventy years of Rajchman measures. *J. Fourier Anal. Appl.* (1995), 363–377 (Kahane Special Issue).
[9] W. Parry. On the β-expansions of real numbers. *Acta Math. Hungar.* **11** (1960), 401–416.
[10] D. J. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergod. Th. & Dynam. Sys.* **10**(2) (1990), 395–406.
[11] J. Schmeling. Symbolic dynamics for β-shifts and self-normal numbers. *Ergod. Th. & Dynam. Sys.* **17**(3) (1997), 675–694.
[12] P. C. Shields. *The Ergodic Theory of Discrete Sample Paths (Graduate Studies in Mathematics, 13)*. American Mathematical Society, Providence, RI, 1996.
[13] P. Varjú and H. Yu. Fourier decay of self-similar measures and self-similar sets of uniqueness. *Anal. PDE* **15**(3) (2020), 843–858.
[14] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79)*. Springer, New York, 2000.

https://doi.org/10.1017/etds.2023.75 Published online by Cambridge University Press