RANDOM WALKS ON TORI AND NORMAL NUMBERS IN SELF-SIMILAR SETS

YIFTACH DAYAN, ARIJIT GANGULY AND BARAK WEISS

Abstract. We study random walks on a d-dimensional torus by affine expanding maps whose linear parts commute. Assuming an irrationality condition on their translation parts, we prove that the Haar measure is the unique stationary measure. We deduce that if $K \subset \mathbb{R}^d$ is an attractor of a finite iterated function system of $n \geqslant 2$ maps of the form $x \mapsto \frac{1}{D} \cdot x + t_i$ $(i = 1, \dots, n)$, where $D \geqslant 2$ is an integer and is the same for all the maps, under an irrationality condition on the translation parts t_i, almost every point in K (w.r.t. any Bernoulli measure) has an equidistributed orbit under the map $x \mapsto Dx$ (multiplication mod $\mathbb{Z}^d$). In the one-dimensional case, this conclusion amounts to normality to base D. Thus for example, almost every point in an irrational dilation of the middle-thirds Cantor set is normal to base 3.

1. Introduction

In this paper we will analyze random walks on a torus. For some random walks, driven by finitely many expanding affine maps of the torus we will show that the only stationary measure is the Haar measure, and hence for any starting point, almost every random trajectory is equidistributed. We will use the random walks results to obtain new results on digital expansion of typical points on self-similar sets. This paper follows a scheme similar to that of [19], where related results about random walks on homogeneous spaces were proved, leading to results on Diophantine properties of typical points on self-similar sets.

1.1. Random walks on tori.

Informally, a random walk on a torus may be described as follows. Suppose G is a semigroup acting on the torus and μ is some probability measure on G. Given a point x in the torus, the random walk proceeds by sampling a random element $g \in G$ according to μ and moving the point x to gx. The process continues indefinitely to obtain an infinite random path in the torus.

More formally, let G be a second countable locally compact semigroup acting on $\mathbb{T}^d \stackrel{\text{def}}{=} \mathbb{R}^d/\mathbb{Z}^d$ and let μ be some Borel probability measure on G. To this system we associate a Bernoulli shift $(B, \beta, \mathcal{B}, T)$, where $B = G^{\mathbb{N}}$, $\beta = \mu^{\otimes \mathbb{N}}$ is the product measure on B, $\mathcal{B}$ is the Borel σ-algebra on B and T is the left shift. For a measure ν on $\mathbb{T}^d$, the convolution of μ with ν is the measure on $\mathbb{T}^d$ which is given by

$$\mu * \nu (A) = \int_G g_* \nu (A) \, d\mu (g) \,,$$

for every measurable set $A \subseteq \mathbb{T}^d$. A probability measure ν for which $\mu * \nu = \nu$ is called *μ-stationary.*

Clearly, every G-invariant measure is μ-stationary, but the converse is often false. The action is called *stiff* if any μ-stationary measure is invariant (this terminology was introduced by Furstenberg in [13]). Recently there have been several breakthroughs for the case where G acts on $\mathbb{T}^d$ by linear automorphisms. Starting with the work of Bourgain, Furman, Lindenstrauss and Mozes [5], followed by a series of papers by Benoist and Quint [1, 2, 3], these results gave certain conditions guaranteeing stiffness.

In this paper we establish stiffness for certain random walks generated by *affine toral endormorphisms*, namely maps $\mathbb{T}^d \to \mathbb{T}^d$ of the form $x \mapsto D(x) + \alpha$, where D is a linear toral endomorphism, $\alpha \in \mathbb{T}$ and addition is the group law on $\mathbb{T}^d$. Moreover we show that these stiff random walks have a unique invariant measure. We recall that a toral endomorphism is a map $\mathbb{T}^d \to \mathbb{T}^d$ of the form $x \mapsto Dx$ (mod
$mathbbZ^d$), where D is a matrix with integer coefficients. In this paper we will use the same letter to denote a toral endomorphism and the corresponding integer matrix. We denote the identity $d \times d$ matrix by $\mathbb{I}_d$. We call a matrix *expanding* if all of its (complex) eigenvalues have modulus greater than 1.

Theorem 1. *Let $D_1, ..., D_n$ be commuting $d \times d$ matrices with coefficients in $\mathbb{Z}$. Assume that all the D_i are expanding. For $i = 1, \ldots, n$, let $\alpha_i \in \mathbb{R}^d$ and let*

$$h_i : \mathbb{T}^d \to \mathbb{T}^d, \quad h_i(x) = D_i(x) + \alpha_i.$$

Assume that

$$\{(\mathbb{I}_d - D_i)\,\alpha_j - (\mathbb{I}_d - D_j)\,\alpha_i : \, i, j \in \{1, ..., n\}\}$$

is not contained in any proper closed subgroup of $\mathbb{T}^d$. Let μ be a probability measure such that $\operatorname{supp}\mu = \{h_1, ..., h_n\}$. *Then Haar measure is the unique μ-stationary measure on $\mathbb{T}^d$.*

The uniqueness property of a stationary measure for a random walk has strong consequences. Indeed, using Breiman's law of large numbers ([6], see also [4, Chap. 2.2]), in the setting of Theorem 1, one obtains that for every $x \in \mathbb{T}^d$ and every $\varphi \in C\left(\mathbb{T}^d\right)$, for β - a.e. $b \in B$,

$$\frac{1}{N}\sum_{k=0}^{N-1} \varphi\left(b_k \cdots b_1 x\right) \underset{N\to\infty}{\longrightarrow} \int\limits_{\mathbb{T}^d} \varphi \, d\mathrm{Haar}.$$

By a standard argument using the separability of the space $C\left(\mathbb{T}^d\right)$, we obtain that for every starting point, a.e. trajectory is uniformly distributed. That is:

Corollary 2. *For every $x \in \mathbb{T}^d$, for β-a.e. $b \in B$,*

$$\frac{1}{N}\sum_{k=0}^{N-1} \delta_{b_k \cdots b_1 x} \underset{N\to\infty}{\longrightarrow} \mathrm{Haar}$$

in the weak- topology.*

1.2. **Normal numbers in self similar sets.**

1.2.1. *Self similar sets.* A *contraction iterated function system (IFS)* is a finite collection of maps $\{\varphi_i\}_{i\in\Lambda}$, where for each $i \in \Lambda$, $\varphi_i : \mathbb{R}^d \to \mathbb{R}^d$ is given by $\varphi_i(x) = r_i \cdot x + \alpha_i$, for some $r_i \in (0, 1)$ which is called the *contraction ratio of* φ_i, and some $\alpha_i \in \mathbb{R}^d$ which is called the *translation of* φ_i. For every such IFS there is a unique non-empty compact set $K \subseteq \mathbb{R}^d$ which satisfies

$$K = \bigcup_{i\in\Lambda} \varphi_i(K),$$

and is called the *attractor* of the IFS. We will refer to an attractor of a contraction IFS as a *self-similar set*. Every point in the attractor K of an IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ has a symbolic coding (possibly more than one), given by the so called coding map $\pi_\Phi : \Lambda^{\mathbb{N}} \to K$, which may be defined by

$$\forall i = (i_1, i_2, ...) \in \Lambda^{\mathbb{N}},\ \pi_\Phi(i) = \lim_{n\to\infty} \varphi_{i_1} \circ \varphi_{i_2} \circ \cdots \circ \varphi_{i_n}(x_0),$$

where x_0 is some arbitrary basepoint. It will be convenient for us to choose $x_0 = 0$. For an introduction see [10].

We will say that a measure on K is a *Bernoulli measure* if it is obtained by pushing forward a Bernoulli measure on the symbol space $\Lambda^{\mathbb{N}}$ by the coding map π_Φ. In other words, the measure is of the form $(\pi_\Phi)_* P^{\otimes\mathbb{N}}$, where P is a probability measure on the finite set Λ and $P^{\otimes\mathbb{N}}$ is the product measure on $\Lambda^{\mathbb{N}}$. Throughout this text, we assume that $P(\{i\}) > 0$ for every $i \in \Lambda$ (otherwise we can replace Λ with $\mathrm{supp}(P)$). This definition depends on the underlying IFS (rather than its attractor), and thus we shall sometimes refer to a measure as a Φ-*Bernoulli measure*, where Φ indicates the IFS.

1.2.2. *Normal numbers.* Let $D \geqslant 2$ be an integer. Recall that $x \in \mathbb{R}$ is called *normal to base* D if for every $n \in \mathbb{N}$, every finite word $\omega \in \{0, ..., D-1\}^n$ occurs in the base D digital expansion of x with asymptotic frequency D^{-n}. Equivalently, $x \in \mathbb{R}$ is normal to base D iff the forward orbit of x under the map $x \mapsto Dx$ (multiplication by D modulo 1) is equidistributed w.r.t. Lebesgue measure in $[0, 1]$. A detailed exposition on normal numbers, and in particular for the equivalence stated above, may be found in [8]. One useful property of normal numbers is that a number $x \in \mathbb{R}$ is normal to some base D iff for every $s, t \in \mathbb{Q}$ s.t. $s \neq 0$, $sx + t$ is normal to base D (this property was proved by Wall in his Ph.D. thesis [20]).

Since the map $x \mapsto Dx$ is ergodic (w.r.t. Lebesgue measure on $[0, 1]$), by Birkhoff's ergodic theorem, a.e. real number is normal to every integer base[1]. Focusing attention to self-similar sets, one may inquire as to the size of the set of all numbers within some self similar set that are normal to a given base. It was proved in [7] that the set of real numbers which are not normal to any integer base (these numbers are called *absolutely non-normal*) intersects any infinite self-similar $K \subseteq \mathbb{R}$, in a set whose Hausdorff dimension is equal to that of K. This extends a result of Schmidt [18], which provides the same conclusion for $K = [0, 1]$, to nice self-similar sets.

On the other hand, in many cases, with respect to natural measures supported on self-similar sets in $\mathbb{R}$, almost every number is normal to a given base D. Of course this is not the case for every self-similar set and every base. For example, no number in the middle-thirds Cantor set is normal to base 3. Several positive results in this direction were obtained in [9, 17, 11, 15, 14]. In all of these papers, some independence is assumed to hold, between the contraction ratios of the IFS and the base D. In the context of self-similar sets, the strongest assertion is the following theorem proved in [14]:

Theorem 3 (Hochman-Shmerkin). *Let $K \subseteq \mathbb{R}$ be the attractor of a contraction IFS $\Phi = \{\varphi_i\}_{i\in\Lambda}$ with contraction ratios r_i (for $i \in \Lambda$), and let μ be a Φ-Bernoulli measure on K. Assume that Φ satisfies the open set condition. Then for every integer $D \geqslant 2$ satisfying*

$$\text{there is } i \in \Lambda \text{ for which } \frac{\log(r_i)}{\log(D)} \notin \mathbb{Q}, \tag{1.1}$$

[1]This fact was first proved by É. Borel in 1909 without using ergodic theory.

μ-a.e. number is normal to base D.

This is a special case of more general results proved in [14].

1.2.3. *New results.* In this work we deal with the opposite situation to the one treated in Theorem 3. Instead of assuming that at least one contraction ratio of the IFS is multiplicatively independent of the base D, we assume that all the contraction ratios of the IFS are integer powers of D. As a substitute for the 'independence' assumption (1.1), we impose an irrationality condition on the translations functions in the IFS. More precisely we prove the following:

Theorem 4. *Let K be the attractor of a contraction IFS $\Phi \overset{\text{def}}{=} \{f_1, ..., f_k\}$, where for each $i \in \Lambda \overset{\text{def}}{=} \{1, ..., k\}$, $f_i : \mathbb{R}^d \to \mathbb{R}^d$ is given by*

$$f_i(x) = \frac{1}{D^{r_i}} x + t_i,$$

for some $t_i \in \mathbb{R}^d$, $r_i \in \mathbb{N}$ and some integer $D \geqslant 2$. Assume that

$$\{D^{r_j} t_i - D^{r_i} t_j : i, j \in \Lambda\} \text{ is not contained in any proper closed subgroup of } \mathbb{T}^d. \tag{1.2}$$

Then, for every Φ-Bernoulli measure μ on K, for μ-almost every $x \in K$, the sequence $\{D^m x\}_{m=1}^{\infty}$ is equidistributed in $\mathbb{T}^d$ with respect to Haar measure.

Note that in the one-dimensional case, and when $r_1 = \cdots = r_k = 1$, assumption (1.2) is equivalent to assuming that there exists a pair $i, j \in \Lambda$ s.t. $t_i - t_j \notin \mathbb{Q}$, and in case this condition holds, then w.r.t. any Φ-Bernoulli measure on K, almost every number is normal to base D.

As an example, denote by $\mathcal{C} \subset [0, 1]$ the middle-thirds Cantor set and let $K \overset{\text{def}}{=} \alpha \cdot \mathcal{C} + \beta$, where α is any irrational number and $\beta \in \mathbb{R}$. Then K is the attractor of $\Phi = \left\{x \mapsto \frac{1}{3}x + \beta,\ x \mapsto \frac{1}{3}x + \frac{2\alpha}{3} + \beta\right\}$ having a uniform contraction ratio of 3^{-1}. While prior results did not provide information regarding normality of typical points in K to base 3, from Theorem 4 one may deduce that w.r.t. any Φ-Bernoulli measure on K, almost every point is normal to base 3. Thus Theorem 4 complements Theorem 3 of Hochman-Shmerkin, and combining them we have the following corollary.

Corollary 5. *With the above notations and assumptions, with respect to any Φ-Bernoulli measure on $K = \alpha \cdot \mathcal{C}$, a.e. point is normal to every integer base $D \geqslant 2$.*

Remark 6. One useful property of normal numbers is that a number x is normal to some integer base $D \geqslant 2$ if and only if it is normal to base D^s for every positive integer s ([8, Thm. 4.4]). Using this fact, given a uniform contraction ratio $\frac{1}{D}$, for every $n \in \mathbb{N}$ one may consider the IFS $\Phi^n \overset{\text{def}}{=} \{f_{i_1} \circ \cdots \circ f_{i_n} : (i_1, ..., i_n) \in \Lambda^n\}$. Obviously, Φ^n and Φ have the same attractor, and it is not hard to show that Φ satisfies (1.2) if and only if Φ^n does. Hence Theorem 4 holds for all Φ^n-Bernoulli measures, for every positive integer n.

Remark 7. It would be interesting to extend Theorem 3 to the more general class of attractors of a *similarity IFS*, in which the maps are of the form $x \mapsto r \cdot O(x) + y$, for $x, y \in \mathbb{R}^d$, $r \in (0, 1)$, and O an orthogonal linear transformation.

In §5 we analyze the case in which $d = 1$ and assumption (1.2) does not hold. In this case we have:

Theorem 8. *Assume that $t_i - t_j \in \mathbb{Q}$ for every $i, j \in \Lambda$, and that t_i is normal to base D for some $i \in \Lambda$. Then for any Bernoulli measure μ, a.e. $x \in K$ is normal to base D.*

For example, if $\mathcal{C}$ is the Cantor middle-thirds set, and α is normal to base 3, then so is a.e. $x \in \mathcal{C} + \alpha$.

1.3. **Acknowledgments.** We gratefully acknowledge the support of BSF grant 2016256 and ISF grant 2095/15.

2. Preliminaries

In this section we collect some preliminary results we will need.

2.1. **Limit measures.** A key ingredient in the proof of Theorem 1 is the following result ([12], see also [4, Lemmas 1.17, 1.19, 1.21]) which applies in the setting of random walks as described at the beginning of section 1.

Proposition 9 (Furstenberg)**.** *Let ν be a μ-stationary probability measure on $\mathbb{T}^d$, then the following hold:*

(1) The limit measures $\nu_b = \lim_{n\to\infty} (b_1 \circ \cdots \circ b_n)_* \nu$ exist for β-a.e. $b \in B$.
(2) $\nu = \int_B \nu_b \, d\beta(b)$.
(3) For all $m \in \mathbb{N}$, for $\beta \times \mu^{*m}$-a.e. $(b, g) \in B \times G^m$, we have $\nu_b = \lim_{n\to\infty} (b_1 \circ \cdots \circ b_n \circ g)_* \nu$.

This is a special case $X = \mathbb{T}^d$ of a much more general result.

2.2. **Commuting expanding matrices.** Recall that a linear transformation (or matrix) is called expanding if all its (complex) eigenvalues have modulus larger than 1. We shall use the following characterization of this property.

Lemma 10. *Let $\mathcal{A}$ be a finite collection of expanding commuting $d \times d$ matrices with entries in $\mathbb{C}$. Then there exists a norm $\|\cdot\|$ on $\mathbb{C}^d$ and some $\rho > 1$ such that for every $A \in \mathcal{A}$ and every $x \in \mathbb{C}^d$, $\|Ax\| \geqslant \rho \|x\|$.*

Proof. Since the matrices commute, there is a basis of $\mathbb{C}^d$ with respect to which they can all be put in an upper triangular form. Thus we may assume that all the matrices are in fact upper triangular complex matrices. Denote by λ the smallest modulus of an eigenvalue of all the matrices in $\mathcal{A}$, and denote by a the largest modulus of all entries of the matrices. Let $m \in \mathbb{R}$ be any number satisfying

$$m > \frac{da}{\lambda - 1},$$

and define a norm by

$$\left\| \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \right\| \stackrel{\text{def}}{=} \max_i \left\{ m^{i-1} |x_i| \right\}.$$

A straightforward computation, which we leave to the reader, shows that for any $x \in \mathbb{C}^d \backslash \{0\}$ and any $A \in \mathcal{A}$ we have $\|Ax\| > \|x\|$. Since the unit sphere in $\mathbb{C}^d$ is compact and $\mathcal{A}$ is finite, the minimum

$$\rho \stackrel{\text{def}}{=} \min \left\{ \frac{\|Ax\|}{\|x\|} : A \in \mathcal{A},\ x \in \mathbb{C}^d,\ \|x\| = 1 \right\}$$

exists, is greater than one, and has the required property. □

2.3. **Invariance of the set of accumulation points of random trajectories.** The following proposition is a general observation about random walks on a second countable space, generated by finitely many continuous maps. It states that for almost every trajectory, the set of accumulation points along the trajectory is invariant under each one of the functions.

Proposition 11. *Let X be a second countable space and $f_j : X \to X,\ j = 1, \ldots, k$ be continuous. Let $\mu \stackrel{\text{def}}{=} \sum_j p_j \delta_j$, where $p_j > 0$ for all j and $\sum p_j = 1$, and let $\mathbb{P} \stackrel{\text{def}}{=} \mu^{\otimes \mathbb{N}}$. Given $x_0 \in X$ and $\underline{i} = (i_1, i_2, \cdots) \in \{1, \cdots, k\}^{\mathbb{N}}$, we set $x_n(\underline{i}) \stackrel{\text{def}}{=} f_{i_n} \circ \cdots \circ f_{i_1}(x_0)$, and denote the set of all accumulation points of $\{x_n(\underline{i})\}_{n=1}^{\infty}$ by $L(\underline{i})$. Then for $\mathbb{P}$-a.e. $\underline{i} \in \{1, \ldots, k\}^{\mathbb{N}}$ we have $f_j(L(\underline{i})) \subseteq L(\underline{i})$ for $j = 1, \ldots, k$.*

For the proof of Proposition 11 we will need the following three lemmas. We retain the same notation as in the Proposition. We think of $\{1, \ldots, k\}^{\mathbb{N}}$ as a probability space and use probabilistic notation when discussing the $\mathbb{P}$-measure of subsets of this space.

Lemma 12. *For $\underline{i} \in \{1, \ldots, k\}^{\mathbb{N}}$, $N \in \mathbb{N}$ and $\varnothing \neq U \subseteq X$, let $K_U^N = K_U^N(\underline{i}) \stackrel{\text{def}}{=} \{n > N : x_n(\underline{i}) \in U\}$. Then*

$$\mathbb{P}\left(\left|K_U^N\right| = \infty \text{ and } \forall n \in K_U^N, i_{n+1} \neq 1\right) = 0. \tag{2.1}$$

Proof. We will write the elements of K_U^N in increasing order, as $N < n_1 < n_2 < \cdots$. For any $j \in \mathbb{N}$, let M_j denote the following event:

$$\left|K_U^N\right| \geqslant j \text{ and } i_{n+1} \neq 1,\ \forall n \in \{n_1, \ldots, n_j\}.$$

That is, M_j is the set of $\underline{i}$ for which the sequence $x_n(\underline{i})$ visits U at least j times, and the next element of the sequence $\underline{i}$ following each of the first j visits to U, is not equal to 1. Observe that $\mathbb{P}(M_{j+1}) = \mathbb{P}(M_{j+1}|M_j)\mathbb{P}(M_j)$ because $M_{j+1} \subseteq M_j$ for any j. We have

$$\begin{aligned}
\mathbb{P}\left(M_{j+1} \middle| M_j\right) =& \mathbb{P}\left(\left|K_U^N\right| \geqslant j+1 \text{ and } \forall n \in \{n_1, \ldots, n_{j+1}\}, i_{n+1} \neq 1 \middle| M_j\right) \\
=& \mathbb{P}\left(\forall n \in \{n_1, \ldots, n_{j+1}\}, i_{n+1} \neq 1 \middle| \left|K_U^N\right| \geqslant j+1 \text{ and } i_{n+1} \neq 1,\ \forall n \in \{n_1, \ldots, n_j\}\right) \\
\times& \mathbb{P}\left(\left|K_U^N\right| \geqslant j+1 \middle| M_j\right).
\end{aligned}$$

Since the entry $i_{n_{j+1}+1}$ does not depend on the previous entries,

$$\begin{aligned}
&\mathbb{P}\left(i_{n+1} \neq 1,\ \forall n \in \{n_1, \ldots, n_{j+1}\} \middle| \left|K_U^N\right| \geqslant j+1 \text{ and } i_{n+1} \neq 1,\ \forall n \in \{n_1, \ldots, n_j\}\right) \\
\leqslant& \mathbb{P}\left(i_{n_{j+1}+1} \neq 1\right) = 1 - p_1.
\end{aligned}$$

Hence, by induction, one has $\mathbb{P}(M_j) \leqslant (1 - p_1)^j$, for all $j \in \mathbb{N}$. The conclusion of the lemma is now immediate. □

Denote by $\mathscr{B}$ a countable base of the topology of X. It follows at once that

Lemma 13. $\mathbb{P}\left(\exists B \in \mathscr{B}, \exists N \in \mathbb{N} \text{ s.t. } \left|K_B^N\right| = \infty \text{ and } \forall n \in K_B^N, i_{n+1} \neq 1\right) = 0.$

Lemma 14. *Fix $\underline{i} \in \{1, \ldots, k\}^{\mathbb{N}}$ and $a \in L(\underline{i})$. If $f_1(a) \notin L(\underline{i})$ then there exist $B \in \mathscr{B}$ containing a, and $N \in \mathbb{N}$ such that $\left|K_B^N\right| = \infty$ and $\forall n \in K_B^N, i_{n+1} \neq 1$.*

Proof. Assume the contrary. That leads to a subsequence $\{x_{n_k}(\underline{i})\}_{k=1}^{\infty}$ converging to a such that $i_{n_k+1} = 1$, for all $k \in \mathbb{N}$. This shows that $f_1(x_{n_k}) = x_{n_k+1}$ for any k and hence, from the continuity of f_1, one obtains $x_{n_k+1} \to f(a)$ as $k \to \infty$. Thus $f_1(a) \in L(\underline{i})$ which contradicts our assumption. □

Proof of Proposition 11. Combining Lemmas 13 and 14 we obtain that $f_1(L(\underline{i})) \subseteq L(\underline{i})$ for $\mathbb{P}$-almost all $\underline{i} \in \{1, \ldots, k\}^{\mathbb{N}}$. By similar arguments, one can prove the same for any f_j, $j = 1, \ldots, k$. □

3. Random walks on tori

In this section we asume the notations and assumptions of Theorem 1: we are given n commuting expanding $d \times d$ integer matrices $D_1, ..., D_n$, real numbers $\alpha_1, \ldots, \alpha_n$, and define the affine endomorphisms $h_1, \ldots, h_n$ by $h_i(x) = D_i(x) + \alpha_i$. We are also given a probability measure μ whose support is the finite set $\{h_1, ..., h_n\}$. We set $\beta = \mu^{\otimes \mathbb{N}}$, and assume that

$$\{(\mathbb{I}_d - D_i)\,\alpha_j - (\mathbb{I}_d - D_j)\,\alpha_i : 1 \leqslant i, j \leqslant n\} \text{ is not contained in a proper closed subgroup of } \mathbb{T}^d. \tag{3.1}$$

We need the following Lemma.

Lemma 15. *If a finite set $\{\mathbf{z}_1, \mathbf{z}_2, \ldots, \mathbf{z}_\ell\} \subseteq \mathbb{T}^d$ is not contained in a proper closed subgroup of $\mathbb{T}^d$, then for β-a.e. $\underline{i} = (i_1, i_2, \ldots) \in \{1, \ldots, n\}^{\mathbb{N}}$ the set of all accumulation points of*

$$\mathscr{S}(\underline{i}) \overset{\text{def}}{=} \{D_{i_m} \circ \cdots \circ D_{i_1}(\mathbf{z}_j) : m \in \mathbb{N}, 1 \leqslant j \leqslant \ell\}$$

is not contained in a proper closed subgroup of $\mathbb{T}^d$.

Proof. Denote the set of $\underline{i} = (i_1, i_2, \cdots) \in \{1, \ldots, n\}^{\mathbb{N}}$ for which the conclusion fails by Λ_0, and assume by contradiction that $\beta(\Lambda_0) > 0$. For $\underline{i} \in \Lambda_0$, let us denote the closed subgroup generated by all the accumulation points of $\mathscr{S}(\underline{i})$ by $K(\underline{i})$. Since a torus contains countably many closed subgroups K, we can pass to a subset of Λ_0 (which we continue to denote by Λ_0) such that $K = K(\underline{i})$ is the same for all $\underline{i}$. Let $p : \mathbb{R}^d \to \mathbb{T}^d$ be the natural projection map. Then $p^{-1}(K)$ is a closed subgroup of $\mathbb{R}^d$. Let S be the connected component of the identity in $p^{-1}(K)$, i.e. the largest subspace contained in $p^{-1}(K)$. Then S is a rational subspace of $\mathbb{R}^d$, $\dim S < d$, and $p^{-1}(K)/S$ is discrete in $\mathbb{R}^d/S$. Consider the following commutative diagram, where the vertical maps are the canonical projection maps and $\bar{p}$ is the map induced by p:

$$\begin{array}{ccc} \mathbb{R}^d & \xrightarrow{\ p\ } & \mathbb{T}^d \\ \downarrow & & \downarrow \\ \mathbb{R}^d/p^{-1}(K) & \xrightarrow{\ \bar{p}\ } & \mathbb{T}^d/K \end{array}$$

The diagram shows that $(\mathbb{R}^d/S)/(p^{-1}(K)/S) \cong \mathbb{R}^d/p^{-1}(K)$ is isomorphic to $\mathbb{T}^d/K$. Hence $\mathbb{T}^d/K$ is a torus of dimension $d - \dim S$, $\mathbb{R}^d/S$ is the universal cover of $\mathbb{T}^d/K$, and the covering map is given by

$$\pi : \mathbb{R}^d/S \longrightarrow \mathbb{T}^d/K, \quad \pi(\mathbf{x} + S) \overset{\text{def}}{=} p(\mathbf{x}) + K.$$

Denote by $\mathscr{S}'(\underline{i})$ the set of accumulation points of $\mathscr{S}(\underline{i})$, and write $\mathscr{S}'(\underline{i}) = \bigcup_{j=1}^{k} \mathscr{S}'_j(\underline{i})$, where $\mathscr{S}'_j(\underline{i})$ is the set of accumulation points of $\{D_{i_m} \cdots D_{i_1}\mathbf{z}_j \in \mathbb{T}^d : m \in \mathbb{N}\}$, for fixed $j \in \{1, \ldots, \ell\}$. It follows from Proposition 11 that for all j, there is a subset of Λ_0 of full measure, of $\underline{i}$ for which $D_1(\mathscr{S}'_j(\underline{i})), \ldots, D_n(\mathscr{S}'_j(\underline{i}))$ are all contained in $\mathscr{S}'_j(\underline{i})$. We replace Λ_0 with its subset of full measure for which this holds for all j (and continue to denote this set by Λ_0). For each $\underline{i} \in \Lambda_0$ and each $r \in \{1, \ldots, n\}$, we have $D_r(K) \subseteq K$. Since S is the largest subspace contained in $p^{-1}(K)$, this implies $D_r(S) \subseteq S$. Being an expanding map, D_r is invertible and hence it preserves dimensions of linear spaces, and so we must have $D_r(S) = S$ (recall that we denote by D_r both a toral endomorphism and the corresponding linear transformation). Therefore we may view D_r as inducing a map on both $\mathbb{T}^d/K$ and $\mathbb{R}^d/S$, which we continue to denote by D_r, and we have the following commutative diagram:

$$\begin{array}{ccc} \mathbb{R}^d/S & \xrightarrow{D_r} & \mathbb{R}^d/S \\ \downarrow{\scriptstyle \pi} & & \downarrow{\scriptstyle \pi} \\ \mathbb{T}^d/K & \xrightarrow{D_r} & \mathbb{T}^d/K \end{array}$$

We have that both horizontal maps in this diagram are surjective endomorphisms, D_r commutes with the natural projection map from $\mathbb{T}^d \longrightarrow \mathbb{T}^d/K$, and the projection of $\mathscr{S}(\underline{i})$ on $\mathbb{T}^d/K$ has $\overline{\mathbf{0}} = K$ as the only accumulation point for all $\underline{i} \in \Lambda_0$.

Fix $\underline{i} = (i_1, i_2 \cdots) \in \Lambda_0$. We first observe that the projections $\overline{\mathbf{z}}_j$ of $\mathbf{z}_j$ in $\mathbb{T}^d/K$ satisfy

$$D_{i_m} \ldots D_{i_1}\overline{\mathbf{z}_j} \neq \overline{\mathbf{0}}, \quad \text{for some } j \in \{1, \ldots, \ell\} \text{ and all } m \in \mathbb{N}.$$

Otherwise, there would exist $M \in \mathbb{N}$ for which $D_{i_M} \circ \cdots \circ D_{i_1}\overline{\mathbf{z}_j} = \overline{\mathbf{0}}$, for all $j \in \{1, 2, \ldots, \ell\}$. We then consider the closed subgroup in $\mathbb{T}^d/K$ generated by $\{\overline{\mathbf{z}_i} : 1 \leqslant i \leqslant \ell\}$. It is obvious that this subgroup is contained in the kernel of the surjective endomorphism $D_{i_M} \cdots D_{i_1} : \mathbb{T}^d \longrightarrow \mathbb{T}^d$. Since the kernel is finite, so is the subgroup we considered and hence proper in $\mathbb{T}^d/K$. Pulling it back to $\mathbb{T}^d$, one obtains a proper closed subgroup that contains both K and all $\overline{\mathbf{z}_j}$'s. This contradicts our hypothesis. Since $\overline{\mathbf{0}}$ is the only accumulation point of the sequence $\{D_{i_m} \cdots D_{i_1}\overline{\mathbf{z}_i}\}_{m=0}^{\infty}$, it follows from compactness that $\lim_{m\to\infty} D_{i_m} \cdots D_{i_1}\overline{\mathbf{z}_j} = \overline{\mathbf{0}}$ in $\mathbb{T}^d/K$, for each j.

Our next observation is as follows. If we choose a basis of S and then extend it to a basis $\mathcal{B}$ of $\mathbb{R}^d$, then for each r, the matrix representation of D_r, with respect to this basis, is of the form

$$\begin{pmatrix} D_S^{(r)} & * \\ 0 & D^{(r)} \end{pmatrix}.$$

By hypothesis, all eigenvalues of $D_S^{(r)}$ and $D^{(r)}$ have absolute value > 1. The matrix representing D_r with respect to the projection of $\mathcal{B}$ to $\mathbb{R}^d/S$ is $D^{(r)}$. Using Lemma 10, there is a norm $||\cdot||$ on $\mathbb{R}^d/S$ such that

$$\rho \stackrel{\text{def}}{=} \inf_{1\leqslant r\leqslant n,\, \mathbf{v}\in\mathbb{R}^d/S,\, ||\mathbf{v}||=1} ||D_r\mathbf{v}|| > 1.$$

We let $||D_r||_{\text{op}}$ denote the operator norm of the linear map D_r, in $\mathbb{R}^d/S$ with respect to the norm on $\mathbb{R}^d/S$ chosen above, and let $R \stackrel{\text{def}}{=} \max_{1\leqslant r\leqslant k} ||D_r||_{\text{op}}$.

As π is a covering homomorphism, we can take a small enough open ball $\mathfrak{B}$ in $\mathbb{R}^d/S$ centered at $\mathbf{0}$ which evenly covers $\pi(\mathfrak{B})$, i.e. $\pi|_{\mathfrak{B}}$ is a homeomorphism onto its image. Since $\lim_{m\to\infty} D_{i_m}\cdots D_{i_1}\overline{\mathbf{z}_j} = \overline{\mathbf{0}}$, there is $N \in \mathbb{N}$ such that $D_{i_m}\cdots D_{i_1}\overline{\mathbf{z}_j}$ lies in $\pi\left(\frac{1}{R}\mathfrak{B}\right)$ for all $m \geqslant N$. We denote the lift of $D_{i_m}\cdots D_{i_1}\overline{\mathbf{z}_j}$ in $\frac{1}{R}\mathfrak{B}$ by $\mathbf{t}_m$, for all $m \geqslant N$. Note that $\mathbf{t}_m \neq \overline{\mathbf{0}}$ for any $m \geqslant N$. Since each application of D_r increases the norm of a nonzero vector by at least ρ and at most R, there exists $s \in \mathbb{N}$ such that $D_{i_{m+s}}\cdots D_{i_{m+1}}\mathbf{t}_N \in \mathfrak{B}\backslash\frac{1}{R}\mathfrak{B}$. Since $\pi|_{\mathfrak{B}}$ evenly covers its image, we see that $\pi(D_{i_{N+s}}\cdots D_{i_{N+1}}\mathbf{t}_N) \in \pi(\mathfrak{B})\backslash\pi\left(\frac{1}{R}\mathfrak{B}\right)$, i.e., $D_{i_{N+s}}\cdots D_{i_{N+1}}D_{i_N}\cdots D_{i_1}\overline{\mathbf{z}_j} \notin \pi\left(\frac{1}{R}\mathfrak{B}\right)$. This contradicts our choice of N. □

We are now ready for the

Proof of Theorem 1. A straightforward induction shows that for any finite sequence $j_1, j_2, \ldots, j_m \in \{1, 2, \ldots, n\}$,

$$h_{j_1} \circ \cdots \circ h_{j_m}(\mathbf{x}) = D_{j_1} \circ \cdots \circ D_{j_m}(\mathbf{x}) + \sum_{s=1}^{m} D_{j_1} \circ \cdots \circ D_{j_{s-1}}(\alpha_{j_s}) \tag{3.2}$$

(with + denoting addition in $\mathbb{T}^d$). From this it is clear that, for any finite sequence $j_1, \ldots, j_m$ and any pair of indices $\ell, s \in \{1, 2, \ldots, n\}$, one has

$$\begin{aligned} &h_{j_1} \circ \cdots \circ h_{j_m} \circ h_\ell \circ h_s(\mathbf{x}) \\ &= h_{j_1} \circ \cdots \circ h_{j_m} \circ h_s \circ h_\ell(\mathbf{x}) + D_{j_1} \circ \cdots \circ D_{j_m}((\mathbb{I}_d - D_\ell)(\alpha_s) - (\mathbb{I}_d - D_s)(\alpha_\ell)). \end{aligned} \tag{3.3}$$

For a given vector $\mathbf{a} \in \mathbb{T}^d$, let

$$R_{\mathbf{a}} : \mathbb{T}^d \to \mathbb{T}^d, \quad R_{\mathbf{a}}(x) \stackrel{\text{def}}{=} x + \mathbf{a}.$$

We also denote, for $m \in \mathbb{N}$ and $\underline{j} \in \{1, \ldots, n\}^{\mathbb{N}}$,

$$\mathbf{a}_{m,\underline{j}}^{\ell,s} \stackrel{\text{def}}{=} D_{j_1} \circ \cdots \circ D_{j_m}((\mathbb{I}_d - D_\ell)(\alpha_s) - (\mathbb{I}_d - D_s)(\alpha_\ell)).$$

With these notations we can rewrite (3.3) as

$$h_{j_1} \circ \cdots \circ h_{j_m} \circ h_\ell \circ h_s(\mathbf{x}) = R_{\mathbf{a}_{m,\underline{j}}^{\ell,s}} \circ h_{j_1} \circ \cdots \circ h_{j_m} \circ h_s \circ h_\ell(\mathbf{x}). \tag{3.4}$$

Suppose now that ν is a μ-stationary measure on $\mathbb{T}^d$. From (3.4) and Proposition 9(3) we obtain that there is a subset $B_0 \subset B$, with $\beta(B_0) = 1$, such that for any $b \in B_0$ and any $\ell, s \in \{1, 2, \ldots, n\}$,

$$\nu_b = \lim_{m\to\infty} (b_1 \circ \cdots \circ b_m \circ h_\ell \circ h_s)_*\nu = \lim_{m\to\infty} \left(R_{\mathbf{a}_{m,\underline{j}}^{\ell,s}}\right)_* (b_1 \circ \cdots \circ b_m \circ h_s \circ h_\ell)_*\nu\,. \tag{3.5}$$

Here we identify B with $\{1, \ldots, n\}^{\mathbb{N}}$ in the obvious way and use the same notation β for the Bernoulli measure on the symbol space $\{1, \ldots, n\}^{\mathbb{N}}$, and identify $\underline{j}$ with b; this should cause no confusion. Let $\mathrm{Prob}(\mathbb{T}^d)$ denote the space of Borel probability measures on $\mathbb{T}^d$, equipped with the weak-* topology. The addition map

$$\mathbb{T}^d \times \mathbb{T}^d \to \mathbb{T}^d, \quad (\mathbf{a}, x) \mapsto R_{\mathbf{a}}(x)$$

is continuous, and thus the induced map

$$\mathbb{T}^d \times \mathrm{Prob}(\mathbb{T}^d) \to \mathrm{Prob}(\mathbb{T}^d), \quad (\mathbf{a}, \theta) \mapsto (R_{\mathbf{a}})_*\theta$$

is also continuous. Thus, by passing to subsequences and using (3.5), we find that for any $b \in B_0$, for any accumulation point $\mathbf{a}'$ of the sequence $\{\mathbf{a}^{\ell,s}_{m,\underline{j}}\}_{m=0}^{\infty}$, the measure ν_b is invariant under $R_{\mathbf{a}'}$.

We now appeal to condition (3.1) and Lemma 15, which ensure that for β-a.e. $\underline{j} \in \{1,\dots,n\}^{\mathbb{N}}$, the accumulation points of the set

$$\mathbf{A}_{\underline{j}} \overset{\text{def}}{=} \left\{\mathbf{a}^{\ell,s}_{m,\underline{j}} : \ell, s \in \{1,...,n\}, \, m \in \mathbb{N}\right\}$$

generate a dense subgroup of $\mathbb{T}^d$. Upon possibly replacing B_0 with its subset, we still have $\beta(B_0) = 1$, and for all $\underline{j} \in B_0$ we also have that the accumulation points of $\mathbf{A}_{\underline{j}}$ generate a dense subgroup of $\mathbb{T}^d$. We obtain that for $b \in B_0$, ν_b is invariant under a dense subgroup of $\mathbb{T}^d$, and since the stabilizer of a measure is a closed subgroup, that ν_b is the Haar measure on $\mathbb{T}^d$. The conclusion of Theorem 1 now follows from Proposition 9(2). ☐

3.1. **Further remarks.** We first note that in the one-dimensional case, a stronger version of Lemma 15 is true: its conclusion holds for *every* sequence $\underline{i} = (i_1, i_2, \cdots) \in \{1, \cdots, k\}^{\mathbb{N}}$.

Lemma 16. *Let $\alpha \in \mathbb{T}\backslash\mathbb{Q}$, and let $D_1, ..., D_n \geqslant 2$ be integers, which we think of also as toral endomorphisms $D_i : \mathbb{T} \to \mathbb{T}$. Given a sequence $i_1, i_2, ... \in \{1,...,n\}^{\mathbb{N}}$, denote $x_k \overset{\text{def}}{=} D_{i_1} \circ \cdots \circ D_{i_k}(\alpha)$. Then the set of all accumulation points of the sequence $(x_k)_{k\in\mathbb{N}}$ is infinite.*

We will not be using this result, and we leave the proof to the reader.

The following shows that condition (3.1) cannot be relaxed, at least in case $d = 1$:

Proposition 17. *Assume that for some $i_0 \in \Lambda$, $(1 - D_{i_0})\,\alpha_j - (1 - D_j)\,\alpha_{i_0} \in \mathbb{Q}$ for every $j \in \Lambda$. Then there exists a finitely supported μ-stationary measure on $\mathbb{T}$.*

Proof. Without loss of generality assume that $i_0 = 1$. Denote

$$\beta_j = \alpha_j - \frac{D_j - 1}{D_1 - 1}\alpha_1.$$

By assumption $\beta_j \in \mathbb{Q}$ for every $j \in \Lambda$. Let $q \in \mathbb{N}$ be a common denominator for all the β_i, and denote $A = \left\{0, \frac{1}{q}, ..., \frac{q-1}{q}\right\}$. Denote also $x_0 = -\dfrac{\alpha_1}{D_1 - 1}$, so that $h_1(x_0) = x_0$.

We now claim that for all $i \in \Lambda$, $h_i(A + x_0) \subseteq A + x_0$. Indeed, for all $a \in A$ and all $i \in \Lambda$,

$$\begin{aligned} h_i(a + x_0) &= D_i(a) + \alpha_i + D_i(x_0) \\ &= D_i(a) + \alpha_i + (D_i - 1)(x_0) + x_0 \\ &= D_i(a) + \alpha_i - \frac{D_i - 1}{D_1 - 1}\alpha_1 + x_0 \\ &= D_i(a) + \beta_i + x_0 \in A + x_0 \end{aligned}$$

Hence $A + x_0$ supports a μ-stationary measure. ☐

Our method of proof also gives another proof of the following well-known fact:

Proposition 18. *Suppose $D_i = 1$ for every $i \in \Lambda$, that is, each h_i is a translation by α_i. Then Haar measure is the unique μ-stationary probability measure on $\mathbb{T}$ if and only if α_i is irrational for some $i \in \Lambda$.*

Proof. Assume first (without loss of generality) that $\alpha_1 \notin \mathbb{Q}$. Since the functions h_i are now only rotations, they commute with each other. Hence, if ν is some μ-stationary measure, by Theorem 9, for β-a.e. $b \in B$,

$$\nu_b = \lim_{k\to\infty} (b_1 \circ \cdots \circ b_k \circ h_1)_* \nu = \lim_{k\to\infty} (h_1)_* (b_1 \circ \cdots \circ b_k)_* \nu = (h_1)_* \nu_b \ .$$

Since h_1 is an irrational rotation, ν_b has to be Haar measure and hence ν is Haar measure.
The other implication is trivial. □

We remark that Proposition 18 could also be obtained as a corollary of a result of Choquet and Deny (see [4, §1.5] for the statement and a similar argument).

4. Normal numbers in fractals

Using an idea of [19], we will show how to derive information on normal numbers in self-similar sets, from random walks on tori. We will then use the results of §2 to prove Theorem 4. In order to make the idea more transparent, we first prove a special case of Theorem 4, namely we assume $r_1 = \cdots = r_k = 1$, or in other words,

$$\text{there is an integer } D \geqslant 2 \text{ such that } f_i(x) = \frac{1}{D} \cdot x + t_i, \quad i = 1, \dots, k. \tag{4.1}$$

In this case the irrationality assumption (1.2) simplifies to

$$\{t_i - t_j : 1 \leqslant i, j \leqslant k\} \text{ is not contained in any proper closed subgroup of } \mathbb{T}^d. \tag{4.2}$$

We will need the following result[2]:

Proposition 19 ([19], Prop. 5.1)**.** *Let Γ be a semigroup acting on a space X, let μ be a probability measure on Γ, and denote the infinite product measure $\mu^{\otimes\mathbb{N}}$ by β. Given any $x_0 \in X$, assume that for β-a.e. $b \in B$, the random path $(b_n \cdots b_1 x_0)_{n\in\mathbb{N}}$ is equidistributed w.r.t. a measure ν on X. Then for β-a.e. $b \in B$, the sequence*

$$(b_n \cdots b_1 x_0,\, T^n b)_{n\in\mathbb{N}}$$

is equidistributed w.r.t. the product measure $\nu \otimes \beta$ on $X \times B$.

Proof of Theorem 4 assuming (4.1). Let μ be a Bernoulli measure on K given by $\mu = \pi_* \sigma$, where σ is a Bernoulli measure on the symbolic space $\Lambda^{\mathbb{N}}$ and $\pi : \Lambda^{\mathbb{N}} \to K$ is the coding map.
By a routine induction, we find that for any $n, m \in \mathbb{N}$ with $n < m$,

$$f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_m}(0) = \frac{f_{i_2} \circ \cdots \circ f_{i_m}(0)}{D} + t_{i_1} = \cdots$$
$$\cdots = \frac{f_{i_{n+1}} \circ \cdots \circ f_{i_m}(0)}{D^n} + \frac{t_{i_n}}{D^{n-1}} + \frac{t_{i_{n-1}}}{D^{n-2}} + \cdots + t_{i_1}.$$

Multiplying by D^n and taking the limit as $m \to \infty$ we obtain that for any $x \in K$ and $n \in \mathbb{N}$,

$$D^n(x) = \sum_{j=1}^{n} D^{n-(j-1)}(t_{i_j}) + \pi(T^n(\underline{i})), \tag{4.3}$$

where $x = \pi(\underline{i})$ for some $\underline{i} = i_1, i_2, ... \in \Lambda^{\mathbb{N}}$.

[2]We use here and further below results of [19, §5], where Γ is taken to be a group rather than a semigroup. However the more general case of semigroups follows from the same proof.

Define, for each $s \in \Lambda$, $h_s : \mathbb{T}^d \to \mathbb{T}^d$ by $h_s(x) \stackrel{\text{def}}{=} D(x + t_s)$. Note that

$$h_{i_n} \circ \cdots \circ h_{i_1}(0) = \sum_{j=1}^{n} D^{n-(j-1)}(t_{i_j}). \tag{4.4}$$

Note that $(h_{i_n} \circ \cdots \circ h_{i_1}(0))_{n \in \mathbb{N}}$ is in fact a random walk trajectory governed by the probability measure σ on $\Lambda^{\mathbb{N}}$, and that condition (1.2) in this case implies condition (3.1) of Theorem 1. Applying Corollary 2, we get that for σ-a.e. $\underline{i} \in \Lambda^{\mathbb{N}}$, the sequence (4.4) equidistributes in $\mathbb{T}^d$ w.r.t. to Haar. Next, we apply Proposition 19, with $X = \mathbb{T}^d$ and $\nu =$ Haar, and obtain the equidistribution of the sequence $(h_{i_n} \circ \cdots \circ h_{i_1}(0), T^n(\underline{i}))_{n=1}^{\infty}$ w.r.t. the product measure Haar $\otimes \sigma$ on $\mathbb{T}^d \times B$, for σ-a.e. $\underline{i} \in \Lambda^{\mathbb{N}}$. Since the coding map π is continuous, this implies that for σ-a.e. $\underline{i} \in \Lambda^{\mathbb{N}}$, the joint sequence

$$(h_{i_n} \circ \cdots \circ h_{i_1}(0), \pi(T^n(\underline{i})))_{n=1}^{\infty} \tag{4.5}$$

is equidistributed in $\mathbb{T}^d \times K$ with respect to the product measure Haar $\otimes \mu$.

Consider the addition map

$$F : \mathbb{T}^d \times K \longrightarrow \mathbb{T}^d, \quad F(x, y) = x + y.$$

It follows easily from the fact that Haar measure is invariant under addition in $\mathbb{T}^d$, that

$$F_*(\lambda \times \mu) = \text{Haar}. \tag{4.6}$$

The equidistribution of $(D^n x)_{n=1}^{\infty}$ for μ-almost every $x \in K$ follows immediately from (4.6), from (4.3), (4.4) and from the equidistribution of (4.5). ☐

For the general case of Theorem 4, we will need an extension of Proposition 19. We let $C_r \stackrel{\text{def}}{=} \mathbb{Z}/r\mathbb{Z}$ denote the cyclic group of order r, and let θ denote the uniform measure on C_r. Recall that we have a random walk on $\mathbb{T}^d$ driven by a finitely supported measure μ on a semigroup Γ of affine toral endomorphisms. We write $\bar{B} \stackrel{\text{def}}{=} (\text{supp}\mu)^{\otimes\mathbb{Z}}$ and $\bar{\beta} \stackrel{\text{def}}{=} \mu^{\otimes\mathbb{Z}}$. Note this is the two-sided shift space. We continue to denote by T the shift map (but this time on $\bar{B}$). We say that a map $\kappa : \Gamma \to C_r$ is a *morphism* if it satisfies

$$\forall \gamma_1, \gamma_2 \in \Gamma, \; \kappa(\gamma_1\gamma_2) = \kappa(\gamma_1) + \kappa(\gamma_2).$$

Proposition 20. *With the notation above, let $r \in \mathbb{N}$ and let $\kappa : \Gamma \to C_r$ be a surjective morphism. Given any $x_0 \in \mathbb{T}^d$, assume that for $\bar{\beta}$-a.e. $b \in \bar{B}$, the random path $(b_n \cdots b_1 x_0)_{n \in \mathbb{N}}$ is equidistributed w.r.t. a measure ν on $\mathbb{T}^d$. Then for $\bar{\beta}$-a.e. $b \in \bar{B}$, the sequence*

$$(b_n \cdots b_1 x_0, \, \kappa(b_n \cdots b_1), \, T^n b)_{n \in \mathbb{N}}$$

is equidistributed w.r.t. the product measure $\nu \otimes \theta \otimes \bar{\beta}$ on $\mathbb{T}^d \times C_r \times \bar{B}$.

Proof. This is not stated explicitly in [19] but is proved along the same lines as [19, Thm. 2.2]. For completeness we sketch the proof.

An extension of Proposition 19 to the bi-infinite sequence space $\bar{B}$ is given in [19, Prop. 5.2]. Putting this result to use, it is enough to show that for any $x_0 \in \mathbb{T}^d$, for $\bar{\beta}$-a.e. b, the sequence $(b_n \cdots b_1 x_0, \kappa(b_n \cdots b_1))$ is equidistributed, with respect to $\nu \otimes \theta$ on $\mathbb{T}^d \times C_r$. Indeed, once this is established we can apply [19, Prop. 5.2] with $X = \mathbb{T}^d \times C_r$.

Now note that the action of Γ on $\mathbb{T}^d \times C_r$ by $\gamma(x, s) \mapsto (\gamma x, s + \kappa(\gamma))$ is ergodic. This follows from the fact that each individual element of γ is an expansive toral endomorphism, and hence is mixing. Thus the required equidistribution statement follows from [19, Cor. 5.5]. ☐

Proof of Theorem 4, general case. We have a self-similar set K which is the attractor of the maps $f_i(x) = \frac{1}{D^{r_i}} \cdot x + t_i$ ($i = 1, \ldots, k$), for an integer $D \geqslant 2$. We can assume with no loss of generality that

$$\gcd(r_1, \ldots, r_k) = 1; \tag{4.7}$$

indeed, if this does not hold, we can replace D with a power of D and (using [8, Thm. 4.4]) reduce to the situation in which (4.7) holds. Write

$$D_i \stackrel{\text{def}}{=} D^{r_i}, \quad r \stackrel{\text{def}}{=} \max r_i, \quad \text{and } \bar{D} \stackrel{\text{def}}{=} D^r.$$

Again by [8, Thm. 4.4], it suffices to show that a.e. $x \in K$ is normal to base $\bar{D}$. Computing as in (4.3) and (4.4) we see that for any $b = (i_1, \ldots) \in B$ and $n \in \mathbb{N}$,

$$D_{i_1} \circ \cdots \circ D_{i_n} \pi(b) = \pi(T^n b) + h_{i_n} \circ \cdots \circ h_{i_1}(0). \tag{4.8}$$

For each n and b, let $\ell = \ell_{b,n} \in \mathbb{N}$ satisfy $\bar{D}^{\ell-1} < D_{i_1} \cdots D_{i_n} \leqslant \bar{D}^{\ell}$, and let $s = s_{b,n}$ so that

$$\bar{D}^{\ell} = D_{i_1} \cdots D_{i_n} D^s. \tag{4.9}$$

That is,

$$\ell \stackrel{\text{def}}{=} \left\lceil \frac{r_{i_1} + \cdots + r_{i_n}}{r} \right\rceil \quad \text{and} \quad s \stackrel{\text{def}}{=} r\ell - (r_{i_1} + \cdots + r_{i_n}) \in \{0, \ldots, r-1\}. \tag{4.10}$$

If we consider s as an element of C_r (i.e. consider its class modulo r) then we see that $s = \kappa(b_n \cdots b_1)$ where κ is the morphism mapping i_j to $-r_j$, considered as an element of C_r. The morphism κ is surjective in view of (4.7).

By Corollary 2 and Proposition 20, the sequence

$$(h_{i_n} \circ \cdots \circ h_{i_1}(0),\ s_{b,n},\ T^n b) \in \mathbb{T}^d \times C_r \times \bar{B}$$

is equidistributed with respect to Haar $\otimes\, \theta \otimes \bar{\beta}$ for $\bar{\beta}$-a.e. b.

For $b \in \bar{B}$ we continue to use the notation π to denote the map $\pi(b) = \pi(i_1, i_2, \ldots)$ where $b = (i_{-1}, i_0, i_1, \ldots)$. Let $x_0 = \pi(b)$, where b belongs to the subset of $\bar{B}$ of full measure for which this equidistribution result holds, we have from (4.8) and (4.9) that

$$\bar{D}^{\ell_{b,n}}(x_0) = D^{s_{b,n}} \left(\pi(T^n b) + h_{i_n} \circ \cdots \circ h_{i_1}(0) \right). \tag{4.11}$$

We consider this sequence as a sequence depending on the index n, and note that it is the image of an equidistributed sequence under the continuous map

$$\Psi : \mathbb{T}^d \times C_r \times \bar{B} \to \mathbb{T}^d, \quad \Psi(x, c, b) \stackrel{\text{def}}{=} D^c(\pi(b) + x).$$

Thus it is equidistributed with respect to Haar measure on $\mathbb{T}^d$, as a sequence of the parameter n.

However, our objective is to show equidistribution of the sequence $(\bar{D}^z(x_0))_{z=1}^{\infty}$. Note that given b as above, the sequence $(\ell_{b,n})_{n \in \mathbb{N}}$ is monotonically increasing, but might have repetitions. To handle these repetitions we introduce the following notation: for fixed b, and for each $n \in \mathbb{N}$, let

$$t_b(n) \stackrel{\text{def}}{=} \left| \{ m \in \mathbb{N} : \ell_{b,n} = \ell_{b,m} \} \right|^{-1}.$$

It is clear from (4.10) and the definition of r that $t_b(n) \in \left\{ \frac{1}{r}, \frac{2}{r}, \ldots, 1 \right\}$. It is not hard to see that we can compute $t_b(n)$ from $s_{b,n}$ and from the symbols $(T^n b)_j$ where $|j| \leqslant r$. That is, there is a continuous function $\hat{t}$ on $C_r \times \bar{B}$ such that $t_b(n) = \hat{t}(s_{b,n}, T^n b)$. Now given $f \in C(\mathbb{T}^d)$, we define

$$F : \mathbb{T}^d \times C_r \times \bar{B} \quad \text{by} \quad F(x, c, b) \stackrel{\text{def}}{=} (f \circ \Psi) \cdot \hat{t}(c, b)$$

so that

$$F(x, s_{b,n}, T^n b) = \frac{f\left(D^s(\pi(T^n b) + x)\right)}{|\{m \in \mathbb{N} : \ell_{b,n} = \ell_{b,m}\}|}.$$

This definition ensures that the Birkhoff sums of the two sides of (4.11) satisfy

$$\sum_{z=0}^{L-1} f\left(\bar{D}^z x_0\right) = \sum_{n=0}^{N_L-1} F\left(h_{i_n} \circ \cdots \circ h_{i_1}(0), s_{b,n}, T^n b\right) + O(1),$$

where $N_L = |\{m \in \mathbb{N} : \ell_{b,m} < L\}|$. Thus

$$\frac{1}{N_L}\sum_{z=0}^{L-1} f\left(\bar{D}^z x_0\right) \underset{L\to\infty}{\longrightarrow} \int_{\mathbb{T}^d \times C_r \times \bar{B}} F \, d\mathrm{Haar} \otimes \theta \otimes \bar{\beta}. \tag{4.12}$$

Applying this with the constant function $f \equiv 1$ gives the existence of the limit

$$\lambda \overset{\text{def}}{=} \int \hat{t}\, d\theta \otimes d\bar{\beta} = \lim_{L\to\infty} \frac{L}{N_L}.$$

Dividing both sides of (4.12) by λ and using Fubini to compute the right hand side, we see that

$$\frac{1}{L}\sum_{z=0}^{L-1} f\left(\bar{D}^z x_0\right) \underset{L\to\infty}{\longrightarrow} \int_{\mathbb{T}^d} f \, d\mathrm{Haar},$$

as required. □

5. When all the differences are rational.

In this section we shall analyze the situation in which condition (1.2) in Theorem 4 does not hold. We focus on the one-dimensional case, thus we assume throughout this section that all the differences $t_i - t_j$ are rational. For all $i \in \Lambda$, denote $\delta_i \overset{\text{def}}{=} t_i - t_1 \in \mathbb{Q}$. Note that $\delta_1 = 0$. Given $x \in K$, suppose that $x = \pi\left(\underline{i}\right)$ for $\underline{i} \in \Lambda^{\mathbb{N}}$, i.e.

$$x = \lim_{n\to\infty} f_{i_1} \circ \cdots \circ f_{i_n}(0).$$

Then by equation (4.3), for every $m \in \mathbb{N}$,

$$D^m x = \sum_{j=1}^{m} D^j t_1 + \sum_{j=1}^{m} D^{m-(j-1)} \delta_{i_j} + \pi(T^m(\underline{i})).$$

Denote

$$\alpha_m \overset{\text{def}}{=} \sum_{j=1}^{m} D^j t_1, \quad \eta_m\left(\underline{i}\right) \overset{\text{def}}{=} \sum_{j=1}^{m} D^{m-(j-1)} \delta_{i_j}$$

so that

$$D^m x = \alpha_m + \eta_m\left(\underline{i}\right) + \pi(T^m(\underline{i})).$$

Note that $\eta_m\left(\underline{i}\right)$ stays inside the finite set $\Delta = \left\{0, \frac{1}{q}, \ldots, \frac{q-1}{q}\right\}$, where q is a common denominator for $\delta_2, \delta_3, \ldots, \delta_k$. Also note that α_m is a deterministic sequence (does not depend on $\underline{i}$), and

$$\alpha_m = \sum_{j=1}^{m} D^j t_1 = \frac{D^{m+1} - D}{D-1} t_1 = D^m \frac{D}{D-1}\, t_1 - \frac{D}{D-1}\, t_1.$$

Lemma 21. *$\eta_m(\underline{i})$ is an aperiodic, irreducible Markov process with a finite state space.*

Proof. For each j denote $\tilde{\delta}_j \overset{\text{def}}{=} D \cdot \delta_j \in \Delta$. Then

$$\eta_m(\underline{i}) = \sum_{j=1}^{m} D^{m-j} \tilde{\delta}_{i_j}.$$

This process may be represented as follows:

$$\begin{aligned} &\eta_1 = \tilde{\delta}_{i_1} \\ &\forall m > 1,\, \eta_{m+1} = D \cdot \eta_m + \tilde{\delta}_{i_{m+1}} \end{aligned}$$

Let $\tilde{\Delta} \subseteq \Delta$ be defined as $\tilde{\Delta} = \{a \in \Delta : \exists m \in \mathbb{N}, \mathbb{P}(\eta_m = a) > 0\}$. Since the variables $\left(\tilde{\delta}_{i_j}\right)_{j=1}^{\infty}$ are IID, this is indeed a Markov process on the finite state space $\tilde{\Delta}$.

Since,$\mathbb{P}\left(\eta_{m+q} = 0 \,\middle|\, \eta_m = a\right) > \mathbb{P}\left(\tilde{\delta}_{i_{m+1}} = \cdots = \tilde{\delta}_{i_{m+q}} = 0\right) > 0$ for every $a \in \tilde{\Delta}$, and $\eta_{m+1} \,\middle|\, \eta_m = 0 \sim \eta_1$, the Markov process is irreducible, and it is also aperiodic since $\mathbb{P}\left(\eta_{m+1} = 0 \,\middle|\, \eta_m = 0\right) > 0$. ☐

From Lemma 21 it follows that the process η_m has a unique stationary measure p.

Theorem 22. *Assume that α_m is equidistributed w.r.t. some measure ν on $\mathbb{T}$. Then for μ_K-a.e. $x \in K$, the orbit $(D^m x)_{m=0}^{\infty}$ is equidistributed w.r.t. the measure $\nu * p * \widetilde{\mu}_K$ (where $\widetilde{\mu}_K$ is the projection of μ_K to $\mathbb{T}$).*

In order to prove Theorem 22, we will need the following property of aperiodic, irreducible Markov chains. The proof of the following proposition uses some of the ideas in the proof of Proposition 19, given in [19].

Proposition 23. *Let $x = (x_1, x_2, ...) \in \Omega^{\mathbb{N}}$ be an aperiodic, irreducible Markov chain with a finite state space Ω and a transition matrix P. Let p be the unique stationary measure for the process and let μ be the corresponding measure on $\Omega^{\mathbb{N}}$ w.r.t. p as the starting probability for the process (i.e., $\mu\left(\left\{\omega \in \Omega^{\mathbb{N}} : \omega_1 \in A\right\}\right) = p(A)$ for every $A \subseteq \Omega$). Then for every strictly increasing sequence of positive integers $(n_k)_{k=1}^{\infty}$, for μ-a.e. $x \in \Omega^{\mathbb{N}}$, the sequence $(T^{n_k}(x))_{k=1}^{\infty}$ is equidistributed w.r.t. μ.*

Proof. For the rest of the paper, given a finite sequence $\omega = (\omega_1, ..., \omega_\ell) \in \Omega^\ell$, we denote the corresponding cylinder set by $[\omega] \overset{\text{def}}{=} \left\{\xi \in \Omega^{\mathbb{N}} : (\xi_1, ..., \xi_\ell) = (\omega_1, ..., \omega_\ell)\right\}$.

Let $\mathcal{B}_k$ be the σ-algebra generated by the first n_k coordinates of $\Omega^{\mathbb{N}}$. Given $\omega = (\omega_1, ..., \omega_\ell) \in \Omega^\ell$ for any $\ell \in \mathbb{N}$, define

$$\begin{aligned} \varphi_{k,m} &\overset{\text{def}}{=} \mathbb{E}\left[\mathbf{1}_{[\omega]}(T^{n_k}x) - \mu([\omega]) \,|\, \mathcal{B}_m\right] \\ M_m &\overset{\text{def}}{=} \sum_{k=1}^{\infty} \varphi_{k,m}. \end{aligned}$$

Note the following:

- For $k \leqslant m - \ell$, $\varphi_{k,m} = \mathbf{1}_{[\omega]}(T^{n_k}x) - \mu([\omega])$
- For $k > m$,

$$|\varphi_{k,m}| = \left|\mu\left(T^{-n_k}[\omega] \,|\, \mathcal{B}_m\right) - \mu([\omega])\right| \leqslant C \cdot \alpha^{n_k - n_m} \leqslant C \cdot \alpha^{k-m}$$

for some constants $C > 0$, $\alpha \in (0,1)$ (see e.g. [16, Theorem 4.9]).

In particular M_m is a finite number for all m. Writing M_m as

$$M_m = \sum_{k=1}^{m-\ell} \left[\mathbf{1}_{[\omega]}\left(T^{n_k}x\right) - \mu\left([\omega]\right)\right] + \sum_{k=m-\ell+1}^{m} \varphi_{k,m} + \sum_{k=m+1}^{\infty} \varphi_{k,m}$$

and noting that $\left|\sum_{k=m-\ell+1}^{m} \varphi_{k,m}\right| \leqslant \ell$, we have

$$M_m - L \leqslant \sum_{k=1}^{m-\ell} \left[\mathbf{1}_{[\omega]}\left(T^{n_k}x\right) - \mu\left([\omega]\right)\right] \leqslant M_m + L$$

for some constant $L > 0$. Since M_m is a martingale w.r.t. the increasing sequence of σ-algebras $(\mathcal{B}_m)_{m\in\mathbb{N}}$, by Doob's martingale convergence theorem M_m converges almost surely to some number, which implies that a.s.

$$\frac{1}{m}\sum_{k=1}^{m-\ell} \left[\mathbf{1}_{[\omega]}\left(T^{n_k}x\right) - \mu\left([\omega]\right)\right] \underset{m\to\infty}{\longrightarrow} 0.$$

Since the countable family of cylinder sets generates the Borel σ-algebra of subsets of $\Omega^{\mathbb{N}}$, we get that a.s., for every function $f \in C\left(\Omega^{\mathbb{N}}\right)$,

$$\frac{1}{m}\sum_{k=1}^{m} f\left(T^{n_k}x\right) \underset{m\to\infty}{\longrightarrow} \int f d\mu.$$

□

Corollary 24. *Let γ_m be an equidistributed sequence w.r.t. some Borel probability measure σ on a compact second countable space X, and let $\left(\Omega^{\mathbb{N}}, T, \mu\right)$ be as above. Then for μ-a.e. $i \in \Omega^{\mathbb{N}}$, the sequence $(\gamma_m, T^m i)_{m=1}^{\infty}$ is equidistributed w.r.t. $\sigma \otimes \mu$.*

Proof. Given a set $I \times [\omega] \in X \times \Omega^{\mathbb{N}}$ where $I \subseteq X$ is some open set and $\omega \in \Omega^{\ell}$ for some $\ell \in \mathbb{N}$, consider the sum

$$\frac{1}{N}\sum_{m=1}^{N} \mathbf{1}_{I\times[\omega]}\left(\alpha_m, T^m i\right).$$

Define $A = \{m \in \mathbb{N} : \gamma_m \in I\}$ and let $(m_k)_{k\in\mathbb{N}}$ be an increasing enumeration of all elements in A. By equidistribution of γ_m we know that

$$\lim_{N\to\infty} \frac{|A \cap \{1, ..., N\}|}{N} = \sigma(I).$$

By Proposition 23, $(T^{m_k} i)_{k=1}^{\infty}$ is a.s. equidistributed w.r.t. μ, hence a.s.

$$\frac{1}{k}\sum \mathbf{1}_{[\omega]}\left(T^{m_k} i\right) \xrightarrow{a.s.} \mu\left([\omega]\right)$$

Denote $B \stackrel{\text{def}}{=} \{m \in A : T^m i \in [\omega]\}$. Then

$$\lim_{N\to\infty} \frac{|B \cap \{1, ..., N\}|}{|A \cap \{1, ..., N\}|} = \mu\left([\omega]\right) \text{ a.s.}$$

Hence,

$$\begin{aligned}\frac{1}{N}\sum_{m=1}^{N}\mathbf{1}_{I\times[\omega]}\left(\gamma_m,T^m i\right) &= \frac{\left|B\cap\{1,...,N\}\right|}{N}\\ =&\frac{\left|A\cap\{1,...,N\}\right|}{N}\cdot\frac{\left|B\cap\{1,...,N\}\right|}{\left|A\cap\{1,...,N\}\right|}\underset{N\to\infty}{\longrightarrow}\sigma\left(I\right)\cdot\mu\left([\omega]\right)\end{aligned}$$

From this we deduce that a.s., for every basic open set $I\subseteq X$, every $\ell\in\mathbb{N}$ and every $\omega\in\Omega^{\ell}$,

$$\frac{1}{N}\sum_{m=1}^{N}\mathbf{1}_{I\times[\omega]}\left(\gamma_m,T^m i\right)\underset{N\to\infty}{\longrightarrow}\sigma\left(I\right)\cdot\mu\left([\omega]\right)=\int\mathbf{1}_{I\times[\omega]}d(\sigma\otimes\mu).$$

By linearity of integration and summation, a.s. the same property holds for any linear combination of indicator functions as above, and thus on a dense subset of $C\left(X\right)$. □

Proof of Theorem 22. By Corollary 24, for β-a.e. $i\in\Lambda^{\mathbb{N}}$, the sequence $\left(\alpha_m,\eta_m\left(i\right)\right)$ is equidistributed w.r.t. $\nu\otimes p$, which implies that the sequence $\alpha_m+\eta_m\left(i\right)$ is equidistributed w.r.t. $\nu * p$. Using Proposition 19 exactly in the same way as in the proof of Theorem 4, we may deduce that for β-a.e. $i\in\Lambda^{\mathbb{N}}$, $\alpha_m+\eta_m\left(i\right)+\pi(T^m(i))$ is equidistributed w.r.t. the measure $\nu * p * \widetilde{\mu_K}$. □

From Theorem 22 we readily obtain:

Proof of Theorem 8. Suppose t_1 is normal to base D. Then so is $\dfrac{D}{D-1}t_1$, and therefore α_m is equidistributed w.r.t. Haar measure. The conclusion now follows immediately from Theorem 22. □

It is not hard to find an example in which the t_i are irrational and not normal to base D, and the conclusion of Theorem 8 fails. For example, this will hold when the t_i have many appearances of long strings of the digit 0 in base D. Nevertheless, the converse to Theorem 8 is also false. Here is a counterexample.

Example 25. Denote

$$f_1^{\alpha}\left(x\right)\overset{\text{def}}{=}\frac{1}{4}x+\alpha,\quad f_2^{\alpha}\left(x\right)\overset{\text{def}}{=}\frac{1}{4}x+\frac{1}{2}+\alpha.$$

Let K_α be the attractor of the IFS $\{f_1^{\alpha},f_2^{\alpha}\}$ for a given value of α. Note that changing α corresponds to translating the fractal K_0. More precisely, $K_\alpha=K_0+c_\alpha$ where $c_\alpha=\dfrac{4}{3}\alpha$. Let μ_α be the $\left(\frac{1}{2},\frac{1}{2}\right)$-Bernoulli measure on K_α. Note that for all $n\in\mathbb{Z}$,

$$\hat{\mu}_\alpha\left(n\right)=e^{2\pi i n c_\alpha}\hat{\mu}_0\left(n\right),$$

hence $\hat{\mu}_\alpha\left(n\right)=0$ if and only if $\hat{\mu}_0\left(n\right)=0$.

Denoting $\Delta_1=0$, $\Delta_2=\frac{1}{2}$ and $\Lambda=\{1,2\}$, the Fourier transform of μ_0 may be calculated as follows (see [8, proof of Theorem 6.1]):

$$\begin{aligned}\hat{\mu}_0\left(n\right)&=\lim_{N\to\infty}2^{-N}\sum_{j\in\Lambda^N}\exp\left(2\pi i n\sum_{s=1}^{N}4^{-s+1}\Delta_{j_s}\right)\\ &=\lim_{N\to\infty}2^{-N}\prod_{s=0}^{N-1}\left(1+\exp\left(2\pi i n 4^{-s}\frac{1}{2}\right)\right).\end{aligned}$$

Therefore

$$|\hat{\mu}_0(n)| = \prod_{s=0}^{\infty} \left| \cos\left(4^{-s}\frac{1}{2}\pi n\right) \right|. \tag{5.1}$$

Using equation 5.1, we see that for all $k, m \in \mathbb{Z}$ with $k \geqslant 0$, we have $\hat{\mu}_0\left(4^k(2m+1)\right) = 0$.

Let ν be the $\left(\frac{1}{2}, \frac{1}{2}\right)$-Bernoulli measure defined on the attractor of the IFS $\left\{x \mapsto \frac{1}{4}x,\, x \mapsto \frac{1}{4}x + \frac{1}{4}\right\}$. Analyzing $\hat{\nu}$ in the same way we analyzed $\hat{\mu}_0$, we see that

$$|\hat{\nu}(n)| = \prod_{s=0}^{\infty} \left| \cos\left(4^{-s}\frac{1}{4}\pi n\right) \right|,$$

and hence for all $k, m \in \mathbb{Z}$ with $k \geqslant 0$, we have $\hat{\nu}\left(4^k 2(2m+1)\right) = 0$.

Since ν is ergodic for the $\times 4$ map, it has generic points. Let t be a generic point for ν, and $\tilde{t} \stackrel{\text{def}}{=} \frac{3}{4}t$. By equation (4.3), we see that for every $x \in K_{\tilde{t}}$, if $x = \lim_{n\to\infty} f_{i_1}^{\tilde{t}} \circ \cdots \circ f_{i_n}^{\tilde{t}}(0)$ then for every $n \in \mathbb{N}$,

$$4^n x = \sum_{j=1}^{n} 4^j \tilde{t} + \pi\left(T^n(i)\right) = 4^n \frac{4}{3}\tilde{t} - \frac{4}{3}\tilde{t} + \pi\left(T^n(i)\right) = 4^n t + \pi\left(T^n(i)\right) - t,$$

where π is the coding map for the IFS $\left\{f_1^{\tilde{t}}, f_2^{\tilde{t}}\right\}$. By Corollary 24 and the computation above, we get that for $\mu_{\tilde{t}}$-a.e. $x \in K_{\tilde{t}}$, the orbit $(4^n x)_{n=1}^{\infty}$ is equidistributed w.r.t. the measure $\nu * \mu_{\tilde{t}}$ translated by $-t$.

Claim. For all nonzero $w \in \mathbb{Z}$, there exist $k \in \mathbb{N} \cup \{0\}$ and $m \in \mathbb{Z}$ such that either $w = 4^k(2m+1)$ or $w = 4^k(4m+2)$.

Proof of Claim. Clearly, it's enough to prove the statement for the case $4 \nmid w$. If w is odd, then $w = 4^0(2m+1)$ for some $m \in \mathbb{Z}$. Otherwise, w is even, and since we assume $4 \nmid w$, then $w = 4^0(4m+2)$ for some $m \in \mathbb{Z}$. ☐

By the Claim and the analysis of the Fourier transforms of ν and μ_0 given above, we get that for every $0 \neq n \in \mathbb{Z}$,

$$\widehat{\nu * \mu_{\tilde{t}}}(n) = \hat{\nu}(n) \cdot \hat{\mu}_{\tilde{t}}(n) = 0.$$

This implies that $\nu * \mu_{\tilde{t}}$ is Haar measure on $\mathbb{T}$, and ultimately we get that for $\mu_{\tilde{t}}$-a.e. $x \in K_{\tilde{t}}$, the orbit $(4^n x)_{n=1}^{\infty}$ is equidistributed w.r.t. Haar measure on $\mathbb{T}$ although $\tilde{t}$ is not normal to base 4.

Remark 26. The convolution in the example above may also be viewed as follows. The probability measure μ_0 is the law of the random variable $\sum_{j=1}^{\infty} 4^{-j}\xi_j$, where the ξ_j are IID variables which assume the values 0, 2 with probability $\frac{1}{2}$. The measure ν is the law of the random variable $\sum_{j=1}^{\infty} 4^{-j}\chi_j$, where the χ_j are IID variables which assume the values 0, 1 with probability $\frac{1}{2}$. Hence, $\nu * \mu_0$ is the law of the random variable $\sum_{j=1}^{\infty} 4^{-j}\chi_j + \sum_{j=1}^{\infty} 4^{-j}\xi_j$. But

$$\sum_{j=1}^{\infty} 4^{-j}\chi_j + \sum_{j=1}^{\infty} 4^{-j}\xi_j = \sum_{j=1}^{\infty} 4^{-j}\left(\chi_j + \xi_j\right)$$

and since $\chi_j + \xi_j$ are IID random variables that take the values $0, 1, 2, 3$ with probability $\frac{1}{4}$ each, $\nu * \mu_0$ is actually Haar measure on $\mathbb{T}$. Therefore, $\nu * \mu_{\tilde{t}}$ is also Haar measure.

References

[1] Y. Benoist and J.-F. Quint. Mesures stationnaires et fermés invariants des espaces homogènes. *Comptes Rendus Mathematique*, 347(1-2):9–13, 2009.

[2] Y. Benoist and J.-F. Quint. Stationary measures and invariant subsets of homogeneous spaces (ii). *Journal of the American Mathematical Society*, 26(3):659–734, 2013.

[3] Y. Benoist and J.-F. Quint. Stationary measures and invariant subsets of homogeneous spaces (iii). *Annals of Mathematics*, pages 1017–1059, 2013.

[4] Y. Benoist and J.-F. Quint. *Random walks on reductive groups*. 2016. published online.

[5] J. Bourgain, A. Furman, E. Lindenstrauss, and S. Mozes. Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. *Journal of the American Mathematical Society*, 24(1):231–280, 2011.

[6] L. Breiman. The strong law of large numbers for a class of markov chains. *The Annals of Mathematical Statistics*, 31(3):801–803, 1960.

[7] R. Broderick, Y. Bugeaud, L. Fishman, D. Kleinbock, and B. Weiss. Schmidt's game, fractals, and numbers normal to no base. *Math. Res. Lett.*, 17(2):307–321, 2010.

[8] Y. Bugeaud. *Distribution Modulo One and Diophantine Approximation*. Cambridge Tracts in Mathematics. Cambridge University Press, 2012.

[9] J. Cassels. On a problem of steinhaus about normal numbers. In *Colloquium Mathematicum*, volume 1, pages 95–101, 1959.

[10] K. J. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*. Wiley, 2013.

[11] J. Feldman and M. Smorodinsky. Normal numbers from independent processes. *Ergodic Theory and Dynamical Systems*, 12(4):707–712, 1992.

[12] H. Furstenberg. Noncommuting random products. *Transactions of the American Mathematical Society*, 108(3):377–428, 1963.

[13] H. Furstenberg. Stiffness of group actions. In *Lie groups and ergodic theory (Mumbai, 1996)*, volume 14 of *Tata Inst. Fund. Res. Stud. Math.*, pages 105–117. Tata Inst. Fund. Res., Bombay, 1998.

[14] M. Hochman and P. Shmerkin. Equidistribution from fractal measures. *Inventiones Mathematicae*, 202(1):427–479, 2015.

[15] B. Host. Nombres normaux, entropie, translations. *Israel Journal of Mathematics*, 91(1-3):419–428, 1995.

[16] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov chains and mixing times*. American Mathematical Society, Providence, RI, 2009. With a chapter by James G. Propp and David B. Wilson.

[17] W. M. Schmidt. On normal numbers. *Pacific Journal of Mathematics*, 10(2):661–672, 1960.

[18] W. M. Schmidt. On badly approximable numbers and certain games. *Transactions of the American Mathematical Society*, 123(1):178–199, 1966.

[19] D. Simmons and B. Weiss. Random walks on homogeneous spaces and diophantine approximation on fractals. *Inventiones mathematicae*, 216(2):337–394, 2019.

[20] D. D. Wall. *Normal Numbers*. ProQuest LLC, Ann Arbor, MI, Thesis (Ph.D.) - University of California, Berkeley, 1950.