ENTROPY AND UNIFORM DISTRIBUTION OF ORBITS IN $\mathbb{T}^d$

BY

DAVID MEIRI

Institute of Mathematics, The Hebrew University of Jerusalem
Jerusalem 91904, Israel
e-mail: dafid@math.huji.ac.il

ABSTRACT

We present a class of integer sequences $\{c_n\}$ with the property that for every p-invariant and ergodic positive-entropy measure μ on $\mathbb{T}$, $\{c_n x \pmod 1\}$ is uniformly distributed for μ-almost every x. This extends a result of B. Host, who proved this for the sequence $\{q^n\}$, for q relatively prime to p. Our class of sequences includes, for instance, the sequence $c_n = \sum f_i(n) q_i^n$, where the numbers q_i are distinct and are relatively prime to p and f_i are any polynomials. More generally, recursion sequences for which the free coefficient of the recursion polynomial is relatively prime to p are in this class as well, provided they satisfy a simple irreducibility condition.

In the multi-dimensional case we derive sufficient conditions for a pair of endomorphisms $A, B \in \mathrm{End}(\mathbb{T}^d)$ (with A diagonal) and an A-invariant and ergodic measure μ, such that B-orbits of the form $\{B^n\omega\}$ are uniformly distributed for μ-almost every $\omega \in \mathbb{T}^d$.

1. Introduction

The focus of this article is the following property, motivated by Host [5]:

Definition 1.1: Given an integer $p > 1$, a sequence of integers $\{c_n\}$ is called a **p-Host sequence** if it has the following property:

If μ is a Borel measure on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, invariant for σ_p: $x \mapsto px$ (mod 1), ergodic with positive entropy, then $\{c_n x \pmod 1\}$ is uniformly distributed for μ-almost every x.

(See more details in the background below.)

Received September 6, 1996

Denoting by λ Lebesgue measure and by $\sigma_c\mu = \mu \circ \sigma_c^{-1}$ the image of μ under the map $\sigma_c : x \mapsto c \cdot x \pmod 1$, an easy application of Weyl's criterion gives

$$\frac{1}{N}\sum_{n=0}^{N-1} \sigma_{c_n}\mu \longrightarrow \lambda.$$

Host [5] establishes a connection between property 1.1 and the asymptotic distribution of $\{c_k \pmod{p^n}\}_{k=0}^{p^n-1}$ when $n \longrightarrow \infty$. The main theorem there is that $\{q^n\}$ is a p-Host sequence, whenever p and q are relatively prime. In particular, Host concludes that a p- and q-invariant measure with positive entropy must be Lebesgue measure, thus giving a short and elegant proof of Rudolph's theorem [17].

The key to extending Host's theorem is a refinement of the combinatorial properties $\{c_k \pmod{p^n}\}_{k=0}^{p^n-1}$ is required to possess. We concentrate on sequences with *almost-sub-exponential collisions* mod p (see Definition 2.1). In essence, this means that after dropping a small set, the number of "collisions" (coincidences) between elements in $\{c_k \pmod{p^n}\}_{k=0}^{p^n-1}$ is not exponentially bigger than p^n. Host's theorem can be easily extended to sequences with this property (Theorem 3.1); this property, in turn, is valid for any sequence that is the restriction to $\mathbb{N}$ of a non-constant p-adic analytic function. More precisely, a **p-adic interpolation** of a sequence $\{c_k\}$ is a continuous function $G(x)$, defined in the unit disk of $\mathbb{Q}_p$ (or some finite extension of $\mathbb{Q}_p$), with $c_k = G(k)$ for all $k \in \mathbb{N}$. Since $\mathbb{N}$ is dense in the disk, such an interpolation, if it exists at all, is unique. It turns out that various analytic properties of G are reflected in various combinatorial properties of $\{c_k\}$; for an exact statement see Theorem 3.2. The two theorems together give:

THEOREM A: *Let $\{c_k\}$ be a non-constant sequence of integers, and p a prime number. Suppose $\{c_k\}$ has a smooth p-adic interpolation G, defined in the unit disk of $\mathbb{Q}_p$ or some finite extension of $\mathbb{Q}_p$. Suppose also that G has finitely many critical points. (This holds, in particular, if G is analytic.) Then $\{c_k\}$ is a p-Host sequence.*

If p is not prime, the same holds true if $\{c_k\}$ has a smooth interpolation with the same property in $\mathbb{Q}_{p'}$, for every prime p' dividing p.

A basic example of p-Host sequences arise from linear recursions (Theorem 5.2):

THEOREM B: *Let $\{c_k\}$ be a sequence of integers satisfying a linear recursion*

$$c_k = a_1c_{k-1} + a_2c_{k-2} + \cdots + a_Lc_{k-L}$$

for some integer constants $a_1, \dots, a_L$ ($a_L \neq 0$) and for all $k > L$. Assume that:

(i) *$\{c_k\}$ has no constant arithmetic subsequences (this is always the case if no roots of the recursion polynomial, or their ratios, are roots of unity).*

(ii) *a_L and p are relatively prime.*

Then $\{c_k\}$ is a p-Host sequence.

By composing analytic functions it is also possible to derive some examples of p-Host sequences which are not recursive; e.g., the sequences 2^{2^k}, 2^{F_k} and F_{2^k} are all 3-Host, when $\{F_k\}$ is the Fibonacci sequence.

The conditions of the last theorem are clear for $c_k = \sum_{i=1}^{L} f_i(k) q_i^k$, when $|q_1|, \dots, |q_L| > 1$ are distinct and are relatively prime to p. Using this theorem, we turn to study orbits of endomorphism in $\mathbb{T}^2$. We have:

THEOREM C: *For $A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$, let μ be an A-invariant and ergodic measure with positive entropy. Let $B \in \operatorname{End}(\mathbb{T}^2)$ be another endomorphism.*

Assume that for every d_1, d_2 not both zero, there exists $(a, b) \in \mathbb{Z}^2$ such that:

(i) *Projecting μ by $\pi_{a,b} : (x, y) \mapsto ax + by \pmod 1$, the entropy of the projected measure $h(\pi_{a,b}\mu, \sigma_p)$ is strictly smaller than $h(\mu, A)$.*

(ii) *$c_k = (d_1, d_2) B^k \binom{b}{-a}$ has almost-sub-exponential collisions $(\operatorname{mod} p)$.*

Then for μ-almost every $\omega \in \mathbb{T}^2$, the orbit $\{B^n \omega\}$ is uniformly distributed.

The theorem we prove is somewhat more general; see Theorem 6.1. For a general version in $\mathbb{T}^d$ we refer the reader to Meiri and Peres [14]. In Theorem 6.4 we use the results on linear recursions to give conditions on B in terms of its eigenvalues such that (ii) holds. The main idea here is the simple observation that a sequence $\{c_k\}$ defined by (ii) satisfies a recursion polynomial whose characteristic roots are the eigenvalues of B. We end this section with a discussion of maximal-entropy measures on Cantor sets. Here condition (i) above — a decrease in the entropy of μ — is equivalent to a decrease in the Hausdorff dimension, when the set in question is projected by $\pi_{a,b}$. As before, the conditions we derive also ensure that $\frac{1}{N} \sum B^n \mu$ converges weakly to Lebesgue measure on $\mathbb{T}^2$.

These measure-theoretic results can be used to examine bi-invariant sets in $\mathbb{T}^2$, leading to reducible analogues of Berend's results (see background). In particular we prove in 7.1 the following result of B. Kra:

THEOREM D ([11]): *For $i = 1, 2$, let p_i, q_i be a pair of relatively prime integers > 1. Suppose that $p_1 \neq p_2$ or $q_1 \neq q_2$. Then for every irrational α, β, the set $\{p_1^n q_1^m \alpha + p_2^n q_2^m \beta\}_{n,m \geq 1}$ is dense* mod *1.*

ORGANIZATION. After some background in the rest of this introduction, we define in §2 the combinatorial conditions we need of integer sequences when projected to $\mathbb{Z}/p^n\mathbb{Z}$.

In §3 we formulate and prove an extension of Host's theorem (3.1), and formulate our main Theorem 3.2. The theorem is proved in §4. Together they yield Theorem A above.

In §5 we investigate linear recursions, and prove Theorem B.

In §6 we study orbits of endomorphisms in $\mathbb{T}^2$, and derive some multi-dimensional analogues to Host's theorem, including Theorem C. This is used in §7 to examine bi-invariant sets in $\mathbb{T}^2$, and to prove Kra's theorem D.

Finally, §8 contains concluding remarks and questions.

BACKGROUND. H. Furstenberg [4] showed that an infinite closed set in $\mathbb{T}$, invariant under multiplication by two multiplicatively independent integers p and q ($\log p/\log q \notin \mathbb{Q}$), must be all of $\mathbb{T}$ (cf. [2]). He also asked if a similar result holds for bi-invariant measures on $\mathbb{T}$, i.e., if the only non-atomic Borel measure μ on $\mathbb{T}$ invariant and ergodic under multiplication by p and q is Lebesgue measure.

Berend [1] extended Furstenberg's topological results to $\mathbb{T}^d$, establishing necessary and sufficient conditions on a commutative semigroup $G \subset \mathrm{End}(\mathbb{T}^d)$, such that the only closed infinite G-invariant set in $\mathbb{T}^d$ is $\mathbb{T}^d$ itself. The measure-theoretic question remained open for a longer period, and in fact, in the form presented by Furstenberg is still unsolved, since all authors also need to assume μ has positive entropy. Lyons [13] obtained a first partial result under the extra assumption that μ is exact and p and q are relatively-prime. Rudolph [17] used symbolic dynamics to give a new proof without the exactness assumption, and two years later Johnson [6] extended the theorem for p and q multiplicatively independent. Feldman [3], motivated by Lyons' argument, gave another proof.

A different approach was taken recently by B. Host [5]. Assuming only that μ is p-invariant, ergodic and with positive entropy, Host noticed that the desired result would follow for an integer q from the stronger statement that for μ-almost every $x \in \mathbb{T}$, the sequence $\{q^n x \pmod 1\}$ is *uniformly distributed*, i.e., that

$$\frac{1}{N}\#\{0 \leq k < N : q^k x \in I\} \longrightarrow |I|,$$

for all intervals $I \subset \mathbb{T}$. In other words, the requirement is that μ-almost every x is *normal* in base q. By Weyl's criterion this is equivalent to

$$\forall a \in \mathbb{Z}, a \neq 0, \quad \frac{1}{N}\sum_{n=0}^{N-1} e(aq^n x) \longrightarrow 0 \quad \mu\text{-a.e.,} \tag{1}$$

writing $e(t) = e^{2\pi \imath t}$. This condition implies

$$\frac{1}{N}\sum_{n=0}^{N-1} \sigma_q^n \mu \longrightarrow \lambda \quad \text{(in the weak* topology)}. \tag{2}$$

Johnson and Rudolph [7] noticed that from (2) it follows that

$$\sigma_q^n \mu \longrightarrow \lambda, \qquad \text{when } n \longrightarrow \infty \text{ and } n \notin J \tag{3}$$

for some set J with zero density.

Finally, (3) implies

$$\sigma_q \mu = \mu \quad \Longrightarrow \quad \mu = \lambda. \tag{4}$$

Thus, in a sense, Furstenberg's conjecture is the weakest among several possible statements. On the other hand, knowing that every bi-invariant positive-entropy measure μ ergodic to the action of the semigroup $\langle \sigma_p, \sigma_q \rangle$ is Lebesgue, guarantees (2) for every p-invariant ergodic positive-entropy measure μ. Indeed, if ν is some weak* limit point of the averages in (2), by the upper semi-continuity of entropy for expansive maps (cf. Walters [19], Theorem 8.2), $h(\nu, \sigma_p) \geq h(\mu, \sigma_p) > 0$. Surely, ν is σ_p and σ_q invariant. Let $\nu = \int \nu_z \, dz$ be the ergodic decomposition with respect to $\langle \sigma_p, \sigma_q \rangle$. By our assumption, every component with positive entropy must be Lebesgue. An argument by Johnson and Rudolph [7] now shows that almost all ergodic components have positive entropy, hence $\nu = \lambda$.

Several authors have investigated normality of numbers in different bases (cf. [5] for some references and discussion). Host used this approach to prove (1) when p and q are relatively prime, thus deriving Rudolph's theorem (cf. Theorem 3.1 below). Unfortunately his method does not seem to extend to multiplicatively independent pairs of integers (cf. [12]).

2. Combinatorial properties mod p^n

We denote by $\#A$ or $|A|$ the cardinality of a finite set A. We will use $\mathbb{N}$ to denote the set of non-negative integers.

Looking at the first N values of $\{c_k \pmod{p^n}\}$, we examine *sizes of cells*

$$\Delta_t(N, n) = \#\{k : 0 \leq k < N, c_k \equiv t \pmod{p^n}\}$$

for a cell number t, $0 \leq t < p^n$. Assembling these together gives the **collision number**

$$\Gamma(N, n) = \sum_{t=0}^{p^n - 1} \Delta_t(N, n)^2 = \#\{0 \leq k, l < N \colon c_k \equiv c_l \pmod{p^n}\}.$$

We will be interested in computing this number for $N \approx p^n$. Obviously $\Gamma(N,n) \geq N$, and we will need to know when $\Gamma(N,n)$ is not much bigger. A set $A \subset \mathbb{N}$ is said to have density $\leq \gamma$ if, for every N,

$$\frac{|A \cap \{0,1,\ldots,N-1\}|}{N} \leq \gamma. \tag{5}$$

Definition 2.1: Let $\{c_k\}$ denote some integer-valued sequence, and fix an integer $p > 1$.

(i) $\{c_k\}$ has **bounded cells** $(\mathrm{mod}\, p)$ if for some M we have $\Delta_t(p^n,n) \leq M$ for all $n \geq 1$ and $0 \leq t < p^n$. (In this case $\Gamma(p^n,n) \leq p^n M^2$.)

(ii) $\{c_k\}$ has **sub-exponential collisions** $(\mathrm{mod}\, p)$ if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \Gamma(p^n,n)/p^{n(1+\varepsilon)} = 0.$$

(iii) $\{c_k\}$ has **almost-bounded cells** $(\mathrm{mod}\, p)$ if for every $\gamma > 0$ there exists a set $A \subset \mathbb{N}$ with density $\leq \gamma$ and some M (which might depend on γ) such that for every $n \geq 1$ and every $0 \leq t < p^n$,

$$\#\{k : 0 \leq k < p^n,\ k \notin A,\ c_k \equiv t \pmod{p^n}\} \leq M.$$

(iv) Similarly, $\{c_k\}$ has **almost-sub-exponential collisions** $(\mathrm{mod}\, p)$ if for every $\gamma > 0$ there exists a set A with density $\leq \gamma$, such that for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \#\{k,l : 0 \leq k,l < p^n,\ k,l \notin A,\ c_k \equiv c_l \pmod{p^n}\}/p^{n(1+\varepsilon)} = 0. \tag{6}$$

Examples:

1. If q is relatively prime to p, it is easy to see that the order of q in $(\mathbb{Z}/p^n\mathbb{Z})^*$ has magnitude p^n; hence $c_k = q^k$ has bounded cells mod p.

2. $c_k = k^2$ has sub-exponential collisions, but does not have bounded cells; in fact one can show that its collision number $\Gamma(p^n,n)$ is of order np^n. For prime p the exact result is

$$\Gamma(p^n,n) = \begin{cases} np^n & \text{if } p = 2, \\ (1 + \frac{p-1}{p}n)p^n & \text{if } p > 2,\ p \text{ prime.} \end{cases}$$

3. $c_k = k^r$, $r \geq 3$, does not have sub-exponential collisions; indeed,

$$\Delta_0(p^n, n) = \#\{0 \leq k < p^n \colon k^r \equiv 0 \pmod{p^n}\} = p^n / p^{\lceil \frac{n}{r} \rceil} \approx p^{n(1-1/r)},$$

   hence $\Gamma(p^n, n) \geq \Delta_0(p^n, n)^2 \approx p^{n \cdot 2(1-1/r)}$. Since $2(1-1/r) > 1$, the zero cell alone is already too large. However, from Theorem 5.2 it will follow that $\{k^r\}$ has almost-bounded cells.

4. According to Host [5], the Fibonacci sequence has a collision number $\Gamma(p^n, n)$ of magnitude np^n (a result of T. Kamae), and so has sub-exponential collisions. The fact that it is a p-Host sequence also follows from our theorems.

5. The following example (by Y. Peres) shows that a sequence with sub-exponential collisions need not have almost-bounded cells. Let $p = 5$, and build $\{c_k\}$ from increasing powers of 3, where 3^j appears $\lfloor \log j \rfloor$ times. When comparing $\{c_k\}_{k=0}^{5^n-1}$ to $\{3^k\}$, it is easy to see that about half the cells expanded in approximately factor n. Remembering that $\{3^k\}$ has bounded cells mod 5, the claim is now clear.

The examples above show that the only implications among the four properties are the trivial ones: (i)$\Longrightarrow$(ii)$\Longrightarrow$(iv) and (i)$\Longrightarrow$(iii)$\Longrightarrow$(iv).

3. Classes of p-Host sequences

The main result we present assembles these examples and many others. We can summarize it in two theorems, the first a reformulation of a theorem of Host, originally proved for sequences whose period when projected to $\mathbb{Z}/p^n\mathbb{Z}$ is of magnitude p^n.

THEOREM 3.1: *Given an integer $p > 1$, if $\{c_k\}$ has almost-sub-exponential collisions* mod p, *then it is a p-Host sequence.*

Our main theorem presents sufficient analytic conditions as to when a sequence $\{c_k\}$ has bounded cells or almost-bounded cells. (Remember that both properties imply almost-sub-exponential collisions.) For a prime p, we denote by $\mathbb{Q}_p$ the p-adic number field. Given a finite extension $\mathbb{K} \supseteq \mathbb{Q}_p$, let

$$B(\mathbb{K}) = \{x \in \mathbb{K} : |x|_p \leq 1\}$$

denote the closed unit disk of $\mathbb{K}$. Henceforth $|\cdot|_p$ will denote the extension of the p-adic norm from $\mathbb{Q}_p$ to $\mathbb{K}$. By a smooth p-adic interpolation of $\{c_k\}$ we mean a continuously differentiable function G, defined in $B(\mathbb{K})$ for some extension $\mathbb{K}$, such that $G(k) = c_k$ for all $k \in \mathbb{N}$.

THEOREM 3.2: *Let $\{c_k\}$ be a sequence of integers, and let $q > 1$ denote a fixed integer. Suppose that for each prime p dividing q, the sequence $\{c_k\}$ has a smooth p-adic interpolation G_p, defined in the unit disk $B(\mathbb{K}_p)$ of some finite extension $\mathbb{K}_p \supseteq \mathbb{Q}_p$.*

(i) *If each G_p has no critical points, i.e., $G'_p(x) \neq 0$ for all $p|q$ and $x \in B(\mathbb{K}_p)$, then $\{c_k\}$ has bounded cells* $\bmod q$.

(ii) *If each G_p has only a finite number of critical points, then $\{c_k\}$ has almost-bounded cells* $\bmod q$. *(This condition is clear when G_p is a non-constant analytic function.)*

We postpone the proof of Theorem 3.2 to the next section.

Remark: An immediate example of a p-Host sequence is a polynomial sequence, i.e., a sequence of the form $c_k = \sum_{i=0}^{d} a_i k^i$. Indeed, for every irrational α, a classic result of Weyl states that $\{c_k\alpha \pmod 1\}$ is uniformly distributed. Since an ergodic measure μ with positive entropy must be non-atomic, $\mu(\mathbb{Q} \cap \mathbb{T}) = 0$, hence $\{c_k x\}$ is uniformly distributed μ-a.e. From the combinatorial point of view, all polynomial sequences have almost-bounded cells.

Proof of Theorem 3.1: We rephrase here the proof from [5] with the necessary modifications. Fix an integer $a \neq 0$. Define $e(x) = e^{2\pi i x}$ for $x \in \mathbb{T} = \mathbb{R}/\mathbb{Z}$, and let

$$g_N(x) = \frac{1}{N} \sum_{k=0}^{N-1} e(ac_k x).$$

We need to prove $g_N \longrightarrow 0$ μ-a.e., when μ is as in Definition 1.1.

Define $\omega_n = \sum_{j=0}^{p^n-1} \mu * \delta_{jp^{-n}}$, and denote by $\varphi_n = d\mu/d\omega_n$ the Radon–Nikodym derivative.

LEMMA 3.3: $\varphi_n^{1/n} \longrightarrow e^{-h}$ *μ-a.e., where $h = h(\mu, \sigma_p)$.*

Proof: Denote by α the p-partition:

$$\alpha = \left\{ \left[\frac{j}{p}, \frac{j+1}{p} \right) \right\}_{j=0}^{p-1}.$$

The lemma follows from the observation that $-\log \varphi_n = I_{\alpha_0^{n-1}|\alpha_n^\infty}$, the conditional information function. (See a detailed proof in [5], or more hints in the proof of Theorem 6.1 below.) Now,

$$\begin{aligned}-\frac{1}{n}\log\varphi_n &= \frac{1}{n}I_{\alpha_0^{n-1}|\alpha_n^\infty}\\ &= \frac{1}{n}[I_{\alpha|\alpha_1^\infty}+I_{\alpha_1|\alpha_2^\infty}+\cdots+I_{\alpha_{n-1}|\alpha_n^\infty}]\\ &= \frac{1}{n}[I_{\alpha|\alpha_1^\infty}+\sigma_p I_{\alpha|\alpha_1^\infty}+\cdots+\sigma_p^{n-1}I_{\alpha|\alpha_1^\infty}]\\ &\longrightarrow \int I_{\alpha|\alpha_1^\infty}\,d\mu = h,\end{aligned}$$

by the Ergodic Theorem. ∎

Note that $\{ac_k\}$ has almost-sub-exponential collisions as well. Fix $\gamma > 0$, and let $A \subset \mathbb{N}$ be a set with density $\le \gamma$, as in Definition 2.1 applied to $\{ac_k\}$. Define

$$\bar{g}_N(x) = \frac{1}{N}\sum_{\substack{k=0\\ k\notin A}}^{N-1} e(ac_k x).$$

Since $\sup_{x\in\mathbb{T}}|g_N(x)-\bar{g}_N(x)| \le |A\cap\{0,\dots,N-1\}|/N \le \gamma$, and γ can be made arbitrarily small, it suffices to prove that $\bar{g}_N \longrightarrow 0$ μ-a.e.

Lemma: *For every $\varepsilon > 0$ and every n large enough, for all $p^{n-1} \le N \le p^n$*

$$\int \frac{|\bar{g}_N(x)|^2}{\varphi_n(x)}\,d\mu \le N^{2\varepsilon}. \tag{7}$$

Proof of lemma: Since $\frac{1}{\varphi_n}\,d\mu \le d\omega_n$, we have

$$\begin{aligned}\int \frac{|\bar{g}_N(x)|^2}{\varphi_n(x)}\,d\mu \le \int |\bar{g}_N(x)|^2\,d\omega_n &= \int \sum_{j=0}^{p^n-1} |\bar{g}_N(x+jp^{-n})|^2\,d\mu(x)\\ &\le \frac{1}{N^2}\sum_{\substack{k,l=0\\ k,l\notin A}}^{N-1}\sum_{j=0}^{p^n-1} e((ac_k-ac_l)jp^{-n})\cdot\left|\int e(a(c_k-c_l)x)\,d\mu(x)\right|.\end{aligned}$$

Since the summation over j is zero when $ac_k \not\equiv ac_l \pmod{p^n}$ and is p^n otherwise, we get

$$\begin{aligned}\int \frac{|\bar{g}_N(x)|^2}{\varphi_n(x)}\,d\mu &\le \frac{1}{N^2}p^n\cdot\#\{0\le k,l<N\colon k,l\notin A,\ ac_k\equiv ac_l \pmod{p^n}\}\\ &\le \frac{p^n}{N^2}p^{n(1+\varepsilon)} \le N^{2\varepsilon},\end{aligned}$$

using (6), for every n large enough. ∎

To complete the proof of the theorem, take $\delta < h/2\log p$ and $\varepsilon < \delta/2$. Choose some k with $k(\delta - 2\varepsilon) > 1$. For each N, let $n = n(N)$ be the unique integer such that $p^{n-1} < N^k \leq p^n$. Assuming without loss of generality that (7) holds for $n \geq 1$ we have

$$\sum_{N=1}^{\infty} \int \frac{|\bar{g}_{N^k}|^2}{N^{k\delta}\varphi_{n(N)}} \, d\mu \leq \sum_N \frac{N^{2k\varepsilon}}{N^{k\delta}} = \sum_N \frac{1}{N^{k(\delta-2\varepsilon)}} < \infty.$$

Hence

$$\frac{|\bar{g}_{N^k}|^2}{N^{k\delta}\varphi_{n(N)}} \longrightarrow 0 \quad \mu\text{-a.e.} \qquad \text{when } N \to \infty.$$

For almost every x, $\varphi_n(x) \leq e^{-nh/2}$ for all n sufficiently large, and so for our choice of δ one can see that $N^{k\delta}\varphi_{n(N)}(x)$ is bounded. Thus $\bar{g}_{N^k}(x) \longrightarrow 0$ as well. From here it is easy to conclude that $\bar{g}_N \longrightarrow 0$ μ-a.e. ∎

To conclude this section, and before proving Theorem 3.2, we list a few properties of $\mathcal{H}_p$, the class of p-Host sequences.

PROPOSITION 3.4: *$\mathcal{H}_p$ is closed under multiplication by a scalar, addition of a scalar, and translation. But a sum of two sequences in $\mathcal{H}_p$ might fail to be a p-Host sequence.*

Proof: If $\{c_k x\}$ is uniformly distributed (mod 1), so are the sequences $\{ac_k x\}$, $\{(c_k + a)x\}$ and $\{c_{k+a}x\}$ (for a non-zero integer a). On the other hand, if q is relatively prime to p, the sequences $\{q^n\}$ and $\{p^n - q^n\}$ have bounded cells mod p, and so are in $\mathcal{H}_p$, but their sum $\{p^n\}$ is obviously not. ∎

Next we give an example of a sequence that does not have almost-sub-exponential collisions, yet is a p-Host sequence. This shows that the converse to Theorem 3.1 is false.

Example 3.5: The sequence $\{2^n 3^m\}$ for $p = 2, 3$.

Arrange the numbers $\{2^n 3^m\}_{n,m=0}^{\infty}$ in increasing order: $1 = c_0 < c_1 < c_2 \cdots$, and let p be 2 or 3. It is easy to see that $G(k) = c_k$ cannot be extended continuously to the unit disk of (any extension of) $\mathbb{Q}_p$. In fact, $\{c_k\}$ lacks any of the properties of Definition 2.1 for $p = 2, 3$, since most values of $c_k \pmod p$ are zero. We claim, however, that $\{c_k\}$ is a p-Host sequence for $p = 2, 3$. We show this for $p = 2$. We need the following lemma, whose proof is part of Breiman's proof of the Shannon–McMillan–Breiman Theorem (cf. [16], p. 261).

LEMMA 3.6: *In an ergodic measure preserving system $(X, \mathfrak{B}, \mu, T)$, if $f_n \longrightarrow f$ a.e. and in L^1, and if $\sup_n |f_n|$ is in $L^1(\mu)$, then*

$$\frac{1}{n}\sum_{k=0}^{n-1} T^k f_{n-k} \longrightarrow \int f\, d\mu \quad \text{a.e. and in } L^1.$$

Fix some $a \neq 0$. Given a σ_2-invariant and ergodic measure μ with $h(\mu, \sigma_2) > 0$, we need to show that

$$g_N(x) \overset{\text{def}}{=} \frac{1}{N}\sum_{k=0}^{N-1} e(ac_k x) \longrightarrow 0 \quad \mu\text{-a.e.}$$

Write $\alpha = \log_3 2$, $L = \lfloor \log_2 N \rfloor$, $M = \lfloor CL^2 \rfloor$. Then

$$\begin{aligned} g_M(x) &\approx \frac{1}{M}\sum_{n\log 2 + m\log 3 \leq N} e(a2^n3^m x) \\ &= \frac{1}{M}\sum_{n=0}^{\lfloor \log_2 N\rfloor}\sum_{m=0}^{\lfloor \log_3 N - n\alpha\rfloor} e(a2^n3^m x) \\ &\approx \frac{1}{CL^2}\sum_{n=0}^{L}\sigma_2^n\left(\sum_{m=0}^{(L-n)\alpha} e(a3^m x)\right) \\ &= \frac{1}{CL}\sum_{n=0}^{L}\sigma_2^n f_{L,n} \end{aligned}$$

where

$$f_{L,n}(x) = \frac{1}{L}\sum_{m=0}^{(L-n)\alpha} e(a3^m x), \qquad \text{for } 0 \leq n \leq L.$$

Noting that

$$f_{L,n} = \frac{L-n}{L} f_{L-n,0}$$

and writing $F_k = |f_{k,0}|$, we have

$$|g_M(x)| \leq \frac{1}{CL}\sum_{n=0}^{L}\sigma_2^n |f_{L,n}| \leq \frac{1}{CL}\sum_{n=0}^{L}\sigma_2^n F_{L-n}.$$

From Theorem 3.1 we know that $F_k \longrightarrow 0$ μ-a.e. Also, $|F_k| \leq 1$. From Lemma 3.6 it follows that $\frac{1}{L}\sum_{n=0}^{L}\sigma_2^n F_{L-n} \longrightarrow 0$ μ-a.e., and our claim is proved.

Example 3.7: Composition sequences.

Let B denote the unit ball of a finite extension $\mathbb{K} \supset \mathbb{Q}_p$, and suppose that $f, g : B \longrightarrow B$ are analytic. Then $f \circ g$ need not be analytic (cf. Schikhof [18], Ex. 41.1), but is still smooth and with a finite number of critical points. Using Theorem 5.2, this gives examples of p-Host sequences which do not satisfy a linear recursion. For instance, for $p = 3$, the following sequences have almost-bounded cells (mod 3): 2^{2^k}, 2^{5^k}, 2^{F_k} and F_{2^k}, where $\{F_k\}$ is the Fibonacci sequence.

4. Proof of Theorem 3.2

We follow the notation of Theorem 3.2. For a prime $p|q$, let $B_p = B(\mathbb{K}_p)$ be the domain of G_p, and let $Z_p = \{x \in B_p\colon G'_p(x) = 0\}$.

Let $q = \prod_{i=1}^{s} p_i^{\beta_i}$ be the decomposition of q to prime factors, and set $\hat{q} = \prod p_i$. We initially treat the first case, where $Z_p = \emptyset$ for all primes $p|q$.

CLAIM 4.1: *If for every prime $p|q$ we have $G'_p(x) \neq 0$ for all $x \in B_p$, then $\{c_k\}$ has bounded cells.*

The key lemma is

LEMMA: *For every prime $p|q$ there exist integers $d, r > 0$ such that for every pair of distinct $x, y \in B_p$*

$$(8) \qquad |x - y|_p \leq p^{-d} \quad \Longrightarrow \quad |G_p(x) - G_p(y)|_p \geq p^{-r}|x - y|_p.$$

Proof of lemma: To ease the notation we fix p, and write G for G_p.

If (8) were not true, we could find for every n a pair of distinct points x_n, y_n satisfying

$$|x_n - y_n|_p \leq \frac{1}{n}, \qquad |G(x_n) - G(y_n)|_p < \frac{1}{n}|x_n - y_n|_p.$$

Recalling that $\mathbb{K}_p$ is a *finite* extension of $\mathbb{Q}_p$, its unit disk B_p is compact. Hence $\{x_n\}$ has a convergent subsequence, so we can assume $x_n \longrightarrow x$. We must have $y_n \longrightarrow x$ as well, and so

$$\frac{G(x_n) - G(y_n)}{x_n - y_n} \longrightarrow G'(x).$$

But

$$\left|\frac{G(x_n) - G(y_n)}{x_n - y_n}\right|_p < \frac{1}{n},$$

hence $G'(x) = 0$, a contradiction. ∎

By taking d, r large enough we can assume that for all distinct $k, l \in \mathbb{N}$

$$(9) \qquad |k-l|_p \le p^{-d} \quad \Longrightarrow \quad |G(k)-G(l)|_p \ge p^{-r}|k-l|_p$$

for every prime $p|q$. Here we write $G(k) = c_k$ for $k \in \mathbb{N}$.

Proof of Claim 4.1: To prove the claim, we show that the size of each cell is bounded by $\hat{q}^{r+d}$. So pick some arbitrary t-cell

$$C = \{k : 0 \le k < q^n,\ G(k) \equiv t \pmod{q^n}\}.$$

Break C into $\hat{q}^d$ sub-cells, according to the value of $k \pmod{\hat{q}^d}$. Suppose that k, l are in the same sub-cell. For $i = 1, \ldots, s$, since $\hat{q}^d | k-l$ we have $|k-l|_{p_i} \le p_i^{-d}$, and so, by (9), $|k-l|_{p_i} \le p_i^r |G(k)-G(l)|_{p_i}$. But $G(k) \equiv G(l) \pmod{q^n}$, hence $|G(k)-G(l)|_{p_i} \le p_i^{-n\beta_i}$, and we conclude that $|k-l|_{p_i} \le p_i^{-(n\beta_i - r)}$, or in other words $p_i^{n\beta_i - r}|k-l$. Since this is true for all $i = 1, \ldots, s$, we conclude that $k-l$ is divisible by $\prod p_i^{n\beta_i - r} = q^n/\hat{q}^r$. Thus the size of each sub-cell of C is bounded by $\hat{q}^r$. Summing over the $\hat{q}^d$ sub-cells we see that the size of the t-cell is bounded by $\hat{q}^{r+d}$. ∎

We turn now to the general case, where Z_p is only assumed to be finite. Let

$$A_m = \{k \in \mathbb{N} : |k-z|_p \le p^{-m} \text{ for some prime } p|q \text{ and some } z \in Z_p\}.$$

We claim that $\{A_m\}$ can serve as the exceptional sets of Definition 2.1. We first note that

$$\begin{aligned} |A_m \cap \{0, \ldots, q^n - 1\}| &\le \sum_{p|q} \sum_{z \in Z_p} \#\{0 \le k < q^n : |k-z|_p \le p^{-m}\} \\ &\le \sum_{p|q} |Z_p| \frac{q^n}{p^m} \end{aligned}$$

and so $|A_m \cap \{0, \ldots, q^n - 1\}|/q^n \longrightarrow 0$ when $m \longrightarrow \infty$.

LEMMA: *For every m and every prime $p|q$ there exist positive constants d, r such that equation* (8) *holds for every pair of distinct x, y in $B_p \setminus A_m(p)$, where*

$$A_m(p) \overset{\text{def}}{=} \{z \in B_p : |x-z|_p \le p^{-m} \textit{ for some } z \in Z_p\}.$$

Proof of lemma: Observe that $\inf_{x \notin A_m(p)} |G_p'(x)| > 0$, by the continuity of G_p' in B_p. ∎

By taking d, r valid for all $p|q$, we can see as before that $\{c_k : k \notin A_m\}$ has bounded cells. This concludes the proof of the theorem. ∎

Remark: An elaborate study of the situation around critical points can be conducted, showing that the sizes of cells in $\{c_k : k \notin A_m\}$ is of magnitude at most $O(\hat{q}^{2mu})$, where u is the maximal order of a critical point. We omit the proof.

The theorem can be refined in several ways: First, from the proof of the general case it is clear that the assertion (ii) of the theorem remains true if we replace the hypothesis of the theorem with the following condition:

> *for each $p|q$ there exists a closed set $N_p \subset B_p$ with zero Haar measure, such that G_p is continuously differentiable and with no critical points in $B_p \setminus N_p$.*

Indeed, all we need to do in this case is to enlarge $A_m(p)$ (and hence A_m) by a finite number of small balls, covering $N_p \cup Z_p$.

Another refinement is the following: suppose that for every k, we choose c_k arbitrarily from $G_1(k), \ldots, G_l(k)$, where each of the l functions $G_1, \ldots, G_l$ satisfies the conditions of the theorem. Then $\{c_k\}$ has almost-bounded cells, and if $G_1, \ldots, G_l$ have no critical points at all, then $\{c_k\}$ has bounded cells. This is clear, since each cell in $\{c_k\}$ is contained in the union of corresponding cells from $\{G_1(k)\}, \ldots, \{G_l(k)\}$. This simple observation will be useful in the next section.

5. Linear recursions

Let $\{c_k\}$ be a sequence of integers satisfying a linear recursion of length L:

$$c_k = a_1 c_{k-1} + a_2 c_{k-2} + \cdots + a_L c_{k-L}$$

for some integer constants $a_1, \ldots, a_L$ ($a_L \neq 0$) and for all $k > L$. We ask in this section when such a sequence is a q-Host sequence. By Theorem 3.2 we can get a sufficient condition once we realize it as the restriction of a smooth (or, as it turns out, an analytic) function in a finite extension of $\mathbb{Q}_p$, for any prime factor p of q.

Let $\lambda_\imath$ be the (complex) roots of the recursion polynomial

$$\lambda^L = a_1 \lambda^{L-1} + \cdots + a_{L-1}\lambda + a_L. \tag{10}$$

Then $\lambda_\imath^k$ are the fundamental solutions of the recursion, and $\{c_k\}$ can be represented as $c_k = G(k) = \sum b_\imath(k)\lambda_\imath^k$, where each $b_\imath$ is a polynomial, and $\deg(b_i) \leq$ multiplicity of $\lambda_\imath$.

Usually this correspondence is used for real or complex parameter, but we can carry it without any change to a finite extension of $\mathbb{Q}_p$ containing all the roots of (10).

Example: Fibonacci sequence: The simplest non-trivial example is of course $c_k = c_{k-1} + c_{k-2}$. Here we should take

$$G(x) = b_1 \left(\frac{1+\sqrt{5}}{2} \right)^x + b_2 \left(\frac{1-\sqrt{5}}{2} \right)^x .$$

In $\mathbb{R}$, $G(x)$ is well-defined for an integer x, but not for a continuous parameter. In $\mathbb{Q}_p$, however, the situation is different (and, eventually, better):

(i) $\mathbb{Q}_p$ contains a square root of 5 only when 5 is a quadratic residue mod p, and this happens when $p \equiv 1 \pmod 5$ or $p \equiv 4 \pmod 5$. For other p we replace $\mathbb{Q}_p$ by $\mathbb{Q}_p(\sqrt{5})$.

(ii) A second problem is that $G(x)$ is not defined for all $x \in B$, the unit disk of $\mathbb{Q}_p$ (or $\mathbb{Q}_p(\sqrt{5})$): for $\lambda^x = \exp_p(x \log_p \lambda)$ to be defined, we must have $|\lambda - 1|_p < 1$ (so that $\log_p \lambda$ is defined) and $|x \log_p \lambda|_p < p^{-1/(p-1)}$ (so that $\exp_p$ is defined).

We will shortly show how to overcome these problems, but first we wish to consider another possible obstacle: if $\{c_k\}$ has a substantial constant subsequence, it obviously cannot have bounded cells (or almost-sub-exponential collisions). In particular this is true for an arithmetic subsequence, i.e., a subsequence along an arithmetic progression.

Definition 5.1: Let $x_k = a_1 x_{k-1} + a_2 x_{k-2} + \cdots + a_L x_{k-L}$ be a recursion scheme, and assume $a_L \neq 0$. We call the scheme **non-degenerate** if the only sequence $\{c_k\}$ satisfying the scheme and having a constant arithmetic subsequence is the zero sequence.

Let $\lambda_1, \ldots, \lambda_{L'}$ be the $L' \leq L$ *distinct* roots of (10). When all the roots are simple, $L' = L$. The roots are all non-zero, since $a_L \neq 0$.

PROPOSITION 5.1: *In the above notations, the recursion scheme is non-degenerate iff*

(11) *λ_i and λ_i/λ_j are not roots of unity, for every $1 \leq i < j \leq L'$.*

Proof: To see the necessity of the condition, pick some $d \geq 1$ and expand the

Van der Monde determinant (denoting $\lambda_0 = 1$)

$$\det\begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & \lambda_1^d & \dots & \lambda_{L'}^d \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_1^{L'd} & \dots & \lambda_{L'}^{L'd} \end{pmatrix} = \prod_{0\le i<j\le L'} (\lambda_j^d - \lambda_i^d).$$

If (11) does not hold, the determinant is zero for some d, hence $\{1\}_k, \{\lambda_1^{kd}\}_k, \dots, \{\lambda_{L'}^{kd}\}_k$ are linearly dependent, and so there are constants $b_0, b_1, \dots, b_{L'}$, not all of them zero, such that

$$b_1\lambda_1^{kd} + \cdots + b_{L'}\lambda_{L'}^{kd} = b_0 \tag{12}$$

holds for all k. The left hand side is a constant subsequence of $c_k = b_1\lambda_1^k + \cdots + b_{L'}\lambda_{L'}^k$, and $\{c_k\}$ is a non-trivial sequence satisfying the recursion.

To prove sufficiency of the condition, first note that if λ_i were all simple roots, we would be done by the same calculation as above. So it is enough to show how, given a constant subsequence $c_{dk+h} = \sum b_i(dk+h)\lambda_i^{dk+h}$, we can find $b_0, \dots, b_{L'}$ (not all zero) such that (12) holds. This is done by induction on $m = \max\deg(b_i)$. In each step we pick some j with $\deg(b_j) = m$, and replace $\{c_{dk+h}\}$ by $\{c_{d(k+1)+h} - \lambda_j^d c_{dk+h}\}$. This process decreases the degree of the coefficient of λ_j^d, without making the equation identically zero. ∎

Note: E. Lindenstrauss showed that given a degenerate recursion, one can always find a non-constant *integer*-valued sequence that vanishes along an arithmetic progression.

We can now solve the question alluded to in the beginning of this section:

THEOREM 5.2: *Let $\{c_k\}$ be an integer-valued recursion sequence with no constant arithmetic subsequences. Let $q > 1$ be an integer that is relatively prime to a_L, the free coefficient in the recursion polynomial. Then $\{c_k\}$ has almost-bounded cells* mod q. *In particular it is a q-Host sequence.*

Proof: Since we are going to utilize Theorem 3.2, it is enough to consider the case where $q = p$ is prime. Let λ_i be the characteristic roots of the recursion, and let $\mathbb{K}_p = \mathbb{Q}_p(\lambda_1, \dots, \lambda_L)$ denote a finite extension of $\mathbb{Q}_p$ containing all of them. The recursion polynomial is defined over $\mathbb{Z}_p$, the ring of integers of $\mathbb{Q}_p$, hence it decomposes over this ring to a product of irreducible factors. Since $|a_L|_p = 1$, the free coefficient of every such factor has also p-adic norm 1, and so all the roots of each factor have norm 1 in the extension of the p-adic norm to $\mathbb{K}_p$. We conclude that $|\lambda_i|_p = 1$ in $\mathbb{K}_p$, for $i = 1, \dots, L$.

We divide the rest of the proof into two cases.

FIRST CASE: $\mathbb{K} = \mathbb{Q}_p$. We will need the following properties of the $\log_p$ and $\exp_p$ functions in $\mathbb{Q}_p$:

1. $\log_p(1+x) = \sum_{n=1}^{\infty}[(-1)^{n+1}x^n/n]$ converges in the disk $\{|x|_p < 1\}$, and $|\log_p(1+x)|_p \leq |x|_p$.
2. $\exp_p(x) = \sum_{n=0}^{\infty}(x^n/n!)$ converges in the disk $\{|x|_p < p^{-1/(p-1)}\}$.
3. $\log_p$ and $\exp_p$ are local inverses of each other.

Write $c_k = \sum_{i=1}^{L'} b_i(k)\lambda_i^k$ as above. For $0 \leq h < p-1$ define

$$G_h(k) = \sum_i b_i(k)\lambda_i^h\left(\lambda_i^{p-1}\right)^k.$$

Pick some λ_i and write it in $\mathbb{Q}_p$ as $\lambda_i = \sum_{j=0}^{\infty} d_j p^j, 0 \leq d_j < p$. Since $|\lambda_i|_p = 1$, $d_0 \neq 0$, and so $d_0^{p-1} \equiv 1 \pmod{p}$, hence $|\lambda_i^{p-1} - 1|_p \leq p^{-1}$. It follows that $\log_p(\lambda_i^{p-1})$ is well-defined (cf. [10], p. 26). Now, assuming for a moment that $p > 2$, for $|x|_p \leq 1$ we get

$$|x\log_p(\lambda_i^{p-1})|_p \leq |x|_p \cdot |\lambda_i^{p-1} - 1|_p \leq p^{-1} < p^{-1/(p-1)},$$

and so $\exp_p(x\log_p(\lambda_i^{p-1}))$ is well-defined in the unit disk $B = \{x : |x|_p \leq 1\}$. The last expression will serve as the definition of $\left(\lambda_i^{p-1}\right)^x$. This definition agrees with the rational one for a rational integer $x = k$. We thus see that each function $G_h(k)$ can be extended to a function $G_h(x) = \sum_i b_i(x)\lambda_i^h\left(\lambda_i^{p-1}\right)^x$, analytic in B. By our assumptions on $\{c_k\}$, the function $G_h(k) = c_{(p-1)k+h}$ is non-constant, and hence all the conditions of Theorem 3.2 are satisfied. It follows that each of the $p-1$ subsequences $\{G_h(k)\}_k$ has almost-bounded cells. Looking back at Definition 2.1 we see that $\{c_k\}$ has this property as well.

The case $p = 2$ should be handled separately, since in this case $G_h(x)$ is defined only for $|x|_p \leq p^{-1}$. But we can just split $\{c_k\}$ into more subsequences, defining $G_{h,j}(x) = G_h(px+j)$, for $0 \leq j < p$. $G_{h,j}$ is defined in the whole unit disk, and so again we conclude that each of the finitely-many subsequences $\{G_{h,j}(k)\}_k$ has almost-bounded cells.

THE GENERAL CASE: $\mathbb{K} = \mathbb{Q}_p(\lambda_1, \ldots, \lambda_L)$. $\log_p$ and $\exp_p$ are defined as above and have the same properties; the only problem is that our observation that $|\lambda_i^{p-1} - 1|_p \leq p^{-1}$ is no longer necessarily true. However, we can use the following simple lemma:

LEMMA: *For some $m > 1$, $|\lambda_i^m - 1|_p \leq p^{-1}$ for $i = 1, \ldots, L$.*

Proof: By considering powers of λ_i and using compactness, for every i there exists some m_i with $|\lambda_i^{m_i} - 1|_p \leq p^{-1}$. Now take $m = m_1 \cdot \ldots \cdot m_L$, and note that $|\lambda_i^m - 1|_p \leq p^{-1}$ as well. ∎

(We remark that it is actually true that there exists some m such that $|\lambda^m - 1|_p \leq p^{-1}$ for *all* $\lambda \in \mathbb{K}$ with $|\lambda|_p = 1$.)

Using this lemma we can split $\{c_k\}$ to m subsequences (instead of $p-1$), and continue as before. This concludes the proof of Theorem 5.2. ∎

COROLLARY 5.3: *Let $q_1, \ldots, q_L$ be integers $\neq -1, 0, 1$, such that $|q_1|, \ldots, |q_L|$ are distinct and are relatively prime to some fixed integer $p > 1$. Let $f_1, \ldots, f_L$ be arbitrary non-zero polynomials. Then $c_k = f_1(k)q_1^k + \cdots + f_L(k)q_L^k$ has almost-bounded cells* (mod p)*, and so is a p-Host sequence.*

6. Uniform orbits in the d-dimensional torus

The general problem we wish to answer is the following: Let A, B be two epimorphisms of the d-torus. Given an A-invariant and ergodic Borel measure μ on $\mathbb{T}^d$ with positive entropy, when can we guarantee that for μ-almost every $\omega \in \mathbb{T}^d$, the orbit $\{B^n\omega\}$ is uniformly distributed in $\mathbb{T}^d$?

For simplicity, we will demonstrate the results in the 2-dimensional torus, $\mathbb{T}^2$. Our methods apply to the case where A is diagonal. We start with treating the case where A is a scalar matrix: $A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$. We refer the reader to [14] for a formulation and proof of the general d-dimensional case.

We begin with some notations and definitions. Let a, b be two relatively prime integers. Denote by $\pi_{a,b} : \mathbb{T}^2 \longrightarrow \mathbb{T}$ the projection map $(x, y) \mapsto ax + by$ (mod 1). Let $\pi_{a,b}\mu$ be the image of μ under the map $\pi_{a,b}$. Since μ is A-invariant, $\pi_{a,b}\mu$ is σ_p-invariant. We call the pair (a, b) an **entropy-decreasing direction** if $h(\pi_{a,b}\mu, \sigma_p) < h(\mu, A)$.

THEOREM 6.1: *Let $A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$, for p an integer. Suppose that μ is a Borel A-invariant measure on $\mathbb{T}^2$, ergodic and with positive entropy $h(\mu, A)$. Let $B = (b_{ij})$ be a matrix of integers. Then:*

(a) *If for a pair of integers d_1, d_2 there exists some entropy-decreasing direction (a, b) such that the sequence $c_k = (d_1, d_2)B^k\binom{b}{-a}$ has almost-sub-exponential collisions* mod p, *then*

$$\frac{1}{N}\sum_{n=0}^{N-1} e\left((d_1, d_2)B^k\begin{pmatrix} x \\ y \end{pmatrix}\right) \longrightarrow 0 \quad \mu\text{-a.e.}$$

(b) *If this is true for all* $(d_1, d_2) \neq (0,0)$, *then for* μ*-almost every* $(x,y) \in \mathbb{T}^2$, *the sequence* $\{B^n\binom{x}{y}\}$ *is uniformly distributed on* $\mathbb{T}^2$. *Consequently,* $\frac{1}{N}\sum_0^{N-1} B^n\mu \longrightarrow m_L$, *Lebesgue measure on* $\mathbb{T}^2$. *In particular, if* μ *is also* B*-invariant then* $\mu = m_L$.

For the proof of the theorem we will need the following lemma, which is an analogue of Lemma 3.3. An idea of B. Weiss simplified the original proof.

LEMMA 6.2: *Let* $A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}$, *for* p *an integer, and let* μ *denote an ergodic Borel* A*-invariant measure on* $\mathbb{T}^2$. *Define measures* $\omega_n = \sum_{j=0}^{p^n-1} \mu * \delta_{(jp^{-n},0)}$, *and let* $\varphi_n = d\mu/\,d\omega_n$ *be the Radon–Nikodym derivative. Then* $\varphi_n^{1/n} \longrightarrow e^{-h'}\mu$*-a.e., where* $h' = h(\mu, A) - h(\pi_{0,1}\mu, \sigma_p)$.

Proof: Denote by α and β the p-partitions corresponding to the X and Y axes, respectively:

$$\alpha = \left\{\left[\frac{j}{p}, \frac{j+1}{p}\right) \times \mathbb{T}\right\}_{j=0}^{p-1}, \qquad \beta = \left\{\mathbb{T} \times \left[\frac{j}{p}, \frac{j+1}{p}\right)\right\}_{j=0}^{p-1}.$$

As in Host's proof of Lemma 3.3, we denote by $\mathcal{B}_n$ the μ-completion of the σ-algebra of Borel sets in $\mathbb{T}^2$ invariant to $(x,y) \mapsto (x + p^{-n}, y)$, and so $\mathcal{B}_n = \alpha_n^\infty \vee \beta_0^\infty$. Applying a similar argument to the one in [5], one sees that $\varphi_n(x,y) = \mathbb{P}_\mu(J \,|\, \mathcal{B}_n)(x,y)$, if $J = [\frac{j}{p^n}, \frac{j+1}{p^n}) \times \mathbb{T}$ is the atom of α_0^{n-1} containing (x,y). (This follows from observing that

$$E_\mu(f \,|\, \mathcal{B}_n)(x,y) = \sum_{j=0}^{p^n-1} f(x + jp^{-n}, y)\varphi_n(x + jp^{-n}, y),$$

and taking $f = 1_J$.) Thus $-\log\varphi_n = I_{\alpha_0^{n-1}|\alpha_n^\infty \vee \beta_0^\infty}$. Since $h(\pi_{0,1}\mu, \sigma_p) = H_\mu(\beta|\beta_1^\infty)$, the assertion is a result of the following general lemma:

LEMMA 6.3: *Let* α, β *denote two finite partitions in a measure-preserving space* $(X, \mathfrak{B}, \mu, A)$, *and suppose that* $\alpha \vee \beta$ *is a generator for* A. *Then*

$$\frac{1}{n} I_{\alpha_0^{n-1}|\alpha_n^\infty \vee \beta_0^\infty} \longrightarrow h(\mu, A) - H_\mu(\beta|\beta_1^\infty) \quad \mu\text{-a.e.}$$

Proof of Lemma 6.3: Conditional probability considerations yield

$$I_{(\alpha\vee\beta)_0^{n-1}|(\alpha\vee\beta)_n^\infty} = I_{\beta_0^{n-1}|(\alpha\vee\beta)_n^\infty} + I_{\alpha_0^{n-1}|\alpha_n^\infty\vee\beta_0^\infty}$$

and so

$$-\frac{1}{n}\log\varphi_n = \frac{1}{n} I_{\alpha_0^{n-1}|\alpha_n^\infty\vee\beta_0^\infty} = \frac{1}{n} I_{(\alpha\vee\beta)_0^{n-1}|(\alpha\vee\beta)_n^\infty} - \frac{1}{n} I_{\beta_0^{n-1}|(\alpha\vee\beta)_n^\infty}.$$

As in Lemma 3.3, the first right-hand term converges a.e. to $h(\mu, A)$, and so it remains to show that $\frac{1}{n} I_{\beta_0^{n-1}|(\alpha\vee\beta)_n^\infty} \longrightarrow H_\mu(\beta|\beta_1^\infty)$.

Use a conditional probabilities argument to see that

$$\begin{aligned} I_{\beta_0^{n-1}|(\alpha\vee\beta)_n^\infty}(x,y) &= \sum_{k=0}^{n-1} I_{\beta_k|\beta_{k+1}^\infty \vee \alpha_n^\infty}(x,y) \\ &= \sum_{k=0}^{n-1} I_{\beta_0|\beta_1^\infty \vee \alpha_{n-k}^\infty}(A^k(x,y)) \\ &= \sum_{k=0}^{n-1} f_{n-k} \circ A^k(x,y) \end{aligned}$$

where $f_n = I_{\beta|\beta_1^\infty \vee \alpha_n^\infty}$.

We proceed as in Breiman's proof of the Shannon–McMillan–Breiman Theorem. From the Martingale Convergence Theorem we have $f_n \longrightarrow f = I_{\beta|\beta_1^\infty \vee \alpha_\infty}$ μ-a.e. and in L^1, where $\alpha_\infty = \bigcap \alpha_n^\infty$ is the tail σ-algebra. Applying Lemma 3.6 we get

$$\frac{1}{n} \sum_{k=0}^{n} f_{n-k} \circ A^k \longrightarrow \int f \, d\mu = H_\mu(\beta \mid \beta_1^\infty \vee \alpha_\infty) = H_\mu(\beta \mid \beta_1^\infty) \quad \mu\text{-a.e.},$$

as claimed (the last equality is part of Pinsker Lemma, see [15], Lemma 7, p. 65). This concludes the proof of the two lemmas. ∎

Remark: It might be interesting to note that the limit constant h' can also be identified as $H(\alpha \mid \alpha_1^\infty \vee \beta_{-\infty}^\infty)$ in the one-to-one extension of $(\mathbb{T}^2, \mu, A)$ to an invertible system. This is an easy consequence of Pinsker Formula (see for instance [15], Theorem 8, p. 66).

Proof of Theorem 6.1: Once we prove part (a), part (b) follows immediately. So pick integers c, d such that $ac + bd = 1$, and change coordinate using

$$U = \begin{pmatrix} d & -c \\ a & b \end{pmatrix},$$

i.e., take

$$\mu' = U\mu \quad \text{and} \quad \omega_n = \sum_{j=0}^{p^n-1} \mu' * \delta_{(jp^{-n},0)}.$$

Also, let $\varphi_n = d\mu' / d\omega_n$ be the Radon–Nikodym derivative. Since

$$h' = h(\mu, A) - h(\pi_{a,b}\mu, \sigma_p) = h(\mu', A) - h(\pi_{0,1}\mu', \sigma_p),$$

from Lemma 6.2 we see that

$$\varphi_n^{1/n} \longrightarrow e^{-h'} \quad \mu'\text{-a.e.}$$

Fix some integers $(d_1, d_2) \neq (0,0)$, and consider

$$g_N(x,y) = \frac{1}{N} \sum_{k=0}^{N-1} e\left((d_1,d_2)B^k \binom{x}{y}\right),$$

where $e(t) = e^{2\pi i t}$. We need to show that $g_N \longrightarrow 0$ μ-a.e. As in the proof of Theorem 3.1, we estimate the $d\omega_n$ integral, while changing coordinates:

$$\begin{aligned}
&\int \frac{|g_N|^2}{\varphi_n \circ U}\, d\mu \\
&= \int \frac{|g_N|^2 U^{-1}}{\varphi_n}\, d\mu' \\
&\leq \int |g_N|^2 U^{-1}\, d\omega_n \\
&= \sum_{j=0}^{p^n-1} \int \left| g_N\left(U^{-1}\binom{x+jp^{-n}}{y}\right)\right|^2 d\mu'(x,y) \\
&= \frac{1}{N^2} \sum_{k,l=0}^{N-1} \sum_{j=0}^{p^n-1} \int e\left((d_1,d_2)(B^k - B^l)U^{-1}\binom{x+jp^{-n}}{y}\right) d\mu'(x,y) \\
&\leq \frac{1}{N^2} \sum_{k,l=0}^{N-1} \sum_{j=0}^{p^n-1} e\left((d_1,d_2)(B^k - B^l)\begin{pmatrix} b & c \\ -a & d \end{pmatrix}\binom{jp^{-n}}{0}\right) \\
&= \frac{p^n}{N^2} \#\{0 \leq k,l < N\colon c_k \equiv c_l \pmod{p^n}\}
\end{aligned}$$

where $c_k \overset{\text{def}}{=} (d_1,d_2)B^k \begin{pmatrix} b & c \\ -a & d \end{pmatrix}\binom{1}{0} = (d_1,d_2)B^k\binom{b}{-a}$.

By our assumptions, $\{c_k\}$ has almost-sub-exponential collisions mod p. On the other hand, by Lemma 6.2, $(\varphi_n \circ U)^{1/n} \longrightarrow e^{-h'} < 1$ μ-a.e. We are now exactly in the same situation as in the proof of Theorem 3.1, and we conclude in the same manner that $g_N \longrightarrow 0$ μ-a.e. ∎

Note that part (a) implies that $\{(d_1,d_2)B^k\omega\}$ is uniformly distributed for μ-almost every $\omega \in \mathbb{T}^2$. We also have the following corollary:

THEOREM 6.4: *Let A, B and μ be as in Theorem 6.1. Assume that B has two eigenvalues $\lambda_1, \lambda_2 \neq 0$ such that*

(13) $\det B, p$ *are relatively prime, and* $\lambda_1, \lambda_2, \lambda_1/\lambda_2$ *are not roots of unity.*

Furthermore, assume that there exists some entropy-decreasing direction (a,b) *for which* $\binom{b}{-a}$ *is not an eigenvector of* B. *Then all the consequences in* (b) *of Theorem* 6.1 *hold.*

Proof: Fix some integers $(d_1,d_2) \neq (0,0)$, and let $c_k = (d_1,d_2)B^k\binom{b}{-a}$.

$\{c_k\}$ satisfies a recursion scheme with characteristic roots λ_1,λ_2. By Theorem 5.2 we can conclude that $\{c_k\}$ has almost-bounded cells mod p, once we show that it has no constant arithmetic subsequences. We can then apply Theorem 6.1.

Since the scheme is non-degenerate (using the condition on λ_1,λ_2 and Proposition 5.1), if $\{c_k\}$ had a constant arithmetic subsequence, $\{c_k\}$ would have to be identically zero. Write $\binom{b}{-a} = a_1u_1+a_2u_2$ where u_1,u_2 are the eigenvectors of B corresponding to λ_1,λ_2. We have $c_k = (d_1,d_2)\cdot(a_1\lambda_1^k u_1 + a_2\lambda_2^k u_2)$. Assuming that $c_k = 0$ for all k, we must have $\dim \operatorname{span}_k\{a_1\lambda_1^k u_1 + a_2\lambda_2^k u_2\} \leq 1$, hence $\{a_1u_1, a_2u_2\}$ are linearly dependent. But then a_1 or a_2 are zero, and so $\binom{b}{-a}$ is proportional to some u_i, contradicting the last assumption of the theorem. ∎

We add one more corollary:

COROLLARY 6.5: *For* A,μ *as in Theorem* 6.4, *if in addition* $h(\mu,A) > \log p$, *then for any matrix* B *satisfying* (13) *all the consequences in* (b) *of Theorem* 6.1 *hold.*

Proof: Take some $\binom{b}{-a}$ which is not an eigenvector of B. As

$$h(\pi_{a,b}\mu,\sigma_p) \leq \log p < h(\mu,A),$$

the direction (a,b) is entropy-decreasing. Now apply Theorem 6.4. ∎

Examples: 1. *Product measures.* Let $\mu = \mu_1 \times \mu_2$, with μ_1,μ_2 some p-invariant measures with entropies h_1,h_2, respectively. The total entropy of μ is $h = h_1+h_2$, and we assume $h > 0$.

If $h_1 > 0$, take $(a,b) = (0,1)$. As $h(\pi_{0,1}\mu,\sigma_p) = h_2 < h$, this direction is indeed entropy-decreasing. $(1,0)$ is an eigenvector of B iff B is upper-triangular. If this is the case, we surely can't guarantee in general uniformly distributed orbits; for instance, if we take $\mu_2 = \delta_0$, an upper-triangular B would leave the X-axis invariant, insuring that μ-almost every point does *not* have a uniform orbit. (It is easy to see in this example that $(0,1)$ is the only entropy-decreasing direction.) On the other hand, if B is not upper-triangular, we are in a position to apply the last theorem.

The situation is similar if $h_2 > 0$: in this case B must not be lower-triangular, and then $(a,b) = (1,0)$ will do.

If both $h_1, h_2 > 0$, we claim that any B satisfying (13) will do. Indeed, if B is not diagonal, one of the two arguments above will work. But choosing $(a,b) = (0,1)$ or $(a,b) = (1,0)$ will also work If B is diagonal, although both are eigenvectors: use Theorem 6.1, taking $(a,b) = (0,1)$ if $d_1 \neq 0$ and $(a,b) = (1,0)$ otherwise.

2. *Diagonal measures.* Let μ be a σ_p-invariant measure on the diagonal $\Delta = \{(x,x) \colon x \in \mathbb{T}\}$. Given a direction (a,b),

$$h(\pi_{a,b}\mu, \sigma_p) = h(\sigma_{a+b}\mu, \sigma_p) = \begin{cases} 0, & a = -b, \\ h(\mu, A), & a \neq -b. \end{cases}$$

Thus we are forced to take $(a,b) = (1,-1)$. To use Theorem 6.4, $(1,1)$ must not be an eigenvector of B. Again, this is a natural condition, since otherwise the diagonal is invariant under the action of B, and no orbit will have a uniform distribution.

3. *Measures on Cantor sets.* The rest of this section is devoted to pointing out a connection between entropy-decreasing directions and a similar phenomenon in Hausdorff dimension of Cantor sets.

To build a Cantor set in $\mathbb{T}^d$, pick an integer p (*the base*), and a *set of allowed digits* $\Lambda \subset \{0, \dots, p-1\}^d$. The set

$$K_\Lambda = \{\omega \in \mathbb{T}^d : \omega = \sum_{i=1}^{\infty} \mathrm{v}_i p^{-i}, \forall i \; \mathrm{v}_i \in \Lambda\}$$

is called a *simple p-invariant set in* $\mathbb{T}^d$. It is well known that its Hausdorff dimension is $\dim_H K_\Lambda = \log r / \log p$, where $r = |\Lambda|$. Let μ denote the natural measure on K_Λ, i.e. the one induced from the product measure on $\prod_1^\infty \Lambda$ which assigns equal probability to each element of Λ. The measure μ is A-invariant and ergodic, where $A = \mathrm{diag}(p, \dots, p)$. Also, μ is supported on K_Λ, i.e., $\mu(\mathbb{T}^2 \setminus K_\Lambda) = 0$. A theorem of Furstenberg ([4], prop. 3.1; for an indirect but shorter proof see [9], prop. 2.2) asserts that

$$\frac{h(\mu, A)}{\log p} = \frac{\log r}{\log p} = \dim_H K_\Lambda = \frac{h_{\mathrm{top}}(K_\Lambda, A)}{\log p}.$$

Thus μ has maximal-entropy among the measures supported on K_Λ.

We will suppose from now on that $d = 2$, and B is a matrix of integers. We wish to use Theorem 6.4 to find sufficient conditions under which $\{B^n\binom{x}{y}\}$ is uniformly distributed for μ-almost every $\binom{x}{y}$. We first need to find entropy-decreasing directions:

PROPOSITION 6.6: *If a direction (a,b) decreases Hausdorff dimension, i.e., $\dim_H(\pi_{a,b}K_\Lambda) < \dim_H K_\Lambda$, then it is entropy-decreasing for the natural measure μ on K_Λ.*

Proof: Suppose $\dim_H K' < \dim_H K_\Lambda$, where $K' = \pi_{a,b}K_\Lambda$. Let $\nu = \pi_{a,b}\mu$. Then

$$h(\nu,\sigma_p) \le h_{\text{top}}(K',\sigma_p) < h_{\text{top}}(K_\Lambda, A) = h(\mu, A),$$

hence (a,b) is entropy-decreasing. ∎

We remark that if $\dim_H K_\Lambda > 1$ every direction is entropy-decreasing, so the interesting case is when $\dim_H K_\Lambda \le 1$.

Typically, there will be an abundance of entropy- (or dimension-) decreasing directions. As an example, let us consider the case $p = 3$ and

$$\Lambda = \{(0,1),(1,0),(1,1)\}.$$

It is easy to see that $(1,0)$, $(0,1)$ and $(1,1)$ are all entropy-decreasing; indeed here $h(\mu, A) = \log 3$, while $h(\pi_{a,b}\mu,\sigma_p) = H(\frac{1}{3},\frac{2}{3}) < \log 3$, for all three cases. Kenyon [8] analyzes this particular case in depth, and proves that $\dim_H(\pi_{a,b}K_\Lambda) < 1 \iff 3 \nmid a+b$. (Kenyon also calculates the dimension of projections in irrational directions.) In general,

$$\pi_{a,b}K_\Lambda = \{\sum a_i p^{-i} : \forall i\ a_i \in \Lambda'\},$$

where Λ' is the image of the map $\Lambda \longrightarrow \mathbb{Z}$ defined by $(x,y) \mapsto ax+by$ (cf. [9]). If this map is not one-to-one, then clearly

$$\dim_H(\pi_{a,b}K_\Lambda) \le \frac{\log|\Lambda'|}{\log p} < \frac{\log|\Lambda|}{\log p} = \dim_H(K_\Lambda).$$

Returning to the last example, if B satisfies (13), the three vectors $\binom{-1}{0}, \binom{0}{1}, \binom{-1}{1}$ cannot all be eigenvectors (remembering the two eigenvalues of B are distinct), and so the conditions of the theorem are fulfilled.

We summarize this discussion with the following theorem.

THEOREM 6.7: *Let K_Λ be a simple p-invariant set in $\mathbb{T}^2$, for some integer $p > 1$. Let B denote a matrix satisfying (13). Suppose that the points of $\Lambda \subset \mathbb{Z}^2$ are not all on one line. Then $\{B^n\binom{x}{y}\}$ is uniformly distributed for μ-a.e. $\binom{x}{y}$, where μ is the natural measure on K_Λ.*

Proof: $|\Lambda| \geq 3$ from the assumptions. Draw in $\mathbb{Z}^2$ all lines passing through two or more points of Λ. Denote them by $l_\imath = \{(x,y) : a_\imath x + b_\imath y = c_\imath\}$, when each $a_\imath, b_\imath$ is a pair of relatively prime integers. From our previous remarks, each direction $(a_\imath, b_\imath)$ is entropy-decreasing. Among all lines, find three directions u_1, u_2, u_3 such that u_1, u_2 are linearly independent. Let v_1, v_2, v_3 be the orthogonal vectors. If all three were eigenvectors of B, we would have $v_3 \in \operatorname{span}\{v_1, v_2\}$, hence the eigenvalues of B would satisfy $\lambda_1 = \lambda_2$, a contradiction. Now use Theorem 6.4. ∎

7. Multi-invariant sets in $\mathbb{T}^d$

H. Furstenberg [4] proved that whenever $G \subset \mathbb{N}$ is a multiplicative semigroup which is non-lacunary (i.e., not all elements of G are powers of one integer), then for every irrational α, $\{G\alpha\}$ is dense in $\mathbb{T}$. An equivalent statement is that if a pair of integers p, q are multiplicatively independent (i.e., $\log p/\log q \notin \mathbb{Q}$), then for every irrational α, $\{p^n q^m \alpha\}_{n,m\geq 1}$ is dense mod 1. Furstenberg conjectured that under these conditions, $\{(p^n + q^m)\alpha\}_{n,m\geq 1}$ is dense as well. While this conjecture is still open, there was some advancement in related questions. Recently, B. Kra has proved the following:

THEOREM 7.1 ([11]): *For $i = 1, 2$, let $p_\imath, q_\imath$ be two multiplicatively independent integers whose absolute value is bigger than 1. Assume that $p_1 \neq p_2$ or $q_1 \neq q_2$. Then for every pair of irrational numbers α, β, the set $\{p_1^n q_1^m \alpha + p_2^n q_2^m \beta\}$ is dense* mod 1.

We wish to present a short proof of a somewhat weaker version of this fact using the ideas from the last section. This proof also employs ideas of Kra and Furstenberg.

We can assume without loss of generality that $p_1 > p_2$, but we will also need to assume that p_1 and q_1 are relatively prime, so that $\{q_1^n\}$ has almost-bounded cells mod p_1.

We first need the following version of Theorem 6.1:

PROPOSITION 7.2: *Let $A = \begin{pmatrix} p_1 & 0 \\ 0 & p_2 \end{pmatrix}, B = \begin{pmatrix} q_1 & 0 \\ 0 & q_2 \end{pmatrix}$ for some integers p_1, p_2, q_1, q_2, and let μ be an A-invariant and ergodic Borel measure on $\mathbb{T}^2$. Assume that $(0,1)$ is an entropy-decreasing direction for μ, and that $\{q_1^n\}$ has almost-bounded cells* mod p_1. *Then for all $d_1, d_2 \in \mathbb{Z}, d_1 \neq 0$ we have*

$$\frac{1}{N}\sum_{n=0}^{N-1} e\left((d_1, d_2)B^n \begin{pmatrix} x \\ y \end{pmatrix}\right) \longrightarrow 0 \quad \mu\text{-a.e.} \tag{14}$$

Similarly, if $(1,0)$ *is entropy-decreasing and* $\{q_2^n\}$ *has almost-bounded cells* mod p_2, *equation* (14) *holds for any* d_1 *and any* $d_2 \neq 0$.

The proof is completely analogous to the proof of Theorem 6.1, and we will not repeat it. The main difference is that $\pi_{a,b}\mu$ is not (in general) invariant. But if we restrict the allowed directions (a,b) to $(0,1)$ or $(1,0)$ only, it is. In this case the argument of the theorem works without a change; for $(a,b)=(0,1)$, the sequence $\{c_k\}$ defined in the statement of Theorem 6.1 is $c_k = (d_1,d_2)B^k\binom{1}{0} = d_1 q_1^k$.

We will denote by λ_d Lebesgue measure on $\mathbb{T}^d$.

Proof of Theorem 7.1 when p_i, q_i *are relatively prime:* Let S be the closure in $\mathbb{T}^2$ of $\{(p_1^n q_1^m \alpha, p_2^n q_2^m \beta)\}$. By Furstenberg's theorem, $\pi_X S = \pi_Y S = \mathbb{T}$, where π_X, π_Y denote projections on the axes. We claim that there is a Borel measure μ on $\mathbb{T}^2$, supported on S, which is A, B-invariant and $\pi_X \mu = \lambda_1$. To see this consider

$$M = \{\nu : \nu \text{ is a Borel probability measure, } A\nu = \nu,\ \nu(\mathbb{T}^2 \setminus S) = 0,\ \pi_X \nu = \lambda_1\}.$$

M is non-empty, since it contains the lifting of λ_1 from the X-axis to S. The set M is also convex, closed (in the weak* topology), and invariant under the action of B. Hence there is a fixed point μ for the B-action, which will be the desired measure.

Denote by $\mu = \int \mu_\theta \, d\nu(\theta)$ the ergodic decomposition of $(\mathbb{T}^2, \mu, A)$. Each component μ_θ is A-invariant and ergodic, but is not necessarily B-invariant. However, $\pi_X \mu_\theta = \lambda_1$ for ν-almost every θ, by ergodicity. Also, $(0,1)$ is an entropy-decreasing direction for μ_θ, since

$$h(\pi_Y \mu_\theta, \sigma_{p_2}) \leq \log p_2 < \log p_1 = h(\lambda, \sigma_{p_1}) = h(\pi_X \mu_\theta, \sigma_{p_1}) \leq h(\mu_\theta, A).$$

By the last proposition we get for every integer d_1, d_2 with $d_1 \neq 0$:

$$g_N = \frac{1}{N} \sum_{n=0}^{N-1} e\left((d_1,d_2)B^n \begin{pmatrix} x \\ y \end{pmatrix}\right) \longrightarrow 0 \quad \mu_\theta\text{-a.e.},$$

for almost every θ. Using the B-invariance of μ we get

$$\hat{\mu}(-d_1, -d_2) = \int g_N \, d\mu = \iint g_N \, d\mu_\theta \, d\nu(\theta) \longrightarrow 0.$$

In particular, for all $m \neq 0$, $(\pi_{1,1}\mu)\hat{}(m) = \hat{\mu}(m,m) = 0$. We conclude that $\pi_{1,1}\mu = \lambda_1$. But μ is supported on S, and so $\lambda_1(\pi_{1,1}S) = 1$, hence $\pi_{1,1}S = \mathbb{T}$. Recalling the definition of S, the claim of the theorem is proved. ∎

More information on bi-invariant sets and measures is contained in the following theorem. Recall that the Hausdorff dimension of a measure μ is defined by

$$\dim \mu = \inf\{\dim_H S \colon \mu(S) = 1\}.$$

THEOREM 7.3 ([14]): *Let $A = \operatorname{diag}(a_1, \dots, a_d)$ and $B = \operatorname{diag}(b_1, \dots, b_d)$ be two diagonal endomorphisms of $\mathbb{T}^d$. Suppose that*

$$(15) \qquad |a_i| > 1, \quad |b_i| > 1 \quad \text{and} \quad \gcd(a_i, b_i) = 1 \quad \text{for } i = 1, \dots, d.$$

(i) *If μ is a Borel measure on $\mathbb{T}^d$ which is invariant for both A and B, then μ has integer Hausdorff dimension. Moreover, if μ is ergodic for the action of the semigroup $\langle A, B\rangle$ and $\dim \mu = r$, then there exists a projection $P : \mathbb{T}^d \longrightarrow \mathbb{T}^r$ such that $\dim P\mu = \dim \mu$ and $P\mu$ is Lebesgue measure on $\mathbb{T}^r$.*

(ii) *Every closed set $S \subset \mathbb{T}^d$ such that $AS \subset S$ and $BS \subset S$ has an integer Hausdorff dimension. If $\dim_H S = 0$ then S is a finite set, and if $\dim_H S = d$ then $S = \mathbb{T}^d$.*

If A is conformal, i.e., $a_1 = \cdots = a_d$, then the assumptions on B can be relaxed: it suffices that B has integer eigenvalues $b_1, \dots, b_d$ such that (15) holds.

[14] also contains a pointwise version of this theorem.

8. Concluding remarks and questions

In this section we list some problems we were not able to solve.

PROBLEM 1. In Proposition 3.4 we saw that $\mathcal{H}_p$, the class of p-Host sequences, is not closed under sums, but is closed to translations, as well as addition or multiplication by a scalar. Is $\mathcal{H}_p$ closed under multiplication?

If two sequences are restrictions to $\mathbb{N}$ of p-adic analytic functions, then their product is obviously in $\mathcal{H}_p$, using this theorem. However, Example 3.5 shows that there are p-Host sequences that lack any of the combinatorial properties of Definition 2.1. In particular, such a sequence cannot be the restriction of a continuously differentiable p-adic function with a finite number of critical points.

PROBLEM 2. Is there a sequence $\{c_k\}$ such that $\frac{1}{N}\sum_{k=0}^{N-1} \sigma_{c_k}\mu \longrightarrow \lambda$ for every σ_p-invariant and ergodic μ with positive entropy, yet $\{c_k\}$ is not a p-Host sequence?

If Host's theorem is false for multiplicatively independent p, q, then a sequence like $c_k = 6^k$ might be an example, for $p = 2$. In this case, the weak property

$\frac{1}{N}\sum \sigma_6^k\mu \longrightarrow \lambda$ follows already from Rudolph's theorem, since any limit measure is both σ_2- and σ_3-invariant.

PROBLEM 3. Is it possible to prove Host's theorem assuming only that p, q are multiplicatively independent? Johnson and Rudolph [7] proved the following weaker result: for every p-invariant and ergodic measure μ with positive entropy,

$$\frac{1}{N}\sum_{k=0}^{N-1} \sigma_q^k\mu \longrightarrow \lambda. \tag{16}$$

The stronger pointwise result follows from Host's argument in case $h = h(\mu, \sigma_p) \geq C$, where $0 < C < 1$ is a constant depending on p, q. To see that, assume first that some prime factor of p does not divide q. Then the collision number of $\{q^k\}$ is of magnitude $p^{n(1+\varepsilon)}$ for some $0 < \varepsilon < 1$. If $h/\log p > \varepsilon$, the proof of Theorem 3.1 will work just the same.

If all prime factors of p divide q, we can proceed in the following way: it is enough to prove (16) when q is replaced by some fixed power q^l (the sequence in (16) simply breaks to l similar subsequences with q replaced by q^l). Also, if $p|q$, we can replace q by q/p. By replacing q by a suitable expression of the form q^l/p^n we can return to the previous case.

It is also possible to derive from here (16) assuming only positive entropy, for a large class of sequences $\{c_n\}$ replacing $\{q^n\}$; this is carried out in Lindenstrauss, Meiri and Peres [12].

PROBLEM 4. As we remarked in the introduction, D. Rudolph noticed that from (16) it follows that for every such μ there exists a subsequence n_k such that $\sigma_q^{n_k}\mu \longrightarrow \lambda$.

Is it true that $\sigma_q^n\mu \longrightarrow \lambda$ as well, at least for relatively prime p and q? Rudolph's proof implies that the convergence holds outside a set with zero upper-density. Lyons [13] showed that $\sigma_q^n\mu \longrightarrow \lambda$ when p and q are relatively prime, and μ is either a Riesz product or a Bernoulli measure.

ACKNOWLEDGEMENT: This work is part of the author's Ph.D. thesis. I am indebted to Hillel Furstenberg for his guidance and constant encouragement, and in particular for pointing out the connections between p-adic analysis and combinatorial properties. I wish to thank Yuval Peres for many helpful discussions and suggestions that improved the overall presentation, and in particular for simplifying the proof of the second part of Theorem 3.2. I also wish to thank Benjamin Weiss for simplifying the proof of Lemma 6.3.

References

[1] D. Berend, *Multi-invariant sets on tori*, Transactions of the American Mathematical Society **280** (1983), 509-532.

[2] M. D. Boshernitzan, *Elementary proof of Furstenberg's diophantine result*, Proceedings of the American Mathematical Society **122** (1994), 67–70.

[3] J. Feldman, *A generalization of a result of Lyons about measures in* [0, 1], Israel Journal of Mathematics **81** (1993), 281–287.

[4] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Mathematical Systems Theory **1** (1967), 1–49.

[5] B. Host, *Nombres normaux, entropie, translations*, Israel Journal of Mathematics **91** (1995), 419–428.

[6] A. Johnson, *Measures on the circle invariant under multiplication by a nonlacunary subgroup of the integers*, Israel Journal of Mathematics **77** (1992), 211–240.

[7] A. Johnson and D. Rudolph, *Convergence under $\times_p$ of $\times_q$ invariant measures on the circle*, Advances in Mathematics **115** (1995), 117–140.

[8] R. Kenyon, *Projecting the one-dimensional Sierpinski gasket*, Israel Journal of Mathematics **97** (1997), 221–238.

[9] R. Kenyon and Y. Peres, *Intersecting random translates of invariant Cantor sets*, Inventiones Mathematicae **104** (1991), 601–629.

[10] N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta-functions*, second edition, Springer-Verlag, Berlin, 1977.

[11] B. Kra, *A generalization of Furstenberg's Diophantine theorem*, Proceedings of the American Mathematical Society, to appear.

[12] E. Lindenstrauss, D. Meiri and Y. Peres, *Entropy of convolutions on the circle*, preprint (1997).

[13] R. Lyons, *On measures simultaneously 2- and 3-invariant*, Israel Journal of Mathematics **61** (1988), 219–224.

[14] D. Meiri and Y. Peres, *Bi-invariant sets and measures have integer Hausdorff dimension*, Ergodic Theory and Dynamical Systems, to appear.

[15] W. Parry, *Topics in Ergodic Theory*, Cambridge University Press, 1981.

[16] K. Petersen, *Ergodic Theory*, Cambridge Studies in Advanced Mathematics **2**, Cambridge University Press, 1983.

[17] D. J. Rudolph, *×2 and ×3 invariant measures and entropy*, Ergodic Theory and Dynamical Systems **10** (1990), 395–406.

[18] W. H. Schikhof, *Ultrametric Calculus*, Cambridge Studies in Advanced Mathematics **4**, Cambridge University Press, 1984.

[19] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag, Berlin, 1982.