Ergod. Th. & Dynam. Sys. (First published online 2020), page 1 of 11*

doi:10.1017/etds.2019.117
*Provisional—final page numbers to be inserted when paper edition is published

Bernoulli decomposition and arithmetical independence between sequences

HAN YU

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, CB3 0WB, UK
(e-mail: hy351@maths.cam.ac.uk)

(*Received* 11 *April* 2019 *and accepted in revised form* 4 *December* 2019)

Abstract. In this paper, we study the set

$$A = \{p(n) + 2^n d \bmod 1 : n \geq 1\} \subset [0, 1],$$

where p is a polynomial with at least one irrational coefficient on non-constant terms, d is any real number and, for $a \in [0, \infty)$, $a \bmod 1$ is the fractional part of a. With the help of a method recently introduced by Wu, we show that the closure of A must have full Hausdorff dimension.

Key words: independence of sequences, Bernoulli decomposition, disjointness between dynamical systems
2010 Mathematics Subject Classification: 11J71 (Primary); 28D20, 28A80 (Secondary)

1. *Introduction and background*

In this paper, we follow a Bernoulli decomposition method developed in [**W16**]. This method combines Sinai's factor theorem with some properties of Bernoulli shifts and solves a dimension version of Furstenberg's intersection problem. Here, we will consider a very different number-theoretic problem with a similar method. Let α be an irrational number, and we know that the sequence (irrational rotation orbit) $\{n\alpha \bmod 1\}_{n\geq 1}$ equidistributes in $[0, 1]$. Let X_n, $n \geq 1$, be a sequence of independent and identically distributed real-valued random variables. For convenience, let X_1 be uniformly distributed in $[0, 1]$. In this setting, one can show that $\{n\alpha + X_n \bmod 1\}_{n\geq 1}$ equidistributes almost surely, and in particular its closure contains intervals. We now replace the random sequence X_n with a deterministic sequence $\{2^n d \bmod 1\}_{n\geq 1}$ by choosing an arbitrary real number d. On the one hand, if d is 'simple' enough, say, a rational number, then it is straightforward that $\overline{\{2^n d + n\alpha \bmod 1\}_{n\geq 1}}$ contains intervals. On the other hand, if d is 'random' enough, say, chosen randomly according to the Lebesgue measure, then by simple probabilistic arguments one can show that almost surely $\{2^n d + n\alpha \bmod 1\}_{n\geq 1}$ again equidistributes and its closure contains intervals. This consideration leads us to the following conjecture.

CONJECTURE 1.1. *Let α be an irrational number and d be a real number. Then the topological closure of the sequence $\{2^n d + n\alpha \bmod 1\}_{n\geq 1}$ contains intervals.*

In this paper, we prove the following partial result towards the above conjecture.

THEOREM 1.2. *Let α be an irrational number and d be a real number. Then the topological closure of the sequence $\{2^n d + n\alpha \bmod 1\}_{n\geq 1}$ has Hausdorff dimension 1.*

In fact, we will prove a stronger result, Theorem 1.4. Before we state this theorem, we provide some more background. Given two sequences $x = \{x_n\}_{n\geq 1}$, $y = \{y_n\}_{n\geq 1}$ in $[0, 1]$, it is often interesting to study their independence. In terms of sequences with dynamical background, this can be also understood as the disjointness between dynamical systems; see [**F67**] for more details. Intuitively, we want to say that two sequences x, y are independent if $\{(x_n, y_n)\}_{n\geq 1}$ is in some sense close to the product set $X \times Y$, where X, Y are the sets of numbers in the sequence x, y, respectively. We give a natural way of expressing this idea.

Definition 1.3. Let $x = \{x_n\}_{n\geq 1}$, $y = \{y_n\}_{n\geq 1}$ be two sequences in $[0, 1]$. We denote by X, Y the sets of numbers in the sequence x, y, respectively. Then we say that x and y are arithmetically independent if the set $H(x, y)$ of numbers in the sequence $\{x_n + y_n\}_{n\geq 1}$ attains the largest possible box dimension, namely,

$$\underline{\dim}_{\mathrm{B}} H(x, y) = \min\{1, \underline{\dim}_{\mathrm{B}} X + \underline{\dim}_{\mathrm{B}} Y\}.$$

As an easy example, we see that $\{n\alpha\}_{n\geq 1}$ and $\{n\beta\}_{n\geq 1}$ are arithmetically independent if $1, \alpha, \beta$ are linearly independent over the field $\mathbb{Q}$. It is also possible to study the independence between $\{n\alpha\}_{n\geq 1}$ and $\{n^2\beta\}_{n\geq 1}$ based on Weyl's equidistribution theorem. Then it is natural to ask about the independence between $\{n\alpha\}_{n\geq 1}$ and $\{2^n d\}_{n\geq 1}$, where d is any real number. For a polynomial p with degree k with real coefficients, we write $p(n) = \sum_{i=0}^{k} a_i n^i$. We say that p is irrational if at least one of the numbers $a_1, \ldots, a_k$ is irrational. In this paper, we show the following result. See §2.3 for a clarification of the notation that appears below.

THEOREM 1.4. *Let p be an irrational polynomial and let d be any real number. Then the sequences $\{p(n) \bmod 1\}_{n\geq 1}$ and $\{2^n d \bmod 1\}_{n\geq 1}$ are arithmetically independent. In fact, we have the stronger result*

$$\dim_{\mathrm{H}} \overline{\{p(n) + 2^n d \bmod 1\}_{n\geq 1}} = 1.$$

We note that there is a curious connection between sequences of form $\{p(n) + 2^n d \bmod 1\}_{n\geq 1}$ and $\alpha\beta$-sequences. Let α, β be two real numbers; an $\alpha\beta$-sequence $\{x_n\}_{n\geq 1}$ is such that $x_1 = 0$ and, for each $i \geq 1$, we can choose $x_{i+1} = x_i + \alpha \bmod 1$ or $x_{i+1} = x_i + \beta \bmod 1$ freely. We have the following problem.

CONJECTURE 1.5. *Let α, β be such that $1, \alpha, \beta$ are independent over the field of rational numbers. Then any $\alpha\beta$-sequence has full box dimension.*

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

This conjecture is related to affine embeddings between Cantor sets, symbolic dynamics and Diophantine approximation; see [**K79**, **FX18**, **Y18**]. A lot of ideas for proving Theorem 1.4 appeared in [**Y18**] for $\alpha\beta$-sets. For this reason, we can consider Theorem 1.4 as a cousin of Conjecture 1.5. Although the method in this paper cannot be used directly for $\alpha\beta$-sequences, it still sheds some light on Conjecture 1.5. However, at this stage, we mention that in [**K79**] there is a construction of an $\alpha\beta$-sequence whose closure does not have full Hausdorff dimension.

We also consider here a number-theoretic result which is closely related to what has been discussed. Let m be an odd number. We consider the ring $R[m]$ of residues modulo m. It is the finite set $\{0, \ldots, m-1\}$ together with integer multiplication and addition modulo m. In this setting, we can also consider the sequence $\{2^n + cn \bmod m\}_{n\geq 0}$, where c is an integer such that $\gcd(c, m) = 1$. On the one hand, the $+c \bmod m$ action on $R[m]$ can be seen as uniquely ergodic, which is analogous to the $+\alpha \bmod 1$ action on the unit interval with an irrational number α. On the other hand, $\{2^n \bmod m\}_{n\geq 0}$ is an orbit under the $\times 2 \bmod m$ action. An analogy of Theorem 1.4 would be that $\{2^n + cn \bmod m\}_{n\geq 0}$ is large in $R[m]$. We show the following result, which confirms this intuition. We remark that the method for proving the following result shares some strategies for proving Theorem 1.4.

THEOREM 1.6. *Let $m \geq 3$ be an odd number and c be such that* $\gcd(c, m) = 1$. *Let $D(m)$ be the number of residue classes visited by* $\{2^n + cn \bmod m\}_{n\geq 0}$. *Then $D(m) = m$. In other words, for each $r \in R[m]$, there is an integer n_r such that $2^{n_r} + cn_r \equiv r \bmod m$.*

The above result is a special case of Problem 6 in the third round of the 27th Brazilian Mathematical Olympiad; see [**27BMO**].

2. *Definitions and notation*

2.1. *Logarithm.* We make the convention that the log function has base 2.

2.2. *Dimensions.* We list here some basic definitions of dimensions mentioned in the introduction. For more details, see [**F05**, Chs. 2, 3] and [**M99**, Chs. 4, 5]. We shall use $N(F, r)$ for the minimal covering number of a set F in $\mathbb{R}^n$ with closed balls of side length $r > 0$.

2.2.1. *Hausdorff dimension.* Let $g : [0, 1) \to [0, \infty)$ be a continuous function such that $g(0) = 0$. Then, for all $\delta > 0$, we define the quantity

$$\mathcal{H}^g_\delta(F) = \inf\left\{\sum_{i=1}^{\infty} g(\operatorname{diam}(U_i)) : \bigcup_i U_i \supset F, \operatorname{diam}(U_i) < \delta\right\}.$$

The g-Hausdorff measure of F is

$$\mathcal{H}^g(F) = \lim_{\delta\to 0} \mathcal{H}^g_\delta(F).$$

When $g(x) = x^s$ we have that $\mathcal{H}^g = \mathcal{H}^s$ is the s-Hausdorff measure, and the Hausdorff dimension of F is

$$\dim_{\mathrm{H}} F = \inf\{s \geq 0 : \mathcal{H}^s(F) = 0\} = \sup\{s \geq 0 : \mathcal{H}^s(F) = \infty\}.$$

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

2.2.2. *Box dimensions.* The upper box dimension of a bounded set F is

$$\overline{\dim}_{\mathrm{B}} F = \limsup_{r\to 0}\left(-\frac{\log N(F,r)}{\log r}\right).$$

Similarly, the lower box dimension of F is

$$\underline{\dim}_{\mathrm{B}} F = \liminf_{r\to 0}\left(-\frac{\log N(F,r)}{\log r}\right).$$

If the limsup and liminf are equal, we call this value the box dimension of F and we denote it by $\dim_{\mathrm{B}} F$.

2.3. *The unconventional fractional part symbol.* For a real number α, it is conventional to use $\{\alpha\}$ for its fractional part. It is unfortunate that $\{\cdot\}$ is also used to denote a set or a sequence as well. For this reason we will use mod 1 for the fractional part. More precisely, for a real number x we write x mod 1 to denote the unique number a in $[0, 1)$ such that $a - x$ is an integer.

2.4. *Sets and sequences.* We write $\{x_n\}_{n\geq 1}$ for the sequence $x_1x_2x_3\ldots$. Sometimes it is convenient to use $\{x_n\}_{n\geq 1}$ to denote the set

$$\{x : \exists n \in \mathbb{N}, x = x_n\}.$$

Thus $\overline{\{x_n\}_{n\geq 1}}$ and $\underline{\dim}_{\mathrm{B}}\{x_n\}_{n\geq 1}$ should be understood in this way.

2.5. *Filtrations, atoms and entropy.* Let X be a set with σ-algebra $\mathcal{X}$. A filtration of σ-algebras is a sequence $\mathcal{F}_n \subset \mathcal{X}$, $n \geq 1$, such that

$$\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{X}.$$

Given a measurable map $S : X \to X$ and a finite measurable partition $\mathcal{A}$ of X, we denote by $S^{-n}\mathcal{A}$ the finite collection of sets

$$\{S^{-n}(A) : A \in \mathcal{A}\}$$

(notice that S might not be invertible). Then we write $\bigvee_{i=0}^{n-1} S^{-i}\mathcal{A}$ for the σ-algebra generated by $S^{-i}\mathcal{A}$, $i \in [0, n-1]$. An atom in $\bigvee_{i=0}^{n-1} S^{-i}\mathcal{A}$ is a set A that can be written as

$$A = \bigcap_i C_i$$

where, for each $i \in \{0, \ldots, n-1\}$, $C_i \in S^{-i}\mathcal{A}$. In this sense $\bigvee_{i=0}^{n-1} S^{-i}\mathcal{A}$ is generated by a finite partition $\mathcal{A}_{n-1}$ of X which is finer than $\mathcal{A}$. Let μ be a probability measure. Then we define the Shannon entropy of μ with respect to a finite partition $\mathcal{A}$ as

$$H(\mu, \mathcal{A}) = -\sum_{A\in\mathcal{A}} \mu(A)\log\mu(A).$$

We define the entropy of S as

$$h(S, \mu) = \lim_{n\to\infty} \frac{1}{n} H(\mu, \mathcal{A}_{n-1}),$$

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

where $\mathcal{A}$ is a partition such that $\bigvee_{i=1}^{\infty} S^{-i}\mathcal{A} = \mathcal{X}$. Here we have implicitly assumed that such a generating partition exists and used Sinai's entropy theorem; see [**PY98**, Lemma 8.8].

Let $\mathcal{Y} \subset \mathcal{X}$ be an S-invariant σ-algebra, that is, $S^{-1}(\mathcal{Y}) \subset \mathcal{Y}$. Let $n \geq 1$ be an integer. We define the conditional information function of $\mathcal{A}_n$ conditioned on $\mathcal{Y}$ as

$$I_{\mu,\mathcal{A}_n|\mathcal{Y}}(x) = -\log E_\mu[\mathbb{1}_{A_n(x)}|\mathcal{Y}](x).$$

Here, $A_n(x)$ is the atom of $\mathcal{A}_n$ which contains $x \in X$. Then we define the conditional Shannon entropy of $\mathcal{A}_n$ conditioned on $\mathcal{Y}$ as

$$H(\mu, \mathcal{A}_n|\mathcal{Y}) = \int I_{\mu,\mathcal{A}_n|\mathcal{Y}}(x)\, d\mu(x).$$

Finally, we define the conditional entropy of S conditioned on $\mathcal{Y}$ as

$$h(S|\mathcal{Y}, \mu) = \lim_{n\to\infty} \frac{1}{n} H(\mu, \mathcal{A}_{n-1}|\mathcal{Y}).$$

All the above quantities are well defined; see [**D11**, Chs. 1, 2] for more details.

2.6. *Factors.* A measurable dynamical system is in general denoted by $(X, \mathcal{X}, S, \mu)$, where X is a set with σ-algebra $\mathcal{X}$, a measure μ (in this paper, μ will be a probability measure) and a measurable map $S : X \to X$. If $\mathcal{X}$ is clear from the context we do not explicitly write it down. Given two dynamical systems $(X, \mathcal{X}, S, \mu)$, $(X_1, \mathcal{X}_1, S_1, \mu_1)$, a measurable map $f : X \to X_1$ is called a factorization map and $(X_1, \mathcal{X}_1, S_1, \mu_1)$ is called a factor of $(X, \mathcal{X}, S, \mu)$ if $\mu_1 = f\mu$ and $f \circ S(x) = S_1 \circ f(x)$ holds for μ-almost all $x \in X$.

Another way of viewing factors is via invariant sub-σ-algebras. Let $\mathcal{Y} \subset \mathcal{X}$ be a sub-σ-algebra which is invariant under the map S. Then $(X, \mathcal{Y}, S, \mu)$ can be seen as a factor of $(X, \mathcal{X}, S, \mu)$ via the identity map. We can take $\mathcal{Y} = f^{-1}(\mathcal{X}_1)$ in the previous paragraph. In this measure-theoretic sense, $(X_1, \mathcal{X}_1, S_1, \mu_1)$ and $(X, \mathcal{Y}, S, \mu)$ can be viewed as the same dynamical system.

2.7. *Bernoulli system.* Let Λ be a finite set of symbols and let $\Omega = \Lambda^{\mathbb{N}}$ be the space of one-sided infinite sequences over Λ. We define S to be the shift operator, namely, for $\omega = \omega_1\omega_2 \cdots \in \Omega$,

$$S(\omega) = \omega_2\omega_3 \ldots.$$

We take the σ-algebra on Ω generated by cylinder subsets. A cylinder subset $Z \subset \Omega$ is such that $Z = \prod_{i\in\mathbb{N}} Z_i$ and $Z_i = \Lambda$ for all but finitely many integers $i \in \mathbb{N}$. We construct a probability measure μ on Ω by giving a probability measure $\mu_\Lambda = \{p_\lambda\}_{\lambda\in\Lambda}$ on Λ and set $\mu = \mu_\Lambda^{\mathbb{N}}$. We require here that $p_\lambda \neq 0$ for all $\lambda \in \Lambda$. Then this system is weak-mixing and has entropy $h(S, \mu) = \sum_{\lambda\in\Lambda} -p_\lambda \log p_\lambda$. We call this system a Bernoulli system.

2.8. *Joinings.* Let $(X, \mathcal{X}, S, \mu)$ and $(Y, \mathcal{Y}, T, \nu)$ be two measurable dynamical systems. A joining between those two dynamical systems is an $S \times T$-invariant probability measure ρ on $X \times Y$ (with respect to the product σ-algebra $\sigma(\mathcal{X} \times \mathcal{Y})$) such that $\pi_X\rho = \mu$, $\pi_Y\rho = \nu$. The two systems $(X, \mathcal{X}, S, \mu)$ and $(Y, \mathcal{Y}, T, \nu)$ are *disjoint* if the only joining is the product measure $\mu \times \nu$. The follow example can be found in [**F67**, Theorem I.4].

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

Example 2.1. Let $(X, \mathcal{X}, S, \mu)$ be a measure-theoretically distal ergodic system with finite height. Let $(Y, \mathcal{Y}, T, \nu)$ be a weakly mixing system. Then $(X, \mathcal{X}, S, \mu)$ and $(Y, \mathcal{Y}, T, \nu)$ are disjoint.

A measure-theoretically distal ergodic system with finite height is obtained from a Kronecker system with finitely many ergodic group extensions. For example, irrational rotations on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with the Lebesgue measure are Kronecker systems. The transformation $(x, y) \in \mathbb{T}^2 \to (x + \alpha, x + y)$ on $\mathbb{T}^2$ with $\alpha \notin \mathbb{Q}$ is obtained from an irrational rotation with an ergodic group extension. In this paper, we will also consider the transformation $(x_1, \ldots, x_n) \in \mathbb{T}^n \to (x_1 + \alpha, x_2 + x_1, x_3 + x_2, \ldots, x_n + x_{n-1})$ on $\mathbb{T}^n$. The above are examples of measure-theoretically distal ergodic systems with finite height.

3. *A mathematical Olympiad problem*

We first provide a short proof of Theorem 1.6, which provides us with some motivation.

Proof of Theorem 1.6. Let $l = \operatorname{ord}(2, m)$ be the order of 2 in the multiplication group $(\mathbb{Z}/m\mathbb{Z})^*$. This is permitted because $\gcd(2, m) = 1$. For convenience, we consider $c = 1$ and note that other cases can be shown with the same method. Since $l = \operatorname{ord}(2, m)$ we consider the sequence

$$\{2^{nl} + nl \bmod m\}_{n \geq 0}.$$

We see that $2^{nl} \equiv 1 \bmod m$ for all $n \geq 0$. However, $H = \{nl \bmod m\}_{n \geq 0}$ is a subgroup of $\mathbb{Z}/m\mathbb{Z}$ of order $m/\gcd(l, m)$. For convenience we write $\Delta = \gcd(l, m)$. This Δ plays the same role of the entropy in the proof of Theorem 4.2 which leads to Theorem 1.4. If $\Delta = 1$ then $D(m) = m$ follows automatically. We consider the case where $\Delta > 1$. Now for each integer r we consider the sequence

$$\{2^{r+nl} + r + nl \bmod m\}.$$

This sequence forms a coset of H. More precisely, it is $2^r + r + H$. Now if $\{2^r + r \bmod \Delta\}_{r \geq 0}$ visited all residue classes modulo Δ, then $2^r + r + H$, $r \geq 0$, would visit all cosets of H in $\mathbb{Z}/m\mathbb{Z}$ and $\{2^n + n\}_{n \geq 1}$ would visit all residue classes modulo m. Since Δ is an odd number as well, we see that we have reduced the problem for m to the problem for Δ which is strictly smaller than m. We can iterate this reduction procedure. Since we are considering a positive integer set, either we eventually obtain $\Delta = 1$ or else we can consider further $\gcd(\Delta, \operatorname{ord}(2, \Delta)) < \Delta$. The latter cannot happen infinitely often. This concludes the proof. □

4. *A consequence of Sinai's factor theorem*

In this section we discuss a consequence of Sinai's factor theorem. As mentioned in the introduction, this section is strongly influenced by [**W16**, §6]. To some extent, the idea resembles the arguments in the previous section. We start this section by introducing the set-ups and making some standard considerations.

Let $(X, \mathcal{X}, S, \mu)$ be a measure-theoretically distal ergodic system with finite height. Here we assume that μ is a probability measure on the σ-algebra $\mathcal{X}$. Let $(Y, \mathcal{Y}, T, \nu)$ be an ergodic measurable dynamical system. Furthermore, we require that T admits a finite

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

generator, that is, a finite measurable partition $\mathcal{A}_0$ of Y such that $\bigvee_{i=0}^{\infty} T^{-i}\mathcal{A}_0$ is $\mathcal{Y}$. For convenience, we make the following definition.

Definition 4.1. Let (Y, T, ν), $\mathcal{A}_0$ be as given in above. Let $B \subset Y$. For each integer $n \geq 1$, we define $N_{\mathcal{A}_0,S,n}(B)$ to be the number of atoms in $\mathcal{A}_n$ intersecting B. Then we define the quantities

$$\overline{\dim}_{\mathcal{A}_0,S} B = \limsup_{n\to\infty} \frac{\log N_{\mathcal{A}_0,S,n}(B)}{n},$$
$$\underline{\dim}_{\mathcal{A}_0,S} B = \liminf_{n\to\infty} \frac{\log N_{\mathcal{A}_0,S,n}(B)}{n}.$$

For example, given $\lambda > 0$, if $Y \subset \mathbb{R}$ and $\operatorname{diam}(A_n(x)) = O(2^{-\lambda n})$ uniformly for all n, x then

$$N(B, 2^{-\lambda n}) = O(N_{\mathcal{A}_0,S,n}(B)).$$

In this case, if $\overline{\dim}_{\mathcal{A}_0,S} B = 0$ then $\overline{\dim}_{\mathrm{B}} B = 0$. The main goal of this section is to show the following result† which is a variant of Wu's ergodic theoretic result in [**W16**, §6].

THEOREM 4.2. *Let (X, S, μ), (Y, T, ν) be as stated above. Let ρ be a joining between those two systems. Then ρ admits a $\sigma(\mathcal{X} \times \mathcal{Y})$-measurable measure disintegration*

$$\rho = \int_{\Omega} \rho_\omega \, d\omega,$$

where $(\Omega, d\omega)$ is a probability space such that, for each $\epsilon > 0$, there is a set E with positive $d\omega$ measure and, for $\omega \in E$,

- *$\pi_X \rho_\omega = \mu$;*
- *there is a $\mathcal{Y}$-measurable set $B_\omega \subset Y$ such that $\overline{\dim}_{\mathcal{A}_0,S} B_\omega \leq \epsilon$ and $\rho_\omega(\pi_Y^{-1}(B_\omega)) > 0$.*

The proof of this theorem is divided into two parts.

4.1. *Step 1: the conditional Shannon–McMillan–Breiman theorem and a counting argument.*

LEMMA 4.3. *Let (Y, T, ν), $\mathcal{A}_0$ be as stated in the beginning of this section. Let $\mathcal{B}$ be a countably generated T-invariant sub-σ-algebra of $\mathcal{Y}$. Suppose that the conditional entropy $h(T|\mathcal{B}, \nu) = 0$. Then, for ν-almost every $y \in Y$ and all $\epsilon > 0$, there is a $\mathcal{Y}$-measurable set $B_{y,\epsilon}$ with $\underline{\dim}_{\mathcal{A}_0,S} B_{y,\epsilon} \leq \epsilon$. Moreover, for each $\epsilon > 0$, there is a $\mathcal{B}$-measurable set E with positive ν measure and $\nu_y^{\mathcal{B}}(B_{y,\epsilon}) > 0$ for $y \in E$.*

Proof. The conditional Shannon–McMillan–Breiman theorem (see [**D11**, Appendix B]) implies that, for ν-almost all $y \in Y$,

$$\lim_{n\to\infty} \frac{1}{n} I_{\nu,\mathcal{A}_n|\mathcal{B}}(y) = h(T|\mathcal{B}, \nu).$$

Let $\epsilon > 0$ be a small number. Let $k \geq 0$ be an integer, and we construct the set

$$B_k = \{y \in Y : \forall n \geq k,\ I_{\nu,A_n|\mathcal{B}}(y) \leq n(h(T|\mathcal{B}, \nu) + \epsilon)\}.$$

† Later on, we only use this result with X, Y as compact metric spaces with Borel σ-algebras and with $\dim_{\mathcal{A}_0,S}$ equivalent to the box counting dimension on Y.

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

Then we have $\nu(\bigcup_{k\geq 1} B_k) = 1$, and thus there is an integer $n_0 > 0$ such that B_{n_0} has positive ν measure. We can choose n_0 to be sufficiently large to ensure that $\nu(B_{n_0})$ is very close to 1. However, positivity here is enough for later use.

Suppose that $\nu = \int \nu_y^{\mathcal{B}}\, d\nu(y)$ is the measure disintegration of ν against the factor $\mathcal{B}$; see [**EW11**, Theorem 5.14] (system of conditional measures). Then we see that, for ν-almost every $y \in Y$,

$$E_\nu[\mathbb{1}_{A_n(y)}|\mathcal{B}](y) = \nu_y^{\mathcal{B}}(A_n(y)).$$

Thus we have

$$B_{n_0} = \{y \in Y : \forall n \geq n_0, \log \nu_y^{\mathcal{B}}(A_n(y)) \geq -n(h(T|\mathcal{B}, \nu) + \epsilon)\}.$$

Let A_n be an atom in $\mathcal{A}_n$ intersecting B_{n_0} with $n \geq n_0$. Then we see that, for ν-almost every $y \in A_n \cap B_{n_0}$, we have

$$\nu_y^{\mathcal{B}}(A_n) = \nu_y^{\mathcal{B}}(A_n(y)) \geq 2^{-n(h(T|\mathcal{B},\nu)+\epsilon)}.$$

Those ν-almost everywhere choices of y form a $\mathcal{B}$-measurable set. Thus, by omitting a $\mathcal{B}$-measurable set with zero ν measure we can assume that the above holds whenever $y \in A_n \cap B_{n_0}$.

Since $\mathcal{B}$ is countably generated, we see that the fibre $[y]_{\mathcal{B}} = \bigcap_{F\in\mathcal{B}, y\in F} F$ is well defined and $\mathcal{B}$ measurable. For ν-almost every $y \in Y$ the measure $\nu_y^{\mathcal{B}}$ is in fact a well-defined probability measure supported on $[y]_{\mathcal{B}}$, and this measure is determined by the atom $[y]$; see [**EW11**, Theorem 5.14(2)]. In what follows, we arbitrarily fix such a $y \in Y$. Suppose that A_n is an atom in $\mathcal{A}_n$ intersecting B_{n_0}. Then by the argument above, we see that if $A_n \cap [y]_{\mathcal{B}} \cap B_{n_0} \neq \emptyset$,

$$\nu_y^{\mathcal{B}}(A_n) \geq 2^{-n(h(T|\mathcal{B},\nu)+\epsilon)}.$$

This implies that the number of atoms in $\mathcal{A}_n$ intersecting $[y]_{\mathcal{B}} \cap B_{n_0}$ is at most

$$2^{n(h(T|\mathcal{B},\nu)+\epsilon)}.$$

We note that the above arguments hold for a set of ν-almost every $y \in Y$. Since we have $h(T|\mathcal{B}, \nu) = 0$, there is an integer $n_0 \geq 1$ such that, for ν-almost every $y \in Y$, all $n \geq n_0$,

$$N_{\mathcal{A}_0,T,n}(B_{n_0} \cap [y]_{\mathcal{B}}) \leq 2^{n\epsilon}.$$

Thus $\overline{\dim}_{\mathcal{A}_0,T} B_{n_0} \cap [y]_{\mathcal{B}} \leq \epsilon$. Moreover, we have $\nu(B_{n_0}) > 0$, therefore we see that there is a $\mathcal{B}$-measurable set E with positive ν measure such that, for $y \in E$,

$$\nu_y^{\mathcal{B}}(B_{n_0} \cap [y]_{\mathcal{B}}) > 0.$$

Note that $B_{n_0} \cap [y]_{\mathcal{B}}$ is $\mathcal{Y}$-measurable but not necessarily $\mathcal{B}$-measurable. This is the set $B_{y,\epsilon}$ as required. □

4.2. *Bernoulli factors: the Ornstein–Weiss unilateral Sinai factor theorem.* For step 2, we need to use the unilateral Sinai factor theorem which was proved in [**OW75**]. Let $h = h(T, \nu)$ be the dynamical entropy of (Y, T, ν). Suppose that $h > 0$. Then the unilateral Sinai factor theorem says that any Bernoulli system (Ω, S_B, ν_B) with entropy at most h is a factor of (Y, T, ν). In particular, we can find a Bernoulli system as a factor of (Y, T, ν) with entropy h.

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

4.3. *Step 2: Wu's ergodic theoretic result revisited.*

Proof of Theorem 4.2. First, suppose that $h = h(T, \nu) = 0$. In this case we will see that the trivial disintegration $\rho = \rho$ works. Indeed, we have $\pi_X \rho = \mu$, $\pi_Y \rho = \nu$ since ρ is a joining. As $h = 0$, we see by Lemma 4.3, with $\mathcal{B}$ being the trivial σ-algebra, that, for each $\epsilon > 0$, there is a Borel set B with positive ν measure such that

$$\overline{\dim}_{\mathcal{A}_0,T} B \leq \epsilon.$$

Then we see that $\rho(\pi_Y^{-1}(B)) = \nu(B) > 0$. This finishes the proof in the case where $h = 0$.

Now suppose that $h > 0$. In this case, let (Ω, S_B, μ_B) be a Bernoulli factor of (Y, T, ν) with entropy h. This Bernoulli factor can be viewed as a T-invariant sub-σ-algebra $\mathcal{B}$ in view of §2.6. This σ-algebra $\mathcal{B}$ is countably generated. Then we see that $\mathcal{C} = \pi_Y^{-1}(\mathcal{B})$ is an $S \times T$-invariant sub-σ-algebra. Then we have the system of conditional measures $\rho^{\mathcal{C}}_{(x,y)}$ which are probability measures for ρ-almost every $(x, y) \in X \times Y$. Essentially, $\rho^{\mathcal{C}}_{(x,y)}$ does not depend on the choice of x. More precisely, we see that $[(x, y)]_{\mathcal{C}} = X \times [y]_{\mathcal{B}}$.

By construction, $\pi_Y(\rho^{\mathcal{C}}_{(x,y)}) = \nu^{\mathcal{B}}_y$ for ρ-almost every (x, y), or equivalently for ν-almost every $y \in Y$. Since $\mathcal{B}$ is obtained via a Bernoulli factor with entropy h, we see that $h(T|\mathcal{B}, \nu) = 0$ (Abramov–Rokhlin formula [**D11**, Fact 4.1.6]). Then, for ν-almost every $y \in Y$ and all $\epsilon > 0$, we see from Lemma 4.3 that there is a $\mathcal{Y}$-measurable set $B_{y,\epsilon}$ (which could be empty) with

$$\overline{\dim}_{\mathcal{A}_0,T} B_{y,\epsilon} \leq \epsilon.$$

Moreover, for each $\epsilon > 0$, for a $\mathcal{B}$-measurable set E with positive ν measure, we have

$$\nu^{\mathcal{B}}_y(B_{y,\epsilon}) > 0$$

whenever $y \in E$.

Let us take a measure $\rho^{\mathcal{C}}_{(x,y)}$ by taking a point (x, y) (where $\rho^{\mathcal{C}}_{(x,y)}$ is defined as a probability measure) such that $y \in E$ and

$$\rho^{\mathcal{C}}_{(x,y)}(\pi_Y^{-1}(B_{y,\epsilon})) = \nu^{\mathcal{B}}_y(B_{y,\epsilon}) > 0.$$

Such choices of (x, y) form a $\mathcal{C}$-measurable set E' with positive ρ measure. In order to finish the proof, we need to show that $\pi_X \rho^{\mathcal{C}}_{(x,y)} = \mu$. To check this, let f be a continuous function from X to $\mathbb{R}$. Then we see that by possibly dropping a $\mathcal{C}$-measurable ρ-null subset from E',

$$\int f(x')\, d\pi_X \rho^{\mathcal{C}}_{(x,y)}(x') = \int f(x')\, d\rho^{\mathcal{C}}_{(x,y)}(x', y') = E_\rho[f|\mathcal{C}](x, y)$$

for $(x, y) \in E'$. Observe that ρ is $S \times T$-invariant. By construction, $(Y, \mathcal{B}, T, \nu)$ is in fact a Bernoulli system. Observe that ρ is also a joining between (X, S, μ) and $(Y, \mathcal{B}, T, \nu)$. As Bernoulli system is weakly mixing, by Example 2.1, we see that ρ must be equal to $\mu \times \nu$ viewed as a probability measure on the product σ-algebra $\sigma(\mathcal{X} \times \mathcal{B})$. Since $\mathcal{C} = \pi_Y^{-1}(\mathcal{B})$ and f is a function on X, we see that, for $(x, y) \in E'$,

$$E_\rho[f|\mathcal{C}](x, y) = \int f\, d\mu.$$

As the above holds for all continuous functions on X, we see that $\pi_X \rho^{\mathcal{C}}_{(x,y)} = \mu$ for $(x, y) \in E'$. In other words, we have shown that $\rho = \int \rho^{\mathcal{C}}_{(x,y)}\, d\rho(x, y)$ is a measure disintegration satisfying the statements of this theorem. □

https://doi.org/10.1017/etds.2019.117 Published online by Cambridge University Press

5. *On sequences* $\{p(n)+2^n d \bmod 1\}_{n\geq 1}$

We now prove Theorem 1.4.

Proof of Theorem 1.4. First, let $\alpha \in (0, 1)$ be an irrational number. We consider the sequence $\{n\alpha + 2^n d\}$. Consider the topological dynamical system $(\mathbb{T}\times\mathbb{T}, S = R_\alpha \times T_2)$ where R_α is the $+\alpha$ mod 1 map and T_2 is the doubling map: $T_2(x) = 2x \bmod 1$. Let $Z = \overline{\{S^n(0, d)\}_{n\geq 0}}$. As S is continuous, by the Bogoliubov–Krylov theorem and ergodic decomposition, we can find an S-ergodic probability measure ρ supported on Z. Let $\mathcal{M}$ be the Borel σ-algebra on $\mathbb{T}$. Then we see that ρ is a joining between $(\mathbb{T}, \mathcal{M}, R_\alpha, \mu)$ and $(\mathbb{T}, \mathcal{M}, T_2, \nu)$ where $\mu = \pi_1\rho$, $\nu = \pi_2\rho$. Note that μ is the Lebesgue measure.

We now use Theorem 4.2. For each $\epsilon > 0$, we can find a probability measure ρ' supported on Z such that $\pi_1\rho'$ is the Lebesgue measure on $\mathbb{T}$ and there is a Borel set B_ϵ such that $\overline{\dim}_{\mathrm{B}} B_\epsilon \leq \epsilon$ and $\rho'(\pi_2^{-1}(B_\epsilon)) > 0$. Here, we choose $\mathcal{A}_0 = \{[0, 0.5), [0.5, 1)\}$ for the doubling map. For this choice, we see that $\mathcal{A}_n$ consists of dyadic intervals of length 2^{-n-1}. Then it is possible to see that $\overline{\dim}_{\mathcal{A}_0,T_2}$ coincides with the upper box dimension. Consider $A = \pi_2^{-1}(B_\epsilon) \cap Z$. As ρ' supports on Z, we see that

$$\rho'(A) > 0.$$

Since A is Borel, we see that $\pi_1(A)$ is Lebesgue measurable. However, as $\pi_1(A)$ might not be Borel measurable, we cannot use the fact that $\pi_1\rho' = \mu$ to deduce that $\pi_1(A)$ has positive Lebesgue measure since all measures here are only defined on Borel sets. If $\pi_1(A)$ has zero Lebesgue measure, then as it is Lebesgue measurable, we see that, for each $\delta > 0$, we can cover $\pi_1(A)$ with open intervals with total length at most δ. Denote the union of those intervals as A^δ. Then $\pi_1^{-1}(A^\delta)$ is Borel and we have $\rho'(\pi_1^{-1}(A^\delta)) = \mu(A^\delta) \leq \delta$. However, as $A \subset \pi_1^{-1}(A^\delta)$, we see that δ cannot be chosen arbitrarily small. Therefore $\pi_1(A)$ has positive Lebesgue measure and hence full Hausdorff dimension. Let Σ denote the arithmetic sum map, that is, $\Sigma(x, y) = x + y$ for $(x, y) \in \mathbb{T}\times\mathbb{T}$. We have

$$\begin{aligned}1 &= \dim_{\mathrm{H}}(\pi_1(A)) \leq \dim_{\mathrm{H}}(\Sigma(A) - \pi_2(A)) \leq \dim_{\mathrm{H}}(\Sigma(A)\times\pi_2(A))\\ &\leq \dim_{\mathrm{H}}(\Sigma(A)) + \overline{\dim}_{\mathrm{B}}\pi_2(A).\end{aligned}$$

Here we have used the fact that

$$\pi_1(A) \subset \Sigma(A) - \pi_2(A) = \{a - b : (a, b) \in \Sigma(A)\times\pi_2(A)\}.$$

We have also used the fact that Σ is a Lipschitz map. The rightmost inequality is a standard result in geometric measure theory; see [**M99**, Theorem 8.10]. Thus we see that

$$\dim_{\mathrm{H}} \overline{\{n\alpha + 2^n d \bmod 1\}_{n\geq 0}} = \dim_{\mathrm{H}} \Sigma(Z) \geq \dim_{\mathrm{H}} \Sigma(A) \geq 1 - \overline{\dim}_{\mathrm{B}}\pi_2(A) \geq 1 - \epsilon.$$

As the above holds for all $\epsilon > 0$, we see that $\dim_{\mathrm{H}} \overline{\{n\alpha + 2^n d \bmod 1\}_{n\geq 0}} = 1$.

We now let p be a polynomial with at least one irrational coefficient. Then the argument above for the special case $p(n) = n\alpha$ can be used here. We need to choose the X component in Theorem 4.2 to be the transformation

$$(x_1, \ldots, x_n) \in \mathbb{T}^n \to (x_1 + \alpha, x_2 + x_1, x_3 + x_2, \ldots, x_n + x_{n-1})$$

on $\mathbb{T}^n$ with a suitably chosen number α, and Σ to be the map

$$(x_1, \ldots, x_n, y) \to \Sigma(x_1, \ldots, x_n, y) = x_n + y.$$

See also [**EW11**, Theorem 1.4] and its proof therein. □

Remark 5.1. In fact, the above proof shows that, for any non-empty closed $R_\alpha \times T_2$ invariant set Z, $\Sigma(Z)$ has full Hausdorff dimension.

Acknowledgements. The author was financially supported by the University of St Andrews, the University of Cambridge and Corpus Christi College, Cambridge. The author has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 803711). The proof of a weaker version of the main result was much more complicated in a previous version of this manuscript. Thanks to an anonymous referee, almost ten pages of technical proofs were distilled into the simple-to-state, simple-to-prove and stronger Theorem 4.2. The author thanks De-Jun Feng, J. Fraser, T. Keleti for fruitful discussions.

References

[**27BMO**] 27th Brazilian Mathematical Olympiad, third round, problem 6, 2005.

[**D11**] T. Downarowicz. *Entropy in Dynamical Systems*. Cambridge University Press, Cambridge, 2011.

[**EW11**] M. Einsiedler and T. Ward. *Ergodic Theory: With a View towards Number Theory (Graduate Texts in Mathematics)*. Springer, London, 2011.

[**F05**] K. Falconer. *Fractal Geometry: Mathematical Foundations and Applications*, 2nd edn. John Wiley & Sons, Chichester, 2005.

[**F67**] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation. *Math. Syst. Theory* **1**(1) (1967), 1–49.

[**FX18**] D.-J. Feng and Y. Xiong. Affine embeddings of Cantor sets and dimension of $\alpha\beta$-sets. *Israel J. Math.* **226**(2) (2018), 805–826.

[**K79**] Y. Katznelson. On $\alpha\beta$-sets. *Israel J. Math.* **33**(1) (1979), 1–4.

[**M99**] P. Mattila. *Geometry of Sets and Measures in Euclidean Spaces: Fractals and Rectifiability (Cambridge Studies in Advanced Mathematics)*. Cambridge University Press, Cambridge, 1999.

[**OW75**] D. Ornstein and B. Weiss. Unilateral codings of Bernoulli systems. *Israel J. Math.* **21** (1975), 159–166.

[**PY98**] M. Pollicott and M. Yuri. *Dynamical Systems and Ergodic Theory (London Mathematical Society Student Texts)*. Cambridge University Press, Cambridge, 1998.

[**W16**] M. Wu. A proof of Furstenberg's conjecture on the intersections of $\times p$ and $\times q$-invariant sets. *Ann. of Math. (2)* **189**(3) (2019), 707–751.

[**Y18**] H. Yu. Multi-rotations on the unit circle. *J. Number Theory* **200** (2019), 316–328.