Ergod. Th. & Dynam. Sys. (1999), **19**, 523–534
Printed in the United Kingdom © *1999 Cambridge University Press*

Bi-invariant sets and measures have integer Hausdorff dimension

DAVID MEIRI and YUVAL PERES†

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

(*Received* 11 *December* 1996 *and accepted in revised form* 18 *June* 1997)

Abstract. Let A, B be two diagonal endomorphisms of the d-dimensional torus with corresponding eigenvalues relatively prime. We show that for any A-invariant ergodic measure μ, there exists a projection onto a torus $\mathbb{T}^r$ of dimension $r \geq \dim \mu$, that maps μ-almost every B-orbit to a uniformly distributed sequence in $\mathbb{T}^r$. As a corollary we obtain that the Hausdorff dimension of any bi-invariant measure, as well as any closed bi-invariant set, is an integer.

1. *Introduction*

Let $a, b > 1$ be relatively prime integers, and let μ be a measure on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ which is invariant and ergodic for multiplication mod 1 by a. Host [**4**] proved that if μ has positive entropy, then the sequence $\{b^n x\}_{n\in\mathbb{N}}$ is uniformly distributed mod 1 for μ-almost every $x \in \mathbb{T}$. Here we extend this to higher dimensions, where the Hausdorff dimension of the measure plays a crucial role. All measures we consider are Borel probability measures. Denote by $\dim_H S$ the Hausdorff dimension of a set S, and for any measure μ, let

$$\dim \mu \stackrel{\text{def}}{=} \inf\{\dim_H S \mid \mu(S) = 1\}.$$

Denote Lebesgue measure on $\mathbb{T}^d$ by λ_d.

THEOREM 1.1. *Let*

$$A = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_d \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} b_1 & & \\ & \ddots & \\ & & b_d \end{pmatrix}$$

be two diagonal endomorphisms of $\mathbb{T}^d$. *Suppose that*

$$|a_i| > 1, \quad |b_i| > 1 \quad \textit{and} \quad \gcd(a_i, b_i) = 1 \quad \textit{for } i = 1, \ldots, d. \tag{1}$$

† Present address: Department of statistics, 367 Evans Hall, University of California, Berkeley CA 94720-3860, USA. (e-mail: peres@stat.berkeley.edu)

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

(i) *Let μ be an A-invariant ergodic measure with* $\dim \mu > 0$. *Then there exist an integer* $r \geq \dim \mu$ *and a projection* $P : \mathbb{T}^d \longrightarrow \mathbb{T}^r$ *such that* $\{PB^n\mathbf{x}\}_{n\in\mathbb{N}}$ *is uniformly distributed for μ-almost every* $\mathbf{x} \in \mathbb{T}^d$. *In particular,*

$$\frac{1}{N}\sum_{n=0}^{N-1} PB^n\mu \longrightarrow \lambda_r$$

weakly. Consequently, $PB^n\mu \longrightarrow \lambda_r$ *weakly as* $n \underset{n\notin J}{\longrightarrow} \infty$, *for some zero-density set* $J \subset \mathbb{N}$.

(ii) *Let* $S \subset \mathbb{T}^d$ *be a closed set such that* $AS \subset S$ *and* $\dim_H S > 0$. *Then there exist an integer* $r \geq \dim S$ *and a projection* $P : \mathbb{T}^d \longrightarrow \mathbb{T}^r$ *with the following property: for some zero-density set* $J \subset \mathbb{N}$ *we have* $PB^n S \longrightarrow \mathbb{T}^r$ *in the Hausdorff metric as* $n \underset{n\notin J}{\longrightarrow} \infty$.

The second part of Theorem 1.1 is related to the following conjecture of Furstenberg, mentioned in Lyons [**11**]:

> *If p and q are multiplicatively independent integers and $S \subset \mathbb{T}$ is a p-invariant infinite closed set, then $q^n S \longrightarrow \mathbb{T}$ as $n \longrightarrow \infty$.*

Invariant sets for a single toral endomorphism can have fractional Hausdorff dimension. Using Theorem 1.1 we show that the situation is different for sets invariant under two diagonal endomorphisms.

THEOREM 1.2. *Let* $A, B \in \mathrm{End}(\mathbb{T}^d)$ *be as in Theorem 1.1.*

(i) *If μ is a measure on $\mathbb{T}^d$ which is invariant under both A and B, then μ has integer 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 $P\mu$ is a 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* $S = \mathbb{T}^d$.

Remark. In both theorems, if A is scalar, i.e. $A = aI$, then the assumptions on B can be relaxed: it suffices that B has integer eigenvalues $b_1, \ldots, b_d$ such that (1) holds.

For $d = 1$, Theorem 1.2 is well known: Furstenberg [**3,** Theorem IV.1] proved a strong version of the topological assertion, and Rudolph [**15**] proved the measure-theoretic assertion, extending earlier work of Lyons [**11**]. Berend [**1**] considered a similar problem, and gave conditions on toral endomorphisms $A_1, \ldots, A_n$ under which every infinite closed multi-invariant subset must be $\mathbb{T}^d$. Berend's work, as well as the recent work of Katok and Spatzier [**7**], concern the irreducible case (no jointly invariant subspaces) and are thus complementary to the present paper.

The following proposition relates bi-invariant sets and measures.

PROPOSITION 1.3. *Let* $A = \mathrm{diag}(a_1, \ldots, a_d)$ *be a diagonal endomorphism with* $|a_i| > 1$ *for* $i = 1, \ldots, d$. *Let B denote any non-singular endomorphism commuting with A. Suppose that* $S \subset \mathbb{T}^d$ *is a closed set such that* $AS \subset S$ *and* $BS \subset S$. *Then there exists an $\langle A, B\rangle$-invariant and ergodic measure μ, supported on S, with* $\dim \mu = \dim_H S$.

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

The proof of all our results are simpler when one of the matrices is scalar, e.g. $A = aI$. On first reading, the reader may want to restrict attention to this case.

The paper is organized as follows. In §2, we show how the two theorems can be deduced from certain lemmas on dimension and entropy and a generalization of a theorem in Meiri [**13**] concerning the uniform distribution of orbits of an endomorphism. We also prove Proposition 1.3. In the non-scalar case we use a formula from Ledrappier and Young [**10**] to relate entropy and dimension. We also need a theorem from Kenyon and Peres [**9**] ensuring the existence of invariant measures of full dimension. In §3 we prove the various lemmas. §4 contains the detailed formulation and proof of the aforementioned theorem on uniform distribution. §5 contains concluding remarks and unsolved problems.

2. *Proof of the main results*

The proof depends upon the following lemmas, which will be proved in §§3 and 4.

LEMMA 2.1. *Let $(X, \mathfrak{B}, \mu)$ be a probability space and let $T, \widetilde{T}$ be two commuting measure-preserving transformations. Assume that μ is ergodic for the joint action of $\langle T, \tilde{T}\rangle$. Let*

$$\mu = \int \mu_\theta \, d\nu(\theta)$$

be the ergodic decomposition of μ with respect to T. Then the entropy $h(\mu_\theta, T)$ is constant ν-a.e.

For the rest of this section, we specialize to $X = \mathbb{T}^d$ and T a transformation induced by $A = \operatorname{diag}(a_1, \ldots, a_d)$. We may assume that $a_1 \le a_2 \le \cdots \le a_d$. The following notation will be used throughout this paper.

Notation. Denote by $\pi_i : \mathbb{T}^d \longrightarrow \mathbb{T}^i$ the projection to the first i coordinates, i.e. $\pi_i(x_1, \ldots, x_d) = (x_1, \ldots, x_i)$. Also, denote $A_i = \operatorname{diag}(a_1, \ldots, a_i)$.

In the scalar case, i.e. when $A = aI$ for some integer a with $|a| > 1$, the relation between entropy and dimension is simple, and was found by Billingsley [**2**]: if μ is A-invariant and ergodic, then

$$\dim \mu = \frac{h(\mu, A)}{\log a}. \tag{2}$$

When the a_i are not all equal, the relation is less transparent. The correct analogue was found by Ledrappier and Young [**10**] in a more general setting. Here we use the notation of Kenyon and Peres [**9**].

Definition 2.1. Let μ be an A-invariant measure on $\mathbb{T}^d$. Then

$$d_{\mathrm{LY}}(\mu, A) = \sum_{i=1}^{d} \frac{1}{\log a_i} [h(\pi_i \mu, A_i) - h(\pi_{i-1}\mu, A_{i-1})]$$

is called the *Ledrappier–Young dimension* of μ. (By convention $h(\pi_0\mu, A_0) = 0$.)

Note that when $A = aI$ we have $d_{\mathrm{LY}}(\mu, A) = h(\mu, A)/\log a$. Kenyon and Peres [**9**] show that when μ is A-ergodic, $d_{\mathrm{LY}}(\mu, A) = \dim \mu$. Our next two lemmas examine the situation in the non-ergodic case.

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

LEMMA 2.2. *Suppose that μ is an $\langle A, B\rangle$-invariant measure on $\mathbb{T}^d$, where*

$$A = \operatorname{diag}(a_1, \ldots, a_d) \quad \text{and} \quad B = \operatorname{diag}(b_1, \ldots, b_d)$$

and assume μ is ergodic for the joint action of $\langle A, B\rangle$. Let $\mu = \int \mu_\theta \, d\nu(\theta)$ be the ergodic decomposition of μ with respect to A. Then

$$\dim \mu_\theta = d_{\mathrm{LY}}(\mu_\theta, A) = d_{\mathrm{LY}}(\mu, A) \quad \nu\text{-a.e.} \tag{3}$$

LEMMA 2.3. *Let μ be an A-invariant measure on $\mathbb{T}^d$. Let $\mu = \int \mu_\theta \, d\nu(\theta)$ denote the ergodic decomposition for the action of A, or a coarser decomposition. Then*

$$\dim \mu = \text{ess-sup}\, d_{\mathrm{LY}}(\mu_\theta, A).$$

The following Lemma is the main result of Kenyon and Peres [**9**].

LEMMA 2.4. *Given an A-invariant and closed set $S \subset \mathbb{T}^d$, there exists an A-invariant ergodic measure μ on S with full dimension, i.e.* $\dim_H S = \dim \mu$.

In the scalar case this result is well known. Indeed, the argument of Furstenberg [**3,** Proposition III.3] yields that for $A = aI$, any A-invariant closed set S satisfies

$$\dim_H S = \frac{h_{\mathrm{top}}(S, A)}{\log a}, \tag{4}$$

where h_{top} is topological entropy. (For another proof, see §2 in Kenyon and Peres [**8**].) By the variational principle for expansive maps (see Walters [**16**]), there exists an ergodic measure μ with $h(\mu, A) = h_{\mathrm{top}}(S, A)$. This measure has full dimension by (2).

Finally, we formulate a variant of Theorem 6.1 from Meiri [**13**].

LEMMA 2.5. *Let A, B be diagonal endomorphisms satisfying (1). Suppose that μ is an A-invariant ergodic measure on $\mathbb{T}^d$ such that $h(\pi_{d-1}\mu, A_{d-1}) < h(\mu, A)$. Then for every $\mathbf{v} = (v_1, \ldots, v_d) \in \mathbb{Z}^d$ with $v_d \neq 0$, for μ-almost every $\mathbf{x} \in \mathbb{T}^d$ the sequence $\{\mathbf{v}^{\mathrm{tr}} B^n \mathbf{x}\}_{n\in\mathbb{N}}$ is uniformly distributed in $\mathbb{T}$. In other words,*

$$\frac{1}{N} \sum_{n=0}^{N-1} e(k\mathbf{v}^{\mathrm{tr}} B^n \mathbf{x}) \longrightarrow 0 \quad \text{for } \mu\text{-a.e. } x \in \mathbb{T}^d \text{ and all } k \in \mathbb{Z} \setminus \{0\}, \tag{5}$$

where $e(t) = \exp(2\pi i t)$.

See §4 for a proof of this lemma in a more general form.

Proof of Theorem 1.1. (i) Since $d_{\mathrm{LY}}(\mu, A) = \dim \mu > 0$, we have $h(\mu, A) > 0$. The proof proceeds by induction on d. The case $d = 1$ is covered by Host [**4**].

Assume that the theorem holds in dimension $d-1$, and take $\pi = \pi_{d-1}$. Let $\pi\mu$ denote the image of μ under this map, and note that $\pi\mu$ is an A_{d-1}-invariant ergodic measure on $\mathbb{T}^{d-1}$. Let $B_{d-1} = \operatorname{diag}(b_1, \ldots, b_{d-1})$. Applying the induction hypothesis to $\pi\mu$, we can find an integer $s \geq \dim \pi\mu$ and a projection $P' : \mathbb{T}^{d-1} \longrightarrow \mathbb{T}^s$ such that

$$\{P' B_{d-1}^n \mathbf{x}\}_{n\in\mathbb{N}} \text{ is uniformly distributed for } \pi\mu\text{-almost every } \mathbf{x} \in \mathbb{T}^{d-1}. \tag{6}$$

We consider two possible cases. If $h(\pi\mu, A_{d-1}) = h(\mu, A)$, then $d_{\mathrm{LY}}(\mu, A) = d_{\mathrm{LY}}(\pi\mu, A_{d-1})$ and so $\dim \pi\mu = \dim \mu$. Here we are done, taking $r = s$ and $P = P' \circ \pi$.

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

In the other case, $h(\pi\mu, A_{d-1}) < h(\mu, A)$. By Lemma 2.5 it follows that for every $\mathbf{v} \in \mathbb{Z}^d$ with $v_d \neq 0$, the sequence $\{\mathbf{v}^{\mathrm{tr}} B^n \mathbf{x}\}_{n\in\mathbb{N}}$ is uniformly distributed for μ-almost every $\mathbf{x} \in \mathbb{T}^d$. Now take $r \stackrel{\mathrm{def}}{=} s + 1$ and $P(\mathbf{x}) = (P'(\pi(\mathbf{x})), x_d)$. In conjunction with (6), this implies that for every non-zero $\mathbf{v} \in \mathbb{Z}^r$, the sequence $\{\mathbf{v}^{\mathrm{tr}} P B^n \mathbf{x}\}_{n\in\mathbb{N}}$ is uniformly distributed for μ-a.e. $\mathbf{x} \in \mathbb{T}^d$. Since $r = s + 1 \geq \dim \pi\mu + 1 \geq \dim \mu$, we are done.

By Weyl's criterion, $(1/N)\sum_{n=0}^{N-1} PB^n\mu \longrightarrow \lambda_r$. From a general argument of Johnson and Rudolph [**6**] if follows that $PB^n\mu \longrightarrow \lambda_r$ as $n \underset{n\notin J}{\longrightarrow} \infty$, for some zero-density set J.

(ii) By Lemma 2.4, there exists an A-invariant and ergodic measure μ supported on S with $\dim \mu = \dim_H S$. Since by the first part $PB^n\mu \longrightarrow \lambda_r$ as $n \underset{n\notin J}{\longrightarrow} \infty$, we obtain

$$\mathrm{supp}(PB^n\mu) \longrightarrow \mathrm{supp}(\lambda_r) = \mathbb{T}^r \quad \text{as } n \underset{n\notin J}{\longrightarrow} \infty,$$

and we are done. □

Remark. The projection $P : \mathbb{T}^d \longrightarrow \mathbb{T}^r$ constructed above satisfies

$$P(x_1, \ldots, x_d) = (x_i)_{i\in R}, \text{ where } R = \{i : h(\pi_{i-1}\mu, A_{i-1}) < h(\pi_i\mu, A_i)\}. \tag{7}$$

Proof of Theorem 1.2. (i) We first consider the case where μ is ergodic with respect to the joint action of $\langle A, B\rangle$.

Let $\mu = \int \mu_\theta \, d\nu(\theta)$ be the ergodic decomposition of μ with respect to A. Almost every component μ_θ is an A-invariant and ergodic measure (but is not necessarily B-invariant). By Lemma 2.2, $\dim \mu_\theta = d_{\mathrm{LY}}(\mu_\theta, A)$ is constant ν-a.e. Lemma 2.3 implies that $\dim \mu = d_{\mathrm{LY}}(\mu, A) = d_{\mathrm{LY}}(\mu_\theta, A)$ ν-a.e.

Applying Theorem 1.1 to each ergodic component of μ, we can find for a projection $P : \mathbb{T}^d \longrightarrow \mathbb{T}^r$ for some $r \geq \dim \mu_\theta = \dim \mu$ such that

$$\frac{1}{N}\sum_{n=0}^{N-1} PB^n\mu_\theta \longrightarrow \lambda_r. \tag{8}$$

Strictly speaking, r and P depend on θ, but by (7) they are the same for ν-almost every θ. (The entropies $h(\pi_i\mu_\theta, A_i)$ are constant ν-a.e. for $i = 1, \ldots, d$ by Lemma 2.1).

Integrating (8) with respect to ν we get

$$\frac{1}{N}\sum_{n=0}^{N-1} PB^n\mu \longrightarrow \lambda_r.$$

As μ is B-invariant, $P\mu = \lambda_r$.

Since a projection cannot increase dimension, $\dim \mu \geq \dim \lambda_r = r$. But $r \geq \dim \mu$, hence $\dim \mu = r$ is an integer. Note that for $d = r$ we get $\mu = \lambda_d$.

It remains to consider the case where μ is not ergodic for the action of $\langle A, B\rangle$. Let $\mu = \int \mu_\alpha \, d\eta(\alpha)$ be the ergodic decomposition of μ with respect to $\langle A, B\rangle$. The argument above implies that $\dim \mu_\alpha$ is an integer for η-almost every α. By Lemma 2.3, $\dim \mu = \text{ess-sup} \dim \mu_\alpha$, so $\dim \mu$ is an integer in this case as well.

(ii) Let S be a closed bi-invariant set. If S is infinite, then the projection of S on at least one of the coordinate axes is closed, infinite and bi-invariant. By Furstenberg [**3**] it is the whole circle, hence $\dim_H S \geq 1$.

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

In the general case, by Theorem 1.1(ii) there exist some integer $r \geq \dim_H S$ and a projection $P : \mathbb{T}^d \longrightarrow \mathbb{T}^r$ such that $PB^nS \longrightarrow \mathbb{T}^r$, as $n \underset{n \notin J}{\longrightarrow} \infty$. Since $BS \subset S$, we have $r \leq \dim_H(PS) \leq \dim_H S$, and so $\dim_H S = r$ is an integer. When $\dim_H S = d$, the projection map P is the identity, and so $S = \mathbb{T}^d$. □

The proof when B is not diagonal. As remarked before, when $A = aI$ Theorems 1.1 and 1.2 extend to any matrix B such that all its eigenvalues are integers greater than one and relatively prime to a. In the proof of Theorem 1.1 we may assume that B is upper-triangular, by replacing it with an appropriate conjugate QBQ^{-1}. The measures $\pi_i\mu$ are still A_i-invariant, and Theorem 4.2 shows that (5) remains valid. The proof of Theorem 1.2 does not change, noting that Lemma 2.2 remains valid.

Proof of Proposition 1.3. Denote $\gamma = \dim_H S$. By Lemma 2.4, S supports some A-invariant ergodic measure μ such that $\dim \mu = d_{\mathrm{LY}}(\mu, A) = \gamma$. Thus,

$$I_A = \{\mu : \mu \text{ is a probability measure, } A\mu = \mu, \mu(S) = 1, d_{\mathrm{LY}}(\mu, A) = \gamma\}$$

is non-empty.

We claim that I_A is also convex, weak*-closed and B-invariant. The first assertion follows from the affinity of entropy. As for the second, since entropy is upper semi-continuous for expansive maps, so is $d_{\mathrm{LY}}(\cdot, A)$. Thus it can only increase when we pass to a limit point; but every A-invariant measure μ on S satisfies $d_{\mathrm{LY}}(\mu, A) \leq \gamma$ by Lemma 2.3, and therefore $d_{\mathrm{LY}}(\mu, A) = \gamma$ for any μ which is a limit of measures in I_A. Finally, for $\mu \in I_A$, the system $(S, B\mu)$ is a factor of (S, μ) since B commutes with A. The factor map B is finite-to-one, whence $h(B\mu, A) = h(\mu, A)$. The same consideration applied to the other entropies appearing in the definition of Ledrappier–Young dimension shows that $d_{\mathrm{LY}}(\mu, A) = d_{\mathrm{LY}}(B\mu, A)$, so I_A is B-invariant.

By Kakutani's fixed point theorem, I_A contains a fixed point ν for the action of B. This is a bi-invariant measure supported on S with $d_{\mathrm{LY}}(\nu, A) = \gamma$. Since $d_{\mathrm{LY}}(\cdot, A)$ is affine, there exists an ergodic component $\tilde{\nu}$ of ν (with respect to $\langle A, B\rangle$), such that $d_{\mathrm{LY}}(\tilde{\nu}, A) \geq \gamma$. As $d_{\mathrm{LY}}(\tilde{\nu}, A) = \dim \tilde{\nu} \leq \dim_H S = \gamma$, we conclude that $\dim \tilde{\nu} = \gamma$. □

3. *Entropy and dimension*

In this section we prove Lemmas 2.1–2.3.

Let $(X, \mathfrak{B}, \mu, T)$ be a measure-preserving system, and denote by $\mathfrak{B}_T$ the μ-completion of the σ-algebra of T-invariant sets. Let α be some finite measurable partition. Denote by

$$\alpha_0^{n-1} = \alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha$$

the generated n-step past algebra. Given a partition β, we denote by $\beta(x)$ the atom of β containing x, and let $I_\beta(x) = -\log \mu(\beta(x))$ be the corresponding information function. We have $h_\mu(\alpha, T) = \int I_{\alpha|\alpha_1^\infty}\, d\mu$.

The Shannon–McMillan–Breiman (SMB) theorem states that when T is ergodic,

$$\frac{1}{n} I_{\alpha_0^{n-1}}(x) \longrightarrow h_\mu(\alpha, T) \quad \mu\text{-a.e. and in } L^1.$$

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

Without assuming ergodicity we have (cf. Parry [**14,** p. 39]):

$$\frac{1}{n} I_{\alpha_0^{n-1}}(x) \longrightarrow \psi(x) \quad \mu\text{-a.e. and in } L^1,$$

where $\psi = E(I_{\alpha|\alpha_1^\infty}|\mathfrak{B}_T)$. The function ψ is T-invariant, and corresponds to the entropy in ergodic fibres of μ. More precisely, let $\mu = \int \mu_\theta \, d\nu(\theta)$ be the ergodic decomposition of μ with respect to T. Then $\psi(x) = h(\mu_\theta, T)$ μ_θ-a.e. for almost every θ, and therefore

$$\int \psi \, d\mu = \int h(\mu_\theta, T) \, d\nu(\theta) = h(\mu, T), \tag{9}$$

by Rokhlin's theorem. We will call ψ the *SMB function* determined by the transformation T and the partition α.

Proof of Lemma 2.1. We use the above notation. Since ψ is T-invariant, once we show that it is $\widetilde{T}$-invariant as well, we will be done by the ergodicity assumption. Since $\{\widetilde{T}\mu_\theta\}$ is another set of ergodic T-invariant fibres, this set coincides with $\{\mu_\theta\}$ by the uniqueness of the ergodic decomposition. Hence $\psi(\widetilde{T}x) = h(\widetilde{T}\mu_\theta, T)$ is the entropy of a factor of (X, μ_θ, T), and therefore $\psi(\widetilde{T}x) \leq \psi(x)$. Since $\widetilde{T}$ is measure-preserving, $\int \psi(\widetilde{T}x) \, d\mu = \int \psi(x) \, d\mu$, so ψ must be $\widetilde{T}$-invariant. □

To compute the Hausdorff dimension of the measures we need the following lemma, which can be found in Billingsley [**2**].

BILLINGSLEY'S LEMMA. *Let μ be a positive finite measure on the d-torus $\mathbb{T}^d$. Assume $K \subset \mathbb{T}^d$ is a Borel set satisfying $\mu(K) > 0$ and*

$$K \subset \left\{ x \in \mathbb{T}^d : \liminf_{\varepsilon \downarrow 0} \frac{\log \mu[Q_\varepsilon(x)]}{\log \varepsilon} \leq \gamma \right\}.$$

Then $\dim_H K \leq \gamma$. *If the* lim inf *is γ a.e. then* $\dim_H K = \gamma$.

Here, following McMullen [**12**], we can take $Q_\varepsilon(x)$ to be a box which is an 'approximate cube', i.e. $Q_\varepsilon(x)$ contains a cube of side ε containing x, and $\operatorname{diam}[Q_\varepsilon(x)] \leq C\varepsilon$ for some constant C.

Proof of Lemma 2.2. The left-hand equality is true for every θ such that μ_θ is ergodic, by Kenyon and Peres [**9**]. For the right-hand equality, note that for $i = 1, \ldots, d$ the measure $\pi_i \mu$ is invariant for $A_i = \operatorname{diag}(a_1, \ldots, a_i)$ and $B_i = \operatorname{diag}(b_1, \ldots, b_i)$, and is ergodic for $\langle A_i, B_i \rangle$. From Lemma 2.1 we deduce that $h(\pi_i \mu_\theta, A_i)$ is constant ν-a.e., and therefore so is $d_{\mathrm{LY}}(\mu_\theta, A)$. By the affinity of entropy, $d_{\mathrm{LY}}(\mu, A) = \int d_{\mathrm{LY}}(\mu_\theta, A) \, d\nu(\theta)$, and (3) follows. □

Proof of Lemma 2.3. It is enough to prove the lemma for the ergodic decomposition; once we show this, for a coarser decomposition we use the expression:

$$\dim \mu_\theta = \text{ess-sup}\{\dim \nu : \nu \text{ is an ergodic component of } \mu_\theta\},$$

to derive the assertion.

If $\mu(S) = 1$, then $\mu_\theta(S) = 1$ for almost every θ. Hence $\dim \mu_\theta \leq \dim_H S$ a.e., and so ess-sup $\dim \mu_\theta \leq \dim_H S$. By definition we get

$$\dim \mu = \inf\{\dim_H S : \mu(S) = 1\} \geq \text{ess-sup} \dim \mu_\theta.$$

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

The inequality in the other direction is more difficult, and we first show it in the scalar case $A = aI$. Let α denote the partition of $\mathbb{T}^d$ to a^d cubes with side $1/a$. Denote by $Q_n(\mathbf{x})$ the cube in α_0^{n-1} containing $\mathbf{x}$ (i.e. the mesh cube with edge a^{-n} containing $\mathbf{x}$). Applying the SMB theorem to μ we get

$$\frac{1}{n} I_{\alpha_0^{n-1}}(\mathbf{x}) = -\frac{1}{n} \log \mu(Q_n(\mathbf{x})) \longrightarrow \psi(\mathbf{x}) \qquad \mu\text{-a.e.}$$

Now, $\psi(\mathbf{x}) = h(\mu_\theta, A)$ μ_θ-a.e., and since μ_θ is ergodic, by Billingsley [**2**]

$$\dim \mu_\theta = h(\mu_\theta, A)/\log a.$$

Writing $\gamma \stackrel{\text{def}}{=} \text{ess-sup} \dim \mu_\theta$ we get

$$\lim_{n\to\infty} -\frac{\log \mu(Q_n(\mathbf{x}))}{n \log a} \le \gamma \quad \mu\text{-a.e.} \tag{10}$$

From Billingsley's lemma we conclude that $\dim \mu \le \gamma$.

The proof in the general case follows Kenyon and Peres [**9**]. We will indicate the proof in detail in the two-dimensional case $A = \left(\begin{smallmatrix} m & 0 \\ 0 & n \end{smallmatrix}\right)$ with $m \le n$. Here

$$d_{\text{LY}}(\mu, A) = \frac{h(\mu, A)}{\log n} + \left(\frac{1}{\log m} - \frac{1}{\log n}\right) h(\pi_1 \mu, A_1). \tag{11}$$

We have $\gamma = \text{ess-sup} \dim \mu_\theta = \text{ess-sup}\, d_{\text{LY}}(\mu_\theta, A)$, and we need to prove that $\dim \mu \le \gamma$.

Let $\tau = \log m / \log n$. For $(x, y) \in \mathbb{T}^2$, denote by $Q_k(x, y)$ the rectangle of size $m^{-k} \times n^{-\lfloor \tau k \rfloor}$ consisting of points (x', y') in $\mathbb{T}^2$ for which the first k digits in the base m expansions of x and x' agree, and similarly for the first $\lfloor \tau k \rfloor$ digits in the base n expansions of y and y'.

By Billingsley's lemma it suffices to show that

$$\liminf_{k\to\infty} -\frac{\log \mu(Q_k(x, y))}{k \log m} \le \gamma \quad \mu\text{-a.e.}$$

Thus Lemma 2.3 reduces to the following variant of the SMB theorem.

LEMMA 3.1. *Let $W_1, W_2, W_3, \ldots$ be a stationary process taking values in*

$$D = \{0, 1, \ldots, m-1\} \times \{0, 1, \ldots, n-1\},$$

with SMB function $\psi(\omega)$. Let $\pi : D \to \{0, 1, \ldots, m-1\}$ denote projection to the first coordinate and write $X_j = \pi W_j$. Let $\chi(\omega)$ denote the SMB function of the X process. Then for any $0 < \tau \le 1$, almost surely

$$\lim_{k\to\infty} \frac{1}{k} \log \mathbb{P}[W_1, W_2, \ldots, W_{\tau k}; X_{\tau k+1}, X_{\tau k+2}, \ldots, X_k] = -\tau\psi(\omega) - (1-\tau)\chi(\omega). \tag{12}$$

Here $\mathbb{P}[w_1, w_2, \ldots, w_\ell; x_{\ell+1}, \ldots, x_k]$ is shorthand for

$$\mathbb{P}\{[W_1 = w_1, W_2 = w_2, \ldots, W_\ell = w_\ell] \cap [\pi W_{\ell+1} = x_{\ell+1}, \ldots, \pi W_k = x_k]\}.$$

In a customary abuse of notation, we then substitute random variables for w_i and x_j; we also write τk instead of $\lfloor \tau k \rfloor$ in (12).

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

Proof of Lemma 3.1. By SMB

$$\lim_{k\to\infty}\frac{1}{k}\log\mathbb{P}[W_1, W_2, \ldots, W_{\tau k}] = -\tau\psi(\omega) \quad \text{a.e.},$$

and

$$\lim_{k\to\infty}\frac{1}{k}\log\mathbb{P}[X_{\tau k+1}, \ldots, X_k] = -(1-\tau)\chi(\omega) \quad \text{a.e.}$$

Now one shows as in Lemma 3.2 of Kenyon and Peres [**9**] that

$$\frac{\mathbb{P}[W_1, \ldots, W_{\tau k}]\mathbb{P}[X_{\tau k+1}, \ldots, X_k]}{\mathbb{P}[W_1, \ldots, W_{\tau k}; X_{\tau k+1}, \ldots, X_k]} \longrightarrow 1 \quad \text{a.e.}$$

This concludes the proof of Lemma 3.1. □

To complete the proof of Lemma 2.3, observe that

$$\frac{\tau\psi(\omega) + (1-\tau)\chi(\omega)}{\log m} = d_{\mathrm{LY}}(\mu_\theta, A) \quad \mu_\theta\text{-a.e.}$$

for almost every θ by (11), and we are done in the same way as in the scalar case.

In the d-dimensional case, where $A = \mathrm{diag}(a_1, \ldots, a_d)$, we proceed as follows. We can assume that $a_1 \le a_2 \le \cdots \le a_d$. For $i = 1, \ldots, d$, let $\tau_i = \log a_1/\log a_i$. For $\mathbf{x} \in \mathbb{T}^d$ denote by $Q_k(\mathbf{x})$ the collection of points $\mathbf{x}' \in \mathbb{T}^d$ such that the first $\lfloor \tau_i k \rfloor$ digits in the base a_i expansions of x_i and x_i' agree, for $i = 1, \ldots, d$. As before, by Billingsley's lemma it suffices to show that

$$\liminf_{k\to\infty} -\frac{\log\mu(Q_k(\mathbf{x}))}{k\log a_1} \le \gamma \quad \mu\text{-a.e.}$$

This reduces to the following statement. Let $\{W_n\}$ be a stationary process taking values in $D = \prod_{i=1}^d \{0, 1, \ldots, a_i - 1\}$. For $i = 1, \ldots, d$, denote by $\psi_i(\omega)$ the SMB function of the process $\{\pi_i W_n\}$. Then for any $0 = \tau_{d+1} < \tau_d \le \tau_{d-1} \le \cdots \le \tau_1 = 1$, almost surely

$$\lim_{k\to\infty}\frac{1}{k}\log\mathbb{P}[\pi_i W_j : 1 \le i \le d, \tau_{i+1}k < j \le \tau_i k] = -\sum_{i=1}^d (\tau_i - \tau_{i+1})\psi_i(\omega).$$

This statement is proved by combining the SMB theorem for the process $\{\pi_i W_n\}_{n\ge 1}$ for $i = 1, \ldots, d$, and the proof of Lemma 4.4 in Kenyon and Peres [**9**]. □

4. *Uniform distribution of orbits and proof of Lemma 2.5*

In this section we present a d-dimensional analogue of Theorem 6.1 of Meiri [**13**], dealing with uniform distribution of orbits of endomorphisms. In that paper only the two-dimensional case was considered in detail.

First, some definitions. By a 'direction vector' we mean a vector $\mathbf{v} = (v_1, \ldots, v_d) \in \mathbb{Z}^d$ other than zero, such that $\gcd(\mathbf{v}) \stackrel{\text{def}}{=} \gcd(v_1, \ldots, v_d) = 1$. For $\mathbf{z} \in \mathbb{Z}^d$ denote by $\pi_{\mathbf{z}}$ the orthogonal projection

$$\pi_{\mathbf{z}} : \mathbb{T}^d \longrightarrow \mathbf{z}^\perp = \{\mathbf{x} \in \mathbb{T}^d : \langle \mathbf{z}, \mathbf{x}\rangle = 0\}$$

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

given by $\pi_{\mathbf{z}} : \mathbf{x} \mapsto |\mathbf{z}|^2\mathbf{x} - \langle \mathbf{x}, \mathbf{z}\rangle \mathbf{z}$. Here $\langle \mathbf{z}, \mathbf{x}\rangle \stackrel{\text{def}}{=} \sum z_i x_i$.

Say that a sequence of integers $\{c_k\}$ has *almost-sub-exponential collisions (mod a)* if for every $\gamma > 0$ there exists a set $D \subset \mathbb{N}$ satisfying

$$\forall N \geq 1, \quad \frac{|D \cap \{0, 1, \ldots, N-1\}|}{N} \leq \gamma \tag{13}$$

such that for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \#\{k, l : 0 \leq k, l < a^n,\ k, l \notin D,\ c_k \equiv c_l \pmod{a^n}\}/a^{n(1+\varepsilon)} = 0. \tag{14}$$

We quote the following theorem.

THEOREM 4.1. (Meiri [**13**]) *Given an integer a, and numbers $b_1, \ldots, b_d$ such that*

$$|a| > 1, |b_i| > 1 \text{ and } \gcd(a, b_i) = 1 \text{ for } i = 1, \ldots, d \tag{15}$$

the sequence $c_k = \sum f_i b_i^k$, if not identically zero, has almost-sub-exponential collisions (mod a) for every choice of integers $f_1, \ldots, f_d$.

Note that these sequences arise naturally in the form $c_k = \mathbf{v}^{\text{tr}} B^k \mathbf{w}$, where $\mathbf{v}, \mathbf{w}$ are any fixed integer vectors, and B is any $d \times d$ integer matrix with eigenvalues $b_1, \ldots, b_d$. For a general $d \times d$ integer matrix B we quote the following result from Meiri [**13**].

> *Suppose that $\gcd(a, \det B) = 1$ and none of the eigenvalues of B or their ratios are roots of unity (in $\mathbb{C}$). Then for any two non-zero vectors $\mathbf{v}, \mathbf{w} \in \mathbb{Z}^d$ such that $\mathbf{w}$ is not an eigenvector of B orthogonal to $\mathbf{v}$, the sequence $c_k = \mathbf{v}^{\text{tr}} B^k \mathbf{w}$ has almost-sub-exponential collisions (mod a).*

THEOREM 4.2. *Let $A = \text{diag}(a_1, \ldots, a_d)$, and suppose that μ is an A-invariant ergodic measure on $\mathbb{T}^d$. Let B be any endomorphism of $\mathbb{T}^d$.*

Assume that for some $\mathbf{v} \in \mathbb{Z}^d$ there exists a direction vector $\mathbf{z}$ such that:

(i) *if A is not scalar, $\mathbf{z}$ is a unit vector;*

(ii) *$h(\pi_{\mathbf{z}}\mu, A) < h(\mu, A)$;*

(iii) *$c_k \stackrel{\text{def}}{=} \mathbf{v}^{\text{tr}} B^k \mathbf{z}$ has almost-sub-exponential collisions (mod a), where a is well defined if $A = aI$, and we take $a = a_i$ if A is non-scalar and $\mathbf{z} = \mathbf{e}_i$.*

Then $\{\mathbf{v}^{\text{tr}} B^n \mathbf{x}\}$ is uniformly-distributed in $\mathbb{T}$ for μ-almost every $\mathbf{x} \in \mathbb{T}^d$, i.e.

$$g_N(\mathbf{x}) \stackrel{\text{def}}{=} \frac{1}{N} \sum_{n=0}^{N-1} e(m\mathbf{v}^{\text{tr}} B^n \mathbf{x}) \longrightarrow 0 \quad \mu\text{-a.e.} \quad \text{for all } m \in \mathbb{Z} \setminus \{0\}, \tag{16}$$

where $e(t) = e^{2\pi i t}$.

Note that if (16) holds for all non-zero vectors $\mathbf{v} \in \mathbb{Z}^d$, then $\{B^n \mathbf{x}\}$ is uniformly-distributed for μ-almost every $\mathbf{x} \in \mathbb{T}^d$, by Weyl's criterion.

Proof. Since we need this theorem only for unit directions, i.e. when $\mathbf{z}$ is one of the vectors $\mathbf{e}_1, \ldots, \mathbf{e}_d$ from the standard basis, we will first prove it in this case, and then indicate how the general case can be handled. For this particular case we may assume, without loss of generality, that $\mathbf{z} = \mathbf{e}_1$. If $\mathbf{v}$ satisfies (iii), then so does every multiple $m\mathbf{v}$; thus it suffices to prove (16) for $m = 1$.

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

Define measures $\omega_n = \sum_{j=0}^{a^n-1} \mu * \delta_{ja^{-n}\mathbf{e}_1}$ and let $\varphi_n = d\mu/d\omega_n$ be the Radon–Nikodym derivative. Let $h' = h(\mu, A) - h(\pi_{\mathbf{z}}\mu, A) > 0$. By Lemmas 6.2 and 6.3 in Meiri [**13**],

$$\varphi_n^{1/n} \longrightarrow e^{-h'} \quad \mu\text{-a.e.} \tag{17}$$

Thus, following Host [**4**],

$$\begin{aligned}
\int \frac{|g_N|^2}{\varphi_n}\, d\mu &\le \int |g_N|^2\, d\omega_n \\
&= \sum_{j=0}^{a^n-1} \int \left|g_N(\mathbf{x} + ja^{-n}\mathbf{e}_1)\right|^2 d\mu(\mathbf{x}) \\
&= \frac{1}{N^2} \sum_{k,l=0}^{N-1} \sum_{j=0}^{a^n-1} \int e(\mathbf{v}^{\mathrm{tr}}(B^k - B^l)(\mathbf{x} + ja^{-n}\mathbf{e}_1))\, d\mu \\
&\le \frac{1}{N^2} \sum_{k,l=0}^{N-1} \sum_{j=0}^{a^n-1} e(\mathbf{v}^{\mathrm{tr}}(B^k - B^l)\mathbf{e}_1 ja^{-n}) \\
&= \frac{a^n}{N^2} \#\{0 \le k, l < N : c_k \equiv c_l \pmod{a^n}\}.
\end{aligned}$$

As in Meiri [**13**], taking $N \approx a^n$ it follows that $g_N(\mathbf{x}) \longrightarrow 0$ μ-a.e.

When $A = aI$ and $\mathbf{z}$ is an arbitrary integer vector with $\gcd(\mathbf{z}) = 1$, we proceed as follows. Find some integer vector $\mathbf{u}_1$ such that $\langle \mathbf{z}, \mathbf{u}_1 \rangle = 1$. There exist integer vectors $\mathbf{u}_2, \ldots, \mathbf{u}_d$, orthogonal to $\mathbf{z}$, such that the matrix U with rows $\mathbf{u}_1, \ldots, \mathbf{u}_d$ is unimodular. Let $\mu' = U\mu$. Clearly, $h(\mu, A) = h(\mu', A)$, and we claim that also $h(\pi_{\mathbf{e}_1}\mu', A) = h(\pi_{\mathbf{z}}\mu, A)$. Indeed, letting V denote the regular matrix with rows $\mathbf{z}, \mathbf{u}_2, \ldots, \mathbf{u}_d$, we see that $V\pi_{\mathbf{z}} = |\mathbf{z}|^2 \pi_{\mathbf{e}_1} U$, and so

$$h(\pi_{\mathbf{z}}\mu, A) = h(V\pi_{\mathbf{z}}\mu, A) = h(|\mathbf{z}|^2 \pi_{\mathbf{e}_1} U\mu, A) = h(\pi_{\mathbf{e}_1}\mu', A).$$

Now, as in Meiri [**13**], write:

$$\int \frac{|g_N|^2}{\varphi_n \circ U}\, d\mu = \int \frac{|g_N|^2 \circ U^{-1}}{\varphi_n}\, d\mu'.$$

From (17) we see that $\varphi_n^{1/n} \longrightarrow e^{-h'}$ μ'-a.e., and repeating the estimate we computed before, we should replace $\mathbf{v}^{\mathrm{tr}} B^k \mathbf{e}_1$ by $\mathbf{v}^{\mathrm{tr}} B^k U^{-1} \mathbf{e}_1 = \mathbf{v}^{\mathrm{tr}} B^k \mathbf{z} = c_k$. Thus the proof extends to this case as well. □

5. *Concluding remarks and questions*

1. In Theorem 1.2, when $\dim \mu = 0$, is μ atomic? This part of Furstenberg's conjecture from 1967 [**3**] is still unsolved.

2. When $S \subset \mathbb{T}^d$ is a closed bi-invariant set and $0 < \dim_H S < d$, can more be said on the structure of S? Is S necessarily a finite union of rational tori?

3. For $d = 1$, the case of bi-invariant sets was understood long before measures (cf. Furstenberg [**3**]; Rudolph [**15**]). The proof of Theorem 1.1(ii) in the scalar case relies on the well-known variational principle for entropy. However, in the non-scalar case it is based

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press

on Theorem 1.1 of Kenyon and Peres [**9**], which in turn relies on an explicit calculation due to McMullen [**12**] of the dimension of self-affine sets. This also applies to Theorem 1.2(ii). Is there a simpler way to derive these topological results?

4. Can the assumptions on A, B be relaxed, e.g. does it suffice to assume that a_i, b_i are multiplicatively independent instead of relatively prime? After seeing a previous version of this paper, J-P. Thouvenot indicated to us an argument that extends Theorem 1.2(i) to this case, using the main result of Johnson [**5**] as the basis for the induction. Together with our Proposition 1.3 this implies that Theorem 1.2(ii) holds for this case as well. However, even the one-dimensional case of Theorem 1.1 is still open when a_1 and b_1 are not relatively prime.

5. More generally, define two endomorphisms A, B to be *strongly multiplicatively independent* if for any common eigenvector v, the moduli of the corresponding eigenvalues are multiplicatively independent. We expect that Theorems 1.1 and 1.2 should be valid for all such A, B.

6. Does every bi-invariant measure have the same dimension as its closed support?

Acknowledgements. We thank Hillel Furstenberg and Dan Rudolph for helpful discussions.

References

[1] D. Berend. Multi-invariant sets on tori. *Trans. Amer. Math. Soc.* **280** (1983), 509–532.
[2] P. Billingsley. *Ergodic Theory and Information.* Wiley, New York, 1965.
[3] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. *Math. Sys. Th.* **1** (1967), 1–49.
[4] B. Host. Nombres normaux, entropie, translations. *Israel J. Math.* **91** (1995), 419–428.
[5] A. Johnson. Measures on the circle invariant under multiplication by a nonlacunary subgroup of the integers. *Israel J. Math.* **77** (1992), 211–240.
[6] A. Johnson and D. Rudolph. Convergence under $\times_p$ of $\times_q$ invariant measures on the circle. *Adv. Math.* **115** (1995), 117–140.
[7] A. Katok and R. Spatzier. Invariant measures for higher-rank hyperbolic abelian actions. *Ergod. Th. & Dynam. Sys.* **16** (1996), 751–778.
[8] R. Kenyon and Y. Peres. Intersecting random translates of invariant Cantor sets. *Invent. Math.* **104** (1991), 601–629.
[9] R. Kenyon and Y. Peres. Measures of full dimension on affine invariant sets. *Ergod. Th. & Dynam. Sys.* **16** (1996), 307–323.
[10] F. Ledrappier and L. S. Young. The metric entropy of diffeomorphisms II. *Ann. Math.* **122** (1985), 540–574.
[11] R. Lyons. On measures simultaneously 2- and 3- invariant. *Israel J. Math.* **61** (1988), 219–224.
[12] C. McMullen. The Hausdorff dimension of general Sierpinski carpets. *Nagoya Math. J.* **96** (1984), 1–9.
[13] D. Meiri. Entropy and uniform distribution of orbits in $\mathbb{T}^d$. *Israel J. Math.* **105** (1998), 155–183.
[14] W. Parry. *Topics in Ergodic Theory*. Cambridge University Press, 1981.
[15] D. J. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergod. Th. & Dynam. Sys.* **10** (1990), 395–406.
[16] P. Walters. *An Introduction to Ergodic Theory*. Springer-Verlag, Berlin, 1982.

https://doi.org/10.1017/S014338579912100X Published online by Cambridge University Press