A simultaneous version of Host’s equidistribution Theorem

Amir Algom

Abstract

Let \( \mu \) be a probability measure on \( \mathbb{R}/\mathbb{Z} \) that is ergodic under the \( \times p \) map, with positive entropy. In 1995, Host [14] showed that if \( \gcd(m, p) = 1 \) then \( \mu \) almost every point is normal in base \( m \). In 2001, Lindenstrauss [19] showed that the conclusion holds under the weaker assumption that \( p \) does not divide any power of \( m \). In 2015, Hochman and Shmerkin [13] showed that this holds in the "correct" generality, i.e. if \( m \) and \( p \) are independent. We prove a simultaneous version of this result: for \( \mu \) typical \( x \), if \( m > p \) are independent, we show that the orbit of \( (x, x) \) under \( (\times m, \times p) \) equidistributes for the product of the Lebesgue measure with \( \mu \). We also show that if \( m > n > 1 \) and \( n \) is independent of \( p \), then the orbit of \( (x, x) \) under \( (\times m, \times n) \) equidistributes for the Lebesgue measure.

1 Introduction

1.1 Background and main results

Let \( p \) be an integer greater or equal to 2. Let \( T_p \) be the \( p \)-fold map of the unit interval, \( T_p(x) = p \cdot x \mod 1 \).

Let \( m > 1 \) be an integer independent of \( p \), that is, \( \frac{\log p}{\log m} \notin \mathbb{Q} \). Henceforth, we will write \( m \not\sim p \) to indicate that \( m \) and \( p \) are independent. In 1967 Furstenberg [7] famously proved that if a closed subset of \( T := \mathbb{R}/\mathbb{Z} \) is jointly invariant under \( T_p \) and \( T_m \), then it is either finite or the entire space \( T \). A well known Conjecture of Furstenberg about a measure theoretic analogue of this result, is that the only continuous probability measure jointly invariant under \( T_p \) and \( T_m \), and ergodic under the \( \mathbb{Z}_+^2 \) action generated by these maps, is the Lebesgue measure. The best results towards this Conjecture, due to Rudolph [26] for \( p, m \) such that \( \gcd(p, m) = 1 \) and later to Johnson [17] for \( p \not\sim m \), is that it holds if in addition the measure has positive entropy with respect to the \( \mathbb{Z}_+ \) action generated by \( T_p \) (see also the earlier results of Lyons [21]).

In 1995 Host proved the following pointwise strengthening of Rudolph’s Theorem. Recall that a number \( x \in [0, 1] \) is said to be normal in base \( p \) if the sequence \( \{T_p^k x\}_{k \in \mathbb{Z}_+} \) equidistributes for the Lebesgue measure on \([0,1]\). Equivalently, the sequence of digits in the base \( p \) expansion of \( x \) has the same limiting statistics as an IID sequence of digits with uniform marginals.

**Theorem.** (Host, [14]) Let \( p, m \geq 2 \) be integers such that \( \gcd(p, m) = 1 \). Let \( \mu \) be \( T_p \) invariant ergodic measure with positive entropy. Then \( \mu \) almost every \( x \) is normal in base \( m \).

Host’s theorem can be shown to imply Rudolph’s Theorem, but is more constructive in the sense that it proves that a large collection of measures satisfy a certain regularity property. Host’s Theorem is also closely related to classical results of Cassels [3] and Schmidt [27] from around 1960,
that proved a similar result for certain Cantor-Lebesgue type measures. This was later generalized by Feldman and Smorodinsky [6] to all non-degenerate Cantor-Lebesgue measures (in fact, weakly Bernoulli) with respect to any base \( p \) (though applying to a less general class of measures, the latter results nonetheless hold for any independent integers \( p, m \geq 2 \)). We remark that the works of Meiri [23] and of Hochman and Shmerkin [13] contain excellent expositions on Host’s Theorem, and on the results of Cassels and Schmidt and some of the research that followed, respectively.

The assumption made on the integers \( p, m \) in Host’s Theorem, however, is stronger than it “should” be. Namely, it is stronger than assuming that \( p \not\sim m \). In 2001, Lindenstrauss [19] showed that the conclusion of Host’s Theorem holds under the weaker assumption that \( p \) does not divide any power of \( m \). Finally, in 2015, Hochman and Shmerkin [13] proved that Host’s Theorem holds in the “correct” generality, i.e. when \( p \not\sim m \).

Now, let \( \mu \) be a measure as in Host’s Theorem, with \( p \not\sim m \). Then, on the one hand, by the results of Hochman and Shmerkin, for \( \mu \) almost every \( x \), its orbit under \( T_m \) equidistributes for the Lebesgue measure. On the other hand, for \( \mu \) almost every \( x \), its orbit under \( T_p \) equidistributes for \( \mu \) (this is just the ergodic Theorem). The main result of this paper is that this holds simultaneously.

**Theorem 1.1.** Let \( \mu \) be a \( T_p \) invariant ergodic measure with \( \dim \mu > 0 \). Let \( m > n > 1 \) be integers such that \( m \not\sim p \).

1. If \( n = p \) then
   \[
   \frac{1}{N} \sum_{i=0}^{N-1} \delta_{(T_{m}^{i}(x), T_{n}^{i}(x))} \to \lambda \times \mu, \quad \text{for } \mu \text{ almost every } x,
   \]
   where the convergence is in the weak-* topology, and \( \lambda \) is the Lebesgue measure on \([0,1]\).

2. If \( n \not\sim p \) then
   \[
   \frac{1}{N} \sum_{i=0}^{N-1} \delta_{(T_{m}^{i}(x), T_{n}^{i}(x))} \to \lambda \times \lambda, \quad \text{for } \mu \text{ almost every } x.
   \]

Several remarks are in order. First, the assumption that \( \mu \) has positive dimension and the assumption that it has positive entropy are equivalent, so there is no discrepancy between the assumptions in Host’s Theorem and those in Theorem 1.1 (see Section 2.1 for a discussion on the dimension theory of measures). Secondly, in the second part of the Theorem we do not need that \( m \) and \( n \) are independent, only that \( m > n \). In addition, we can prove a version of Theorem 1.1 where the initial point \((x, x)\) is replaced with \((f(x), g(x))\) for \( f, g \) that are non singular affine maps of \( \mathbb{R} \) that satisfy some extra mild conditions. This is explained in Section 6.

Theorem 1.1 can also be considered as part of the following general framework. Let \( S, T \in \text{End}(T^2) \), and let \( \nu \) be an \( S \) invariant probability measure. The idea is to study the orbits \( \{T^k x\}_{k \in \mathbb{Z}_+} \) for \( \mu \) typical \( x \). In our situation,

\[
S = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}, \quad T = \begin{pmatrix} m & 0 \\ 0 & n \end{pmatrix}
\]

and the measure \( \nu = \Delta \mu \), where \( \Delta : T \to T^2 \) is the map \( \Delta(x) = (x, x) \).

Problems around this framework were studied by several authors. Notable related examples are the works of Meiri and Peres [24], and the subsequent work of Host [15]. Meiri and Peres prove a Theorem similar to ours, with the following differences:

- They work with two general diagonal endomorphisms \( S \) and \( T \), but they require that the corresponding diagonal entries \( S_{i,i} \) and \( T_{i,i} \) be larger than 1 and co-prime.
They allow for more general measures than the one dimensional measures that we work with (in Theorem 1.1 we work with measures on the diagonal of $\mathbb{T}^2$).

Host in turn has some requirements on $S$ and the measure that are more general than ours, but also requires that $\det(S)$ and $\det(T)$ be co-prime, and that for every $k$ the characteristic polynomial of $T^k$ is irreducible over $\mathbb{Q}$ (clearly this is not the case here). The results of both Host, and Meiri and Peres, extend to any $d$ dimensional torus.

Our proof of Theorem 1.1 is inspired by the work of Hochman and Shmerkin [13]. In particular, the scenery of the measure $\mu$ at typical points plays a pivotal role in our work. We devote the next Section to defining this scenery and some related notions, and to formulating the main technical tool used to prove Theorem 1.1.

1.2 On sceneries of measures and the proof of Theorem 1.1

We first recall some notions that were defined in ([13], Section 1.2). However, we remark that these notions and ideas have a long history, going back variously to Furstenberg ([8], [9]), Zähle [31], Bedford and Fisher [1], Mörters and Preiss [25], and Gavish [10]. See ([13], Section 1.2) and [11]

for some further discussions and comparisons.

For a compact metric space $X$ let $\mathcal{P}(X)$ denote the space of probability measures on $X$. Let

$$\mathcal{M}^\square = \{ \mu \in \mathcal{P}([-1, 1]) : 0 \in \text{supp}(\mu) \}. $$

For $\mu \in \mathcal{M}^\square$ and $t \in \mathbb{R}$ we define the scaled measure $S_t \mu \in \mathcal{M}^\square$ by

$$S_t \mu(E) = c \cdot \mu(e^{-t}E \cap [-1, 1]), \quad \text{where } c \text{ is a normalizing constant.}$$

For $x \in \text{supp}(\mu)$ we similarly define the translated measure by

$$\mu_x(E) = c' \cdot \mu((E + x) \cap [-1, 1]), \quad \text{where } c' \text{ is a normalizing constant.}$$

The scaling flow is the Borel $\mathbb{R}^+$ flow $S = (S_t)_{t \geq 0}$ acting on $\mathcal{M}^\square$. The scenery of $\mu$ at $x \in \text{supp}(\mu)$ is the orbit of $\mu^x$ under $S$, that is, the one parameter family of measures $\mu_{x,t} := S_t(\mu^x)$ for $t \geq 0$. Thus, the scenery of the measure at some point $x$ is what one sees as one ”zooms” into the measure.

Notice that $\mathcal{P}(\mathcal{M}^\square) \subseteq \mathcal{P}(\mathcal{P}([-1, 1]))$. As is standard in this context, we shall refer to elements of $\mathcal{P}(\mathcal{P}([-1, 1]))$ as distributions, and to elements of $\mathcal{P}(\mathbb{R})$ as measures. A measure $\mu \in \mathcal{P}(\mathbb{R})$ generates a distribution $P \in \mathcal{P}(\mathcal{P}([-1, 1]))$ at $x \in \text{supp}(\mu)$ if the scenery at $x$ equidistributes for $P$ in $\mathcal{P}(\mathcal{P}([-1, 1]))$, i.e. if

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(\mu_{x,t})dt = \int f(\nu)dP(\nu), \quad \text{for all } f \in C(\mathcal{P}([-1, 1])).$$

and $\mu$ generates $P$ if it generates $P$ at $\mu$ almost every $x$.

If $\mu$ generates $P$, then $P$ is supported on $\mathcal{M}^\square$ and is $S$-invariant ([11], Theorem 1.7). We say that $P$ is trivial if it is the distribution supported on $\delta_0 \in \mathcal{M}^\square$ - a fixed point of $S$. To an $S$-invariant distribution $P$ we associate its pure point spectrum $\Sigma(P, S)$. This set consists of all the $\alpha \in \mathbb{R}$ for which there exists a non-zero measurable function $\phi : \mathcal{M}^\square \to \mathbb{C}$ such that $\phi \circ S_t = \exp(2\pi i t \alpha)\phi$ for every $t \geq 0$, on a set of full $P$ measure. The existence of such an eigenfunction indicates that some non-trivial feature of the measures of $P$ repeats periodically under magnification by $e^{\alpha}$.

Finally, we say that a measure $\mu \in \mathcal{P}([0, 1])$ is pointwise generic under $T_n$ for a measure $\rho \in \mathcal{P}([0, 1])$ if $\mu$ almost every $x$ equidistributes for $\rho$ under $T_n$, that is,

$$\frac{1}{N} \sum_{k=0}^{N-1} f(T^k_n x) \to \int f(x)d\rho(x), \quad \forall f \in C([0, 1]).$$
We are now ready to state our second main result, which is the technical tool that shall be employed to prove Theorem 1.1.

**Theorem 1.2.** Let $\mu \in \mathcal{P}([0,1])$ and let $m > n > 1$ be integers, such that:

1. The measure $\mu$ generates a non-trivial $S$-ergodic distribution $P \in \mathcal{P}(\mathcal{P}([-1,1]))$.
2. The pure point spectrum $\Sigma(P, S)$ does not contain a non-zero integer multiple of $\frac{1}{\log m}$.
3. The measure $\mu$ is pointwise generic under $T_n$ for an ergodic and continuous measure $\rho$.

Then

$$\frac{1}{N} \sum_{i=0}^{N-1} \delta_{(T_{im}(x), T_{in}(x))} \to \lambda \times \rho, \quad \text{for } \mu \text{ almost every } x. \tag{1}$$

Notice that under assumption (3) of Theorem 1.2 the measure $\rho$ is $T_n$ invariant, so its ergodicity is with respect to this map. Theorem 1.2 together with the machinery developed by Hochman and Shmerkin in [13], Section 8 imply Theorem 1.1; this is explained in Section 5.

We end this introduction with a brief overview of the proof of Theorem 1.2. First, we note that if we only assume (1) and (2) in Theorem 1.2, then

$$\frac{1}{N} \sum_{i=0}^{N-1} \delta_{T_{im}(x)} \to \lambda, \quad \text{for } \mu \text{ almost every } x, \tag{2}$$

according to the main result of Hochman and Shmerkin [13]. This is proved by roughly following three steps: first, using the spectral condition, they show that any accumulation point $\nu$ of measures as in (2) can be represented as an integral over measures that are closely related to those drawn according to $P$. They proceed to use this representation to show that

There exists some $\epsilon > 0$ such that for any $\tau \in \mathcal{P}([0,1])$ with $\dim \tau \geq \epsilon, \dim \tau \ast \nu = 1$.

They conclude by showing that the only $T_m$ invariant measure $\nu$ satisfying the latter property is the Lebesgue measure.

Our strategy is to first show that a $T_m \times T_n$ invariant measure $\nu$, that projects to $\lambda$ and to $\rho$ in the first and second coordinate respectively, must be $\lambda \times \rho$ if it satisfies the following condition:

There exists some $\epsilon > 0$ such that for any $\tau \in \mathcal{P}([0,1])$ with $\dim \tau \geq \epsilon, \dim \tau \ast P_1 \nu_y = 1$,

for $\rho$ almost every $y$, where $\nu_y$ is the conditional measure of $\nu$ on the fiber $\{ (x, z) : z = y \}$, and $P_1(x, y) = x$. This is Claim 3.4 in Section 3. Afterwards, we show that this property holds for all the accumulation points of the measures from (1). This is done via a corresponding integral representation, see Claim 4.1 in Section 4.

**Notation** We shall use the letter $\lambda$ to indicate both the Lebesgue measure on $[0,1]$ and the Lebesgue measure on $\mathbb{T}$. Which is meant will be clear from context. Also, whenever we have a finite product space, we denote by $P_i$ the projection to the $i$-th coordinate.

**Acknowledgements** I would like to thank Mike Hochman for suggesting the problem studied in this paper to me, and for his various insightful comments on earlier versions of this manuscript. I would also like to thank Shai Evra for some helpful conversations related to this paper.
2 Preliminaries

2.1 Dimension theory of measures, and their Fourier coefficients

For a Borel set $A$ in some metric space $X$, we denote by $\dim A$ its Hausdorff dimension, and by $\dim_P A$ its packing dimension (see Falconer’s book \[5\] for an exposition on these concepts). Now, let $\mu \in \mathcal{P}(X)$. The (lower) Hausdorff dimension of the measure $\mu$ is defined as

$$\dim \mu := \inf \{\dim A : \mu(A) > 0, \ A \text{ is Borel}\},$$

and the upper Hausdorff dimension of the measure $\mu$ is defined as

$$\overline{\dim} \mu := \inf \{\dim A : \mu(A) = 1, \ A \text{ is Borel}\},$$

The (upper) packing dimension of the measure $\mu$ is defined as

$$\dim_P \mu = \inf \{\dim_P A : \mu(A) = 1, \ A \text{ is Borel}\}.$$

An alternative characterization of the dimension of $\mu$ that we shall often use is given in terms of their local dimensions: For every $x \in \text{supp}(\mu)$ we define the local (pointwise) dimension of $\mu$ at $x$ as

$$\dim(\mu, x) = \liminf_{r \to 0} \frac{\log \mu(B(x, r))}{\log r},$$

where $B(x, r)$ denotes the closed ball of radius $r$ about $x$. The Hausdorff dimension of $\mu$ is equal to

$$\dim \mu = \text{ess-inf}_{x \sim \mu} \dim(\mu, x),$$

and the upper Hausdorff dimension of $\mu$ is equal to

$$\overline{\dim} \mu = \text{ess-sup}_{x \sim \mu} \dim(\mu, x).$$

see e.g. \[4\]. If $\dim(\mu, x)$ exists as a limit at almost every point, and is constant almost surely, we shall say that the measure $\mu$ is exact dimensional. In this case, most metric definitions of the dimension of $\mu$ coincide (e.g. lower and upper Hausdorff dimension and packing dimension).

Let us now collect some known facts regarding dimension theory of measures:

**Proposition 2.1.** 1. Let $\mu \in \mathcal{P}(\mathbb{R}^d)$ and suppose that there is a distribution $Q \in \mathcal{P}(\mathcal{P}(\mathbb{R}^d))$ such that

$$\mu = \int \nu dQ(\nu).$$

Then

$$\dim \mu \geq \text{ess-inf}_{\nu \sim Q} \dim \nu$$

If $\mu = p_1 \mu_1 + p_2 \mu_2$ where $\mu_1, \mu_2 \in \mathcal{P}(\mathbb{R}^d)$ and $(p_1, p_2)$ is any probability vector, then

$$\dim \mu = \min\{\dim \mu_1, \dim \mu_2\}.$$

2. Let $f : X \to Y$ be a Lipschitz map between complete metric spaces. Then for any $\mu \in \mathcal{P}(X)$,

$$\dim f \mu \leq \dim \mu$$

with an equality if $f$ is locally bi-Lipschitz.
3. Let \( \mu \in \mathcal{P}(\mathbb{T}) \) be exact dimensional, and \( \nu \in \mathcal{P}(\mathbb{T}) \) be a measure supported on finitely many atoms. Then \( \dim \mu \ast \nu = \dim \mu \), and moreover, \( \mu \ast \nu \) is exact dimensional.

The next Lemma is essentially Lemma 3.5 in [13], with a minor modification which follows e.g. from Lemma 6.13 in [11].

**Lemma 2.2.** ([13], Lemma 3.5) Let \( \mu \in \mathcal{P}(\mathbb{R}^2) \).

1. Suppose that for \( P_2 \mu \) almost every \( y \), \( \dim \mu_y \geq \alpha \) (where \( \mu_y \) is the conditional measure on the fiber \( \{ (x, z) : z = y \} \)). Then \( \dim \mu \geq \dim P_2 \mu + \alpha \).

2. For an upper bound, we have \( \dim \mu \leq \dim P_1 \mu + \dim P_2 \mu \).

We end this section with a brief discussion of the Fourier coefficients of measures on \( \mathbb{T}^d \). These are defined as follows. First, given \( \mu \in \mathcal{P}(\mathbb{T}^d) \) we define for any \( k \in \mathbb{Z}^d \) the corresponding Fourier coefficient by

\[
\hat{\mu}(k) := \int e^{2\pi ik \cdot x} d\mu(x).
\]

The following relations are easily verified for two measures \( \mu, \nu \in \mathcal{P}(\mathbb{T}) \):

\[
\widehat{\mu \ast \nu}(k) = \hat{\mu}(k) \ast \hat{\nu}(k), \quad k \in \mathbb{Z}.
\]

\[
\widehat{\mu \times \nu}(k, j) = \hat{\mu}(k) \ast \hat{\nu}(j), \quad (k, j) \in \mathbb{Z}^2.
\]

The following Lemma is standard:

**Lemma 2.3.** ([22], Section 3.10 ) Let \( \mu, \nu \in \mathcal{P}(\mathbb{T}^2) \). If \( \hat{\mu}(k) = \hat{\nu}(k) \) for all \( k \in \mathbb{Z}^2 \) then \( \mu = \nu \).

Finally, let \( m \geq 2 \) and let \( \mu \) be the Cantor-Lebesgue measure corresponding to the non-degenerate probability vector \((p_0, ..., p_{m-1})\). That is, \( \mu \) is the distribution of the Random sum \( \sum_{k=1}^{m} \frac{X_k}{m^k} \), where \( X_k \) are IID random variables with \( P(X_k = i) = p_i \). It is a well known fact that for every \( k \in \mathbb{Z} \),

\[
\hat{\mu}(k) = \prod_{j=1}^{\infty} \left( \sum_{u=0}^{m-1} p_u \exp(2\pi i u \frac{k}{m^j}) \right).
\]

### 2.2 Dimension theory of invariant measures

#### 2.2.1 Some notions from ergodic theory

In this paper, a dynamical system is a quadruple \((X, B, T, \mu)\), where \( X \) is a compact metric space, \( B \) is the Borel sigma algebra, and \( T : X \rightarrow X \) is a measure preserving map, i.e. \( T \) is Borel measurable and \( T \mu = \mu \). Since we always work with the Borel sigma-algebra, we shall usually just write \((X, T, \mu)\). For example one may consider \( X = \mathbb{T} \), the Borel map \( T_p \) for some \( p \geq 2 \), and some Cantor-Lebesgue measure with respect to base \( p \).

A dynamical system is ergodic if and only if the only invariant sets are trivial. That is, if \( B \in B \) satisfies \( T^{-1}(B) = B \) then \( \mu(B) = 0 \) or \( \mu(B) = 1 \). A dynamical system is called weakly mixing if for any ergodic dynamical system \((Y, S, \nu)\), the product system \((X \times Y, T \times S, \mu \times \nu)\) is also ergodic. In particular, weakly mixing systems are ergodic. Moreover, If both \((X, T, \mu)\) and \((Y, S, \nu)\) are weakly mixing, then their product system is also weakly mixing. A class of examples of weakly mixing systems is given \((\mathbb{T}, T_p, \mu)\) where \( \mu \) is a Cantor-Lebesgue measure with respect to base \( p \).

We will have occasion to use the ergodic decomposition Theorem: Let \((X, T, \mu)\) be a dynamical system. Then there is a map \( X \rightarrow \mathcal{P}(X) \), denoted by \( \mu \mapsto \mu^x \), such that:
1. The map $x \mapsto \mu^x$ is measurable with respect to the sub-sigma algebra $\mathcal{I}$ of $T$ invariant sets.

2. $\mu = \int \mu^x d\mu(x)$

3. For $\mu$ almost every $x$, $\mu^x$ is $T$ invariant and ergodic. The measure $\mu^x$ is called the ergodic component of $x$.

Finally, we shall say that a point $x \in X$ is generic with respect to $\mu$ if

$$\frac{1}{N} \sum_{i=0}^{N-1} \delta_{T^i x} \to \mu,$$

where $\delta_y$ is the Dirac measure on $y \in X$, in the weak-* topology. By the ergodic Theorem, if $\mu$ is ergodic then $\mu$ a.e. $x$ is generic for $\mu$.

2.2.2 Dimension theory of invariant measures

Recall that in general, if $\mu \in \mathcal{P}(X)$ is a $T$ invariant measure, we may define its entropy with respect to $T$, a quantity that we shall denote by $h(\mu, T)$. As there is an abundance of excellent texts on entropy theory (e.g. [30]), we omit a discussion on entropy here. We now restrict our attention to dynamical systems of the form $(\mathbb{T}, \mu, T_p)$ or $(\mathbb{T}^2, \mu, T_m \times T_n)$, where we always assume that $m > n > 1$. In the one dimensional case, the dimension of $\mu$ may be computed via the entropies of its ergodic components:

**Theorem 2.4.** ([20], Theorem 9.1) Let $\mu \in \mathcal{P}(\mathbb{T})$ be a $T_p$ invariant and ergodic measure. Then $\mu$ is exact dimensional and

$$\dim \mu = \frac{h(\mu, T_p)}{\log p}.$$  

In general, if $\mu \in \mathcal{P}(\mathbb{T})$ is a $T_p$ invariant measure with ergodic decomposition $\mu = \int \mu^x d\mu(x)$, then

$$\dim \mu = \text{ess-inf}_{x \sim \mu} \dim \mu^x$$  \hspace{1cm} (8)

and

$$\overline{\dim \mu} = \text{ess-sup}_{x \sim \mu} \dim \mu^x$$  \hspace{1cm} (9)

The situation for dynamical systems of the form $(\mathbb{T}^2, \mu, T_m \times T_n)$ is more complicated. This may be attributed to the fact that the map $T_m \times T_n$ is not conformal. There is, however, a way to compute the dimension of $\mu$ in this situation via entropy theory, using a suitable version of the Ledrappier-Young formula. This was first done by Kenyon and Peres in [18] for ergodic measures. The general case may be treated using similar methods, as observed by Meiri and Peres in ([24], Lemma 3.1).

**Theorem 2.5.** ([24]) Let $\mu \in \mathcal{P}(\mathbb{T}^2)$ be a $T_m \times T_n$ invariant measure. Then for $\mu$ almost every $x$ the local dimension $\dim(\mu, x)$ exists as a limit and

$$\dim(\mu, x) = \frac{h(\mu^x, T_m \times T_n) - h((P_2 \mu)^{P_2 x}, T_n)}{\log m} + \frac{h((P_2 \mu)^{P_2 x}, T_n)}{\log n}$$

where $\mu^x$ and $(P_2 \mu)^{P_2 x}$ denote the corresponding ergodic components of $\mu$, and of $P_2 \mu$, respectively.

Finally, we will require the following result of Meiri, Lindenstrauss and Peres from [20]:

7
Theorem 2.6. Let $\mu \in \mathcal{P}(\mathbb{T})$ be a $T_\mu$ invariant weakly mixing measure, such that $\dim \mu > 0$. Let $\mu^k$ denote the convolution of $\mu$ with itself $k$-times. Then

$$\dim(\mu^k) \to 1, \quad \text{monotonically as } k \to \infty$$

We remark that we have only cited a special case of this result. Indeed, Meiri, Lindenstrauss and Peres deal with the growth of the entropy of more general convolutions of $T_\mu$ ergodic measures. We refer the reader to [20] for the full statement.

2.3 Relating the distribution of orbits to the measure

Let $X$ be a compact metric space, $T : X \to X$ a Borel measurable map, and let $\mu, \nu \in \mathcal{P}(X)$. Following Hochman and Shmerkin [13], we shall say that $\mu$ is pointwise generic for $\nu$ under $T$ if $\mu$ almost every $x$ equidistributes for $\nu$ under $T$, that is,

$$\frac{1}{N} \sum_{k=0}^{N-1} f(T^k x) \to \int f \, d\nu, \quad \forall f \in C(X).$$

This notion is closely related to the main results of this paper. Indeed, let $X = \mathbb{T}^2$, $T = T_m \times T_p$ for $m > p > 1$ and $m \not\sim p$, and $\alpha$ be the pushforward of a $T_p$ invariant ergodic positive dimensional measure $\mu \in \mathcal{P}(\mathbb{T})$ to the diagonal of $\mathbb{T}^2$. Then Theorem 2.6 part (1) for example may be stated as "$\alpha$ is pointwise generic for $\lambda \times \mu$ under $T$".

In [13], the authors obtain a criteria for this to occur, one that shall play a central role in this paper as well. We now recall its formulation. Let $A$ be a finite partition of $X$, and for every $i \in \mathbb{N} \cup \{0\}$ let $T^i A = \{T^{-i} A : A \in \mathcal{A}\}$. Let $A^k = \bigvee_{i=0}^{k-1} T^i A$ denote the coarsest common refinement of $A, T^1 A, \ldots, T^{k-1} A$. Now, if the smallest sigma algebra that contains $A^k$ for all $k$ is the Borel sigma algebra, we say that $A$ is a generator for $T$. We say that $A$ is a topological generator if $\sup\{\text{diam} A : A \in A^k\} \to 0$ as $k \to \infty$. A topological generator is clearly a generator.

Let us give two examples of topological generators that shall be used in this paper: for every $p \in \mathbb{N}$ let $D_p$ be the $p$-adic partition of $\mathbb{T}$ (and of $\mathbb{R}$), that is,

$$D_p = \left\{ \left( \frac{z}{p}, \frac{z+1}{p} \right) : \; z \in \mathbb{Z} \right\}.$$ 

Then, under the map $T_p$, we see that

$$D_p^k = D_{p^k} = \left\{ \left( \frac{z}{p^k}, \frac{z+1}{p^k} \right) : \; z \in \mathbb{Z} \right\}.$$ 

It is thus easy to see that $D_p$ is a generator for $T_p$. Similarly, if $m > n$ then the partition $D_m \times D_n$ of $\mathbb{T}^2$ is a generator under $T_m \times T_n$.

Finally, in general, for every $k \geq 1$ and $x \in X$, let $A^k(x)$ denote the unique element of $A^k$ that contains $x$. Given $\mu \in \mathcal{P}(X)$ and $x \in X$ such that $\mu(A^k(x)) > 0$, let

$$\mu_{A^k(x)} = c \cdot T^k(\mu|_{A^k(x)}), \quad \text{where } c = \mu(A^k(x))^{-1},$$

which is well defined almost surely.

Theorem 2.7. (13, Theorem 2.1) Let $T : X \to X$ be a Borel measurable map of a compact metric space, $\mu \in \mathcal{P}(X)$ and $A$ a generating partition. Then for $\mu$ almost every $x$, if $x$ equidistributes for $\nu \in \mathcal{P}(X)$ along some $N_k \to \infty$, and if $\nu(\partial A) = 0$ for all $A \in A^k$, $k \in \mathbb{N}$, then

$$\nu = \lim_{k \to \infty} \frac{1}{N_k} \sum_{k=0}^{N_k-1} \mu_{A^k(x)}, \quad \text{weak-* in } \mathcal{P}(X).$$
A crucial ingredient in our application of Theorem 2.7 is the following Claim. Let $m > n$, and define for every $k \in \mathbb{N}$

$$A_k = \{x \in \mathbb{R} : \mathcal{D}_{mk}(x) \not\subseteq \mathcal{D}_{nk}(x)\}$$

Also, recall that the density of a sequence $S \subseteq \mathbb{N}$ (if it exists) is the limit of the sequence $\frac{|S\cap[1,N]|}{N}$ as $N \to \infty$. If the limit does not exist, the corresponding lim sup is called the upper density of $S$.

**Claim 2.8.** Suppose that $\mu \in \mathcal{P}([0,1])$ is a measure that is pointwise generic under $T_n$ for a continuous measure $\rho$. Then for $\mu$ almost every $x$, if $x \in \limsup A_k$ and $\{n_k\}$ represents the times when $x \in A_{n_k}$, then the density of $\{n_k\}$ is zero.

**Proof.** Choose $x \sim \mu$, and if $x \in \limsup A_k$ let $\{n_k\}$ be the sequence as in the statement of the Claim. Let $\epsilon > 0$. We will show that the upper density of $\{n_k\}$ is at most $\epsilon$. First, since $\rho$ is a continuous measure, there exists some $\delta > 0$ such that $\rho(B(0,\delta)) < \epsilon$, where $B(0,\delta)$ is the ball about 0 in $T$. By our assumption that $\mu$ is pointwise generic under $T_n$ for $\rho$, and since $\rho$ is a continuous measure,

$$V_\delta = \{i \mid T^i_n(x) \in B(0,\delta)\}$$

has density $\rho(B(0,\delta)) < \epsilon$.

Now, let us decompose our sequence

$$\{n_k\} = (\{n_k\} \cap V_\delta) \cup (\{n_k\} \cap (\mathbb{N} \setminus V_\delta)).$$

Then the upper density of $\{n_k\} \cap V_\delta$ is at most $\epsilon$. We now show that the density of the sequence $\{\ell_k\} := \{n_k\} \cap (\mathbb{N} \setminus V_\delta)$ is 0. In fact, we will show that this is a finite sequence.

Indeed, let $K > \frac{\log \delta}{\log m}$. We claim that $\{\ell_k\} \subseteq [0,K]$. Assume towards a contradiction that there exists some $q > K$ such that $\ell_k = q$ for some $k$. Then there is a unique $n^q$-adic number $a$ (an endpoint of an $\mathcal{D}_{n^q}$ cell) such that $a \in \mathcal{D}_{m^s}(x)$. Write $a = \frac{s}{n^q}$ for some integer $s$. Then we have

$$|x - \frac{s}{n^q}| \leq \frac{1}{m^q},$$

which implies that $T^q_n(x) \in B(0,\frac{n^q}{m^q}) \subset B(0,\delta)$, by the choice of $K$. Thus, $q \in V_\delta$, contradicting the choice of the sequence $\{\ell_k\}$. Thus, $\{\ell_k\} \subseteq [0,K]$, which is sufficient for us.

We will also require the following Lemma.

**Lemma 2.9.** Let $x \in [0,1]$ be such that it equidistributes for a continuous measure $\rho$ under $T_n$. Let $D \subset [0,1]$ be some interval. Let $\{n_k\}$ be the sequence of times when $T^n_{nk}x \notin D$ but $d(T^n_{nk}(x),\partial D) \leq (\frac{n}{m})^{nk}$. Then the density of the sequence $\{n_k\}$ is 0.

**Proof.** Let $\epsilon > 0$. Since $\rho$ is continuous, there exists some $\delta > 0$ such that

$$\rho(\{y : d(y,\partial D) \leq \delta\}) < \epsilon.$$

Let

$$V_\delta = \{k \mid T^k_n x \in \{y : d(y,\partial D) \leq \delta\}\}.$$ 

Then by our assumption on $x$, the density of $V_\delta$ is at most $\epsilon$. However, the sequence $\{n_k\} \subseteq V_\delta$, apart from maybe finitely many indices. It follows that the upper density of $\{n_k\}$ is at most the density of $V_\delta$, and therefore is at most $\epsilon$. This proves the Lemma. □
2.4 Ergodic fractal distributions

Recall the definitions introduced in Section 1.2. In this Section we discuss some other related results of [13] that we shall require. First, we cite a result about the implication of not having some element $t_0 > 0$ in the pure point spectrum of a distribution generated by a measure.

**Proposition 2.10.** ([13], Section 4) Suppose that $\mu$ generates an $S$-ergodic distribution $P$ and that no non-zero integer multiple of $t_0 > 0$ is in $\Sigma(P,S)$. Then $P$ is $t_0$-generated by $\mu$ at almost every $x$, i.e. the sequence $\{\mu_{x,kt_0}\}_{k=0}^{\infty}$ equidistributes for $P$.

The next result says that distributions $P \in \mathcal{P}([0,1])$ that are generated by a given measure $\mu$ have some additional invariance properties:

**Theorem 2.11.** ([13], Theorem 4.7) Suppose that $\mu$ generates an $S$-invariant distribution $P$. Then $P$ is supported on $M^{\square}$ and satisfies the $S$-quasi-Palm property: for every Borel set $B \subseteq M^{\square}$, $P(B) = 1$ if and only if for every $t > 0$, $P$ almost every measure $\eta$ satisfies that $\eta_{x,t} \in B$ for $\eta$ almost every $x$ such that $[x-e^{-t}, x+e^{-t}] \subseteq [-1,1]$.

We shall refer henceforth to $S$-ergodic distributions $P$ supported on $M^{\square}$ that satisfy the conclusion of Theorem 2.11 as EFD’s (Ergodic Fractal Distributions), a term coined by Hochman in [11]. The next Proposition says that typical measures with respect to a non-trivial EFD have positive dimension (recall the definition of non-triviality in this situation from Section 1.2):

**Proposition 2.12.** ([13], Proposition 4.12) Let $P$ be an EFD. Then there exists some $\delta \geq 0$ such that $P$ almost every $\nu$ has $\dim \nu = \delta$. If $P$ is non-trivial then $\delta > 0$.

We will also need to know that $P$-typical measures are not ”one sided at small scales”

**Proposition 2.13.** ([13], Proposition 4.13) Let $P$ be an EFD. For every $\rho > 0$, for $P$ almost every $\nu$, we have $\inf \nu(I) > 0$, where $I \subseteq [-1,1]$ ranges over closed intervals of length $\rho$ containing $0$.

The next Proposition follows from the $S$-invariance of EFD’s, and from a Theorem of Hunt and Kaloshin [16]:

**Proposition 2.14.** ([13], Lemma 5.8) Let $P$ be a non trivial EFD such that $P$ typical measures have dimension $\delta > 0$. Let $\tau \in \mathcal{P}(\mathbb{R})$ be such that $\dim \tau \geq 1 - \delta$. Then $\dim \tau * \nu = 1$ for $P$ almost every $\nu$.

Finally, the next Proposition shows that ergodic $T_p$ invariant measures of positive dimension generate non-trivial EFD’s:

**Theorem 2.15.** ([12] Let $\mu \in \mathcal{P}([0,1])$ be a $T_p$ invariant ergodic measure with $\dim \mu > 0$. Then $\mu$ generates a non-trivial $S$ ergodic distribution $P$ (which is an EFD by Theorem 2.11).

Let $m \not\sim p$. We remark that while non-degenerate Cantor-Lebesgue measures with respect to base $p$ do generate EFD’s $P$ such that $\frac{k}{\log m} \notin \Sigma(P,S)$ for every non zero integer $k$, this is not true in general. Thus, in order to deduce Theorem 1.1 from Theorem 1.2, we shall require some additional machinery developed by Hochman and Shmerkin in [13] for a similar purpose. This is discussed in Section 5.
3 Some properties of \((\times m, \times n)\) invariant measures

Throughout this section we fix integers \(m > n > 1\). We begin with an elementary Lemma from entropy theory. Recall that we denote the coordinate projections by \(P_1, P_2\).

**Lemma 3.1.** Let \(\alpha \in \mathcal{P}(\mathbb{T}^2)\) be a \(T_m \times T_n\) invariant measure such that \(P_2\alpha = \rho\). If

\[
h(T_m \times T_n, \alpha) = \log m + h(T_n, \rho)
\]

then \(\alpha = \lambda \times \rho\).

**Proof.** Let \(\mathcal{E}\) be the invariant sigma algebra that corresponds to the second coordinate of \(\mathbb{T}^2\). Then, by the Abramov-Rokhlin Lemma (see [2] for the non-invertible case),

\[
h(T_m \times T_n, \alpha) = h(T_m \times T_n, \alpha|\mathcal{E}) + h(T_n, \rho).
\]

Combining this with our condition, we see that

\[
h(T_m \times T_n, \alpha|\mathcal{E}) = \log m.
\]

Recall that the partition \(A = D_m \times D_n\) is a generating partition of \(\mathbb{T}^2\) (see Section 2.3). Then it follows from Fekete’s Lemma and the Kolmogorov-Sinai Theorem that

\[
\inf_k \frac{1}{k} H_{\alpha}(\bigvee_{i=0}^{k-1} (T_m \times T_n)^i A|\mathcal{E}) = h(T_m \times T_n, A|\mathcal{E}) = h(T_m \times T_n, \alpha|\mathcal{E}) = \log m.
\]

As \(\log m\) is also an upper bound for the sequence \(\{\frac{1}{k} H_{\alpha}(\bigvee_{i=0}^{k-1} (T_m \times T_n)^i A|\mathcal{E})\}\), we find that for every \(k \in \mathbb{N}\),

\[
\frac{1}{k} H_{\alpha}(\bigvee_{i=0}^{k-1} (T_m \times T_n)^i A|\mathcal{E}) = \log m.
\]

So, by the formula for conditional entropy as average of the conditional measures \(\{\alpha_x^\mathcal{E}\}\),

\[
\log m^k = H_{\alpha}(\bigvee_{i=0}^{k-1} (T_m \times T_n)^i A|\mathcal{E}) = \int H_{\alpha_x^\mathcal{E}}(\bigvee_{i=0}^{k-1} (T_m \times T_n)^i A) d\rho(x) = \int H_{\alpha_x^\mathcal{E}}(\bigvee_{i=0}^{k-1} T_m^i D_m) d\rho(x),
\]

where the partition in the last term on the RHS should be understood as the corresponding partition on the fiber \([0, 1] \times \{P_2(x)\}\). We also have \(H_{\alpha_x^\mathcal{E}}(\bigvee_{i=0}^{k-1} T_m^i D_m) \leq \log m^k\) almost surely, since \(\bigvee_{i=0}^{k-1} T_m^i D_m\) has \(m^k\) atoms. Therefore,

\[
H_{\alpha_x^\mathcal{E}}(\bigvee_{i=0}^{k-1} T_m^i D_m) = \log m^k
\]

almost surely. Such an equality is possible only if \(\alpha_x^\mathcal{E}\) is the uniform measure on \(D_m\). It follows that almost surely the measure \(\alpha_x^\mathcal{E}\) is the uniform measure on \(D_m\) for every \(k\). By the Kolmogorov consistency Theorem, \(\alpha_x^\mathcal{E} = \lambda\) almost surely. Since \(\alpha = \int \alpha_x^\mathcal{E} d\rho(x)\), this proves the result. \(\square\)

**Claim 3.2.** Let \(\theta \in \mathcal{P}(\mathbb{T}^2)\) be a \(T_m \times T_n\) invariant measure such that \(P_2\theta = \rho\) is exact dimensional. If \(\dim \theta = 1 + \dim \rho\) then \(\theta = \lambda \times \rho\).
Proof. By equation (3), and by Theorem 2.5

\[ 1 + \dim \rho = \dim \theta = \text{ess-inf}_{x \sim \theta} \dim(\theta, x) = \text{ess-inf}_{x \sim \theta} \left( \frac{h(\theta^x, T_m \times T_n) - h((P_2 \theta)^{P_2 x}, T_n)}{\log m} + \frac{h((P_2 \theta)^{P_2 x}, T_n)}{\log n} \right) + \frac{h(P^2 \theta, T_n)}{\log n} \]

(10)

(recall that \( \theta^x \) and \( \rho^{P_2 x} \) denote the corresponding ergodic components of \( \theta \) and of \( \rho \), respectively).

Now, by Theorem 2.4, and since \( \rho \) has exact dimension

\[ \text{ess-sup}_{x \sim \rho} \dim (\rho^{P_2 x}) = \text{ess-sup}_{y \sim \rho} \dim (\rho^y) = \dim \rho = \text{ess-inf}_{x \sim \theta} \dim (\rho^{P_2 x}). \]

So, for \( \theta \) almost every \( x \) we have

\[ \frac{h(\rho^{P_2 x}, T_n)}{\log n} = \dim \rho^{P_2 x} = \dim \rho. \]

(11)

Combining (11) with (10), we find that

\[ 1 = \text{ess-inf}_{x \sim \theta} \frac{h(\theta^x, T_m \times T_n) - h(\rho^{P_2 x}, T_n)}{\log m}. \]

(12)

Therefore, by (12), the formula for entropy as an average over ergodic components, the Abramov-Rokhlin Lemma, and the formula for entropy as the average of conditional measures (as in Lemma 3.1), we have

\[ \log m \leq \int \left( h(\theta^x, T_m \times T_n) - h(\rho^{P_2 x}, T_n) \right) d\theta(x) \]

\[ = \int h(\theta^x, T_m \times T_n)d\theta(x) - \int h(\rho^{P_2 x}, T_n)d\theta(x) \]

\[ = h(\theta, T_m \times T_n) - \int h(\rho^y, T_n)d\rho(y) \]

\[ = h(\theta, T_m \times T_n) - h(\rho, T_n) \]

\[ = h(\theta, T_m \times T_n | \mathcal{E}) \]

\[ \leq \log m \]

where \( \mathcal{E} \) be the invariant sigma algebra that corresponds to the second coordinate of \( \mathbb{T}^2 \). Thus, we have that \( \theta \) almost surely,

\[ \frac{h(\theta^x, T_m \times T_n) - h(\rho^{P_2 x}, T_n)}{\log m} = 1, \quad \text{and} \quad h(\theta, T_m \times T_n | \mathcal{E}) = \log m. \]

(13)

Now, (13) and the Abramov-Rokhlin Lemma imply that \( \theta \) almost surely

\[ \log m + h(\rho^{P_2 x}, T_n) = h(\theta^x, T_m \times T_n) = h(\theta^x, T_m \times T_n | \mathcal{E}) + h(P_2 \theta^x, T_n). \]

(14)

By (13) and the formula for entropy and convex combinations,

\[ \log m = h(\theta, T_m \times T_n | \mathcal{E}) = \int h(\theta^x, T_m \times T_n | \mathcal{E})d\theta(x). \]
Since \(0 \leq h(\theta^x, T_m \times T_n | \mathcal{E}) \leq \log m\) almost surely, we must have \(h(\theta^x, T_m \times T_n | \mathcal{E}) = \log m\) almost surely. By this equality and (14), we see that for \(\theta\) almost every \(x\),
\[
h(P_2 \theta^x, T_n) = h(\rho^{P_2x}, T_n).
\]
Finally, by (15) and (14),
\[
h(\theta^x, T_m \times T_n) = \log m + h(P_2 \theta^x, T_n) \quad \text{for almost every }x.
\]
By Lemma 3.1, almost every ergodic component \(\theta^x\) equals \(\lambda \times P_2 \theta^x\). Thus,
\[
\theta = \int \theta^x d\theta(x) = \int \lambda \times P_2 \theta^x d\theta(x) = \lambda \times \left( \int P_2 \theta^x d\theta(x) \right) = \lambda \times P_2 \left( \int \theta^x d\theta(x) \right) = \lambda \times \rho.
\]

Next, we make a short digression to discuss the relation between the conditional measures of a convolution of measures, and the conditional measures of the individual measures convolved, in some special cases. In the following, the convolution of the two measures on the unit square \([0,1]^2\) is understood to take place in \(\mathbb{R}^2\). For a measure \(\nu \in \mathcal{P}(\mathbb{R}^2)\), Let \(\nu = \int \nu_y dP_2 \nu(y)\) be the disintegration of \(\nu\) with respect to the projection \(P_2\).

**Claim 3.3.** Let \(\theta, \nu \in \mathcal{P}([0,1]^2)\) be two measure such that \(\theta = \tau \times \alpha\), where the measure \(\alpha\) is a convex combination of finitely many atomic measures. Then for \(P_2(\nu \ast \theta)\) almost every \(z\), the conditional measure \((\nu \ast \theta)_z\) with respect to the projection \(P_2\) is a finite convex combination of measures of the form \(\nu_{z-z_i} \ast (\tau \times \delta_{z_i})\), where \(z_i\) is an atom of \(\alpha\) and \(\nu_{z-z_i}\) is a conditional measure of \(\nu\) with respect to the projection \(P_2\).

**Proof.** If \(\alpha = \delta_y\) for some \(y\) then the result is straightforward. For the general case, notice that if \(\theta = \tau \times \alpha\) and \(\alpha = \sum p_i \delta_{z_i}\) then by the linearity of both convolution and of taking product measures,
\[
(\nu \ast \theta)_z = \sum p_i \nu \ast (\tau \times (\delta_{z_i})) = \sum p_i \nu \ast (\tau \times \delta_{z_i}).
\]
In general, if \(\mu = \mu_1 \cdot p_1 + \mu_2 \cdot p_2\) is a convex combination of probability measures and \(\mathcal{C}\) is some sigma algebra, then the following holds almost surely for every \(f \in L^1:\)
\[
E_{\mu}(f|\mathcal{C}) = p_1 \cdot E_{\mu_1}(f|\mathcal{C}) \cdot \frac{d\mu_1}{d\mu} + p_2 \cdot E_{\mu_2}(f|\mathcal{C}) \cdot \frac{d\mu_2}{d\mu}.
\]
We remark that in the above equation, the Radon-Nikodym derivatives \(\frac{d\mu}{d\mu}\) in fact stand for the Radon-Nikodym derivatives when the measures are restricted to the sigma-algebra \(\mathcal{C}\), i.e. \(\frac{d\mu|\mathcal{C}}{d\mu}\). However, we suppress this in our notation. So, for \(B_2\) the Borel sigma algebra on the \(y\)-axis, for every \(f \in L^1\) and for almost every \(z\)
\[
\int f d(\nu \ast \theta)_z = E_{\nu \ast \theta}(f|P_2^{-1}B_2)(z)
\]
\[
= \sum_i p_i \cdot E_{\nu \ast (\tau \times \delta_{z_i})}(f|P_2^{-1}B_2)(z) \cdot \frac{d\nu \ast (\tau \times \delta_{z_i})}{d\nu \ast \theta}(z)
\]
\[
= \sum_i p_i \cdot \int f d(\nu \ast (\tau \times \delta_{z_i}))(z) \cdot \frac{d\nu \ast (\tau \times \delta_{z_i})}{d\nu \ast \theta}(z)
\]
\[
= \sum_i p_i \cdot \int f d(\nu_{z-z_i} \ast (\tau \times \delta_{z_i}))(z) \cdot \frac{d\nu \ast (\tau \times \delta_{z_i})}{d\nu \ast \theta}(z)
\]
\[
= \int f d\left( \sum_i (\nu_{z-z_i} \ast (\tau \times \delta_{z_i})) \cdot p_i \cdot \frac{d\nu \ast (\tau \times \delta_{z_i})}{d\nu \ast \theta}(z) \right)
\]

13
It follows that almost surely,

$$ (\nu * \theta)_z = \sum_i \nu_{z-y_i} \ast (\tau \times \delta_{z_i}) \cdot p_i \cdot \frac{d\nu \ast (\tau \times \delta_{z_i})}{d\nu \ast \theta}(z) $$

The following Claim, which forms the main result of this section, is also the key for our argument.

**Claim 3.4.** Let $\nu \in \mathcal{P}([0,1]^2)$ be a $T_m \times T_n$ invariant measure such that:

1. We have $P_2 \nu = \rho$, where $\rho$ is a continuous ergodic measure, and $P_1 \nu = \lambda$.
2. There exists some $\delta > 0$ such that:

   For every probability measure $\tau \in \mathcal{P}([0,1])$ with $\dim \tau \geq 1 - \delta$, for $\rho$ almost every $y$, we have $\dim \tau \ast P_1 \nu_y = 1$.

Then $\nu = \lambda \times \rho$.

**Proof.** Suppose towards a contradiction that $\nu \neq \lambda \times \rho$. Let us first identify $\nu$ with the corresponding measure on $\mathbb{T}^2$ (i.e. we project $\nu$ to $\mathbb{T}^2$ but we keep the notation $\nu$), which cannot be $\lambda \times \rho$ either. Then, by Lemma 2.3 there exists $(i,j) \in \mathbb{Z}^2 \setminus \{(0,0)\}$ such that

$$ \hat{\nu}(i,j) \neq \lambda \times \rho(i,j). $$

Now, as $P_2 \nu = \rho$ and $P_1 \nu = \lambda$ we must have $i,j \neq 0$, since if e.g. $i = 0$ then, using (6),

$$ \hat{\nu}(0,j) = P_2 \nu(j) = \hat{\rho}(j) = 1 \cdot \hat{\rho}(j) = \hat{\lambda}(0) \hat{\rho}(j) = \hat{\lambda} \times \rho(0,j) $$

a contradiction. Thus, we may assume both $i,j \neq 0$, and since $\hat{\lambda}(i) = 0$ we have $\hat{\nu}(i,j) \neq 0$ by (6).

Now, let $k \in \mathbb{N}$ be such that $2|j| < n^k + 1$. We construct two measures $\tau, \alpha \in \mathcal{P}(\mathbb{T})$ such that:

1. The measure $\alpha$ is a uniform measure on a finite (periodic) $T_{n^k}$ orbit such that $\hat{\alpha}(j) \neq 0$.

   To find such a measure, we take the $T_{n^k}$ periodic orbit ${x_0, x_1}$ where $x_0 = \frac{1}{n^k - 1}$ and $x_1 = \frac{n^k}{n^k - 1}$. Define a measure $\alpha = \frac{1}{2} \delta_{x_0} + \frac{1}{2} \delta_{x_1}$ on this orbit. Then

   $$ \hat{\alpha}(x) = \frac{1}{2} \cdot e^{2\pi i \frac{x}{n^k - 1}} (1 + e^{2\pi i \frac{1}{n^k - 1}}) $$

   Now, if $\hat{\alpha}(j) = 0$ then $e^{2\pi i \frac{j}{n^k - 1}} = -1$, which can only happen if $2j(n^k - 1) \in 1 + 2\mathbb{Z}$. However, $2j(n^k - 1) = 2\frac{j}{n^k + 1}$ and $\frac{j}{n^k + 1} < 1$. Thus, it is impossible that $\hat{\alpha}(j) = 0$.

2. The measure $\tau$ is $T_{m^k}$ invariant, dim $\tau \geq 1 - \delta$ and $\hat{\tau}(i) \neq 0$.

   To find such a measure, let $\beta$ be the Cantor-Lebesgue measure with respect to base $m$ and the non-degenerate probability vector $(\frac{1}{3}, \frac{2}{3}, 0, ..., 0)$ (see the end of Section 2.1). Then $\beta$ is a weakly mixing $T_m$ invariant measure (a Bernoulli measure). By (7),

   $$ \hat{\beta}(x) = \prod_{j=1}^{\infty} (\frac{1}{3} + \frac{2}{3} \exp(2\pi i x \frac{1}{m^j})) $$

   $$ = \prod_{j=1}^{\infty} \left( 1 + \frac{2}{3} \cdot (\exp(2\pi i x \frac{1}{m^j}) - 1) \right) $$

   14
By looking at the corresponding power series expansion, we see that for every $j$

$$|\exp(2\pi i x \frac{1}{m^j}) - 1| = \sum_{k=1}^{\infty} \frac{(2\pi i x \frac{1}{m^j})^k}{k!} \leq \sum_{k=1}^{\infty} \frac{(2\pi |x|)^k}{k!} \leq \frac{1}{m^j} \cdot \exp(2\pi |x|).$$

By Proposition 3.1 in Chapter 5 of [29], we conclude that $\hat{\beta}(i) = 0$ if and only if one of its factors has a zero at $i$. Since $0 \neq i \in \mathbb{Z} \subset \mathbb{R}$, this clearly does not happen, and we conclude that $\hat{\beta}(i) \neq 0$.

Also, notice that $\dim \beta = \frac{H(\frac{1}{\log m})}{\log m} > 0$, where $H(p_1, p_2)$ is the Shannon entropy of the probability vector $(p_1, p_2)$. Finally, by Theorem 2.6 there exists some $q \in \mathbb{N}$ such that $\dim \beta^q > 1 - \delta$, where by $\beta^q$ we mean that we convolve $\beta$ with itself $q$ times. Recalling (5), we see that $\hat{\beta^q}(i) = \left(\hat{\beta}(i)^q\right) \neq 0$. Thus, we may take $\tau = \beta^q$. Notice that $\tau$ is $T_m$ invariant, so it is also $T_{m^k}$ invariant.

Thus, by (5) and (6), the Fourier coefficients of the measure $\nu * (\tau \times \alpha) \in \mathcal{P}(\mathbb{T}^2)$ satisfy

$$(\nu * (\tau \times \alpha))(i, j) = \hat{\nu}(i, j) \cdot (\tau \times \alpha)(i, j) = \hat{\nu}(i, j) \cdot \hat{\tau}(i) \cdot \hat{\alpha}(j) \neq 0.$$ 

Therefore, as $i \neq 0$, we have by Lemma 2.3

$$\nu * (\tau \times \alpha) \neq \lambda \times (\rho \ast \alpha), \quad (16)$$

since

$$(\lambda \times (\rho \ast \alpha))(i, j) = \lambda(i) \cdot \rho \ast \alpha(j) = 0.$$ 

On the other hand, let us now lift all our measures to corresponding measures on $[0, 1]$ and $[0, 1]^2$. Since $\nu$ is already defined on the unit square, we take this representative for our lift. Since $\tau$ cannot be atomic we can take our lift as the corresponding measure on $[0, 1]$, and for the measure $\alpha$ we can take essentially the same measure. By Claim 3.3 the conditional measures of $\nu * (\tau \times \alpha)$ with respect to the projection $P_2$ are almost surely finite convex combinations of measures of the form $\nu_{y-x_1} * (\tau \times \delta_{x_1})$, where $i = 0, 1$ are the atoms of the measure $\alpha$, with weights $p_i(y)$ for $i = 0, 1$. So, for $P_2(\nu * (\tau \times \alpha))$ almost every $y$,

$$\dim (\nu * (\tau \times \alpha))_y = \dim (\nu_{y-x_0} * (\tau \times \delta_{x_0})) \cdot p_0(y) + \dim (\nu_{y-x_1} * (\tau \times \delta_{x_1})) \cdot p_1(y)$$

$$= \min(\dim \nu_{y-x_0} * (\tau \times \delta_{x_0}) \cdot p_0(y), \quad \dim \nu_{y-x_1} * (\tau \times \delta_{x_1}) \cdot p_1(y))$$

$$\geq \min(\dim P_1(\nu_{y-x_0} * (\tau \times \delta_{x_0})), \quad \dim P_1(\nu_{y-x_1} * (\tau \times \delta_{x_1})))$$

$$= 1$$

where we have used condition (2) in the statement of the Claim, the lower bound on $\dim \tau$ and Proposition 2.1. Since the opposite inequality is always true, we conclude that

$$\dim (\nu * (\tau \times \alpha))_y = 1, \quad \text{for } P_2(\nu * (\tau \times \alpha)) \text{ almost every } y. \quad (17)$$

Since $P_2(\nu * (\tau \times \alpha)) = \rho * \alpha$, and by (17), we see via Lemma 2.2 part (1) that

$$\dim \nu * (\tau \times \alpha) \geq 1 + \dim \rho * \alpha.$$
On the other hand, by part (2) of Lemma 2.2 and since $P_1\nu = \lambda$,

$$\dim \nu \ast (\tau \times \alpha) \leq \dim_p \lambda \ast \tau + \dim \rho \ast \alpha = 1 + \dim \rho \ast \alpha.$$}

We conclude that $\dim \nu \ast (\tau \times \alpha) = 1 + \dim \rho \ast \alpha$.

Finally, we project $\nu \ast (\tau \times \alpha)$ to $T^2$. Since this projection is a local diffeomorphism, it preserves dimension. Thus, the convolved measure $\nu \ast (\tau \times \alpha)$, with the ambient group being $T^2$, has dimension $1 + \dim \rho \ast \alpha$. Moreover, by Theorem 2.4 since $\rho$ is ergodic it is exact dimensional. Since $\alpha$ is a discrete measure (supported on two atoms), the convolution $\rho \ast \alpha$ remains exact dimensional (Proposition 2.1).

Therefore, we may apply Claim 3.2 for the measure $\nu \ast (\tau \times \alpha)$, since this is a $T^m \times T^n$ invariant measure (as the convolution of such measures), and the assumptions on the dimension of $\nu \ast (\tau \times \alpha)$ and on $P_2(\nu \ast (\tau \times \alpha)) = \rho \ast \alpha$ are met by the previous paragraph. Thus, we may conclude that $\nu \ast (\tau \times \alpha) = \lambda \times (\rho \ast \alpha)$. Via (16), this yields our desired contradiction.

4 Proof of Theorem 1.2

Let $\lambda$ be as in Theorem 1.2, and let $\nu$ be some accumulation point of the sequence of measures as in (1) (where we pick a typical $x$ according to $\mu$), along a subsequence $N_k$. Our goal is to show that $\nu = \lambda \times \rho$, and we shall do this by showing that $\nu$ meets the conditions of Claim 3.4.

By our assumptions and Theorem 1.1 in [13] it follows that $P_1\nu = \lambda$ and $P_2\nu = \rho$. Thus, $\nu$ satisfies condition (1) in Claim 3.4. Notice that this implies that $\nu$ gives zero mass to the points of discontinuity of $T_m \times T_n$. So, $\nu$ is $T_m \times T_n$ invariant. For the second condition of Claim 3.4, we require the following analogue of Theorem 5.1 in [13]. Recall that $P$ is the EFD generated by $\mu$ (see Section 2.4).

Claim 4.1. (Conditional integral representation) For $P_2\nu = \rho$ almost every $y$ there is a probability space $(\Omega, \mathcal{F}, Q(y))$ and measurable functions

$$c : \Omega \to (0, \infty), \quad x : \Omega \to [-1, 1], \quad \eta : \Omega \to \mathcal{P}([-1, 1])$$

such that:

1. $P_1\nu_y = \int c(\omega) \cdot (\delta_x(\omega) \ast \eta(\omega))|_{[0, 1]}dQ(y)(\omega)$

2. Let $P_y$ denote the distribution of random variable $\eta$ as above. Then $P = \int P_y d\rho(y)$.

Proof. We dedicate the first part of the proof to finding a disintegration of $P$ according to the measure $\rho$. To this end, consider the following sequence of distributions $R_{N_k} \in \mathcal{P}(\mathcal{P}([0, 1]) \times [0, 1])$, defined by

$$R_{N_k} = \frac{1}{N_k} \sum_{i=0}^{N_k-1} \delta(\mu_x, i \log m, T^n x).$$

Let $R$ be some accumulation point of this sequence. Without the loss of generality, let us assume the limit already exists along the sequence $N_k$. Then we may assume that $P_1R = P$ and $P_2R = \rho$, since we are considering a $\mu$ typical point $x$, making use of the fact that $\mu$ is pointwise generic under $T_n$ for $\rho$, and of the spectral condition on $P$ via Proposition 2.10.
Next, we disintegrate the distribution $R$ via the projection $P_2$:

$$R = \int R_y d\rho(y).$$

Applying the map $P_1$ to this disintegration, we see that

$$P = P_1 R = \int P_1 R_y d\rho(y).$$

Thus, the family of measures $\{P_1 R_y\}$ forms our desired disintegration.

Let us study this family of distributions a little further: It is well known (see e.g. [9] or [28]) that for $\rho$ almost every $y$, we may write

$$R_y = \lim_{p \to \infty} \frac{R(\cdot \cap P_2^{-1}(D_{2p}(y)))}{\rho(D_{2p}(y))}.$$

Therefore,

$$P_1 R_y = \lim_{p \to \infty} \frac{P_1 R(\cdot \cap P_2^{-1}(D_{2p}(y)))}{\rho(D_{2p}(y))} = \lim_{p \to \infty} \lim_{k \to \infty} \frac{1}{\rho(D_{2p}(y))} \cdot \sum_{0 \leq i \leq N_k - 1: T_n x \in D_{2p}(y)} \delta_{\mu_x,i \log m}.$$

Finally, we note that for $\rho$ almost every $y$, for every $p$,

$$\frac{R(\cdot \cap P_2^{-1}(D_{2p}(y)))}{\rho(D_{2p}(y))} \ll R.$$

Therefore, for $\rho$ almost every $y$, for every $p$,

$$\frac{P_1 R(\cdot \cap P_2^{-1}(D_{2p}(y)))}{\rho(D_{2p}(y))} \ll P_1 R = P. \quad (18)$$

We now turn our attention to the main assertions of the Claim. First, we embed $\mu$ (the measure from Theorem 1.2) on the diagonal of the unit square by pushing it forward via the map $x \mapsto (x, x)$. We call this new measure $\tilde{\mu}$. For $k \in \mathbb{N}$ let $\mathcal{A}^k$ denote the partition of $[0, 1]^2$ given by

$$D_{m^k} \times D_{n^k} = \bigvee_{i=0}^{k-1} (T_m \times T_n)^i (D_m \times D_n).$$

Given a point $z \in [0, 1]^2$ such that $\tilde{\mu}((\mathcal{A}^k(z)) > 0$ we define a probability measure

$$\tilde{\mu}_{\mathcal{A}^k(z)} := c \cdot (T_m \times T_n)^k(\tilde{\mu}|_{\mathcal{A}^k(z)}),$$

where $c$ is a normalizing constant. By applying Claim 2.8 we see that there is a set $S \subseteq \mathbb{N}$ of density 1 (possibly depending on the $x$ we chose according to $\mu$), such that for every $k \in S$ the measure $\tilde{\mu}|_{\mathcal{A}^k(x,x)}$ is an affine image of the measure $\mu|_{D_{m^k}(x)}$. Since we are only interested in the limiting behaviour of these measures, we may assume $S = \mathbb{N}$. Also, $\nu(\partial A) = 0$ for all $A \in \mathcal{A}^k, k \in \mathbb{N}$ since $P_1 \nu$ and $P_2 \nu$ are both continuous measures. Thus, by Theorem 2.7

$$\nu = \lim_{k \to \infty} \frac{1}{N_k} \sum_{i=0}^{N_k-1} \tilde{\mu}_{\mathcal{A}^k(x,x)} \quad (19)$$

17
Now, for $P_2\nu = \rho$ almost every $y$ the conditional measure $\nu_y$ can be obtained as the weak-* limit $\lim_{p \to \infty} \nu_{P_2^{-1}D_{2p}(y)}$, where for every Borel set $A \subset [0,1]^2$ and $p \in \mathbb{N}$,

$$\nu_{P_2^{-1}D_{2p}(y)}(A) := \frac{\nu(A \cap P_2^{-1}D_{2p}(y))}{\nu(P_2^{-1}D_{2p}(y))} = \frac{\nu(A \cap P_2^{-1}D_{2p}(y))}{\rho(D_{2p}(y))}.$$

Fix $p \in \mathbb{N}$. By (19) and since $\tilde{\mu}|_{D_{mk}(x)}$ is an affine image of the measure $\mu|_{D_{mk}(x)}$ for every $k$, the projection of $\nu_{P_2^{-1}D_{2p}(y)}$ to the $x$-axis (i.e. via $P_1$) equals $\tilde{\mu}$.

$$\lim_{k \to \infty} \frac{1}{\rho(D_{2p}(y))} \cdot \sum_{i:0 \leq i \leq N_k - 1} T_{n_i}^{-1}(x) \in D_{2p}(y) \{P_1 \circ \left( L_{n_i^{\alpha}}(T_{n_i}(x), T_{n_i}(x)) \right) c_k (\tau_{x,i} * \mu_{x,i} \log m) \} \bigg| [0,1] \bigg)$$

where $L_{\alpha,z}$ is the unique affine map taking the $x$-axis to the line with slope $\alpha$ through the point $z$. Notice that in the first equation above we only take note of the indices such that $T_{n_i}(x) \in D_{2p}(y)$, and this is justified by Lemma 2.9.

We thus see, as in Theorem 5.1 in [13] and its preceding discussion, that there is a distribution $Q_{D_{2p}(y)} \in \mathcal{P} \left( \mathbb{R} \times [-1,1] \times \mathcal{P}(\mathcal{P}([0,1])) \right)$ such that we have an integral representation (that depends on both $p$ and $y$)

$$P_1 \nu_{P_2^{-1}D_{2p}(y)} = \int g(\omega) dQ_{D_{2p}(y)}(\omega)$$

where $g : \mathbb{R} \times [-1,1] \times \mathcal{P}(\mathcal{P}([0,1])) \to \mathcal{P}([0,1])$ is the map $g(c, x, \eta) = c \cdot (\delta_{x} * \eta)|[0,1]$. Moreover, the distribution of $Q_{D_{2p}(y)}$ on the measure component $\mathcal{P}(\mathcal{P}([0,1]))$ is given by

$$\lim_{k \to \infty} \frac{1}{\rho(D_{2p}(y))} \cdot \sum_{0 \leq i \leq N_k - 1} \delta_{\mu_{x,i} \log m},$$

and by equation (18) and its preceding discussion, this distribution is absolutely continuous with respect to $P$.

Notice that for $Q_{D_{2p}(y)}$ almost every $(c, x, \eta)$, $c$ is the normalizing constant making $(x * \eta)|[0,1]$ a probability measure. Also, the map $g$ is continuous almost surely. Moreover, by moving to a subsequence, we may assume the weak-* limit $\lim_{p \to \infty} Q_{D_{2p}(y)}$ exists, call it $Q_y$. For these assertions, we argue, as in ([13], Theorem 5.1), that the distribution $\{P_1(Q_{D_{2p}(y)})\}$ is tight. Indeed, for measures drawn according to $P$ this follows from Proposition 2.13 and in our case the distribution of $Q_{D_{2p}(y)}$ on the measure component is absolutely continuous with respect to $P$. Finally,

$$P_1 \nu_y = P_1 \lim_{p \to \infty} \nu_{P_2^{-1}D_{2p}(y)} = \lim_{p \to \infty} P_1 \nu_{P_2^{-1}D_{2p}(y)} = \lim_{p \to \infty} \int g(\omega) dQ_{D_{2p}(y)}(\omega) = \int g(\omega) dQ_y(\omega).$$

1Recall that by equation (8) in [13] Section 5.2,

$$\mu_{D_{mk}(x)} := c \cdot T_{mk}(\mu|_{D_{mk}(x)}) = c_k (\tau_{x,k} * \mu_{x,k} \log m)|[0,1]$$

for corresponding parameters.
This completes the proof of part (1). For part (2), it remains to note that by our construction, for \( \rho \) almost every \( y \) the distribution of \( Q_y \) on the measure component \( \mathcal{P}(\mathbb{P}) \) is given by \( P_1 R_y \), as in the first part of the proof, by the discussion preceding \( \footnote{13} \).

**Proof of Theorem 1.2** We are now in position to show that \( \nu \) satisfies all conditions in Claim \footnote{3.4} We have already established that condition (1) holds. As for condition (2), we choose \( \delta = \dim P > 0 \) (here we use the assumption that \( P \) is non trivial, and Proposition 2.12). Let \( \tau \in \mathcal{P}(\mathbb{P}) \) be such that \( \dim \tau \geq 1 - \delta \). We now show that \( \dim \tau \ast P_1 \nu_y = 1 \) for \( P_2 \nu = \rho \) almost every \( y \).

First, by Lemma 2.14 and Claim 4.1 part (2)

\[
1 = \int \dim(\tau \ast \eta) dP(\eta) = \int \dim(\tau \ast \eta) dP(y) d\rho(y).
\]

Therefore, for \( \rho \) almost every \( y \), for \( P_y \) almost every \( \eta \), \( \dim \tau \ast \eta = 1 \) (since the integrand is always \( \leq 1 \)). Thus, by Claim 4.1 part (1), for \( \rho \) almost every \( y \),

\[
\dim \tau \ast P_1 \nu_y = \dim \tau \ast \int c(\omega) \cdot (\delta_{x(\omega)} \ast \eta(\omega)) |_{[0,1]} dQ(y)(\omega)
\]

\[
= \dim \int c(\omega) \cdot \tau \ast (\delta_{x(\omega)} \ast \eta(\omega)) |_{[0,1]} dQ(y)(\omega)
\]

\[
\geq \operatorname{ess-inf}_{\omega \sim Q(y)} \dim \tau \ast (\delta_{x(\omega)} \ast \eta(\omega)) |_{[0,1]}
\]

\[
\geq \operatorname{ess-inf}_{\eta \sim P_y} \dim \tau \ast \delta_{x(\omega)} \ast \eta(\omega)
\]

\[
= \operatorname{ess-inf}_{\eta \sim P_y} \dim \tau \ast \eta = 1
\]

Since \( \dim \tau \ast P_1 \nu_y \leq 1 \) is always true, we find that \( \dim \tau \ast P_1 \nu_y = 1 \) for \( \rho \) almost every \( y \).

We conclude that \( \nu \) satisfies the conditions of Claim \footnote{3.4} Therefore, \( \nu = \lambda \times \rho \), as desired.

5 Proof of Theorem 1.1

Let \( \mu \) be a \( T_\rho \)-invariant and ergodic measure with positive dimension. Then \( \mu \) generates an EFD \( P \) with \( \dim P > 0 \) by Theorem 2.15 Let \( m > n > 1 \). The pure point spectrum \( \Sigma(P, S) \) can contain non zero integer multiples of \( \frac{1}{\log m} \) only if either \( m \sim p \) (in Theorem 1.1 we assume this is not the case), or if \( \frac{\log p}{\log m} \in \Sigma(T, \mu) \), see \footnote{12}. We shall prove Theorem 1.1 by using Theorem 1.2 and following the analysis of Hochman and Shmerkin from \footnote{13}, Section 8 in order to relax the spectral condition (i.e. deal with the latter case). We begin by treating the case \( n = p \).

Suppose first that \( \Sigma(P, S) \) does not contain a non-zero integer multiple of \( \frac{1}{\log m} \). By the ergodic Theorem, \( \mu \) is pointwise \( T_n \) generic for \( \mu \). Also, since \( \mu \) generates an EFD such that \( \frac{k}{\log m} \notin \Sigma(P, S) \) for every \( k \in \mathbb{Z} \setminus \{0\} \), we may apply Theorem 1.2 and obtain

\[
\frac{1}{N} \sum_{i=0}^{N-1} \delta_{(T_{m_i}(x), T_{m_i}(x))} \rightarrow \lambda \times \mu,
\]

for \( \mu \) almost every \( x \).

Suppose now that there exists some \( k \in \mathbb{Z} \setminus \{0\} \) such that \( \frac{k}{\log m} \in \Sigma(P, S) \), so \( P \) is not \( S_{\log m} \) ergodic by Proposition 4.1 in \footnote{13}. By the results discussed in \footnote{13}, Sections 8.2 and 8.3 there is a probability space \( (\Omega, \mathcal{F}, Q) \) and a measurable family of measures \( \{\mu_\omega\}_{\omega \in \Omega} \) such that:
1. The measures \( \{\mu_\omega\}_{\omega \in \Omega} \) form a disintegration of \( \mu \), that is, \( \mu = \int \mu_\omega dQ(\omega) \).

2. For \( Q \) almost every \( \omega \), \( \mu_\omega \) generates \( P \).

3. For \( Q \) almost every \( \omega \), \( \mu_\omega \log m \)-generates an \( S_{\log m} \) ergodic distribution \( P_x \) at almost every point \( x \) (see Proposition 2.10 for the definition of \( \log m \)-generation).

Let \( \delta = \dim P > 0 \) denote the almost sure dimension of the measures drawn by \( P \). Then the following holds:

**Lemma 5.1.** \([13], \text{Lemma 8.3}\) Let \( \tau \in \mathcal{P}(\mathbb{R}) \) be such that \( \dim \tau \geq 1 - \delta \). Then \( \dim \tau \ast \eta = 1 \) for \( Q \) almost every \( \omega \), \( \mu_\omega \) almost every \( x \), and \( P_x \) almost every \( \eta \).

Now, we may finish the proof in a similar fashion to the proof of Theorem 1.1.\(^2\) Namely, For \( Q \) almost every \( \omega \) and for \( \mu_\omega \) almost every \( x \), let \( \nu \) be such that \( (x, x) \) equidistribute for it sub-sequentially under \( T_m \times T_n \). Then we may assume \( P_2 \nu = \mu \) by the ergodic Theorem, and that \( \mu_\omega \log m \)-generates \( P_x \), where \( P_x \) is typical with respect to Lemma 5.1. Then we have a conditional integral representation as in Claim 4.1 but now we can only disintegrate \( P_x = \int (P_x)_y d\mu(y) \). Since we have Lemma 5.1 at our disposal (so that an analogue of (21) holds for \( P_x \) instead of \( P \)), we still have that for every \( \tau \in \mathcal{P}(\mathbb{R}) \) with \( \dim \tau \geq 1 - \delta \), for \( \mu \) almost every \( y \), \( \dim \tau \ast P_1 \nu_y = 1 \) as the calculation carried out during the last stage of the proof of Theorem 1.1 follows through in this case as well. It follows that \( \nu = \lambda \times \mu \). Finally, since this is true for \( Q \) almost every \( \mu_\omega \) and for \( \mu_\omega \) almost every \( x \), this is also true for \( \mu \) almost every \( x \) (recall that \( \mu = \int \mu_\omega dQ(\omega) \)).

The case when \( n \not\sim p \) follows by a similar argument, only here for \( Q \) almost every \( \mu_\omega \), \( \mu_\omega \) is pointwise \( n \)-normal, since this is true for \( \mu \) by Theorem 1.10 in [13].

**6 Perturbing the Initial Point**

In this section we prove the following generalization of Theorem 1.1.

**Theorem 6.1.** Let \( \mu \) be a \( T_p \) invariant ergodic measure with \( \dim \mu > 0 \). Let \( m > n > 1 \) be integers such that \( m \not\sim p \), and let \( f, g \in \mathcal{A} \) be such that \( f([0,1]), g([0,1]) \subseteq [0,1] \).

1. If \( n = p \) then
   \[
   \frac{1}{N} \sum_{i=0}^{N-1} \delta_{(T_{n} f(x), T_{n} g(x))} \rightarrow \lambda \times \mu, \quad \text{for } \mu \text{ almost every } x,
   \]

2. If \( n \not\sim p \) then
   \[
   \frac{1}{N} \sum_{i=0}^{N-1} \delta_{(T_{n} f(x), T_{n} g(x))} \rightarrow \lambda \times \lambda, \quad \text{for } \mu \text{ almost every } x,
   \]

The proof is similar to the proof of Theorem 1.1. In particular, it relies on the following generalization of Theorem 1.2.

**Theorem 6.2.** Let \( \mu \in \mathcal{P}([0,1]) \) be a probability measure, \( f, g \in \mathcal{A} \), and \( m > n > 1 \) be integers, such that:

1. The measure \( \mu \) generates a non-trivial \( S \)-ergodic distribution \( P \in \mathcal{P}(\mathcal{P}([-1,1])) \).

\(^2\)Here, we use the fact that the commutative phase measure from Theorem 8.2 in [13] has dimension 1, as proven in Section 8.3.
2. The pure point spectrum \( \Sigma(P, S) \) does not contain a non-zero integer multiple of \( \frac{1}{\log m} \).

3. The measure \( g\mu \) is pointwise generic under \( T_n \) for an ergodic and continuous measure \( \rho \), and \( f([0, 1]), g([0, 1]) \subseteq [0, 1] \).

Then

\[
\frac{1}{N} \sum_{i=0}^{N-1} \delta(T_m^i f(x), T_n^i g(x)) \rightarrow \lambda \times \rho, \quad \text{for } \mu \text{ almost every } x.
\]

(22)

For this to work, we need the following version of Claim 2.8. Let \( f, g \in \text{Aff}(\mathbb{R}) \). For every \( k \in \mathbb{N} \), define

\[
A_k = \{ x \in \mathbb{R} : f^{-1}D_m^k(f(x)) \nsubseteq g^{-1}D_n^k(g(x)) \}
\]

Claim 6.3. Suppose that \( \mu \in \mathcal{P}([0, 1]) \) is a measure such that \( g\mu \) is pointwise generic under \( T_n \) for a continuous measure \( \rho \). Then for \( \mu \) almost every \( x \), if \( x \in \lim \sup A_k \) and \( \{n_k\} \) represents the times when \( x \in A_{n_k} \), then the density of \( \{n_k\} \) is zero.

The proof is analogous to that of Claim 2.8.

Proof of Theorem 6.2 The proof follows essentially the same steps as the proof of Theorem 1.2. Let \( \nu \) be some accumulation point of the orbit under \( T_m \times T_n \) of \( \delta(f(x), g(x)) \), where \( x \) is drawn according to \( \mu \).

- By (13), Theorem 1.1, we have \( P_1 \nu = \lambda \). By our assumption on \( g\mu \), \( P_2 \nu = \rho \).

- A complete analogue of Claim 4.1 holds in this case as well. First, we disintegrate \( P \) according to \( \rho \), in a similar manner to the first part of the proof of Claim 4.1. Here, we make use of the fact that \( f\mu \) generates and \( \log m \) generates \( P \), which follows by (13), Lemma 4.16.

Secondly, we embed \( \mu \) on a line in \( \mathbb{T}^2 \) by pushing it forward via the map \( x \mapsto (f(x), g(x)) \) (recall that we are assuming that both \( f \) and \( g \) map \([0, 1]\) to \([0, 1]\)). Calling this measure \( \tilde{\mu} \), and using the same notation as in Claim 4.1, we have

\[
\nu = \lim_{k \to \infty} \frac{1}{N_k} \sum_{i=0}^{N_k-1} \tilde{\mu}_{\mathcal{A}_i(f(x), g(x))}
\]

by an application of Theorem 2.7. Also, by applying Claim 6.3, we see that there is a set \( S \subseteq \mathbb{N} \) of density 1 (possibly depending on the \( x \) we chose according to \( \mu \)), such that for every \( i \in S \) the measure \( \tilde{\mu}_{\mathcal{A}_i(f(x), g(x))} \) is an affine image of the measure \( \mu|f^{-1}P_{m,i}(f(x)) \). Thus, we obtain an analogue of (20). From here, we complete the proof as in the proof of Claim 4.1.

- We finish the proof of the Theorem by showing that \( \nu \) meets the conditions of Claim 3.4. The proof is essentially the same as in the case of Theorem 1.2.

Proof of Theorem 6.1 Since we have Theorem 6.2 at our disposal, the proof is now essentially the same as the proof of Theorem 1.1. We remark that an analogue of Lemma 5.1 remains true in this case as well, which may be deduced from the results of (13), Section 8.4.

References

[1] Tim Bedford and Albert M. Fisher. Ratio geometry, rigidity and the scenery process for hyperbolic Cantor sets. Ergodic Theory Dynam. Systems, 17(3):531–564, 1997.
[2] Thomas Bogenschütz and Hans Crauel. The Abramov-Rokhlin formula. In Ergodic theory and related topics, III (Güstrow, 1990), volume 1514 of Lecture Notes in Math., pages 32–35. Springer, Berlin, 1992.

[3] J. W. S. Cassels. On a problem of Steinhaus about normal numbers. Colloq. Math., 7:95–101, 1959.

[4] Kenneth Falconer. Techniques in fractal geometry. John Wiley & Sons, Ltd., Chichester, 1997.

[5] Kenneth J Falconer. The geometry of fractal sets, volume 85. Cambridge university press, 1986.

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

[7] Harry Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation. Theory of Computing Systems, 1(1):1–49, 1967.

[8] Harry Furstenberg, Intersections of Cantor sets and transversality of semigroups. In Problems in analysis (Sympos. Salomon Bochner, Princeton Univ., Princeton, N.J., 1969), pages 41–59. Princeton Univ. Press, Princeton, N.J., 1970.

[9] Hillel Furstenberg. Ergodic fractal measures and dimension conservation. Ergodic Theory and Dynamical Systems, 28(02):405–422, 2008.

[10] Matan Gavish. Measures with uniform scaling scenery. Ergodic Theory Dynam. Systems, 31(1):33–48, 2011.

[11] Michael Hochman. Dynamics on fractals and fractal distributions. arXiv preprint arXiv:1008.3731, 2010.

[12] Michael Hochman. Geometric rigidity of \( \times m \) invariant measures. J. Eur. Math. Soc. (JEMS), 14(5):1539–1563, 2012.

[13] Michael Hochman and Pablo Shmerkin. Equidistribution from fractal measures. Inventiones mathematicae, 202(1):427–479, 2015.

[14] Bernard Host. Nombres normaux, entropie, translations. Israel J. Math., 91(1-3):419–428, 1995.

[15] Bernard Host. Some results of uniform distribution in the multidimensional torus. Ergodic Theory Dynam. Systems, 20(2):439–452, 2000.

[16] Brian R. Hunt and Vadim Yu. Kaloshin. How projections affect the dimension spectrum of fractal measures. Nonlinearity, 10(5):1031–1046, 1997.

[17] Aimee S. A. Johnson. Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers. Israel J. Math., 77(1-2):211–240, 1992.

[18] R. Kenyon and Y. Peres. Measures of full dimension on affine-invariant sets. Ergodic Theory Dynam. Systems, 16(2):307–323, 1996.

[19] Elon Lindenstrauss. \( p \)-adic foliation and equidistribution. Israel J. Math., 122:29–42, 2001.
[20] Elon Lindenstrauss, David Meiri, and Yuval Peres. Entropy of convolutions on the circle. *Ann. of Math. (2)*, 149(3):871–904, 1999.

[21] Russell Lyons. On measures simultaneously 2- and 3-invariant. *Israel J. Math.*, 61(2):219–224, 1988.

[22] Pertti Mattila. *Fourier analysis and Hausdorff dimension*, volume 150 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2015.

[23] David Meiri. Entropy and uniform distribution of orbits in $T^d$. *Israel J. Math.*, 105:155–183, 1998.

[24] David Meiri and Yuval Peres. Bi-invariant sets and measures have integer Hausdorff dimension. *Ergodic Theory Dynam. Systems*, 19(2):523–534, 1999.

[25] Peter M"orters and David Preiss. Tangent measure distributions of fractal measures. *Math. Ann.*, 312(1):53–93, 1998.

[26] Daniel J. Rudolph. $\times 2$ and $\times 3$ invariant measures and entropy. *Ergodic Theory Dynam. Systems*, 10(2):395–406, 1990.

[27] Wolfgang M. Schmidt. On normal numbers. *Pacific J. Math.*, 10:661–672, 1960.

[28] David Simmons. Conditional measures and conditional expectation; Rohlin’s disintegration theorem. *Discrete Contin. Dyn. Syst.*, 32(7):2565–2582, 2012.

[29] Elias M. Stein and Rami Shakarchi. *Complex analysis*, volume 2 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003.

[30] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

[31] U. Z"ahle. Self-similar random measures. I. Notion, carrying Hausdorff dimension, and hyperbolic distribution. *Probab. Theory Related Fields*, 80(1):79–100, 1988.

EINSTEIN INSTITUTE OF MATHEMATICS, EDMOND J. SAFRA CAMPUS, THE HEBREW UNIVERSITY OF JERUSALEM, GIVAT RAM, JERUSALEM, 9190401, ISRAEL.

E-mail address amir.algom@mail.huji.ac.il