Equidistribution from fractal measures

Michael Hochman[*] and Pablo Shmerkin[†]

Abstract

We give a fractal-geometric condition for a measure on $[0,1]$ to be supported on points x that are normal in base n, i.e. such that $\{n^k x\}_{k \in \mathbb{N}}$ equidistributes modulo 1. This condition is robust under C^1 coordinate changes, and it applies also when n is a Pisot number rather than an integer. As applications we obtain new results (and strengthen old ones) about the prevalence of normal numbers in fractal sets, and new results on measure rigidity, specifically completing Host's theorem to multiplicatively independent integers and proving a Rudolph-Johnson-type theorem for certain pairs of beta transformations.

1 Introduction

1.1 Background

A number $x \in [0,1]$ is called n-normal, or normal in base n, if $\{n^k x\}_{k \in \mathbb{N}}$ equidistributes modulo 1 for Lebesgue measure. This is the same as saying that the sequence of digits in the base-n expansion of x has the same limiting statistics as an i.i.d. sequence of digits with uniform marginals. It was E. Borel who first showed that Lebesgue-a.e. x is normal (in every base); thus the n-ary expansion of a typical number is maximally random. It is generally believed that, absent obvious obstructions, this phenomenon persists when it is relativised to "naturally" defined subsets of the reals, i.e. that typical elements of well-structured sets, with respect to appropriate measures, are normal, unless the set displays an obvious obstruction. Taking this to the extreme and applying it to singletons one arrives at the folklore conjecture that natural constants such as $\pi, e, \sqrt{2}$ are normal in every base. While the last conjecture seems very much out of reach of current methods, there are various positive results known for more substantial sets, often "fractal" sets. The present paper is a contribution in this direction.

It is better to work with measures than with sets, and it will be convenient to say that a measure μ is **pointwise n-normal** if it is supported on n-normal numbers. The first results on the problem above were obtained independently by Cassels and W. Schmidt in the late 1950s [13, 51]. Motivated by a question of Steinhaus, who asked whether normality in infinitely many bases implies it for all bases, they showed that the Cantor-Lebesgue measure μ on the middle-$\frac{1}{3}$ Cantor is pointwise m-normal whenever m is not

a power of 3. This answers Steinhaus's question negatively since no number in the middle-$\frac{1}{3}$ Cantor set is 3-normal.

The proofs of Cassels and Schmidt are analytical: they establish rapid decay, as $N \to \infty$, of the $L^2(\mu)$ norms of the trigonometric polynomials $\frac{1}{N} \sum_{k=0}^{N-1} e(mn^k t)$ appearing in Weyl's equidistribution criterion (here and in what follows, $e(s) = \exp(2\pi i s)$). An essentially sharp condition for pointwise n-normality in terms of these norms was provided a few years later by Davenport, Erdős and LeVeque [14]. The latter theorem underlies most subsequent work on the subject and is particularly effective when the measures are constructed with this method in mind, for example Riesz products, which are defined in terms of their Fourier transform. Many results have been obtained in this way by Brown, Pearce, Pollington, and Moran [9, 10, 47, 11, 46, 42]. However, for most "natural" measures the required norm bounds are nontrivial to obtain, if they can be obtained at all. They also are fragile in the sense that they do not persist when the measure is perturbed. The book [12] contains a thorough overview of many classical equidistribution results.

1.2 Main results

In this paper we give a new sufficient condition for pointwise n-normality, which is more dynamical and geometric in nature, and captures the spirit of the conjecture stated at the beginning of this introduction. Roughly speaking, we show that if the process of continuously magnifying the measure around a typical point does not exhibit any almost-periodic features at frequency $1/\log n$, then the measure is pointwise n-normal. While the condition is not a necessary one, it is a natural one in many of the most interesting examples, and can be verified relatively easily in many cases where other methods fail. It also leads to many applications which we discuss below.

The condition is formulated in terms of an auxiliary measure-valued flow which arises from the process of "zooming in" on μ-typical points. This procedure has a long history, going back variously to Furstenberg [22, 23], Zähle [55], Bedford and Fisher [3], Mörters and Preiss [43], and Gavish [25]; the following definitions are adapted from [27], where further references can be found. Let $\mathcal{P}(X)$ denote the space of Borel probability measures on a metric space X; when X is compact we equip it with the Borel structure, and then the space $\mathcal{P}(X)$ is then compact and metrizable in the weak-* topology. Write $\mathcal{M}$ for the space of Radon (locally finite Borel) measures on $\mathbb{R}$ and $\operatorname{supp} \mu$ for the topological support of a measure $\mu \in \mathcal{M}$. Let

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

and for $\mu \in \mathcal{M}^{\square}$ and $t \in \mathbb{R}$, define $S_t\mu \in \mathcal{M}^{\square}$ by[1]

$$S_t\mu(E) = c \cdot \mu(e^{-t}E \cap [-1,1])$$

where $c = c(\mu, t)$ is a normalizing constant. For $x \in \operatorname{supp} \mu$, similarly define the translated measure by $\mu^x(E) = c' \cdot \mu((E+x) \cap [-1,1])$. The **scaling flow** is the Borel $\mathbb{R}^+$-flow

---

[1] In [27] S_t was denoted $S_t^{\square}$ to emphasize that it was acting on $\mathcal{M}^{\square}$, and similarly in some of the later definitions, but this is not needed here and we drop the extra notation.

$S = (S_t)_{t>0}$ acting on $\mathcal{M}^{\square}$. The **scenery** of μ at $x \in \operatorname{supp} \mu$ is the orbit of μ^x under S, that is, the one-parameter family of measures $\mu_{x,t} = S_t(\mu^x)$, $t \geq 0$.

Write $\mathcal{D} = \mathcal{P}(\mathcal{P}([-1,1]))$, which is again compact and metrizable and $\mathcal{P}(\mathcal{M}^{\square}) \subseteq \mathcal{D}$.[2] For clarity we refer to elements of $\mathcal{D}$ as **distributions**, whereas we continue to refer to the elements of $\mathcal{M}^{\square}$ as measures. A measure $\mu \in \mathcal{P}(\mathbb{R})$ **generates a distribution** $P \in \mathcal{D}$ **at** $x \in \operatorname{supp} \mu$ if the scenery at x equidistributes for P in $\mathcal{D}$, i.e. if

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

and μ generates P if it generates P at μ-a.e. x.

If μ generates P, then P is supported on $\mathcal{M}^{\square}$ and S-invariant (while unsurprising this is not completely trivial since S acts discontinuously, see [27, Theorem 1.7] for the proof). We say that P is **trivial** if it is the distribution supported on the measure $\delta_0 \in \mathcal{M}^{\square}$, which is a fixed point of S. It can be shown that if μ generates a distribution, then it is the trivial one if and only if μ gives full mass to a set of zero Hausdorff dimension (this follows from [27, Proposition 1.19]).

To an S-invariant distribution P we associate its **pure-point spectrum** $\Sigma(P, S)$. This is the set of $\alpha \in \mathbb{R}$ for which there exists a non-zero measurable function $\varphi : \mathcal{M}^{\square} \to \mathbb{C}$ satisfying $\varphi \circ S_t = e(\alpha t)\varphi$, $t \in \mathbb{R}$, 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 when the measures are magnified by a factor of e^{α}.

Finally, let $f\mu$ denote the push-forward of the measure μ, i.e. $(f\mu)(A) = \mu(f^{-1}A)$. We note this is sometimes denoted $f_{\#}\mu$.

Theorem 1.1. *Let $\mu \in \mathcal{M}$ be a measure generating a non-trivial S-ergodic distribution $P \in \mathcal{D}$, and let $n \in \mathbb{N}$, $n \geq 2$. If $\Sigma(P, S)$ does not contain a non-zero integer multiple of $1/\log n$, then μ is pointwise n-normal. Furthermore, the same is true for $f\mu$ for all $f \in \operatorname{diff}^1(\mathbb{R})$.*

The non-triviality assumption means that the theorem does not apply to measures supported on zero-dimensional sets. This limitation is intrinsic to our methods.

The hypotheses of the theorem may seem restrictive, since general measures do not generate any distribution, let alone an ergodic one satisfying the spectral condition. However, "natural" measures arising in dynamics, fractal geometry or arithmetic, very often do generate an S-ergodic distribution (see e.g. [25, 26, 27] for many examples), and the important hypothesis becomes the spectral one. It is possible to formulate a version of the theorem that applies to measures which do not generate a distribution in the above sense, but the result is less useful. See remark at the end of Section 5.4. In Section 8 we give some stronger versions of the theorem which are used in some of the later applications.

Finally, note that the theorem is not a characterization, and the presence of $k/\log n$ in the pure point spectrum of P does not rule out pointwise n-normality. Indeed, if

---

[2]We would have liked to define $\mathcal{D} = \mathcal{P}(\mathcal{M}^{\square})$, but while $\mathcal{M}^{\square}$ is a Borel set it is not topologically nice. This is why we define $\mathcal{D}$ as above, and why the test functions in the definition of equidistribution in $\mathcal{D}$ are taken from $C(\mathcal{P}([-1,1])$ and not from $C(\mathcal{M}^{\square})$.

a measure is translated by a random, uniformly chosen distance, then the sceneries are not affected, but almost surely the measure becomes pointwise normal in every base (see also Theorem 1.7 below). It is worth mentioning though that the canonical example of a measure that is *not* pointwise n-normal is that of a singular measure on $[0, 1]$ invariant and ergodic for $x \mapsto nx \bmod 1$. For such μ, the first author showed in [26] that, when the entropy is positive, the generated distribution indeed has a multiple of $1/\log n$ in its spectrum.

There is some interest also in expansions of numbers in non-integer bases. Following Rényi [49], for $\beta > 1$ we define the β-expansion of $x \in [0, 1)$ to be the lexicographically least sequence $x_n \in \{0, 1, \ldots, \lceil \beta - 1 \rceil\}$ such that $x = \sum_{n=1}^{\infty} x_n \beta^{-n}$. This sequence is obtained from the orbit of x under $T_\beta : x \mapsto \beta x \bmod 1$ in a manner similar to the integer case. It is known that T_β has a unique absolutely continuous invariant measure, called the Parry measure, and we shall say that x is β-normal if under T_β it equidistributes for this measure.

Recall that $\beta > 1$ is called a Pisot number if it is an algebraic integer whose algebraic conjugates are of modulus strictly smaller than 1. We adopt the convention that integers ≥ 2 are Pisot numbers. The dynamics of T_β is best understood for this class of numbers, and our results extend to them:

Theorem 1.2. *Theorem 1.1 holds as stated for a Pisot number $\beta > 1$ in place of n.*

It is possible that the Pisot assumption is unnecessary but currently we are unable to prove this (but see also the discussion following Corollary 1.11 below). On the other hand, Bertrand-Mathis [5] proved that if β is Pisot and x is β-normal, then $\{\beta^n x\}_{n=1}^{\infty}$ equidistributes on the circle. Hence we have:

Corollary 1.3. *If $\beta > 1$ is Pisot and μ satisfies the hypothesis of the Theorem 1.1 with β in place of n, then $\{\beta^n x\}_{n=1}^{\infty}$ equidistributes modulo 1 for μ-a.e. x.*

Before turning to applications let us say a few words about what goes into the proof of Theorem 1.2 (a more detailed sketch of the proof is given in Section 5.1). There are two main ingredients. The first involves the behavior of the dimension of measure under convolution. Specifically, among the measures of positive dimension invariant under $x \mapsto \beta x \bmod 1$, one can characterize Lebesgue measure (or the Parry measure) in terms of its dimension growth under convolutions. This part of the argument is special to the dynamics of $x \mapsto \beta x \bmod 1$ and is the main place where the Pisot property is used in the non-integer case. Most of the work then goes into showing that, if there were a measure μ satisfying the hypothesis of the theorems above but not their conclusion, then one could concoct an invariant measure η violating the characterization alluded to above. The scheme above is a refinement of ideas we have used before in [29] and [27].

The second ingredient in the proof, and one of the main innovations in this paper, applies in a more general setting than invariant measures for piecewise-affine maps of $[0, 1]$. The proper context is that of a Borel map T of a compact metric space X. Roughly speaking, we show how to relate the small-scale structure of a measure μ on X to the distribution of T-orbits of μ-typical points. This result, while classical in nature, appears

4

to be new and we believe it may find further applications. We leave the discussion and precise statement to Section 2.

1.3 Applications

1.3.1 Normal numbers in fractals

As our first application we consider sets arising as attractors of iterated function systems, or, equivalently, repellers of uniformly expanding maps on the line (see below for definitions). We show that, under some weak regularity assumptions, if such a set is defined by nonlinear dynamics, or if the contraction rates of the defining maps satisfy a natural algebraic condition, then typical points in the set are n-normal. This should be interpreted in terms of the conjecture stated earlier: indeed, it implies that if such a set contains no n-normal numbers, then the set is essentially defined by linear[3] maps whose slopes are rational powers of n, and in this sense the dynamics is similar to the canonical examples of sets without n-normal numbers, namely closed subsets of $[0, 1]$ that are invariant under the piecewise-linear maps $x \mapsto nx \bmod 1$.

We start with the relevant definitions. An **iterated function system** (IFS) is a finite family $\mathcal{I} = \{f_0, \ldots, f_{r-1}\}$ of strictly contracting maps $f_i : I \to I$ for a compact interval $I \subset \mathbb{R}$ (of course one can define IFSs in general metric spaces). The IFS is of class C^α if all the f_i are. We shall say that the IFS $\mathcal{I}$ is **regular** if the maps f_i are orientation-preserving injections, and the intervals $f_i(I)$ are disjoint except possibly at their endpoints so, in particular, the so-called open set condition is satisfied. In this article we will only consider $C^{1+\varepsilon}$ regular IFSs, but some of the assumptions can be relaxed. For example, the orientation-preserving assumption is just for simplicity and can be easily dropped.

The **attractor**[4] of $\mathcal{I}$ is the unique nonempty compact set $X \subseteq I$ satisfying

$$X = \bigcup_{i \in [r]} f_i(X)$$

(here and throughout the paper, $[r] = \{0, \ldots, r - 1\}$). There are a number of natural measures one can place on X. One is the $\dim X$-dimensional Hausdorff measure, which for a $C^{1+\varepsilon}$-IFS is positive and finite on X. Another good class are the **self-conformal measures** (also called **self-similar measures** if the maps f_i are linear), that is, measures satisfying the relation

$$\mu = \sum_{i \in [r]} p_i \cdot f_i \mu$$

for a positive probability vector $(p_0, \ldots, p_{r-1})$. Both of the examples above are special cases of **Gibbs measures** for Hölder potentials $\varphi : X \to \mathbb{R}$. We will not define Gibbs measures, but rather rely on a standard property of such measures μ, namely, that there is a constant $C > 1$ such that for all finite sequences $i_1, \ldots, i_k, j_1, \ldots, j_\ell \in [r]$,

$$C^{-1} \leq \frac{\mu(f_{i_1} \cdots f_{i_k} I) \mu(f_{j_1} \cdots f_{j_\ell} I)}{\mu(f_{i_1} \cdots f_{i_k} f_{j_1} \cdots f_{j_\ell} I)} \leq C. \tag{1}$$

---

[3]We shall follow the convenient but imprecise convention of using the term linear also for affine maps.

[4]There is an equivalent dynamical description of attractors of C^α-IFSs, namely, as the maximal compact invariant sets of expanding C^α maps $I \mapsto \mathbb{R}$.

We shall call measures satisfying this property **quasi-product measures** (or quasi-Bernoulli measures). This is a broader class than Gibbs measures for Hölder potentials; for example, it contains Gibbs measures for almost-additive sequences of potentials, see [2].

Our first result assumes an algebraic condition on the contractions. For a C^1-contraction f on $\mathbb{R}$, we define its **(asymptotic) contraction ratio** to be $\lambda(f) = f'(p)$, where p is the unique fixed point of f. For affine f this is just the usual contraction ratio; to justify the name in the nonlinear case note that for every distinct pair of points x, y,

$$\lambda(f) = \lim_{n \to \infty} -\frac{\log |f^n(x) - f^n(y)|}{n}.$$

Write $a \sim b$ if a, b are integer powers of a common number, equivalently $\log a / \log b \in \mathbb{Q}$; otherwise write $a \not\sim b$, in which case a, b are said to be **multiplicatively independent**.

Theorem 1.4. *Let $\mathcal{I}$ be a $C^{1+\varepsilon}$ IFS that is regular in the sense above, and $\beta > 1$ a Pisot number.[5] If there exists an $f \in \mathcal{I}$ with $\lambda(f) \not\sim \beta$, then any quasi-product measure μ for $\mathcal{I}$ is pointwise β-normal, and so is $g\mu$ for all $g \in \mathrm{diff}^1(\mathbb{R})$.*

The classical results of Cassels and Schmidt are special cases of this for certain IFSs consisting of affine maps with the same contraction ratio. We note that the result above is new even when the IFS is affine and contains maps with two multiplicatively independent contraction ratios; classical methods break down since nothing seems to be known about the decay (or lack thereof) of the Fourier transform of natural measures on such attractors.

Our second result says that nonlinearity and enough regularity are sufficient for pointwise normality, irrespective of algebraic considerations. More precisely, we say an IFS $\mathcal{I} = \{f_i\}$ is **linear** if all of the maps in $\mathcal{I}$ are affine maps, and **non-linear** otherwise. We say that $\mathcal{I}$ is **totally non-linear** if it is not conjugate to a linear IFS via a C^1 map; here an IFS is $\mathcal{J}$ is C^α-conjugate to $\mathcal{I}$ if it has the form $\mathcal{J} = g\mathcal{I} = \{gf_ig^{-1}\}$ for a C^α-diffeomorphism g.

Theorem 1.5. *Let $\mathcal{I}$ be a C^ω IFS that is regular in the sense above and $\beta > 1$ a Pisot number. If $\mathcal{I}$ is totally non-linear, then any quasi-product measure μ for $\mathcal{I}$ is pointwise β-normal, and so is $g\mu$ for all $g \in \mathrm{diff}^1(\mathbb{R})$.*

The two theorems above have substantial overlap and each of them is generic in the appropriate space of IFSs. The algebraic condition is generally the easier one to verify, and the regularity assumptions are weaker, though it seems very probable that weaker regularity assumptions are sufficient in the totally non-linear case also.

It also seems very likely that non-linearity, rather than total non-linearity, should suffice in Theorem 1.5. We are able to prove such a result for a smaller class of measures. Namely,

Theorem 1.6. *Let $\mathcal{I}$ be a C^ω IFS that is regular in the sense above, and $\beta > 1$ a Pisot number. If $\mathcal{I}$ is non-linear, then every self-conformal measure for $\mathcal{I}$ is pointwise β-normal.*

---

[5] We remark again that in this and subsequent statements, the Pisot assumption includes the possibility $\beta \in \mathbb{N}$.

In this theorem, the totally non-linear case is covered by Theorem 1.5. In the conjugate-to-linear case, if $g \in \text{diff}^\omega(\mathbb{R})$ conjugates $\mathcal{I}$ to a linear IFS $\mathcal{J} = g\mathcal{I}$, then $\mu = g^{-1}\nu$ where ν is a self-similar measure for $\mathcal{J}$, and g is not affine (since $\mathcal{J}$ is linear and $\mathcal{I} = g^{-1}\mathcal{J}$ is not). Also, it is a remarkable consequence of the work of Sullivan [53] and Bedford-Fisher [3] that if a C^α-IFSs, $\alpha > 1$, is C^1-conjugate to a linear IFS then it is also C^α-conjugate to a linear IFS (see [3, Theorem 7.5]). Thus, Theorem 1.6 follows from the following one:

Theorem 1.7. *Let μ be a self-similar measure for a linear IFS that is regular in the sense above. Then for any non-affine real-analytic $g \in \text{diff}^\omega(\mathbb{R})$, $g\mu$ is pointwise β-normal for every Pisot $\beta > 1$.*

Here is one concrete consequence of the results above.

Corollary 1.8. *Let μ denote the Cantor-Lebesgue measure on the middle-1/3 Cantor set. Then x^2 is 3-normal for μ-a.e. x.*

The point is, of course, that no points in the middle-1/3 Cantor set are 3-normal themselves.

The corollary above is immediate from the previous theorem and the use of the square function is incidental. In fact for we could replace x^2 with $f(x)$ for $f \in \text{diff}^2$. From Theorem 1.4 we can reduce the regularity to diff^1 if we only want n-normality for $n \sim 3$. These differences perhaps indicate that our regularity assumptions may be sub-optimal. Note that for $f = \text{identity}$, this again is the theorem of Cassels and Schmidt, and their spectral methods carry over to translations, but the stability under perturbation is new even for affine f. Related to this question, we note that Bugeaud, Fishman, Kleinbock and Weiss [8] have shown that for many fractals sets, including self-similar sets satisfying the open set condition, there is a full-dimension subset consisting of numbers which are *not* normal in any integer base. Moreover their result holds for any bi-Lipschitz image of the set. The stability of our results under bi-Lipschitz transformations remains open.

While this paper was in revision we learned of Kaufman's paper [35]. Kaufman studies differentiable images of certain Bernoulli convolutions, and obtains polynomial decay of the Fourier transform of their image under C^2 diffeomorphisms, implying in particular pointwise normality of the images. His results apply to linear self-similar measures defined by two maps with the same contraction ratio, and equal weights (it is likely the method can be adapted to more than two maps, but unlikely that the equicontraction assumption can be dropped with current methods). In particular the last corollary follows from Kaufman's work.

1.3.2 Host's theorem and measure rigidity

Let $n \in \mathbb{N}$ and let $T_n : [0,1] \to [0,1]$ denote the map $T_n x = nx \mod 1$. An important phenomenon concerning these maps is measure rigidity: a well-known conjecture of Furstenberg states that, if $m \not\sim n$, then the only probability measures jointly invariant under T_m and T_n are combinations of Lebesgue measure and atomic measures on rational points. This conjecture, known as the times-2, times-3 conjecture, is the prototype

for many similar conjectures in other contexts, see e.g. [38]. The best result towards it is due to Rudolph and Johnson [50, 33]: if a measure has positive entropy and is jointly invariant and ergodic under T_m, T_n for $m \not\sim n$, then it is Lebesgue. Although nothing is known about the zero-entropy case, in the positive entropy case there is a pointwise strengthening of the Rudolph-Johnson theorem for $\gcd(m, n) = 1$, due to B. Host [31, Théorème 1]:

Theorem 1.9. *Let $m, n \geq 2$ be integers and $\gcd(m, n) = 1$. Suppose μ is an invariant and ergodic measure for T_n of positive entropy. Then μ is pointwise m-normal.*

This implies the Rudolph-Johnson theorem in the case $\gcd(m, n) = 1$: if μ is a jointly T_m, T_n invariant measure and all T_n ergodic components have positive entropy, then by the theorem μ-a.e. point equidistributes for Lebesgue under T_m. But by the ergodic theorem, it also equidistributes for the ergodic component of μ to which it belongs; hence μ is Lebesgue.

The hypothesis of Host's theorem, however, is stronger than it "should" be, i.e. it is stronger than the hypothesis of the Rudolph-Jonson Theorem. Lindenstrauss [37] showed that the conclusion holds under the weaker assumption that n does not divide any power of m, but this is still too strong.[6] On the other hand, Feldman and Smorodinsky [20] had earlier proved a similar result assuming only that $m \not\sim n$, but under the strong assumption that the measure μ is weak Bernoulli. In that work it is conjectured that the same holds assuming only that μ is ergodic and has positive entropy. The following theorem gives the result in its "correct" generality and for some non-integer bases, and also shows that it is stable under smooth enough perturbation.

Theorem 1.10. *Let $\beta, \gamma > 1$ with β a Pisot number, and $\beta \not\sim \gamma$. Then any T_γ-invariant and ergodic measure μ with positive entropy is pointwise β-normal. Furthermore the same remains true for $g\mu$ for any $g \in \mathrm{diff}^2(\mathbb{R})$.*

Of course, the same is true under the assumption that all T_γ-ergodic components of μ have positive entropy. Note the asymmetry in the requirement from β, γ. We do not know whether the Pisot assumption is unnecessary, but we note that Bertrand-Mathis [4] has obtained some complementary results for γ Pisot and β arbitrary, though only for measures that satisfy the weak-Bernoulli property with respect to the natural symbolic coding of T_γ.

From this one derives a new measure rigidity result for β-maps.

Corollary 1.11. *Let $\beta, \gamma > 1$ with $\beta \not\sim \gamma$ and β Pisot. If μ is jointly invariant under T_β, T_γ, and if all ergodic components of μ under T_γ have positive entropy, then μ is the common Parry measure for β and γ; in particular, it is absolutely continuous. The same holds if T_β, T_γ are conjugated separately by C^2-diffeomorphisms.*

Proof. If μ is as in the statement, then by Theorem 1.10, μ-almost all x equidistribute under T_β for the β-Parry measure (i.e. an absolutely continuous measure). On the other

---

[6]Host's theorem has also been generalized in some other directions, see Meiri [41]

hand, by the ergodic theorem μ-a.e. x equidistributes for the T_γ-ergodic component to which it belongs; hence μ is also the Parry measure for T_γ.

The latter assertion follows in the same way, using that $g\mu$ is pointwise β-normal for all $g \in \mathrm{diff}^2(\mathbb{R})$. $\qquad\square$

We hope to be able to eliminate the Pisot assumption in this result; this will be addressed in a forthcoming paper. Corollary 1.11 also improves [26, Corollary 1.5] by eliminating the ergodicity assumption. We do not know for what pairs (β, γ) the Parry measures coincide, or even whether this may happen for different non-integer β, γ.

1.3.3 Badly approximable normal numbers

Another application concerns continued fraction representations and their relation to integer expansions. Let $\Lambda \subseteq \mathbb{N}$ be a finite set with at least two elements, and set

$$C_\Lambda = \{x \in [0,1] : x \text{ has only symbols from } \Lambda \text{ in its continued fraction expansion}\}.$$

These sets are natural in Diophantine approximation since their union over all finite $\Lambda \subseteq \mathbb{N}$ is the set of badly approximable numbers. The question of whether there are badly approximable normal numbers reduces to asking whether any of the C_Λ contain normal numbers. An affirmative answer follows from work of Kaufman [34], who, assuming $\dim C_\Lambda > 2/3$, constructed probability measures on C_Λ whose Fourier transform decays polynomially. The bound on the dimension was relaxed to $\dim C_\Lambda > 1/2$ by Queffélec and Ramaré [48]. Thus, for example, there are normal numbers whose continued fraction expansions consist only of the digits $1, 2$ (because $\dim C_{\{1,2\}} > 1/2$). However, the methods from those papers fail below dimension $1/2$, so, for example, it was not known whether there are normal numbers with continued fraction coefficients $5, 6$.

We note that C_Λ is the attractor of a regular IFS, namely $\{f_i \circ f_j : i, j \in \Lambda\}$, where $\{f_i\}$ are the inverse branches of the Gauss map (the reason for the compositions is that, although f_1 is not a strict contraction, all the compositions $f_i \circ f_j$ are). As an application of Theorem 1.4, we have:

Theorem 1.12. *Any quasi-product measure on C_Λ (in particular the $\dim C_\Lambda$-dimensional Hausdorff measure) is pointwise β-normal for any Pisot $\beta > 1$.*

Even when $\dim C_\Lambda > 1/2$, this improves the results of Kaufman, Queffélec and Ramaré, in that the result holds for a broader and more natural class of measures. The result on normality in non-integer Pisot bases is new in all cases. It seems very likely that the result holds also for Gibbs measures when $\Lambda \subseteq \mathbb{N}$ is infinite, under standard assumptions on the Gibbs potential, but we do not pursue this.

One natural question is whether a reciprocal of Theorem 1.12 holds. For example, is it true that almost all points in the middle-$1/3$ Cantor set are normal with respect to the Gauss map G? (i.e. they equidistribute under G for the Gauss measure, which is the only absolutely continuous G-invariant measure). To the best of our knowledge, it is not even known whether there exists a point which is Gauss normal but not n-normal for

any n (in the positive direction, Einsiedler, Fishman and Shapira [15] recently proved that almost all points in the middle-1/3 Cantor set have unbounded partial quotients, i.e. are not contained in any C_Λ). Unfortunately, our methods do not seem to help with this problem. The (piecewise) linearity of T_β is strongly used in the part of the proof that deals with the geometric behavior of invariant measures under convolution. In particular, Theorem 5.5 seems to fail for the Gauss map and likely for most non-linear and many piecewise linear maps.

1.4   Organization of the paper

In the next section we state and prove a general result relating orbits of μ-typical points to the structure of μ; this is the second main component of the proof of Theorem 1.2 referred to above. Section 3 collects some background on dimension. In section 4 we recall some background on the pure point spectrum and eigenfunctions of flows, and discuss the class of distributions arising from scenery flows, called ergodic fractal distributions. We also introduce the concept of phase measure and its main properties. We prove Theorem 1.2 in Section 5 (with a key component postponed to Section 6). In Section 7 we derive Theorems 1.4, 1.5 and 1.12. Finally, in Section 8 we prove some variants of Theorem 1.1, and employ them to prove Theorems 1.10 and 1.7.

Acknowledgment

2   Relating the distribution of orbits to the measure

While most of our considerations in this paper are special to $\mathbb{R}$, those in this section apply in the following very general setting. Let X be a compact metric space and $T : X \to X$ a Borel measurable map.[7] For Borel probability measures μ, ν on X, let us say that a measure μ is **pointwise generic** for ν if μ-a.e. x equidistributes for ν under T, that is,

$$\frac{1}{N} \sum_{n=0}^{N-1} f(T^n x) \to \int f \, d\nu \quad \text{for every } f \in C(X). \tag{2}$$

This notion appears in many contexts, although the name is not standard. Clearly when $Tx = nx \bmod 1$ and ν is Lebesgue measure on $[0, 1]$, this is the same as pointwise n-normality. A well-known variant appears in smooth dynamics: when X is a manifold, a measure ν is called the Sinai-Ruelle-Bowen (SRB) measure if the volume measure on X is pointwise generic for ν. Other examples include the study of badly approximable points on analytic curves in $\mathbb{R}^d$, and similar applications in arithmetic contexts.

While one does not expect to be able to say very much for arbitrary maps and measures, there is an obvious formal strategy to follow if one wants to prove that μ is point-

---

[7]Compactness is only required in order to define weak-* convergence (i.e. provide a natural algebra of test function), but the core of the discussion below is purely measure-theoretic.

wise generic for ν: it is sufficient to show that for μ-a.e. x, if x equidistributes for a measure η along some subsequence of times (i.e. (2) holds along some $N_k \to \infty$), then $\eta = \nu$.

To go any further with this scheme, one needs a way to relate measures η arising as above to the original measure μ. It is not obvious that such a relation exists: η is determined primarily by the point x, and although x is μ-typical, once it is selected, it would appear that the role of μ has ended. However, it turns out that there is a very close connection between η and μ, provided by the theorem below. Roughly speaking, it shows that, under a mild technical condition, one can express η as a weak limit of "pieces" of μ, "magnified" via the dynamics.

For a finite measurable partition $\mathcal{A}$ of X, write $T^i \mathcal{A} = \{T^{-i}A \ : \ A \in \mathcal{A}\}$ and $\mathcal{A}^n = \bigvee_{i=0}^{n} T^i \mathcal{A}$ for the coarsest common refinement of $\mathcal{A}, T\mathcal{A}, \ldots, T^n \mathcal{A}$. Also let $\mathcal{A}^\infty = \bigvee_{i=0}^{\infty} T^i \mathcal{A}$ denote the σ-algebra generated by the partitions $\mathcal{A}^n$, $n \geq 0$. We say that $\mathcal{A}$ is a **generator** for T if $\mathcal{A}^\infty$ is the full Borel algebra. When T is invertible, we similarly define $\mathcal{A}^{\pm n} = \bigvee_{i=-n}^{n} T^i \mathcal{A}$ and $\mathcal{A}^{\pm\infty} = \bigvee_{i=-\infty}^{\infty} T^i \mathcal{A}$, and say that $\mathcal{A}$ is a generator if $\mathcal{A}^{\pm\infty}$ is the full Borel algebra. Finally, we say that $\mathcal{A}$ is a **topological generator** if $\sup\{\operatorname{diam} A \ : \ A \in \mathcal{A}^n\} \to 0$ as $n \to \infty$ (or, in the invertible case, the sup is over $A \in \mathcal{A}^{\pm n}$). A topological generator is clearly a generator.

Write $\mathcal{A}(x) \in \mathcal{A}$ for the unique element $A \in \mathcal{A}$ containing x. Given $\mu \in \mathcal{P}(X)$ and a point $x \in X$ such that $\mu(\mathcal{A}^n(x)) > 0$, let

$$\mu_{\mathcal{A}^n(x)} = c \cdot T^n(\mu|_{\mathcal{A}^n(x)})$$

where $c = \mu(\mathcal{A}^n(x))^{-1}$ is a normalizing constant. For a.e. x, this is well-defined for all n.

Theorem 2.1. *Let $T : X \to X$ be a Borel-measurable map of a compact metric space, μ be a Borel probability measure on X and $\mathcal{A}$ a generating partition. Then for μ-a.e. x, if x equidistributes for $\nu \in \mathcal{P}(X)$ along some $N_k \to \infty$, and if $\nu(\partial A) = 0$ for all $A \in \mathcal{A}^n, n \in \mathbb{N}$, then*

$$\nu = \lim_{k \to \infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \mu_{\mathcal{A}^n(x)} \qquad \text{weak-* in } \mathcal{P}(X). \tag{3}$$

If, furthermore, $\mathcal{A}$ is a topological generator, then the hypothesis on ν follows if, for all m,

$$\limsup_{k \to \infty} \frac{1}{N_k} \sum_{n=1}^{N_k} \mu_{\mathcal{A}^n(x)}(C_m^{(\varepsilon)}) = o(1) \qquad \text{as } \varepsilon \to 0, \tag{4}$$

where $C_m = \bigcup_{A \in \mathcal{A}^m} \partial A$, and $C_m^{(\varepsilon)}$ is its ε-neighborhood.

Note that if in the right hand side of (3) we replace $\mu_{\mathcal{A}^n(x)}$ by $\delta_{T^n x}$, then the convergence to ν is just a reformulation of the definition of equidistribution. Generally $\mu_{\mathcal{A}^n(x)}$ and $\delta_{T^n x}$ are very different measures and the content of the theorem is that these two sequences are nevertheless asymptotic in the Cesàro sense. This is quite surprising, and such a general fact can only be due to very general principles, as we shall see in the proof.

Proof. We give the proof assuming that $\mathcal{A}$ is forward generating and comment on the invertible case at the end.

Let $\mathcal{F}$ denote the set of linear combinations of indicator functions of $A \in \mathcal{A}^n$, $n \in \mathbb{N}$, with coefficients in $\mathbb{Q}$. This is a countable algebra and, for $x \in X$ and $\nu \in \mathcal{P}(X)$ such that $\nu(\partial A) = 0$ for $A \in \mathcal{A}^n$, it is well known that x equidistributes for ν along N_i if and only if $\lim \frac{1}{N_i} \sum_{n=1}^{N_i} f(T^n x) = \int f \, d\nu$ for every $f \in \mathcal{F}$ (it is here that we use the assumption that ν gives zero mass to the boundaries of $A \in \mathcal{A}^n$). Similarly, the limit in the conclusion of the theorem holds if and only if $\lim \frac{1}{N_i} \sum_{n=1}^{N_i} \int f(x) \, d\mu_{\mathcal{A}^n(x)} = \int f \, d\nu$ for all $f \in \mathcal{F}$. It follows, then, that to prove the theorem it suffices for us to show that for μ-a.e. x,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \left(\int f \, d\mu_{\mathcal{A}^n(x)} - f(T^n x) \right) \to 0 \qquad \text{for every } f \in \mathcal{F}. \tag{5}$$

Suppose that $f = \sum a_i 1_{A_i}$ where $a_i \in \mathbb{Q}$ and $A_i \in \mathcal{A}^k$ for some k. Notice that by definition of $\mu_{\mathcal{A}^n(x)}$,

$$\int f \, d\mu_{\mathcal{A}^n(x)} = \frac{1}{\mu(\mathcal{A}^n(x))} \int_{\mathcal{A}^n(x)} T^n f \, d\mu$$
$$= \mathbb{E}_\mu(T^n f \mid \mathcal{A}^n)(x)$$

Writing $g_n = \mathbb{E}_\mu(T^n f \mid \mathcal{A}^n) - T^n f$, it suffices to show that $\lim \frac{1}{N} \sum_{n=0}^{N-1} g_n = 0$ μ-a.e., and for this it clearly suffices to prove that $\lim \frac{1}{N} \sum_{n=0}^{N-1} g_{kn+p} = 0$ μ-a.e. for $0 \le p \le k - 1$.

Now, g_n is $\mathcal{A}^{n+k}$-measurable (because $T^n f$ is $\mathcal{A}^{n+k}$-measurable); and on the other hand

$$\mathbb{E}_\mu(g_n \mid \mathcal{A}^n) = (\mathbb{E}_\mu(T^n f \mid \mathcal{A}^n) - \mathbb{E}_\mu(T^n f \mid \mathcal{A}^n)) = 0$$

Therefore, $\{g_{p+kn}\}_{n=0}^\infty$ is an orthogonal system in $L^2(\mu)$, since if $j > i$ then

$$\int g_{p+ki} \, g_{p+kj} \, d\mu = \int \mathbb{E}_\mu(g_{p+ki} \, g_{p+kj} | \mathcal{A}^{p+k(i+1)}) \, d\mu$$
$$= \int g_{p+ki} \cdot \mathbb{E}_\mu(g_{p+kj} | \mathcal{A}^{p+k(i+1)}) \, d\mu$$
$$= \int g_{p+ki} \cdot 0 \, d\mu$$
$$= 0.$$

Since the sequence $\{g_{p+kn}\}_{n=0}^\infty$ is also uniformly bounded in $L^2(\mu)$, we conclude that $\frac{1}{N} \sum_{n=0}^{N-1} g_{p+kn} \to 0$ a.e., see for instance [39]. (Alternatively, $\{g_{p+kn}\}_{n=1}^\infty$ form a sequence of bounded martingale differences for the filtration $\{\mathcal{A}^{p+kn}\}$, hence their averages converges a.e. to 0, see [21, Chapter 9, Theorem 3].)

We turn to the second statement. Assume that $\mathcal{A}$ is a topological generator. We will show that the assumption (4) implies that $\nu(\partial A) = 0$ for $A \in \mathcal{A}^n$, $n \in \mathbb{N}$. Fix n and

$C = C_n$ as in the statement. For $\varepsilon > 0$ let $f_\varepsilon \in \mathcal{F}$ be such that $1_C \leq f_\varepsilon \leq 1_{C^{(\varepsilon)}}$. Then, using (5) and the hypothesis (4), we get

$$
\begin{aligned}
\limsup_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f_\varepsilon(T^n x) &= \limsup_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} \int f_\varepsilon \, d\mu_{\mathcal{A}^n(x)} \\
&\leq \limsup_{k\to\infty} \frac{1}{N_k} \sum_{n=0}^{N_k-1} \mu_{\mathcal{A}^n(x)}(C^{(\varepsilon)}) \\
&= o(1) \qquad \text{as } \varepsilon \to 0.
\end{aligned}
$$

Since $\mathcal{A}$ is a topological generator, $\mathcal{F}$ is uniformly dense in $C(X)$, so the above conclusion holds also for $f \in C(X)$ satisfying $1_C \leq f \leq 1_{C^{(\varepsilon)}}$. Since x equidistributes for ν along $\{N_k\}$, this implies that $\nu(C) = 0$. $\qquad\square$

In the case that T is invertible we consider instead the algebra $\mathcal{F}^\pm$ of $\mathbb{Q}$-linear combinations of indicators of sets from $\mathcal{A}^{\pm n} = \bigvee_{i=-n}^n T^i \mathcal{A}$. The rest of the proof proceeds as before using the filtration $\mathcal{A}^{\pm n}$.

3 Preliminaries on dimension

In this section we summarize some standard and some less well known facts about dimension.

3.1 Dimension of measures

The **(lower) Hausdorff dimension** of a finite non-zero Borel measure θ on some metric space is defined by

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

Here $\dim A$ is the Hausdorff dimension of A. We note that this is only one of many possible concepts of dimension of a measure, but it turns out to be the appropriate one for our purposes because of the way it behaves under convolutions, i.e. the resonance and dissonance phenomena discussed in the following sections.

An alternative characterization that we will have occasion to use is given in terms of local dimensions:

$$
\dim \theta = \operatorname*{essinf}_{x\sim\theta} \underline{\dim}(\theta, x), \tag{6}
$$

where

$$
\underline{\dim}(\theta, x) = \liminf_{r\downarrow 0} \frac{\log \theta(B(x, r))}{\log r}
$$

is the lower local dimension of θ at x. The equivalence is a version of the mass distribution principle, see [19, Proposition 4.9]. Note that this characterization shows that (when the underlying space is compact) the dimension is a Borel function of the measure in the weak* topology.

We briefly recall some other properties of the dimension which will be used throughout the paper without further reference. Clearly $\dim(\theta|_E) \geq \dim \theta$ for any set E of positive measure, and $\dim$ is invariant under bi-Lipschitz maps (since this is true for the

13

dimension of sets); in particular it is invariant under diffeomorphisms. Dimension also satisfies the relations

$$\dim \sum_i \theta_i = \inf_i \dim \theta_i,$$

$$\dim \int \theta_\omega \, dQ(\omega) \geq \operatorname{essinf}_{\omega \sim Q} \dim \theta_\omega.$$

In particular, $\dim \mu = \dim T\mu$ for any map T between intervals that is a piecewise diffeomorphism (such as the maps T_β or the Gauss map), as can be seen by writing the measure as a sum over countably many domains where the map is bi-Lipschitz. The same argument shows that dimension is invariant under the quotient map $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$.

Finally, we note that (6) implies that $\dim \mu \times \nu \geq \dim \mu + \dim \nu$ (strict inequality is possible).

3.2 Projection theorems

It is a general principle that if μ is a measure on some space X and $f : X \to Y$ is a "typical" Lipschitz map, then the image measure $f\mu$ will have dimension that is "as large" as possible: namely, it will have the same dimension as μ itself if Y is large enough to accommodate this, and otherwise it will be as large as a subset of Y can possibly be, that is, it will have the same dimension as Y. Thus one expects $\dim f\mu = \min\{\dim \mu, \dim Y\}$. There are many precise versions of this fact. The most classical is Marstrand's projection theorem, concerning linear images of sets and measures on $\mathbb{R}^2$. The following version is due to Hunt and Kaloshin [32, Theorem 4.1].

Theorem 3.1. *If η is a probability measure on $\mathbb{R}^2$, then for a.e. $\alpha \in [0, \pi)$, $\dim \pi_\alpha \eta = \min\{1, \dim \eta\}$, where π_α is the orthogonal projection onto a line making angle α with the x-axis.*

In our applications, θ will be a product $\mu \times \nu$. In this particular case, we obtain

Corollary 3.2. *Let $\mu, \nu \in \mathcal{P}(\mathbb{R})$. Then for almost all $t \in \mathbb{R}$,*

$$\dim(\mu * S_t \nu) \geq \min(1, \dim \mu + \dim \nu).$$

Proof. The family of linear maps $\{P_t(x, y) = x + ty\}$ is a smooth reparametrization of the orthogonal projections $\{\pi_\alpha\}$, up to affine changes of coordinates which do not affect dimension. Hence, by Theorem 3.1,

$$\dim P_t(\mu \times \nu) = \min(1, \dim(\mu \times \nu)) \geq \min(1, \dim \mu + \dim \nu) \quad \text{for a.e. } t.$$

The corollary follows since $\mu * S_t \nu$ is a restriction of $P_t(\mu \times \nu)$ to a set of positive measure, and restriction does not decrease dimension. $\square$

We will have occasion to use the following refinement of the above.

Theorem 3.3. *If μ, ν are Borel probability measures on $\mathbb{R}$ such that $\dim \mu + \dim \nu > 1$, then*

$$\dim\{t \in \mathbb{R} : \dim(\mu * S_t \nu) < 1\} < 1.$$

Proof. Falconer [17] essentially proved the corresponding result for Hausdorff dimensions of sets, we indicate how to modify his proof to work with dimension of measures (the argument is standard). Let $P_t(x, y) = x + ty$. In the course of the proof of [17, Theorem 1] it is shown that if η is a Borel probability measure on $\mathbb{R}^2$ such that

$$\int \int \frac{d\eta(x) d\eta(y)}{|x - y|^s} < \infty \tag{7}$$

for some $s > 1$, then the set E of parameters t such that the projection $P_t \eta$ is not absolutely continuous, satisfies $\dim(E) \leq 2 - s < 1$ (as above, Falconer worked with orthogonal projections, but by reparametrization the same holds for the family $\{P_t\}$).

Let $\rho = \mu \times \nu$. We only need to show that $\dim(E) < 1$, where

$$E = \{\alpha : P_\alpha \rho \text{ is not absolutely continuous}\}.$$

We have $\dim \rho \geq \dim \mu + \dim \nu > 1$. Using Equation (6), it follows that there is $s_0 > 1$ such that

$$\liminf_{r \downarrow 0} \frac{\log \rho(B(x, r))}{\log r} \geq s_0 \quad \text{for } \rho\text{-a.e. } x.$$

By Egorov's Theorem, for any $\varepsilon > 0$ there are a set A_ε with $\rho(A_\varepsilon) > 1 - \varepsilon$ and a constant $r_\varepsilon > 0$ such that

$$\rho(B(x, r)) \leq r^{(1+s_0)/2} \quad \text{for all } x \in A_\varepsilon, 0 < r < r_\varepsilon.$$

It follows that $\eta := \rho|_{A_\varepsilon}$ satisfies (7) with $s = 1 + (s_0 - 1)/4$ (say). Hence $\dim(E_\varepsilon) \leq 2 - s$, where $E_\varepsilon = \{\alpha : P_\alpha \rho|_{A_\varepsilon} \text{ is singular}\}$. Since $E \subset \bigcup_{n \in \mathbb{N}} E_{1/n}$ the result follows.

$\square$

3.3 Further facts on dimension

For the part of the proof of Theorem 1.10 dealing with invariance under C^2 diffeomorphisms, we will need some classical but perhaps less well-known facts about dimension. The material below will not be used anywhere except in this application.

It is always true that $\dim(A \times B) \geq \dim(A) \times \dim(B)$ for Borel sets A, B; however, the inequality may be strict. More generally, there is a "Cavalieri inequality" for Hausdorff dimensions. To get inequalities in the opposite direction, one needs to consider also packing dimension $\dim_P$. The interested reader may consult e.g. [40] for its definition, but we shall only require the property given in the following proposition.

Proposition 3.4. *Let $E \subset \mathbb{R}^{d_1 + d_2}$ be a Borel set.*

1. *Suppose there is a set $A \subset \mathbb{R}^{d_1}$ of positive Lebesgue measure, such that for $x_0 \in A$, the fiber $\{y : (x_0, y) \in E\}$ has Hausdorff dimension at least α. Then $\dim(E) \geq d_1 + \alpha$.*

2. *Let P_i be the coordinate projection onto $\mathbb{R}^{d_i}$. Then $\dim(E) \leq \dim(P_1 E) + \dim_P(P_2 E)$.*

The first part follows from [40, Theorem 7.7], and the second from [40, Theorem 8.10]

We now turn to measures. In a similar way to our definition of lower Hausdorff dimension dim, we may define **upper packing dimension** $\dim_P$ as

$$\dim_P(\mu) = \inf\{\dim_P(E) : \mu(E) = 1\}.$$

(Note that in the definition of dim the infimum is taken over sets of positive measure; here, it is taken over sets of full measure.) The following is an analog of Proposition 3.4 for measures.

Lemma 3.5. *Let μ be a measure on $\mathbb{R}^{d_1+d_2}$, and let P_i be the coordinate projection onto $\mathbb{R}^{d_i}$.*

1. *Suppose $P_1\mu$ is absolutely continuous, and $\dim(\mu_{x_0}) \geq \alpha$ for $P_1\mu$-a.e. x_0, where μ_{x_0} is the conditional measure on the fiber $\{(x,y) : x = x_0\}$. Then $\dim\mu \geq d_1 + \alpha$.*

2. *$\dim\mu \leq \dim(P_1\mu) + \dim_P(P_2\mu)$.*

Proof. For the first part, suppose $\mu(E) > 0$. Then there is a set A with $P_1\mu(A) > 0$ (and hence A has positive Lebesgue measure) such that $\mu_x(E) > 0$ for almost all $x \in A$. The claim then follows from the corresponding statement for sets. The second part is established in a similar manner. □

Finally, recall that a measure μ is **exact dimensional** if the local dimension

$$\lim_{r\downarrow 0} \frac{\log\mu(B(x,r))}{\log r}$$

exists and is μ-a.e. constant. For exact dimensional measures μ, it is well known that $\dim\mu = \dim_P\mu$, with both dimensions agreeing with the almost sure value of the local dimension; see [18, Proposition 2.3]. If $\beta > 1$ is Pisot and μ is T_β-ergodic, then μ is exact dimensional. This well known fact follows from the Shannon-McMillan-Breiman Theorem and, in the Pisot case, a classical Lemma of Garsia (see Lemma 6.2 below).

4 Ergodic fractal distributions, spectra and phase

4.1 Ergodicity and spectrum

Below we prove some basic facts relating the spectrum of a flow to the equidistribution properties of points under individual maps in the flow. The discussion is mostly valid for general flows on metric spaces but for simplicity we formulate them for $(\mathcal{M}^\square, S)$.

Proposition 4.1. *If $P \in \mathcal{D}$ is S-ergodic and $t_0 > 0$, then P is S_{t_0}-ergodic if and only if no non-zero multiple of $1/t_0$ is in the pure point spectrum of P.*

Proof. S acts on the ergodic decomposition with respect to S_{t_0}: $P = \int P_\mu \, dP(\mu)$. Clearly this action is t_0-periodic. Thus the factor of (P, S) with respect to the σ-algebra $\mathcal{E}$ of invariant sets factors through the standard translation action of $\mathbb{R}$ on $\mathbb{R}/t_0\mathbb{Z}$. The only

factors of the latter action are the trivial one, in which case $\mathcal{E}$ is trivial and P is S_{t_0}-ergodic, or an action isomorphic to the translation action of $\mathbb{R}$ on $\mathbb{R}/(t_0/k)\mathbb{Z}$ for some $k \in \mathbb{Z} \setminus \{0\}$, in which case this factor map defines an eigenfunction with eigenvalue k/t_0. $\qquad \square$

Lemma 4.2. *Let P be S-ergodic and $t_0 > 0$. Then P-a.e. μ equidistributes under S_{t_0} for an S_{t_0}-ergodic distribution P_μ, and $P = \int P_\mu \, dP(\mu)$ is the ergodic decomposition of P under S_{t_0}. If no multiple of $1/t_0$ is in $\Sigma(P, S)$, then $P_\mu = P$ a.s.*

Proof. Let $P = \int P_\mu \, dP(\mu)$ be the ergodic decomposition of P with respect to the measure-preserving map S_{t_0}. By the ergodic theorem, for P-a.e. μ, P_μ-a.e. ν equidistributes for P_ν. the first statement follows. For the second statement, if $k/t_0 \notin \Sigma(P, S)$ for all non-zero integers k, then by the previous Proposition P is S_{t_0}-ergodic, and so $P_\mu = P$ a.s. $\qquad \square$

We turn to distributions generated by a measure $\mu \in \mathcal{P}(\mathbb{R})$. Given $t_0 > 0$, we say that a distribution P is t_0-generated by μ at x if μ^x equidistributes for P under the discrete semigroup $\{S_{kt_0}\}_{k \in \mathbb{N}}$, that is, the sequence $\{\mu_{x,kt_0}\}_{k=0}^\infty$ equidistributes for P.

We have seen that if $k/t_0 \notin \Sigma(P, S)$ for all non-zero integers k, then P-a.e. μ t_0-equidistributes for P. The next result says that the same is true for any measure μ that generates P.

Lemma 4.3. *Suppose μ generates an S-ergodic distribution P and no non-zero integer multiple of t_0 is an eigenvalue of (P, S). Then P is t_0-generated at μ-a.e. x.*

This is, essentially, the following well-known fact from ergodic theory, whose proof we provide for completeness:

Lemma 4.4. *Let $W = (W_t)_{t>0}$ be a continuous flow on a compact metric space X. Suppose θ is a W-invariant and ergodic measure which does not have k/t_0 is its pure point spectrum for any $k \in \mathbb{Z} \setminus \{0\}$. Then any point x which equidistributes for θ under W equidistributes for θ also under the "time t_0" map W_{t_0}.*

Proof. As in Lemma 4.2, the spectral hypothesis implies ergodicity of θ under the map W_{t_0}. Now suppose that x equidistributes for a measure θ' under W_{t_0} along a sequence $N_k \to \infty$; it suffices to prove $\theta' = \theta$. By continuity, θ' is W_{t_0}-invariant, so $W_t\theta'$ is W_{t_0}-invariant for every t. Let $\rho = \frac{1}{t_0} \int_0^{t_0} W_t\theta' \, dt$. Then for every $f \in C(X)$,

$$
\begin{aligned}
\int f \, d\rho &= \lim_{k \to \infty} \frac{1}{t_0} \int_0^{t_0} \frac{1}{N_k} \sum_{n=0}^{N_k-1} f(W_{t_0}^n W_t x) dt \\
&= \lim_{k \to \infty} \frac{1}{N_k t_0} \int_0^{N_k t_0} f(W_t x) \, dt \\
&= \int f \, d\theta,
\end{aligned}
$$

17

where the last equality is because x equidistributes for θ. Thus $\rho = \theta$, i.e. $\frac{1}{t_0} \int_0^{t_0} W_t \theta' dt = \theta$. Since θ is W_{t_0}-ergodic and this is a representation of θ as the integral of W_{t_0}-invariant measures, we conclude that $W_t \theta' = \theta$ for a.e. t. Since θ is W-invariant this holds for $t = 0$, i.e. $\theta' = \theta$, as desired. $\qquad\square$

A priori this does not apply in our situation, because the topological assumptions are not satisfied (S acts discontinuously, and is not everywhere defined on $\mathcal{P}(\mathcal{P}([-1,1]))$). However, the only place in the proof that continuity was used was in the assertion that θ' is W_{t_0}-invariant. In the context of Lemma 4.3 this is true at μ-a.e. point by [27, Theorem 1.7]. Thus, we have proved Lemma 4.3.

4.2   Ergodic fractal distributions

Definition 4.5. *An S-invariant distribution $P \in \mathcal{D}$ is **S-quasi-Palm** if for every Borel set $B \subset \mathcal{M}^\square$, $P(B) = 1$ if and only if for every $t > 0$, P-almost every measure η satisfies $\eta_{x,t} \in B$ for η-almost all x such that $[x - e^{-t}, x + e^{-t}] \subset [-1, 1]$.*

Definition 4.6. *A distribution $P \in \mathcal{D}$ which is supported on $\mathcal{M}^\square$, S-invariant and satisfies the S-quasi-Palm property is called a **fractal distribution**, or **FD**. If, in addition, P is S-ergodic, then P is called an **ergodic fractal distribution**, or **EFD**.*

This definition differs slightly from the one introduced and studied in [27]. More precisely, the notion of quasi-Palm in [27] is suited for distributions on Radon measures on $\mathbb{R}$, rather than distributions on probability measures on $[-1, 1]$, and the notion of EFDs there is for distributions on Radon measures that are invariant under the action of a semigroup S^*, which is defined similarly to S but without restricting the measures to a bounded interval, so that S^* acts on measures of unbounded support (our S is denoted by $S^\square$ in [27]). For this reason, in the definition of quasi-palm measure given in [27] there is no need to assume that $[x - e^{-t}, x + e^{-t}] \subset [-1, 1]$, and it has μ^x in place of $\mu_{x,t}$. However, it is proved in [27, Lemma 3.1] that S-invariant and S^*-invariant distributions are canonically in one-to-one correspondence. Hence any EFD according to our definition arises as the push-forward of an EFD in the sense of [27] under the map $\mu \mapsto \mu|_{[-1,1]}$. Therefore all results proved for EFDs in [27] continue to be valid with our definition of EFD. In particular, the following is proved in [27, Theorem 1.7].

Theorem 4.7. *For μ almost all x, any distribution P generated by μ at x along a sequence of times T_i is a FD (i.e. it is S-invariant and automatically satisfies the S-quasi-Palm property).*
In particular, if μ generates an S-ergodic distribution P, then P is an EFD.

For the rest of the section we fix an EFD P, and shall draw some simple but important conclusions about it. We will repeatedly use the following consequence of the S-quasi-palm property:

Lemma 4.8. *Let P be an EFD, and $B \subseteq \mathcal{M}^\square$ a Borel set with the property that $\eta \in B$ whenever $S_t \eta \in B$ for some t. Then $P(B) = 1$ if and only if for P-almost all η and η-almost all x, the translation η^x is in B.*

As a first application, we have:

18

Lemma 4.9. *Let P be an EFD. Fix $t_0 > 0$. P-a.e. μ generates P, and t_0-generates an S_{t_0}-ergodic component of P at μ-a.e. point x.*

Proof. Let B be the set of μ such that μ generates P and t_0-generates an S_{t_0}-ergodic component of P at 0; then $P(B) = 1$ by ergodicity. Also, $\eta \in B$ whenever $S_t\eta \in B$, so the lemma follows from Lemma 4.8. $\square$

Recall that an S-invariant distribution is trivial if it is supported on the S-fixed point δ_0.

Lemma 4.10. *If P is a non-trivial EFD then P-almost all measures are non-atomic.*

Proof. Let $a(\mu) = \mu(\{0\})$. It is clear that $a(S_t\mu) \geq a(\mu)$, by definition of S, so a is a.s. constant. Also it is clear that if $a(\mu) > 0$ then $a(S_t\mu) \to 1$ as $t \to \infty$, so if that were the case, $a = 1$ P-a.s. But this would imply that $\mu = \delta_0$ a.s. and so P is trivial, contrary to assumption. Hence $a = 0$ P-a.s.; using Lemma 4.8 applied to the set $\{\nu : a(\nu) = 0\}$, we find that P-a.e. ν satisfies $a(\nu^x) = 0$ for ν-a.e. x, so ν is non-atomic. $\square$

Lemma 4.11. *Suppose that μ t_0-generates P and P is supported on non-atomic measures. For every $\varepsilon > 0$ there is a $\rho > 0$ such that*

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \sup \mu_{x,t_0 n}(I) < \varepsilon \quad \text{for } \mu\text{-almost all } x.$$

where the supremum is over intervals $I \subseteq [-1, 1]$ of length $|I| < \rho$.

Proof. Fix $\varepsilon > 0$ and let $\mathcal{C}_\rho$ denote the set of measures η such that $\eta(I) < \varepsilon$ for every open interval of length $|I| < \rho$. Note that $\mathcal{C}_\rho$ is open.

By the fact that P gives no mass to measures with atoms, for P-a.e. η there is a $\rho = \rho_\eta > 0$ depending on η such that $\sup \eta(I) < \varepsilon$ where I ranges over open intervals of length ρ_η. It follows that there is a ρ such that with P-probability $> 1 - \varepsilon$ we have $\rho < \rho_\eta$, and in particular $P(\mathcal{C}_\rho) > 1 - \varepsilon$. Since μ t_0-generates P, we find that

$$\limsup_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \delta_{\mu_{x,t_0 n} \in \mathcal{C}_\rho} \geq P(\mathcal{C}_\rho) > 1 - \varepsilon$$

as required. $\square$

It is not hard to show that the same conclusion holds if one assumes only that μ generates a non-trivial P (without necessarily t_0-generating it), but we will not use this fact.

In fact, not only are P-typical measures non-atomic; they also have positive dimension:

Proposition 4.12. *Let P be an EFD. There is a number δ such that P-a.e. ν has $\dim \nu = \delta$. If P is nontrivial then $\delta > 0$.*

19

Proof. This follows from [27, Lemma 1.18]; we include a proof for completeness. For the first statement, restriction can only increase dimension, and scaling does not affect it, so for any measure ν we have $\dim S_t \nu \geq \dim \nu$. By ergodicity, the dimension is P-a.s. equal to some constant $\delta \geq 0$.

Now assume that P is nontrivial, we need to show that $\delta > 0$. We will use the characterization of dimension using local dimension, recall Equation (6). Write

$$f(\nu) = \liminf_{r \downarrow 0} \frac{\log \nu([-r, r])}{\log r}.$$

By Lemma 4.8, it is enough to verify that there is $\delta > 0$ such that $f(\nu) \geq \delta$ for P-a.e. ν (note that the set $B = \{\nu : f(\nu) \geq \delta\}$ satisfies $S^t \nu \in B \Rightarrow \nu \in B$). But f is S-invariant, whence by ergodicity we only need to check that $f(\nu) > 0$ on a set of positive P-measure.

Now Lemma 4.10 and S-invariance ensure that $g(\nu) = -\log \nu([-1/2, 1/2])$ satisfies $\int g \, dP > 0$. By the ergodic theorem applied to the (possibly non-ergodic) discrete-time system $S_{\log 2}$,

$$\lim_{N \to \infty} \frac{\log \nu([-2^{-N}, 2^{-N}])}{N \log 2} = \lim_{N \to \infty} \frac{1}{N \log 2} \sum_{n=0}^{N-1} g(S^{n \log 2} \nu)$$

converges almost everywhere to a function of ν with strictly positive integral; but the left-hand side equals $f(\nu)$, so this completes the proof. $\square$

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

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

Proof. Let $B = \{\nu : \nu[-\varepsilon, 0] = 0 \text{ for some } \varepsilon > 0\}$. It is enough to show that $P(B) = 0$. Indeed, if this is true then by symmetry also $P(B') = 0$ where $B' = \{\nu : \nu([0, \varepsilon]) = 0 \text{ for some } \varepsilon > 0\}$, and the claim follows since any interval of length ρ containing 0 contains either $[-\rho/2, 0]$ or $[0, \rho/2]$.

Since B is S-invariant, by ergodicity we only need to show that $P(B) < 1$. Suppose otherwise. Since $S^t \mu \in B$ implies that $\mu \in B$, it follows from Lemma 4.8 that, for P-typical ν and ν-typical x, there is $\varepsilon(x)$ such that $\nu([x - \varepsilon(x), x]) = 0$. Take $\varepsilon > 0$ such that $\nu(A) > 0$, where $A = \{x : \varepsilon(x) \geq \varepsilon\}$. The restriction $\nu|_A$ has the property that the distance between any two distinct points in its support is at least ε. However this can only happen for discrete measures, and we have already established in Lemma 4.10 that P-typical measures have no atoms. Hence $P(B) < 1$ and therefore $P(B) = 0$, as claimed.

$\square$

4.3 Phase and synchronization

Suppose that μ generates P and t_0-generates an S_{t_0}-ergodic distribution P_x at μ-typical points x. Let φ be an eigenfunction of the flow (P, S) for some eigenvalue k/t_0. Since φ is S_{t_0}-invariant, it is almost surely constant on each ergodic component of P under S_{t_0}, hence it is P_x-a.s. constant for μ-a.e. x. This allows us to define the **phase** of μ at x to be the a.s. value of φ on P_x. We denote the phase by $\varphi_\mu(x)$, and claim that it is a measurable function of x. Indeed, write φ as the increasing limit of simple functions φ_n, and note that $\varphi_\mu(x) = \int \varphi \, dP_x = \lim_{n\to\infty} \int \varphi_n \, dP_x$. The map $x \mapsto P_x$ is measurable in x, since P_x arises as an almost-sure limit of measurable functions of x, and hence $x \mapsto \int \phi_n \, dP_x$ is measurable for each n. By the limit above, also ϕ_μ is.

The push-forward of μ to the unit circle by $x \mapsto \varphi_\mu(x)$ gives a measure $\theta = \theta_\mu$ which describes the distribution of phases, and is called the **phase measure**.[8]

Lemma 4.14. *For P-typical ν, let P_ν denote the S_{t_0}-ergodic component of P to which ν belongs. Then for P-a.e. ν, the phase of ν is well defined at 0 and is equal to $\varphi(\nu)$.*

Proof. Fix an S_{t_0}-ergodic component P' of P. Let z denote the P'-a.s. value of φ. Now, for P'-a.e. ν we know that $\varphi(\nu) = z$ and, by the ergodic theorem, that ν equidistributes for P' under S_{t_0}. This shows that the phase of ν is well defined at 0 and equal to z. Since P is the integral of its ergodic components, the claim follows. $\square$

Proposition 4.15. *For P-a.e. ν, the function φ_ν is ν-a.e. constant and $\theta_\nu = \delta_{\varphi(\nu)}$.*

Proof. By the S-quasi-Palm property and the last lemma, it is clear that for P-a.e. ν and ν-a.e. x, the eigenfunction φ is well-defined on $S_t(\nu^x)$ for all large enough t, and that this value is the phase of the distribution that is t_0-generated by ν^x. Since $x \mapsto \varphi_\nu(x)$ is measurable, it is enough to show that P-almost all ν and all $\varepsilon > 0$,

$$\int \int |\varphi_\nu(y') - \varphi_\nu(y'')| \, d\nu(y') \, d\nu(y'') < \varepsilon$$

Write A_ε for the set of ν for which the above holds; we aim to show $P(A_\varepsilon) = 0$. Let

$$B_\varepsilon = \{\nu : S_t\nu \in A_\varepsilon \text{ for sufficiently large } t\}.$$

By invariance, it is enough to show that $P(B_\varepsilon) = 1$.

By the Besicovitch differentiation theorem [40, Corollary 2.14(2)], for ν-almost all x,

$$\lim_{t\to\infty} \frac{\int_{[x-e^{-t}, x+e^{-t}]} |\varphi_\nu(x) - \varphi_\nu(y)| d\nu(y)}{\nu([x - e^{-t}, x + e^{-t}])} \to 0 \quad \text{as } t \to \infty,$$

---

[8]This definition of the phase and phase measure differs from that in [26, Definition 2.6], but the two definitions coincide for measures for which both definitions apply. One might say that the definition given here is absolute (given φ), while the definition in [26] is relative, as it measures difference in phase between sceneries at pairs of μ-typical points.

and therefore

$$\lim_{t\to\infty} \frac{\int_{[x-e^{-t},x+e^{-t}]^2} |\varphi_\nu(y') - \varphi_\nu(y'')| d\nu(y') d\nu(y'')}{\nu([x-e^{-t},x+e^{-t}])^2} \to 0 \quad \text{as } t \to \infty.$$

From the eigenfunction property, $\varphi_{\nu_{x,t}}(y) = e(-\alpha t)\varphi_\nu(x + e^{-t}y)$ for all t, ν-a.e. x and $\nu_{x,t}$-a.e. y. It follows that for ν-a.e. x, the measure $\nu_{x,t}$ is in A_ε for sufficiently large t, i.e. $\nu^x \in B_\varepsilon$. But then we conclude from Lemma 4.8 that $P(B_\varepsilon) = 1$, as desired. $\qquad\square$

Finally we consider the effect of perturbation on the generated distributions and the phase of a measure.

Lemma 4.16. *Let $\nu \in \mathcal{P}(\mathbb{R})$.*

1. *Let $f \in L^1(\nu)$, $f \geq 0$ and $\int f d\nu > 0$, and write $d\nu' = f\, d\nu$. Then for ν'-a.e. x, the sceneries of ν and of ν' at x are asymptotic. In particular, if ν generates P, then so does ν'.*

2. *Let I be an interval and $f : I \to J$ an orientation-preserving diffeomorphism. Let $\nu' = f(\nu)$. Then for ν-a.e. x, the sceneries $\nu_{x,t}$ and $\nu'_{f(x),t-\log f'(x)}$ are mean-asymptotic in $\mathcal{P}([-1,1])$ in the sense that*

$$\lim_{T\to\infty} \frac{1}{T} \left(\int_0^T F(\nu_{x,t})\, dt - \int_0^T F(\nu'_{f(x),t-\ln f'(x)})\, dt \right) = 0 \quad \text{for all } F \in C([-1,1])$$

   and similarly when one averages at discrete time steps of some size t_0. In particular, if ν generates P at x then ν' generates P at $f(x)$.

Proof. The first part is an immediate consequence of the Besicovitch differentiation theorem (see Mattila [40, Corollary 2.14(2)], or [27] for more detail).

The second part can be proved by adapting the argument in Proposition 1.9 of [27] or the forthcoming paper of Aspenberg, Ekström, Persson and Shmeling [1].[9] Here we only give a sketch. Consider the maps $g_t(y) = e^t \cdot (y - x)$ and $h_t(y) = e^t \cdot (f(y) - f(x))$, so that $\nu_{x,t} = a_t \cdot g_t(\nu)|_{[-1,1]}$ and $\nu'_{x,t} = b_t \cdot h_t(\nu)|_{[-1,1]}$ for normalizing constants a_t, b_t (we suppress the dependence on x in the notation). Using the linear approximation of f at x, we see that the uniform distance $\varepsilon(t)$ between the maps g_t and $h_{t-\log f'(x)} \circ f$ on $[x - 2e^{-t}, x + 2e^{-t}]$ tends to 0 as $t \to \infty$. Thus we will be done if we show that for ν-a.e. x we have $a_t/b_t \to 1$ in the mean (Cesaro) sense. Now, for a given $\delta > 0$, in order to have $|a_t/b_t - 1| > \delta$, we must have $\left| \frac{\nu(B_{e^{-t-\varepsilon(t)}}(x))}{\nu(B_{e^{-t+\varepsilon(t)}}(x))} - 1 \right| > \frac{\delta}{100}$. If this were to happen for a non-negligible proportion of ts in arbitrarily long intervals $[0, T_i]$ we would conclude that there is a distribution P generated by ν at x along the times T_i, such that, with positive P-probability, a measure θ satisfies $\theta(\{\pm 1\}) > 0$. This is impossible by Theorem 4.7, Lemma 4.10 and the ergodic decomposition.

---

[9]This lemma first appeared as Lemma 2.3 of [26] but the statement there incorrectly omits the "almost every" quantifier over x.

In the discrete time case, suppose that when averaged at steps of size t_0 the two sceneries are not a.s. mean-asymptotic. Passing to a subsequence, the we find that for a positive μ-proportion of x, there is a subsequence along which μ generates some distribution P_x t_0-discretely at x, and P_x gives positive mass to measures with atoms at ± 1. But then for μ-a.e. such x one sees that $P'_x = \int_{-t_0}^0 S_t P_x \, dt$ is a FD supported on measures that have atoms at non-zero points, and we know this is impossible, because each ergodic component of P is an EFD [26] and is either trivial, in which case its measures have an atom only at 0, or non-trivial, in which case Lemma 4.10 applies (since the space of atomic measures with atoms is not closed, some more care must be taken in the last step, and one needs to use the fact that for P there is already a positive probability of finding atoms of mass bounded away from zero at locations bounded away from $0, \pm 1$, and this translates to P'. We omit the details). $\qquad \square$

Corollary 4.17. *If μ generates P t_0-discretely, P is S_{t_0}-ergodic and $t_0 \in \Sigma(P, S)$ with eigenfunction φ, then*

1. *If $\nu \ll \mu$, then θ_ν is well defined and $\theta_\nu \ll \theta_\mu$.*

2. *If $f \in \mathrm{diff}^1(\mathbb{R})$ and $\nu = f(\mu)$, then θ_ν is well defined and*

$$\theta_\nu = \int \delta_{e(-t_0 \log f'(x))\varphi_\mu(x)} \, d\mu(x).$$

Proof. For (1), by the previous lemma, if $\nu \ll \mu$ then for ν-a.e. y, the distribution t_0-generated by μ and ν at y is the same, and the claim follows. For (2), fixing a μ-typical x, by the second part of the previous lemma, $\mu_{t,x}$ and $\nu_{f(x),-\log f'(x)}$ generate that same distribution t_0-discretely. Hence[10] ν generates $S_{-\log f'(x)}P$ t_0-discretely at $f(x)$, and by the eigenfunction property,

$$\varphi_\nu(fx) = e(-t_0 \log f'(x))\varphi_\mu(x),$$

from which we deduce (2). $\qquad \square$

5 Proof of theorem 1.2

5.1 A sketch of the proof

We start by explaining the main steps involved in the proof of Theorem 1.2. This strategy will also apply for the generalizations considered in Section 8, with suitable modifications.

We start with a measure μ on $[0, 1]$ generating an EFD P such that $k/\log \beta \notin \Sigma(P, S)$ for $k \in \mathbb{Z} \setminus \{0\}$ for Pisot $\beta > 1$. We fix a μ-typical x and suppose that x equidistributes under T_β for a measure ν along some subsequence N_j; our job is to show that ν is in fact the Parry measure λ_β. To accomplish this, there are three main steps involved:

---

[10]Here we use the fact that although $S_{-\log f'(x)}$ is not continuous, it is continuous on the set of non-atomic measures, and hence on a set of full measure for P, since the non-triviality of $\Sigma(P)$ implies that P is non-trivial, hence supported on non-atomic measures.

1. The first step is to use Theorem 2.1 and the spectral hypothesis to establish that ν can be represented as a superposition of measures drawn according to P, each of them suitably translated, restricted and normalized. See Theorem 5.1 and the ensuing discussion for the general Pisot case.

2. We show that any T_β-invariant measure of positive dimension, other than the Parry measure, resonates with measures of arbitrarily large dimension (see Section 5.3 for the definition of resonance and dissonance). This is stated in Theorem 5.5 and proved in Section 6.

3. Using the first step, the S-invariance of P and Marstrand's Theorem, we show that ν dissonates with arbitrary measures of sufficiently large dimension (this step uses the nontriviality of P). Hence, in light of the second step, ν must be the Parry measure. This step is carried over in Section 5.4.

We note that both the first and second steps use the algebraic assumption on β (in each case it can be slightly relaxed, but in different directions).

5.2 An integral representation

We begin with the details. From now on, we specialize to the interval $[0,1]$ and to maps of the form $T_n : x \mapsto nx \bmod 1$ for an integer $n \geq 2$. We comment on the Pisot case afterwards. Let $\mu \in \mathcal{P}([0,1])$ be a measure that generates a distribution P satisfying the spectral hypothesis in Theorem 1.1 (we do not assume that μ is T_β-invariant; in fact, we will eventually apply the result of this section to measures μ which are invariant under a *different* dynamics). We shall obtain a certain integral representation of the measures for which μ-typical points equidistribute along sub-sequences.

Let $\mathcal{A}$ denote the partition of $[0,1]$ into n-adic intervals, $[j/n, (j+1)/n)$. Note that $\delta_y * \nu$ is the translate of the measure ν by y. Fixing x, we claim that

$$\mu_{\mathcal{A}^k(x)} = c_k \cdot (\delta_{y_k} * \mu_{x,k\log n})|_{[0,1]} \tag{8}$$

for some normalizing constant c_k and a number $y_k \in [0,1]$. Indeed, $\mu_{x,k\log n}$ is the restriction of μ to the interval I of side $2 \cdot n^{-k}$ centered at x, re-scaled to $[-1,1]$ and normalized; while $\mu_{\mathcal{A}^k(x)}$ is obtained similarly from the restriction of μ to an interval $J = \mathcal{A}^k(x)$ of length n^{-k} around x, re-scaled to the interval $[0,1]$ and normalized. Since $J \subseteq I$, the representation (8) follows.

For a μ-typical x, suppose that x equidistributes for some measure ν under T_n, along a sequence N_j. Since P is nontrivial, Lemma 4.11 applied with $t_0 = \log n$, together with the representation (8), imply that the condition (4) in Theorem 2.1 holds. Thus for some sequence $N_j \to \infty$,

$$\nu = \lim_{j \to \infty} \frac{1}{N_j} \sum_{k=1}^{N_j} c_k \cdot (\delta_{y_k} * \mu_{x,k\log n})|_{[0,1]} \qquad \text{weak-* in } \mathcal{P}([0,1]) \tag{9}$$

Passing to a further subsequence we may assume that the joint distribution of c_k, y_k and the measures converges, i.e. that $\frac{1}{N_j} \sum_{k=0}^{N_j-1} \delta_{(c_k, y_k, \mu_{x,k\log n})}$ converges to a probability

measure Q on $\Omega = \mathbb{R} \times [-1, 1] \times \mathcal{P}(\mathcal{P}([-1, 1]))$ (to see that the distribution of the c_k's is tight, we use Proposition 4.13). Moreover, thanks to Lemma 4.3, the measure marginal of Q is P: this is the point of the proof where the spectral assumption is used.

Taking stock, we have proved the following representation of ν.

Theorem 5.1. *Let μ be a measure on $[0, 1]$ which generates a distribution P at a.e. point, and $\Sigma(P, S) \cap \frac{1}{\log n}\mathbb{Z} = \{0\}$. Then for μ-a.e. x, if x equidistributes under T_n for ν along some subsequence, then there is an auxiliary probability space $(\Omega, \mathcal{F}, Q)$ and measurable functions $c : \Omega \to (0, \infty)$, $y : \Omega \to [-1, 1]$ and $\eta : \Omega \to \mathcal{P}[-1, 1])$, such that η is distributed according to P, and*

$$\nu = \int c_\omega \cdot (\delta_{y_\omega} * \eta_\omega)|_{[0,1]} \, dQ(\omega).$$

Invoking Proposition 4.12, we immediately get:

Corollary 5.2. *A measure ν as in the theorem is of dimension at least δ (the a.s. dimension of measures drawn according to P); in particular, $\dim \nu > 0$.*

The changes needed to prove this for T_β and non-integral Pisot β are minimal. In this case one uses the partition $\mathcal{A}$ of $[0, 1]$ into intervals $[j/\beta, (j + 1)/\beta) \cap [0, 1]$. The main difference is that now the identity (8) is not always true because the length of $\mathcal{A}^k(x)$ is no longer constant, and so the left hand side of (8) is generally the restriction of the right hand side to a shorter interval (followed by normalization). If in (8) we replace the restriction on the right hand side with restriction to the appropriate interval $I_k(x) \subseteq [0, 1]$, then we obtain a representation of the same kind as in Theorem 5.1 but of the form

$$\nu = \int c_\omega \cdot (\delta_{y_\omega} * \eta_\omega)|_{I_\omega} \, dQ(\omega) \tag{10}$$

where $I_\omega \subseteq [0, 1]$ is a random interval. The missing ingredient in this argument is that a-priori the intervals I_k may be vanishingly short for a positive frequency of k, and we must ensure that the distribution of lengths does not concentrate on 0, i.e. we must ensure that I_ω is a.s. of positive length. This is where the Pisot property of β comes into play, via

Lemma 5.3. *There is a constant $c > 1$ such that for any k and interval $I \in \mathcal{A}^k$, the length of I satisfies $c^{-1}\beta^{-k} < |I| < c\beta^{-k}$.*

This is a consequence of a classical lemma of Garsia [24], stated more completely below, see Lemma 6.2. We note that the weaker version stated here continues to hold for the larger class of β for which the β-shift T_β satisfies the specification property, but for these numbers the results of the next section do not appear to hold.

5.3 Resonance and dissonance

As indicated in Section 5.1, the second idea we need for the proof of Theorem 1.1 is that, among invariant measures for T_β of positive dimension, the Parry measure can be

identified by the behavior of its dimension under convolutions. Following terminology of Peres and Shmerkin [45], we say that measures $\mu, \nu \in \mathcal{P}(\mathbb{R})$ **resonate** if

$$\dim \mu * \nu < \min\{1, \dim \mu + \dim \nu\} \tag{11}$$

otherwise they **dissonate**.

As a general rule, measures should dissonate; resonance requires, heuristically, that they have some common structure. This heuristic can be made precise in many ways. For example, as an immediate consequence of Corollary 3.2, we have

Theorem 5.4. *If μ, ν are Borel probability measures on $\mathbb{R}$, then for Lebesgue-a.e. $t \in \mathbb{R}$, the measures μ and $S_t \nu$ dissonate.*

Moreover, suppose that $\dim \mu|_I = \dim \mu$ for any interval I of positive μ-measure. Then for a.e. t, if I is any set of positive $S_t\mu$-measure, then $(S_t\mu)|_I$ and ν dissonate.

Proof. This is a consequence of Theorem 3.1, and elementary properties of $\dim$. $\qquad\square$

Unlike the "generic" case, where dissonance is the rule, for integer n, T_n-invariant measures of dimension strictly between 0 and 1 do resonate, often with themselves and always with other T_n invariant measures. For example consider a T_n-invariant measure μ with $1/2 < \dim \mu < 1$. That μ resonates with itself can be seen as follows. First, $\mu * \mu$ has the same dimension as the dimension of the self-convolution $\nu = \mu * \mu$ with the convolution taken in $\mathbb{R}/\mathbb{Z}$ (this is because the map $\mathbb{R} \to \mathbb{R}/\mathbb{Z}$ is a countable to 1 local isometry). Consider the Fourier transform: $\hat{\nu}(k) = \hat{\mu}(k)^2$. Since μ is not Lebesgue measure it has a non-zero coefficient, hence so does ν, and therefore ν is not Lebesgue measure. But it is a well known fact that the only T_n-invariant measure of dimension 1 is Lebesgue measure, and ν is T_n-invariant; hence $\dim \nu = \dim \mu * \mu < 1 = \min\{1, \dim \mu + \dim \mu\}$.

We will require the following strengthening of the fact above.

Theorem 5.5. *Let $\beta > 1$ be a Pisot number. Then there is a sequence of probability measures $\tau_1, \tau_2, \ldots$ on $\mathbb{R}$ with*

$$\dim \tau_n \to 1 \quad \text{as } n \to \infty,$$

such that any T_β-invariant measure ν with $0 < \dim \nu < 1$ resonates with τ_n for all large enough n.

In order not to interrupt the main line of argument, we postpone the proof to Section 6.

5.4 Proof of Theorem 1.1

Let $\beta > 1$ be a Pisot number. Let $\mu \in \mathcal{P}([0,1))$ generate an S-ergodic and non-trivial distribution P, and suppose that $k/\log \beta$ is not in $\Sigma(P, S)$ for any $k \in \mathbb{Z} \setminus \{0\}$.

Let ν_β be the unique absolutely continuous invariant measure for T_β (the Parry measure). The following fact is standard, but we include a proof as we have not been able to find a reference.

Lemma 5.6. *The measure ν_β is also the unique invariant measure of maximal dimension 1.*

Proof. It is well known that ν_β is the only measure of maximal entropy $\log \beta$ ([30], see also [52, Remark 2.4]). Let $\theta \neq \nu_\beta$ be another invariant measure. By the Shannon-McMillan-Breiman applied to the (generating) partition $\{[k/\beta, (k+1)/\beta) \cap [0,1]\}$, and Lemma 5.3,

$$\underline{\dim}(\theta, x) \leq \lim_{n \to \infty} \frac{\log \theta([x - \beta^{-n}, x + \beta^{-n}])}{n \log \beta} = \frac{h(\theta, x)}{\log \beta},$$

for θ-almost all x, where $h(\theta, x)$ is the entropy of the ergodic component of x. Since $h(\theta) < h(\nu_\beta) = \log \beta$, there is a set of positive measure where the right-hand side above is < 1. In light of the characterization of $\dim$ using local dimensions given in Equation (6), $\dim \theta < 1$, as desired. $\square$

Fix a μ-typical x. It suffices to show that if x equidistributes under T_β along a subsequence for a measure ν, then ν is the unique absolutely continuous T_β-invariant measure ν_β.

From Theorem 5.1 (and the discussion following it for the general Pisot case), we have the representation

$$\nu = \int c_\omega \cdot (\delta_{y_\omega} * \eta_\omega)|_{I_\omega} \, dQ(\omega)$$

where $c_\omega, y_\omega, \eta_\omega, I_\omega$ are defined for ω in some auxiliary probability space $(\Omega, \mathcal{F}, Q)$, and the distribution of η_ω is P. Recalling Proposition 4.12, let $\delta > 0$ denote the a.s. dimension of measures drawn according to P, so also $\dim \eta_\omega = \delta$ a.s. In particular, $\dim \nu > 0$ and ν is non-atomic.

Lemma 5.7. *ν is T_β-invariant.*

Proof. T_β has finitely many discontinuities, and ν is non-atomic, so the set of discontinuities has ν-measure zero. Since ν arises as the measure for which x equidistributes subsequentially, it is T_β-invariant. $\square$

Lemma 5.8. *Let τ be a probability measure on $\mathbb{R}$ with $\dim \tau \geq 1 - \delta$. Then $\dim \tau * \eta_\omega = 1$ for Q-a.e. ω.*

Proof. Using S-invariance of P, Fubini and Theorem 5.4,

$$
\begin{aligned}
\int \dim(\tau * \eta_\omega) \, dQ(\omega) &= \int \dim(\tau * \eta) \, dP(\eta) \\
&= \int_0^1 \int \dim(\tau * \eta) \, dS_t P(\eta) \, dt \\
&= \int \int_0^1 \dim(\tau * S_t \eta) \, dt \, dP(\eta) \\
&= \int \min\{1, \dim \tau + \dim \eta\} \, dP(\eta) \\
&= 1.
\end{aligned}
$$

Since the integrand on the left hand side is ≤ 1, it is a.s. equal to 1, as claimed. $\square$

Now let $\{\tau_n\}$ be the sequence of resonant measures provided by Theorem 5.5. Then $\dim \tau_n \to 1$, so for n large enough we have $\dim \tau_n > 1 - \delta$, hence by linearity of convolution, basic properties of dimension, and the previous lemma,

$$
\begin{aligned}
\dim \tau_n * \nu &= \dim \left(\tau_n * \int c_\omega \cdot (\delta_{y_\omega} * \eta_\omega)|_{[0,1]} \, dQ(\omega) \right) \\
&= \dim \left(\int c_\omega \cdot (\tau_n * \delta_{y_\omega} * \eta_\omega)|_{[0,1]} \, dQ(\omega) \right) \\
&\geq \operatorname{essinf}_{\omega \sim Q} \dim(\tau_n * \delta_{y_\omega} * \eta_\omega|_{[0,1]}) \\
&\geq \operatorname{essinf}_{\omega \sim Q} \dim(\tau_n * \eta_\omega) \\
&= \operatorname{essinf}_{\eta \sim P} \dim(\tau_n * \eta) \\
&= 1.
\end{aligned}
$$

But by choice of τ_n, this is possible only if $\dim \nu = 0$ or 1. Since $\dim \nu > 0$, we must have $\dim \nu = 1$. Lemma 5.6 then allows us to conclude that ν is the Parry measure for T_β, as desired.

This completes the proof of Theorem 1.1

There is a version of Theorem 1.1 for measures which do not generate a distribution. For a measure μ and a typical point x let $\mathcal{D}(\mu, x) \subseteq \mathcal{D}$ denote the set of accumulation points of $\frac{1}{T} \int_0^T \delta_{\mu_{x,t}} dt$ as $t \to \infty$. In [27, Theorem 1.7] it was shown that for μ-a.e. x, this set consists EFDs. An easy adaptation of the proof of the theorem above shows that if μ is a measure such that a.s., $\mathcal{D}(\mu, x)$ contains only non-trivial ergodic distributions which do not have $k / \log n$ in their spectrum, then μ is pointwise n-normal. We shall not give the proof of this in detail.

6 Construction of resonant measures

The proof of Theorem 5.5 is slightly more transparent in the case that β is an integer. After some preliminaries we will prove this case, since it is shorter and may shed light on the general case.

6.1 Preliminaries on entropy

We use standard notation and properties for the entropy $H(\mu, \mathcal{P})$ of a measure μ with respect to a partition $\mathcal{P}$. See [54] or any textbook in ergodic theory for details.

Let $\mathcal{A}^k$ be the partition of $\mathbb{R}$ into k-generation n-adic intervals, that is, intervals $[r/n^k, (r+1)/n^k)$ for $r \in \mathbb{N}$. For a T_n-invariant measure μ, the Kolmogorov-Sinai entropy is given by

$$
h(\mu) = \lim_{k \to \infty} \frac{1}{k} H(\mu, \mathcal{A}^k)
$$

and the limit is also the infimum. In general, $h(\mu) \leq \log n$, with equality if and only if μ

is Lebesgue measure λ. We also have

$$\frac{1}{\log n} h(\mu) \geq \dim \mu$$

with equality if μ is ergodic; in general $\dim \mu$ is the essential infimum over the dimensions (=normalized entropies) of the ergodic components of μ. This follows e.g. from the proof of Lemma 5.6.

The quantity $H(\mu, \mathcal{A}^k)$ is not continuous in μ, however we have the following approximate continuity under translation: If η is a measure supported on an interval of length $< 1/n^k$, then

$$|H(\eta * \mu, \mathcal{A}^k) - H(\mu, \mathcal{A}^k)| < c$$

where c is a universal constant.

6.2   The integer case

Fix an integer $n \geq 2$. Our goal is to construct a sequence of probability measures $\tau_1, \tau_2, \ldots$ on $\mathbb{R}$ such that $\dim \tau_i \to 1$ and any T_n-invariant measure ν with $0 < \dim \nu < 1$ resonates with τ_i for all large enough i.

We will use the standard identification of the map T_n on $[0,1]$ with the shift map on the sequence space $\{0, \ldots, n-1\}^{\mathbb{N}}$, given by the base-$n$ expansion. This is defined uniquely off a countable set of points and hence for non-atomic measures is an a.e. isomorphism, so we will not distinguish between the models.

Let N be an integer and define a measure ν_N on infinite sequences of digits $\{0, \ldots, n-1\}$ as follows. Set the first N digits to be 0. Let the next N^2 digits be chosen independently and equiprobably from $\{0, \ldots, n-1\}$. Repeat this procedure, independently of previous choices, for each subsequent block of $N + N^2$ symbols. Write ν_N also for the corresponding measure on $[0,1]$. Now, this measure is not T_n-invariant but it is $T_n^{N+N^2}$-invariant, so the measure

$$\tau_N = \frac{1}{N + N^2} \sum_{i=0}^{N+N^2-1} T_n^i \nu_N$$

is T_n-invariant.

It is elementary to use Equation (6) to show that $\dim \nu_N = N^2/(N + N^2)$, and so the same is true for $T^i \nu_N$, and hence for τ_N. Thus $\dim \tau_N \to 1$ as $N \to \infty$.

Now let μ be a T_n-invariant measure and suppose that it is not Lebesgue measure. We aim to show that $\dim \tau_N * \mu < 1$ for large enough N. Using the T_n-invariance of μ

and the fact that T_n^i is piecewise affine with constant expansion, we have

$$
\begin{aligned}
\dim \tau_N * \mu &= \dim \left(\frac{1}{N+N^2} \sum_{i=0}^{N+N^2-1} (T_n^i \nu_N) * \mu \right) \\
&= \inf_{0 \le i < N+N^2} \dim (T_n^i \nu_N) * \mu \\
&= \inf_{0 \le i < N+N^2} \dim (T_n^i \nu_N) * (T_n^i \mu) \\
&= \inf_{0 \le i < N+N^2} \dim (\nu_N * \mu) \\
&= \dim \nu_N * \mu.
\end{aligned}
$$

Thus it is enough to show that $\dim(\nu_N * \mu) < 1$ for large enough N, and since $\nu_N * \mu$ is $T_n^{N+N^2}$-invariant, we only need to show that $\nu_N * \mu$ is not Lebesgue. ν_N is concentrated on the interval $[0, n^{-N})$, so we know that

$$
H(\nu_N * \mu, \mathcal{A}^N) < H(\mu, \mathcal{A}^N) + c
$$

where c is a universal constant. Since μ is not Lebesgue, it has less than full entropy, and hence $H(\mu, \mathcal{A}^N) < (1-\varepsilon)N \log n$ for some $\varepsilon > 0$ independent of N. Thus

$$
H(\nu_N * \mu, \mathcal{A}^N) < (1-\varepsilon)N \log n + c < N \log n
$$

for large enough N. Dividing by N and taking the infimum over N we find that $h(\tau_N * \nu) < \log n = h(\lambda)$, where λ is Lebesgue measure, so $\tau_N * \nu \ne \lambda$, as desired.

6.3 Dynamics of beta transformations

We review some basic facts about the beta transformations $T_\beta : x \mapsto \beta x \mod 1$. We refer the reader to the surveys [7, 52] for further information and references.

Recall that $[m] = \{0, \dots, m-1\}$. For each $\beta > 1$, there is a T-invariant closed subset $X_\beta \subset [[\beta]]^{\mathbb{N}}$ (known as the β-**shift**) such that the beta expansion map $\pi : X_\beta \to [0, 1]$, $\pi(x) = \sum_{n=1}^{\infty} x_n \beta^{-n}$ semi-conjugates the action of the shift map T on X_β with the action of T_β on $[0, 1]$. Further, π is injective on X_β, except at countably many points on which it is two-to-one. In particular, any non-atomic T_β invariant measure lifts uniquely to a shift-invariant measure on the β-shift.

We will require the following lemma on the structure of the β-shift for Pisot β.

Lemma 6.1. *Let β be a Pisot number. There exists $N_0 = N_0(\beta) \in \mathbb{N}$ with the following property: let $\{x_i\}$ be finite words in X_β (i.e. X_β contains infinite words starting with each of the x_i). Then the infinite concatenation $(0^N x_1 0^N x_2 \dots)$ is in X_β for $N > N_0$.*

Proof. The following characterization of X_β is essentially due to Parry [44], see also [7, Proposition 2.3]. Let a be the lexicographically least β-expansion of 1, i.e. the lexicographically smallest sequence $a \in [[\beta]]^{\mathbb{N}}$ such that $1 = \sum_{i=1}^{\infty} a_i \beta^i$. Then $x \in X_\beta$ if and only if $T^k x \prec a$ for all k, where $\prec$ denotes lexicographically smaller or equal.

On the other hand, if β is Pisot, then the sequence a is eventually periodic, see [7, Section 4.1]. It cannot end in infinitely many zeros because $(a_1 \ldots a_{k-1}(a_k - 1))^\infty \prec (a_1 \ldots a_k 0^\infty)$ and both sequences represent the same number in base β. It follows that the number of consecutive zeros in a is bounded by some integer N_0. But then it is clear that for any finite words $\{y_i\}$ in X_β, any $N > N_0$ and any $\ell \geq 0$, we have $(0^\ell y_1 0^N y_2 0^N \ldots) \prec a$. This gives the claim. $\qquad\square$

6.4   Resonance in the Pisot case

The proof of the Pisot case is not unlike the integer one. The main difference is that convolutions of T_β-invariant measures are no longer invariant or related in any obvious way to an invariant measure. This makes estimating their dimension more involved.

For the rest of this section we fix a Pisot number $\beta > 1$, and write $B = \lceil \beta \rceil$. Given an integer $D \geq B$ (often implicit), and a $[D]$-valued finite or infinite sequence x of length $|x|$, we let π be the β expansion map, i.e. $\pi(x) = \sum_{k=1}^{|x|} x_k \beta^{-k}$. If $|x| = \infty$, we also write $\pi_k(x) = \pi(x|_{\{1,\ldots,k\}})$. A key role will be played by the following partition of D^k:

$$\mathcal{P}_k = \{\pi^{-1}(\pi x) : x \in [D]^k\}.$$

The property of Pisot numbers that will be used in the proof is given in the following classical Lemma of Garsia [24, Lemma 1.51]:

Lemma 6.2. *There exists $c > 0$ (depending on β and D) such that for any $x, y \in [D]^k$, either $\pi(x) = \pi(y)$, or $|\pi(x) - \pi(y)| \geq c\beta^{-k}$.*

We quote a basic fact for later reference:

Lemma 6.3. *Let $\widetilde{\mu}$ be any measure on $[D]^\mathbb{N}$, and set*

$$a = a_{\beta,D} = \frac{(D-1)\beta^{-1}}{1 - \beta^{-1}}.$$

Then for any set Borel $A \subset \mathbb{R}$ and any $k \in \mathbb{N}$,

1. *$\pi\widetilde{\mu}(A) \leq \pi_k\widetilde{\mu}(A^{(a\beta^{-k})})$,*

2. *$\pi\widetilde{\mu}(A^{(a\beta^{-k})}) \geq \pi_k\widetilde{\mu}(A)$,*

where $A^{(\delta)}$ denotes the δ-neighborhood of A.

Proof. Immediate from the fact that if $x \in [D]^\mathbb{N}$, then

$$|\pi(x) - \pi_k(x)| \leq \sum_{i=k+1}^{\infty} (D-1)\beta^{-i} = a\,\beta^{-k}.$$

$\qquad\square$

The following lemma is similar to [36, Lemma 3].

31

Lemma 6.4. *Let $\widetilde{\mu}$ be an T-invariant measure on $[D]^{\mathbb{N}}$ (as before T is the shift map). Then*

$$\dim \pi\widetilde{\mu} \le \lim_{k\to\infty} \frac{H(\widetilde{\mu}, \mathcal{P}_k)}{k \log \beta} = \inf_{k\ge 1} \frac{H(\widetilde{\mu}, \mathcal{P}_k)}{k \log \beta}.$$

Proof. Write $\mu = \pi\widetilde{\mu}$. For the first inequality, note first that

$$\mu(B(\pi x, (1+a)\beta^{-k})) \ge \widetilde{\mu}(\mathcal{P}_k(x)),$$

for any $x \in [D]^{\mathbb{N}}$, where a is the constant from Lemma 6.3. The inequality follows by combining this and Fatou's lemma applied to the sequence

$$g_k(x) = \frac{\log \widetilde{\mu}(\mathcal{P}_k(x))}{-k \log \beta}.$$

For the second equality, it is enough to show that the sequence $H(\mathcal{P}_k, \widetilde{\mu})$ is subadditive. The partition $\mathcal{P}_k \vee T^{-k}\mathcal{P}_m$ is a refinement of $\mathcal{P}_{m+k}$, since $\pi_k(x)$ and $\pi_m(T^k x)$ determine $\pi_{m+k}(x)$. Thus

$$
\begin{aligned}
H(\mathcal{P}_{m+k}, \widetilde{\mu}) &\le H(\mathcal{P}_k \vee T^{-k}\mathcal{P}_m, \widetilde{\mu}) \\
&\le H(\mathcal{P}_k, \widetilde{\mu}) + H(\mathcal{P}_m, \widetilde{\mu}),
\end{aligned}
$$

using the invariance of μ. $\qquad\square$

Note that in the above we do not assume that $\pi\widetilde{\mu}$ is T_β-invariant.

Lemma 6.5. *Let $\widetilde{\mu}$ be the lift to $X_\beta \subset [B]^{\mathbb{N}}$ of a non-atomic T_β-invariant measure μ. If μ is not the Parry measure, then*

$$\lim_{k\to\infty} \frac{H(\mathcal{P}_k, \widetilde{\mu})}{k} < \log \beta.$$

Proof. We claim that the limit in the left-hand side equals the entropy of μ under T_β; this will imply the lemma since the Parry measure is the unique measure of maximal entropy $\log \beta$.

As before, let $\mathcal{A}$ be the partition of $[0,1]$ into intervals $[j/\beta, (j+1)/\beta) \cap [0,1]$, and $\mathcal{A}^k = \mathcal{A} \vee \cdots \vee T_\beta^{k-1}\mathcal{A}$. Let also $\mathcal{Q}_k$ be the partition of $[0,1]$ into half-open intervals determined by the points $\{\pi(x) : x \in D^n\}$. Since $\widetilde{\mu}$ is supported on X_β, $H(\mathcal{P}_k, \widetilde{\mu}) = H(\mathcal{Q}_k, \mu)$ (the correspondence between the elements of both partitions follows from the fact that X_β is composed of the lexicographically least sequences with a given β expansion, which implies tha the lexicographic order on X_β projects onto the usual order of $[0,1]$, see e.g. [52]).

On the other hand, it is easy to see that $\mathcal{Q}_k$ refines $\mathcal{A}^k$ and, thanks to Garsia's Lemma, each atom of $\mathcal{A}^k$ is the union of a uniformly bounded number of atoms of $\mathcal{Q}_k$. Hence

$$\lim_{k\to\infty} \frac{1}{k} H(\mathcal{P}_k, \widetilde{\mu}) = \lim_{k\to\infty} \frac{1}{k} H(\mathcal{A}^k, \mu) = h(\mu),$$

as claimed. $\qquad\square$

Proof of Theorem 5.5. Let $M \gg N \gg 1$ be large numbers; M will be chosen as a function of N later. We construct a measure $\tau = \tau_{M,N}$ on $[B]^{\mathbb{N}}$ as follows. Let $\widetilde{\lambda}$ be the lift of the Parry measure λ_β to the code space. Now let $\widetilde{\tau}_0$ be the measure on $[B]^{\mathbb{N}}$ defined as follows (compare with the measure constructed in the integer case). The first N digits are 0. The next M digits are chosen according to $\widetilde{\lambda}$. Continue this procedure for each block of $N + M$ digits, with all the choices independent.

As in the integer case, this measure is T^{M+N}-invariant but not T-invariant, so we define

$$\widetilde{\tau} = \widetilde{\tau}_{M,N} = \frac{1}{M+N} \sum_{i=0}^{M+N-1} T^i \widetilde{\tau}_0,$$

which is shift-invariant and ergodic. Lemma 6.1 shows that, provided N is large enough, $\widetilde{\tau}_0$ and hence also $\widetilde{\tau}$ are defined on the β-shift X_β. In particular $\tau = \tau_{N,M} := \pi\widetilde{\tau}$ is T_β-invariant.

Let $\tau_0 = \pi\widetilde{\tau}_0$. The Parry measure λ_β has a bounded density with respect to Lebesgue measure (in the Pisot case it is actually piecewise constant). It follows that if I is an interval determined by two consecutive points of the form $\sum_{i=1}^{(N+M)k} x_i \beta^{-i}$, then $\tau_0(I) \le c^k \beta^{-Mk}$, where $c > 0$ is a constant that depends only on β (in particular, it is independent of M, N and I). By Garsia's Lemma 6.2, any interval of length $2\beta^{-(M+N)k}$ can be covered by a uniformly bounded number of such I, and we conclude that

$$\liminf_{r\downarrow 0} \frac{\log \tau_0([x-r, x+r])}{\log r} \ge \frac{\log c + M \log \beta}{(M+N) \log \beta}.$$

Thus for any N, by taking $M = M(N, c)$ large enough, we can ensure that $\dim \tau = \dim \tau_0 > 1 - 1/N$.

It remains to show that if N is large enough, then for any M, $\dim(\mu * \tau) < 1$. Since μ is invariant, arguing as in the integer case we see that it suffices to show this with τ_0 in place of τ (note that the argument does not use invariance of the convolved measure, only the identity $\dim(T_\beta^i \mu * T_\beta^i \nu) = \dim(\mu * \nu)$, which holds for any map that is piecewise affine with constant slope, in particular T_β).

Note that $\mu * \tau_0$ is the projection of $\mu \times \tau_0$ under the addition map $(x, y) \to x + y$, and hence $\mu * \tau_0 = \pi\widetilde{\rho}$, where $\widetilde{\rho}$ is the image of $\widetilde{\mu} \times \widetilde{\tau}$ on $[B]^{\mathbb{N}} \times [B]^{\mathbb{N}}$ under the map $(x, y) \to (x_i + y_i)_i$ (so that $\widetilde{\rho}$ is defined in $[2B-1]^{\mathbb{N}}$, and it is T-invariant). It follows from Lemma 6.4 (applied with $D = 2B - 1$ and the partitions $\mathcal{P}_k$ defined in terms of D) that

$$\dim(\mu * \tau_0) \le \frac{H(\mathcal{P}_{M+N}, \widetilde{\rho})}{(M+N) \log \beta}. \tag{12}$$

Since, by assumption, $\dim \mu < 1$, we know from Lemma 6.5 that there is $\varepsilon > 0$ such that, if N is large enough, then

$$H(\mathcal{P}_N, \widetilde{\mu}) < (1 - \varepsilon)N \log \beta.$$

Using this, the fact that $\mathcal{P}_N \vee T^{-N}\mathcal{P}_M$ refines $\mathcal{P}_{M+N}$, that $|\mathcal{P}_M| \le C\beta^M$ (by Garsia's Lemma), and that τ is concentrated on $\{x \in [B]^{\mathbb{N}} : x_1 = \cdots = x_N = 0\}$ (which implies

that $\widetilde{\rho}$ and $\widetilde{\mu}$ coincide on $\mathcal{P}_N$), we estimate

$$
\begin{aligned}
H(\mathcal{P}_{M+N}, \widetilde{\rho}) &\leq H(\mathcal{P}_N, \widetilde{\rho}) + H(T^{-N}\mathcal{P}_M, \widetilde{\rho}) \\
&\leq H(\mathcal{P}_N, \widetilde{\rho}) + \log |\mathcal{P}_M| \\
&\leq H(\mathcal{P}_N, \widetilde{\mu}) + M \log \beta + \log C \\
&\leq ((1 - \varepsilon)N + M + \log C) \log \beta.
\end{aligned}
$$

Recalling (12), we conclude that there is N_0 such that for all $N \geq N_0$ and all $M \in \mathbb{N}$,

$$
\dim(\mu * \tau_0) < 1.
$$

This completes the proof. $\qquad\square$

7 Application to iterated function systems: Theorems 1.4, 1.5 and 1.12

7.1 Limit geometries

In this section we fix the following notation. Let $\mathcal{I} = \{f_0 \ldots f_{r-1}\}$ be an IFS on an interval which, without loss of generality, we assume is $[0, 1]$. We will henceforth assume that $\mathcal{I}$ is C^α for some $\alpha > 1$ or $\alpha = \omega$, and regular as defined in the introduction.

Let μ be a quasi-product measure for $\mathcal{I}$. The next lemma contains the key structural information we shall require about μ; it is a manifestation of ideas that go back to Sullivan [53]. We write $\nu_1 \sim_C \nu_2$ to denote that the measures ν_1, ν_2 are mutually absolutely continuous with both Radon-Nikodym densities bounded by C, i.e. $1/C \leq d\nu_1/d\nu_2 \leq C$.

Lemma 7.1. *Then there is $C = C(\mu) > 0$ such that the following holds. Let $x \in \operatorname{supp} \mu$ and let ν be an accumulation point of $\mu_{x,t}$ as $t \to \infty$. Then $\nu \sim_C (g\mu)|_{[-1,1]}$ and $\mu \sim_C (h\nu)|_{[0,1]}$ for some $g, h \in \operatorname{diff}^\alpha(\mathbb{R})$.*

Proof. Given a finite sequence $y \in [r]^n$, let $f_y = f_{y_1} \circ \cdots \circ f_{y_n}$ and $f_y^* = A_y f_y$, where A_y is the renormalizing homothety mapping $f_y([0, 1])$ back to $[0, 1]$. It is proved in [3, Theorems 5.9 and 6.1] that for a left-infinite sequence $y = (y_i)_{i=-\infty}^0$, the sequence $f_{y_{-n} \ldots y_0}^*$ converges, in the C^1 topolofy, to a C^α diffeomorphism F_y^* (these are known as *limit diffeomorphisms*). Moreover, the dependence of F_y^* on y is uniformly continuous. In particular, the family $\{f_y^*\}$, where y ranges over all finite words, is relatively compact in the C^1 topology.

For the first part, $\nu \sim_C (g\mu)|_{[-1,1]}$, one notes that $\mu_{x,t}$ is C-equivalent to a bounded translation, a restriction and normalization of $S_s f_y^* \mu$ for an appropriate word $y = y(t)$, whose length tends to ∞ with t, and some $s = s(t) \in [0, L]$, where L depends only on the IFS (one can take L to be the maximum of $|\log f_i'(x)|$ over $i \in [r]$ and $x \in [0, 1]$). Also, the space of measures C-equivalent to μ is weak-* closed. Thus, up to passing to a subsequence, ν is C-equivalent to a translation, restriction and normalization of $S_s F \mu$ for some limit diffeomorphism F and $s \in [0, L]$.

For the second statement, note that it follows from the first part that $\nu|_I \sim_C g(\mu|_J)$ for suitable intervals I, J. Moreover, we can take $J = f_y([0,1])$ for some word y. In this case $\mu|_J \sim_C f_y\mu$ by the quasi-product property, so the claim follows with $h = (gf_y)^{-1}$ (and a different value of C). $\qquad\square$

Observe that by Lemmas 4.16 and 7.1, if ν is any accumulation point of $\mu_{x,t}$, and ν generates P, then so does μ. This fact will be key in the proof of the following theorem.

Theorem 7.2. *μ generates an S-ergodic, non-trivial distribution P. Furthermore, there is $C > 0$, such that P-a.e. ν satisfies $\nu \sim_C (g\mu)|_{[-1,1]}$ for some $g \in \mathrm{diff}^\alpha(\mathbb{R})$. .*

Proof. The proof of the first part is essentially identical to [27, Proposition 1.36]. We sketch the details for completeness.

For μ-a.e. x, if the scenery at x equidistributes for P along a subsequence of times $T_k \to \infty$, then P is S-invariant and S-quasi-Palm [27, Theorem 1.7]. By the ergodic theorem and the S-quasi-Palm property, P-almost all measures ν generate the S-ergodic component P_ν of ν (this argument holds for general measures μ with no additional assumptions).

Let x be a μ-typical point, and let the scenery at x equidistribute for P along a subsequence (such subsequence exists by compactness). By the previous paragraph, a P-typical measure ν generates an S-ergodic distribution P_ν. By the remark preceding the theorem this means that μ generates P_ν, and the generated distribution is S-ergodic. The second part is just Lemma 7.1 and the fact that P is supported on accumulation points of sceneries of μ. $\qquad\square$

7.2 Proofs of Theorems 1.4, 1.5 and 1.12

Proof of Theorem 1.4. Let μ be a quasi-product measure for a $C^{1+\varepsilon}$-IFS $\mathcal{I}$ such that $\lambda(f) \not\sim \lambda(g)$ for some $f, g \in \mathcal{I}$. We want to show that μ is pointwise β-normal for any Pisot $\beta > 1$. We have already established in Theorem 7.2 that μ generates an EFD P, and our aim is to apply Theorem 1.2 to μ, so we must show that $\Sigma(S, P) \cap \frac{1}{\beta}\mathbb{Z} = \{0\}$. To this end, we shall show that if $k/\log\beta \in \Sigma(P, S)$, then $\lambda(f) \sim \beta$ for all $f \in \mathcal{I}$.

We argue by contradiction. Let $t_0 = k/\log\beta \in \Sigma(P, S)$ such that $\lambda(f) \not\sim \beta$ for some $f \in \mathcal{I}$. Hence

$$e(t_0 \log \lambda(f)) \neq 1. \tag{13}$$

Let ν be a P-typical measure; we know from Proposition 4.15 that the phase measure θ_ν of the eigenvalue corresponding to t_0 is (well defined, and) a single atom. Now by Lemma 7.1, ν is also (the restriction of) a quasi-product measure for a conjugated C^α IFS $\mathcal{J} = g\mathcal{I} = \{gfg^{-1} : f \in \mathcal{I}\}$. Since $\lambda(gfg^{-1}) = \lambda(f)$, we may assume without loss of generality that, already for the original measure μ, the phase measure is an atom, say δ_z.

Let x_0 be the fixed point of f (this is in the support of μ). Let U be a small interval centered at x_0. Since $\mu_U \ll \mu$ and $f(\mu_U) \ll \mu$, by the first part of Corollary 4.17, the phase measures of $\frac{1}{\mu(U)}\mu|_U$ and $\nu_U := \frac{1}{\mu(U)}f(\mu|_U)$ are (well defined and) equal to δ_z. By the second part of Corollary 4.17, the phase measure of ν_U equals

35

$\frac{1}{\mu(U)} \int_U \delta_{e(-t_0 \log f'(x))z} d\mu(x)$. Since f' is continuous, this shows that as the size of U tends to 0, the support of θ_{ν_U} tends to $\{e(-t_0 \log \lambda(f))z\}$. In light of (13), this is a contradiction, as desired. $\qquad \square$

We sketch an alternative proof of Theorem 1.4 which does not directly use EFDs and instead relies on the results from [29] on dissonance between quasi-product measures for regular IFSs. The overall strategy is the same, with the main difference coming in the part that establishes dissonance. Namely, for a μ-typical x, suppose x equidistributes for ν under T_β along a sequence $N_k \to \infty$. Using Lemma 7.1, one can check that a representation (10) still holds, except that a priori we do not know that the measure marginal of Q is P; however, it is easily seen to be supported on limits of sceneries of μ which, as we know from Lemma 7.1, are restrictions of quasi-product measures for a smoothly conjugated IFS. Now since the measures τ_n constructed in Theorem 5.5 are (convex combinations of) quasi-product measures for a homogeneous affine IFS with contraction ratio $\sim \beta$, it follows from [29, Theorem 1.4] that the measures η_ω in (10) dissonate with the τ_n (this is the step that uses that $\lambda(f) \approx \beta$ for some $f \in \mathcal{I}$), and hence so does ν if $\dim(\tau_n)$ is large enough. This contradicts Theorem 5.5 unless ν is the Parry measure.

Proof of Theorem 1.5. Let μ be a quasi-product measure for a real-analytic, totally non-linear IFS $\mathcal{I}$. We know that μ generates an EFD P; we will again show that P is weak-mixing, i.e. $\Sigma(P, S) = \{0\}$. Together with Theorem 1.2, this will yield the result.

Suppose for contradiction that $0 \neq t_0 \in \Sigma(P, S)$. We follow the scheme of the proof of Theorem 1.4: instead of working with the original measure μ, we consider a P-typical measure ν such that the phase measure is an atom δ_z. This measure is a restriction of the attractor of a conjugated IFS $g\mathcal{I}$, which is also real-analytic. Since $\mathcal{I}$ is totally non-linear, $g\mathcal{I}$ is nonlinear, hence it contains a non-affine analytic map h. We can then find a non-trivial interval U meeting the support of ν, on which h' is strictly monotone (this is the point where analyticity gets used; if the IFS was merely C^2, a priori h' might have no point of strict monotonicity on the Cantor set $\operatorname{supp} \nu$). Now arguing as in the proof of Theorem 1.4, on one hand the phase measure of $\frac{1}{\nu(U)} h(\nu|_U)$ is δ_z, and on the other hand it equals $\frac{1}{\nu(U)} \int_U \delta_{e(-t_0 \log h'(x))z} d\nu(x)$. The latter measure clearly cannot be atomic, so we have reached the desired contradiction. $\qquad \square$

The facts on the spectrum $\Sigma(P, S)$ that emerged in the above proofs may find other applications, so we summarize them below.

Theorem 7.3. *Let $\mathcal{I}$ be a $C^{1+\varepsilon}$ IFS, μ a quasi-product measure for it, and P the distribution generated by μ.*

1. *Suppose that $\lambda(f) \approx t_0$ for some $f \in \mathcal{I}$. Then $k/\log t_0 \notin \Sigma(P, S)$ for any $k \in \mathbb{Z} \setminus \{0\}$. In particular, if $\lambda(f_1) \approx \lambda(f_2)$ for $f_1, f_2 \in \mathcal{I}$, then $\Sigma(P, S) = \{0\}$.*

2. *If $\mathcal{I}$ is C^ω and totally non-linear or, more generally, if $\mathcal{I}$ has the property that for any limit diffeomorphism g, the conjugated IFS $g\mathcal{I}$ contains a map h such that h' is a local diffeomorphism, then $\Sigma(P, S) = \{0\}$.*

To conclude this section, we present the deduction of Theorem 1.12 from Theorem 1.2.

Proof of Theorem 1.12. Let a, b be two distinct elements of Λ. Write x_i for the fixed point of the inverse branch of the Gauss map $f_i(x) = 1/(x+i)$. Then x_a and x_b are quadratic numbers generating distinct quadratic fields. It follows that $\lambda(f_a^2) = x_a^4 \not\sim x_b^4 = \lambda(f_b^2)$. Hence for any Pisot $\beta > 1$, either $\beta \not\sim \lambda(f_a^2)$ or $\beta \not\sim \lambda(f_b^2)$; by Theorem 1.4, any quasi-product measure on C_Λ is pointwise β-normal. $\square$

8   A refinement of Theorem 1.1 and applications

8.1   Relaxing the spectral hypothesis

For an integer n, any T_n-invariant and ergodic measure μ generate an EFD P, see [26]. This P can be rather explicitly described, and its spectrum can be shown to contain non-zero integer multiples of $\frac{1}{\log m}$ only if either $m \sim n$ or $\log n / \log m \in \Sigma(T, \mu)$. Thus in many cases the pointwise m-normality of μ follows directly from Theorem 1.2. In order to deal with the remaining cases we now present some refinements of Theorem 1.2, in which $k/\log \beta$ is present in the spectrum of P, but instead we assume that the phase is "sufficiently spread out". We give two versions, the first being simpler to state:

Theorem 8.1. *Let $\beta > 1$ be a Pisot number. Let $\mu \in \mathcal{P}([0,1])$ and suppose that μ generates an S-ergodic and non-trivial distribution P which is not $S_{\log \beta}$-ergodic (so that $k/\log \beta \in \Sigma(P, S)$). Further, assume that $\mu \log \beta$-generates an $S_{\log \beta}$-ergodic distribution P_x at μ-a.e. point x. Let $\theta = \theta_\mu$ denote the associated phase measure as described in Section 4.3. If $\dim \theta = 1$, then μ is pointwise β-normal.*

One consequence is that if $\int P_x d\mu(x)$ is S-invariant, then μ is pointwise β-normal, as it is clear that in this case the phase measure is invariant under rotations of the circle hence is normalized length measure. Although the theorem above is strong enough for applications, the proofs become simpler using the following variant:

Theorem 8.2. *Let $\beta > 1$ be a Pisot number. Let $\{\mu_\omega\}_{\omega \in \Omega} \subseteq \mathcal{P}(\mathbb{R})$ be a measurable family defined on a probability space $(\Omega, \mathcal{F}, Q)$. Suppose that there is an S-ergodic and non-trivial distribution P, which is not $S_{\log \beta}$-ergodic, and such that Q-a.e. μ_ω generates P and at a.e. point $\log \beta$-generates an $S_{\log \beta}$-ergodic distribution. Let θ_{μ_ω} denote the associated phase measures and $\theta = \int \theta_{\mu_\omega} \, dQ(\omega)$ the "cumulative" phase measure. If $\dim \theta = 1$, then μ_ω is pointwise β-normal for Q-a.e. ω, and hence also $\mu = \int \mu_\omega \, dQ(\omega)$ is pointwise β-normal.*

It is clear that the first theorem follows from the second by taking $\mu_\omega = \mu$ for all ω. Nevertheless we shall prove the first, and then explain the changes needed for the second. The proof of Theorem 8.1 follows the scheme of the proof of Theorem 1.2 detailed in Section 5.1, with a minimal change to the first step and a more significant change in the proof of the third step, in particular making use of the stronger version of Marstrand's projection theorem given in Theorem 3.3.

For the rest of this section, suppose that β and μ are as in the statement of Theorem 8.1. In particular, let $\theta = \theta_\mu$ be the associated phase measure with respect to an

appropriate eigenfunction φ of (P, S), as in Section 4.3. Fix a μ-typical x_0 for which P_{x_0} is defined. For μ-typical x, define a function $\ell(x) \in [0, 1)$ by $\varphi_\mu(x) = e(\ell(x))\varphi_\mu(x_0)$. It follows from the fact that $P = \int_0^{\log \beta} S_t P_x dt$ and the eigenfunction property that $P_x = S_{\ell(x)} P_{x_0}$. Since $\ell(x)$ depends only on $\varphi_\mu(x)$, we will also denote $\ell(z) = \ell(x)$ where $z = \varphi_\mu(x)$, or in other words $e(\ell(z)) = z/\varphi_\mu(x_0)$. In particular, $\dim(\ell\theta) = \dim \theta = 1$.

Let δ denote the almost-sure dimension of measures drawn from P; it is also the a.s. dimension of measures drawn from P_x for μ-almost all x. Recall from Proposition 4.12 that $\delta > 0$. The following is a refined version of Lemma 5.8.

Lemma 8.3. *Let τ be a probability measure on $\mathbb{R}$ with $\dim \tau \geq 1 - \delta$. Then $\dim \tau * \eta = 1$ for μ-a.e. x and P_x-a.e. η.*

Proof. Using Fubini, the fact that $\dim(\ell\theta) = 1$, and Theorem 3.3,

$$\int \int \dim(\tau * \eta)\, dP_x(\eta)\, d\mu(x) = \int \int \dim(\tau * \eta)\, dS_{\ell(x)} P_{x_0}(\eta)\, d\mu(x) \qquad (14)$$

$$= \int \int \dim(\tau * \eta)\, dS_{\ell(z)} P_{x_0}(\eta)\, d\theta(z)$$

$$= \int \int \dim(\tau * S_{\ell(z)}\eta)\, d\theta(z)\, dP_{x_0}(\eta)$$

$$= \int \int \dim(\tau * S_t\eta)\, d\ell\theta(t)\, dP_{x_0}(\eta)$$

$$\geq \int \min\{1, \dim \tau + \dim \eta\}\, dP_{x_0}(\eta)$$

$$= \min\{1, \dim \tau + \delta\}$$

$$= 1.$$

But the integrand on the left hand side is ≤ 1, so it is a.s. equal to 1, as claimed. $\qquad\square$

We can now finish the proof of the theorem.

Proof of Theorem 8.1. For μ-typical x, the analog of Lemma 4.3 holds for the distributions P_x by assumption. It follows that for μ-a.e. x and any measure ν for which x equidistributes under T_β sub-sequentially, we have a representation similar to Theorem 5.1:

$$\nu = \int c_\omega \cdot (\delta_{y_\omega} * \eta_\omega)|_{I_\omega}\, dQ(\omega)$$

where $c_\omega, y_\omega, \eta_\omega, I_\omega$ are defined on some auxiliary probability space $(\Omega, \mathcal{F}, Q)$, and η_ω is distributed as P_x (rather than P).

The proof is now concluded exactly in the same way as in Theorem 1.1. Combining the integral representation with Lemma 8.3, for μ-a.e. x and any ν for which x equidistributes sub-sequentially under T_β, we have that ν dissonates with every measure of large enough dimension, and also that $\dim \nu \geq \delta$. But, by Theorem 5.5, this is possible only if ν is of dimension 1, hence the unique absolutely continuous measure for T_β. This completes the proof. $\qquad\square$

38

As for Theorem 8.2, the argument is identical, except that in equation (14) one replaces μ by μ_ω and integrates $dQ(\omega)$. We leave the remaining details to the reader.

8.2 Distributions associated to T_γ invariant measures

Let $\gamma > 1$ and μ a T_γ-invariant and ergodic measure with $\dim \mu > 0$. In this section we develop some background about such measures and distributions associated to them. This is a minor adaptation of [26, Section 3], which dealt with the integer case (though the language we employ here is slightly different).

Let $G = \lceil \gamma \rceil$. We have already met the γ-shift $X_\gamma \subseteq [G]^{\mathbb{N}}$, which, together with the shift map T, factors onto $([0,1], T_\gamma)$, and have noted that μ lifts uniquely to X_γ. We also will need the so-called natural extension: let $\widetilde{X}_\gamma \subseteq [G]^{\mathbb{Z}}$ denote two-sided γ-shift, i.e. the set of bi-infinite sequences all of whose subwords appear in the one-sided γ-shift X_γ. For $\omega \in \widetilde{X}_\gamma$ let $\omega^+ = (\omega_1, \omega_2, \ldots)$ and $\omega^- = (\ldots, \omega_{-1}, \omega_0)$, and also write $x(\omega) = \pi(\omega^+)$, where $\pi : X_\gamma \to [0,1]$ is the usual base-γ coding map. It is a standard fact that μ lifts uniquely to a T-invariant measure $\widetilde{\mu}$ on $(\widetilde{X}_\gamma, T)$.

For $\widetilde{\mu}$-typical ω, let μ_ω denote the conditional measure of $\widetilde{\mu}$ given the "past" $(\ldots, \omega_{-1}, \omega_0)$. These conditional measures can be defined abstractly as the disintegration of $\widetilde{\mu}$ given the measurable and countably generated partition into different pasts, see [16, Theorem 5.14], or more concretely by the conditions

$$\mu_\omega[i_1 \cdots i_k] = \lim_{n \to \infty} \frac{\mu[\omega_{-n} \ldots \omega_0 i_1 \ldots i_k]}{\mu[\omega_{-n} \ldots \omega_0]}.$$

(That the limit exists for $\widetilde{\mu}$-a.e. ω can be seen from a martingale argument.)

These conditional measures are measures on the "future" $[G]^{\mathbb{N}}$ and almost surely are supported on the one-sided γ-shift. We silently shall identify μ_ω with the corresponding measure $\pi \mu_\omega$ on $[0,1]$. It is well known that $\dim \mu_\omega = \dim \mu$ a.s.

Definition 8.4. *A distribution $P_0 \in \mathcal{D}$ is S_{t_0}-**quasi-Palm** if it is S_{t_0}-invariant, gives full mass to $\mathcal{M}^{\square}$, and for every Borel set $B \subset \mathcal{M}^{\square}$ with $P(B) = 1$ and every $k \in \mathbb{N}$, P-almost every measure η satisfies $\eta_{x, kt_0} \in B$ for η-almost all x such that $[x - e^{-kt_0}, x + e^{kt_0}] \subset [-1, 1]$.*

If P_0 is S_{t_0}-quasi-Palm, it is easy to see that $P = \frac{1}{t_0} \int_0^{t_0} S_t P_0 \, dt$ is S-quasi-Palm.

Definition 8.5. *An S_{t_0}-invariant and ergodic distribution which is also S_{t_0}-quasi-Palm is a t_0-**discrete ergodic fractal distribution**.*

It is again clear that if P_0 is such a distribution then $P = \frac{1}{t_0} \int_0^{t_0} S_t P_0 \, dt$ is an EFD.

Theorem 8.6. *Let μ be T_γ-invariant and ergodic. Then there is a $1/\log \gamma$-discrete EFD P_0 and a factor map $\sigma : (\widetilde{X}_\gamma, \widetilde{\mu}, T) \to (\mathcal{M}^{\square}, P_0, S_{\log \gamma})$, such that for $\widetilde{\mu}$-a.e. ω,*

$$(\mu_\omega)_{x(\omega)} \ll \sigma(\omega). \tag{15}$$

(actually the two measures are proportional on the interval $[-x(\omega), 1 - x(\omega)]$).

Proof. The factor map in question is defined by

$$\sigma(\omega) = \lim_{n \to \infty} S_{n \log \gamma}((\mu_{T^{-n}\omega})_{x(T^{-n}\omega)}),$$

and P_0 is the push-forward of $\widetilde{\mu}$ through this map. The S-quasi-Palm property is a consequence of the fact that the distribution of $\mu_{T\omega}$ for $\omega \sim \widetilde{\mu}$ is equal in distribution to μ_ω for $\omega \sim \widetilde{\mu}$. For a more detailed verification of the integer case, see [26, Theorem 3.1]; there are no substantial changes when passing to a general $\gamma > 1$. □

Let P_0 be as in Theorem 8.6, and

$$P = \frac{1}{\log \gamma} \int_0^{\log \gamma} S_t P_0 \, dt$$

which, as was already noted, is an EFD.

Proposition 8.7. *1. P_0 is $\log \gamma$-generated by μ_ω at μ_ω-a.e. point, for $\widetilde{\mu}$-a.e. ω.*

2. μ_ω generates P for $\widetilde{\mu}$-a.e. ω.

Proof. (1) By the ergodic theorem, P_0-a.e. measure ν generates P_0 $\log \gamma$-discretely at 0, and by the $S_{\log \gamma}$-quasi-Palm property of P_0, 0 can be replaced by ν-typical x (this argument is the same as the proof of Lemma 4.9). Thus $\sigma(\omega)$ generates P_0 $\log \gamma$-discretely, and using (15), the same is true for μ_ω.

(2) is a consequence of (1). Let $f \in C(\mathcal{P}([-1,1]))$ and let

$$F(\nu) = \frac{1}{\log \gamma} \int_0^{\log \gamma} f(S_t \nu) \, dt.$$

Let $\nu = \mu_\omega$ for a typical ω. We must show that

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(\nu_{x,t}) \, dt = \int f \, dP_0 \qquad \text{for } \nu\text{-a.e. } x.$$

But

$$\frac{1}{T} \int_0^T f(\nu_{x,t}) \, dt = \frac{1}{\lfloor T \rfloor} \sum_{n=0}^{\lfloor T \rfloor} F(\nu_{x,n}) + O\left(\frac{\|f\|_\infty}{T}\right).$$

If F were continuous on $\mathcal{P}([0,1])$, convergence above would follow immediately from the fact that P_0 is $\log \gamma$-discretely generated by ν at ν-a.e. point. In fact, F is defined only on $\mathcal{M}^\square$, but it is continuous on a P_0-full measure set (such as the set of atomless measures in $\mathcal{M}^\square$), so the result still follows, see e.g. [6, Theorem 2.7]. □

8.3 Proof of Theorem 1.10: normality of μ

Let $\gamma > 1$, let μ be a T_γ-invariant and ergodic measure, and let $\beta > 1$ be a Pisot number with $\gamma \nsim \beta$. Our goal is to show that μ is pointwise β-normal, and so is $f\mu$ when $f \in \mathrm{diff}^2(\mathbb{R})$. We continue with the notation of the previous section: $\widetilde{\mu}, \mu_\omega, P_0, P$ etc.

We will first show that μ is pointwise β-normal; the case of $f\mu$ for $f \in \mathrm{diff}^2(\mathbb{R})$ will be handled in the next section.

Suppose first that $\Sigma(P, S)$ does not contain non-zero integer multiples of $1/\log\beta$. Then by Proposition 8.7(2) and Theorem 1.2, for $\widetilde{\mu}$ a.e. ω the conditional measure μ_ω is pointwise β-normal. But then so is μ since $\mu = \int \mu_\omega \, d\widetilde{\mu}(\omega)$.

Therefore, assume that there is some integer $k \neq 0$ with $k/\log\beta \in \Sigma(P, S)$, or, equivalently, that P is not $S_{\log\beta}$-ergodic. Our goal is to verify that the assumptions of Theorem 8.2 are met. In light of Proposition 8.7, and setting $\Omega = \widetilde{X}_\gamma$ and $Q = \widetilde{\mu}$, we see that all that remains to be checked is that the cumulative phase measure $\theta = \int \theta_{\mu_\omega} \widetilde{\mu}(\omega)$ has full dimension.

Recall that P-a.e. ν equidistributes under $S_{\log\beta}$ for an $S_{\log\beta}$-ergodic distribution (Lemma 4.2). This is an S-invariant property, so it holds for P_0-a.e. ν, and by the $S_{\log\lambda}$-quasi-Palm property, the relation (15), and Lemma 8.7, the same holds for ν^x at ν-a.e. x for P_0-a.e. ν.

Fix an eigenfunction φ for the eigenvalue $k/\log\beta$. The assumption $\beta \nsim \gamma$ comes in during the proof of the next lemma.

Lemma 8.8. $\theta' = \int \theta_\nu \, dP_0\nu$ is Lebesgue measure on the circle.

Proof. To begin, note that either $P_0 = P$ or else P_0 is the level set of an eigenfunction ψ with eigenvalue $m/\log\gamma$ for some non-zero integer m. In the first case the assertion is clear, since θ' is a translation-invariant measure on the circle, so we consider the second case only. Switching to additive notation, the map $\nu \mapsto (\varphi(\nu), \psi(\nu))$ defines a factor map from (P, S) to the torus equipped with translation by $(k/\log\beta, m/\log\gamma)$. Since $\beta \nsim \gamma$, Lebesgue measure is the unique invariant measure for this translation, and we deduce that the distribution of φ conditioned on any level set of ψ is uniform on the circle; in particular, the distribution of φ on P_0 is uniform on the circle. Now the lemma follows since, by Proposition 4.15, $\theta_\nu = \delta_{\varphi(\nu)}$ for P-a.e. ν and hence, by S-invariance, for P_0-a.e. ν. $\qquad\square$

Since $\mu_\omega \ll \sigma(\omega)$ for $\widetilde{\mu}$-a.e. ω, and θ_ν is a single atom, we have $\theta_{\mu_\omega} = \theta_{\sigma(\omega)}$, hence $\theta = \int \theta_{\mu_\omega} \, d\widetilde{\mu}(\omega)$ is uniform on the circle, in particular of dimension 1. We have shown all the hypotheses of Theorem 8.2 hold, so this proves the pointwise β-normality of μ.

8.4 Conclusion of the proof of Theorem 1.10: normality of $f\mu$

It remains for us to prove pointwise β-normality of $f\mu$ for $f \in \mathrm{diff}^2(\mathbb{R})$; again we will do so by applying Theorem 8.2. Since μ_ω and $f\mu_\omega$ generate and $\log\beta$-generate the same distributions, the task is to show that the cumulative phase measure has dimension 1. By Corollary 4.17, this measure is given by

$$\theta' = \int\int \delta_{e(\log(-f'(x))/\log\gamma)\cdot\varphi(\mu_\omega)} \, d\mu_\omega(x) \, d\widetilde{\mu}(\omega). \tag{16}$$

In order to be able to apply projection results, we pass from multiplicative to additive notation; in particular the range of φ becomes the unit interval $[0, 1]$. Define the measure η on $[0, 1]^2$ by

$$\eta = \int \mu_\omega \times \delta_{\varphi(\mu_\omega)} \, d\widetilde{\mu}(\omega)$$

Then θ' is the projection of η by the map

$$\pi(x, y) = y - \log f'(x)/\log\gamma.$$

Now, by definition the projection $P_2\eta$ of η to the y-axis is $\int \delta_{\varphi(\mu_\omega)} d\widetilde{\mu}(\omega)$, which we have seen is Lebesgue measure. Also, since $\dim\mu_\omega = \dim\mu$ a.s., by Lemma 3.5(1) we find that

$$\dim\eta \geq 1 + \dim\mu.$$

Note that the map $F(x, y) = (x, \pi(x, y))$ preserves dimension: indeed, since $f \in \mathrm{diff}^2(\mathbb{R})$, we have that f' is differentiable, and one easily computes and finds that F is nonsingular. Thus the image $\widetilde{\eta} = F\eta$ has dimension $\dim\eta \geq 1 + \dim\mu$. Also note that $P_1\widetilde{\eta} = P_1\eta = \int \mu_\omega \, d\widetilde{\mu}(\omega) = \mu$. Since μ is exact-dimensional, $\dim_P(P_1\widetilde{\eta}) = \dim\mu$ (see the discussion at the end of Section 3.3).

On the other hand, $P_2\widetilde{\eta} = \theta'$ by definition. Thus, applying Lemma 3.5(2) to $\widetilde{\eta}$, we conclude

$$\dim\theta' \geq \dim\widetilde{\eta} - \dim_P(P_1\widetilde{\eta})$$
$$\geq \dim\eta - \dim\mu$$
$$\geq 1.$$

Summarizing, we have shown that $\dim\theta' = 1$, hence we can apply Theorem 8.2 to $f\mu$. This completes the proof of Theorem 1.10.

8.5 Proof of Theorem 1.7

We now assume that $\mathcal{I}$ consists of linear maps, μ is self-similar, and $f \in \mathrm{diff}^\omega(\mathbb{R})$ is not affine. Our aim is to prove Theorem 1.7, asserting the β-normality of $f\mu$ for all Pisot $\beta > 1$. The argument is similar to what we have already seen, except that the classical projection theorems are not strong enough and we rely instead on a recent result from [28] that gives stronger bounds for self-similar measures.

If $\mathcal{I}$ contains two maps with contraction ratios $\lambda_1 \nsim \lambda_2$, then it follows from Theorem 1.4 that $f\mu$ is normal to all Pisot bases for all $f \in \mathrm{diff}^2(\mathbb{R})$ (and in fact $f \in \mathrm{diff}^1(\mathbb{R})$ is enough in this case). Thus we may assume $\lambda(f_i) \sim \gamma^{p_i}$ for all $f_i \in \mathcal{I}$ and integers p_i. Let $\beta > 1$ be Pisot. Again, if $\gamma \nsim \beta$ then we are done by Theorem 1.4, so we assume that $\beta \sim \gamma$.

Lemma 8.9. *μ generates an ergodic distribution P with $\Sigma(P, S) \subseteq (1/\log\gamma)\mathbb{Q}$. For each eigenvalue, the phase measure is well defined and consists of a single atom.*

Proof. The first statement follows from the fact that the measure in question is a quasi-product measure, and from Theorems 7.2 and 7.3. Because the contractions are linear, it is elementary to see that for every accumulation point ν of $\mu_{x,t}$, there is a linear map f and interval I with $f\mu = c \cdot \nu|_I$ for a normalizing constant c. The statement about the phase then follows from Proposition 4.15. □

Let P be as in the lemma, and fix an eigenvalue α of P with associated eigenfunction φ. Let $f \in \operatorname{diff}^\omega(\mathbb{R})$ be non-linear. By Corollary 4.17, we know that the phase distribution of $f\mu$ is, up smooth coordinate change and in additive notation, the push-forward of μ through f'. Since f is real analytic and non-linear, f' is a piecewise diffeomorphism and so $\dim f'\mu = \dim \mu > 0$.

Now let τ_n be the sequence of eventually-resonant measures for T_β-invariant measures provided by Theorem 5.5, and observe from the construction of τ_n that they are in fact affine combinations of self-similar measures with uniform contraction ratio $\sim \beta$ satisfying the open set condition. Arguing through the proof of Theorem 8.1, we find that the following lemma, which replaces Lemma 8.3, allows the proof to carry through.

Lemma 8.10. *Let $\nu = f\mu$ and $\theta = \theta_\nu$. Let P_x denote the distribution that is $\log\gamma$-generated by ν at x. Let τ be a self-similar measure for an IFS with uniform contraction ratio a power of γ, satisfying the open set condition, and satisfying $\dim \tau + \dim \nu \geq 1$. Then for ν-a.e. x and P_x-a.e. η, we have $\dim \tau * \eta = \min\{1, \dim \tau + \dim \eta\}$.*

Proof. We have already noted that $\dim \theta > 0$. We now calculate exactly as in the proof of Lemma 8.3. The only change is that we cannot use Theorem 3.3 to deduce that $\dim \tau * S_{\ell(z)}\eta = 1$ for θ-almost all z, since all we know about θ is that $\dim \theta > 0$.

Instead, note that up to a translation, $S_t\eta$ is absolutely continuous with respect to the measure η scaled by e^{-t}, whence $\tau * S_t\eta$ is absolutely continuous with respect to the image of the self-similar measure $\tau \times \mu$ via the linear map $(x,y) \mapsto x + e^{-t}y$. Now, for $\tau = \tau_n$, both μ and τ are self-similar with contraction ratios $\sim \beta$. When the contraction ratios of $\mathcal{I}$ are uniform, $\tau \times \mu$ is also self-similar and of dimension > 1 (for large n), and we can invoke Theorem [28, Theorem 1.8], which implies that $\dim \tau * S_t\eta = 1$ for all t outside a set of Hausdorff dimension 0, and hence of θ-measure 0. In the non-uniformly contracting case a minor (but not short) modification of the arguments in [28] is needed; this will appear separately. □

This completes the proof of Theorem 1.7.

References

[1] Magnus Aspenberg, Fredrik Ekström, Tomas Persson, and Jörg Schmeling. On the asymptotics of the scenery flow. preprint, available at `http://arxiv.org/abs/1309.1619`, 2013.

[2] Luis Barreira. Nonadditive thermodynamic formalism: equilibrium and Gibbs measures. *Discrete Contin. Dyn. Syst.*, 16(2):279–305, 2006.

[3] 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.

[4] Anne Bertrand. Développements en base de Pisot et répartition modulo 1. *C. R. Acad. Sci. Paris Sér. A-B*, 285(6):A419–A421, 1977.

[5] Anne Bertrand. Répartition modulo 1 et développement en base θ. *C. R. Acad. Sci. Paris Sér. A-B*, 289(1):A1–A4, 1979.

[6] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[7] F. Blanchard. β-expansions and symbolic dynamics. *Theoret. Comput. Sci.*, 65(2):131–141, 1989.

[8] Ryan Broderick, Yann Bugeaud, Lior Fishman, Dmitry Kleinbock, and Barak Weiss. Schmidt's game, fractals, and numbers normal to no base. *Math. Res. Lett.*, 17(2):307–321, 2010.

[9] Gavin Brown, William Moran, and Charles E. M. Pearce. Riesz products and normal numbers. *J. London Math. Soc. (2)*, 32(1):12–18, 1985.

[10] Gavin Brown, William Moran, and Charles E. M. Pearce. Riesz products, Hausdorff dimension and normal numbers. *Math. Proc. Cambridge Philos. Soc.*, 101(3):529–540, 1987.

[11] Gavin Brown, William Moran, and Andrew D. Pollington. Normality to noninteger bases. *C. R. Acad. Sci. Paris Sér. I Math.*, 316(12):1241–1244, 1993.

[12] Yann Bugeaud. *Distribution Modulo one and Diophantine Approximation*, volume 193 of *Cambridge Tracts in Mathematics*. Cambridge University Press, 2012.

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

[14] H. Davenport, P. Erdős, and W. J. LeVeque. On Weyl's criterion for uniform distribution. *Michigan Math. J.*, 10:311–314, 1963.

[15] Manfred Einsiedler, Lior Fishman, and Uri Shapira. Diophantine approximations on fractals. *Geom. Funct. Anal.*, 21(1):14–35, 2011.

[16] Manfred Einsiedler and Thomas Ward. *Ergodic theory with a view towards number theory*, volume 259 of *Graduate Texts in Mathematics*. Springer-Verlag London Ltd., London, 2011.

[17] K. J. Falconer. Hausdorff dimension and the exceptional set of projections. *Mathematika*, 29(1):109–115, 1982.

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

[19] Kenneth Falconer. *Fractal geometry*. John Wiley & Sons Inc., Hoboken, NJ, second edition, 2003. Mathematical foundations and applications.

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

[21] William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

[22] 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.

[23] Hillel Furstenberg. Ergodic fractal measures and dimension conservation. *Ergodic Theory Dynam. Systems*, 28(2):405–422, 2008.

[24] Adriano M Garsia. Arithmetic properties of Bernoulli convolutions. *Transactions of the American Mathematical Society*, 102:409–432, 1962.

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

[26] Michael Hochman. Geometric rigidity of times-m invariant measures. *Journal of the European Mathematical Society*, 14(5):1539–1563, 2012.

[27] Michael Hochman. Dynamics on fractals and fractal distributions. Preprint, available at http://arxiv.org/abs/1008.3731v2, 2013.

[28] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy. *Ann. of Math.*, 180(2):773–822, 2014.

[29] Michael Hochman and Pablo Shmerkin. Local entropy averages and projections of fractal measures. *Ann. of Math. (2)*, 175(3):1001–1059, 2012.

[30] Franz Hofbauer. β-shifts have unique maximal measure. *Monatsh. Math.*, 85(3):189–198, 1978.

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

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

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

[34] R. Kaufman. Continued fractions and Fourier transforms. *Mathematika*, 27(2):262–267 (1981), 1980.

[35] Robert Kaufman. On Bernoulli convolutions. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 217–222. Amer. Math. Soc., Providence, RI, 1984.

[36] Steven P Lalley. Random series in powers of algebraic integers: Hausdorff dimension of the limit distribution. *Journal of the London Mathematical Society. Second Series*, 57(3):629–654, 1998.

[37] Elon Lindenstrauss. p-adic foliation and equidistribution. *Israel J. Math.*, 122:29–42, 2001.

[38] Elon Lindenstrauss. Rigidity of multiparameter actions. *Israel J. Math.*, 149:199–226, 2005. Probability in mathematics.

[39] Russell Lyons. Strong laws of large numbers for weakly correlated random variables. *Michigan Math. J.*, 35(3):353–359, 1988.

[40] Pertti Mattila. *Geometry of sets and measures in Euclidean spaces*, volume 44 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1995. Fractals and rectifiability.

[41] David Meiri. Entropy and uniform distribution of orbits in T^d. *Israel Journal of Mathematics*, 105:153–183, 1998.

[42] W. Moran and A. D. Pollington. Metrical results on normality to distinct bases. *J. Number Theory*, 54(2):180–189, 1995.

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

[44] W. Parry. On the β-expansions of real numbers. *Acta Math. Acad. Sci. Hungar.*, 11:401–416, 1960.

[45] Yuval Peres and Pablo Shmerkin. Resonance between Cantor sets. *Ergodic Theory Dynam. Systems*, 29(1):201–221, 2009.

[46] A. D. Pollington. The Hausdorff dimension of a set of normal numbers. *Pacific J. Math.*, 95(1):193–204, 1981.

[47] A. D. Pollington. The Hausdorff dimension of a set of normal numbers. II. *J. Austral. Math. Soc. Ser. A*, 44(2):259–264, 1988.

[48] Martine Queffélec and Olivier Ramaré. Analyse de Fourier des fractions continues à quotients restreints. *Enseign. Math. (2)*, 49(3-4):335–356, 2003.

[49] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar*, 8:477–493, 1957.

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

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

[52] Nikita Sidorov. Arithmetic dynamics. In *Topics in dynamics and ergodic theory*, volume 310 of *London Math. Soc. Lecture Note Ser.*, pages 145–189. Cambridge Univ. Press, Cambridge, 2003.

[53] Dennis Sullivan. Differentiable structures on fractal-like sets, determined by intrinsic scaling functions on dual Cantor sets. In *The mathematical heritage of Hermann Weyl (Durham, NC, 1987)*, volume 48 of *Proc. Sympos. Pure Math.*, pages 15–23. Amer. Math. Soc., Providence, RI, 1988.

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

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

Email: mhochman@math.huji.ac.il

Address: Einstein Institute of Mathematics, Givat Ram, Jerusalem 91904, Israel

Email: pshmerkin@utdt.edu

Address: Department of Mathematics, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, GU2 7XH, United Kingdom

Current address: Department of Mathematics and Statistics, Torcuato Di Tella University, Av. Figueroa Alcorta 7350 (1425), Buenos Aires, Argentina.