Proceedings of Symposia in Pure Mathematics

Notes on Bernoulli convolutions

Boris Solomyak

This paper is dedicated to Benoit Mandelbrot

ABSTRACT. Bernoulli convolutions are examples of self-similar measures. They have been studied (under different names) since the beginning of the 20th century. We focus on the question of absolute continuity for Bernoulli convolutions and discuss some of their applications.

Table of Contents

1. Introduction

A Borel probability measure ν on $\mathbb{R}^d$ is called **self-similar** if it satisfies the equation

$$\nu = \sum_{i=1}^{m} p_i(\nu \circ S_i^{-1}),$$

for some $m \geq 2$, probability vector $\mathbf{p} = (p_1, \ldots, p_m)$, and contractive similitudes $S_1, \ldots, S_m$. This terminology is due to Hutchinson [**15**] who proved that there exists a unique such measure. Self-similar measures arise in fractal geometry, dynamical systems, probability, and harmonic analysis.

The simplest non-trivial case is when $d = 1$, $m = 2$, and S_1, S_2 have the same linear part. This way we get a remarkable family of measures, called Bernoulli convolutions. They include the classical Cantor-Lebesgue measure and its variants,

2000 *Mathematics Subject Classification.* Primary 26A46, Secondary 26A30, 28A78, 28A80, 39B12.

Key words and phrases. Bernoulli convolutions, self-similar measures.

Supported in part by NSF Grant #DMS 0099814.

which are well understood. However, in the "overlapping" case, the situation is much more delicate; there has been a lot of progress over the years, but many natural questions remain open. The main question that we address is: when is the measure ν absolutely continuous? We present a detailed proof of absolute continuity for almost every parameter, following [**30**]. As an application, we compute the Hausdorff dimension of some self-affine sets following [**34**]. Some other applications and generalizations are discussed, for which we either sketch a proof or provide a reference.

These notes are based on a series of lectures given at the School on Real Analysis and Measure Theory in Grado in September 2001. I am very grateful to Aljosa Volcic for his hospitality.

2. Background and Historical Notes

Let $\lambda \in (0,1)$. Consider the random series

$$Y_\lambda = \sum_{n=0}^{\infty} \pm \lambda^n$$

where the "+" and "−" signs are chosen independently with probability 1/2. Let ν_λ be its distribution:

$$\boxed{\nu_\lambda(E) = \mathrm{Prob}(Y_\lambda \in E)}$$

Thus,

$$\boxed{\nu_\lambda = \textbf{infinite convolution product of } \tfrac{1}{2}(\delta_{-\lambda^n} + \delta_{\lambda^n})}$$

This is the reason why ν_λ are called "infinite Bernoulli convolutions" (or just "Bernoulli convolutions"). Taking the Fourier transform, we get

$$\widehat{\nu}_\lambda(\xi) = \int_{\mathbb{R}} e^{it\xi}\, d\nu_\lambda(t) = \prod_{n=0}^{\infty} \tfrac{1}{2}\widehat{(\delta_{\lambda^n} + \delta_{-\lambda^n})}(\xi) = \prod_{n=0}^{\infty} \cos(\lambda^n \xi). \tag{2.1}$$

Bernoulli convolutions have been studied since the 1930's, in connection with harmonic analysis, fractal geometry, dynamical systems, and probability. Their very important property is **self-similarity**: for any Borel set E,

$$\nu_\lambda(E) = \tfrac{1}{2}[\nu_\lambda(S_1^{-1}E) + \nu_\lambda(S_2^{-1}E)], \tag{2.2}$$

where $S_1(x) = 1 + \lambda x$ and $S_2(x) = -1 + \lambda x$.

THEOREM 2.1 (Jessen and Wintner 1935). *The measure ν_λ is either absolutely continuous or purely singular, depending on λ.*

Jessen and Wintner proved this statement (called "the law of pure types") for more general distributions. Now there are several easy proofs, e.g. one can deduce it by decomposing the self-similar measure into absolutely continuous and singular parts and using uniqueness of the probability measure satisfying (2.2).

A question arises:

for which λ is ν_λ singular? absolutely continuous?

(Unless stated otherwise, "absolutely continuous" and "singular" is understood with respect to the Lebesgue measure $\mathcal{L}$.)

LEMMA 2.2 (Kershner and Wintner 1935). *The support of ν_λ is a Cantor set of zero length for $\lambda < 1/2$, and the interval $[-(1-\lambda)^{-1}, (1-\lambda)^{-1}]$ for $\lambda \in (1/2, 1)$.*

Thus, ν_λ is certainly singular for $\lambda < 1/2$ (in fact, ν_λ is the standard Cantor-Lebesgue measure on its support). It is easy to see that $\nu_{1/2}$ is the uniform measure on $[-2, 2]$. It was thought for a while that ν_λ is absolutely continuous for all $\lambda \in (1/2, 1)$ until Erdős proved the following surprising result.

DEFINITION 2.3. An algebraic integer $\theta > 1$ is called a **Pisot number** (or PV-number, for Pisot and Vijayaraghavan), if all Galois conjugates (i.e. the other roots of the minimal polynomial) of θ are less than 1 in modulus.

THEOREM 2.4 (Erdős 1939). *If $\lambda \neq 1/2$ and $1/\lambda$ is a Pisot number, then ν_λ is singular. Moreover, the Fourier transform $\widehat{\nu}_\lambda(\xi)$ does not tend to 0 as $\xi \to \infty$.*

The fundamental property of Pisot numbers is that $\mathrm{dist}(\theta^n, \mathbb{Z})$ tends to 0 at a geometric rate. In fact, let $\theta_2, \dots, \theta_k$ be the Galois conjugates of θ. Then $\theta^n + \sum_{j=2}^k \theta_j^n$ is an integer for all $n \in \mathbb{N}$, being a symmetric function of the roots of the minimal polynomial for θ. By assumption, $\max_{j \geq 2} |\theta_j| =: \rho \in (0, 1)$, hence

$$\mathrm{dist}(\theta^n, \mathbb{Z}) \leq (k-1)\rho^n, \quad n \in \mathbb{N}. \tag{2.3}$$

Proof of Theorem 2.4. Denote $\theta = 1/\lambda$. By (2.1), for $N \geq 1$,

$$\begin{aligned} \widehat{\nu}_\lambda(\pi\theta^N) &= \cos(\pi\theta^N)\cos(\pi\theta^{N-1}) \cdot \ldots \cdot \cos(\pi)\cos(\lambda\pi)\cos(\lambda^2\pi)\ldots \\ &= \prod_{n=1}^{N} \cos(\pi\theta^n) \cdot \widehat{\nu}_\lambda(\pi). \end{aligned} \tag{2.4}$$

Since θ^n is an algebraic integer and $\theta \neq 2$, we have $\theta^n \notin (1/2)\mathbb{Z}$, for all $n \in \mathbb{Z}$, which implies that $\widehat{\nu}_\lambda(\pi\theta^N) \neq 0$. Now (2.4) and (2.3) imply

$$|\widehat{\nu}_\lambda(\pi\theta^N)| \geq \prod_{n=1}^{\infty} |\cos((k-1)\rho^n)| \cdot |\widehat{\nu}_\lambda(\pi)| =: \delta > 0,$$

for all $N \geq 1$. Thus, $\widehat{\nu}_\lambda(\xi) \not\to 0$ as $\xi \to \infty$. By the Riemann-Lebesgue lemma, ν_λ is not absolutely continuous, so ν_λ is singular in view of Theorem 2.1. $\square$

OPEN PROBLEM. Is ν_λ absolutely continuous for all $\lambda \in (1/2, 1)$ other than reciprocals of Pisot numbers?

Digression: some basic facts about Pisot numbers (see [**35, 5**]). Denote by $\mathcal{P}$ the set of Pisot numbers. Here is the fundamental result obtained by Pisot in 1938 (of course, he didn't use the term "Pisot numbers"):

$$\theta > 1,\ \alpha > 0,\ \sum_{n=1}^{\infty} \mathrm{dist}(\theta^n\alpha, \mathbb{Z})^2 < \infty \ \Rightarrow\ \theta \in \mathcal{P}. \tag{2.5}$$

It is a famous open problem whether the set

$$\Phi := \{\theta > 1 :\ \exists\, \alpha > 0,\ \lim_{n\to\infty} \mathrm{dist}(\theta^n\alpha, \mathbb{Z}) = 0\}$$

coincides with $\mathcal{P}$. It is known that Φ is countable and all algebraic numbers in Φ are Pisot.

More facts about $\mathcal{P}$: it is a closed subset of $\mathbb{R}$; the smallest Pisot number is the positive root $\theta_1 \approx 1.324718$ of $x^3 - x - 1 = 0$, and the second smallest element of $\mathcal{P}$ is the positive root $\theta_2 \approx 1.3802777$ of $x^4 - x^3 - 1$. These are the only Pisot numbers in $(1, \sqrt{2})$, so we know only two examples of singular ν_λ for $\lambda \in (2^{-1/2}, 1)$. The golden ratio $(1+\sqrt{5})/2$ is the only quadratic Pisot number in $(1,2)$, and it is also the smallest limit point of $\mathcal{P}$.

The early study of Bernoulli convolutions ν_λ was related to some questions of harmonic analysis. Using (2.5), Salem proved the following

THEOREM 2.5 (Salem 1944). *If $\lambda \in (0,1)$ and $1/\lambda$ is not a Pisot number, then* $\lim_{\xi\to\infty} \widehat{\nu}_\lambda(\xi) = 0$.

A set $E \subset (0, 2\pi)$ is a **set of uniqueness** for trigonometric series, if

$$\sum_{n=0}^{\infty}(a_n \cos(nx) + b_n \sin(nx)) = 0,\ \forall\, x \in E \ \Rightarrow\ a_n = b_n = 0,\ \forall\, n.$$

Salem (see [**35**]) used Theorem 2.5 to show that a Cantor set with constant ratio of dissection $\lambda \in (0, 1/2)$ (e.g., the support of ν_λ) is a set of uniqueness if and only if $1/\lambda$ is a Pisot number.

Results on absolute continuity for ν_λ. In this direction the first important result is also due to Erdős. Denote by $S_\perp$ the set of $\lambda \in (1/2, 1)$ such that ν_λ is singular.

THEOREM 2.6 (Erdős 1940). *There exists $a_0 < 1$ such that $S_\perp \cap (a_0, 1)$ has zero Lebesgue measure. Moreover, there exist $a_k \uparrow 1$ such that ν_λ has a density in $\mathcal{C}^k(\mathbb{R})$ for Lebesgue-a.e. $\lambda \in (a_k, 1)$.*

Erdős used a beautiful combinatorial argument, inspired by the work of Pisot, to show that there exists $\gamma > 0$ such that

$$|\widehat{\nu}_\lambda(\xi)| = O(|\xi|^{-\gamma}) \quad \text{for a.e. } \lambda \in (2^{-1}, 2^{-1/2}). \tag{2.6}$$

After that, the proof is concluded as follows. Notice that (2.1) implies

$$\widehat{\nu}_\lambda(\xi) = \widehat{\nu}_{\lambda^m}(\xi)\widehat{\nu}_{\lambda^m}(\lambda\xi)\cdot \ldots \cdot \widehat{\nu}_{\lambda^m}(\lambda^{m-1}\xi) \quad \text{for } m \ge 2. \tag{2.7}$$

Combining this formula with (2.6) yields

$$|\widehat{\nu}_\lambda(\xi)| = O(|\xi|^{-m\gamma}) \quad \text{for a.e. } \lambda \in (2^{-1/m}, 2^{-1/(2m)}).$$

If $|\widehat{\nu}_\lambda(\xi)| = O(|\xi|^{-\alpha})$ for $\alpha > 1$, then $\widehat{\nu}_\lambda \in L^1(\mathbb{R})$, whence ν_λ is absolutely continuous with a density $\frac{d\nu_\lambda}{dx} \in \mathcal{C}(\mathbb{R})$. If $|\widehat{\nu}_\lambda(\xi)| = O(|\xi|^{-\alpha})$ for $\alpha > k+1$, then $\widehat{\frac{d^k}{dx^k}(\frac{d\nu_\lambda}{dx})} = \xi^k \widehat{\nu}_\lambda(\xi)$ (in the distributional sense) is in $L^1(\mathbb{R})$, and therefore, $\frac{d\nu_\lambda}{dx} \in \mathcal{C}^k(\mathbb{R})$. This implies the assertion of Theorem 2.6 with $a_k = 2^{-\gamma/(k+1)}$. $\square$

Two comments are in order. First, Erdős did not give an explicit value of γ and a_0, but it was clear that γ in (2.6) is rather small, and consequently, a_0 is rather close to 1. Second, Kahane [**17**] observed that the argument of [**11**] actually implies that the Hausdorff dimension of $S_\perp \cap (b, 1)$ tends to 0 as $b \uparrow 1$. In [**28**] there is an exposition of the Erdős-Kahane argument with explicit constants $\gamma = 6 \cdot 10^{-4}$ and $a_0 = .99933$.

In 1962 Garsia [**14**] found the largest explicitly given set of λ known to date, for which ν_λ is absolutely continuous (actually, it has bounded density). This set consists of reciprocals of algebraic integers in $(1,2)$ whose minimal polynomial has other roots outside the unit circle and the constant coefficient ± 2. Such are for instance the polynomials $x^{n+p} - x^n - 2$ where $p, n \geq 1$ and $\max\{p, n\} \geq 2$. Another example is the polynomial $x^3 - 2x - 2$; the reciprocal of its largest root is $\approx .565198$.

There was a renewed interest in Bernoulli convolutions in the 1980's, after their importance in various problems of dynamics and dimension was discovered by Alexander and Yorke [**1**], Przytycki and Urbański [**34**], and Ledrappier [**19**].

The latest stage in the study of Bernoulli convolutions started with a seemingly unrelated development: the formulation of the "$\{0,1,3\}$-problem" by Keane and Smorodinsky in the early 1990's (see [**18**]). Motivated by questions of Palis and Takens on sums of Cantor sets, they asked how the dimension and morphology of the set $\{\sum_{n=0}^{\infty} a_n \lambda^n : \ a_n = 0, 1, \text{ or } 3\}$ depends on the parameter λ. Pollicott and Simon [**32**] proved that the Hausdorff dimension equals $\frac{\log 3}{\log(1/\lambda)}$ for a.e. $\lambda \in (\frac{1}{4}, \frac{1}{3})$ using self-similar measures obtained by taking the digits $a_n \in \{0,1,3\}$ independently with equal probabilities. Pollicott and Simon were influenced by the important paper of Falconer [**12**], where methods originating from geometric measure theory were used to obtain "almost sure" results on the dimension of self-affine sets. A crucial new idea in [**32**] was the notion of **transversality** for power series. Combining this idea with some techniques from [**25**] and [**36**] led to

THEOREM 2.7 ([**37**]). **(i)** *$\nu_\lambda \ll \mathcal{L}$ with a density in $L^2(\mathbb{R})$ for a.e. $\lambda \in (\frac{1}{2}, 2^{-1/2})$.*

(ii) *$\nu_\lambda \ll \mathcal{L}$ with a density in $\mathcal{C}(\mathbb{R})$ for a.e. $\lambda \in (2^{-1/2}, 1)$.*

A simpler proof was found jointly with Y. Peres in [**30**], see Sections 4 and 5. Recently, Peres and Schlag [**27**] established, as a corollary of a more general result, that the Hausdorff dimension of $S_\perp \cap (b, 1)$ is less than one for any $b > 1/2$. A self-contained proof of this fact is given in [**28**]. All of these papers use some versions of transversality.

REMARK 2.8. There has been a lot of recent work on Bernoulli convolutions which we do not mention here, in particular, the study of various dimensions of ν_λ when $1/\lambda$ is Pisot. A bibliography up to 1999 can be found in [**28**]. Diaconis and Freedman [**7**] discuss Bernoulli convolutions in connection with fractal image generation. Strichartz investigated self-similar measures and their Fourier transforms, see [**40**] and references there.

3. Interlude: some Pictures

Here we present a "gallery" of histograms for ν_λ approximations. The pictures were created as follows. Fix $\lambda \in (1/2, 1)$. Starting with the set $\{0\}$ we performed the iterations of the map $A \mapsto (\lambda A - 1) \cup (\lambda A + 1)$, counting points with multiplicities. After n iterations we got a data set with 2^n entries, for which we made a histogram with *Mathematica*. These histograms are, in fact, rescaled probability mass functions for discrete measures which converge weak* to ν_λ, as $n \to \infty$. We used $n = 20$ (so there are about 10^6 data points) and 2^{10} "bins" for the histogram. Of course, these approximations are very crude, but they give a rough idea of what the measures look like.

In Figures 1 and 2 we show the histograms of ν_λ approximations for λ indicated in the tables.

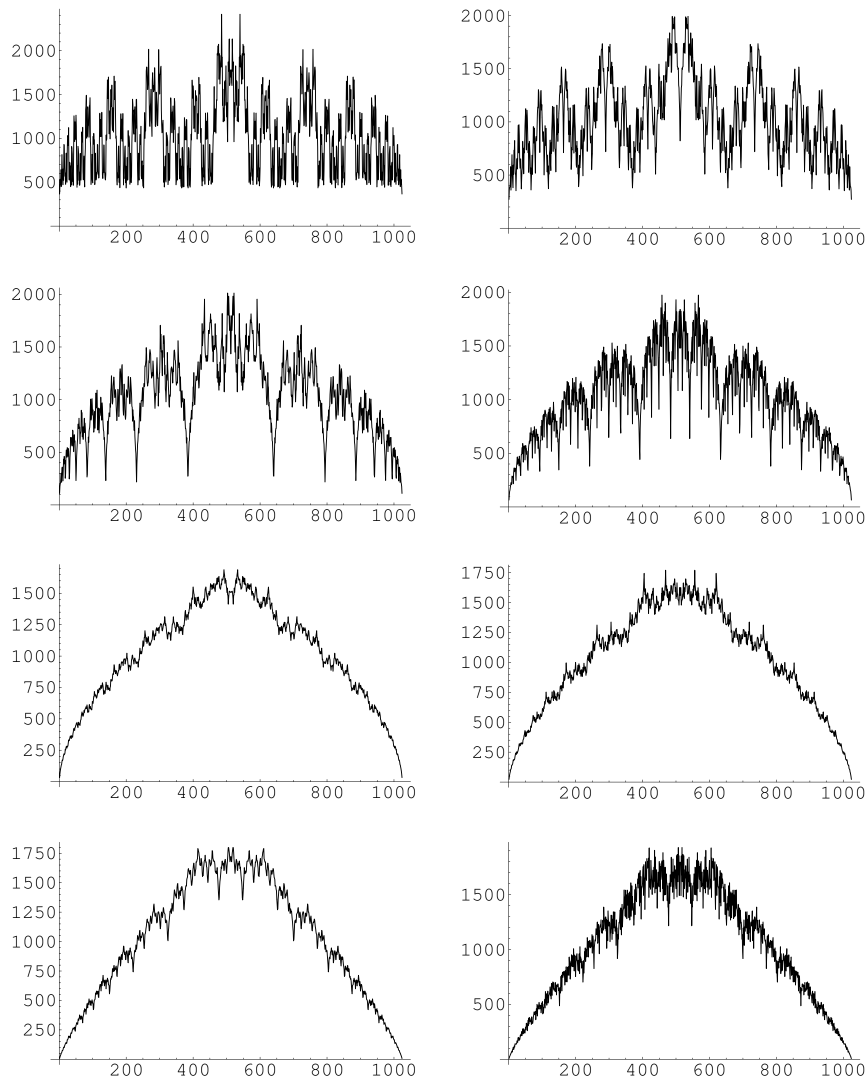

FIGURE 1. The eight graphs correspond to the following λ (left to right, up to down): $.55$, $.565198 \approx \lambda_7$, $.6$, $.618034 \approx \lambda_6$, $.64$, $.651822 \approx \lambda_5$, $.68$, $.682328 \approx \lambda_4$.

The numbers λ_k, $1 \le k \le 7$, were chosen as follows: λ_7^{-1} is a Garsia number, the largest root of $x^3 - 2x - 2 = 0$; $\lambda_6^{-1} = \frac{\sqrt{5}+1}{2}$ is the golden ratio; λ_5^{-1} is the 6-th smallest Pisot number, the largest root of $x^5 - x^3 - x^2 - x - 1 = 0$; λ_4^{-1} is the 4-th smallest Pisot number, the largest root of $x^3 - x^2 - 1 = 0$; λ_3^{-1} is the 3-d smallest

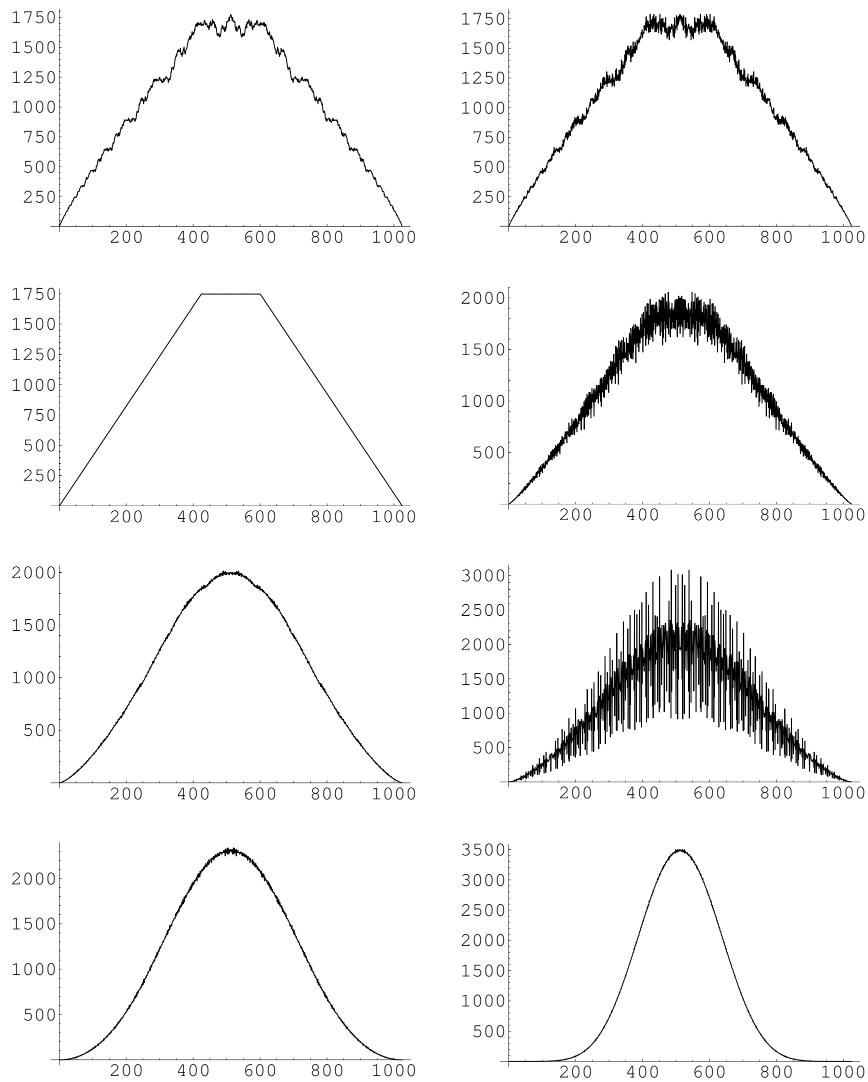

FIGURE 2. The eight graphs correspond to the following λ (left to right, up to down): $.69$, $.692871 \approx \lambda_3$, $2^{-1/2}$, $.724492 \approx \lambda_2$, $.75$, $.754877 \approx \lambda_1$, $.8$, $.9$.

Pisot number, the largest root of $x^5 - x^4 - x^3 + x^2 - 1 = 0$; λ_2^{-1} is the 2-nd smallest Pisot number, the largest root of $x^4 - x^3 - 1 = 0$; λ_1^{-1} is the smallest Pisot number, the largest root of $x^3 - x - 1 = 0$. (See [**5**] for the complete list of Pisot numbers smaller than the golden ratio.)

Here are some observations:

1. One cannot really infer that a certain measure is singular from these pictures. The "ruggedness" of histograms for low λ only suggests that the density, if it exists, is not smooth. Recall that the measures ν_λ have a density in L^2 for a.e. $\lambda \in (1/2, 2^{-1/2})$ (and even some "fractional" smoothness [**27**]), but we do not know if they (ever? typically?) have a continuous density. Note the case of $\lambda = \lambda_7$ (reciprocal of a Garsia number), for which ν_λ is known to have a bounded density.

2. The approximate densities generally get "smoother" monotonically with λ, however, there are exceptions, detected around the reciprocals of small Pisot numbers (they are more pronounced for small Pisot numbers of small degree).

3. For λ closer to 1, the histogram starts to resemble the normal distribution. This is not surprising since in the random series $\sum_{n=0}^\infty \pm\lambda^n$ we are adding independent, "almost" identically distributed random variables. For $\lambda = 2^{-1/2}$ the picture corresponds to the exact density of ν_λ, which is easily computed. Recall that ν_λ has a continuous density for a.e. $\lambda \in (2^{-1/2}, 1)$ by Theorem 2.7(ii).

4. Absolute Continuity for a Typical λ

Here and in the following section we prove Theorem 2.7, following [**30**]. Actually, we will establish a more general result. Let $\mathbf{p} = (p_1, \ldots, p_m)$ be a probability vector and let $D = \{d_1, \ldots, d_m\} \subset \mathbb{R}$ be a finite set of "digits." Let $\nu_\lambda^{D,\mathbf{p}}$ be the distribution of the random series $\sum_{n=0}^\infty a_n\lambda^n$, where the coefficients a_n are chosen from D randomly, with probabilities p_i, and independently. Here $\lambda \in (0,1)$ is a parameter. Alternatively, we can define $\nu_\lambda^{D,\mathbf{p}}$ as the unique probability measure satisfying the equation

$$\nu_\lambda^{D,\mathbf{p}} = \sum_{i=1}^m p_i(\nu_\lambda^{D,\mathbf{p}} \circ S_i^{-1}), \quad \text{where} \quad S_i(x) = \lambda x + d_i,$$

see Hutchinson [**15**]. Such measures are called **self-similar**; our measures are special, since they have **uniform contraction ratios**.

We will need a more concrete realization of the measures $\nu_\lambda^{D,\mathbf{p}}$. Let $\Omega = D^{\mathbb{N}}$ be the sequence space with the product topology. The elements of Ω are infinite sequences $\omega = \omega_0\omega_1\ldots$ There is a natural metric compatible with the topology:

$$\varrho(\omega,\tau) = m^{-|\omega\wedge\tau|} \ \text{ for } \ \omega,\tau \in \Omega,$$

where

$$|\omega \wedge \tau| = \min\{n : \ \omega_n \neq \tau_n\}.$$

Let μ be the Bernoulli measure $(p_1, \ldots, p_m)^{\mathbb{N}}$ on Ω. Consider the map $\Pi_\lambda : \Omega \to \mathbb{R}$, given by

$$\Pi_\lambda(\omega) = \sum_{n=0}^\infty \omega_n\lambda^n. \tag{4.1}$$

This map is continuous, for any $\lambda \in (0,1)$; in fact, it is Hölder-continuous:

$$\begin{aligned} |\Pi_\lambda(\omega) - \Pi_\lambda(\tau)| &\leq \max_{i,j}|d_i - d_j| \cdot (1-\lambda)^{-1} \cdot \lambda^{|\omega\wedge\tau|} \\ &= \text{const} \cdot \varrho(\omega,\tau)^{\frac{-\log\lambda}{\log m}}. \end{aligned} \tag{4.2}$$

LEMMA 4.1.

$$\nu_\lambda^{D,\mathbf{p}} = \mu \circ \Pi_\lambda^{-1}. \tag{4.3}$$

Proof is immediate from definitions. □

Next we define δ-**transversality** for power series. For $\gamma > 0$ let

$$\mathcal{B}_\gamma = \Big\{ g(x) = 1 + \sum_{n=1}^{\infty} a_n x^n : \ |a_n| \le \gamma, \ n \ge 1 \Big\}. \tag{4.4}$$

DEFINITION 4.2. Let J be a closed subinterval of $[0,1)$ and let $\gamma, \delta > 0$. We say that $\mathcal{B}_\gamma$ satisfies the δ-**transversality condition** on J if

$$\forall\, g \in \mathcal{B}_\gamma, \ (\lambda \in J, \ g(\lambda) < \delta) \ \Rightarrow \ g'(\lambda) < -\delta.$$

δ-transversality means that the graph of each $g \in \mathcal{B}_\gamma$ crosses transversally, with slope at most $-\delta$, all horizontal lines below height δ that it meets.

The two main ingredients of the proof are: (a) proving absolute continuity for a.e. λ in a certain interval where δ-transversality holds, and (b) checking δ-transversality. Part (a) is contained in the next theorem; part (b) is dealt with in the next section.

For a digit set $D = \{d_1, \ldots, d_m\}$ let

$$\gamma(D) = \max \Big\{ \frac{|d_i - d_j|}{|d_k - d_\ell|} : \ 1 \le i,j,k,\ell \le m, \ k \ne \ell \Big\}. \tag{4.5}$$

THEOREM 4.3. *Let $\nu_\lambda^{D,\mathbf{p}}$ be the family of self-similar measures corresponding to a digit set D and a probability vector $\mathbf{p} = (p_1, \ldots, p_m)$. Let J be a closed subinterval of $[0,1)$, such that $\mathcal{B}_{\gamma(D)}$ satisfies the δ-transversality condition on J for some $\delta > 0$. Then $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous with a density in $L^2(\mathbb{R})$ for a.e. $\lambda \in J \cap (\sum_{i=1}^m p_i^2, 1)$.*

Proof. Heuristically, we want to think of $\nu_\lambda^{D,\mathbf{p}}$ as "nonlinear projections" of the measure μ on the sequence space. We apply the differentiation method used by Mattila [**20**] to prove projection theorems in Euclidean space. Let $J = [\lambda_0, \lambda_1]$. Without loss of generality, we can assume that $\lambda_0 > \sum_{i=1}^m p_i^2$. Let $B_r(x) = [x - r, x + r]$. Consider the lower derivative of $\nu_\lambda^{D,\mathbf{p}}$:

$$\underline{D}(\nu_\lambda^{D,\mathbf{p}}, x) = \liminf_{r \downarrow 0} \, (2r)^{-1} \nu_\lambda^{D,\mathbf{p}}(B_r(x)).$$

A standard application of the Vitali covering theorem shows that $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous if and only if $\underline{D}(\nu_\lambda^{D,\mathbf{p}}, x) < \infty$ for $\nu_\lambda^{D,\mathbf{p}}$ almost all $x \in \mathbb{R}$, see e.g. [**21**, 2.12]. If we show that

$$\mathcal{J} := \int_J \int_{\mathbb{R}} \underline{D}(\nu_\lambda^{D,\mathbf{p}}, x) \, d\nu_\lambda^{D,\mathbf{p}}(x) \, d\lambda < \infty, \tag{4.6}$$

then this criterion yields that $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous for a.e. $\lambda \in J$. Notice that for λ such that $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous, $\underline{D}(\nu_\lambda^{D,\mathbf{p}}, x)$ is the Radon-Nikodym derivative $\frac{d\nu_\lambda^{D,\mathbf{p}}}{dx}$, so

$$\int_{\mathbb{R}} \underline{D}(\nu_\lambda^{D,\mathbf{p}}, x) \, d\nu_\lambda^{D,\mathbf{p}}(x) = \int_{\mathbb{R}} \Big(\frac{d\nu_\lambda^{D,\mathbf{p}}}{dx} \Big)^2 dx,$$

and (4.6) will imply that $\frac{d\nu_\lambda^{D,\mathbf{p}}}{dx} \in L^2(\mathbb{R})$ for a.e. $\lambda \in J$.

By Fatou's lemma,

$$\mathcal{J} \le \liminf_{r\downarrow 0} (2r)^{-1} \int_J \int_{\mathbb{R}} \nu_\lambda^{D,\mathbf{p}}(B_r(x))\, d\nu_\lambda^{D,\mathbf{p}}(x)\, d\lambda.$$

Using (4.3) we can change variables to get

$$\mathcal{J} \le \liminf_{r\downarrow 0} (2r)^{-1} \int_J \int_{\Omega} \nu_\lambda^{D,\mathbf{p}}(B_r(\Pi_\lambda(\omega)))\, d\mu(\omega)\, d\lambda. \tag{4.7}$$

Next, denote by $\mathbf{1}_A$ the characteristic function of a set A and use (4.3) again:

$$\begin{aligned} \nu_\lambda^{D,\mathbf{p}}(B_r(\Pi_\lambda(\omega))) &= \int_{\mathbb{R}} \mathbf{1}_{B_r(\Pi_\lambda(\omega))}(x)\, d\nu_\lambda^{D,\mathbf{p}}(x) \\ &= \int_\Omega \mathbf{1}_{\{\tau\in\Omega:\ |\Pi_\lambda(\tau)-\Pi_\lambda(\omega)|\le r\}}\, d\mu(\tau). \end{aligned}$$

Substitute this into (4.7), exchange the order of integration, and integrate with respect to λ to obtain

$$\mathcal{J} \le \liminf_{r\downarrow 0} (2r)^{-1} \int_\Omega \int_\Omega \mathcal{L}\Big\{\lambda \in J:\ |\Pi_\lambda(\tau) - \Pi_\lambda(\omega)| \le r\Big\}\, d\mu(\tau) d\mu(\omega), \tag{4.8}$$

where $\mathcal{L}$ denotes Lebesgue measure. Let

$$\phi_{\tau,\omega}(\lambda) := \Pi_\lambda(\tau) - \Pi_\lambda(\omega) = \sum_{n=0}^\infty (\tau_n - \omega_n)\lambda^n.$$

We need to estimate $\mathcal{L}\{\lambda \in J:\ |\phi_{\tau,\omega}(\lambda)| \le r\}$. Let

$$k = |\omega \wedge \tau| = \min\{n:\ \omega_n \ne \tau_n\}.$$

Then

$$\phi_{\tau,\omega}(\lambda) = (\tau_k - \omega_k)\lambda^k g(\lambda), \tag{4.9}$$

where

$$g(\lambda) = 1 + \sum_{n=1}^\infty \frac{\tau_{k+n} - \omega_{k+n}}{\tau_k - \omega_k}\lambda^n \in \mathcal{B}_{\gamma(D)},$$

see (4.4) and (4.5). Recall that J is an interval of δ-transversality for $\mathcal{B}_{\gamma(D)}$. We claim that for all $g \in \mathcal{B}_{\gamma(D)}$ and all $\rho > 0$,

$$\mathcal{L}\{\lambda \in J:\ |g(\lambda)| \le \rho\} \le 2\delta^{-1}\rho. \tag{4.10}$$

This is obvious if $\rho \ge \delta$; otherwise, $g'(\lambda) < -\delta$ whenever $|g(\lambda)| \le \rho$ by δ-transversality. Thus, g is monotone decreasing on the set in (4.10) with $|g'| > \delta$, which implies the inequality (4.10).

Now we obtain from (4.9) and (4.10), keeping in mind that $J = [\lambda_0, \lambda_1]$:

$$\begin{aligned} \mathcal{L}\{\lambda \in J:\ |\phi_{\tau,\omega}(\lambda)| \le r\} &= \mathcal{L}\{\lambda \in J:\ |g(\lambda)| \le (\tau_k - \omega_k)^{-1}\lambda^{-k} r\} \\ &\le \mathcal{L}\{\lambda \in J:\ |g(\lambda)| \le (\tau_k - \omega_k)^{-1}\lambda_0^{-k} r\} \\ &\le C\delta^{-1}\lambda_0^{-k} r, \end{aligned}$$

where $C = 2(\min_{i\neq j}|d_i - d_j|)^{-1}$. We can substitute this into (4.8):

$$\begin{aligned} \mathcal{J} &\le \liminf_{r\downarrow 0}\,(2r)^{-1}\int_\Omega\int_\Omega C\delta^{-1}\lambda_0^{-|\omega\wedge\tau|}r\,d\mu(\tau)\,d\mu(\omega) \\ &= (C/2)\delta^{-1}\sum_{k=0}^{\infty}\lambda_0^{-k}(\mu\times\mu)\{(\omega,\tau):\ |\omega\wedge\tau| = k\} \qquad (4.11) \\ &\le (C/2)\delta^{-1}\sum_{k=0}^{\infty}\lambda_0^{-k}(p_1^2+\cdots+p_m^2)^k < \infty, \end{aligned}$$

where the series converges since $\lambda_0 > \sum_{i=1}^m p_i^2$. This proves (4.6), and we are done. □

Before we turn to checking the transversality condition, let us pause for a moment and see which properties of μ we have used. The fact that μ is a product measure was only used in the last step. The argument goes through for an arbitrary measure and its "projections," as long as the series in (4.11) converges. Let us make this precise.

Definition 4.4. Let μ be a Borel probability measure on the sequence space Ω and let

$$\mu_k := (\mu\times\mu)\{(\omega,\tau):\ |\omega\wedge\tau| = k\}.$$

The **lower correlation dimension** of μ is defined by

$$D_2(\mu) = \liminf_{k\to\infty}\frac{-\log\mu_k}{k\log m}.$$

Theorem 4.5. *Let $\Omega = D^{\mathbb{N}}$ where $D = \{d_1,\dots,d_m\}$, and let Π_λ be the corresponding "projection" map (4.1). Let μ be a probability measure on Ω and let J be a closed subinterval of $[0,1)$, such that $\mathcal{B}_{\gamma(D)}$ satisfies the δ-transversality condition on J for some $\delta > 0$. Then the measure $\mu\circ\Pi_\lambda^{-1}$ is absolutely continuous with a density in $L^2(\mathbb{R})$ for a.e. $\lambda\in J\cap(m^{-D_2(\mu)},1)$.*

Proof. We proceed exactly as in the proof of Theorem 4.3, assuming that $J = [\lambda_0,\lambda_1]$, with $\lambda_0 > m^{-D_2(\mu)}$, until the second line of (4.11), which yields

$$\mathcal{J} \le (C/2)\delta^{-1}\sum_{k=0}^{\infty}\lambda_0^{-k}\mu_k.$$

We have $\frac{\log(1/\lambda_0)}{\log m} < D_2(\mu)$. Therefore, there exists $\varepsilon > 0$ such that

$$\frac{\log(1/\lambda_0)}{\log m} < \frac{-\log\mu_k}{k\log m} - \varepsilon,$$

for all k sufficiently large. Then $\lambda_0^{-k} < \mu_k^{-1}\cdot m^{-\varepsilon k}$ for k sufficiently large, so $\mathcal{J} < \infty$, concluding the proof. □

5. Establishing Transversality

Here we finish the proof of Theorem 2.7, following [**30, 31**].

Definition 5.1. Let $\gamma > 0$. A power series $h(x)$ is called a $(*)$-function for $\mathcal{B}_\gamma$ if for some $k \ge 1$ and $a_k \in \mathbb{R}$,

$$h(x) = 1 - \gamma \sum_{i=1}^{k-1} x^i + a_k x^k + \gamma \sum_{i=k+1}^{\infty} x^i. \tag{5.1}$$

Notice that h is an element of $\mathcal{B}_\gamma$ if and only if $a_k \in [-\gamma, \gamma]$, but this is not required in the definition.

Lemma 5.2. *Suppose that h is a $(*)$-function for $\mathcal{B}_\gamma$ such that*

$$h(x_0) > \delta \quad \textit{and} \quad h'(x_0) < -\delta \tag{5.2}$$

for some $x_0 \in (0,1)$ and $\delta \in (0,\gamma)$. Then $\mathcal{B}_\gamma$ satisfies the δ-transversality condition on $[0, x_0]$.

Proof. First we claim that

$$h(x) > \delta \quad \text{and} \quad h'(x) < -\delta \quad \text{for all} \;\; x \in [0, x_0]. \tag{5.3}$$

We have to consider the cases $k = 1$ and $k \ge 2$ separately. If $k = 1$, then $h(x) = 1 + a_1 x + \gamma \sum_{i=2}^\infty x^i$, with $a_1 < 0$ by (5.2). Obviously, h' is increasing on $[0, 1)$, hence (5.2) implies (5.3). Now, if $k \ge 2$ then $h'(0) = -\gamma < -\delta$. If $h'(x) \ge -\delta$ for any $x \in (0, x_0)$, then h' has a local maximum in $(0, x_0)$ and a local minimum in $(x_0, 1)$, due to the fact that $\lim_{x\uparrow 1} h'(x) = +\infty$. Then h'' has at least two zeros in $(0, 1)$, which is impossible for a power series with one coefficient sign change. Thus, $h'(x) < -\delta$ for $x \in [0, x_0]$, so h is decreasing on $[0, x_0]$, and (5.3) follows.

Now let $g \in \mathcal{B}_\gamma$. Consider $f(x) = g(x) - h(x)$. Then (5.1) and (4.4) imply that $f(x) = \sum_{i=1}^{\ell} c_i x^i - \sum_{i=\ell+1}^{\infty} c_i x^i$, where $c_i \ge 0$ and $\ell = k - 1$ or $\ell = k$, depending on the sign of a_k. We have for any $x \in [0, x_0]$, by (5.3),

$$g(x) < \delta \;\;\Rightarrow\;\; f(x) < 0 \;\;\Rightarrow\;\; f'(x) < 0 \;\;\Rightarrow\;\; g'(x) < -\delta,$$

proving δ-transversality, see Definition 4.2. The middle implication is a consequence of one coefficient sign change:

$$\begin{aligned} f(x) < 0 \quad &\Rightarrow \quad \sum_{i=1}^{\ell} c_i x^i < \sum_{i=\ell+1}^{\infty} c_i x^i \\ &\Rightarrow \quad \sum_{i=1}^{\ell} c_i i x^{i-1} < \sum_{i=\ell+1}^{\infty} c_i i x^{i-1} \;\;\Rightarrow\;\; f'(x) < 0. \end{aligned}$$

$\square$

Corollary 5.3. There exists $\delta > 0$ such that
(i) $\mathcal{B}_1$ satisfies the δ-transversality condition on $[0, .64)$.
(ii) $\mathcal{B}_2$ satisfies the δ-transversality condition on $[0, x_0)$ for any $x_0 < \frac{1}{2}$.

Proof (i) Let

$$h_1(x) = 1 - x - x^2 - x^3 + 0.1x^4 + \sum_{n=5}^{\infty} x^n.$$

This is a $(*)$-function for $\mathcal{B}_1$, and one can check that $h_1(.64) > 0$, $h_1'(.64) < 0$. The δ-transversality condition on $[0, .64)$ follows by Lemma 5.2. (The function h_1 was found with the help of a computer, but once it is given, the inequalities can be checked even by hand.)

(ii) Let

$$\begin{aligned} h_2(x) &= 1-2x-2x^2+2\sum_{n=3}^{\infty} x^n \\ &= 1-2x-2x^2+\frac{2x^3}{1-x} \\ &= \frac{4x^3-3x+1}{1-x} \\ &= \frac{(2x-1)^2(x+1)}{1-x}\,. \end{aligned}$$

We see that h_2 is a $(*)$-function for $\mathcal{B}_2$ and $h_2(x_0)>0,\ h_2'(x_0)<0$ for all $x_0\in[0,\frac{1}{2})$. It remains to apply Lemma 5.2. □

Proof of Theorem 2.7. (i) For the classical Bernoulli convolutions ν_λ we have $D=\{-1,1\}$, so $\gamma(D)=1$, see (4.5). By Theorem 4.3 and Corollary 5.3(i), $\nu_\lambda \ll \mathcal{L}$ with a density in $L^2(\mathbb{R})$ for a.e. $\lambda\in(1/2,.64)$. We need to establish the analogous statement for the interval $(.64,2^{-1/2})$. There are at least two ways to do this; here we follow [**31**] rather than [**30**]. Consider

$$\eta_\lambda=\nu_\lambda*\nu_\lambda.$$

This is the distribution of the sum of two independent realizations of the random series $\sum_{n=0}\pm\lambda^n$. It follows that η_λ is the distribution of $Z_\lambda=\sum_{n=0}^{\infty}c_n\lambda^n$ where $c_n\in\{-2,0,2\}$ with probabilities $\mathbf{p}=(\frac14,\frac12,\frac14)$. Thus, $\eta_\lambda=\nu_\lambda^{D,\mathbf{p}}$, where $D=\{-2,0,2\}$, so $\gamma(D)=2$. We compute $p_1^2+p_2^2+p_3^2=3/8$. By Theorem 4.3 and Corollary 5.3(ii), $\eta_\lambda\ll\mathcal{L}$ with a density in $L^2(\mathbb{R})$ for a.e. $\lambda\in(3/8,1/2)$. Recall that by Plancherel's theorem, a function is in $L^2(\mathbb{R})$ if and only if its Fourier transform is in $L^2(\mathbb{R})$. Thus, $\widehat{\eta}_\lambda\in L^2(\mathbb{R})$ for a.e. $\lambda\in(3/8,1/2)$. But $\widehat{\eta}_\lambda=(\widehat{\nu}_\lambda)^2$, so $\widehat{\nu}_\lambda\in L^4(\mathbb{R})$ for a.e. $\lambda\in(3/8,1/2)$. Recall (2.7), which for $m=2$ becomes

$$\widehat{\nu}_\lambda(\xi)=\widehat{\nu}_{\lambda^2}(\xi)\widehat{\nu}_{\lambda^2}(\lambda\xi).$$

It follows that $\widehat{\nu}_\lambda\in L^2(\mathbb{R})$ for a.e. $\lambda\in((3/8)^{1/2},2^{-1/2})$. Since $(3/8)^{1/2}\approx .61237<.64$, the proof of (i) is complete.

(ii) By part (i) and Plancherel's theorem, $\widehat{\nu}_\lambda\in L^2(\mathbb{R})$ for a.e. $\lambda\in(1/2,2^{-1/2})$. In view of (2.7), $\widehat{\nu}_\lambda\in L^{2/m}(\mathbb{R})$ for a.e. $\lambda\in(2^{-1/m},2^{-1/(2m)})$ where $m=2,3,\ldots$ This implies that $\widehat{\nu}_\lambda\in L^1(\mathbb{R})$ for a.e. $\lambda\in(2^{-1/2},1)$ (here we use that $|\widehat{\nu}_\lambda|$ is always bounded by one, being the Fourier transform of a probability measure). For such λ the density of ν_λ is continuous, being the inverse Fourier transform of an L^1 function. The proof is finished. □

Remark 5.4. There is still something mysterious about the role of transversality. In [**37**] we noted that the reciprocal of the fourth Pisot number $\lambda_4\approx 0.68233<2^{-1/2}$, satisfying $\lambda_4^3+\lambda_4=1$, is a double zero of a power series with coefficients $0,\pm1$. This implies that δ-transversality fails on any interval containing λ_1, for any positive δ. One can show that δ-transversality fails *on any* interval contained in $[2^{-1/2},1)$. Thus, transversality is not necessary for almost sure absolute continuity of ν_λ, but we still do not know how to avoid using it.

6. Application: Dimension of Self-Affine Sets

Definition 6.1. Let $(\mathcal{X}, \varrho)$ be a metric space. A map $S: \mathcal{X} \to \mathcal{X}$ is called a **contraction** if there exists $c \in (0,1)$ such that $\varrho(S(x), S(y)) \le c\varrho(x,y)$ for all $x, y \in \mathcal{X}$. An **iterated function system** (IFS) is a family $\{S_1, \ldots, S_m\}$ of contractions from $\mathcal{X}$ into itself. A non-empty compact set $\mathcal{K} \subset X$ is called an **attractor** of the IFS $\{S_1, \ldots, S_m\}$ if

$$\mathcal{K} = \bigcup_{i \le m} S_i(\mathcal{K}).$$

By a theorem of Hutchinson [**15**], every IFS on a complete metric space has a unique attractor.

Definition 6.2. If $\mathcal{X} = \mathbb{R}^d$ and the maps in the IFS have the form $S_i(x) = T_i x + \mathbf{a}_i$ for some linear contractions T_i and $\mathbf{a}_i \in \mathbb{R}^d$, $i \le m$, then the attractor is called a **self-affine set**.

If, moreover, the maps S_i are similitudes, that is, $|S_i(x) - S_i(y)| = r_i|x-y|$ for all $x, y \in \mathbb{R}^d$, then the attractor is called a **self-similar set**. The unique number $\alpha > 0$ such that $\sum_{i=1}^m r_i^\alpha = 1$, is called the **similarity dimension** of the IFS $\{S_1, \ldots, S_m\}$.

The dimension and measure properties of self-similar sets are fairly well understood, especially when the "pieces" of the attractor do not overlap.

Definition 6.3. The IFS $\{S_i\}_{i \le m}$ is said to satisfy the **open set condition** if there is a non-empty open set U such that $S_i(U) \subset U$ and $S_i(U) \cap S_j(U) = \emptyset$ for all $i \ne j$. We say that the **strong separation condition** holds if the attractor $\mathcal{K}$ satisfies $S_i(\mathcal{K}) \cap S_j(\mathcal{K}) = \emptyset$ for all $i \ne j$.

Theorem 6.4 (Moran 1946, Hutchinson 1981). *If the IFS of similitudes satisfies the open set condition, then the Hausdorff dimension of the attractor equals the similarity dimension.*

Notice, in particular, that the Hausdorff dimension of a self-similar set depends only on the linear parts of S_i and not on the translations parts, as long as the open set condition holds. For a general self-affine set the situation is much more complicated.

Digression: when does (and what if) the open set condition fails? The IFS $\{\lambda x - 1, \lambda x + 1\}$, corresponding to the Bernoulli convolution measure ν_λ, does not satisfy the open set condition for $\lambda > \frac{1}{2}$. More generally, Theorem 6.4 implies that the open set condition fails whenever the similarity dimension is greater than d, the dimension of the whole space. In other cases, it may appear that the IFS has an "overlap," but it is non-trivial to check that the open set condition actually fails, see [**29**] and references therein.

Sometimes the dimension of the attractor of an "overlapping" IFS may be found easily, as in the Bernoulli convolution case, when the attractor is an interval for $\lambda > \frac{1}{2}$, but there are many non-trivial examples and open problems. An interesting family of self-similar sets $\mathcal{C}_\lambda$ arises as attractors of the IFS $\{\lambda x, \lambda x + 1, \lambda x + 3\}$. The open set condition fails for all $\lambda \in (\frac{1}{3}, 1)$, since the similarity dimension is greater than 1. These sets were investigated in [**18, 32, 37, 38, 27**]; in fact, it was the study of $\{\mathcal{C}_\lambda\}$ by Pollicott and Simon [**32**] that initiated the recent advances on

Bernoulli convolutions. Consider the self-similar measure $\nu_\lambda^{D,\mathbf{p}}$ where $D = \{0, 1, 3\}$ and $\mathbf{p} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. One can deduce from Theorem 4.3 and Lemma 5.2 that $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous for a.e. $\lambda \in (\frac{1}{3}, \frac{2}{5})$, which implies that $\mathcal{C}_\lambda$ has positive Lebesgue measure for such λ, see [**37**]. (It is easy to check that $\mathcal{C}_\lambda$ is an interval for all $\lambda \in (\frac{2}{5}, 1)$.) An interesting open problem is to decide whether $\mathcal{C}_\lambda$ has **nonempty interior** for a typical $\lambda > \frac{1}{3}$. (In [**18**] a countable set of $\lambda > \frac{1}{3}$ for which $\mathcal{C}_\lambda$ has measure zero, and even dimension less than 1, was found.)

Now we return to self-affine sets and consider the special case of an IFS $\{S_1, S_2\}$ in $\mathbb{R}^2$, with $S_1(x) = Tx+\mathbf{a}_1$, $S_2(x) = Tx+\mathbf{a}_2$ (that is, the two contractions have the same linear part). First, we make a change of variable to reduce the IFS to as simple form as possible. It is clear that applying an invertible affine linear map preserves the Hausdorff dimension (this is true for any bi-Lipschitz map). If $\mathcal{K}$ is the attractor of the IFS $\{S_i\}_{i\le m}$, then $F^{-1}(\mathcal{K})$ is the attractor of the IFS $\{F^{-1} \circ S_i \circ F\}_{i\le m}$.

Let $F(x) = Ax + \mathbf{b}$ for some invertible linear $A : \mathbb{R}^2 \to \mathbb{R}^2$ and $\mathbf{b} \in \mathbb{R}^2$. Then for $S_i : x \mapsto Tx + \mathbf{a}_i$ we have

$$F^{-1} \circ S_i \circ F : x \mapsto A^{-1}TAx + A^{-1}((T - I)\mathbf{b} + \mathbf{a}_i). \tag{6.1}$$

Assume, in addition, that T is diagonalizable over $\mathbb{C}$. If the eigenvalues of T are complex, then we can find A so that $A^{-1}TA$ corresponds to the multiplication by a complex number, which is a similitude, hence the attractor is a self-similar set. In the genuinely self-affine case, T can be reduced to a diagonal matrix with two distinct real eigenvalues. Thus, without loss of generality, we can assume that $T = \begin{bmatrix} \lambda & 0 \\ 0 & \gamma \end{bmatrix}$, with $|\gamma| < |\lambda| < 1$. For simplicity we will assume that the eigenvalues are positive. Further, choosing $\mathbf{b} = (I - T)^{-1}\mathbf{a}_1$ in (6.1), we make the translation part of the first contraction zero, that is, reduce the IFS to $\{Tx, Tx+\mathbf{a}\}$. We can simplify the IFS even further, making a transformation $F(x) = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix} x$, which keeps the linear part T unchanged (it being already diagonalized). Thus, we can reduce the translation part $\mathbf{a}$ of the second contraction to one of the three cases: $\mathbf{a} = (1, 0)$, $\mathbf{a} = (0, 1)$, and $\mathbf{a} = (1, 1)$. (There is also the trivial case $\mathbf{a} = (0, 0)$ when the attractor is the single point $\{(0, 0)\}$.)

Denote the attractor of the IFS by $\mathcal{K}_{\lambda,\gamma}$.

Case 1: $\mathbf{a} = (1, 0)$. We have $\mathcal{K}_{\lambda,\gamma} \subset \mathbb{R} \times \{0\}$. It is a Cantor set for $0 < \lambda < \frac{1}{2}$ and the line segment $[0, \frac{1}{1-\lambda}] \times \{0\}$ for $\frac{1}{2} \le \lambda < 1$. We have

$$\dim_H \mathcal{K}_{\lambda,\gamma} = \begin{cases} \frac{\log 2}{\log(1/\lambda)}, & \text{if } 0 < \lambda < \frac{1}{2}; \\ 1, & \text{if } \frac{1}{2} \le \lambda < 1. \end{cases}$$

Notice that γ does not play any role here.

Case 2: $\mathbf{a} = (0, 1)$. We have $\mathcal{K}_{\lambda,\gamma} \subset \{0\} \times \mathbb{R}$. This case is similar to the previous one. We have

$$\dim_H \mathcal{K}_{\lambda,\gamma} = \begin{cases} \frac{\log 2}{\log(1/\gamma)}, & \text{if } 0 < \gamma < \frac{1}{2}; \\ 1, & \text{if } \frac{1}{2} \le \gamma < 1. \end{cases}$$

Notice that λ does not play any role here.

Case 3: $\mathbf{a} = (1, 1)$. This is the generic case. The sub-case $\lambda \le \frac{1}{2}$ is easy, see Figure 3. It turns out that the answer is the same as in the first part of Case 1, γ again does not play any role.

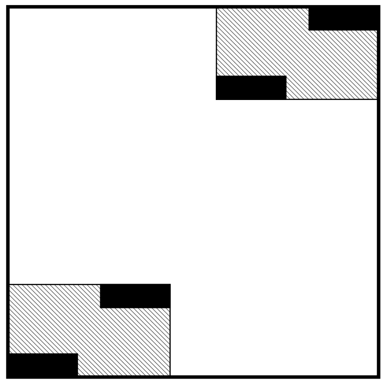

FIGURE 3. $\mathbf{a} = (1,1)$, $\lambda < \frac{1}{2}$

PROPOSITION 6.5. **(folklore)** If $0 < \gamma < \lambda \le \frac{1}{2}$, then

$$\dim_H \mathcal{K}_{\lambda,\gamma} = \frac{\log 2}{\log(1/\lambda)}\,.$$

Proof sketch. The Hausdorff dimension does not increase when the set is projected on the x-axis, and this projection is a Cantor set of Hausdorff dimension $\frac{\log 2}{\log(1/\lambda)}$ (or the full interval if $\lambda = \frac{1}{2}$). It remains to prove the upper bound. Applying the maps of the IFS n times, we see that $\mathcal{K}_{\lambda,\gamma}$ is covered by the union of 2^n rectangles of size $c\lambda^n \times c'\gamma^n$. Using these rectangles as our cover by sets of diameter $\sim \lambda^n$, we obtain that $\mathcal{H}^{\log 2/\log(1/\lambda)}(\mathcal{K}_{\lambda,\gamma}) < \infty$, hence $\dim_H \mathcal{K}_{\lambda,\gamma} \le \frac{\log 2}{\log(1/\lambda)}$, as desired. □

Already the cases we considered show that for self-affine sets the dimension depends not only on the linear part, but also on the translation parts of the contractions, even when the strong separation condition holds. The remaining subcase $\mathbf{a} = (1,1), \lambda \in (\frac{1}{2}, 1)$, is more delicate, see Figure 4. Perhaps surprisingly, the Hausdorff dimension of $\mathcal{K}_{\lambda,\gamma}$ depends on the properties of the Bernoulli convolution measure ν_λ! (One can motivate the connection, however, by observing that the IFS $\{\lambda x, \lambda x + 1\}$, which is conjugate to the one defining ν_λ, arises by projecting the planar IFS onto the x-axis.)

The Hausdorff dimension of a finite Borel measure ν on $\mathbb{R}$ is defined by

$$\dim_H \nu = \inf\{\dim_H A : \ A \subset \mathbb{R},\, \nu(A) > 0\}.$$

It is immediate that if $\dim_H \nu < 1$ then ν is singular.

THEOREM 6.6 (Przytycki and Urbański [**34**]). *Let $\mathcal{K}_{\lambda,\gamma}$ be the attractor of the IFS $\{Tx, Tx + \mathbf{a}\}$, where $T = \begin{bmatrix} \lambda & 0 \\ 0 & \gamma \end{bmatrix}$, with $0 < \gamma \le \frac{1}{2} < \lambda < 1$, and $\mathbf{a} = (1,1)$. Then*

(i) *if $\dim_H \nu_\lambda = 1$, then*

$$\dim_H \mathcal{K}_{\lambda,\gamma} = 1 + \frac{\log(2\lambda)}{\log(1/\gamma)}\,; \tag{6.2}$$

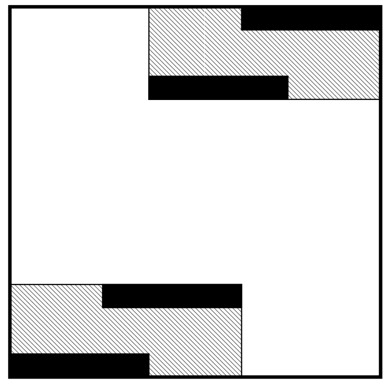

FIGURE 4. $\mathbf{a} = (1, 1)$, $\lambda > \frac{1}{2}$

(ii) *if* $1/\lambda$ *is a Pisot number, then*

$$\dim_H \mathcal{K}_{\lambda,\gamma} < 1 + \frac{\log(2\lambda)}{\log(1/\gamma)}.$$

This was proved in [**34**] for $\gamma = \frac{1}{2}$, but the proof readily extends to stated generality, as we do below. In [**33**] a weaker statement than Theorem 6.6(i) is proved, since there the dimension formula is established under the assumption that $\frac{d\nu_\lambda}{dx} \in L^\infty(\mathbb{R})$. On the other hand, [**33**] makes an interesting observation that the sets $\mathcal{K}_{\lambda,\gamma}$ (for $\gamma < \frac{1}{2}$) arise from certain "horseshoes" in $\mathbb{R}^3$ by intersecting them with the stable manifold.

Combining Theorem 6.6(i) with Theorem 2.7 we immediately obtain the following.

COROLLARY 6.7. Under the assumptions of Theorem 6.6, the dimension formula (6.2) holds for a.e. $\lambda \in (\frac{1}{2}, 1)$, for any $\gamma \in (0, \frac{1}{2})$.

REMARK 6.8. Since $\dim_H \nu_\lambda = 1$ for a *residual* set of $\lambda \in (\frac{1}{2}, 1)$, see [**28**], it follows also that (6.2) holds for a residual set of $\lambda \in (\frac{1}{2}, 1)$, for any $\gamma \in (0, \frac{1}{2})$. It is an open problem whether $\nu_\lambda \ll \mathcal{L}$ for a residual set of $\lambda \in (\frac{1}{2}, 1)$.

Next we give a complete proof of Theorem 6.6(i). For the proof of (ii) the reader is referred to [**34**].

Proof of Theorem 6.6(i). The upper estimate in (6.2) is standard and holds without any assumptions on λ. In fact, the set $\mathcal{K}_{\lambda,\gamma}$ is covered by 2^n rectangles of size $c\lambda^n \times c'\gamma^n$. Each of these rectangles may be covered by $\sim \frac{c\lambda^n}{c'\gamma^n}$ balls of diameter $2c'\gamma^n$. Thus,

$$\begin{aligned} \dim_H \mathcal{K}_{\lambda,\gamma} &\leq \lim_{n\to\infty} \frac{\log(2^n \frac{c\lambda^n}{c'\gamma^n})}{-\log(2c'\gamma^n)} \\ &= 1 + \frac{\log(2\lambda)}{\log(1/\gamma)}. \end{aligned}$$

For the lower estimate in (6.2) we need the following classical result in fractal geometry, see, e.g., [**13**, Prop. 2.3].

LEMMA 6.9. *Let $E \subset \mathbb{R}^d$ be a Borel set and let η be a finite Borel measure on $\mathbb{R}^d$.*

(i) *If* $\liminf_{r\downarrow 0} \dfrac{\log \eta(B_r(x))}{\log r} \ge s$ *for η-a.e. $x \in E$ and $\eta(E) > 0$ then* $\dim_H E \ge s$.

(ii) *If* $\liminf_{r\downarrow 0} \dfrac{\log \eta(B_r(x))}{\log r} \le s$ *for all $x \in E$ then* $\dim_H E \le s$.

Let $\Omega = \{0,1\}^{\mathbb{N}}$. Define $\Pi:\ \Omega \to \mathbb{R}^2$ by

$$\Pi(\omega) = (\Pi_\lambda(\omega), \Pi_\gamma(\omega)) = \Big(\sum_{n=0}^\infty \omega_n \lambda^n, \sum_{n=0}^\infty \omega_n \gamma^n\Big)$$

(the numbers λ and γ are fixed throughout the proof). Then $\mathcal{K}_{\lambda,\gamma} = \Pi(\Omega)$. The natural measure on the attractor is $\eta = \mu \circ \Pi^{-1}$, where $\mu = (\frac12, \frac12)^{\mathbb{N}}$ on Ω. Note that projecting η on the x-axis yields the Bernoulli convolution ν_λ (or rather, its rescaled copy on the interval $[0, \frac{1}{1-\lambda}]$). We will apply Lemma 6.9(i) to the measure η. The idea is that a ball $B_r(x)$, with $r \sim \gamma^n$, centered at some $x \in \mathcal{K}_{\lambda,\gamma}$, will cover the chunk $c'\gamma^n \times c'\gamma^n$ of an n-th generation rectangle (whose shorter side is $c'\gamma^n$). Using the fact that the measure η is invariant under the IFS, we will obtain that the η measure of this chunk equals 2^{-n} times the η measure of a rectangle with horizontal side $\sim \lambda^{-n}\gamma^n$, which vertically extends all the way from top to bottom of the square in Figure 4. Its η measure is just the ν_λ measure of the projection, and we can use the dimension assumption on the latter to get the needed estimate. This is a rough idea; now for the details.

For $\omega, \tau \in \Omega$, with $|\omega \wedge \tau| = k$, we have

$$\begin{aligned} |\Pi(\omega) - \Pi(\tau)| &\ge \Big|\sum_{n=0}^\infty (\omega_n - \tau_n)\gamma^n\Big| \\ &\ge \gamma^k - \sum_{n=k+1}^\infty |\omega_n - \tau_n|\gamma^n \\ &\ge \gamma^k - \frac{\gamma^{k+1}}{1-\gamma} \\ &\ge \gamma^k \cdot \frac{1-2\gamma}{1-\gamma} > 0, \end{aligned} \tag{6.3}$$

since $\gamma \in (0, \frac12)$. Let $c_1 = \frac{1-2\gamma}{1-\gamma}$. It follows from (6.3) that for any $\omega \in \Omega$, $r > 0$, and $\ell \in \mathbb{N}$ such that $\gamma^\ell \ge c_1 r$,

$$\begin{aligned} \eta(B_r(\Pi(\omega))) &= \mu\{\tau \in \Omega:\ |\Pi(\omega) - \Pi(\tau)| \le r\} \\ &\le \mu\{\tau \in \Omega:\ |\omega \wedge \tau| \ge \ell,\ |\Pi_\lambda(\omega) - \Pi_\lambda(\tau)| \le r\} \\ &= \mu\{\tau \in \Omega:\ |\omega \wedge \tau| \ge \ell,\ \lambda^\ell|\Pi_\lambda(\sigma^\ell\omega) - \Pi_\lambda(\sigma^\ell\tau)| \le r\} \\ &= 2^{-\ell}\mu\{\tau' \in \Omega:\ |\Pi_\lambda(\sigma^\ell\omega) - \Pi_\lambda(\tau')| \le \lambda^{-\ell} r\}\,. \end{aligned} \tag{6.4}$$

The last equality follows from the fact that μ is a product measure; here σ is the left shift on the sequence space Ω.

Observe that $\mu \circ \Pi_\lambda^{-1}$ is a self-similar measure on $\mathbb{R}$, which is just a rescaled copy of the Bernoulli convolution ν_λ. Abusing notation slightly, we denote $\nu_\lambda := \mu \circ \Pi_\lambda^{-1}$. Thus, the last line in (6.4) equals $2^{-\ell}\nu_\lambda(B_{\lambda^{-\ell}r}(\Pi_\lambda(\sigma^\ell\omega)))$, hence

$$\gamma^\ell \ge c_1 r \ \Rightarrow\ \eta(B_r(\Pi(\omega))) \le 2^{-\ell}\nu_\lambda(B_{\lambda^{-\ell}r}(\Pi_\lambda(\sigma^\ell\omega))). \tag{6.5}$$

Since $\dim_H \nu_\lambda = 1$, Lemma 6.9 implies

$$\liminf_{r\downarrow 0} \frac{\log \nu_\lambda(B_r(x))}{\log(r)} = 1 \quad \text{for } \nu_\lambda\text{-a.e. } x,$$

hence, for μ-a.e. $\omega \in \Omega$,

$$\liminf_{r\downarrow 0} \frac{\log \nu_\lambda(B_r(\Pi_\lambda(\omega)))}{\log(r)} = 1 \quad \text{for } \mu\text{-a.e. } \omega \in \Omega.$$

By Egorov's theorem, there is a set $\Omega' \subset \Omega$ such that $\mu(\Omega') > 0$ and

$$\lim_{r\downarrow 0} \inf_{\rho \le r} \frac{\log \nu_\lambda(B_\rho(\Pi_\lambda(\omega)))}{\log(\rho)} = 1 \tag{6.6}$$

uniformly on Ω'.

Next we apply Birkhoff's Ergodic Theorem to the Bernoulli system (Ω, σ, μ), see, e.g., [**26**]. This theorem asserts that for any $f \in L^1(\Omega, \mu)$,

$$\lim_{N\to\infty} \frac{1}{N+1} \sum_{n=0}^{N} f(\sigma^n \omega) = \int_\Omega f(\tau)\, d\mu(\tau),$$

for μ-a.e. $\omega \in \Omega$. Taking $f = \mathbf{1}_{\Omega'}$, we obtain

$$\lim_{N\to\infty} (N+1)^{-1} \#\{n \le N : \ \sigma^n \omega \in \Omega'\} = \mu(\Omega') > 0, \tag{6.7}$$

for μ-a.e. $\omega \in \Omega$. Let Ω'' be the set of $\omega \in \Omega$ for which (6.7) holds. We know that $\mu(\Omega'') = 1$, and our goal is to verify the condition in Lemma 6.9(i), with appropriate s, for an arbitrary $x \in \Pi(\Omega'')$.

Fix $\omega \in \Omega''$. Let n_j be the j-th element of the set $\{n \in \mathbb{N} : \ \sigma^n \omega \in \Omega'\}$. By (6.7), this set has a well-defined positive density in $\mathbb{N}$, so

$$\lim_{j\to\infty} \frac{n_{j+1}}{n_j} = 1. \tag{6.8}$$

Fix $\varepsilon > 0$. Since the convergence in (6.6) is uniform on Ω' and $\sigma^{n_j}\omega \in \Omega'$, $j \ge 1$, there exists $r_0 > 0$ such that for all $j \ge 1$ and all $\rho \le r_0$,

$$\frac{\log \nu_\lambda(B_\rho(\Pi_\lambda(\sigma^{n_j}\omega)))}{\log \rho} \ge 1 - \varepsilon. \tag{6.9}$$

For a small $r > 0$, let $j = j(r)$ be such that

$$\gamma^{n_{j+1}} < c_1 r \le \gamma^{n_j}. \tag{6.10}$$

Clearly, $j(r) \to \infty$ as $r \to 0$. By (6.5),

$$\eta(B_r(\Pi(\omega))) \le 2^{-n_j} \nu_\lambda(B_{\lambda^{-n_j} r}(\Pi_\lambda(\sigma^{n_j}\omega))). \tag{6.11}$$

By (6.10), we have $\lambda^{-n_j} r \le c_1^{-1} (\frac{\gamma}{\lambda})^{n_j} \to 0$, as $r \to 0$. Thus, for r sufficiently small we obtain from (6.11) and (6.9):

$$\begin{aligned} \log \eta(B_r(\Pi(\omega))) &\le -n_j \log 2 + (1-\varepsilon) \log(\lambda^{-n_j} r) \\ &= -n_j[\log 2 + (1-\varepsilon)\log\lambda] + (1-\varepsilon)\log r. \end{aligned} \tag{6.12}$$

By (6.10), we have $r^{-1} < c_1\gamma^{-n_{j+1}}$, hence $\log(1/r) < \log c_1 - n_{j+1}\log\gamma$, and we obtain from (6.12), keeping in mind that $\log r < 0$:

$$\begin{aligned}\frac{\log\eta(B_r(\Pi(\omega)))}{\log r} &\geq (1-\varepsilon) + \frac{n_j(\log 2 + (1-\varepsilon)\log\lambda)}{\log(1/r)} \\ &\geq (1-\varepsilon) + \frac{n_j(\log 2 + (1-\varepsilon)\log\lambda)}{\log c_1 - n_{j+1}\log\gamma}.\end{aligned}$$

The last expression approaches $(1-\varepsilon) + \frac{\log 2 + (1-\varepsilon)\log\lambda}{-\log\gamma}$, as $r \to 0$, in view of (6.8). By Lemma 6.9, it follows that

$$\dim_H \mathcal{K}_{\lambda,\gamma} \geq (1-\varepsilon) + \frac{\log 2 + (1-\varepsilon)\log\lambda}{-\log\gamma},$$

and since $\varepsilon > 0$ was arbitrary, $\dim_H \mathcal{K}_{\lambda,\gamma} \geq 1 + \frac{\log(2\lambda)}{\log(1/\gamma)}$, as desired. □

7. Some Generalizations and Applications

7.1. Biased Bernoulli convolutions. For $p \in (0,1)$ let ν_λ^p be the distribution of the random series $\sum_{n=0}^\infty a_n\lambda^n$, where $a_n \in \{-1,1\}$ independently with probabilities $(p, 1-p)$. It turns out that "biased" Bernoulli convolutions ν_λ^p, with $p \neq \frac{1}{2}$, exhibit some features absent in the classical case $p = \frac{1}{2}$.

Corollary 7.1. The measure ν_λ^p is absolutely continuous with a density in $L^2(\mathbb{R})$ for a.e. $\lambda \in (p^2 + (1-p)^2, .64)$.

Proof. This follows from Theorem 4.3 and Corollary 5.3. □

Note that the interval of λ in the corollary is non-empty for $p \in (.5 - \sqrt{.7}, .5 + \sqrt{.7}) \approx (.235, .765)$. If p is not in this interval, the corollary does not say anything. One can ask what happens for $\lambda \in (.64, 1)$. Sometimes absolute continuity can be established for a.e. $\lambda \in (.64, 1)$, using tricks like at the end of Section 5, see [**31**]. But what happens for $\lambda \in (\frac{1}{2}, p^2 + (1-p)^2)$? For a fixed λ, all the measures ν_λ^p, $p \in (0,1)$, have the same support; it is the interval $[-(1-\lambda)^{-1}, (1-\lambda)^{-1}]$ for $\lambda > \frac{1}{2}$.

Proposition 7.2. **(i)** $\nu_\lambda^p \ll \mathcal{L}$ for a.e. $\lambda \in (p^p(1-p)^{(1-p)}, .64)$.
(ii) Let $p \in (0,1)$. The measure ν_λ^p is singular for all $\lambda < p^p(1-p)^{(1-p)}$.

Note that the interval $(p^p(1-p)^{(1-p)}, .64)$ is non-empty for $p \in (.165, .835)$ (approximately). Proposition 7.2 follows from a more general result, which also contains information on when the density belongs to $L^q(\mathbb{R})$.

Theorem 7.3 ([**31**]). *Let $\nu_\lambda^{D,\mathbf{p}}$ be the family of self-similar measures corresponding to a digit set $D = \{d_1, \ldots, d_m\}$ and a probability vector $\mathbf{p} = (p_1, \ldots, p_m)$. Let J be a closed subinterval of $[0,1)$, such that $\mathcal{B}_{\gamma(D)}$ satisfies the δ-transversality condition on J for some $\delta > 0$. Then*

(i) *The measure $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous for a.e. $\lambda \in J \cap (\prod_{i=1}^m p_i^{p_i}, 1)$, and singular for all $\lambda < \prod_{i=1}^m p_i^{p_i}$.*

(ii) *Let $q \in (1,2]$. Then the measure $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous with a density in $L^q(\mathbb{R})$ for a.e. $\lambda \in J \cap ((\sum_{i=1}^m p_i^q)^{\frac{1}{q-1}}, 1)$.*

(iii) *For any $q > 1$ and all $\lambda \in (0,1)$, if $\nu_\lambda^{D,\mathbf{p}}$ is absolutely continuous with respect to Lebesgue measure and its density $\frac{d\nu_\lambda}{dx}$ is in $L^q(\mathbb{R})$, then $\lambda \geq (\sum_{i=1}^m p_i^q)^{\frac{1}{q-1}}$.*

REMARK 7.4. Notice that the function $q \mapsto (\sum_{i=1}^m p_i^q)^{\frac{1}{q-1}}$ is increasing in q for $q > 1$, and it converges to $\prod_{i=1}^m p_i^{p_i}$ as $q \to 1$. It is strictly increasing if and only if $\mathbf{p} \neq (\frac{1}{m}, \ldots, \frac{1}{m})$. Thus, for such measures, e.g. for the biased Bernoulli convolutions, the thresholds for almost sure existence of L^q densities change with q. As a result, we get a gradual change in the typical properties of the density of $\nu_\lambda^{D,\mathbf{p}}$ as λ increases.

OPEN QUESTION. Find a specific λ and $p \neq \frac{1}{2}$ so that ν_λ^p is absolutely continuous. In the unbiased case at least we know absolute continuity for Garsia numbers—no such example is known in the biased case!

7.2. Application to functional equations. Consider the functional equation

$$f(x) = \sum_{i=1}^m c_i f(\alpha x - \beta_i), \tag{7.1}$$

where $\alpha > 1$, $\beta_1 < \beta_2 < \ldots < \beta_m$, and c_i are constants. It is called a **two-scale difference equation**, or **refinement equation**. Such equations have been much studied in the theory of wavelets, see [**8**] and references therein. Usually one looks for L^1 or continuous solutions with compact support. The case of most interest is

$$\sum_{i=1}^m c_i = \alpha,$$

when (7.1) has at most one non-trivial L^1 solution, up to a multiplicative constant, see [**8**]. If, in addition, $c_i > 0$ for all i, we can consider the self-similar measure

$$\nu = \sum_{i=1}^m p_i(\nu \circ S_i^{-1}), \text{ with } p_i = c_i/\alpha \text{ and } S_i(x) = \alpha^{-1}(x + \beta_i). \tag{7.2}$$

If $\nu \ll \mathcal{L}$, then the density is a compactly supported L^1 solution of (7.1). One special case that attracted attention, is the so-called "Schilling equation"

$$4\lambda f(\lambda t) = f(t+1) + f(t-1) + 2f(t), \tag{7.3}$$

which has its origins in physics. It reduces to the form in (7.1) by taking $\alpha = \frac{1}{\lambda}$ and $t = \alpha x$. The equation (7.3) was studied in [**9**] and references therein. The corresponding self-similar measure is $\nu_\lambda^{D,\mathbf{p}}$, where $D = \{-1, 0, 1\}$ and $\mathbf{p} = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. Notice that $\nu_\lambda^{D,\mathbf{p}}$ is (up to scaling) the convolution product $\nu_\lambda * \nu_\lambda$ where ν_λ is the classical Bernoulli convolution associated with $D = \{-1, 1\}$ and $\mathbf{p} = (\frac{1}{2}, \frac{1}{2})$.

COROLLARY 7.5 ([**9**],[**31**]). **(i)** There is a compactly supported L^1 solution of (7.3) for a.e. $\lambda \in (\frac{1}{2\sqrt{2}}, 1)$.

(ii) There is a continuous compactly supported solution of (7.3) for a.e. $\lambda \in (\frac{1}{2}, 1)$.

(iii) There is no compactly supported L^∞ solution of (7.3) for any $\lambda \in (0, \frac{1}{2})$.

Proof. For $D = \{-1, 0, 1\}$ we have $\gamma(D) = 2$, see (4.5); hence by Theorem 7.3(ii) and Corollary 5.3(ii), $\nu_\lambda^{D,\mathbf{p}} \ll \mathcal{L}$ for a.e. $\lambda \in (\frac{1}{2\sqrt{2}}, \frac{1}{2})$ (note that $\frac{1}{2\sqrt{2}} = (\frac{1}{2})^{\frac{1}{2}}(\frac{1}{4})^{\frac{1}{4}}(\frac{1}{2})^{\frac{1}{2}}$). For a.e. $\lambda \in (\frac{1}{2}, 1)$ we have that ν_λ has a density in L^2, hence $\nu_\lambda * \nu_\lambda$ has a continuous density. This proves (ii) and concludes the proof of (i). Part (iii) follows from Theorem 7.3(iii). $\square$

7.3. Self-similar measures with non-uniform contraction ratios. Let $\nu_{\lambda_1,\lambda_2}$ be the self-similar measure corresponding to the IFS $\{S_1, S_2\}$ on $\mathbb{R}$ with

$$S_1(x) = \lambda_1 x, \quad S_2(x) = \lambda_2 x + 1, \quad p_1 = p_2 = 1/2.$$

If $\lambda_1 \neq \lambda_2$, then we say that $\nu_{\lambda_1,\lambda_2}$ has non-uniform contraction ratios. Such measures are harder to study than the ones with uniform contraction ratios, discussed so far. Recently, Neunhäuserer [**23**] and Ngai and Wang [**24**] independently obtained some results on absolute continuity. They followed the general scheme of [**30, 31**] presented here, but had to introduce some new ideas as well.

THEOREM 7.6 ([**23**], [**24**]). **(i)** *If $\lambda_1\lambda_2 < \frac{1}{4}$, then $\nu_{\lambda_1,\lambda_2} \perp \mathcal{L}$.*
(ii) *$\nu_{\lambda_1,\lambda_2} \ll \mathcal{L}$ for a.e. pair $(\lambda_1, \lambda_2) \in (0, .64)^2$ such that $\lambda_1\lambda_2 > \frac{1}{4}$.*

In fact, the absolute continuity is proved for the family $\nu_{a\lambda,\lambda}$ for a fixed $a \in (0, 1)$ and a.e. λ in some interval. The number 0.64 appears here because of the transversality condition for $\mathcal{B}_1$, see Corollary 5.3. In [**24**] the region of almost sure absolute continuity is slightly larger, but it is still far from covering the whole set $\{(\lambda_1, \lambda_2) \in (0, 1)^2 : \ \lambda_1\lambda_2 > \frac{1}{4}\}$ where it is expected to hold. The difficulty here is that $\nu_{\lambda_1,\lambda_2}$ is not an infinite convolution if $\lambda_1 \neq \lambda_2$, so the tricks based on (2.7) do not work.

7.4. Complex-valued case. Consider the IFS $\{\lambda z - 1, \lambda z + 1\}$ in the complex plane $\mathbb{C}$ for $\lambda \in \mathbb{C}$, $|\lambda| < 1$, and let ν_λ be the corresponding self-similar measure (with probabilities $(\frac{1}{2}, \frac{1}{2})$). We can view the family of measures $\{\nu_\lambda : \ |\lambda| < 1\}$ as the analytic continuation of the family of classical Bernoulli convolutions $\{\nu_\lambda : \ \lambda \in (0, 1)\}$. It is natural to ask for which $\lambda \notin \mathbb{R}$ the measure ν_λ is absolutely continuous with respect to the area measure $\mathcal{L}_2$. Perhaps surprisingly, in the complex-valued case even the structure of $\mathcal{K}_\lambda := \mathrm{supp}(\nu_\lambda)$ is non-trivial. The set $\mathcal{K}_\lambda$ is self-similar; in fact,

$$\mathcal{K}_\lambda = (\lambda\mathcal{K}_\lambda - 1) \cup (\lambda\mathcal{K}_\lambda + 1).$$

When λ is real, it is obvious that $\mathcal{K}_\lambda$ is an interval for $|\lambda| \geq \frac{1}{2}$ and a Cantor set for $|\lambda| < \frac{1}{2}$, but for $\lambda \notin \mathbb{R}$ we get much more complicated behavior. There are many interesting problems concerning the set

$$\mathcal{M} := \{\lambda \in \mathbb{C} : \ |\lambda| < 1, \ \mathcal{K}_\lambda \text{ is connected}\}.$$

It was studied by Barnsley and Harrington [**4**] (see also [**3**, Ch.8]), who called it the "Mandelbrot set for the pair of transformations" $\{\lambda z - 1, \lambda z + 1\}$. There are indeed some parallels with the famous Mandelbrot set in complex dynamics; for instance, $\mathcal{K}_\lambda$ is totally disconnected for all $\lambda \notin \mathcal{M}$. There are pictures of $\mathcal{M}$ which show fractal-looking boundary, but to our knowledge, there are no rigorous results on its dimension. Notice that $\lambda \notin \mathcal{M}$ if and only if the IFS $\{\lambda z - 1, \lambda z + 1\}$ satisfies the strong separation condition.

One striking feature of $\mathcal{M}$ is that it has two "spikes," or "antennas," on the real line. In fact, $\mathcal{M} \cap \mathbb{R}_+ = [\frac{1}{2}, 1)$, but $\mathrm{clos}(\mathcal{M} \setminus \mathbb{R}) \cap \mathbb{R}_+ = [\alpha, 1)$ with $\alpha > \frac{1}{2}$. It is shown in [**38**] that α is greater or equal (conjecturally equal) to the smallest double zero of a power series in $\mathcal{B}_1$, so $\alpha \geq .64$ by Corollary 5.3(i) (numerical calculations suggest that $\alpha \approx .67$).

Bousch [**6**] proved that $\mathcal{M}$ is locally path connected, even by Hölder continuous paths. (For the classical Mandelbrot set, local connectivity is a famous open problem.) Bandt [**2**] has recently studied the set $\mathcal{M}$ as well; he asked several interesting questions motivated by computer pictures. In particular, Bandt asked whether the set $\mathcal{M}$ is the closure of its interior (apart from the two antennas). A partial result in this direction (that a "chunk" of $\mathcal{M}$ near the imaginary axis is the closure of its interior) was obtained in [**39**].

References

[1] J. C. Alexander, J. A. Yorke, Fat Baker's transformations, *Ergodic Th. Dynam. Sys.* **4** (1984), 1–23.

[2] C. Bandt, On the Mandelbrot set for pairs of linear maps, *Nonlinearity* **15** (2002), 1127–1147.

[3] M.F. Barnsley, *Fractals Everywhere*, Academic Press, 1988, 1993.

[4] M.F. Barnsley and A.N. Harrington, A Mandelbrot set for pairs of linear maps, *Physica* **15D** (1985), 421–432.

[5] M. J. Bertin, A. Decomps-Guilloux, M. Grandet-Hugot, M. Pathiaux-Delefosse, J. P. Shreiber, *Pisot and Salem Numbers.* Birkhäuser, 1992.

[6] T. Bousch, Connexité locale et par chemins höldériens pour les systèmes itérés de fonctions, *Preprint*, 1993.

[7] P. Diaconis, D. Freedman, Iterated random functions, *SIAM Review* **41**, No. 1, 45–76.

[8] I. Daubechies, J. Lagarias, Two-scale difference equations, I. Existence and global regularity of solutions, *SIAM J. Math. Anal.* **22** (1991), 1388-1410.

[9] G. Derfel, R. Schilling, Spatially chaotic configurations and functional equations, *J. Phys. A: Math. Gen.* **29** (1996), 4537–4547.

[10] P. Erdős, On a family of symmetric Bernoulli convolutions, *Amer. J. Math.* **61** (1939), 974–975.

[11] P. Erdős, On the smoothness properties of Bernoulli convolutions, *Amer. J. Math.* **62** (1940), 180–186.

[12] K. J. Falconer, The Hausdorff dimension of self-affine fractals, *Math. Proc. Camb. Phil. Soc.* **103** (1988), 339–350.

[13] K.J. Falconer, *Techniques in Fractal Geometry*, Wiley, 1997.

[14] A. M. Garsia, Arithmetic properties of Bernoulli convolutions, *Trans. Amer. Math. Soc.* **102** (1962), 409–432.

[15] J. E. Hutchinson, Fractals and self-similarity, *Indiana Univ. Math. J.* **30** (1981), 713–747.

[16] B. Jessen, A. Wintner, Distribution functions and the Riemann zeta function, *Trans. Amer. Math. Soc.* **38** (1935), 48–88.

[17] J.-P. Kahane, Sur la distribution de certaines séries aleatoires, *Colloque Th. Nombres [1969, Bordeaux], Bull. Soc. Math. France*, Mémoire **25** (1971), 119–122.

[18] M. Keane, M. Smorodinsky, B. Solomyak, On the morphology of γ–expansions with deleted digits, *Trans. Amer. Math. Soc.* **347**, no. 3 (1995), 955–966.

[19] F. Ledrappier, On the dimension of some graphs, *Contemporary Math.* **135** (1992), 285–293.

[20] P. Mattila, Orthogonal projections, Riesz capacities and Minkowski content, *Indiana Univ. Math. J.* **39** (1990), 185–198.

[21] P. Mattila, *Geometry of Sets and Measures in Euclidean Spaces*, Cambridge University Press, 1995

[22] P. A. P. Moran, Additive functions of intervals and Hausdorff measure, *Math. Proc. Cambridge Phil. Soc.* **42** (1946), 15–23.

[23] J. Neunhäuser, Properties of some overlapping self-similar and some self-affine measures, *Acta Math. Hungar.* **92** (2001), 143–161.

[24] S.-M. Ngai, Y. Wang, Self-similar measures associated with IFS with non-uniform contraction ratios, *Preprint.*

[25] J. Palis, J.-C. Yoccoz, Homoclinic tangencies for hyperbolic sets of large Hausdorff dimension, *Acta Math.* **172** (1994), 91–136.

[26] K. Petersen, *Ergodic Theory*, Cambridge University Press, 1983.

[27] Y. Peres, W. Schlag, Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions, *Duke Math. J.* **102** (2000), 193–251.

[28] Y. Peres, W. Schlag, B. Solomyak, Sixty years of Bernoulli convolutions, *Fractal Geometry and Stochastics* **II**, C. Bandt, S. Graf, and M. Zähle (editors), Progress in Probability Vol. **46**, 39–65, Birkhäuser, 2000.

[29] Y. Peres, K. Simon and B. Solomyak, Self-similar sets of zero Hausdorff measure and positive packing measure, *Israel J. Math.* **117** (2000), 353–379.

[30] Y. Peres, B. Solomyak, Absolute continuity of Bernoulli convolutions, a simple proof, *Math. Research Letters* **3**, no. 2 (1996), 2231–239.

[31] Y. Peres, B. Solomyak, Self–similar measures and intersections of Cantor sets, *Trans. Amer. Math. Soc.* **350**, no. 10 (1998), 4065–4087.

[32] M. Pollicott, K. Simon, The Hausdorff dimension of λ-expansions with deleted digits, *Trans. Amer. Math. Soc.* **347**, no. 3 (1995), 967–983.

[33] M. Pollicott, H. Weiss, The dimension of some self affine limit sets in the plane, *J. Stat. Phys.* **77** (1994), 841–866.

[34] F. Przytycki, M. Urbański, On Hausdorff dimension of some fractal sets, *Studia Math.* **93** (1989), 155–186.

[35] R. Salem, *Algebraic Numbers and Fourier Analysis*, Heath, 1963.

[36] B. Solomyak, Conjugates of β-numbers and the zero-free domain for a class of analytic functions, *Proc. London Math. Soc.* **68** (1994), 477–498.

[37] B. Solomyak, On the random series $\sum \pm\lambda^i$ (an Erdős problem), *Annals of Math.* **142** (1995), 611–625.

[38] B. Solomyak, Measure and dimension of some fractal families, *Math. Proc. Cambridge Phil. Soc.* **124**, no. 3 (1998), 531–546.

[39] B. Solomyak, H. Xu, On the "Mandelbrot set" for a pair of linear maps and complex Bernoulli convolutions, *Preprint,* 2002.

[40] R. S. Strichartz, Self-similarity in harmonic analysis, *J. Fourier Anal. Applic.* **1** (1994), 1–37.

Department of Mathematics, University of Washington, Seattle, WA 98195-4350
E-mail address: `solomyak@math.washington.edu`