SIXTY YEARS OF BERNOULLI CONVOLUTIONS

YUVAL PERES, WILHELM SCHLAG, AND BORIS SOLOMYAK

ABSTRACT. The distribution ν_λ of the random series $\sum \pm\lambda^n$ is the infinite convolution product of $\frac{1}{2}(\delta_{-\lambda^n} + \delta_{\lambda^n})$. These measures have been studied since the 1930's, revealing connections with harmonic analysis, the theory of algebraic numbers, dynamical systems, and Hausdorff dimension estimation. In this survey we describe some of these connections, and the progress that has been made so far on the fundamental open problem: For which $\lambda \in (\frac{1}{2}, 1)$ is ν_λ absolutely continuous?

Our main goal is to present an exposition of results obtained by Erdős, Kahane and the authors on this problem. Several related unsolved problems are collected at the end of the paper.

TABLE OF CONTENTS

1. **Introduction**

Let ν_λ be the distribution of $\sum_0^\infty \pm\lambda^n$ where the signs are chosen independently with probability $\frac{1}{2}$. It is the infinite convolution product of $\frac{1}{2}(\delta_{-\lambda^n} + \delta_{\lambda^n})$, hence the term "infinite Bernoulli convolution". These measures have been studied since the 1930's, revealing surprising connections with harmonic analysis, the theory of algebraic numbers, dynamical systems, and Hausdorff dimension estimation. There are several ways to think about ν_λ which hint at these connections.

Research of Peres was partially supported by NSF grant #DMS-9803597. Research of Solomyak was supported in part by NSF grant #DMS 9800786, the Fulbright foundation, and the Institute of Mathematics at the Hebrew University of Jerusalem.

(i) The Fourier transform $\widehat{\nu_\lambda}(\xi) = \int_{-\infty}^{\infty} e^{it\xi}\, d\nu_\lambda(t)$ is easily computed:

$$\widehat{\nu_\lambda}(\xi) = \prod_{n=0}^{\infty} \cos(\lambda^n \xi), \tag{1.1}$$

and this formula has been crucial for number-theoretic considerations.

(ii) ν_λ can be characterized by the functional equation for its cumulative distribution function $F_\lambda(x) = \nu_\lambda(-\infty, x]$:

$$F_\lambda(x) = \frac{1}{2}\Big[F_\lambda\Big(\frac{x-1}{\lambda}\Big) + F_\lambda\Big(\frac{x+1}{\lambda}\Big)\Big].$$

In other words, ν_λ is the self-similar measure for the iterated function system $\{\lambda x - 1, \lambda x + 1\}$ with probabilities $(\frac{1}{2}, \frac{1}{2})$ (see [17]). This point of view is useful in applications to dynamical systems and dimension estimation.

(iii) ν_λ can be viewed as a "non-linear projection": let $\Omega = \{-1, 1\}^{\mathbb{N}}$ be the sequence space with the Bernoulli measure $\mu = (\frac{1}{2}, \frac{1}{2})^{\mathbb{N}}$. Then

$$\nu_\lambda = \mu \circ \Pi_\lambda^{-1} \qquad \text{where} \qquad \Pi_\lambda(\omega) = \sum_{n=0}^{\infty} \omega_n \lambda^n.$$

This representation has been most useful in the recent work on ν_λ which used ideas of geometric measure theory.

The fundamental question about ν_λ is to decide for which $\lambda \in (\frac{1}{2}, 1)$ this measure is absolutely continuous and for which λ it is singular. If the density exists, it is natural to inquire about its smoothness. If the measure is singular, one would like to compute, or estimate, its dimension.

Denote by $S_\perp$ the set of $\lambda \in (\frac{1}{2}, 1)$ such that ν_λ is singular. The only elements of $S_\perp$ that are known were found in [9] by Erdős (1939): they are the reciprocals of Pisot numbers in $(1, 2)$. It is an open problem whether they constitute all of $S_\perp$. The first important result in the opposite direction is also due to Erdős (1940): he proved in [10] that $S_\perp \cap (a, 1)$ has zero Lebesgue measure for some $a < 1$. Later, Kahane [19] indicated that the argument of [10] actually implies that the Hausdorff dimension of $S_\perp \cap (a, 1)$ tends to 0 as $a \uparrow 1$. We included the Erdős-Kahane argument with explicit numerical bounds since they have never appeared in the literature (see section 6). In [47] Solomyak (1995) proved that ν_λ is absolutely continuous for a.e. $\lambda \in (\frac{1}{2}, 1)$. A simpler proof was found by Peres and Solomyak [39] and we do not reproduce it here. In [38] Peres and Schlag established, as a corollary of a more general result, that the Hausdorff dimension of $S_\perp \cap (a, 1)$ is less than one for any $a > \frac{1}{2}$. Our main goal is to give a self-contained proof of this (see section 7).

The rest of the paper is organized as follows. In section 2 we give some additional historical background. Sections 3 and 4 contain several results which hold for *all* parameters λ. In section 3 we discuss two "laws of pure type" for self similar measures, one classical and one recently proved by Mauldin and Simon [36]. Together, they imply that any self similar measure ν with support K is either singular to Lebesgue measure $\mathcal{L}$, or equivalent to the restriction of $\mathcal{L}$ to K.

In section 4 we discuss the Hausdorff and correlation dimensions of Bernoulli convolutions. We prove that the upper and lower correlation dimensions of ν_λ coincide for all λ and use this to show that the correlation dimension of ν_λ equals one on a residual set of $\lambda \in (\frac{1}{2}, 1)$; these appear to be new results. In section 5 we make some comments on Bernoulli convolutions for Salem numbers. The contents of sections 6 and 7 were described above. We conclude in section 8, with a brief discussion of some applications, generalizations and problems.

2. **Historical notes**

This section is by no means comprehensive; we just add a few remarks to what was said in the introduction. For instance, much of the early work on Bernoulli convolutions covers more general random variables $\sum \pm r_n$ for an arbitrary sequence r_n with $\sum r_n^2 < \infty$, but here we only discuss the case $r_n = \lambda^n$ for $\lambda \in (0, 1)$.

In the 1930's Bernoulli convolutions were studied by Wintner and his co-authors. Jessen and Wintner (1935) showed that ν_λ is either absolutely continuous, or purely singular, depending on λ (see §3). Kershner and Wintner (1935) observed that ν_λ is singular for $\lambda \in (0, \frac{1}{2})$ since it is supported on a Cantor set of zero Lebesgue measure (in fact, ν_λ is the standard Cantor-Lebesgue measure on this Cantor set). Wintner (1935) noted that ν_λ is uniform on $[-2, 2]$ for $\lambda = \frac{1}{2}$, and for $\lambda = 2^{-1/k}$ with $k \geq 2$ it is absolutely continuous, with a density in $C^{k-2}(\mathbb{R})$. For $\lambda \in (\frac{1}{2}, 1)$ the support of ν_λ is the interval $[-(1-\lambda)^{-1}, (1-\lambda)^{-1}]$, so one might surmise that ν_λ is absolutely continuous for all such λ. However, in [9] Erdős (1939) showed that ν_λ is singular when λ is the reciprocal of a Pisot number. Recall that a Pisot number is an algebraic integer all of whose conjugates are less than one in modulus. This gives a closed countable set of $\lambda \in (\frac{1}{2}, 1)$ with ν_λ singular. Curiously, there are only two Pisot numbers in $(1, 2^{1/2})$: the positive root $\theta_1 \sim 1.324718$ of $x^3 - x - 1 = 0$ and the positive root $\theta_2 \sim 1.3802777$ of $x^4 - x^3 - 1$. The golden ratio $\frac{1}{2}(1 + \sqrt{5})$ is the only quadratic Pisot number in $(1, 2)$, and it is also the smallest limit point of Pisot numbers, see [3]. No $\lambda \in S_\perp$, other than reciprocals of Pisot numbers, have been found.

The proof of Erdős [9] proceeds by showing that the Fourier Transform $\widehat{\nu_\lambda}(\xi)$ does not tend to zero at infinity if $\theta = \lambda^{-1}$ is Pisot. Salem [43], using Pisot's Theorem (see [44, p.11]), showed that for all $\lambda \in (0, 1)$ such that λ^{-1} is not a Pisot number, $\widehat{\nu_\lambda}(\xi)$ does tend to zero when $\xi \to \infty$. (I.e., ν_λ is a *Rajchman measure*, see [31].) For $\lambda \in (0, \frac{1}{2})$, this result is related to the fact that a Cantor set with dissection ratio λ is a set of uniqueness for Fourier series if and only if λ^{-1} is a Pisot number, see Salem [44].

Kahane and Salem (1958) obtained criteria for Bernoulli convolutions to have a density in L^2. Although they could not apply these criteria to ν_λ, they analyzed the distributions of certain series $\pm r_n$ where the ratios r_n/r_{n-1} are random.

Garsia (1962) found the largest explicitly given set of λ known to date, for which ν_λ is absolutely continuous (and even 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 $\max\{p, n\} \geq 2$.) Garsia showed that for these λ, all 2^n sums $\sum_0^{n-1} \pm\lambda^k$ are distinct and at least $C2^{-n}$ apart for some $C > 0$. This implies that ν_λ has a bounded density.

Starting with Garsia (1963), many authors studied the measure ν_λ in the Pisot case, computing or estimating various dimensions and giving alternative proofs of singularity. This line of research is not the focus of our paper, and we only discuss it briefly in section 8.

The interest in Bernoulli convolutions was renewed in the 1980's when their importance in various problems of dynamics and dimension was discovered by Alexander and Yorke (1982), Przytycki and Urbański (1989), and Ledrappier (1992) (see section 8 for details)..

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 90's (see [22]). 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 (1995) proved that the Hausdorff dimension equals the similarity dimension 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. A crucial tool in their paper was the notion of *transversality* for power series. This notion turned out to be crucial in all the recent work on Bernoulli convolutions [47, 39, 40, 38]; we discuss (a version of) transversality in section 7. 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.

3. **Laws of pure type**

Jessen and Wintner (1935) showed that any convergent infinite convolution of discrete measures is of *pure type:* it is either singular or absolutely continuous with respect to Lebesgue measure. In particular, this applies to ν_λ for any $\lambda < 1$. The purity of ν_λ can also be obtained from its self similarity, as the following proposition demonstrates.

Proposition 3.1. *Suppose that ν is a self similar probability measure on $\mathbb{R}^d$ corresponding to the contracting similitudes $\{S_j : j = 1, \ldots \ell\}$ and the (positive) probabilities $\{p_j : j = 1, \ldots \ell\}$*, i.e.,

$$\nu = \sum_{j=1}^{\ell} p_j \left(\nu \circ S_j^{-1}\right). \tag{3.1}$$

Let K denote the closed support of ν. If ν is not singular to Lebesgue measure $\mathcal{L}$ on $\mathbb{R}^d$, then

- **(i)** *ν is absolutely continuous with respect to $\mathcal{L}$.*
- **(ii)** *The restriction $\mathcal{L}_K$ of $\mathcal{L}$ to K is absolutely continuous with respect to ν.*

Part (i) of this proposition is folklore; part (ii) was proved by Mauldin and Simon [36] for ν_λ, and we extend their proof to general self similar measures below.

Proof: **(i)** Apply the similitude S_j to the Lebesgue decomposition $\nu = \nu_{\text{ac}} + \nu_{\text{s}}$ of ν into an absolutely continuous part and a singular part. Averaging the result over j with weights p_j, we infer that ν_{ac} and ν_{s} also satisfy (3.1). Since ν is the only probability measure satisfying (3.1) (see [17]), it follows that ν_{ac} and ν_{s} are proportional, hence one of them must vanish.

(ii) Since ν is not singular to $\mathcal{L}$, necessarily $\mathcal{L}(K) > 0$ and for some $\beta < 1$,

$$\sup\left\{\mathcal{L}(A) \,\Big|\, A \text{ Borel, } A \subset K,\ \nu(A) = 0\right\} = \beta \mathcal{L}(K) \tag{3.2}$$

Let A_0 be a Borel subset of K such that $\nu(A_0) = 0$. Denote by c_* the minimal contraction ratio for the maps $\{S_j\}_{j=1}^{\ell}$. Fix $x \in K$ and $r \in (0, \text{diam} K)$. Since $K = \cup_{j=1}^{\ell} S_j(K)$, there exist similitudes S of the form $S = S_{i_1} \circ S_{i_2} \circ \ldots \circ S_{i_m}$ such that $S(K)$ is contained in the open ball $B(x, r)$. Choose such an S with m minimal; clearly $\text{diam} S(K) \ge c_* r$, and therefore $\mathcal{L}[S(K)] \ge \eta \mathcal{L}[B(x,r)]$, where $\eta > 0$ does not depend on x and r (we can take $\eta = c_*^d \mathcal{L}[K]/\mathcal{L}[B(0, \text{diam} K)]$).

The preimage $S^{-1}\left(A_0 \cap S(K)\right)$ is a Borel subset of K, and ν assigns it measure zero, since self similarity of ν implies that $\nu \circ S^{-1}$ is dominated by a constant multiple of ν. By (3.2) and scaling, $\mathcal{L}[A_0 \cap S(K)] \le \beta \mathcal{L}[S(K)]$ and consequently

$$\mathcal{L}\Big[B(x,r) \setminus A_0\Big] \ge (1-\beta)\mathcal{L}[S(K)] \ge (1-\beta)\eta \mathcal{L}[B(x,r)].$$

Thus A_0 cannot have a Lebesgue density point, whence $\mathcal{L}(A_0) = 0$. $\square$

Remark: Suppose that ν satisfies (3.1) and let K be the support of ν. Consider the lower derivative

$$\underline{D}(\nu, x) := \liminf_{r \downarrow 0} (2r)^{-d} \nu[B(x,r)],$$

and denote by A_0 the set of $x \in K$ where $\underline{D}(\nu, x) = 0$. If $K \setminus A_0$ has nonempty interior then a variant of the above argument shows that $\dim(A_0) < 1$.

4. Dimensions of Bernoulli convolutions

Recall that the Hausdorff dimension of a Borel measure ν on $\mathbb{R}$ is defined by

$$\dim \nu = \inf\{\dim A : \ A \text{ Borel}, \ \nu(\mathbb{R} \setminus A) = 0\}.$$

The *correlation dimension* of ν is defined by

$$C_2(\nu) = \lim_{r \to 0} \frac{\log(\nu \times \nu)\{(x,y) : \ |x-y| \le r\}}{\log r}, \tag{4.1}$$

if the limit exists; otherwise, one considers the upper and lower correlation dimensions, obtained by taking lim sup and lim inf.

Proposition 4.1. **(i)** *For any $\lambda \in (0,1)$ the limit in (4.1) with $\nu = \nu_\lambda$ exists.*

(ii) *The set $\{\lambda \in (\frac{1}{2}, 1) : \ C_2(\nu_\lambda) = 1\}$ is a dense G_δ set in $(\frac{1}{2}, 1)$.*

(iii) *The set $\{\lambda \in (\frac{1}{2}, 1) : \ \dim \nu_\lambda = 1\}$ is residual in $(\frac{1}{2}, 1)$ (it contains a dense G_δ set).*

Remarks. **1.** Part (i) of the proposition appears to be new for $\lambda \in (\frac{1}{2}, 1)$; in the literature that we are aware of, this limit is only proved to exist under some separation conditions. We can prove a similar result in much greater generality (for arbitrary self-similar measures and quite general self-conformal measures, with or without overlaps) but we do not describe this here.

2. Part (iii) is immediate from part (ii) since $\dim \nu$ is at least the lower correlation dimension of ν for any measure ν (see, e.g., [45]).

3. Of course, to prove (ii) we only have to show that the set under consideration is a G_δ set, since we already know that ν_λ has a density in $L^2(\mathbb{R})$ for a.e. $\lambda \in (\frac{1}{2}, 1)$.

4. Ledrappier [29] indicated a proof of the statement that the set in (iii) is a G_δ set, based on the fact that every ν_λ is "exact-dimensional", which in turn relies on the Ledrappier-Young theory.

Proof of Proposition 4.1. We begin with the proof of the implication (i) $\Rightarrow$ (ii) since it is easier. Observe that for any positive r and t the following set is open:

$$\{\lambda : \ (\nu_\lambda \times \nu_\lambda)\{(x,y) : \ |x-y| \le r\} < t\} = \{\lambda : \ (\mu \times \mu)\{(\omega, \tau) : \ |\Pi_\lambda(\omega - \tau)| \le r\} < t\}$$

where we denote $\Pi_\lambda(\omega - \tau) = \sum_{j=0}^\infty (\omega_j - \tau_j)\lambda^j$. Let $\varepsilon_m \to 0$. We have

$$\{\lambda : \ C_2(\nu_\lambda) = 1\} = \bigcap_m \bigcap_{n=1}^\infty \bigcup_{k=n}^\infty \Big\{\lambda : \ \frac{\log(\nu_\lambda \times \nu_\lambda)\{(x,y) : \ |x-y| \le 2^{-k}\}}{-k \log 2} > 1 - \varepsilon_m\Big\}$$

which is a G_δ set, and the claim follows.

(i) Let $\phi_r(x) = \max\{0, 1 - \frac{|x|}{r}\}$ be the triangular kernel and define

$$a_n = \int\int \phi_{\lambda^n}(y-x)\, d\nu_\lambda(x) d\nu_\lambda(y) = \int_\Omega \int_\Omega \phi_{\lambda^n}(\Pi_\lambda(\omega - \tau))\, d\mu(\omega) d\mu(\tau). \tag{4.2}$$

Since $\frac{1}{2}\cdot\mathbf{1}_{[-r/2,r/2]} \le \phi_r \le \mathbf{1}_{[-r,r]}$, it is enough to prove that the limit $\lim \frac{1}{n}\log a_n$ exists. Thus, it suffices to show that for some C and k we have

$$a_{m+n} \le C a_m a_{n-k} \quad \text{for} \ \ m \ge 1,\ n \ge k+1$$

(then $b_n = \log(C a_{n-k})$ is sub-additive hence $\lim(b_n/n)$ exists). Denote $\omega' = \sigma^n\omega$ and $\tau' = \sigma^n\tau$ where σ is the left shift on Ω. We have

$$a_{m+n} = \sum_{\substack{\omega_0,\dots,\omega_{n-1}\\ \tau_0,\dots,\tau_{n-1}}} 2^{-2n}\int_\Omega\int_\Omega \phi_{\lambda^{m+n}}\Big[\sum_{j=0}^{n-1}(\omega_j-\tau_j)\lambda^j + \lambda^n\Pi_\lambda(\omega'-\tau')\Big]\,d\mu(\omega')d\mu(\tau')\,. \tag{4.3}$$

The integral above equals

$$\iint \phi_{\lambda^m}(y-x+c)\,d\nu_\lambda(x)d\nu_\lambda(y) = \int \widehat{\phi}_{\lambda^n}(\xi)e^{-ic\xi}|\widehat{\nu}_\lambda(\xi)|^2\,d\xi \tag{4.4}$$

where $c = \lambda^{-n}\sum_{j=0}^{n-1}(\omega_j-\tau_j)\lambda^j$, and we have invoked Plancherel's theorem. Since $\widehat{\phi}_r(\xi) > 0$, the integral (4.4) is maximized for $c=0$, and then it equals a_m, see (4.2). On the other hand, $\phi_{\lambda^m}(y-x+c)\equiv 0$ on the support of $\nu_\lambda\times\nu_\lambda$ if $|c| > C_1 = \frac{3}{1-\lambda}$. Combining the last two observations with (4.3) we obtain

$$\begin{aligned} a_{m+n} &\le a_m\cdot \mathrm{Prob}\Big\{\big|\sum_{j=0}^{n-1}(\omega_j-\tau_j)\lambda^j\big| \le C_1\lambda^n\Big\} \\ &\le a_m\cdot \mathrm{Prob}\Big\{|\Pi_\lambda(\omega-\tau)| \le 2C_1\lambda^n\Big\} \\ &\le a_m\cdot 2\int_\Omega\int_\Omega \phi_{4C_1\lambda^n}(\Pi_\lambda(\omega-\tau))\,d\mu(\omega)d\mu(\tau) \\ &\le 2a_m a_{n-k} \end{aligned}$$

where k is such that $\lambda^{-k} > 4C_1$. The proof is complete. □

The investigation of dimension(s) of ν_λ for particular λ essentially goes back to Garsia (1963). In [15] he considered $H_N(\lambda)$, the entropy of the distribution of the random sum $\sum_{n=0}^N \pm\lambda^n$, and the limit $G_\lambda = \lim_{N\to\infty}\frac{H_N(\lambda)}{N+1}$ which he showed exists. If there are no coincidences among the finite sums, then $G_\lambda = \log 2$. If, on the other hand, there are coincidences, then $G_\lambda < \log 2$. Garsia proved that if $G_\lambda < \log(\frac{1}{\lambda})$ then ν_λ is singular, and this inequality holds for Pisot λ^{-1}. Alexander and Yorke [1] proved that $G_\lambda/\log(\frac{1}{\lambda})$ is always an upper bound for the Rényi (information) dimension of ν_λ, and equality holds in the Pisot case. In fact,

$$\lambda^{-1} \ \text{ is Pisot} \quad\Rightarrow\quad \dim\nu_\lambda = G_\lambda/\log(1/\lambda). \tag{4.5}$$

As observed by Ledrappier and Porzio [30], this follows from ν_λ being "exact-dimensional" and [53]. A direct proof of (4.5) was given by Lalley [25].

In several papers numerical estimates of $\dim \nu_\lambda$ were pursued, especially for the golden ratio case $\lambda_g = \frac{\sqrt{5}-1}{2}$. Although the formula (4.5) looks simple, it is inefficient to use it directly for such estimates. Alexander and Zagier [2] found a formula for $\dim \nu_{\lambda_g}$ by analyzing the "Fibonacci graph", and used it to show that $0.99557 < \dim \nu_{\lambda_g} < 0.99574$. Recently Sidorov and Vershik [46] gave another proof of the Alexander-Zagier formula relating it to the entropy of the random walk on the Fibonacci graph. (They also gave a nice ergodic-theoretic proof of singularity of ν_{λ_g}.) On the other hand, Ledrappier and Porzio [30] and independently, Lalley [25], expressed $\dim \nu_{\lambda_g}$ as the top Lyapunov exponent of certain random matrix products; Lalley covered the general case of Pisot numbers and biased Bernoulli convolutions.

We should also mention the paper by Bovier [5] who gave yet another proof of singularity in the golden mean case using automata theory. Lau [26] and Lau and Ngai [27, 28] computed the L^q-spectrum of Bernoulli convolutions for the golden ratio and other Pisot numbers. The spectrum of local dimensions was investigated by Hu [16].

5. **Bernoulli convolutions and Salem numbers**

Recall that an algebraic integer $\theta > 1$ is a *Salem number* if its Galois conjugates satisfy $|\theta_j| \le 1$ and at least one of the conjugates has modulus equal to one (i.e. θ is not Pisot). The set of Salem numbers is rather poorly understood. In particular, the following is open:

Problem: is there $b > 1$ such that every Salem number is greater than b?

This is related to the well-known Lehmer problem on the range of Mahler measure for integer polynomials, see [7].

Below we show that obtaining a topological analog of "almost sure" results for Bernoulli convolutions (such as Corollary 6.2(ii) below) would settle the problem on Salem numbers. Throughout, fractional derivatives will be expressed in terms of the standard $2,\gamma$–Sobolev space L^2_γ which is defined by the norm $\|\nu_\lambda\|^2_{2,\gamma} = \int_{-\infty}^{\infty} |\widehat{\nu_\lambda}(\xi)|^2 |\xi|^{2\gamma}\, d\xi$.

Proposition 5.1. *If there exist $\gamma > 0$ and $a < 1$ such that the set $\{\lambda \in (a,1) :\ \nu_\lambda \in L^2_\gamma\}$ is residual in $(a,1)$, then Salem numbers do not accumulate to one.*

The proof is based on several easy lemmas.

Lemma 5.2. *Let θ be a Salem number and $\lambda = \theta^{-1}$. Then*

$$\limsup_{\xi\to\infty} |\widehat{\nu_\lambda}(\xi)|\, |\xi|^{\varepsilon} = \infty \quad \text{for all } \ \varepsilon > 0. \tag{5.1}$$

Proof. Let $||x||$ denote the distance of $x \in \mathbb{R}$ to the nearest integer. It has been observed by several authors (see [6], [37, 6.9], [3, 5.5.1]) that for each Salem number θ and any $\delta > 0$ one can find $t \ge 1$

such that

$$||t\theta^n|| \le \delta \quad \text{for all} \quad n \ge 1. \tag{5.2}$$

We have by (1.1) and (5.2) for $\lambda = \theta^{-1}$:

$$\begin{aligned} |\widehat{\nu_\lambda}(\pi t\theta^n)| &\ge |\widehat{\nu_\lambda}(\pi t)| \prod_{k=1}^{n} \cos(\pi||t\theta^k||) \\ &\ge |\widehat{\nu_\lambda}(\pi t)|(1-c\delta^2)^n \ge c'|\widehat{\nu_\lambda}(\pi t)|(\pi t\theta^n)^{-\varepsilon}, \end{aligned}$$

where $\varepsilon = -\frac{\log(1-c\delta^2)}{\log\theta}$. Since $\varepsilon \to 0$ as $\delta \to 0$, it remains to show that $\widehat{\nu_\lambda}(\pi t) \ne 0$.

Suppose $\widehat{\nu_\lambda}(\pi t) = 0$, then $t = \frac{2k+1}{2}\theta^l$ for some $l \ge 0$. But $\{\theta^n\}_{n\ge1}$ is dense mod 1 (see [44]), hence $t\theta^n$ is dense mod $\frac{1}{2}$ which contradicts (5.2) for $\delta < \frac{1}{2}$. □

Lemma 5.3. *Suppose that $f \in L^2(\mathbb{R})$ has compact support and $\|f\|_{2,\gamma}^2 = \int_{-\infty}^{\infty} |\hat{f}(\xi)|^2|\xi|^{2\gamma}\, d\xi < \infty$. Then*

$$|\hat{f}(\xi)| = o(|\xi|^{-\gamma}), \quad |\xi| \to \infty.$$

Proof. This is a standard fact from harmonic analysis but we provide a proof for the reader's convenience. Let $\tilde{f}(x) = f(-x)$ and let ϕ be a smooth compactly supported function on $\mathbb{R}$ equal to one on the support of $f * \tilde{f}$. Then $f * \tilde{f} = (f * \tilde{f})\phi$ hence

$$\begin{aligned} |\hat{f}(\xi)|^2 = (|\hat{f}|^2 * \hat{\phi})(\xi) &= \int_{\mathbb{R}} |\hat{f}(\xi-\eta)|^2 \hat{\phi}(\eta)\, d\eta \\ &\le C\int_{|\eta|<|\xi/2|} |\hat{f}(\xi-\eta)|^2\, d\eta + C\int_{|\eta|\ge|\xi/2|} |\hat{\phi}(\eta)|\, d\eta. \end{aligned} \tag{5.3}$$

Since $|\hat{\phi}(\eta)|$ is rapidly decreasing, the second integral is $O(|\xi|^{-q})$ for any $q > 0$. The first integral can be estimated above by

$$|\xi/2|^{-2\gamma} \int_{|\eta|<|\xi/2|} |\hat{f}(\xi-\eta)|^2|\xi-\eta|^{2\gamma}\, d\eta \le |\xi/2|^{-2\gamma} \int_{|\eta|\ge|\xi/2|} |\hat{f}(\eta)|^2|\eta|^{2\gamma}\, d\eta = o(|\xi|^{-2\gamma}),$$

and the claim follows. □

Lemma 5.4. *Let Γ be the set of $\lambda \in (0,1)$ for which there exists $\varepsilon > 0$ such that $\widehat{\nu_\lambda}(\xi) = O(|\xi|^{-\varepsilon})$, as $|\xi| \to \infty$. Then $(0,1) \setminus \Gamma$ is a G_δ set.*

Proof. Let $\varepsilon_j \to 0$ for $j \in \mathbb{N}$. It is enough to observe that

$$\Gamma = \bigcup_j \bigcup_{k\ge1} \bigcap_{|\xi|\ge k} \{\lambda : \ |\widehat{\nu_\lambda}(\xi)| \le |\xi|^{-\varepsilon_j}\}$$

is an F_σ set. □

Proof of Proposition 5.1. By Lemma 5.2, reciprocals of Salem numbers belong to $(0,1)\setminus\Gamma$ where Γ is defined in Lemma 5.4. Any integer power of a Salem number is a Salem number, by definition. It is easy to see that for any sequence $x_n \downarrow 1$ the union of its (integer) powers is dense in $[1,+\infty)$. Thus, if Salem numbers accumulate to 1, the set $(0,1)\setminus\Gamma$ is dense G_δ in $(0,1)$. If ν_λ is generically in L^2_γ for $\lambda\in(a,1)$ for some $a<1$, then using Lemma 5.3 we see that Γ is residual in $(a,1)$, a contradiction. □

Remark. The statement (5.1) is contained in [19] but with a typo, claiming that it holds with lim rather than lim sup. As a consequence of this typo, the statements in [8] concerning Salem numbers are unjustified.

In fact, for *all* $\lambda\in(0,1)$ and any $\varepsilon>0$ we have

$$\liminf_{|\xi|\to\infty} |\widehat{\nu_\lambda}(\xi)|\cdot|\xi|^{\frac{\log 2}{-\log\lambda}-\varepsilon} = 0. \tag{5.4}$$

Indeed, let $\theta=\lambda^{-1}$. By Koksma's theorem, for a.e. $t>0$ the sequence $\{\theta^n t\}_{n\ge 1}$ is uniformly distributed mod 1 (see [24, Cor. 1.4.3]). Fix such a t. We have by (1.1)

$$\frac{1}{n}\log|\widehat{\nu_\lambda}(\pi t\theta^n)| \le \frac{1}{n}\log\prod_{k=1}^{n}|\cos(\pi t\theta^k)| \to \int_0^1 \log|\cos(\pi u)|\,du = -\log 2,$$

by the definition of the uniform distribution. This clearly implies (5.4).

6. Close to one; the Erdős-Kahane argument

In [10] Erdős proved that ν_λ is absolutely continuous for a.e. λ sufficiently close to one. However, explicit bounds for the neighborhood of one were not given. Kahane [19] gave a brief outline of the argument and indicated that it actually yields that the dimension of $\{\lambda\in(\lambda_0,1):\ \nu_\lambda \text{ is singular }\}$ tends to zero as $\lambda_0\uparrow 1$. Below we give an exposition of this argument since it remains the only way to prove the statement, while Kahane's paper [19] is not widely known and is tersely written. We also give explicit numerical bounds for the neighborhoods where the statements hold.

Proposition 6.1. *Let $1<a<b<\infty$. Fix $k\ge 3$ and define*

$$r=\frac{1}{2}(b+1)^{-2},\qquad A=1+(b+1)^2.$$

Suppose that

$$B<\frac{-\log[\cos(\pi r)]}{\log b}.$$

Then

$$\dim\{\lambda\in[b^{-1},a^{-1}]:\ \widehat{\nu_\lambda}(u)\ne O(u^{-B/k})\}\le\frac{\log[eA^3k]}{k\log a}. \tag{6.1}$$

Corollary 6.2. **(i)** *For any $s > 0$ there exists $\alpha(s) < 1$ such that*

$$\dim\{\lambda \in (\alpha(s), 1) : \ \nu_\lambda \text{ is singular}\} \leq s.$$

(ii) *For any $k \in \mathbb{N}$ and any $s > 0$ there exists $\alpha(k,s) < 1$ such that*

$$\dim\{\lambda \in (\alpha(k,s), 1) : \ \frac{d\nu_\lambda}{dx} \notin C^k(\mathbb{R})\} \leq s.$$

The corollary is immediate from the proposition by the formula $\widehat{\nu_\lambda}(u) = \widehat{\nu_{\lambda^2}}(u)\widehat{\nu_{\lambda^2}}(\lambda u)$ which implies

$$\dim\{\lambda \in [b^{-2^{-m}}, a^{-2^{-m}}] : \ \widehat{\nu_\lambda}(u) \neq O(u^{-2^m B/k})\} \leq \frac{\log[eA^3k]}{k\log a}.$$

To get a concrete numerical estimate, take $a = 2^{\frac{1}{2}}$ and $b = 2$. Then $\frac{\log[eA^3k]}{k\log a} < 1$ for $k = 34$ and $\frac{-\log[\cos(\pi r)]}{k\log b} > 0.0006$, so Proposition 6.1 implies

$$\dim\{\lambda \in [2^{-1}, 2^{-1/2}] : \ \widehat{\nu_\lambda}(u) \neq O(u^{-0.0006})\} < 1,$$

hence

$$\dim\{\lambda \in [2^{-2^{-10}}, 2^{-2^{-11}}] : \ \widehat{\nu_\lambda}(u) \neq O(u^{-0.6})\} < 1.$$

Therefore, by this argument ν_λ has a density in $L^2(\mathbb{R})$ for all $\lambda \in [2^{-2^{-10}}, 1) \supset [0.99933, 1)$ outside a set of dimension less than one. This can be improved somewhat by optimizing the choice of a and b but not very significantly.

Proof of Proposition 6.1. Denote $\theta = \lambda^{-1}$. From (1.1) we have

$$\widehat{\nu_\lambda}(\pi\theta^N t) = \prod_{n=1}^{N} \cos(\pi\theta^n t)\widehat{\nu_\lambda}(\pi t). \tag{6.2}$$

Let $\theta^n t = c_n + \varepsilon_n$ where $c_n \in \mathbb{N}$ and $\varepsilon_n \in [-\frac{1}{2}, \frac{1}{2})$ (the dependence on θ and t is not written explicitly but should be kept in mind). By assumption, we can fix $\delta > 0$ so that

$$\rho := [2(b+1)(b+1+\delta)]^{-1} \quad \text{satisfies} \quad B \leq \frac{-\log[\cos(\pi\rho)]}{\log b}. \tag{6.3}$$

Fix also $1 < a < b$ and $k \geq 3$ and denote by E_N the set of $\theta \in [a, b]$ such that there exists $t \in [1, \theta)$ for which

$$\mathrm{card}\{n \in [1, N] : \ |\varepsilon_n| \geq \rho\} \leq \frac{N}{k}.$$

Further, let $E = \limsup E_N$. Since by (6.2)

$$|\widehat{\nu_\lambda}(\pi\theta^N t)| \leq \prod_{n=1}^{N} |\cos(\pi\varepsilon_n)|,$$

it is immediate that

$$\theta \notin E \ \Rightarrow \ \widehat{\nu_\lambda}(u) = O(u^{-\gamma}),$$

where

$$\gamma = \frac{-\log[\cos(\pi\rho)]}{k\log\theta} \geq \frac{-\log[\cos(\pi\rho)]}{k\log b} \geq \frac{B}{k}\,.$$

Thus, to prove (6.1) one only needs to estimate $\dim E$. This is done with a beautiful argument inspired by a theorem of Pisot (see [44]).

Observe that

$$|\theta - \frac{c_{n+1}}{c_n}| = |\frac{\theta\varepsilon_n - \varepsilon_{n+1}}{c_n}| \leq \frac{b|\varepsilon_n| + |\varepsilon_{n+1}|}{c_n} \leq \frac{const}{a^n}\,. \tag{6.4}$$

We are going to cover E_N by intervals of size $\sim a^{-N}$ centered at $\frac{c_N}{c_{N-1}}$, so we want to estimate the number of possible pairs (c_{N-1}, c_N) corresponding to $\theta \in E_N$ (and some t). In fact, we will estimate the number of sequences $c_1, \dots, c_N$ with the help of the following lemma.

Lemma 6.3. *The following holds for n sufficiently large ($n \geq n_0(a, b, \delta)$).*

(i) *Given c_n, c_{n+1} there are at most $A' := 1 + (b+1)(b+1+\delta)$ possibilities for c_{n+2}, independent of $\theta \in [a, b]$ and $t \in [1, \theta)$.*

(ii) *If*

$$\max\{|\varepsilon_n|, |\varepsilon_{n+1}|, |\varepsilon_{n+2}|\} < \rho = \frac{1}{2(b+1)(b+1+\delta)}\,,$$

then c_{n+2} is uniquely determined by c_n and c_{n+1}, independent of $\theta \in [a, b]$ and $t \in [1, \theta)$.

Proof of the lemma. It is easy to see that that $\frac{c_{n+1}}{c_n} \leq \theta + \delta \leq b + \delta$ for n sufficiently large (depending on a and δ). Using this together with (6.4) we obtain

$$\begin{aligned} |c_{n+2} - \frac{c_{n+1}^2}{c_n}| &\leq c_{n+1}\Big(|\theta - \frac{c_{n+2}}{c_{n+1}}| + |\theta - \frac{c_{n+1}}{c_n}|\Big) \\ &\leq b|\varepsilon_{n+1}| + |\varepsilon_{n+2}| + (c_{n+1}/c_n)(b|\varepsilon_n| + |\varepsilon_{n+1}|) \\ &\leq (b+1)(b+1+\delta)\max\{|\varepsilon_n|, |\varepsilon_{n+1}|, |\varepsilon_{n+2}|\}\,. \end{aligned}$$

Now both (i) and (ii) are immediate since $c_{n+2} \in \mathbb{N}$. $\square$

Proof of Proposition 6.1 concluded. For $\Gamma \subset [1, N] \cap \mathbb{N}$ consider those $\theta \in [a, b]$ for which there exists $t \in [1, \theta)$ such that $|\varepsilon_n| < \rho$ for $n \in [1, N] \setminus \Gamma$. It follows from Lemma 6.3 that the number of sequences $c_1, \dots, c_N$ corresponding to such θ is bounded above by $C_{a,b,\delta}(A')^{3\,\mathrm{card}(\Gamma)}$. Thus, the number of sequences $c_1, \dots, c_N$ corresponding to E_N does not exceed

$$\mathrm{const}\begin{pmatrix} N \\ N/k \end{pmatrix} \cdot (A')^{3N/k}$$

(dealing with the possible non-integrality of N/k is left to the reader). By (6.4), this is the number of intervals of size $const \cdot a^{-N}$ needed to cover E_N. Therefore,

$$\dim E \le \lim_{N\to\infty} \frac{\log\left[\begin{pmatrix} N \\ N/k \end{pmatrix} (A')^{3N/k}\right]}{N\log a} \le \frac{\log[e(A')^3 k]}{k\log a},$$

as one can easily deduce using Stirling's formula. To complete the proof it remains to notice that $A' = A'(\delta) \to A$ as $\delta \to 0$. □

7. Smoothness and the dimension of exceptions

Following [38], we show in this section that the density of ν_λ has fractional derivatives in L^2 for almost all $\lambda \in (\frac{1}{2}, 1)$ and we estimate the dimension of those λ so that ν_λ is singular with respect to Lebesgue measure. Throughout, fractional derivatives will be expressed in terms of the standard $2, \gamma$–Sobolev space L^2_γ which is defined by the norm $\|\nu_\lambda\|^2_{2,\gamma} = \int_{-\infty}^{\infty} |\widehat{\nu_\lambda}(\xi)|^2 |\xi|^{2\gamma}\, d\xi$. First we recall the definition of δ–transversality from [39].

Definition 7.1. Let $\delta > 0$. We say that $J \subset \mathbb{R}$ is an interval of δ–transversality for the class of power series

$$g(x) = 1 + \sum_{n=1}^{\infty} b_n x^n, \quad \text{with} \quad b_n \in \{-1, 0, 1\} \tag{7.1}$$

if $g(x) < \delta$ implies $g'(x) < -\delta$ for any $x \in J$.

A useful criterion for checking δ–transversality was found in [39]. A power series $h(x)$ is called a $(*)$–function if for some $k \ge 1$ and $a_k \in [-1, 1]$,

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

In [47] Solomyak showed that among all convex combinations of series of the form (7.1), the power series with the smallest double zero must be a $(*)$–function. The following lemma from [39] bypasses this fact and reduces the search for intervals of transversality to finding a suitable $(*)$–function.

Lemma 7.2. *Suppose that a $(*)$–function h satisfies*

$$h(x_0) > \delta \quad \text{and} \quad h'(x_0) < -\delta$$

for some $x_0 \in (0, 1)$ and $\delta \in (0, 1)$. Then Definition 7.1 is satisfied on $[0, x_0]$.

In [39] a particular $(*)$–function was found that satisfies $h(2^{-2/3}) > 0.07$ and $h(2^{-2/3}) < -0.09$, so transversality in the sense of 7.1 holds on $[0, 2^{-2/3}]$ by this lemma. On the other hand, in [47] Solomyak proved that there is a power series of the form (7.1) with a double zero at roughly 0.68, whereas $2^{-2/3} \simeq 0.63$. We will return to this issue below.

Following [39] and [38], we consider Bernoulli convolutions from the point of view of "projections" — in an appropriate sense. More precisely, let $\Omega = \{-1, +1\}^{\mathbb{N}}$ be equipped with the product measure $\mu = \prod_0^\infty (\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1)$. For any distinct $\omega, \tau \in \Omega$ we define

$$|\omega \wedge \tau| = \min\{i \geq 0 \,:\, \omega_i \neq \tau_i\}.$$

Fix some interval $J = [\lambda_0, \lambda_1] \subset (0,1)$ and define $\Pi : J \times \Omega \to \mathbb{R}$ via $\Pi_\lambda(\omega) = \sum_{n=0}^\infty \omega_n \lambda^n$. The metric on Ω (depending on J) is given by $d(\omega, \tau) = \lambda_1^{|\omega \wedge \tau|}$. By definition the distribution ν_λ is equal to $\nu_\lambda = \mu \circ \Pi_\lambda^{-1}$. The α–energy of μ is defined as $\mathcal{E}_\alpha(\mu) = \int_\Omega \int_\Omega \frac{d\mu(\omega_1) d\mu(\omega_2)}{d(\omega_1, \omega_2)^\alpha}$. One checks that $\mathcal{E}_\alpha(\mu) < \infty$ if and only if $\lambda_1^\alpha > \frac{1}{2}$. Here we address the following question: How much regularity does ν_λ inherit from μ for a typical value of λ? In [39] it was shown that ν_λ has an L^2–density for a.e. $\lambda > \frac{1}{2}$. This is based on the fact that $\mathcal{E}_1(\mu) < \infty$ for any compact $J \subset (\frac{1}{2}, 1)$. In [38], Peres and Schlag improved this statement in two ways. Firstly, they showed that $\nu_\lambda \in L^2_\gamma$ for a.e. $\lambda^{1+2\gamma} > \frac{1}{2}$ using that $\mathcal{E}_{1+2\gamma}(\mu) < \infty$ on intervals $J = [\lambda_0, \lambda_1]$ with $\lambda_0^{1+2\gamma} > \frac{1}{2}$. In fact, they proved that $\int_J \|\nu_\lambda\|_{2,\gamma}^2 \, d\lambda < \infty$. Secondly, they used this "mean derivative bound" on the densities to show that the Hausdorff dimension of the set of parameters $\lambda \in J$ for which $\widehat{\nu_\lambda} \notin L^2$ is strictly less than one. A rigorous formulation of these principles is given by the following theorem, which is a special case of Theorem 2.8 in [38] (see also section 5.1 in that paper).

Theorem 7.3. *Suppose $J = [\lambda_0, \lambda_0'] \subset (\frac{1}{2}, 1)$ is an interval of δ–transversality for the power series (7.1). Then*

$$\int_J \|\nu_\lambda\|_{2,\gamma}^2 \, d\lambda < \infty \quad \text{if} \quad \lambda_0^{1+2\gamma} > \frac{1}{2}. \tag{7.2}$$

Furthermore,

$$\dim \{\lambda \in J \,:\, d\nu_\lambda / dx \notin L^2(\mathbb{R})\} \leq 2 - \frac{\log 2}{\log \frac{1}{\lambda_0}}. \tag{7.3}$$

The relation between the dimension of a set in Euclidean space and that of a generic projection has been studied by many authors, see [21], [11], [34]. In these works the dimension of the exceptional parameters is typically estimated by averaging a suitable functional (e.g. energy) against a Frostman measure on the set of exceptional parameters, see [11]. This depends crucially on a simple relation between the Fourier transform of a measure and the Fourier transform of its projections. Such a relation is not available in the case of Bernoulli convolutions, and a new idea is required. In [38] this is accomplished by means of Lemma 7.4 below, which allows one to derive (7.3) from (7.2). The idea behind that lemma is as follows: Let $\{h_j\}_{j=0}^{\infty}$ be a family of nonnegative smooth functions on $[0,1]$ whose derivatives grow at most exponentially in j. If $\int_0^1 \sum_0^\infty R^j h_j(x)\,dx < \infty$, then one can bound the dimension of the set of $x \in [0,1]$ for which $\sum_0^\infty r^j h_j(x) = \infty$ for any $1 \le r < R$. To apply this lemma, we rely on the dyadic decomposition of frequency space. More precisely, let $h_j(\lambda) = 2^{-j}\int_{2^{j-1}\le|\xi|\le 2^j} |\widehat{\nu_\lambda}(\xi)|^2\,d\xi$. Then $\int_J \sum_{j=0}^\infty 2^{(1+2\gamma)j} h_j(\lambda)\,d\lambda$ is controlled by the square of the $2,\gamma$–Sobolev norm of ν_λ averaged in λ, whereas $\sum_{j=0}^\infty 2^j h_j(\lambda) = \infty$ characterizes those $\lambda \in J$ so that ν_λ does not have an L^2–density. In order to keep the presentation as transparent as possible, we establish only the weaker bound

$$\dim\{\lambda \in J \,:\, d\nu_\lambda/dx \notin L^2(\mathbb{R})\} \le \frac{3}{2} - \frac{\log 2}{2\log\frac{1}{\lambda_0}} \tag{7.4}$$

and then sketch briefly how (7.3) can be obtained. For more details concerning the full statement of Theorem 7.3 above, we refer the reader to [38], section 3.

The following lemma (essentially Lemma 3.1 in [38]) is the basic tool for bounding the Hausdorff dimension of the exceptional parameters.

Lemma 7.4. *Let $I \subset \mathbb{R}$ be a nonempty open interval and $N \in \mathbb{N}$. Suppose $\{h_j\}_0^\infty \in C^N(I)$ satisfy*

$$\sup_{j\ge 0} A^{-jn}\|h_j^{(n)}\|_\infty \le C_n \quad \text{for all } n \le N \text{ and } \quad \sup_{j\ge 0}\int_I R^j|h_j(\lambda)|\,d\lambda \le C_* < \infty\,, \tag{7.5}$$

where $A > 1$. Suppose that $R > r \ge 1$ satisfy $A^\alpha r^{\alpha/N} = \frac{R}{r} \le A r^{1/N}$ with $0 < \alpha \le 1$. Then

$$\dim\Big\{\lambda \in I \,:\, \sum_{j=0}^\infty r^j|h_j(\lambda)| = \infty\Big\} \le 1-\alpha\,. \tag{7.6}$$

Proof for $N = 1$. Define $E_j = \{\lambda \in I \,:\, |h_j(\lambda)| > j^{-2}r^{-j}\}$. Then

$$\Big\{\lambda \in I \,:\, \sum_{j=0}^\infty r^j|h_j(\lambda)| = \infty\Big\} \subset \limsup_{j\to\infty} E_j. \tag{7.7}$$

Let $s > 1-\alpha$. We will estimate the s–Hausdorff measure of $\limsup E_j$ by covering each E_j with intervals of side length $\simeq (rA)^{-j}$. The idea is that any point in E_j has a neighborhood of size $\simeq (rA)^{-j}$ on which $|h_j|$ is at least $Cj^{-2}r^{-j}$. More precisely, fix some j and let $\{I_{ij}\}_{j=1}^{M_j}$ be a covering of E_j by intervals of size $j^{-2}(rA)^{-j}(2C_1)^{-1}$. We can assume that all $\{I_{ij}\}_{j=1}^{M_j}$ are contained in I and no point of I is covered more than twice. Since $\sup_\lambda |h_j(\lambda+y)-h_j(\lambda)| \le C_1 A^j |y|$, it follows that

$$\bigcup_{i=1}^{M_j} I_{ij} \quad \subset \quad \{\lambda \in I \,:\, |h_j(\lambda)| > \frac{1}{2j^2} r^{-j}\}. \tag{7.8}$$

By Markov's inequality and assumption (7.5), we conclude from (7.8) that

$$M_j \le 8C_1 j^4 (rA)^j r^j \int_I |h_j(\lambda)|\, d\lambda \le 4C_1 C_* j^4 (rA)^j \big(\frac{r}{R}\big)^j . \tag{7.9}$$

Let $s > 1-\alpha$. In view of (7.9),

$$\mathcal{H}^s(\limsup E_j) \quad \le \quad \lim_{k\to\infty} \sum_{j=k}^{\infty}\sum_{i=1}^{M_j} |I_{ij}|^s \le \lim_{k\to\infty} \sum_{j=k}^{\infty} C\, j^4 (rA)^j \big(\frac{r}{R}\big)^j \big(j^{-2}(rA)^{-j}\big)^s = 0,$$

using that $(rA)^\alpha = \frac{R}{r}$ when $N=1$. Thus (7.6) follows from (7.7) by letting $s \uparrow (1-\alpha)$. □

Next we sketch briefly how the cases $N > 1$ are handled. It turns out that any point in E_j (defined in the same way as for $N=1$) has a neighborhood of size $\simeq r^{-j/N} A^{-j}$ on which the *average* of $|h_j|$ is at least $C\frac{1}{j^2} r^{-j}$. This follows by considering Nth order differences. Using a covering of E_j by intervals of this size leads to the desired estimate. Let us make this more precise for $N=2$. We have

$$\sup_\lambda |h_j(\lambda+2y) - 2h_j(\lambda+y) + h_j(\lambda)| \le C_2 A^{2j} |y|^2,$$

hence for $L > 0$ and $\lambda_0 \in E_j$

$$\begin{aligned} \frac{2C_2}{3} A^{2j} L^3 \quad &\ge \quad \int_{[-L,L]} |h_j(\lambda_0+2y) - 2h_j(\lambda_0+y) + h_j(\lambda_0)|\, dy \\ &\ge \quad 2L \cdot |h_j(\lambda_0)| - \int_{[-L,L]} |h_j(\lambda_0+y)|\, dy - \int_{[-L,L]} 2|h_j(\lambda_0+2y)|\, dy \\ &\ge \quad \frac{2L}{j^2} r^{-j} - 2\int_{[-2L,2L]} |h_j(\lambda_0+y)|\, dy. \end{aligned}$$

Therefore,

$$\frac{1}{4L}\int_{[-2L,2L]} |h_j(\lambda_0+y)|\, dy \ge \frac{1}{4j^2} r^{-j} - \frac{C_2}{12} A^{2j} L^2,$$

so taking $L = C_2^{-1/2} j^{-1} r^{-j/2} A^{-j}$ yields that the average of $|h_j|$ on $[\lambda_0 - 2L, \lambda_0 + 2L]$ is at least $\frac{1}{6j^2} r^{-j}$.

The following proposition shows how to obtain a dimension bound from a suitable Sobolev estimate by means of the previous lemma.

Proposition 7.5. *If $\int_I \|\nu_\lambda\|_{2,\gamma}^2\, d\lambda < \infty$ with some $I \subset (0,1)$ and $0 < \gamma < 1/2$, then*

$$\dim\{\lambda \in I : \widehat{\nu_\lambda} \notin L^2(\mathbb{R})\} \le 1 - 2\gamma.$$

Proof. For any $j = 1, 2, \ldots$ define

$$h_j(\lambda) = 2^{-j} \int_{2^{j-1}}^{2^j} |\widehat{\nu_\lambda}(\xi)|^2 \, d\xi = 2^{-j} \int_{2^{j-1}}^{2^j} \int_\Omega \int_\Omega \exp\Big(i\xi\big[\Pi_\lambda(\omega) - \Pi_\lambda(\tau)\big]\Big) \, d\mu(\omega) d\mu(\tau) d\xi.$$

It follows that

$$\|h_j^{(n)}(\lambda)\| \le C_n 2^{jn} \sum_{k=1}^n \Big\| \frac{d^k}{d\lambda^k} \Pi_\lambda(\omega) \Big\|_{L^\infty(I\times\Omega)} \le C_n' 2^{jn}\,.$$

Thus, the first condition of (7.5) is satisfied with $A = 2$, for all $N \in \mathbb{N}$. The second condition of (7.5) holds with $R = 2^{1+2\gamma}$ by the main assumption and the definition of the $L_{2,\gamma}$-norm. Letting $r = 2$ and $N \to \infty$ yields by Lemma 7.4 that

$$\dim\{\lambda \in J : \ \widehat{\nu_\lambda} \notin L^2(\mathbb{R})\} = \dim\{\lambda \in J : \ \sum_{j=0}^\infty 2^j h_j(\lambda) = \infty\} \le 1 - 2\gamma.$$

Observe that the case $N = 1$ of Lemma 7.4 (for which complete details were given) yields (7.4). $\square$

Next we turn to the proof of (7.2). As a preliminary step we present the standard construction of a Littlewood–Paley decomposition, see Stein [50] or Frazier, Jawerth, Weiss [13]. Recall that $\mathcal{S}(\mathbb{R})$ is the Schwartz space of smooth functions all of whose derivatives decay faster than any power. It is a basic property of the Fourier transform that it preserves $\mathcal{S}$.

Lemma 7.6. *There exists $\psi \in \mathcal{S}(\mathbb{R})$ so that $\hat{\psi} \ge 0$,*

$$\operatorname{supp}(\hat{\psi}) \subset \{\xi \in \mathbb{R} \,:\, 1 \le |\xi| \le 4\}, \quad \textit{and} \quad \sum_{j=-\infty}^{\infty} \hat{\psi}(2^{-j}\xi) = 1 \quad \textit{if } \xi \ne 0. \tag{7.10}$$

Moreover, given any finite measure ν on $\mathbb{R}$ and any $\gamma \in \mathbb{R}$

$$\|\nu\|_{2,\gamma}^2 \asymp \sum_{j=-\infty}^{\infty} 2^{2j\gamma} \int_{\mathbb{R}} (\psi_{2^{-j}} * \nu)(x) \, d\nu(x) \tag{7.11}$$

where $\psi_{2^{-j}}(x) = 2^j \psi(2^j x)$.

Proof. Choose $\phi \in \mathcal{S}(\mathbb{R})$ with $\hat{\phi} \ge 0$, $\hat{\phi}(\xi) = 1$ for $|\xi| \le 1$ and $\hat{\phi}(\xi) = 0$ for $|\xi| > 2$. Define ψ via $\hat{\psi}(\xi) = \hat{\phi}(\xi/2) - \hat{\phi}(\xi)$. It is clear that $\hat{\psi}(\xi) \ge 0$ and that $\hat{\psi}(\xi) = 0$ if $|\xi| < 1$ or $|\xi| > 4$. (7.10) holds since the sum telescopes. Moreover, it is clear from (7.10) that there exists some constant C_γ depending only on γ so that for any $\xi \ne 0$

$$C_\gamma^{-1} |\xi|^{2\gamma} \le \sum_{j=-\infty}^{\infty} 2^{2j\gamma} \hat{\psi}(2^{-j}\xi) \le C_\gamma |\xi|^{2\gamma}.$$

Since $\widehat{\psi_{2^{-j}}}(\xi) = \hat{\psi}(2^{-j}\xi)$, Plancherel's theorem implies

$$\int_{\mathbb{R}} (\psi_{2^{-j}} * \nu)(x) \, d\nu(x) = \int \hat{\psi}(2^{-j}\xi) |\widehat{\nu}(\xi)|^2 \, d\xi$$

and (7.11) follows. □

Now we come to the main technical statement needed to prove the Sobolev estimate (7.2).

Lemma 7.7. *Let $J = [\lambda_0, \lambda_1]$ be an interval of δ-transversality for some $\delta > 0$ and let $\beta = \frac{\log \lambda_0}{\log \lambda_1} - 1$. Suppose that $\rho \in C^\infty(\mathbb{R})$ is supported in the interior of J and $\|\rho\|_\infty \le 1$. Let ϕ be any function in the Schwarz space and $\psi(x) = 2\phi(2x) - \phi(x)$. Fix $s \in (1, 2)$. Then for any distinct $\omega, \tau \in \Omega$, with $|\omega \wedge \tau| = k$, and any $R > 0$,*

$$|\mathcal{J}| := \left| \int_{\mathbb{R}} \rho(\lambda)\psi(R[\Pi_\lambda(\omega) - \Pi_\lambda(\tau)])\, d\lambda \right| \le C\left(R\lambda_1^{k(1+3\beta)}\right)^{-s}, \tag{7.12}$$

where C depends only on ρ, β and s.

Proof. Fix $\omega, \tau \in \Omega$ such that $|\omega \wedge \tau| = k$, and $R > 0$. We may assume that

$$R\lambda_1^{k(1+3\beta)} \ge 1 \tag{7.13}$$

since otherwise the estimate is obvious. We can write $\Pi_\lambda(\omega) - \Pi_\lambda(\tau) = 2\lambda^k f(\lambda)$ where $f(\lambda)$ is a power series of the form (7.1). Recall that δ-transversality says that $f'(\lambda) < -\delta$ for $\lambda \in J$ whenever $f(\lambda) < \delta$. Let η be the distance between the support of ρ and the boundary of J. δ-transversality implies that if $|f(\lambda)| < \delta\eta$ for some $\lambda \in \mathrm{supp}(\rho)$, then f has a zero $\overline{\lambda} \in J$ which is the only zero of f on J. Denote $u = \lambda - \overline{\lambda}$. We are going to linearize everything around $\overline{\lambda}$. Clearly,

$$|\lambda^k f(\lambda) - \overline{\lambda}^k f'(\overline{\lambda})u| \le \|f''\|_\infty \lambda_1^k u^2 + \|f'\|_\infty k\lambda_1^{k-1}u^2 \le C_1 k\lambda_1^k u^2, \tag{7.14}$$

with the constant C_1 depending only on J. Let $\chi \in C^\infty$ be non-negative with $\chi = 1$ on $[-\frac{1}{2}, \frac{1}{2}]$ and $\mathrm{supp}(\chi) \subset (-1, 1)$. Then

$$\begin{aligned} \int_{\mathbb{R}} \rho(\lambda)\psi(2R\lambda^k f(\lambda))\, d\lambda &= \int \rho(\lambda)\psi(2R\lambda^k f(\lambda))\chi\Big(\frac{2C_1 k}{\delta^2\eta\lambda_1^{k\beta}} f(\lambda)\Big)\, d\lambda \\ &+ \int \rho(\lambda)\psi(2R\lambda^k f(\lambda))\Big[1 - \chi\Big(\frac{2C_1 k}{\delta^2\eta\lambda_1^{k\beta}} f(\lambda)\Big)\Big]\, d\lambda. \end{aligned} \tag{7.15}$$

The integrand of the second integral is non-zero only if

$$|f(\lambda)| > \frac{\delta^2\eta\lambda_1^{k\beta}}{4C_1 k} \ge C_\beta \lambda_1^{2k\beta}.$$

Then, using that $\lambda^k \ge \lambda_1^{k(1+\beta)}$ for $\lambda \in [\lambda_0, \lambda_1]$ we conclude that $|2R\lambda^k f(\lambda)| \ge 2C_\beta R\lambda_1^{k(1+3\beta)}$. By the rapid decay of ψ, the second integral in (7.15) is therefore less than $C'_\beta(R\lambda_1^{k(1+3\beta)})^{-s}$.

Thus it suffices to estimate the first integral in (7.15), which we denote by $\mathcal{J}_1$. Notice that its integrand is nonzero only if $|f(\lambda)| \le \frac{\delta^2\eta\lambda_1^{k\beta}}{2C_1 k}$, hence

$$|u| = |\lambda - \overline{\lambda}| \le \frac{\delta\eta\lambda_1^{k\beta}}{2C_1 k} < \frac{\delta\lambda_1^{k\beta}}{2C_1 k} \tag{7.16}$$

by δ-transversality. This implies that

$$C_1 k \lambda_1^k |u|^2 < \tfrac{1}{2}\,\overline{\lambda}^k \delta |u| \le \tfrac{1}{2}\,\overline{\lambda}^k |f'(\overline{\lambda})| u,$$

hence by (7.14)

$$\tfrac{1}{2}\,\overline{\lambda}^k |f'(\overline{\lambda})| u \le \lambda^k |f(\lambda)| \le \tfrac{3}{2}\,\overline{\lambda}^k |f'(\overline{\lambda})| u. \tag{7.17}$$

Let $g(\lambda) = \chi\big(\frac{2C_1 k}{\delta^2 \eta \lambda_1^{k\beta}} f(\lambda)\big)\rho(\lambda)$. Since $g(\lambda) = g(\overline{\lambda} + u) = g(\overline{\lambda}) + O(\|g'\|_\infty u)$ we have

$$\begin{aligned} \mathcal{J}_1 &= \int g(\overline{\lambda} + u)\big[\psi(2R\lambda^k f(\lambda)) - \psi(2R\overline{\lambda}^k f(\overline{\lambda})u)\big]\, du \\ &+ \int g(\overline{\lambda})\psi(2R\overline{\lambda}^k f(\overline{\lambda})u)\, du \\ &+ \int O(\|g'\|_\infty u)\, \psi(2R\overline{\lambda}^k f(\overline{\lambda})u)\, du = I_1 + I_2 + I_3. \end{aligned} \tag{7.18}$$

The integral I_2 is the easiest one: since $\psi(t) = 2\phi(2t) - \phi(t)$ we have $\int \psi(t)\, dt = 0$, so $I_2 = 0$.

To estimate I_1 we use that $\|g\|_\infty \le 1$, the mean value theorem, (7.14), (7.17), the rapid decay of ψ', and that $|f'(\overline{\lambda})| > \delta$, to obtain

$$\big|\psi(2R\lambda^k f(\lambda)) - \psi(2R\overline{\lambda}^k f(\overline{\lambda})u)\big| \le CkR\lambda_1^k u^2 \min\{1,\, (R\overline{\lambda}^k u)^{-4}\}\,. \tag{7.19}$$

(The exponent -4 above can be replaced by any negative integer by changing the constant but is sufficient for our purposes.) Since $s < 2$ we can find ε so that

$$0 < \varepsilon < \frac{2 - s}{3}\,. \tag{7.20}$$

We write

$$I_1 \le \int_{\mathbb{R}} CkR\lambda_1^k u^2 \min\{1,\, (R\overline{\lambda}^k u)^{-4}\}\, du = \int_{|u| \le (R\lambda_1^k)^{\varepsilon - 1}} + \int_{|u| \ge (R\lambda_1^k)^{\varepsilon - 1}} = I_{11} + I_{12}.$$

We have

$$I_{11} \le \int_{|u| \le (R\lambda_1^k)^{\varepsilon - 1}} CkR\lambda_1^k u^2\, du = C'k(R\lambda_1^k)^{-2+3\varepsilon} =: C'S_{11}.$$

Estimating I_{12} we use that $\overline{\lambda}^{-1} \le \lambda_1^{-(1+\beta)}$ to get

$$I_{12} \le \int_{|u| \ge (R\lambda_1^k)^{\varepsilon - 1}} CkR\lambda_1^k \cdot R^{-4}\lambda_1^{-4k(1+\beta)} u^{-2}\, du = C''kR^{-2-\varepsilon}\lambda_1^{-k(2+4\beta+\varepsilon)} =: C''S_{12}.$$

A straightforward computation shows that $\max\{S_{11}, S_{12}\} \le C_\beta \big(R\lambda_1^{k(1+3\beta)}\big)^{-s}$. Let us demonstrate this for S_{11}. Since $k \le C_\beta \lambda_1^{-k\beta}$ it is enough to establish that

$$R^{-2+3\varepsilon}\lambda_1^{-k(2+\beta-3\varepsilon)} \le \big(R\lambda_1^{k(1+3\beta)}\big)^{-s}$$

which is equivalent to

$$R^{s-2+3\varepsilon}\lambda_1^{k[(1+3\beta)s-(2+\beta-3\varepsilon)]} \le 1.$$

Applying (7.20) and (7.13) reduces the last inequality to $3(2-3\varepsilon) \ge 1$ which is certainly true. The expression S_{12} is handled similarly, eventually reducing the estimate to $3(2+\varepsilon) \ge 5$. This concludes the estimation of the integral I_1 from (7.18).

It remains to estimate I_3. First recall the definition of $g(\lambda)$ to get

$$\|g'\|_\infty \le Ck\lambda_1^{-k\beta}.$$

By the rapid decay of ψ and since $|f'(\overline{\lambda})| \ge \delta$ we obtain

$$I_3 \le \int Ck\lambda_1^{-k\beta} u \min\{1, (R\overline{\lambda}^k u)^{-3}\}\, du = \int_{|u|\le(R\lambda_1^k)^{\varepsilon-1}} + \int_{|u|\ge(R\lambda_1^k)^{\varepsilon-1}} = I_{31} + I_{32}.$$

where ε is again defined by (7.20). Now

$$I_{31} \le \int_{|u|\le(R\lambda_1^k)^{\varepsilon-1}} Ck\lambda_1^{-k\beta} u\, du = C'kR^{-2+2\varepsilon}\lambda_1^{-k(2+\beta-2\varepsilon)} =: C'S_{31}$$

and

$$I_{32} \le \int_{|u|\ge(R\lambda_1^k)^{\varepsilon-1}} Ck\lambda_1^{-k\beta}R^{-3}\lambda_1^{-3k(1+\beta)}u^{-2}\, du = C''kR^{-2-\varepsilon}\lambda_1^{-k(2+4\beta+\varepsilon)} =: C''S_{32}.$$

It remains to check that $\max\{S_{31}, S_{32}\} \le \left(R\lambda_1^{k(1+3\beta)}\right)^{-s}$ which is done similarly to the above (in fact, $S_{32} = S_{12}$ and the estimate of S_{31} reduces to $s < 2-2\varepsilon$ and $3(2-2\varepsilon) \ge 2$). This concludes the estimation of the integral I_3, and the proof of the lemma is complete. $\square$

Proof of Theorem 7.3. If an interval of δ-transversality is enlarged slightly, it is obviously going to be an interval of $\frac{\delta}{2}$-transversality. Thus, (7.2) will be established if we show that for an arbitrary interval of δ-transversality $J = [\lambda_0, \lambda_1] \in (\frac{1}{2}, 1)$ and any function $\rho \in C^\infty(\mathbb{R})$ supported in its interior, $\int_{\mathbb{R}} \|\nu_\lambda\|_{2,\gamma}^2\, \rho(\lambda)\, d\lambda < \infty$. Let $\lambda_0^\alpha = \frac{1}{2}$. We have $\alpha < 2$ by the discussion of transversality above. Fix some $\gamma < \frac{1}{2}$ with $\lambda_0^{1+2\gamma} > \frac{1}{2}$ and $s \in (1+2\gamma, 2)$. Assume first that J is so short that $0 < (1+2\gamma)(1+3\beta) \le \alpha$ where $\beta = \frac{\log\lambda_0}{\log\lambda_1} - 1$. Let ψ be the Littlewood-Paley function from

Lemma 7.6. In view of (7.11), the definition of ν_λ, and Lemma 7.7

$$\int_{-\infty}^{\infty} \|\nu_\lambda\|_{2,\gamma}^2 \,\rho(\lambda)\,d\lambda \asymp \int \sum_{j=-\infty}^{\infty} 2^{2j\gamma} \int_{-\infty}^{\infty} (\psi_{2^{-j}} * \nu_\lambda)(x)\,d\nu_\lambda(x)\;\rho(\lambda)\,d\lambda$$

$$\begin{aligned}
&\leq \int_\Omega \int_\Omega \sum_{j=-\infty}^{\infty} 2^{j(1+2\gamma)} \left| \int \psi(2^j[\Pi_\lambda(\omega) - \Pi_\lambda(\tau)])\rho(\lambda)\,d\lambda \right| d\mu(\omega)d\mu(\tau) \\
&\leq C_{\beta,\gamma} \int_\Omega \int_\Omega \sum_{-\infty}^{\infty} 2^{j(1+2\gamma)} \min\Big\{1, \big(2^j \lambda_1^{|\omega\wedge\tau|(1+3\beta)}\big)^{-s}\Big\}\,d\mu(\omega)d\mu(\tau) \\
&\leq C_{\beta,\gamma} \int_\Omega \int_\Omega \frac{d\mu(\omega)d\mu(\tau)}{\lambda_1^{(1+3\beta)(1+2\gamma)|\omega\wedge\tau|}} \leq C_{\beta,\gamma}\,\mathcal{E}_\alpha(\mu) < \infty,
\end{aligned}$$

as claimed. This argument depended on β being sufficiently small. In the general case fix any small $\beta > 0$ and partition J into subintervals $J_i = [\lambda_i, \lambda_{i+1}]$ for $i = 0, \ldots, m$ so that $\lambda_i > \lambda_{i+1}^{1+\beta}$. Applying the previous calculation to each of the J_i and summing concludes the proof of (7.2). The dimension estimate (7.3) follows from (7.2) and Proposition 7.5, letting $\gamma \to \frac{1}{2}(\frac{\log 2}{-\log \lambda_0} - 1)$. ☐

It is well–known that for $0 < \lambda < \frac{1}{2}$ the support of ν_λ is a Cantor set of dimension $\frac{\log 2}{-\log \lambda}$. In fact, ν_λ is a Frostman measure on that set, which implies that $\dim(\nu_\lambda) = \frac{\log 2}{-\log \lambda}$ for $0 < \lambda < \frac{1}{2}$. Solomyak [47] showed that the first double zero for a power series of the form (7.1) lies in the interval $[0.649, 0.683]$. In particular, the previous theorem will apply only up to some point in this interval. Nevertheless, one can show that ν_λ has some smoothness for a.e. $\lambda \in (\frac{1}{2}, 1)$. This follows from Theorem 7.3 by "thinning and convolving", see [47] and [39]. As one expects, the number of derivatives tends to ∞ as $\lambda \to 1$.

Lemma 7.8. *For any $\epsilon > 0$ there exists a $\gamma = \gamma(\epsilon) > 0$ so that*

$$\int_{\frac{1}{2}+\epsilon}^{\frac{1}{\sqrt{2}}} \|\nu_\lambda\|_{2,\gamma}^2 \,d\lambda < \infty. \tag{7.21}$$

Furthermore, there exists some $\ell_0 \in (2^{-1/2}, 2^{-1/4})$ and a $\gamma_0 > 0$ so that

$$\int_{\ell_0^2}^{\ell_0} \|\nu_\lambda\|_{2,\gamma_0}^2 \,d\lambda < \infty. \tag{7.22}$$

Proof. As mentioned above, $[0, \lambda_1]$ is an interval of transversality for the power series (7.1) for some $\lambda_1 > 2^{-2/3}$. Fix any $\lambda_0 \in (\frac{1}{2}, 2^{-2/3}]$. Partitioning the interval $[\lambda_0, \lambda_1]$ as in the proof of Theorem 7.3, one obtains from (7.2) that

$$\int_{\lambda_0}^{\lambda_1} \|\nu_\lambda\|_{2,\gamma}^2 \,d\lambda < \infty \quad \text{provided} \quad \lambda_0^{1+2\gamma} > \frac{1}{2}. \tag{7.23}$$

To go beyond $2^{-2/3}$ we remove every third term from the original series. More precisely, let $\widetilde{\Pi}_\lambda(\omega) = \sum_{3\not| n} \omega_n \lambda^n$ and denote the distribution of this series by $\widetilde{\nu_\lambda}$. It was shown in [47] and [39] that the class of power series (7.1) that satisfy either $b_{3j+1} = 0$ for all $j \geq 0$ or $b_{3j+2} = 0$ for all $j \geq 0$ have $[0, \lambda_3]$ as an interval of δ–transversality for some $\lambda_3 > 1/\sqrt{2}$. After some straightforward modifications the argument given above for the full series shows that

$$\int_{\lambda_2}^{\lambda_3} \|\widetilde{\nu_\lambda}\|_{2,\gamma}^2 \, d\lambda < \infty \quad \text{provided} \quad \lambda_2^{1+2\gamma} > 2^{-2/3}. \tag{7.24}$$

For more details we refer the reader to section 5.1 of [38]. Since $|\widehat{\nu_\lambda}| \leq |\widehat{\widetilde{\nu_\lambda}}|$ and $\lambda_1 > 2^{-2/3}$, (7.21) follows from (7.23) and (7.24). Moreover, we have shown (7.22). □

Corollary 7.9. *For any $\lambda_0 > \frac{1}{2}$ there exists $\epsilon(\lambda_0) > 0$ such that*

$$\dim\{\lambda \in (\lambda_0, 1) \,:\, \nu_\lambda \text{ does not have } L^2\text{–density}\} < 1 - \epsilon(\lambda_0).$$

Proof. This follows from the previous lemma and Proposition 7.5 using the identity

$$\widehat{\nu_\lambda}(\xi) = \widehat{\nu_{\lambda^2}}(\xi)\widehat{\nu_{\lambda^2}}(\lambda\xi)\,.$$

□

As observed by Kahane [19], Erdős's argument yields that $\epsilon(\lambda_0) \to 1$ as $\lambda_0 \uparrow 1$ (see section 6).

8. Applications, generalizations and problems

8.1. **Applications to dimension and dynamics.** Alexander and Yorke [1] considered the "fat baker's transformation"

$$T_\lambda(x, y) = \begin{cases} (\lambda x + (1-\lambda), 2y-1) & \text{if } y \geq 0 \\ (\lambda x - (1-\lambda), 2y+1) & \text{if } y < 0 \end{cases}$$

on the square $[-1, 1]^2$. They proved that the Sinai-Bowen-Ruelle measure η_λ for T_λ is the product of ν_λ (more precisely, its affine copy supported on $[-1, 1]$) and the uniform measure in y-direction. They showed further that absolute continuity of ν_λ implies the equality of the information (Rényi) and Lyapunov dimension for η_λ but this breaks down in the Pisot case.

Another application of Bernoulli convolutions has to do with fractal graphs. Let ϕ be a $\mathbb{Z}$-periodic function and $\lambda \in (\frac{1}{2}, 1)$. Define

$$\Gamma_{\lambda,\phi} = \Big\{(x, y) : \; x \in [0, 1], \; y = \sum_{n=0}^{\infty} \lambda^n \phi(2^n x)\Big\}.$$

If $\phi(x) = \cos(2\pi x)$ this defines a family of Weierstrass nowhere differentiable functions. It is an open problem to compute $\dim \Gamma_{\lambda,\cos(2\pi x)}$, even for a typical λ. This problem served as a motivation for studying $\Gamma_{\lambda,\phi}$ with an easier choice of ϕ.

Przytycki and Urbański [42] considered the case $\phi(x) = r(x) = 1$ if $x \in [0, \frac{1}{2})$ mod 1 and $r(x) = -1$ otherwise. This is a discontinuous function with a self-affine graph. It is proved in [42] that if $\dim \nu_\lambda = 1$ then

$$\dim \Gamma_{\lambda,r(x)} = \dim_M \Gamma_{\lambda,r(x)} = 2 - \frac{\log(1/\lambda)}{\log 2}.$$

Here $\dim_M$ is the Minkowski (box) dimension. The equality for Minkowski dimension holds for all $\lambda \in (\frac{1}{2}, 1)$ but the Hausdorff dimension drops for reciprocals of Pisot numbers. We note that the methods of [42] readily extend to the case of more general self-affine sets invariant for the iterated function system $\{(\gamma x, \lambda x - 1), (\gamma x + (1-\gamma), \lambda x + 1)\}$ for $\gamma \in (0, \frac{1}{2})$. In particular, $\dim \nu_\lambda = 1$ suffices for the equality of the Hausdorff and Minkowski dimensions.

Ledrappier [29] studied the family of continuous graphs $\Gamma_{\lambda,\phi}$ where $\phi(x) = \mathrm{dist}(x, \mathbb{Z})$ (sometimes called Takagi graphs). Their analysis is quite a bit harder. Ledrappier proved that if $\dim \nu_{(2\lambda)^{-1}} = 1$ then $\dim \Gamma_{\lambda,\phi} = 2 - \frac{\log(1/\lambda)}{\log 2}$.

Observe that in all applications mentioned here it is the equality $\dim \nu_\lambda = 1$ that gets used, not the absolute continuity of ν_λ.

8.2. **Generalizations.** There are many natural generalizations of Bernoulli convolutions; many of them can be treated similarly to the classical case with some additional work.

(i) Biased Bernoulli convolutions: as in the classical case but the signs are taken with probabilities $(p, 1-p)$. We will generalize a bit further:

(ii) Suppose that $D \subset \mathbb{R}$ is an arbitrary finite set of digits, with $\mathrm{card}(D) = m$, and $\mathbf{p} = (p_1, \dots, p_m)$ is a probability vector. Let $\nu_\lambda^{D,\mathbf{p}}$ be the distribution of the random series $\sum_{n=0}^\infty a_n \lambda^n$ where $a_n \in D$ independently with probabilities p_i.

"Almost sure" results on the existence of a density in $L^q(\mathbb{R})$ for $\nu_\lambda^{D,\mathbf{p}}$, when $q \in [1, 2]$, were obtained in [40], and the dimension of exceptions was estimated in [38] (for $q = 1$ and 2). These results were proved on an interval of transversality which, in this case, means an interval free of double zeros for power series with coefficients in $D - D$. Checking transversality is not always easy, so some of these results are less complete than those for classical Bernoulli convolutions. For instance, it is proved in [40] that the $(p, 1-p)$ Bernoulli convolutions are absolutely continuous for a.e. $\lambda \in (p^p(1-p)^{1-p}, 1)$, but only for $p \in [1/3, 2/3]$.

It is easy to see that the Erdős-Kahane argument transfers to the case of $\nu_\lambda^{D,\mathbf{p}}$. The question of convergence to zero at infinity of $\widehat{\nu_\lambda}^{D,\mathbf{p}}$ was considered by Salem (Borwein and Girgensohn [4] were apparently unaware of this when they discussed some special cases.) Making a linear change of variable we can assume that the first two digits in D are 0 and 1. Then $\widehat{\nu_\lambda}^{D,\mathbf{p}}$ tends to zero at infinity if and only $\theta = \lambda^{-1}$ is Pisot and and D lies in the field of θ, see [44, Ch.VII].

(iii) Consider the same set-up as in (ii) but with complex a_i and λ complex of modulus less than one. Some results were obtained in [49] (and the dimension of exceptions was estimated in [38]) but checking transversality becomes more formidable. Note that here, determining the support of the measure in the two-digit case is non-trivial.

(iv) Convolutions of self-similar measures and arithmetic sums of Cantor sets: see [48, 40, 38] for some "almost sure" results; see also the references in [48] for other work on sums of Cantor sets and the connection with smooth dynamics and the Palis-Takens problem.

8.3. **Problems on the Bernoulli convolutions ν_λ.**

1. Other properties of the density. Since ν_λ has a density $\frac{d\nu_\lambda}{dx}$ in $L^2(\mathbb{R})$ for a.e. $\lambda \in (\frac{1}{2}, 1)$, it follows from the formula

$$\nu_\lambda(\cdot) = \nu_{\lambda^2}(\cdot) * \nu_{\lambda^2}(\lambda\cdot) \tag{8.1}$$

that $\frac{d\nu_\lambda}{dx}$ is continuous for a.e. $\lambda \in (2^{-1/2}, 1)$. It is not known whether $\frac{d\nu_\lambda}{dx}$ is continuous, or even bounded, for a.e. $\lambda \in (\frac{1}{2}, 2^{-1/2})$. Using (8.1) again and the result of Mauldin and Simon [36], we may infer that for a.e. $\lambda \in (2^{-1/2}, 1)$, the density of ν_λ is strictly positive in the interior of its support. We do not know whether for a.e. $\lambda \in (\frac{1}{2}, 2^{-1/2})$, the essential infimum of $\frac{d\nu_\lambda}{dx}$ on any compact subinterval of $\mathrm{supp}(\nu_\lambda)$ is positive.

Numerical approximation of self-similar measures was studied in several papers, among them [51] which contains histograms of ν_λ for some λ.

2. Is absolute continuity generic? We saw in Proposition 4.1 that for $\lambda \in (\frac{1}{2}, 1)$, the Bernoulli convolution generically has correlation and Hausdorff dimension equal to one. The analogous question for absolute continuity is open. We are grateful to Elon Lindenstrauss for simplifying the original proof of the following proposition.

Proposition 8.1. *The set* $S_\perp = \{\lambda \in (\frac{1}{2}, 1) : \ \nu_\lambda \text{ is singular } \}$ *is* G_δ.

Proof. It is easy to see that the function $\lambda \mapsto \nu_\lambda(a, b)$ is continuous for any interval (a, b). Let $\mathcal{G}$ be the collection of all finite unions of open intervals. Fix a sequence ε_n converging to 0. Now observe that

$$S_\perp = \bigcap_n \bigcup_{\mathcal{L}(G)<\varepsilon_n} \{\lambda \in (1/2, 1) : \ \nu_\lambda(G) > 0.5\}$$

where the union is over all $G \in \mathcal{G}$ with $\mathcal{L}(G) < \varepsilon_n$. Thus $S_\perp$ is a G_δ set.

A consequence of this proposition is that if absolute continuity holds on a residual set in $(\frac{1}{2}, 1)$, then the exceptional set $S_\perp$ is nowhere dense, and hence, by (8.1), there is a left neighborhood of 1 which is disjoint from $S_\perp$.

3. Let $\frac{1}{2} < a < 1$. Is it possible to prove that the set $\{\lambda \in (\frac{1}{2}, 1) : \ \nu_\lambda \text{ is singular } \}$ has *packing* dimension strictly less than 1?
(The methods of [38] only give such a bound for the Hausdorff dimension.)

4. Let

$$J_n(\lambda) := \int_{2^n}^{2^{n+1}} \widehat{\nu_\lambda}(\xi)^2 \, d\xi \,.$$

Does the limit $\lim_{n\to\infty} J_n(\lambda)^{1/n}$ exist for *all* $\lambda \in (1/2, 1)$?

A positive answer would imply that some neigborhood of 1 does not contain Salem numbers. To see this, fix $\frac{1}{2} < \lambda_0 < 1$. As proved in [38] (see §7), there exists $\gamma > 0$ such that $\nu_\lambda \in L^2_\gamma$ for a.e. $\lambda \in [\lambda_0, 1)$. The set

$$W_\epsilon := \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} \left\{ \lambda \in [\lambda_0, 1) \,:\, J_n(\lambda) < (1-\epsilon)^n \right\}$$

is a G_δ set in $[\lambda_0, 1)$ for any $\epsilon > 0$. Moreover, W_ϵ is dense in $[\lambda_0, 1)$ provided that $1 - \epsilon > 2^{-2\gamma}$. If existence of $\lim_{n\to\infty} J_n(\lambda)^{1/n}$ could be proved for all $\lambda \in W_\epsilon$, then it would follow that $W_\epsilon \subset \{\lambda \in (\lambda_0, 1) \,:\, \nu_\lambda \in L^2_{\gamma_0}\}$ provided that $2^{-2\gamma_0} > 1 - \epsilon$; Proposition 5.1 could then be invoked to deduce that 1 is not a limit of Salem numbers.

Acknowledgements. We are indebted to Karoly Simon for his many original insights, that have significantly influenced our work on Bernoulli convolutions. We thank the organizers of the conference on Fractals and Stochastics, C. Bandt, S. Graf and M. Zähle, for their hospitality. The title of this survey was inspired by the beautiful survey by Lyons [31] on the related topic of Rajchman measures.

References

1. J. C. Alexander, J. A. Yorke, *Fat Baker's transformations.* Ergodic Th. Dynam. Sys. 4 (1984), 1–23.
2. J. C. Alexander, D. Zagier, *The entropy of a certain infinitely convolved Bernoulli measure.* J. London Math. Soc. (2) 44 (1991), 121–134.
3. M. J. Bertin, A. Decomps-Guilloux, M. Grandet-Hugot, M. Pathiaux-Delefosse, J. P. Shreiber, *Pisot and Salem numbers.* Birkhäuser (1992).
4. J. Borwein, R. Girgensohn, *Functional equations and distribution functions.* Results in Math. 26 (1994), 229–237.
5. A. Bovier, *Bernoulli convolutions, dynamical systems and automata.* Disordered systems (Temuco, 1991/1992), 63–86, Travaux en Cours, 53, Hermann (1996).
6. D. Boyd, *Transcendental numbers with badly distributed powers.* Proc. Amer. Math. Soc. 23 (1969), 424–427.
7. D. Boyd, *Speculations concerning the range of Mahler's measure.* Can. Math. Bull. 24 (1981), 453–469.
8. G. Derfel, R. Schilling, *Spatially chaotic configurations and functional equations.* J. Phys. A: Math. Gen. 29 (1996), 4537–4547.
9. P. Erdős, *On a family of symmetric Bernoulli convolutions.* Amer. J. Math. 61 (1939), 974–975.
10. P. Erdős, *On the smoothness properties of Bernoulli convolutions.* Amer. J. Math. 62 (1940), 180–186.
11. K. J. Falconer, *Hausdorff dimension and the exceptional set of projections.* Mathematika 29, 1982. 109–115.
12. K. J. Falconer, *The Hausdorff dimension of self-affine fractals.* Math. Proc. Camb. Phil. Soc. 103 (1988), 339–350.
13. M. Frazier, B. Jawerth, G. Weiss, *Littlewood–Paley theory and the study of function spaces.* CBMS Regional Conference Ser. 79, American Math. Soc. (1991).
14. A. M. Garsia, *Arithmetic properties of Bernoulli convolutions.* Trans. Amer. Math. Soc. 102 (1962), 409–432.
15. A. M. Garsia, *Entropy and singularity of infinite convolutions.* Pacific J. Math. 13 (1963), 1159–1169.

16. T.-Y. Hu, *The local dimensions of the Bernoulli convolution associated with the golden number.* Trans. Amer. Math. Soc. 349, no. 7 (1997), 2917–2940.
17. J. E. Hutchinson, *Fractals and self-similarity.* Indiana Univ. Math. J. 30 (1981), 713–747.
18. B. Jessen, A. Wintner, *Distribution functions and the Riemann zeta function.* Trans. Amer. Math. Soc. 38 (1935), 48–88.
19. J. P. Kahane, *Sur la distribution de certaines series aleatoires.* Colloque Th. Nombres [1969, Bordeaux], Bull. Soc. math. France, Mémoire 25 (1971), 119–122.
20. J. P. Kahane, R. Salem, *Sur la convolution d'une infinité de distributions de Bernoulli.* Colloq. Math. 6 (1958), 193–202.
21. R. Kaufman, *On Hausdorff dimension of projections.* Mathematika 15 (1968), 153-5.
22. 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.
23. R. Kershner, A. Wintner, *On symmetric Bernoulli convolutions.* Amer. J. Math. 57 (1935), 541–548.
24. L. Kuipers, H. Niederreiter, *Uniform distribution of sequences.* Wiley (1974).
25. S. Lalley, *Random series in powers of algebraic integers: Hausdorff dimension of the limit distribution.* J. London Math. Soc., to appear.
26. K.-S. Lau, *Dimension of a family of singular Bernoulli convolutions.* J. Funct. Anal. 116 (1993), 335–358.
27. K.-S. Lau, S.-M. Ngai, *L^q-spectrum of the Bernoulli convolution associated with the golden ratio.* Studia Math. 131, no. 3, (1998), 225–251.
28. K.-S. Lau, S.-M. Ngai, *L^q-spectrum of the Bernoulli convolutions associated with the P.V. numbers.* Preprint.
29. F. Ledrappier, *On the dimension of some graphs.* Contemporary Math. 135 (1992), 285–293.
30. F. Ledrappier, A. Porzio, *A dimension formula for Bernoulli convolutions.* J. Stat. Phys. 76, nos. 5/6 (1994), 1307–1327.
31. R. Lyons, *Seventy years of Rajchman measures.* J. Fourier Anal. Appl., Kahane Special Issue (1995), 363–377.
32. J. Marstrand, *Some fundamental geometrical properties of plane sets of fractional dimension.* Proc. London Math. Soc. 4 (1954), 257–302.
33. P. Mattila, *Hausdorff dimension, orthogonal projections and intersections with planes.* Ann. Acad. Sci. Fenn. Ser. A Math. 1 (1975), 227–244.
34. P. Mattila, *Orthogonal projections, Riesz capacities and Minkowski content.* Indiana Univ. Math. J. 39 (1990), 185–198.
35. P. Mattila, *Geometry of sets and measures in Euclidean spaces.* Cambridge University Press (1995).
36. D. Mauldin, K. Simon, *The equivalence of some Bernoulli convolutions to Lebesgue measure.* Proc. Amer. Math. Soc. 126, no. 9 (1998), 2733–2736.
37. Y. Meyer, *Algebraic numbers and harmonic analysis.* North-Holland (1972).
38. Y. Peres, W. Schlag, *Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions.* Submitted.
39. Y. Peres, B. Solomyak, *Absolute continuity of Bernoulli convolutions, a simple proof.* Math. Research Letters 3, no. 2 (1996), 231–239.
40. Y. Peres, B. Solomyak, *Self–similar measures and intersections of Cantor sets.* Trans. Amer. Math. Soc. 350, no. 10 (1998), 4065–4087.

41. M. Pollicott, K. Simon, *The Hausdorff dimension of λ-expansions with deleted digits.* Trans. Amer. Math. Soc. 347, no. 3 (1995), 967–983.
42. F. Przytycki, M. Urbański, *On Hausdorff dimension of some fractal sets.* Studia Math. 93 (1989). 155–186.
43. R. Salem, *A remarkable class of algebraic integers, proof of a conjecture by Vijayaraghavan.* Duke Math. J. 11 (1944), 103–108.
44. R. Salem, *Algebraic numbers and Fourier analysis.* Heath (1963).
45. T. D. Sauer and J. A. Yorke, *Are the dimensions of a set and its image equal under typical smooth functions?* Ergod. Th. Dynam. Sys. 17 (1997), 941–956.
46. N. Sidorov and A. Vershik, *Ergodic properties of Erdős measure, the entropy of the goldenshift, and related problems,* Preprint, 1997.
47. B. Solomyak, *On the random series $\sum \pm\lambda^i$ (an Erdős problem).* Annals of Math. 142 (1995), 611–625.
48. B. Solomyak, *On the measure of arithmetic sums of Cantor sets.* Indagationes Math., N.S. 8 (1997), 133–141.
49. B. Solomyak, *Measure and dimension of some fractal families.* Proc. Cambridge Phil. Soc., 124, no.3 (1998), 531–546.
50. E. Stein, *Harmonic Analysis.* Princeton University Press (1993).
51. R. S. Strichartz, A. Taylor, T. Zhang, *Densities of self-similar measures on the line.* Experimental Math. 4, no. 2 (1995), 101–128.
52. A. Wintner, *On convergent Poisson convolutions.* Amer. J. Math. 57 (1935), 827–838.
53. L.-S. Young, *Dimension, entropy and Lyapunov exponents.* Ergodic Th. Dynam. Sys. 2 (1982), 109–124.

Yuval Peres, Department of Mathematics, Hebrew University, Jerusalem and Department of Statistics, University of California at Berkeley. peres@math.huji.ac.il

Wilhelm Schlag, Department of Mathematics, Fine Hall, Princeton University Princeton, NJ 08544, USA schlag@math.princeton.edu

Boris Solomyak, Box 354350, Department of Mathematics, University of Washington, Seattle, WA 98195, USA, solomyak@math.washington.edu

Current address: Institute of Mathematics, The Hebrew University of Jerusalem, Givat Ram, Jerusalem, Israel, solomyak@math.huji.ac.il