

Recent progress on Bernoulli convolutions

Péter P. Varjú*

Abstract. The Bernoulli convolution with parameter $\lambda \in (0, 1)$ is the measure on \mathbf{R} that is the distribution of the random power series $\sum \pm \lambda^n$, where \pm are independent fair coin-tosses. This paper surveys recent progress on our understanding of the regularity properties of these measures.

2010 Mathematics Subject Classification. 11R06, 28A80, 42A85.

Keywords. Bernoulli convolution, self-similar measure, entropy, Lehmer's conjecture

1. Introduction

Fix a parameter $\lambda \in (0, 1)$ and let X_0, X_1, X_2, \dots be a sequence of independent random variables, whose distribution satisfies $\mathbf{P}(X_j = -1) = \mathbf{P}(X_j = 1) = 1/2$. The Bernoulli convolution with parameter λ is the probability measure ν_λ on \mathbf{R} that is the distribution of the random variable

$$\sum_{j=0}^{\infty} X_j \lambda^j. \quad (1)$$

The study of Bernoulli convolutions goes back to Jessen and Wintner [34], Kershner and Wintner [38] and Wintner [64]. The main question of interest is to determine the set of parameters for which the measure is absolutely continuous with respect to the Lebesgue measure. It was proved in [34] that ν_λ is always of pure type, i.e. it is either absolutely continuous or singular with respect to the Lebesgue measure. (The result holds more generally for infinite convolutions of discrete measures.)

We are also interested in the related question of determining the set of parameters for which $\dim \nu_\lambda = 1$. It is known that ν_λ is exact dimensional (see e.g. [21] for a very general result), that is to say, there is a number α such that

$$\lim_{r \rightarrow 0} \frac{\log(\nu_\lambda(x-r, x+r))}{\log r} = \alpha$$

for ν_λ -almost every x , and we call this number α the dimension of ν_λ .

If $\lambda < 1/2$, then ν_λ is the Cantor-Lebesgue measure on a Cantor set, hence it is singular (as observed in [38]). Moreover, the dimension of the measure is $\dim \nu_\lambda = \log 2 / \log \lambda^{-1}$. The reason for this behavior is that any term of the series (1) dominates the combined contribution of the following terms. On the other

*I gratefully acknowledge the support of the Royal Society.

hand, $\nu_{1/2}$ is the normalized Lebesgue measure restricted to $[-2, 2]$. For $\lambda > 1/2$, ν_λ is more difficult to describe.

The aim of this note is to survey recent progress on this problem. In Section 2, we briefly recall some earlier results. We discuss generalizations and an application in Section 3. Finally, we turn to recent developments in Sections 4 and 5.

Acknowledgment. I am grateful to Jean Bourgain, Emmanuel Breuillard, Mike Hochman, Elon Lindenstrauss, Pablo Shmerkin and Boris Solomyak for many helpful and inspiring discussions on the subject of this note.

2. Earlier results

This section is a very brief overview of some of the earlier results on Bernoulli convolutions, which is by no means intended to be comprehensive. A more detailed account of the first sixty years of the problem can be found in [48].

2.1. Fourier transform. Fourier transform is a powerful tool in studying Bernoulli convolutions, and it dominated the early literature including [34], [38] and [64]. Erdős furthered this line of research in two important papers [17], [18]. Since ν_λ is the law of a sum of independent random variables, its Fourier transform is the product of the Fourier transforms of the individual terms. Indeed,

$$\widehat{\nu}_\lambda(t) = \prod_{j=0}^{\infty} \cos(2\pi\lambda^j t). \quad (2)$$

Erdős used this formula to prove the following two results.

Theorem 2.1 (Erdős [17]). *Suppose that $1/2 < \lambda < 1$ is a number such that λ^{-1} is a Pisot number. Then ν_λ is singular with respect to the Lebesgue measure.*

Recall that a Pisot number (or Pisot-Vijayaraghavan number) is an algebraic integer all of whose Galois conjugates are inside the unit disk. One example of such a number is the Golden ratio $(1 + \sqrt{5})/2$. If λ^{-1} is a Pisot number, then λ^j is very closely approximated by integers, in fact $\text{dist}(\lambda^j, \mathbf{Z}) < a^{|j|}$ for all $j \in \mathbf{Z}$ for some number $a < 1$ depending only on λ . This follows from the fact that the sum of all Galois conjugates of λ^j is an integer. Using this, Erdős proved that

$$|\widehat{\nu}_\lambda(\lambda^n)| \geq \prod_{j \in \mathbf{Z}} |\cos(2\pi\lambda^j)|$$

can be bounded below by a positive number independent of $n \in \mathbf{Z}$, hence ν_λ would violate the Riemann Lebesgue lemma if it were absolutely continuous.

The second result of Erdős is the following.

Theorem 2.2 (Erdős [18]). *For every $m \in \mathbf{Z}_{\geq 0}$, there is $a_m < 1$ such that the following holds. For almost every $\lambda \in (a_m, 1)$, ν_λ is absolutely continuous with respect to the Lebesgue measure, and its density is m times continuously differentiable.*

Referring to the original paper for the proof, we restrict ourselves to a brief heuristic. We consider the factors $\cos(2\pi\lambda^j t)$ in (2). For $j > \log t / \log \lambda^{-1}$, the argument of $\cos(\cdot)$ is close to 0, hence the value of these factors is close to 1. On the other hand, for a typical λ , it is reasonable to expect that a positive proportion of the first $\lfloor \log t / \log \lambda^{-1} \rfloor$ factors are bounded away from 1, hence we expect that

$$|\widehat{\nu}_\lambda(t)| < \exp(-c \log t / \log \lambda^{-1}) = t^{-c / \log \lambda^{-1}},$$

i.e. the Fourier transform of ν_λ has an arbitrarily fast power type Fourier decay, provided λ is sufficiently close to 1. Erdős found an ingenious argument to make this heuristic rigorous.

Kahane [36] observed that Erdős's proof of Theorem 2.2 yield a bound on the Hausdorff dimension for the exceptional set of parameters (i.e. for the set of parameters, for which the Bernoulli convolution does not have m times differentiable density). See also [48, Section 6] for a detailed discussion.

Salem proved a partial converse to Theorem 2.1 by showing that $\widehat{\nu}_\lambda(t) \rightarrow 0$ as $t \rightarrow \infty$ if λ^{-1} is not a Pisot number, however, this is not strong enough information to conclude that ν_λ is absolutely continuous. For Pisot parameters, the Fourier transform of ν_λ was studied by Sarnak [54] and Sidorov and Solomyak [58]. They proved that the set of limit points of $\widehat{\nu}_\lambda(n)$ for $n \in \mathbf{Z}$ is countable, a question motivated by the spectra of multipliers.

There is another class of algebraic numbers for which the Bernoulli convolution has slow Fourier decay. A Salem number is an algebraic integer $\xi > 1$ of degree at least 4 whose Galois conjugates include ξ^{-1} and all others are on the unit circle. It is known (see e.g. [48, Section 5]) that ν_λ does not have a power-type Fourier decay if λ^{-1} is a Salem number, that is to say, $\limsup_{t \rightarrow \infty} |t^\varepsilon \widehat{\nu}_\lambda(t)| = \infty$ for every $\varepsilon > 0$. This does not imply that ν_λ is singular, but it implies that ν_λ may not have any fractional derivatives in L^1 .

Salem numbers are poorly understood. It is not known whether they form a dense subset of $\mathbf{R}_{\geq 1}$. In comparison, we know that Pisot numbers form a closed subset of the real numbers, a result due to Salem [53].

Feng and Wang [22] gave examples of Bernoulli convolutions with non-Pisot and non-Salem algebraic parameters whose density, if it exists, does not belong to L^2 .

2.2. Explicit examples of absolutely continuous Bernoulli convolutions.

In the previous section, we have seen examples of singular Bernoulli convolutions in the interesting parameter range $\lambda \in (1/2, 1)$. These are complemented by the examples of $\nu_{2^{-1/k}}$ for $k \in \mathbf{Z}_{>0}$, which are absolutely continuous, as observed by Wintner [64]. Indeed, we note that

$$\nu_{2^{-1/k}} = \nu_{1/2} * \mu$$

for some probability measure μ . To see this decomposition, consider the series defining Bernoulli convolutions and separate the terms divisible by k from the rest. Since $\nu_{1/2}$ is absolutely continuous, so is $\nu_{2^{-1/k}}$.

Garsia extended this set of examples. To state his result, we recall the following notion.

Definition 2.3. Let λ be an algebraic number and write $a \prod (x - \lambda_j)$ for its minimal polynomial in $\mathbf{Z}[X]$, i.e. $a \in \mathbf{Z}_{>0}$ is the leading coefficient and λ_j are its Galois conjugates. The Mahler measure of λ is the number

$$M_\lambda = a \prod_{j: |\lambda_j| > 1} |\lambda_j|.$$

This quantity is widely used to measure the complexity of an algebraic number.

Theorem 2.4 (Garsia [23]). *Let λ be a number such that λ^{-1} is an algebraic integer and $M_\lambda = 2$. Then ν_λ is absolutely continuous with a density in L^∞ .*

The numbers $2^{-1/k}$ all satisfy the hypothesis of the theorem, and Garsia provided the real roots of $x^{p+n} - x^n - 2$ for $\max\{p, n\} \geq 2$ as further examples. Rodemich found additional such polynomials including

$$x^3 + x^2 - x - 2, \quad x^3 - x - 2, \quad x^3 - 2x - 2, \quad x^3 - x^2 + x - 2.$$

Until recently the numbers provided by Theorem 2.4 were the only known explicit examples of absolutely continuous Bernoulli convolutions. See [27] for a study of these numbers.

Garsia's proof is based on the following ideas. He showed that these numbers are not algebraic integers themselves and hence have the property that $P(\lambda) \neq 0$ for any non-zero polynomial P that has coefficients $-1, 0$ or 1 only. This implies that the random variable $\sum_{j=0}^{n-1} X_j \lambda^j$ takes 2^n different values with equal probabilities. Indeed, if $\{a_j\}, \{b_j\} \in \{-1, 1\}^n$ are two different sequences, then

$$\sum_{j=0}^{n-1} a_j \lambda^j - \sum_{j=0}^{n-1} b_j \lambda^j = \sum_{j=0}^{n-1} (a_j - b_j) \lambda^j = 2P(\lambda) \neq 0.$$

The set of polynomials of degree at most d with coefficient $-1, 0$ or 1 play an important role in the study of Bernoulli convolutions, therefore we introduce the notation \mathcal{P}_d for it. The second ingredient in Garsia's proof is the following estimate that also plays an important role in later developments.

Theorem 2.5. *Let λ be an algebraic number and denote by m the number of its Galois conjugates on the unit circle. Then*

$$|P(\lambda)| \geq cd^{-m} M_\lambda^{-d} \quad \text{or} \quad P(\lambda) = 0$$

for any $P \in \mathcal{P}_d$, where c is a constant depending only on λ .

Garsia showed that under the hypothesis of Theorem 2.4, λ has no conjugates on the unit circle, hence there is a constant c such that consecutive points in the support of $\sum_{j=0}^{n-1} X_j \lambda^j$ are separated by at least $c2^{-n}$. Therefore the number of points that fall in a given interval of length a is at most $c^{-1}2^n a + 1$. Since each atom has probability 2^{-n} , this gives an upper bound $c^{-1}a + 2^{-n}$ for the probability that $\sum_{j=0}^{n-1} X_j \lambda^j$ is in the interval in question, and the claim follows if we take the limit $n \rightarrow \infty$.

2.3. Transversality. As we have discussed in Section 2.1, Erdős showed that ν_λ is absolutely continuous for almost all λ in an interval near 1. Whether or not the same holds on $(1/2, 1)$ was open until a remarkable paper of Solomyak [60].

Theorem 2.6 (Solomyak). *The Bernoulli convolution ν_λ is absolutely continuous and has density in L^2 for almost every $\lambda \in (1/2, 1)$.*

Solomyak's original proof has been simplified in [49]. The result has been improved by Peres and Schlag [47], who proved that the density also has fractional derivatives in L^2 and they gave estimates for the Hausdorff dimension of the exceptional set of parameters, for which ν_λ is singular.

For the details we refer to the original papers and the survey [48], which treats the topic extensively. Now we only highlight the main idea, called transversality, which can be traced back to Pollicott and Simon [50]. Denote by \mathcal{P}_∞ the set of analytic functions on the unit disk, whose Taylor series at 0 has the coefficients $-1, 0$ or 1 only. Solomyak proved that $[1/2, 2^{-2/3}]$ is an interval of transversality for \mathcal{P}_d , that is to say, any function in \mathcal{P}_d may have at most 1 zero in the interval $[1/2, 2^{-2/3}]$. Some bound on the number of zeros could be deduced from Jensen's formula, however, it is essential for the argument that there are no more than a single zero, and the proof of this fact requires much more delicate ideas.

We indicate briefly and heuristically why transversality is useful in the study of Bernoulli convolutions. Let X_0, X_1, \dots and X'_0, X'_1, \dots be two independent sequences of independent unbiased ± 1 variables and consider the random variables $\sum_{j=0}^{\infty} X_j \lambda^j$ and $\sum_{j=0}^{\infty} X'_j \lambda^j$, which both have distribution ν_λ . Fix a small number $r > 0$ and consider a subdivision of the support of ν_λ into intervals of length r . If ν_λ is singular, then most of the mass is concentrated on a small proportion of these intervals. This suggests that if ν_λ is singular then

$$\mathbf{P}\left(\left|\sum_{j=0}^{\infty} X_j \lambda^j - \sum_{j=0}^{\infty} X'_j \lambda^j\right| < r\right)$$

is 'much larger' than it would be if the distribution was absolutely continuous.

A nice way to quantify this idea is the following statement. If there is a constant C independent of r such that

$$\mathbf{P}\left(\left|\sum_{j=0}^{\infty} X_j \lambda^j - \sum_{j=0}^{\infty} X'_j \lambda^j\right| < r\right) < Cr,$$

then ν_λ is absolutely continuous with density in L^2 .

Now we fix the values X_j and X'_j and take a random value of λ uniformly in the interval $[1/2, 2^{-2/3}]$. Knowing that $\sum_{j=0}^{\infty} X_j \lambda^j - \sum_{j=0}^{\infty} X'_j \lambda^j$ may have only one zero in this interval allows us to estimate the probability of the event

$$\left|\sum_{j=0}^{\infty} X_j \lambda^j - \sum_{j=0}^{\infty} X'_j \lambda^j\right| < r$$

in terms of the first index j such that $X_j \neq X'_j$. Taking expectation over the X_j and X'_j gives the desired result. For the details of this argument see [49].

3. Generalizations and applications

Bernoulli convolutions are natural objects from several points of view including fractal geometry, dynamics and number theory. We have already seen that the arithmetic properties of λ has a decisive influence on the regularity of Bernoulli convolutions. This is a dominant feature also in the more recent results that we discuss later.

Bernoulli convolutions appear naturally in the context of dynamics. For example, Alexander and Yorke [1] described the Sinai Ruelle Bowen measure for the fat Baker's transformation in terms of Bernoulli convolutions. For further studies of dynamical properties of Bernoulli convolutions see e.g. [59] and [37].

A detailed exposition of these connections would go beyond the scope of these notes. We limit ourselves to a brief discussion of a more general framework to which Bernoulli convolutions belong. Then we briefly mention a conjecture of Breuillard on growth of groups and its connection to Bernoulli convolutions.

3.1. Self-similar and self-affine measures. Given a continuous map $T : X \rightarrow X$ on a σ -compact metric space and a measure μ on X , we denote by $T(\mu)$ the pushforward of μ , that is the unique measure that satisfies

$$\int f dT(\mu) = \int f \circ T d\mu$$

for all compactly supported continuous functions $f : X \rightarrow \mathbf{R}$.

It follows from the definition that Bernoulli convolutions satisfy the identity

$$\nu_\lambda = \frac{1}{2}T_{-1}(\nu_\lambda) + \frac{1}{2}T_1(\nu_\lambda), \quad (3)$$

where $T_j(x) = \lambda x + j$. It is not difficult to see that for any given λ , there is precisely one probability measure that satisfies this identity and this provides an alternative definition of Bernoulli convolutions.

It is easy to see from this definition that Bernoulli convolutions are always of pure type. That is, they are singular or absolutely continuous with respect to the Lebesgue measure. Indeed, if this was not the case, then both the singular and the absolutely continuous parts would satisfy (3). In fact, more is true. Mauldin and Simon [46] proved that, if non-singular, ν_λ is equivalent to the Lebesgue measure restricted to the support of ν_λ , that is to say, the Lebesgue measure is also absolutely continuous with respect to ν_λ .

This definition of Bernoulli convolutions generalizes in a straight forward manner. Let T_1, \dots, T_k be a collection of contracting similarities of Euclidean space \mathbf{R}^d and let p_1, \dots, p_k be a probability vector. Then there is a unique probability measure μ on \mathbf{R}^d that satisfies the identity

$$\mu = p_1 T_1(\mu) + \dots + p_k T_k(\mu).$$

Measures with this property are called self-similar. The requirement that the similarities T_j are contracting may be relaxed.

This concept can be further generalized if we allow the transformations T_j to be in a more general class, for example affine transformations. In the latter case, the measure satisfying the identity is called self-affine.

Certain self-affine structures have intimate connections to Bernoulli convolutions, see for example the work of Przytycki and Urbański [51].

Another class of related measures are the so-called Furstenberg measures. To define them, we consider the projective line $X = \mathbf{P}^1$ and we require T_j to be Möbius transformations. These measures play an important role in the study of random walks in Lie groups and also in the Anderson-Bernoulli model on \mathbf{Z} . See [6], [7], [3] and [32] for some recent results about these measures; some techniques used in these papers have strong analogues in the study of Bernoulli convolutions.

3.2. Growth in groups. Let G be a group and let $S \subset G$ be a finite subset. We denote by S^n the set of elements of G that can be expressed as the product of n elements of S . (Repetition of elements is allowed.)

The speed at which the sequence $|S^n|$ grows and its relation to the structure of the group G has been extensively studied in combinatorial and geometric group theory. A celebrated result of Gromov [26] asserts that $|S^n|$ grows polynomially if and only if the group generated by S is virtually nilpotent. If G is a free group and S contains at least two non-commuting elements, then $|S^n|$ grows exponentially. Grigorchuk [25] gave examples of groups of intermediate growth, that is, when $|S^n|$ grows faster than any polynomial but slower than exponential.

If $G = \mathrm{GL}_d(\mathbf{C})$, then the Tits alternative yields a duality: the growth of $|S^n|$ is either a polynomial or exponential. In this case, the exponential growth rate defined as

$$\rho(S) = \lim_{n \rightarrow \infty} \frac{\log |S^n|}{n}$$

is an important quantity of interest. Breuillard [13] has made the following conjecture.

Conjecture 3.1 (Breuillard). *For every integer $d > 0$, there is $c_d > 0$ such that*

$$\rho(S) \geq c_d \quad \text{or} \quad \rho(S) = 0$$

for any $S \subset \mathrm{GL}_d(\mathbf{C})$.

Breuillard [10], [11] and [12] proved this in the important case, when the group generated by S is not virtually solvable. See [19] and [14] for related earlier work.

Breuillard [9] also observed the importance of the special case

$$S_\lambda = \left\{ \begin{pmatrix} \lambda & 1 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} \lambda & -1 \\ 0 & 1 \end{pmatrix} \right\}. \quad (4)$$

He noted that $\rho(S_\lambda) \leq \log M_\lambda$, where M_λ is the Mahler measure of λ , which we encountered in Section 2.2. This means that Breuillard's growth conjecture implies Lehmer's conjecture, which asserts that the value of M_λ is bounded away from 1 for numbers that are not roots of unity.

The semigroup generated by S_λ is intimately related to Bernoulli convolutions. Indeed, the action of the matrices (4) on the line $(x, 1)^T \in \mathbf{R}^2$ is given by the formula $x \mapsto \lambda x \pm 1$, which are precisely the transformations that appear in (3).

Breuillard and the author [15] gave an estimate for the entropy of the random walk on the semigroup generated by S_λ in terms of M_λ . The proof utilizes measures related to Bernoulli convolutions. Using these estimates, they proved that Conjecture 3.1 is equivalent to Lehmer's conjecture. (The general case of the growth conjecture can be reduced to the case of S_λ .) The estimates for the entropy of the random walk also have consequences for the dimension of Bernoulli convolutions, which we will discuss in the next section.

4. Recent developments

4.1. Results on dimension. Hochman [29] made a recent breakthrough in the study of self-similar measures. (See also [30] for a more elementary exposition.) We discuss his results only in the special case of Bernoulli convolutions, but they are valid in greater generality, in fact, he extended them even to higher dimensions subsequently [31].

One of his results provide the following information about Bernoulli convolutions of dimension less than 1. Recall that \mathcal{P}_d denotes the set of polynomials of degree at most d with coefficients $-1, 0$ or 1 .

Theorem 4.1 (Hochman). *Suppose $\lambda \in (1/2, 1)$ is a number such that $\dim \nu_\lambda < 1$ and let $A \in \mathbf{R}_{>0}$. Then for all sufficiently large (depending on λ and A) integer d , there is a number $\xi \in \mathbf{C}$ that is a root of a polynomial in \mathcal{P}_d such that*

$$|\lambda - \xi| < \exp(-Ad).$$

In particular, the set

$$\{\lambda \in (1/2, 1) : \dim \nu_\lambda < 1\}$$

is of 0 packing dimension.

We note that the packing dimension of a set is always at least as large as its Hausdorff dimension, so the exceptional set is also of 0 Hausdorff dimension. In fact, to conclude the result for Hausdorff dimension it would be enough to know that there are infinitely many integers d such that the algebraic approximation exists at the corresponding scales. Having the approximations at all scales allowed Hochman to also derive the following corollary.

Corollary 4.2 (Hochman). *Suppose that there is a number $A > 0$ such that the following holds for every sufficiently large integer d . Let ξ_1 and ξ_2 be two numbers that are roots of polynomials in \mathcal{P}_d (not necessarily the same one). Then $|\xi_1 - \xi_2| > \exp(-Ad)$.*

Under this hypothesis, we have $\dim \nu_\lambda = 1$ for any transcendental $\lambda \in (1/2, 1)$.

Indeed, this can be deduced from the theorem as follows. If $\dim \nu_\lambda < 1$, then Theorem 4.1 yields a sequence of algebraic approximations $\{\xi_d\}_{d \geq d_0}$ such that $|\lambda - \xi_d| < \exp(-2Ad)$. By induction, and using the hypothesis on the separation between the roots of polynomials in \mathcal{P}_d , one can show that $\xi_d = \xi_{d_0}$ for all d . This leads to the conclusion that $\lambda = \xi_0$ is algebraic.

The hypothesis in the corollary is very reasonable given that the number of roots of all polynomials in \mathcal{P}_d is less than $d \cdot 3^d$. However, the best result available in the literature in this direction is the following.

Theorem 4.3 (Mahler [45]). *There is an absolute constant C such that the following holds for all sufficiently large integers d . Let ξ_1 and ξ_2 be two numbers that are roots of polynomials in \mathcal{P}_d (not necessarily the same one). Then $|\xi_1 - \xi_2| > \exp(-Cd \log d)$.*

One may take e.g. $C = 4$ in this theorem. Breuillard and the author [16] obtain the following information about Bernoulli convolutions of dimension less than 1.

Theorem 4.4 (Breuillard, Varjú). *Suppose $\lambda \in (1/2, 1)$ is a number such that $\dim \nu_\lambda < 1$. Then there are infinitely many integers d , such that there is a number $\xi \in \mathbf{R}$ that is a root of a polynomial in \mathcal{P}_d , $\dim \nu_\xi < 1$ and*

$$|\lambda - \xi| < \exp(-d^{\log^{(3)}(d)}).$$

Unlike Theorem 4.1, this result does not provide an algebraic approximation at every scale. On the other hand, Theorem 4.4 provides much smaller error terms in the approximation. Combining this with results on transcendence measures (see e.g. [62]), one can conclude that ν_λ has dimension 1 for many classical constants in the role of λ , giving the first such explicit examples among transcendental numbers.

Corollary 4.5 (Breuillard, Varjú). *We have $\dim \nu_\lambda = 1$ for any of*

$$\lambda \in \{\ln(2), e^{-1/2}, \pi/4\}.$$

Another important feature of Theorem 4.4 is that we know that $\dim \nu_\xi < 1$ for the approximants. This allows the formulation of the following corollary.

Corollary 4.6 (Breuillard, Varjú). *We have*

$$\{\lambda \in (1/2, 1) : \dim \nu_\lambda < 1\} \subset \overline{\{\xi \in (1/2, 1) \cap \overline{\mathbf{Q}} : \dim \nu_\xi < 1\}},$$

where $\overline{\mathbf{Q}}$ denotes the set of algebraic numbers and $\overline{\{\cdot\}}$ denotes closure in the standard topology of real numbers.

This result suggests that understanding the dimension of Bernoulli convolutions for algebraic parameters may lead to information about the dimension also for transcendental parameters. Here we note that the only known examples of Bernoulli convolutions for $\lambda \in (1/2, 1)$ (algebraic or transcendent) of dimension less than 1 are the inverses of Pisot numbers and that the set of Pisot numbers

is known to be closed [53]. If there are no further examples among the algebraic parameters, then there are no further examples at all.

The fact that Bernoulli convolutions have dimension less than 1 for inverses of Pisot numbers can be traced back to Garsia [24] in some form. (He proved that $h_\lambda < \log \lambda^{-1}$ with the below notation.) The dimension and finer properties of Bernoulli convolutions for special algebraic parameters (for inverses of Pisot numbers in most cases) have been studied by many authors, including [1], [2], [8], [20], [28], [33], [35], [39], [40], [41], [42] and [43].

Hochman [29] also gave a formula for the dimension of Bernoulli convolutions with algebraic parameters. This formula is in terms of the entropy of the random walk on the semigroup generated by the transformations $x \mapsto \lambda x \pm 1$; a quantity introduced by Garsia to the study of Bernoulli convolutions, which we define now. For an integer n , we denote by $\nu_\lambda^{(n)}$ the distribution of the random variable $\sum_{j=0}^{n-1} X_j \lambda^j$, where X_j are independent unbiased ± 1 valued random variables. Denoting by $H(\cdot)$ the Shannon entropy of a discrete probability measure, we define

$$h_\lambda = \lim_{n \rightarrow \infty} \frac{1}{n} H(\nu_\lambda^{(n)}).$$

(It is easy to see that the sequence $H(\nu_\lambda^{(n)})$ is subadditive, hence the limit exists and is equal to the infimum.)

Theorem 4.7 (Hochman). *Let $\lambda \in (0, 1)$ be an algebraic number. Then*

$$\dim \nu_\lambda = \min \left\{ 1, \frac{h_\lambda}{\log \lambda^{-1}} \right\}.$$

It is convenient for us to adopt the normalization that \log (which also appears in the definition of entropy) is the base 2 logarithm. The value of h_λ is maximal, when the semigroup generated by $x \mapsto \lambda x \pm 1$ is free. Then $\nu_\lambda^{(n)}$ is supported on 2^n atoms, each of which have equal weight. With our normalization, $h_\lambda = 1$ in this case. Most algebraic numbers are not roots of polynomials in \mathcal{P}_d , for example this holds for all the rationals (with the exception of ± 1 and 0) and we have $h_\lambda = 1$ for these. Hence Hochman's formula provides plenty of explicit examples of Bernoulli convolutions with dimension 1.

The quantity h_λ has been studied by Breuillard and the author [15]. They gave the following bounds in terms of the Mahler measure M_λ .

Theorem 4.8 (Breuillard, Varjú). *If λ is an algebraic number, we have*

$$0.44 \cdot \min\{1, \log M_\lambda\} \leq h_\lambda \leq \min\{1, \log M_\lambda\}.$$

We add that the upper bound is often strict. Indeed, generalizing Garsia's [23], [24] arguments, Breuillard and the author proved that $h_\lambda < \log M_\lambda$ if λ has no Galois conjugates on the unit circle.

Recall that Lehmer's conjecture asserts the existence of a constant $c > 0$ such that $M_\lambda > 1 + c$ for all algebraic numbers that are not roots of unity. Combining

this with Theorem 4.8, we obtain a positive constant lower bound for h_λ . Then Theorem 4.7 implies that $\dim \nu_\lambda = 1$ for all algebraic $\lambda \in (1 - c, 1)$ for some $c > 0$. Combining this with Corollary 4.6, we obtain the following.

Corollary 4.9 (Breuillard, Varjú). *If Lehmer's conjecture holds, then there is a positive number $c > 0$ such that $\dim \nu_\lambda = 1$ for all $\lambda \in (1 - c, 1)$.*

4.2. Results on absolute continuity. Shmerkin [55] achieved the following important result on absolute continuity of ν_λ for typical parameters.

Theorem 4.10 (Shmerkin). *The set*

$$\{\lambda \in (1/2, 1) : \nu_\lambda \text{ is singular}\}$$

is of 0 Hausdorff dimension.

Shmerkin's proof is based on his observation that the convolution of a self-similar measure of dimension 1 with another one that has a polynomial Fourier decay is absolutely continuous. He then decomposed ν_λ as the convolution of two self-similar measures in a suitable way so that he could apply Hochman's result to one of them and the result of Erdős and Kahane (discussed in Section 2.1) to the other. This argument yields not only that ν_λ is absolutely continuous (if λ is outside the exceptional set) but also that it has some fractional derivatives in L^p for some $p > 1$ (depending on λ). Shmerkin and Solomyak [56], [57] extended these ideas to more general self-similar measures.

We conclude with the following result from [61], which gives the first new examples of absolutely continuous Bernoulli convolutions since [23].

Theorem 4.11 (Varjú). *For every $\varepsilon > 0$, there is $c > 0$ such that the following holds. Let $\lambda < 1$ be an algebraic number and suppose that*

$$\lambda > 1 - c \min\{\log M_\lambda, (\log M_\lambda)^{-1-\varepsilon}\}.$$

Then ν_λ is absolutely continuous and has a density in $L \log L$.

The constant c is effective, that is, it could be computed following the steps of the proof. However, this has not been done.

Unlike Garsia's method in [23], the proof of this result is robust enough that it applies to 'biased' Bernoulli convolutions, i.e. we can allow in the definition that the random variables X_j take the values ± 1 with unequal probabilities.

Specializing the above result to rational numbers, we obtain the following.

Corollary 4.12 (Varjú). *For every $\varepsilon > 0$, there is $c > 0$ such that the following holds. Let p and q be positive integers such that*

$$p < \frac{c}{(\log q)^{1+\varepsilon}} q. \tag{5}$$

Then the Bernoulli convolution $\nu_{1-p/q}$ is absolutely continuous.

The proof relies on the observation that $|P(p/q)| \geq q^{-d}$ for all $P \in \mathcal{P}_d$. A stronger Diophantine input would allow a weaker hypothesis. In particular, a positive answer to the following question would permit us to replace (5) by $p < cq$.

Question 4.13. *Is it true that for all rational $p/q \in [9/10, 19/20]$, there is a constant $c_{p,q} > 0$ such that*

$$\#\{P \in \mathcal{P}_d : P(p/q) < c_{p,q} \exp(-Cd)\} < \exp(d/100)$$

holds for some absolute constant C ?

5. Ideas from the proofs

We present a mixture of ideas from the papers [29], [61] and [16].

5.1. Entropy. Using entropy to study Bernoulli convolutions goes back to the work of Garsia [24]. In his breakthrough [29], Hochman brought in ideas from additive combinatorics to estimate the growth of entropy under convolution of measures. In this section, we explain the variant of entropy that we work with and explain their basic properties. In the next two sections, we give a brief account of the main ideas from the proofs of some of the results we mentioned in Section 4.

Definition 5.1. Let X be a bounded random variable and let $r > 0$ be a number. The entropy of X at scale r is defined by the formula

$$H(X; r) = \int_0^1 H(\lfloor r^{-1}X + t \rfloor) dt,$$

where $H(\cdot)$ denotes the Shannon entropy of a discrete random variable. We also define the conditional entropy between two scales r_1 and r_2 by the formula

$$H(X; r_1|r_2) = H(X; r_1) - H(X; r_2).$$

If μ is the distribution of X , we write $H(\mu; r) = H(X; r)$ and $H(\mu; r_1|r_2) = H(X; r_1|r_2)$.

In words, $H(X; r)$ is defined as follows. We take a partition of the real line into intervals of length r , and compute the Shannon entropy of X with respect to this partition. However, the choice of the partition is not unique, so we take an average over all possible partitions, and define this to be the quantity $H(X; r)$.

The main purpose of this averaging procedure (which goes back to Wang [63]) is technical convenience, because it endows $H(\mu; r)$ with some useful properties that would otherwise fail. In particular, we have

$$0 \leq H(\mu; r_1|r_2) \leq H(\mu; s_1|s_2),$$

whenever $s_1 \leq r_1 \leq r_2 \leq s_2$. That is to say, increasing the gap between the scales may only increase conditional entropy. In addition,

$$H(\mu * \nu; r_1|r_2) \geq H(\mu; r_1|r_2), \tag{6}$$

whenever r_2/r_1 is an integer. That is to say, convolution may only increase the conditional entropy between two scales of integral ratio. Of course, these properties also hold without the averaging procedure at the expense of introducing an absolute constant error term. However, in certain arguments such an error could not be spared. In others, this is not essential. In particular [29] does not use averaging.

5.2. Dimension 1. In this section, we explain some ideas from the proofs of the results that we discussed in Section 4.1.

The entropy dimension of a measure μ with bounded support in \mathbf{R} is defined

$$\lim_{n \rightarrow \infty} \frac{1}{n} H(\mu; 2^{-n}), \quad (7)$$

if the limit exists. If μ is exact dimensional (which is the case for Bernoulli convolutions [21]), then the limit (7) exists and is equal to $\dim \mu$.

Therefore, if $\dim \nu_\lambda < 1$ for some parameter λ , then there is a number $c > 0$ such that $H(\nu_\lambda; 2^{-n}) < (1 - c)n$ for all sufficiently large n . Moreover, one can prove (see [16] for the details) that

$$H(\nu_\lambda; r|2^r) < 1 - c' \quad (8)$$

for all $r > 0$, where $c' > 0$ is another positive constant.

We introduce one more piece of notation that we use in our discussions. For a bounded set $I \subset \mathbf{R}_{>0}$, we denote by ν_λ^I the distribution of the random variable

$$\sum_{j: \lambda^j \in I} X_j \lambda^j.$$

The following observation is key to the arguments. For any partition $I_1 \dot{\cup} \dots \dot{\cup} I_n = (0, 1]$, we have

$$\nu_\lambda = \nu_\lambda^{I_1} * \dots * \nu_\lambda^{I_n}.$$

The proofs of the results on dimension are indirect. We assume that $\dim \nu_\lambda < 1$, yet the conclusion of the theorem in question fails. (In case of Theorem 4.7, this conclusion is that $\dim \nu_\lambda = h_\lambda \log \lambda^{-1}$.) The first step of the proof is to find some non-trivial lower bound on $H(\nu_\lambda^I; r_1|r_2)$ for an appropriate set $I \subset \mathbf{R}_{>0}$ and for appropriate scales r_1 and r_2 . We formulate two such results, which are related to the settings of Theorems 4.1 and 4.7 respectively.

Lemma 5.2. *Let $\lambda \in (1/2, 1)$ and $A > 0$ be numbers. Suppose that $|\lambda - \xi| > \exp(-Ad)$ for all roots ξ of polynomials in \mathcal{P}_{d-1} for some integer d . Then there is a positive integer B depending only on A and λ such that*

$$H(\nu_\lambda^{(\lambda^d, 1]}; \lambda^{Bd}) = d.$$

Lemma 5.3. *Let $\lambda \in (1/2, 1)$ be an algebraic number. Then there is a positive integer B such that*

$$H(\nu_\lambda^{(\lambda^d, 1]}; \lambda^{Bd}) \geq dh_\lambda$$

holds for all positive integers d .

Sketch proof of Lemma 5.2. The proof begins with the observation that the assumption $|\lambda - \xi| > \exp(-Ad)$ for all roots ξ of a polynomial $P \in \mathcal{P}_{d-1}$ implies that $|P(\lambda)| > \lambda^{Bd}$ for some number B ; a fact closely related to the idea of ‘transversality’ discussed in Section 2.3.

To see this, we first note that there is a number m depending only on λ such that each polynomial $P \in \mathcal{P}_{d-1}$ has at most m roots of modulus less than $(1+\lambda)/2$. It is key that m is independent of d . Such a bound can be deduced from Jensen’s formula. Alternatively, one can argue by contradiction and show that a putative sequence of polynomials with unbounded number of roots would converge along a subsequence to an analytic function with infinitely many zeros in a compact subset of its domain.

We factorize P and separate the contribution of the zeros of modulus less than $(1+\lambda)/2$. We obtain

$$|P(\lambda)| = \prod_{\xi: P(\xi)=0} |\lambda - \xi| \geq \exp(-Amd) \left(\frac{1-\lambda}{2} \right)^{d-1-m} \geq \lambda^{Bd}$$

provided B is large enough so that $\lambda^B < \exp(-Am)(1-\lambda)/2$.

We note that the difference between any two numbers of the form $\sum_{j=0}^{d-1} \pm \lambda^j$ is $2P(\lambda)$ for some $P \in \mathcal{P}_{d-1}$. Hence the random variable

$$\left\lfloor \lambda^{-Bd} \sum_{j=0}^{d-1} X_j \lambda^j + t \right\rfloor$$

takes 2^d different values with equal probability for any t . This proves the claim. \square

Sketch proof of Lemma 5.3. This lemma relies on Garsia’s estimate (Theorem 2.5) on the separation between two distinct numbers of the form $\sum_{j=0}^{d-1} \pm \lambda^j$.

If we set B sufficiently large, that estimate implies that the random variable

$$\left\lfloor \lambda^{-Bd} \sum_{j=0}^{d-1} X_j \lambda^j + t \right\rfloor$$

has the same entropy as $\sum_{j=0}^{d-1} X_j \lambda^j$. Thus

$$H\left(\nu_\lambda^{(\lambda^d, 1]}; \lambda^{Bd}\right) \geq H\left(\sum_{j=0}^{d-1} X_j \lambda^j\right) \geq dh_\lambda,$$

and the claim follows. \square

Recall our standing assumption $\dim \nu_\lambda < 1$. Via (8) this implies

$$H(\nu_\lambda^{(\lambda^d, 1]}; \lambda^d) < (1 - c') \cdot \log(\lambda^{-d}) + C$$

for a C depending on λ (more precisely on the diameter of $\text{supp } \nu_\lambda$). We combine this with the conclusion of Lemma 5.2 and find that there is $\beta > 0$ such that

$$H(\nu_\lambda^{(\lambda^d, 1]}, \lambda^{Bd} | \lambda^d) \geq \beta \cdot \log(\lambda^{-d(B-1)}) \quad (9)$$

provided the hypotheses of the lemma is satisfied. In the setting of Lemma 5.3 we can obtain the same conclusion provided $h_\lambda \geq \log \lambda^{-1}$.

Using scaling properties of entropy, we note that (9) also yields

$$H(\nu_\lambda^{(\lambda^{d+a}, \lambda^a]}, \lambda^{Bd+a} | \lambda^{d+a}) \geq \beta \cdot \log(\lambda^{-d(B-1)})$$

for any integer $a \geq 0$.

In the next step of the proofs, we exploit the idea that convolution increases entropy to improve on the bound (9). We noted above that convolution may not decrease entropy (at least not between scales of integral ratio), however, we need now a stronger result, which says that we can obtain a definite entropy increase. We recall the following result from [61], which is a quantitative strengthening of Hochman's original estimate [29].

Theorem 5.4. *For every $\alpha > 0$, there are $C, c > 0$ such that the following holds. Let μ, ν be two compactly supported probability measures on \mathbf{R} . Let $s_1 > s_2 > 0$ and $\beta > 0$ be real numbers. Suppose that*

$$H(\mu; s|2s) < 1 - \alpha \quad (10)$$

for all $s_2 < s < s_1$. Suppose further that

$$H(\nu; s_2 | s_1) > \beta \cdot (\log s_1 - \log s_2).$$

Then

$$H(\mu * \nu; s_2 | s_1) > H(\mu; s_2 | s_1) + c\beta \cdot (\log \beta^{-1})^{-1} (\log s_1 - \log s_2) - C.$$

This result can be thought of as an entropy analogue of the additive part in the proof of Bourgain's discretized sum product theorem [4], [5]. Another variant of Theorem 5.4 can be found in [44]. The proof is beyond the scope of this note. Both Hochman's original proof [29] and the proofs in [44] and [61] are based on a multiscale argument. Hochman used the Berry-Esseen inequalities at multiple scales, while [44] and especially [61] are closer to the ideas of Bourgain.

Observe that this formulation of the theorem is well-adapted for the application of proving dimension 1 for Bernoulli convolutions. Indeed, equations (8) and (6) imply that condition (10) always holds for ν_λ^I with some constant α independent of $I \subset (0, 1]$ provided $\dim \nu_\lambda < 1$.

After these preparations, Theorems 4.1 and 4.7 are reduced to the following.

Theorem 5.5 (Hochman). *Let $\lambda \in (1/2, 1)$ be a number. Suppose that there are $B \in \mathbf{Z}_{>0}$ and $\beta > 0$ such that there are infinitely many integers d that satisfy*

$$H(\nu_\lambda^{(\lambda^d, 1]}, \lambda^{Bd} | \lambda^d) > \beta \cdot \log(\lambda^{-d(B-1)}). \quad (11)$$

Then $\dim \nu_\lambda = 1$.

Sketch proof. Suppose to the contrary that (11) holds, yet $\dim \nu_\lambda < 1$. We fix a small number $\varepsilon > 0$ that we will specify later. Since the limit

$$\gamma = \lim_{d \rightarrow \infty} \frac{1}{d} H(\nu_\lambda; \lambda^d) = \lim_{d \rightarrow \infty} \frac{1}{d} H(\nu_\lambda^{(\lambda^d, 1]}; \lambda^d | 1)$$

exists, there is a number D such that

$$\left| \frac{1}{d} H(\nu_\lambda^{(\lambda^d, 1]}; \lambda^d | 1) - \gamma \right| < \varepsilon \quad (12)$$

for all $d > D$. We have $\gamma = \dim \nu_\lambda \log \lambda^{-1}$, but we do not need to know this.

Let $d > D$ be a number such that (11) holds for this d . We apply (12) with $(B-1)d$ in place of d and scale it by a factor of λ^d :

$$H(\nu_\lambda^{(\lambda^{Bd}, \lambda^d]}; \lambda^{Bd} | \lambda^d) \geq (\gamma - \varepsilon) \cdot (B-1)d. \quad (13)$$

We apply Theorem 5.4 for the measures $\mu = \nu_\lambda^{(\lambda^{Bd}, \lambda^d]}$ and $\nu = \nu_\lambda^{(\lambda^d, 1]}$ between the scales $s_2 = \lambda^{Bd}$ and $s_1 = \lambda^d$. Using (11) and (13), we obtain

$$H(\nu_\lambda^{(\lambda^{Bd}, 1]}; \lambda^{Bd} | \lambda^d) > (\gamma - \varepsilon + c\beta(\log \beta^{-1})^{-1} \log \lambda^{-1}) \cdot (B-1)d - C.$$

We use (12) again and write

$$H(\nu_\lambda^{(\lambda^{Bd}, 1]}; \lambda^d | 1) \geq H(\nu_\lambda^{(\lambda^d, 1]}; \lambda^d | 1) - C \geq (\gamma - \varepsilon) \cdot d - C.$$

We combine this with our previous estimate and find

$$H(\nu_\lambda^{(\lambda^{Bd}, 1]}; \lambda^{Bd} | 1) > (\gamma - \varepsilon + c\beta(\log \beta^{-1})^{-1} \log(\lambda^{-1})(B-1)/B) \cdot Bd - C.$$

We take

$$2\varepsilon < c\beta(\log \beta^{-1})^{-1} \log(\lambda^{-1})(B-1)/B,$$

which leads to a contradiction with (12). \square

We briefly comment on the proof of Theorem 4.4. The argument again begins with considerations about the separation between the points in the support of the measure $\nu_\lambda^{(\lambda^d, 1]}$. Consider the set of polynomials in \mathcal{P}_{d-1} that take very small values (less than d^{-Cd}) at λ . Using Diophantine considerations related to Theorem 4.3, one can show that these polynomials have a common root ξ such that $|\xi - \lambda| < d^{-4d}/2$. Hence we obtain an upper bound on $H(\nu_\xi^{(d)})$ in terms of $H(\nu_\lambda^{(\lambda^d, 1]}; d^{-Cd})$.

We can deduce from this argument that either $h_\xi < \log \xi^{-1}$ or

$$H(\nu_\lambda^{(\lambda^d, 1]}; d^{-Cd}; \lambda^d) \geq cd = \frac{c}{C \log d + \log \lambda} \cdot \log(d^{Cd} \lambda^d) \quad (14)$$

holds for suitable numbers c and C that may depend on λ , but which are independent of d .

If there is no ξ that is a root of a polynomial in \mathcal{P}_{d-1} , such that $|\xi - \lambda| < d^{-4d}/2$ and $h_\xi < \log \xi^{-1}$, then (14) holds for d . If there is such a ξ , then we set d' to be the smallest integer such that $|\xi - \lambda| \geq d'^{-4d'}/2$. By Theorem 4.3, any root ξ' of a polynomial in $\mathcal{P}_{d'-1}$ with $|\xi - \xi'| < (d' - 1)^{-4(d'-1)}$ satisfies $\xi = \xi'$. Therefore, there is no root ξ' of a polynomial in $\mathcal{P}_{d'-1}$ that satisfy $|\xi' - \lambda| < d'^{-4d'}/2$, and by the previous argument, (14) holds with d' in place of d .

This argument provides a sequence of integers d such that (14) holds. We can control the gaps in this sequence in terms of the distances of λ from algebraic numbers ξ that are roots of polynomials in \mathcal{P}_d and satisfy $h_\xi < \log \lambda^{-1}$.

In the next step, we apply Theorem 5.4 to improve on (14). However, this time, the parameter β that we can use in Theorem 5.4 tends to 0, as d grows. This has a number of consequences.

First, the quantitative aspects of Theorem 5.4 become important. (The proof of Theorem 5.5 was not sensitive to the amount of entropy gained in Theorem 5.4, as long as it was a positive constant times the logarithm of the ratio of the scales.)

Second, we have to apply Theorem 5.4 repeatedly sufficiently many times until the combined contribution of the convolution factors to the entropy gain becomes significant. These calculations are delicate. To illustrate this, we note that the success of the argument ultimately relies on the fact that the series $\sum (n \log n \log \log n)^{-1}$ diverges.

For the appropriate decomposition of ν_λ as a convolution and for the details of the calculation, we refer the reader to the paper [16].

5.3. Absolute continuity. The argument discussed in the previous section can be adapted to show that under the relevant hypothesis in the setting of Theorems 4.1, 4.4 or 4.7 we have

$$\lim_{d \rightarrow \infty} H(\nu_\lambda; 2^{-d} | 2^{-d+1}) = 1. \quad (15)$$

We have seen that this implies $\dim \nu_\lambda = 1$.

In order that we can conclude that ν_λ is absolutely continuous, we need to find a rate for the speed of convergence in (15). Indeed, Garsia [24] observed that ν_λ is absolutely continuous if the sequence $d - H(\nu_\lambda; 2^{-d})$ is bounded. Moreover, this condition implies that the density of ν_λ belongs to the Orlicz space $L \log L$. Therefore, it is enough for us to show that the series

$$\sum (1 - H(\nu_\lambda; 2^{-d} | 2^{-d+1}))$$

is convergent.

The proof of Theorem 4.11 follows a similar strategy as discussed in the previous section, but now we cannot rely on Theorem 5.4 alone, because that provides no control on the entropy gain if the parameter α approaches 0. For this reason, we need the following estimate.

Theorem 5.6. *There is an absolute constant $C > 0$ such that the following holds. Let $\mu, \tilde{\mu}$ be two compactly supported probability measures on \mathbf{R} and let $\alpha, r > 0$ be*

real numbers. Suppose that

$$H(\mu; s|2s) \geq 1 - \alpha \quad \text{and} \quad H(\tilde{\mu}; s|2s) \geq 1 - \alpha$$

for all s with $|\log r - \log s| < C \log \alpha^{-1}$.

Then

$$H(\mu * \tilde{\mu}; r|2r) \geq 1 - C(\log \alpha^{-1})^3 \alpha^2.$$

A suitable decomposition of ν_λ as a convolution of measures of the form ν_λ^I and then a repeated application of Theorems 5.4 and 5.6 leads to the estimate

$$H(\nu_\lambda; 2^{-n}|2^{-n+1}) = 1 + O(n^{-a(\lambda)}),$$

whenever λ is an algebraic number such that $h_\lambda > \log \lambda^{-1}$, where $a(\lambda) > 0$ is a number that depends on λ . When $a(\lambda) > 1$, this is enough to conclude that ν_λ is absolutely continuous. The analysis of this inequality leads to the conditions imposed in Theorem 4.11. The details of these calculations are beyond the scope of this note and the interested reader may consult the original paper [61].

References

- [1] J. C. Alexander and J. A. Yorke, Fat baker's transformations. *Ergodic Theory Dynam. Systems* **4** (1984), no. 1, 1–23.
- [2] J. C. Alexander and D. Zagier, The entropy of a certain infinitely convolved Bernoulli measure. *J. London Math. Soc. (2)* **44** (1991), no. 1, 121–134.
- [3] B. Bárány, M. Pollicott and K. Simon, Stationary measures for projective transformations: the Blackwell and Furstenberg measures. *J. Stat. Phys.* **148** (2012), no. 3, 393–421.
- [4] J. Bourgain, On the Erdős-Volkmann and Katz-Tao ring conjectures. *Geom. Funct. Anal.* **13** (2003), no. 2, 334–365.
- [5] J. Bourgain, The discretized sum-product and projection theorems. *J. Anal. Math.* **112** (2010), 193–236.
- [6] J. Bourgain, Finitely supported measures on $SL_2(\mathbb{R})$ which are absolutely continuous at infinity. In *Geometric aspects of functional analysis*, Lecture Notes in Math., 2050, Springer, Heidelberg, 2012, 133–141.
- [7] J. Bourgain, On the Furstenberg measure and density of states for the Anderson-Bernoulli model at small disorder. *J. Anal. Math.* **117** (2012), 273–295.
- [8] A. Bovier, Bernoulli convolutions, dynamical systems and automata. In *Disordered systems (Temuco, 1991/1992)*, Travaux en Cours, 53, Hermann, Paris, 1996, 63–86.
- [9] E. Breuillard, On uniform exponential growth for solvable groups. *Pure Appl. Math. Q.* **3** (2007), no. 4, part 1, 949–967.
- [10] E. Breuillard, A strong Tits alternative. Preprint. arXiv:0804.1395.
- [11] E. Breuillard, A height gap theorem for finite subsets of $GL_d(\overline{\mathbb{Q}})$ and nonamenable subgroups. *Ann. of Math. (2)* **174** (2011), no. 2, 1057–1110.

- [12] E. Breuillard, Heights on SL_2 and free subgroups. In *Geometry, rigidity, and group actions*, Chicago Lectures in Math., Univ. Chicago Press, Chicago, IL, 2011, 455–493.
- [13] E. Breuillard, Diophantine geometry and uniform growth of finite and infinite groups. In *Proceedings of the International Congress of Mathematicians, Seoul 2014, Volume 3* Kyung Moon SA, 2014, 27–50.
- [14] E. Breuillard and T. Gelander, Uniform independence in linear groups. *Invent. math.* **173** (2008), no. 2, 225–263.
- [15] E. Breuillard and P. P. Varjú, Entropy of Bernoulli convolutions and uniform exponential growth for linear groups. Preprint. arXiv:1510.04043.
- [16] E. Breuillard and P. P. Varjú, On the dimension of Bernoulli convolutions. Work in progress.
- [17] P. Erdős, On a family of symmetric Bernoulli convolutions. *Amer. J. Math.* **61** (1939), 974–976.
- [18] P. Erdős, On the smoothness properties of a family of Bernoulli convolutions. *Amer. J. Math.* **62** (1940), 180–186.
- [19] A. Eskin, S. Mozes and H. Oh, On uniform exponential growth for linear groups. *Invent. math.* **160** (2005), no. 1, 1–30.
- [20] D.-J. Feng, Multifractal analysis of Bernoulli convolutions associated with Salem numbers. *Adv. Math.* **229** (2012), no. 5, 3052–3077.
- [21] D.-J. Feng and H. Hu, Dimension theory of iterated function systems. *Comm. Pure Appl. Math.* **62** (2009), no. 11, 1435–1500.
- [22] D.-J. Feng and Y. Wang, Bernoulli convolutions associated with certain non-Pisot numbers. *Adv. Math.* **187** (2004), no. 1, 173–194.
MR2889154 (Reviewed)
- [23] A. M. Garsia, Arithmetic properties of Bernoulli convolutions. *Trans. Amer. Math. Soc.* **102** (1962), 409–432.
- [24] A. M. Garsia, Entropy and singularity of infinite convolutions. *Pacific J. Math.* **13** (1963), 1159–1169.
- [25] R. I. Grigorchuk, On the Milnor problem of group growth. (Russian) *Dokl. Akad. Nauk SSSR* **271** (1983), no. 1, 30–33.
- [26] M. Gromov, Groups of polynomial growth and expanding maps. *Inst. Hautes Études Sci. Publ. Math.* **53** (1981), 53–73.
- [27] K. G. Hare and M. Panju, Some comments on Garsia numbers. *Math. Comp.* **82** (2013), no. 282, 1197–1221.
- [28] K. G. Hare and N. Sidorov, A lower bound for Garsia’s entropy for certain Bernoulli convolutions. *LMS J. Comput. Math.* **13** (2010), 130–143.
- [29] M. Hochman, On self-similar sets with overlaps and inverse theorems for entropy. *Ann. of Math. (2)* **180** (2014), no. 2, 773–822.
- [30] M. Hochman, Self similar sets, entropy and additive combinatorics. In *Geometry and analysis of fractals*, Springer Proc. Math. Stat., 88, Springer, Heidelberg, 2014. 225–252
- [31] M. Hochman, On self-similar sets with overlaps and inverse theorems for entropy in \mathbb{R}^d . Preprint. arXiv:1503.09043.

- [32] M. Hochman, B. Solomyak, Work in progress.
- [33] T.-Y. Hu, The local dimensions of the Bernoulli convolution associated with the golden number. *Trans. Amer. Math. Soc.* **349** (1997), no. 7, 2917–2940.
- [34] B. Jessen and A. Wintner, Distribution functions and the Riemann zeta function. *Trans. Amer. Math. Soc.* **38** (1935), no. 1, 48–88.
- [35] T. Jordan, P. Shmerkin, and B. Solomyak, Multifractal structure of Bernoulli convolutions. *Math. Proc. Cambridge Philos. Soc.* **151** (2011), no. 3, 521–539.
- [36] J.-P. Kahane, Sur la distribution de certaines séries aléatoires. (French) In *Colloque de Thorie des Nombres (Univ. Bordeaux, Bordeaux, 1969)*, Bull. Soc. Math. France, Mém. No. 25, Soc. Math. France Paris, 1971, 119–122.
- [37] T. Kempton and T. Persson, Bernoulli convolutions and 1D dynamics. *Nonlinearity* **28** (2015), no. 11, 3921–3934.
- [38] R. Kershner and A. Wintner, On Symmetric Bernoulli Convolutions. *Amer. J. Math.* **57** (1935), no. 3, 541–548.
- [39] S. P. Lalley, Random series in powers of algebraic integers: Hausdorff dimension of the limit distribution. *J. London Math. Soc.* (2) **57** (1998), no. 3, 629–654.
- [40] K.-S. Lau, Dimension of a family of singular Bernoulli convolutions. *J. Funct. Anal.* **116** (1993), no. 2, 335–358.
- [41] K.-S. Lau and S.-M. Ngai, L^q -spectrum of the Bernoulli convolution associated with the golden ratio. *Studia Math.* **131** (1998), no. 3, 225–251.
- [42] K.-S. Lau and S.-M. Ngai, L^q -spectrum of Bernoulli convolutions associated with P. V. numbers. *Osaka J. Math.* **36** (1999), no. 4, 993–1010.
- [43] F. Ledrappier and A. Porzio, A dimension formula for Bernoulli convolutions. *J. Statist. Phys.* **76** (1994), no. 5-6, 1307–1327.
- [44] E. Lindenstrauss and P. P. Varjú, Work in progress.
- [45] K. Mahler, An inequality for the discriminant of a polynomial. *Michigan Math. J.* **11** (1964), 257–262.
- [46] R. D. Mauldin and K. Simon, The equivalence of some Bernoulli convolutions to Lebesgue measure. *Proc. Amer. Math. Soc.* **126** (1998), no. 9, 2733–2736.
- [47] Y. Peres and W. Schlag, Smoothness of projections, Bernoulli convolutions, and the dimension of exceptions. *Duke Math. J.* **102** (2000), no. 2, 193–251.
- [48] Y. Peres, W. Schlag and B. Solomyak, Sixty years of Bernoulli convolutions. In *Fractal geometry and stochastics, II (Greifswald/Koserow, 1998)*, Progr. Probab. 46, Birkhäuser, Basel, 2000, 39–65.
- [49] Y. Peres and B. Solomyak, Absolute continuity of Bernoulli convolutions, a simple proof. *Math. Res. Lett.* **3** (1996), no. 2, 231–239.
- [50] M. Pollicott and K. Simon, The Hausdorff dimension of λ -expansions with deleted digits. *Trans. Amer. Math. Soc.* **347** (1995), no. 3, 967–983.
- [51] F. Przytycki and M. Urbański, On the Hausdorff dimension of some fractal sets. *Studia Math.* **93** (1989), no. 2, 155–186.
- [52] R. Salem, Sets of uniqueness and sets of multiplicity. *Trans. Amer. Math. Soc.* **54**, (1943), 218–228.

- [53] R. Salem, A remarkable class of algebraic integers. Proof of a conjecture of Vijayaraghavan. *Duke Math. J.* **11**, (1944), 103–108.
- [54] P. Sarnak, Spectra of singular measures as multipliers on L^p . *J. Funct. Anal.* **37** (1980), no. 3, 302–317.
- [55] P. Shmerkin, On the exceptional set for absolute continuity of Bernoulli convolutions. *Geom. Funct. Anal.* **24** (2014), no. 3, 946–958.
- [56] P. Shmerkin and B. Solomyak, Absolute continuity of self-similar measures, their projections and convolutions. *Trans. Amer. Math. Soc.* **368** (2016), no. 7, 5125–5151.
- [57] P. Shmerkin and B. Solomyak, Absolute continuity of complex Bernoulli convolutions. *Math. Proc. Cambridge Philos. Soc.* To appear.
- [58] N. Sidorov and B. Solomyak, Spectra of Bernoulli convolutions as multipliers in L^p on the circle. *Duke Math. J.* **120** (2003), no. 2, 353–370.
- [59] N. Sidorov and A. Vershik, Ergodic properties of the Erdős measure, the entropy of the golden shift, and related problems. *Monatsh. Math.* **126** (1998), no. 3, 215–261.
- [60] B. Solomyak, On the random series $\sum \pm \lambda^n$ (an Erdős problem). *Ann. of Math. (2)* **142** (1995), no. 3, 611–625.
- [61] P. P. Varjú, Absolute continuity of Bernoulli convolutions for algebraic parameters. Preprint. arXiv:1602.00261.
- [62] M. Waldschmidt, Transcendence measures for exponentials and logarithms. *J. Austral. Math. Soc. Ser. A* **25** (1978), no. 4, 445–465.
- [63] Z. Wang, Quantitative density under higher rank abelian algebraic toral actions. *Int. Math. Res. Not. IMRN* 2011, no. 16, 3744–3821.
- [64] A. Wintner, On Convergent Poisson Convolutions. *Amer. J. Math.* **57** (1935), no. 4, 827–838.

Péter P. Varjú, DPMMS, University of Cambridge, Wilberforce Road, Cambridge CB3 0WB, UK

E-mail: pv270@dpmms.cam.ac.uk