Mathematisches Forschungsinstitut Oberwolfach

Report No. 47/2017

DOI: 10.4171/OWR/2017/47

Arbeitsgemeinschaft: Additive Combinatorics, Entropy, and Fractal Geometry

Organised by
Emmanuel Breuillard, Münster
Mike Hochman, Jerusalem
Pablo Shmerkin, Buenos Aires

8 October – 13 October 2017

Abstract. The aim of the workshop was to survey recent developments in fractal geometry, specifically those related to projections and slices of planar self-similar sets, and dimension and absolute continuity of self-similar measures on the line, in particular Bernoulli convolutions. The methods combine ergodic theory, additive combinatorics, and algebraic number theory. Talks were high-level descriptions of the results, aimed at a mixed audience with minimal background in real analysis, ergodic theory and dimension theory.

Mathematics Subject Classification (2010): 28A80, 28A75.

Introduction by the Organisers

The workshop *Additive Combinatorics, Entropy, and Fractal Geometry*, organized by Emmanuel Breuillard (Münster), Michael Hochman (Hebrew University) and Pablo Shmerkin (Universidad Torcuato Di Tella), was well attended by a mix of graduate students, postgraduates and senior mathematicians. It included 21 talks ranging from introductory lectures to presentations of very recent results. The program was intended to give a broad view of the subject, so as to be suitable to people new to the area and also to experts interested in the most recent developments.

The program was divided into four main parts.

Introductory material: Introducing self-similar sets and measures, and basic facts about dimension and entropy.

Ergodic-theoretic methods: This part covered in some detail the dimension conservation theorem of Furstenberg for self-homothetic sets, the projection theorem of Shmerkin-Peres and Hochman-Shmerkin for self-similar sets with dense rotations and products of non-commensurable self-similar sets in the line, and Wu's proof of Furstenberg's slice conjecture for such sets. Underlying all of these is Furstenberg's notion of a CP-processes, a dynamical system capturing the small-scale structure of sets in Euclidean space; this was first introduced. The local entropy averages method was also discussed.

Entropy, algebra and additive combinatorics: This part was devoted to Hochman's work on dimension of self-similar sets and measures with overlaps, Shmerkin's theorem on smoothness of Bernoulli convolutions, and the work of Varjú and Breuillard-Varjú on smoothness and dimension of Bernoulli convolutions. A major tool here are inverse theorems for entropy, which give conditions under which entropy of a convolution grows relative to its factors. Both Hochman's and Varjú's inverse theorems were discussed. A historical introduction to Bernoulli convolutions was given discussing a variety of classical results, and the role Fourier methods and the Erdös-Kahane argument were discussed.

Other methods: The final sequence of talks was devoted to Bourgain's sum-product and projection results, and Shmerkin's slice theorems. These rely more heavily on techniques from "combinatorial" additive combinatorics, such as Plünnecke-Ruzsa inequality and the Balog-Szemerédi-Gowers theorem and its asymmetrical variant. Bourgain's construction of "good" subtrees was discussed, as well as background material on L^q-spectrum of measures and associated notions.

Although the pace was rapid, the organizers felt that the level of the talks was very good, and they would thank the speakers for the careful and often no easy preparation that went into them.

Acknowledgement: The MFO and the workshop organizers would like to thank the National Science Foundation for supporting the participation of junior researchers in the workshop by the grant DMS-1049268, "US Junior Oberwolfach Fellows".

Arbeitsgemeinschaft: Additive Combinatorics, Entropy, and Fractal Geometry

Table of Contents

Abstracts

Self-similar sets and measures

Károly Simon

This was an introductory lecture. The aim of the talk was to give an overall picture about the most important properties of self-similar sets and measures (the major references are: [1], [4]) and to provide the context of some of the most important conjectures and recent results (the major references are: [3], [2] and [6]).

Self-Similar Iterated Function Systems and separation conditions

A self-similar Iterated Function System (SSIFS) is a finite list of contracting similarity transformations on $\mathbb{R}^d$. That is let $m \geq 2$ and $O_1, \ldots, O_m \in O(d)$ orthogonal matrices and $r_1, \ldots, r_m \in (0,1)$ and $t_1, \ldots, t_m \in \mathbb{R}^d$. Then

$$\mathcal{S} := \{S_i(x) = r_i \cdot O_i x + t_i\}_{i=1}^m \tag{1}$$

is called a self-similar Iterated Function System on $\mathbb{R}^d$. Then there exists a unique non-empty compact set Λ which satisfies $\Lambda = \bigcup_{i=1}^m S_i(\Lambda)$. That is why we say that Λ is a self-similar set. We call Λ the attractor of the SSIFS $\mathcal{S}$. There is a natural coding of the elements $x \in \Lambda$ by the infinite sequences above the alphabet $\{1, \ldots, m\}$. Namely, for an $x \in \Lambda$ we can easily find at least one code $\mathbf{i} \in \Sigma = \{1, \ldots, m\}^{\mathbb{N}}$ such that $x = \Pi(\mathbf{i})$, where

$$\Pi(\mathbf{i}) := \lim_{n \to \infty} S_{i_1} \circ \cdots \circ S_{i_n}(0). \tag{2}$$

The natural coding Π is called natural projection from the symbolic space Σ to the attractor Λ. If the coding is unique (that is $S_i(\Lambda) \cap S_j(\Lambda) = \emptyset$) then we say that the Strong Separation Condition (SSC) holds. A less strong requirement for the cylinders being well separated is the so-called Open Set Condition (OSC). This holds if there is a non-empty open set V satisfying $S_i(V) \subset V$ for all i and $S_i(V) \cap S_j(V) = \emptyset$ for all $j \neq j$.

Similarity dimension of an SSIFS and exact overlap condition

In this case, Hutchinson's Theorem states that the Hausdorff- and box-dimensions (which are identical always for any self-similar attractors), can be computed as the unique solution of the equation $\sum_{i=1}^{m} r_i^s = 1$, where we remind that $r_i \in (0,1)$ was the contraction ratio of the map S_i. Moreover, the s-dimensional Hausdorff measure of Λ is positive and finite. The solution of this equation, $\sum_{i=1}^{m} r_i^s = 1$, is the similarity dimension of the SSIFS $\mathcal{S}$ and we denote it by $\dim_{\mathrm{S}}(\mathcal{S})$. Note, that the similarity dimension depends only on the contraction ratios $\{r_i\}_{i=1}^m$ of the similarity transformations $\{S_i\}_{i=1}^m$. It is easy to see that $\dim_{\mathrm{H}} \Lambda \leq \dim_{\mathrm{S}} \mathcal{S}$ holds always. However, it is easy to construct examples where due to heavy overlaps between the cylinders $\Lambda_i := S_i(\Lambda)$, there is a dimension drop that is $\dim_{\mathrm{H}}(\Lambda) < \dim_{\mathrm{S}}(\mathcal{S})$. To study the properties of SSIFS with heavy overlaps M. Keane's introduced the

$\{0,1,3\}$ family: $\mathcal{S}^\lambda := \{\lambda x, \lambda x+1, \lambda x+3\}$ where the parameter of the family is $\lambda \in \left(\frac{1}{4}, \frac{1}{3}\right)$. It is straightforward that the similarity dimension is $s(\lambda) = \frac{\log 3}{-\log \lambda}$.

In this case, using the so-called transversality method, it was proved that the SSC does not hold, but for almost all $\lambda \in \left(\frac{1}{4}, \frac{1}{3}\right)$ we have $\dim_{\mathrm{H}}(\Lambda^\lambda) = s(\lambda)$ (that is the overlaps between the cylinders $\Lambda_i = S_i(\Lambda)$ are not strong enough to cause the drop of Hausdorff dimension of the attractor Λ for a typical parameter $\lambda \in \left(\frac{1}{4}, \frac{1}{3}\right)$). However, the set of those parameters λ for which $\dim_{\mathrm{H}}(\Lambda^\lambda) < s(\lambda)$ is dense in $\left(\frac{1}{4}, \frac{1}{3}\right)$ and also for such a typical λ the Hausdorff measure $\mathcal{H}^s(\lambda)(\Lambda^\lambda) = 0$ (as opposed to the case when the SSC holds as we mentioned above). As a consequence of a groundbreaking recent result of M. Hochman we also know now that the set of those $\lambda \in \left(\frac{1}{4}, \frac{1}{3}\right)$ for which $\dim_{\mathrm{H}}(\Lambda^\lambda) < s(\lambda)$ is a set of Hausdorff dimension zero. The dense exceptional set where $\dim_{\mathrm{H}}(\Lambda^\lambda) < s(\lambda)$ mentioned above contains parameters λ for which there exists $\mathbf{i} = (i_1, \dots, i_n), \mathbf{j} = (j_1, \dots, j_n) \in \{0,1,3\}^n$, with $i_1 \neq j_1$ and with

$$S_{i_1} \circ \cdots \circ S_{i_n}(\Lambda) = S_{j_1} \circ \cdots \circ S_{j_n}(\Lambda). \tag{3}$$

In general, if (3) holds for some n and $\mathbf{i}, \cdots \mathbf{j} \in \Sigma_n$, $i_1 \neq j_1$ then we say that there exact an exact overlap. It is easy to that is the similarity dimension is smaller than the dimension of the space (smaller than d if work in $\mathbb{R}^d$) and if there is an exact overlap then the we have dimension that is the Hausdorff dimension is smaller than the similarity dimension. Now returning to the special case of the M. Keane's $\{0,1,3\}$ problem we do not if there is any $\lambda \in \left(\frac{1}{4}, \frac{1}{3}\right)$ for which $\dim_{\mathrm{H}}(\Lambda^\lambda) < s(\lambda)$. It is the so-called exact overlap conjecture that in general the only reason for dimension drop for an SSIFS on the line with similarity dimension smaller than 1 is the existence of exact overlaps. It is not difficult to construct examples of SSIFS in $\mathbb{R}^d$ with $d \geq 2$ when we have dimension drop but we do not have exact overlaps (and the similarity dimension is smaller than d).

Self-similar measures

Let $\mathcal{S}$ be an SSIFS of the form (3) and let $\mathbf{p} = (p_1, \dots, p_m)$ be a probability vector. We say that a probability measure $\mu = \mu_{\mathcal{F},\mathbf{p}}$ is an invariant measure or stationary measure corresponding to $\mathcal{S}$ and $\mathbf{p}$ if

$$\mu(A) = \sum_{i=1}^{m} p_i \cdot \mu(S_i^{-1}(A)) = \sum_{i=1}^{m} p_i \cdot S_i\mu(A), \tag{4}$$

where $S_i\mu := \mu \circ S_i^{-1}$ is the push forward measure of μ by S_i. If the cylinders are well separated that is the OSC holds then it follows from Birkhoff Ergodic theorem that the Hausdorff dimension of the measure μ (which is defined by $\dim_*(\mu) := \inf\{\dim_{\mathrm{H}} A : \mu(A) > 0\}$) is equal its similarity dimension of the measure μ which is defined by

$$\dim_{\mathrm{S}}(\mu) := \frac{\sum\limits_{i=1}^{m} p_i \log p_i}{\sum\limits_{i=1}^{m} p_i \log r_i}. \tag{5}$$

A recent Theorem of M. Hochman [2] gives a much better condition for $\dim_{\mathrm{H}} \mu = \dim_{\mathrm{S}} \mu$ when the SSIFS is defined on $\mathbb{R}$.

Hochman's Theorem

Let $\mathcal{S} := \{S_i(x) = r_i x + t_i\}_{i=1}^{m}$, $t_i \in \mathbb{R}$ and $r_i \in (-1,1) \setminus \{0\}$ for all i and let $\mathbf{p} := (p_1, \dots p_m)$ be a probability vector. Let $\Delta_n(\mathcal{S})$ be the minimum of $\Delta(\mathbf{i},\mathbf{j})$ for distinct $\mathbf{i},\mathbf{j} \in \{1,\dots,m\}^{\mathbb{N}}$, where

$$\Delta(\mathbf{i},\mathbf{j}) = \begin{cases} \infty & S_{\mathbf{i}}(0) \neq S_{\mathbf{j}}(0) \\ |S_{\mathbf{i}}(0) - S_{\mathbf{j}}(0)| & S_{\mathbf{i}}(0) = S_{\mathbf{j}}(0), \end{cases}$$

here we used the short hand notation: $S_{\mathbf{i}} := S_{i_1} \circ \cdots \circ S_{i_n}$.

We say that the self-similar IFS $\mathcal{S}$ satisfies Hochman's exponential separation condition if there exists an $\varepsilon > 0$ and an $n_k \uparrow \infty$ such that

$$\Delta_{n_k} > \varepsilon^{n_k}. \tag{6}$$

For example, if all parameters in $\mathcal{S}$ are algebraic then either there is an exact overlap or Hochman's exponential separation condition holds. Hochman's Theorem [2] says that if Hochman's exponential separation condition holds then the Hausdorff dimension of μ is equal to the minimum of 1 and the similarity dimension of μ.

Hochman's Theorem for families of SSIFS

Let $I \subset \mathbb{R}$ be a compact parameter interval and $m \geq 2$. For every parameter $t \in I$ given a self-similar IFS on the line:

$$\mathcal{S}_t := \{S_{i,t}(x) = r_i(t) \cdot (x - a_i(t))\}_{i=1}^{m},$$

where

$$r_i : I \to (-1,1) \setminus \{0\} \text{ and } a_i : I \to \mathbb{R}$$

are real analytic functions. Let Π_t be the natural projection (defined in (2)) from $\Sigma := \{1,\dots,m\}^{\mathbb{N}}$ to the attractor Λ_t of $\mathcal{S}_t$. For every probability vector $\mathbf{p} := (p_1,\dots,p_m)$ the associated self-similar measure is $\nu_{\mathbf{p},t} := (\Pi_t)_*(\mathbf{p}^{\mathbb{N}})$. Recall that the similarity dimension of $\nu_{\mathbf{p},t}$ is $\dim_{\mathrm{S}}(\nu_{\mathbf{p},t}) := \frac{\sum_{i=1}^{m} p_i \log p_i}{\sum_{i=1}^{m} p_i \log r_i(t)}$ and the similarity dimension of $\mathcal{S}_t$ is the solution $s(t)$ of the equation $r_1^{s(t)}(t) + \cdots + r_m^{s(t)}(t) = 1$.

We say that a parameter $t \in I$ is exceptional if either $\dim_{\mathrm{H}} \Lambda_t < \min\{1, s(t)\}$ or there exists a probability vector $\mathbf{p} := (p_1,\dots,p_m)$ such that $\dim_{\mathrm{H}}(\nu_{\mathbf{p},t}) < \min\{1, \dim_{\mathrm{S}}(\nu_{\mathbf{p},t})\}$. Hochman's theorem [2] for families of SSIFS is as follows: Assume that for $\mathbf{i},\mathbf{j} \in \Sigma = \{1,\dots,m\}^{\mathbb{N}}$ we have

$$\text{if } \Pi_t(\mathbf{i}) = \Pi_t(\mathbf{j}) \text{ holds for } \textbf{all } t \in I \text{ then } \mathbf{i} = \mathbf{j}.$$

Then the packing dimension (and consequently the Hausdorff dimension) of the set of exceptional parameters is equal to 0.

References

[1] K. Falconer, *Fractal geometry: mathematical foundations and applications*, John Wiley & Sons, 2004

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

[3] M. Hochman, P. Shmerkin, *Local entropy averages and projections of fractal measures*, Annals of Mathematics, **175** 1001-1059.

[4] P. Mattila, *Geometry of sets and measures in Euclidean spaces*, Cambridge University Press, vol. **44**, (1995).

[5] MJ. Marstrand, *Some fundamental geometrical properties of plane sets of fractional dimensions*, Proceedings of the London Mathematical Societ **3** (1954), 257–302.

[6] P. Shmerkin, *On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions*, Preprint 2016.

[7] J. Bourgain, *On the Erdős-Volkmann and Katz-Tao ring conjectures*, Geom. Funct. Anal. **13** (2003), 334–365.

[8] J. Bourgain, *The discretized sum-product and projection theorems*, J. Anal. Math. **112** (2010), 193–1236.

[9] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four.*, Geom. Funct. Anal. **8** (1998), 529–551.

[10] A. Iosevich, O. Roche-Newton, M Rudnev, *On discrete values of bilinear forms*, preprint.

[11] N. Katz, T. Tao, *Some connections between the Falconer and Furstenburg conjectures*, New York J. Math. **7** (2001), 148–187.

[12] S. V. Konyagin, I. D. Shkredov, *New results on sum-products in* $\mathbb{R}$, Tr. Mat. Inst. Steklova **294** (2016), 87–98.

[13] J. Solymosi, *Bounding multiplicative energy by the sumset*, Adv. Math. **222** (2009), 402–4088.

Introduction to Entropy

Agamemnon Zafeiropoulos

We define the notion of entropy of a probability measure with respect to a partition, as well as the entropy dimension of a probability measure. We investigate connections between entropy dimension and other notions of dimension.

1. Definitions and Basic Properties

Let $\mathbf{p} = (p_i)_i$ be a probability vector. The *entropy* of $\mathbf{p}$ is defined by

$$H(\mathbf{p}) = -\sum p_i \log p_i \, .$$

The *entropy of a probability measure* μ *with respect to the partition* α of the underlying space X is defined to be

$$H(\mu, \alpha) = -\sum_{A \in \alpha} \mu(A) \log \mu(A) \, .$$

The *conditional entropy* of μ with respect to a partition α given the partition β is defined by

$$H(\mu, \alpha | \beta) = \sum_{B \in \beta} \mu(B) H(\mu_B, \alpha) \, ,$$

where μ_B denotes the normalised restriction of μ on the set B. Furthermore, given two partitions α, β their *join* $\alpha \vee \beta$ is defined to be the coarsest common refinement, i.e.

$$\alpha \vee \beta = \{A \cap B : A \in \alpha, B \in \beta\}.$$

The entropy as defined above satisfies the following properties:

- $0 \leq H(\mu, \alpha) \leq \log |\alpha|$, with $H(\mu, \alpha) = 0$ iff α is a trivial partition and $H(\mu, \alpha) = \log |\alpha|$ iff $\mu(A) = 1/|\alpha|$ for all $A \in \alpha$.
- $H(\mu, \alpha \vee \beta) = H(\mu, \alpha) + H(\mu, \beta | \alpha)$.
- $H(\mu, \alpha \vee \beta) \leq H(\mu, \alpha) + H(\mu, \beta)$.
- If μ, ν are probability measures and $0 < \lambda < 1$, then

$$H(\lambda\mu + (1-\lambda)\nu, \alpha) \geq \lambda H(\mu, \alpha) + (1-\lambda) H(\nu, \alpha).$$

- If $\mathbf{p} = (p_i)_{i=1}^k$ is a probability vector and $\mu_1, \ldots, \mu_k$ are probability measures, then

$$H\left(\sum_{i=1}^k p_i \mu_i, \alpha\right) \leq \sum_{i=1}^k p_i H(\mu_i, \alpha) + H(\mathbf{p}).$$

2. Entropy Dimension

From now on the positive integer $d \geq 1$ is considered fixed. We define the *n-th level dyadic partition* of $\mathbb{R}^d$ to be

$$\mathcal{D}_n = \left\{ \left[\frac{k_1}{2^n}, \frac{k_1+1}{2^n}\right) \times \ldots \times \left[\frac{k_d}{2^n}, \frac{k_d+1}{2^n}\right) : k_1, \ldots, k_d \in \mathbb{Z} \right\}.$$

The *n-th scale entropy* of a probability measure μ is defined to be

$$H_n(\mu) = \frac{1}{n} H(\mu, \mathcal{D}_n).$$

Finally, we define the *entropy dimension* of μ by

$$\dim_e \mu = \lim_{n \to \infty} H_n(\mu),$$

provided the limit exists. Whenever $\dim_e \mu$ exists, it is a number in $[0, d]$. The following proposition shows the relation between the entropy dimension of a measure and the box dimension of its support set.

Proposition 1. *Let μ be a probability measure. Then*

$$\dim_e \mu \leq \dim_B supp(\mu),$$

provided both dimensions exist.

The following theorem shows the connection between pointwise dimension and entropy dimension of a probability measure.

Theorem 1. *Let μ be a probability measure which is compactly supported in $\mathbb{R}^d$. If μ is exact dimensional with dimension α almost everywhere, then $\dim_e \mu = \alpha$. More generally, if the pointwise dimension of μ at the point $x \in supp(\mu)$ is $\alpha(x)$, then*

$$\dim_e \mu = \int \alpha(x) \mathrm{d}\mu(x).$$

3. Entropy Dimension of Self-Similar Measures

Apart from exact-dimensional measures, entropy dimension also exists for certain self-similar measures.

Theorem 2. *Let $\Phi = \{\phi_i\}_{i \in I}$ be an Iterated Function System of similarities in $\mathbb{R}^d$ and $\mu = \sum_{i \in I} p_i \, \phi_{i*}\mu$ be a self-similar measure. The entropy dimension $\dim_e \mu$ exists.*

References

[1] Ai-Hua Fan, Ka-Sing Lau, Hui Rao: *Relationships between different dimensions of a measure* Monatsch. Math, 135(3): 191-201, 2002

[2] Y. Peres, B. Solomyak: *Existence of L^q Dimensions and Entropy Dimension for Self-conformal Measures.* Indiana University Mathematics Journal, 49(4), 1603-1621 (2000). Retrieved from http://www.jstor.org/stable/24901114

[3] Peter Walters: *An Introduction to Ergodic Theory* volume 79 of graduate Texts in Mathematics, Springer-Verlag, New York, 1982

CP Processes

Daniel Glasscock

A CP process is, roughly speaking, a measure preserving dynamical system on a space of probability measures under zoom–and–scale dynamics. Harry Furstenberg [1] introduced CP processes in 1970 as a tool in the study of the relationship between $2x$ and $3x \pmod 1$ dynamics, and they were recently employed to resolve some of Furstenberg's original conjectures [4, 6] concerning the dimension of projections and slices of product sets invariant under those dynamics. In this talk, we define CP processes, give some basic examples and properties, and outline in broad strokes the way in which they are used.

Dynamics comes to bear on problems in fractal geometry via the repeated action of zooming in on part of a probability measure. To zoom in on $\mu \in \mathcal{P}([0,1])$ on an interval $I \subseteq [0,1]$ for which $\mu(I) > 0$, we restrict μ to I, push μ forward through the unique homothety which sends I to $[0,1]$, and renormalize so that this pushforward becomes a probability measure. The goal is to gain insight into the fine-scale structure of μ by repeatedly zooming in and understanding, for example, the trajectory of μ through $\mathcal{P}([0,1])$. We can realize this goal by constructing a CP process related to μ and transferring nice properties of that process back to μ.

CP processes may be described dynamically as measure preserving systems or probabilistically as random processes. We shall use dynamical language, following

[5]; for an introduction to CP processes in the language of random walks and Markov chains, see [3, Section 6] .

1. CP-processes on trees

We will define CP processes on symbolic trees. The symbolic setting is helpful because the zoom–and–scale map on measures on a totally disconnected space is continuous. Passing results back and forth between the Euclidean setting and the symbolic setting has its own complications, but we will not address those here.

Fix $b \geq 2$ and $d \geq 1$, and let $\Lambda = \{0, \dots, b-1\}^d$. Denote by $\sigma : \Lambda^{\mathbb{N}} \to \Lambda^{\mathbb{N}}$ the *left shift* $(\sigma w)_n = w_{n+1}$. For $v \in \Lambda^n$, let $[v] = \{w \in \Lambda^{\mathbb{N}} \mid w_1 \cdots w_n = v\}$. For $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$ and $v \in \Lambda^n$ for which $\mu[v] > 0$, define the measure $\mu^v \in \mathcal{P}(\Lambda^{\mathbb{N}})$ by zooming in on μ on $[v]$ and scaling:

$$\mu^v = \frac{\sigma_*^n(\mu\big|_{[v]})}{\mu[v]}, \quad \text{that is, for all cylinder sets } [u] \subseteq \Lambda^{\mathbb{N}}, \ \mu^v[u] = \frac{\mu[vu]}{\mu[v]}.$$

The *geometric coding map* $\gamma : \Lambda^{\mathbb{N}} \to [0,1]^d$ defined by $w \mapsto \sum_{n=1}^{\infty} w_n/b^n$ connects the symbolic and Euclidean settings; the measure μ^v corresponds to the Euclidean measure gotten by zooming in on $\gamma_*\mu$ on the b-adic cube $\gamma[v]$ and scaling.

CP processes will be defined to be measure preserving dynamical systems on a subset of $\mathbf{X} = \mathcal{P}(\Lambda^{\mathbb{N}}) \times \Lambda^{\mathbb{N}}$, the space of pairs of a measure on which to zoom in and a point indicating where to zoom. Endowing $\mathcal{P}(\Lambda^{\mathbb{N}})$ with the weak-$*$ topology, the set $\mathbf{X}$ is a compact and metrizable topological space. On the subset

$$X = \big\{(\mu, w) \in \mathbf{X} \bigm| \mu[w_1 \cdots w_n] > 0 \text{ for all } n \in \mathbb{N}\big\} = \big\{(\mu, w) \in \mathbf{X} \mid w \in \mathit{supp}\mu\big\},$$

we define the zoom–and–scale map $T : X \to X$ by

$$T(\mu, w) = \big(\mu^{w_1}, \sigma w\big).$$

A *(base-b) CP distribution* is a Borel probability measure $Q \in \mathcal{P}(X)$ that is T-invariant and adapted (defined in the next paragraph). A *(base-b) CP process* is a measure preserving dynamical system $(X, \mathcal{B}, Q, T)$ where $\mathcal{B}$ is the Borel σ-algebra on X and Q is a CP distribution.

Probability measures on $\mathbf{X}$ or $\mathcal{P}(\Lambda^{\mathbb{N}})$ are called *distributions* in order to distinguish them from measures on smaller spaces such as $\Lambda^{\mathbb{N}}$. The projection $\pi_1 : \mathbf{X} \to \mathcal{P}(\mathcal{P}(\Lambda^{\mathbb{N}}))$ allows us to associate to the distribution $Q \in \mathcal{P}(\mathbf{X})$ its *measure marginal* $\overline{Q} = (\pi_1)_*Q \in \mathcal{P}(\mathcal{P}(\Lambda^{\mathbb{N}}))$. We define Q to be *adapted* if for all $f \in C(\mathbf{X})$,

$$\int f(\mu, w)\ dQ(\mu, w) = \iint f(\mu, w)\ d\mu(w) d\overline{Q}(\mu).$$

Adaptedness means "for Q-a.e. (μ, w)" is interchangeable with "for $\overline{Q}$-a.e. μ, for μ-a.e. w." Given $P \in \mathcal{P}(\mathcal{P}(\Lambda^{\mathbb{N}}))$, there is a unique adapted distribution $Q \in \mathcal{P}(\mathbf{X})$ for which $\overline{Q} = P$; if Q is adapted, then $Q(X) = 1$, so $Q \in \mathcal{P}(X)$. Since CP distributions are (by definition) adapted, it is common to speak of them as being supported on $\mathcal{P}(\mathcal{P}(\Lambda^{\mathbb{N}}))$ and to write $\mathit{supp}Q$ to mean $\mathit{supp}\overline{Q}$.

The simplest examples of CP distributions are those supported on a single measure; it is an easy exercise to show that $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$ is a product measure if and only if there exists a CP distribution Q such that $supp\overline{Q} = \{\mu\}$. This example already demonstrates a basic connection between the T-invariance of Q and the fine-scale structure of measures in $supp\overline{Q}$. As a related example, if $\mu, \nu \in \mathcal{P}(\Lambda^{\mathbb{N}})$ are such that for all $\lambda \in \Lambda$, $\mu^{\lambda} = \nu$ and $\nu^{\lambda} = \mu$, then the adapted distribution with measure marginal $(\delta_{\mu} + \delta_{\nu})/2$ is a CP distribution.

There are two other examples of CP distributions that we will just mention here. For a σ-invariant measure $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$, the adapted distribution Q with $\overline{Q} = \int \delta_{\delta_w}\, d\mu(w)$ is a CP distribution supported entirely on point masses. Furstenberg describes an extension of this example with prediction measures in [2, pg. 409].

2. Micromeasure distributions and dimension

Given a measure $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$ and a point $w \in supp\mu$, the trajectory of μ through $\mathcal{P}(\Lambda^{\mathbb{N}})$ alluded to above is $(\mu^{w_1 \cdots w_n})_{n \in \mathbb{N}}$, the sequence of probability measures seen around w in μ. A *micromeasure of* μ is a limit point of such a trajectory. If μ has dynamical or combinatorial origins, its micromeasures can often be related back to itself. An example of this is given in the following lemma.

Lemma 1. *If $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$ is σ-invariant and ν is a micromeasure of μ, then $supp\nu \subseteq supp\mu$.*

The set $\mathcal{MD}(\mu, w)$ of *micromeasure distributions of μ at w* is the set of limit points of $\left\{\frac{1}{n}\sum_{i=0}^{n-1} \delta_{T^i(\mu,w)}\right\}_{n\in\mathbb{N}}$ in $\mathcal{P}(\mathbf{X})$. Micromeasure distributions are supported on the micromeasures of μ. Just as in the theorem of Krylov and Bogolioubov, the set $\mathcal{MD}(\mu, w)$ is non-empty by the compactness of $\mathcal{P}(\mathbf{X})$ and every element is T-invariant, *provided it is supported on X*. The following theorem says that most of the time, this caveat is satisfied.

Theorem 1 ([5, Theorem 28])**.** *For all $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$, for μ-a.e. $w \in \Lambda^{\mathbb{N}}$, every element of $\mathcal{MD}(\mu, w)$ is a CP distribution.*

Thus micromeasure distributions provide a rich array of CP distributions. Two useful facts about ergodic CP distributions – those Q for which $(X, \mathcal{B}, Q, T)$ is ergodic – follow quickly from Theorem 1: since almost every point is generic for an ergodic measure, if Q is ergodic, then for Q-a.e. (μ, w), $\mathcal{MD}(\mu, w) = \{Q\}$; and the ergodic components in the ergodic decomposition of a CP distribution are themselves CP distributions.

The latter fact is useful as it allows us to concentrate on ergodic CP distributions. The *dimension* of an ergodic CP distribution Q is $\dim Q = \int H\big(\nu, \mathcal{C}_1\big)\, d\overline{Q}(\nu)$, the $\overline{Q}$-average Shannon entropy of measures in $\mathcal{P}(\Lambda^{\mathbb{N}})$ with respect to the partition $\mathcal{C}_1 = \big\{[\lambda] \mid \lambda \in \Lambda\big\}$. Measures which support ergodic CP distributions have nice dimensionality properties, as indicated in the following theorem.

Theorem 2 ([2, Theorem 2.1]). *Let Q be an ergodic CP distribution. The $\overline{Q}$-typical measure μ is exact dimensional with* $\dim\mu = \dim Q$*: for μ-a.e. $w \in \Lambda^{\mathbb{N}}$,*

$$\lim_{n\to\infty} \frac{\log\big(\mu[w_1\cdots w_n]\big)}{n} = \dim Q.$$

Combining the ideas behind micromeasure distributions and Theorem 2, we can construct from a measure μ an ergodic CP process supported on the micromeasures of μ with dimension bounded from below.

Theorem 3 ([4, Theorem 7.10]). *Let $\mu \in \mathcal{P}(\Lambda^{\mathbb{N}})$. There exists an ergodic CP distribution of dimension at least* $\limsup_{n\to\infty} H(\mu,\mathcal{C}_n)/n$*, where $\mathcal{C}_n = \big\{[v] \mid v \in \Lambda^n\big\}$, supported on the micromeasures of μ.*

More can be said about CP processes such as the ones arising in Theorem 3 than about the specific measures from which they arise; this will be, in part, the subject of the following talks. Some properties of these CP processes then pass back to the originating measure μ via results such as the one in Lemma 1.

References

[1] H. Furstenberg, *Intersections of Cantor sets and transversality of semigroups*, in Problems in Analysis (Sympos. Salomon Bochner, R. C. Gunning (Ed.)), Princeton Univ. Press, **50**:7040 (1970), 41–59

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

[3] M. Hochman, *Lectures on dynamics, fractal geometry, and metric number theory*, Journal of Modern Dynamics **8** (2015), 437–497

[4] M. Hochman, P. Shmerkin, *Local entropy averages*, Ann. Math **175** (2012), 1001–1059

[5] P. Shmerkin, *Ergodic geometric measure theory*, Unpublished, available upon request (2012)

[6] M. Wu, *A proof of Furstenberg's conjecture on the intersections of $\times p$ and $\times q$-invariant sets*, ArXiv e-prints, https://arxiv.org/abs/1609.08053

Furstenberg's Dimension Conservation Theorem and Local Entropy Averages

Simion Filip

This talk developed results on CP processes due to Furstenberg [Fur08] and Hochman & Shmerkin [HS12].

The first basic result is concerned with projections of measures on trees. Suppose that X, Y are two trees and $\pi : X \times Y \to X$ is a projection onto the first factor. A measure θ on $X \times Y$ yields the pushed-forward measure $\pi_*\theta$ on X and also conditional measures θ_x for a.e. $x \in X$. Furstenberg's result [Fur08, §3] shows that for random measures coming from CP processes the dimension of the projected measure and the dimension of the conditional measures add up to the total dimension of the measure θ.

Theorem [Furstenberg Dimension Conservation] For an ergodic CP process, for a.e. measure θ the projection $\pi_*\theta$ is exact-dimensional, as are the fiberwise conditional measure θ_x, and we have

$$\dim\theta = \dim\pi_*\theta + \dim\theta_x$$

The proof of the theorem involves the following steps. First, for any CP processes there is an observable (the entropy for the first level partition) whose Birkhoff averages give the dimension of the measure. An adaptation of this construction, using fiberwise entropy, gives a formula for $\dim\theta_x$ in terms of an expression resembling, but not quite equaling a Birkhoff sum. Untangling the expression and applying a variant of the Birkhoff ergodic theorem implies the end result.

Furstenberg used his result to show that dimension conservation holds for self-similar fractals $A \subset \mathbb{R}^{m_1+m_2}$ for a projection $\pi : \mathbb{R}^{m_1+m_2} \to \mathbb{R}^{m_1}$. Namely, starting from a self-similar fractal he builds a CP process, adapted to the projection in question. Applying the theorem above implies dimension conservation in the following sense: there exists $\delta > 0$ such that

$$\delta + \dim\left\{x \in \mathbb{R}^{m_1} : \dim\pi^{-1}(x) \leq \delta\right\} \geq \dim A$$

By convention, the dimension of the empty set is $-\infty$.

The next result discussed is due to Hochman & Shmerkin [HS12, §4] and is a key tool in proving further results on agreement of expected and actual dimension in later talks. The Birkhoff sums used to compute dimension of measures for CP processes are now replaced by local entropy averages. The advantage of entropy averages is that they can be estimated in terms of local quantities. The result applies to any measure on a tree, not just one coming from a CP process.

For a point x in a tree X, which we view as a point at infinity in the tree, denote by $[x_1^n]$ the level n cylinder containing X. Then we have:

Theorem [Local Entropy Averages] Suppose that μ is a measure on a tree X and that for μ-a.e. $x \in X$ we have

$$\liminf_{n\to\infty}\left(\frac{-\log\mu([x_1^n])}{n}\right) \geq \alpha$$

Then we have $\underline{\dim}\mu \geq \alpha$.

There is also a relative version of this theorem which is useful when estimating dimensions of projections. To describe it, denote for $x \in X$ and level n cylinder set $[x_1^n]$ the conditional measure $\mu_{[x_1^n]}$ induced on the symbol set by

$$\mu_{[x_1^n]}(\lambda) := \frac{\mu([x_1^n\lambda])}{\mu([x_1^n])}$$

Theorem [Local Entropy Averages, Relative Version] For a morphism of trees $f : X \to Y$ and a probability measure μ on X, suppose that

$$\liminf \frac{1}{N}\sum_{k=0}^{N-1} H(f_*(\mu_{[x_1^n]})) \geq \alpha$$

for μ-a.e. x. Then we have $\underline{\dim} f_*\mu \geq \alpha$.

The proof of this theorem is by finding a "random" section $\sigma : Y \to X$ of the projection map $f : X \to Y$. The sections are chosen uniformly at random in each fiber and one applies the previous result to the measure $\sigma_* f_* \mu$; the lower bound for the dimension of $f_* \mu$ then follows by averaging.

References

[Fur08] Furstenberg H. — "Ergodic fractal measures and dimension conservation". *Ergodic Theory Dynam. Systems* **28** no. 2, (2008) 405–422. http://dx.doi.org/10.1017/S0143385708000084.

[HS12] Hochman M. & Shmerkin P. — "Local entropy averages and projections of fractal measures". *Ann. of Math. (2)* **175** no. 3, (2012) 1001–1059. http://dx.doi.org/10.4007/annals.2012.175.3.1.

Hochman–Shmerkin projection theorems

Laurent Dufloux

This talk is based on [1]; we restrict ourselves to dimension 2 in order to simplify the exposition, and we skip the part of this paper which deals with products of Gibbs measures. We first state Hochman–Shmerkin projection theorem. This result is then applied to self-similar measures with dense rotations, and products of $\times 2$ and $\times 3$ invariant measures, settling a conjecture of Furstenberg.

1. Projection theorem

In previous talks, CP-distributions were defined in trees; the definition of a CP-distribution in Euclidean spaces is essentially the same, with nested dyadic partitions playing the role of nested cylinders. See [1] for a more general definition.

If P is a CP-distribution in $\mathbb{R}^2$ and π is orthogonal projection from $\mathbb{R}^2$ onto some line of $\mathbb{R}^2$, we let

$$E_P(\pi) = \int \mathrm{d}P(\mu)\ \underline{\dim}(\pi_* \mu).$$

Theorem 1 ([1] Theorems 8.1 and 8.2)**.** *Let P be an ergodic CP-distribution of dimension $\alpha \in [0, 2]$.*

(1) *For almost every π, $E_P(\pi) = \inf\{\alpha, 1\}$.*
(2) *If π is fixed, $E_P(\pi) = \underline{\dim}(\pi_* \mu)$ for P-almost every μ.*
(3) *For P-almost every μ, $\underline{\dim}(\pi_* \mu) \geq E_P(\mu)$ for every π.*
(4) *The mapping $\pi \mapsto E_P(\pi)$ is lower semi-continuous.*

The proof of this result relies on local entropy averages bounds and Marstrand's projection theorem (for the first statement).

The main point is that local entropy averages bounds, along with the (statistical) "self-similarity" property of a random μ, allow to consider the entropy of projected measures *at a fixed scale*, and this is then "essentially" a continuous function of π.

2. Projections of self-similar measures

Consider an IFS $\{f_i \ ; \ i \in \Lambda\}$ where Λ is finite, and the f_i are contracting similarities of $\mathbb{R}^2$. We assume that this IFS satisfies the strong separation condition, *i.e.* is X is the attractor of the IFS, the $f_i(X)$ are pairwise disjoint. Let μ be a self-similar measure, *i.e.* $\mu = \sum_i p_i (f_i)_* \mu$, where (p_i) is a probability vector with strictly positive components.

Theorem 2 ([1], Theorem 1.6)**.** *Assume that the rotation parts of the f_i generate a dense semigroup of* $\mathbf{SO}(2)$*. Then for* any *linear projection π from $\mathbb{R}^2$ onto $\mathbb{R}$,*

$$\dim(\pi_* \mu) = \inf\{1, \dim(\mu)\}.$$

Proof. It is possible to construct an ergodic CP-distribution P satisfying the property that for P-almost every ν, there is an affine similarity S such that μ is absolutely continuous with respect to $S_* \nu$. An application of Theorem 1 then shows that, given $\varepsilon > 0$, the set of projections π such that $\pi_* \mu$ has dimension at least $\inf\{1, \dim(\mu)\} - \varepsilon$ is dense and open, and the hypothesis on the rotation parts of the f_i then implies that actually $\underline{\dim}(\pi_* \mu) > \inf\{1, \dim(\mu)\} - \varepsilon$ for *every* π. □

The corresponding result for projections of self-similar *sets* follows from the theorem for self-similar measures.

3. Furstenberg's conjecture

The following result was stated, but not proved in the course of the talk:

Theorem 3 ([1], Theorem 1.3)**.** *Let μ (resp. ν) be a $\times 2$ (resp $\times 3$) invariant measure on $[0,1]$. Let θ be the product measure $\theta = \mu \otimes \nu$. Then for every orthogonal projection π which is not one of the coordinate projections,*

$$\underline{\dim}(\pi_* \theta) = \inf\{1, \underline{\dim}(\theta)\}.$$

This is a strengthening of a conjecture of Furstenberg dealing with sets rather than measures. The statement for sets follows from the statement for measures, using the variational principle.

The proof is quite technical. It relies on Theorem 1 (more precisely, a version of this result for non-ergodic CP-distributions) and the construction of a "generalized" CP-distribution, where dyadic partitions are replaced with a family of partitions by rectangles of bounded eccentricity. The measure θ is invariant by the product transformation $(\times 2, \times 3)$ which is non-conformal; in order to be able to zoom in on θ in a meaningful way, one is led to construct a dynamical system living *above* an irrationnal rotation, and the zooming process is a skew product over this rotation.

References

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

Wu's Proof of Furstenberg's Intersection Conjecture

Tom Kempton

This talk was based on recent work of Wu [1], in which he proved the following theorem, originally conjectured by Furstenberg.

Theorem 1. *If $A, B \subset [0,1]$ are closed and invariant under $\times p, \times q$ respectively, and if $\frac{\log p}{\log q} \notin \mathbb{Q}$, then for all real numbers u and v,*

$$\dim_H((uA+v)\cap B) \leq \max\{0, \dim_H(A) + \dim_H(B) - 1.$$

We focused on the special case that A is the middle-$\frac{1}{3}$ Cantor set and B the middle-$\frac{1}{2}$ Cantor set, this is notationally simpler and allows for good pictures to be drawn, but is actually not much easier than the proof of the full theorem.

The sets $(uA+v)\cap B$ can be thought of (up to an affine coordinate change) as slices through the product set $A \times B$. Wu's proof involves showing that that, if there is a slice through $A \times B$ of upper box dimension γ, then

(1) For Lebesgue almost every θ there exists a slice l_θ through $A \times B$ with slope θ and $\dim_H(l_\theta \cap (A \times B)) \geq \gamma$.

(2) Furthermore, these slices l_θ can be chosen such that there is a set $C \subset A \times B$ of small box dimension such that each l_θ intersects B. (The real statement is a little more complicated, but follows this idea)

(3) Putting 1 and 2 together gives that $A \times B$ must have dimension at least $1+\gamma$.

Part 1 was originally proved by Furstenberg. The proof involves building CP chains supported on slices through $A \times B$ of Hausdorff dimension at lest γ. The majority of the talk was spent proving part 1 and showing how parts 1 and 2 together are enough to show part 3.

References

[1] M. Wu, *A proof of Furstenberg's conjecture on the intersections of $\times p$ and $\times q$-invariant sets*, (arXiv: 1609.08053)

Some additive combinatorics

Thomas F. Bloom

1. Introduction to additive combinatorics

I give an introduction to some basic tools and concepts of additive combinatorics, in their traditional context of set addition and in terms of entropy, the latter following in particualar the paper of Tao [2]. Let G be an abelian group, which for convenience I will take to be finite, and let $A, B \subset G$. The sumset $A+B$ is defined as

$$A + B = \{a + b : a \in A, b \in B\}.$$

In this talk I will discuss inequalities between sizes of sumsets, and also what kind of structural information can be deduced from knowing that such sizes are small.

(1) **Inequalities from trivial identities:** There is a very useful family of relationships between the sizes of sumsets, of which the most useful is the following sumset inequality due to Ruzsa, often known as Ruzsa's triangle inequality:

$$|A - C||B| \leq |A - B||B - C|,$$

which follows since the trivial identity $(a - b) + (b - c) = a - c$ implies that the map $(a - c, b) \mapsto (a - b, b - c)$ is an injection (choosing a unique representative for each $a - c \in A - C$).

(2) **Covering lemmas:** Often the starting point for the proof of inverse theorems discussed below, these show that sets with small sumset can be efficiently covered by a small number of translates. Again, the classic example is due to Ruzsa: if $|A + B| \leq K|B|$ then A is contained in the union of at most K translates of $B - B$. The proof is just to take a set $X \subset A$ maximal such that the translates $x + B$ are pairwise disjoint for all $x \in X$.

(3) **Plünnecke's inequality:** Suppose that $|A + A|$ is small compared to $|A|$. Since set addition is a smoothing operation, one would hope that this property is preserved under more additions, e.g. $|A + A + A|$ continues to be small, and so on. This is made rigorous by the following inequality of Plünnecke:

$$\text{if } |A + A| \leq K|A| \quad \text{then for all } t \geq 2 \quad |tA| \leq K^t|A|,$$

where $tA = A + \cdots + A$ is the t-fold sumset of A.

(4) **Inverse theorems:** What structural information about A can we deduce if $|A + A| \ll |A|$? Such an inequality clearly holds if A is dense in some multidimensional arithmetic progression. The Freiman-Ruzsa inverse theorems show that this is the only possibility, allowing us to deduce strong algebraic structure from quite weak statistical information.

We will go into the latter two topics in more depth shortly. Two other aspects of additive combinatorics important for this workshop are the Balog-Szemerédi-Gowers theorem and the sum-product phenomenon, which will be discussed in separate talks.

2. Entropy analogues

Let X be a random variable taking values in G. If X is uniformly sampled from some $A \subset G$ then the entropy $H(X)$ is exactly $\log|A|$. This leads us to ask similar questions as in the previous section, but now considering arbitrary random variables. For example, what can we deduce about X if $H(X + X) - H(X)$ is small? Are there sumset inequalities analaogous to the Ruzsa triangle inequality?

It is important to note an important distinction between the combinatorial and entropy worlds – even if X is sampled uniformly from A, $X + X$ is not necessarily uniformly sampled from $A + A$. Instead, the corresponding probability measure is proportional to $1_A * 1_A$, which is much smoother than 1_{A+A}.

For example, consider the set A which is the union of two arithmetic progressions in $\mathbb{Z}^2$ along the orthogonal axes. Sampling uniformly from $A+A$ would almost surely give a point inside the integer lattice on a box, while sampling according to $1_A * 1_A$ would, with positive probability, give a point on the axes.

While the relationship between cardinalities of sumsets and entropy of sums of random variables is close, neither can be deduced from the other in general, and both must be developed in parallel.

The important facts about entropy $H(X)$ we need are that it is a non-negative quantity, however we condition on other random variables, that conditioning decreases entropy, $H(X|Y) \leq H(X)$, and we have the chain rule $H(X|Y)+H(Y) = H(X,Y)$. From this it is straightforward to deduce the submodularity inequality: if X and Y both individually determine Z and the joint distribution (X,Y) determines W then

$$H(Z) + H(W) \leq H(X) + H(Y).$$

This has two important consequences.

(1) **Ruzsa triangle inequality:** Let X, Y, Z be independent random variables. Since $(X-Y, Y-Z)$ and (X,Z) each independently determine $X-Z$ and $(X, X-Y, Z, Y-Z)$ determines (X,Y,Z),

$$H(X-Z) + H(X,Y,Z) \leq H(X-Y, Y-Z) + H(X,Z)$$

whence by independence

$$H(X-Z) + H(Y) \leq H(X-Y) + H(Y-Z),$$

the entropy analogue of Ruzsa's triangle inequality.

(2) **Kaimanovich-Vershik inequality:** This is an entropy analogue of Plünnecke's inequality. Let X, Y, Z be independent random variables. Since $(X+Y, Z)$ and $(X, Y+Z)$ each determine $X+Y+Z$ and $(X+Y, Z, X, Y+Z)$ determines (X,Y,Z), the submodularity inequality yields, after rearrangement,

$$H(X+Y+Z) - H(X+Z) \leq H(X+Y) - H(X).$$

In particular, if $H(X+Y) \leq H(X) + \log K$ then, if $Y_1, \ldots, Y_t$ are independently sampled copies of Y, then iterating the above gives

$$H(X + Y_1 + \cdots + Y_t) \leq H(X) + t \log K.$$

This inequality will be useful many times in the subsequent talks.

3. Plünnecke's inequality

The direct analogue of the previous entropy inequality for cardinalities would be that, if A, B, C are any finite sets, then

$$|A+B+C||A| \leq |A+B||A+C|.$$

This inequality is false, however, since the cardinality of a sumset is much less robust than the entropy of the sum of random variables. Petridis observed that if we are able to pass to some subset $A' \subset A$, however, conditioned only on B, then this becomes true. That is, for any sets A, B there is $A' \subset A$ (depending on B) such that for all sets C

$$|A' + B + C||A| \leq |A + B||A' + C|.$$

Petridis gave a very elegant proof of this fact, using only elementary combinatorics. Plünnecke's inequality is a simple deduction of this inequality.

4. Inverse theorems

How can $|A + A|$ be small compared to $|A|$? For convenience, we will now assume that $G = \mathbb{Z}$. It is easy to check that d-dimensional arithmetic progressions are an example: a progression of rank d is a set of the shape

$$P = \{a_0 + a_1 n_1 + \cdots + a_d n_d : 0 \leq n_i < N_i\}$$

for some integers a_i, N_i. Whatever the choice of parameters, $|P + P| \leq 2^d |P|$. Furthermore, if A is any large subset of P then it must also have small doubling for trivial reasons.

A important and deep result of Freiman and Ruzsa is that the converse also holds. That is, $|A + A| \leq K|A|$ if and only if there is a progression P of rank $d \ll_K 1$ such that $A \subset P$ and $|A| \gg_K |P|$.

Tao proved the following natural entropy analogue: if $H(X+X) \leq H(X)+\log K$ then there is a progression P of rank $d \ll_K 1$ such that X is close in transport distance to μ_P, the uniform measure on P. That is, there is some Z (possibly dependent on X) such that $H(Z) \ll_K 1$ and $X + Z \equiv \mu_P$. The proof first reduces to the case when X is close to the uniform measure on some set A, and then invokes the previous Freiman-Ruzsa inverse result.

References

[1] G. Petridis, *New proofs of Plünnecke-type estimates for product sets in groups*, Combinatorica, **32** (2012), 721–733.

[2] T. Tao, *Sumset and inverse sumset theory for Shannon entropy*, Combin. Probab. Comput. **19** (2010) 603–639.

[3] T. Tao and V. Vu, *Additive combinatorics*, Cambridge University Press (2006).

Hochman's inverse theorem on the growth of entropy under convolutions

Mikolaj Fraczyk

Hochman's inverse theorem, introduced in [1], describes the multiscale structure of probability measures μ, ν on $[0, 1)$ for which scale 2^{-n}−entropy of the convolution $\mu * \nu$ is very close to scale 2^{-n}−entropy of μ. To make the statement precise we define below the scale 2^{-n}−entropy as well as other necessary notions.

0.1. **Notations.**

0.1.1. *Dyadic partitions.* Let I be an interval in $\mathbb{R}$. Write $\mathcal{P}(I)$ for the space of probability measures on I. For $n \in \mathbb{Z}$ we the partition D_n of $\mathbb{R}$ is given as

$$\mathbb{R} = \bigsqcup_{k\in\mathbb{Z}} \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right).$$

If $x \in \mathbb{R}$ we write $D_n(x)$ for the unique cell in D_n containing x. For a measure $\mu \in \mathcal{P}(\mathbb{R})$ and a cell $D \in D_m$ such that $\mu(D) \neq 0)$ we write

$$\mu_D := \frac{1}{\mu(D)}\mu\mid_D \text{ and } \mu^D := (T_D)_*\mu_D$$

where $T_D\colon D \to [0,1)$ is the unique bijective homothety between D and $[0,1)$. Measure μ_D is called the **raw D-component** of μ and μ^D is the **rescaled D-component** of μ. For $x \in R$ we adopt convention $\mu_{x,i} = \mu_{D_i(x)}$ and $\mu^{x,i} = \mu^{D_i(x)}$.

0.1.2. *Entropy, almost atomic and almost uniform measures.* Let $\mu \in \mathcal{P}(\mathbb{R})$. The normalized entropy of μ at scale 2^{-n} is defined as

$$H_n(\mu) := \frac{1}{n}H(\mu, D_n) = \frac{1}{n}\int_{\mathbb{R}} -\log(\mu(D_m(x))dx.$$

Let $\varepsilon > 0, m \geq 0$. We say that measure $\mu \in \mathcal{P}([0,1))$ is (ε, m)**-atomic** if $H_m(\mu) \leq \varepsilon$ and (ε, m)**-uniform** if $H_m(\mu) \geq 1-\varepsilon$. Intuitively, when ε is small the measure μ is (ε, m)-atomic is its mass in concentrated in a single cell of D_m and (ε, m)-atomic if its mass is almost uniformly distributed on the cells of D_m in $[0,1)$.

0.1.3. *Probability and expected value.* We adopt following conventions: Let $A \subset \mathcal{P}([0,1))$ be event and let $f\colon \mathcal{P}(\mathbb{R}) \to \mathbb{R}$ a measurable function. For measures $\mu, \nu \in \mathcal{P}(\mathbb{R})$ we write

$$\mathbb{P}(\mu^{x,i} \in A) := \mu(\{x \in \mathbb{R} \mid \mu^{x,i} \in A\})$$

and

$$\mathbb{E}(f(\mu^{x,i})) := \int f(\mu^{x,i})d\mu(x).$$

When two or more measures are involved we will treat their component measures as independent random variables. In particular

$$\mathbb{E}(f(\mu^{x,i} * \nu^{y,j})) := \int f(\mu^{x,i} * \nu^{j,y})d\mu(x)d\nu(y).$$

0.2. **Inverse theorem.** As we explained in the first paragraph Hochman's inverse theorem describes the structure of measures μ, ν on $[0,1)$ such that for n-big and δ very small we have $H_n(\mu * \nu) \leq H_n(\mu) + \delta$. Note that the inequality is satisfied trivially if μ is uniform (i.e. Lebesgue on $[0,1)$) and ν is supported on a small ball or when ν is atomic and μ is any probability measure. Let us give a more complicated example.

0.2.1. *Motivating example.* Fix $\delta > 0$ and a natural number n, it has to be big when δ is very small, say $n >> \delta^{-1}$. We will construct a probability measure μ on $[0, 1)$ with the property that $H_n(\mu * \mu) \leq H_n(\mu) + \delta$. Choose $k \leq \delta n$ and choose natural numbers $1 = a_1 < b_1 < a_2 < b_2 < \ldots < a_k < b_k = n$. Consider the set Σ of rational numbers of form $q = \frac{m}{2^n}, 0 \leq m < 2^n$ such that q has non-zero binary digits only on positions in the intervals $[a_i, b_i)$ for $i = 1, \ldots k$. We choose the numbers a_i, b_i in such a way that $N := \sum_{i=1}^{k}(b_i - a_i)$ (i.e the number of positions where the digit in not fixed) is roughly of size $n/2$. The measure μ is defined as the normalized counting measure on Σ. The cardinality of Σ is 2^N and the atoms are separated by at least 2^{-n} so we can compute the scale $2^{-n}-$entropy as follows $H_n(\mu) = \frac{1}{n}H(\mu) = \frac{1}{n}\log 2^N = \frac{N}{n} \sim \frac{1}{2}$. To estimate the entropy of the convolution $\mu * \mu$ we will simply bound the cardinality of the support of $\mu * \mu$. Let $x, y \in \Sigma$, we claim than the sum $x + y$ can have non zero binary digits only on places in the intervals $[a_i - 1, b_i)$ for $i = 1, \ldots, k$. Indeed, if the $m - th$ digit of $x + y$ is non-zero then either $m-$th of $m + 1$-th digit of one of x or y had to be non-zero. Hence, the non-zero digits of numbers in Σ can appear only on places in $\bigcup_{i=1}^{k}[a_i - 1, b_i)$ which implies a bound $|\Sigma + \Sigma| \leq 2^{N+k}$. It follows that $H_n(\mu * \mu) = \frac{1}{n}H(\mu) \leq \frac{N+k}{n} = H_n(\mu) + \frac{k}{N}$. Recall that k was chosen so that $k \leq \delta N$ so we have $H_n(\mu * \mu) \leq H_n(\mu) + \delta$.

We end this paragraph with the multiscale analysis of μ. Choose m small relatively to n. We are interested in the scale-2^{-m} structure of the rescaled component measures $\mu^{x,j}, j = 1 \ldots, n$. It turns out that when the scale j is inside $[a_i, b_i - m)$ then the rescaled components $\mu^{x,i}$ are, at scale 2^{-m} as uniform as possible i.e. $H_m(\mu^{x,i}) = 1$. On the other hand when the scale j is in $[b_i, a_{i+1} - m)$ then the measures $\mu^{x,i}$ are atomic at scale 2^{-m} i.e. $H_m(\mu^{x,i}) = 0$. The set of scale where this dychotomy doesn't hold is contained in the union $\bigcup_{i=1}^{k}([a_{i-m}, a_i) \cup [b_i - m, b_i))$ so its cardinality is roughly of size $2km \leq 2\delta nm$. We see that for m small compared to n the dychotomy between uniformness and atomicity holds at most scales between 1 and n. Inverse theorem says that when we relax a bit the notions of uniformness and atomicity, such a dychotomy holds at almost all scales whenever $H_n(\mu * \mu) \leq H_n(\mu) + \delta$.

0.2.2. *Statement of the theorem.*

Theorem 1 (Theorem 2.7 and Theorem 4.11 [1])**.** *For every $\varepsilon > 0$ and integer $m \geq 1$, there is a $\delta = \delta(\varepsilon, m) > 0$ such that for every $n > n(\varepsilon, \delta, m)$, the following holds. If $\mu, \nu \in \mathcal{P}([0, 1))$ and*

$$H_n(\mu * \nu) < H_n(\mu) + \delta,$$

then there are disjoint subsets $I, J \subset \{1, \ldots, n\}$ with $|I \cup J| > (1 - \varepsilon)n$, such that

$$\mathbb{P}\left(\mu^{x,k} \text{ is } (\varepsilon, m) - \text{uniform}\right) > 1 - \varepsilon \text{ for } k \in I,$$

$$\mathbb{P}\left(\nu^{x,k} \text{ is } (\varepsilon, m) - \text{atomic}\right) > 1 - \varepsilon \text{ for } k \in J.$$

References

[1] M. Hochman, *On self-similar sets with overlaps and inverse theorems for entropy*, Annals of Mathematics **180** (2014), 773–822.

Hochman's Theorem on self similar measures with overlap

Amir Algom

The purpose of my talk was to discuss Hochman's method (see [1]) of computing dimensions of self similar measures and sets using the inverse theorem for entropy, proved in the previous lecture.

Let $\Phi = \{\phi_i\}_{i\in\Lambda}, |\Lambda| \geq 2$ be a finite family of real linear contractions; that is, for every $i \in \Lambda$ the map $\phi_i : \mathbb{R} \to \mathbb{R}$ is defined by $\phi_i(x) = r_i \cdot x + a_i$ where $|r_i| < 1$ and $a_i \in \mathbb{R}$ (recall that Φ is called an iterated function system - IFS). To avoid trivialities we assume throughout that there are at least two distinct contractions. Let X denote the attractor of Φ (the existence of such a set was already established in a previous talk). Let μ be the self-similar measure associated with Φ and a non-degenerate probability vector $(p_i)_{i\in\Lambda}$.

For a Borel probability measure θ on $\mathbb{R}$ we denote

$$\dim \theta = \inf\{\dim A : \theta(A) > 0\}.$$

This notion is sometimes known as lower-Hausdorff dimension. There are other notions of dimension, but for self similar measures, that are exact dimensional, most major ones coincide.

The aim of this talk was to discuss Hochman's approach to computing the dimension of the self similar measure μ. The classical approach to dimension theory of self similar measures is to impose some separation condition on Φ (e.g. the strong separation condition), and deduce that $\dim X = s - \dim X$ (the latter denotes similarity dimension). Similarly (assuming again some separation of Φ), $\dim \mu = s - \dim \mu$ (the latter denotes the similarity dimension of the self similar measure μ, introduced in a previous talk). It is when the images $\phi_i(X)$ have significant overlap that computing the dimension becomes difficult, and much less is known. The main strength of Hochman's approach is its ability to yield non trivial information about the small scale geometry of the measure μ in this situation.

Notation For $i = i_1...i_n \in \Lambda^n$ write

- $\phi_i = \phi_{i_1} \circ ... \circ \phi_{i_n}$, and call this a cylinder map.
- $r_i = r_{i_1} \cdots r_{i_n}$, the contraction ratio of ϕ_i.
- Similarly, we given a proabability vector $(p_i)_{i\in\Lambda}$ we write $p_i = p_{i_1} \cdots p_{i_n}$.

Let $n \in \mathbb{N}$, and fix $i \neq j \in \Lambda^n$. We define the distance between the n-generational cylinders i, j by

$$d_n(i,j) = |\phi_i(0) - \phi_j(0)| \quad \text{if } r_i = r_j$$

and otherwise denote $d_n(i,j) = \infty$. We define

$$\Delta_n = \min\{d(i,j) : i \neq j \in \Lambda^n\}.$$

We note the following observations:

- The definition is unchanged if we pick any other point in $\mathbb{R}$.
- The are exact overlaps in Φ if and only if $\Delta_n = 0$ for some n.
- $\Delta_n \to 0$ exponentially.
- There can be an exponential lower bound for Δ_n: this happens if the images $\phi_i(X)$ are disjoint, under the OSC, or for example when the parameters of Φ are algebraic.

The main result on self-similar measures, presented in my talk, was the following:

Theorem 1. *If μ is a self similar measure on $\mathbb{R}$ and* $\dim \mu < \min(1, s - \dim \mu)$ *then $\Delta_n \to 0$ super-exponentially, i.e.* $\lim_n \frac{-\log \Delta_n}{n} = \infty$

The conclusion is about Δ_n, which is determined by the IFS, not by the measure. Thus, if the conclusion fails, then $\dim \mu = s - \dim \mu$ for every self-similar measure of Φ. Also, the same statement remains true for the attractor X, i.e. if $\dim X < \min(1, s - \dim X)$ then $\Delta_n \to 0$ super-exponentially.

Theorem 1 is derived from a more quantitative result about the entropy of finite approximations of μ, which we now describe.

Recall

- We write $H(\mu, E)$ for the Shannon entropy of a measure μ with respect to a partition E, and $H(\mu, E|F)$ for the conditional entropy on F.
- For $n \in \mathbb{Z}$ the dyadic partitions of $\mathbb{R}$ into intervals of length 2^{-n} is
$$D_n = \{[\frac{k}{2^n}, \frac{k+1}{2^n}) : k \in \mathbb{Z}\}$$
For $t \in \mathbb{R}$ we also write $D_t = D_{[t]}$.
- If μ is a self similar measure then $\lim_n \frac{H(\mu, D_n)}{n} = \dim \mu$ (since μ is exact dimensional). In general, we always have $\liminf \frac{H(\mu, D_n)}{n} \geq \dim \mu$.

We shall only consider the case when Φ is uniformly contracting, i.e. the contractions r_i are equal to some fixed r (but the results extend to the non-uniformly contracting case).

Fix a self-similar measure μ defined by a probability vector $(p_i)_{i \in \Lambda}$. We may assume, without the loss of generality, that $0 \in X$. Define the n-generational approximation of μ by
$$\nu^{(n)} = \sum_{i \in \Lambda^n} p_i \delta_{\phi_i(0)}$$
Note that:

- This is a probability measure on X.
- $\nu^{(n)}$ converges to μ weakly.
- Let $n' = n \frac{\log \frac{1}{r}}{\log 2}$ so that $2^{-n'} = r^n$. Then $\nu^{(n)}$ closely resembles μ up to scale $2^{-n'} = r^n$ in the sense that
$$\lim_n \frac{1}{n'} H(\nu^{(n)}, D_{n'}) = \dim \mu$$

- Note that is is possible that $H(\nu^{(n)}, D_{n'})$ is strictly less than $H(\nu^{(n)})$ (w.r.t. partition into points).
- If the above inequality holds, we ask in what scale and in what rate does it appear? it must appear since $\lim_k H(\nu^{(n)}, D_k) = H(\nu^{(n)})$.
- Note that the excess entropy at scale k relative to the entropy at scale n' is just the appropriate conditional entropy:

$$H(\nu^{(n)}, D_k|D_{n'}) = H(\nu^{(n)}, D_k) - H(\nu^{(n)}, D_{n'})$$

Theorem 2. *Let μ be a self similar measure on $\mathbb{R}$ defined by an IFS with uniform contraction ratio r. Let $\nu^{(n)}$ be as above. If $\dim\mu < 1$ then*

$$\lim_{n\to\infty} \frac{1}{n'} H(\nu^{(n)}, D_{qn'}|D_{n'}) = 0 \quad \text{for every } q > 1. \tag{1}$$

Note: we only assume $\dim\mu < 1$. If $\dim\mu = s - \dim\mu$ the statement remians true, though for rather trivial reasons.

We end this brief summary with an interesting corollary of Theorem 1:

Corollary 3. *For IFS's on $\mathbb{R}$ defined by algebraic parameters there is a dichotomy: Either there are exact overlaps or the attractor X satisfies $\dim X = \min\{1, s - \dim X\}$.*

Note that this verifies the exact overlaps conjecture, introduced in a previous talk, for IFS's with algebraic parameters.

References

[1] M. Hochman, *On self-similar sets with overlaps and inverse theorems for entropy*, Annals of Mathematics. Second Series **180** (2014), 773–822.

Background on Bernoulli convolutions

Kornélia Héra

For $\lambda \in (0, 1)$, let ν_λ be the distribution of $\sum_{n=0}^{\infty} \pm\lambda^n$, where the signs are chosen independently with probability $\frac{1}{2}$. It can be written as the infinite convolution product $\nu_\lambda = *_{n=0}^{\infty} \frac{1}{2}(\delta_{-\lambda^n} + \delta_{\lambda^n})$, hence the measures ν_λ are called Bernoulli convolutions (BC). They have been studied since the 1930's, revealing surprising connections with harmonic analysis, the theory of algebraic numbers, dynamical systems, and Hausdorff dimension estimation. In this talk we touched on some of the main classical and modern results about BC, based on [7].

Jessen and Wintner (1935, [4]) showed that ν_λ is either absolutely continuous, or purely singular, depending on λ. Kershner and Wintner (1935, [6]) 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, and $\nu_{\frac{1}{2}}$ is uniform on $[-2, 2]$. The main question about BC is the following:

Question 1. *For which parameters $\lambda \in (\frac{1}{2}, 1)$ is ν_λ absolutely continuous? If absolute continuity holds, what can we say about the density?*

The following properties of BC are useful to understand their structure better. We will use the notation $f\mu$ for the push-forward of μ by f: $f\mu = \mu \circ f^{-1}$.

(1) ν_λ can be characterized as the unique probability measure satisfying
$$\nu_\lambda = \frac{1}{2} S_{-1}\nu_\lambda + \frac{1}{2} S_1 \nu_\lambda,$$
where $S_i(x) = \lambda x + i$ for $i = 1, -1$. Thus ν_λ is a self similar measure for the IFS $\{S_{-1}, S_1\}$ with weights $\frac{1}{2}$.

(2) Let $\Omega = \{-1, 1\}^{\mathbb{N}}$, $\mu = (\frac{1}{2}, \frac{1}{2})^{\mathbb{N}}$ the Bernoulli measure on Ω, and
$$\pi_\lambda : \Omega \to \mathbb{R},\ \omega \to \sum_{n=0}^{\infty} \omega_n \lambda^n.$$
Then $\nu_\lambda = \pi_\lambda \mu$. This point of view is useful when ideas from geometric measure theory are applied, in particular in the transversality method.

(3) The Fourier transform of a finite Borel measure ν on $\mathbb{R}$ is defined as $\hat{\nu}(u) = \int e^{-2\pi i u x} d\nu(x)$. Easy computation gives
$$\hat{\nu}_\lambda(u) = \prod_{n=0}^{\infty} \cos(2\pi \lambda^n u).$$
The formula is important when using methods related to number theory.

The first answer to Question 1 was given by Erdős in 1939.

Definition 1. *$\beta > 1$ is a Pisot number if it is an algebraic integer such that all other roots of its minimal polynomial have modulus less than 1.*

Theorem 1 (Erdős, 1939, [1])**.** *If $\lambda \in (0,1) \setminus \{\frac{1}{2}\}$, $\beta = \frac{1}{\lambda}$ is a Pisot number then ν_λ is singular.*

The next theorem of Erdős revealed that near 1, absolute continuity is generic.

Theorem 2 (Erdős, 1940, [2])**.** *There exists $\varepsilon > 0$ such that for almost all $\lambda \in (1-\varepsilon, 1)$, ν_λ is absolutely continuous.*

However, explicit bounds for the neighborhood of 1 were not given. Kahane (1971, [5]) gave a brief outline of the argument and indicated that
$$\dim\{\lambda \in (1-\varepsilon, 1) : \nu_\lambda \text{ is singular}\} \to 0 \text{ as } \varepsilon \to 0.$$
Here and in the sequel dim always denotes Hausdorff dimension. The following theorem has come to be known as the "Erdős-Kahane argument".

Theorem 3. *Let $a = \sqrt{2}, b = 2, k \geq 3$ be an arbitrary integer. Then*
$$\dim\{\lambda \in [b^{-1}, a^{-1}] : \hat{\nu}_\lambda(u) \neq O(u^{-0.02/k})\} \leq 3\frac{\log(3000k)}{k}.$$

The theorem can be formulated for arbitrary $1 < a < b < \infty$ with appropriate constants in the corresponding places. The proof is combinatorial in nature, an exposition of the argument can be found in [7].

Denote by P the family of measures on $\mathbb{R}$ which have a power Fourier decay at infinity, that is, there exist $\sigma, C > 0$ such that $|\hat{\mu}(u)| \leq C|u|^{-\sigma}$ for all $u \in \mathbb{R}$. Then theorem 3 easily implies:

Corollary 1. $\dim\{\lambda \in (0,1) : \nu_\lambda \notin P\} = 0$.

The following corollary can be derived from Theorem 3 using the convolution structure of ν_λ, namely that $\hat{\nu}_\lambda(u) = \hat{\nu}_{\lambda^2}(u) \cdot \hat{\nu}_{\lambda^2}(\lambda u)$.

Corollary 2. *For any $s > 0$ and $m \in \mathbb{N}$, there exists $\varepsilon > 0$ such that*

$$\dim\{\lambda \in (1-\varepsilon, 1) : \textit{the density } \frac{d\nu_\lambda}{dx} \notin C^m\} < s.$$

Garsia (1962, [3]) found an explicit set of parameters for which absolute continuity holds.

Definition 2. *$\beta > 1$ is a Garsia number if it is an algebraic integer such that all roots of its minimal polynomial have modulus greater than 1, and the minimal polynomial has constant term ± 2.*

Theorem 4 (Garsia, 1962, [3])**.** *If $\lambda \in (\frac{1}{2}, 1)$, $\beta = \frac{1}{\lambda}$ is a Garsia number then ν_λ is absolutely continuous.*

After Erdős's result from 1940 on generic absolute continuity near 1, an obvious question was whether ν_λ is absolutely continuous for almost all λ in the maximal possible interval $(\frac{1}{2}, 1)$. Solomyak [9] gave an affirmative answer in 1995, moreover showing that ν_λ has an L^2 density for almost all $\lambda \in (\frac{1}{2}, 1)$. Soon after, Peres and Solomyak [8] gave a simplified proof which avoided use of the Fourier transform. Both proofs rely on the so-called transversality method.

Theorem 5 (Solomyak, 1995, [9])**.** *For almost all $\lambda \in (\frac{1}{2}, 1)$, ν_λ is absolutely continuous with an L^2 density.*

In this talk we gave a brief outline of the proof from [8], by showing the basic ideas of transversality.

References

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

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

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

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

[5] J. P. Kahane, *Sur la distribution de certaines series aleatoires*, Colloque Th. Nombres [1969, Bordeaux], Bull. Soc. math. France, Memoire textbf25 (1971), 119–122.

[6] R. Kershner, A. Wintner, *On symmetric Bernoulli convolutions*, Amer. J. Math. **57** (1935), 541–548.

[7] Y. Peres, W. Schlag, and B. Solomyak, *Sixty years of Bernoulli convolutions*, In Fractal geometry and stochastics, II (Greifswald/Koserow, 1998), volume 46 of Progr. Probab., pages 39–65. Birkhäuser, Basel, 2000.

[8] Y. Peres, B. Solomyak, *Absolute continuity of Bernoulli convolutions, a simple proof*, Math. Res. Lett., 3(2) (1996), 231–239.
[9] B. Solomyak. *On the random series $\sum \pm\lambda^n$ (an Erdös problem)*, Ann. of Math. **142** (1995), 611–625.

Algebraic properties of Pisot and Salem numbers

Mark Pollicot

1. Definitions: Pisot and Salem numbers and the Mahler measure

Recall that an algebraic number $\beta > 1$ is a (real) root of a polynomial with integer coefficients

$$P(z) := a_d z^d + a_{d-1} z^{d-1} + \cdots + a_1 z + a_0.$$

with $a_0, \cdots, a_d \in \mathbb{Z}$ with $a_d, a_0 \neq 0$. An algebraic integer β is a root of a (minimal) polynomial with integer coefficients, where the leading coefficient is unity, i.e., $a_d = 1$. If $\beta > 1$ is an algebraic integer then its *conjugate roots* are the other $d-1$ roots $\alpha_1, \cdots, \alpha_{d-1}$ of the corresponding polynomial, i.e.,

$$P(z) = (z - \beta) \prod_{i=1}^{d-1} (z - \alpha_i).$$

Definition 1. *We say that λ is a Pisot number if its conjugate roots have modulus strictly smaller than unity, i.e., $|\alpha_i| < 1$, for $i = 1, \cdots, d-1$. A weaker condition is to say that λ is a Salem number if the conjugate roots have modulus less than or equal to unity, i.e., $|\alpha_i| \leq 1$, for $i = 1, \cdots, d-1$.*

We can also define the Mahler measure of any polynomial $P \in \mathbb{C}[z]$ (of degree d with leading coefficient a_d).

Definition 2. *The Mahler measure of $P(z)$ is given by*

$$m(P) = |a_d| \prod_{|\alpha_i|>1} |\alpha_i| \in \mathbb{R}^+$$

The Mahler measure is related to another value associated to P, called the height $h(P) = \max_i |a_i|$, by $m(P) \leq h(p)\sqrt{d+1}$.

Example 1. *The Mahler measure for $P_0(x) = x^{10} + x^9 - x^7 - x^6 - x^5 - x^4 - x^3 + x + 1$ takes the value $m(P) = 1.17628\ldots$.*

Indeed, this is the smallest known Mahler measure of a (non-trivial) polynomial.

2. The Lehmer conjecture

The best known conjecture on Mahler measures is due to Lehmer.

Conjecture 1 (Lehmer Conjecture)**.** *There exists $\delta > 1$ such that for any polynomial $P \in \mathbb{Z}[x]$ we have $m(P) \geq \delta$ or $m(P) = 1$ For trivial cases).*

Moreover, the Salem number in Example 2 is supposed to be the one for which $m(P) > 1$ is least.

3. Bernoulli convolutions

Given $\beta > 1$ we can write the Fourier transform of the Bernoulli convolution measure ν associated to $0 < \lambda = \beta^{-1} < 1$ as:

$$\widehat{\nu}_\lambda(t) = \prod_{n=0}^{\infty} \cos(\lambda^n t) \, t \in \mathbb{R}.$$

Recall that Erdös showed in 1939 that if β is Pisot then $\widehat{\nu}_\lambda(t)$ *does not* tend to zero (and thus by the Riemann Lebesgue lemma ν_λ is not absolutely continuous). For Salem numbers we have the following:

Lemma 1 ($\widehat{\nu}(t)$ doesn't decay polynomially)**.** *If λ is a Salem number then for $\widehat{\nu}(t)$ and any $\epsilon > 0$ we can choose $t_k \nearrow +\infty$ with $\widehat{\nu}(t_k) t_k^\epsilon \nearrow +\infty$.*

4. Garsia's lemma

We have the following bound (see [2])

Lemma 2. *Let $\beta > 1$ be an algebraic integer. Let $\alpha_1, \cdots, \alpha_{s-1}$ be the conjugate roots. Let $\sigma = \#\{1 \leq i \leq s-1 : |\alpha_i| = 1\}$. Assume $P(x) \in \mathbb{Z}[x]$ is a polynomial and height M and degree at most d with $P(\alpha) \neq 0$ then*

$$|P(\alpha)| \geq \frac{\prod_{|\alpha_i| \geq 1} ||\alpha_i| - 1|}{(d+1)^\sigma \left(\prod_{|\alpha_i|>1} |\alpha_i|\right)^{d+1} M^s}$$

Theorem 1 (Garsia's separation theorem)**.** *Let $1 < \beta < 2$ be a Pisot number corresponding to a polynomial of degree d, say. There exists $C = C(\beta, m) > 0$ such that if*

$$\sum_{j=1}^{n} \epsilon_j \beta^{-j} \neq \sum_{j=1}^{n} \epsilon_j' \beta^{-j}$$

for some $\epsilon_j, \epsilon_j' \in \{-1, 1\}$ then

$$\left|(\epsilon_j - \epsilon_j')\beta^{-j}\right| \geq C\beta^{-n}.$$

5. Garsia entropy

Let $0 < \lambda < 1$ be any real number. Let $0 < p_{-1} < 1$ and $p_1 = 1 - p_{-1}$ Consider the finite convolutions

$$\begin{aligned}
&(p_{-1}\delta_{-\lambda} + p_1\delta_\lambda) * (p_{-1}\delta_{-\lambda^2} + p_1\delta_{\lambda^2}) * \cdots * (p_{-1}\delta_{-\lambda^n} + p_1\delta_{\lambda^n}) \\
&= \sum_{i_1, \cdots, i_n \in \{-1,1\}} (p_{i_1} \cdots p_{i_n}) \delta_{(-1)^{i_1}\lambda + (-1)^{i_2}\lambda^2 + \cdots + (-1)^{i_n}\lambda^n}
\end{aligned}$$

which is supported on (distinct) points $x_1, x_2, \cdots, x_r$ $(r \leq 2^n)$ with weights $q_1, q_2, \cdots, q_r$, say.

Definition 3 (Garsia Entropy (see [3])). *We then define*

$$H_n = -\sum_{j=1}^{r} q_j \log q_j$$

and

$$H_\lambda = \lim_{N\to+\infty} \frac{H_N}{N}$$

The limit exists by a subaditivity argument. Clearly $H_\lambda \leq \log \lambda$.

Let us restrict to the case $p_{-1} = p_1 = \frac{1}{2}$ for simplicity.

Theorem 1 (Garsia). *If $\beta > 1$ is a Pisot number then for $\beta = \lambda^{-1}$ we have $H_\beta < \log \lambda$.*

In fact stronger inequalities are possible for Pisot numbers and the entropy is related to dimension of ν_λ.

6. Mahler's seperation lemma

If $P \in \mathbb{C}[z]$ is an complex polynomial of degree d with distinct roots then $z_1, \cdots, z_d$ then Mahler gave a lower bound on the seperation of roots

$$\sup_{i\neq j} |z_i - z_j| \geq \frac{\sqrt{3}\sqrt{|D(P)|}}{d^{d/2+1} M(P)^{d-1}}$$

where $\sqrt{|D(P)|} = |a_d^{d-1}| \prod_{i<j} |z_i - z_j|$.

Let $\beta > 1$ be an algebraic number.

Definition 4. *Let $\mathcal{P}_d$ denote the set of polynomials of degree at most d all of whose coefficients are $-1, 0, 1$.*

This has the following corollary.

Lemma 3 (see [1]). *Let $d \geq 9$. Let $\eta \neq \eta'$ be two algebraic numbers each of which is a root of a polynomial in $\mathcal{P}_d$. Then $|\eta - \eta'| > 2n^{-4n}$.*

7. Hochman's Theorem

We can consider the following theorem and question (see [4], Theorem 1.9 and Question 1.10).

Theorem 2 (Hochman). $\dim \nu_\lambda = 1$ *outside a set of λ of dimension 0. Furthermore, the exceptional parameters for which* $\dim \nu_\lambda < 1$ *are "nearly algebraic" in the sense that for every $0 < \theta < 1$ and all large enough d, there is a polynomial $p_d(x) \in \mathcal{P}_n$ such that $p_d(\lambda) < \theta^d$.*

Question 2 (Hochman). *Does there exist a constant $s > 0$ such that for α, β that are roots of polynomials in $\mathcal{P}_d$ either $\alpha = \beta$ or $|\alpha - \beta| > s^d$?*

The Lemma 3 above at least gives some lower bound (with $s = \frac{1}{d^4}$, dependent on d).

References

[1] Emmanuel Breuillard and Pter Varjú. On the dimension of Bernoulli convolutions. Preprint, arXiv:1610.09154, 2016.
[2] Adriano M. Garsia. Arithmetic properties of Bernoulli convolutions. Trans. Amer. Math. Soc., 102:409-432, 1962.
[3] Adriano M. Garsia. Entropy and singularity of infnite convolutions. Pacific J. Math., 13:1159-1169, 1963.
[4] Michael Hochman. On self-similar sets with overlaps and inverse theorems for entropy. Ann. of Math. (2), 180(2):773-822, 2014.
[5] Raphaël Salem. Algebraic numbers and Fourier analysis. D. C. Heath and Co., Boston, Mass., 1963

Shmerkin's theorem on smoothness of BC

Ariel Rapaport

1. The main Theorem

Given $\lambda \in (0,1)$ denote by ν_λ the unbiased Bernoulli convolution corresponding to the parameter λ.

Theorem 1 (P.Shmerkin, [1])**.** *There exists $E \subset (\frac{1}{2}, 1)$, with $\dim_H E = 0$, such that ν_λ is absolutely continuous for all $\lambda \in (\frac{1}{2}, 1) \setminus E$.*

2. Correlation dimension, energies, and the Fourier transform

Denote by $\mathcal{P}(\mathbb{R})$ the collection of all compactly supported Borel probability measures on $\mathbb{R}$. Given $\mu \in \mathcal{P}(\mathbb{R})$ the (lower) correlation dimension of μ is defined by,

$$\dim_c \mu = \liminf_{r \downarrow 0} \frac{\log \int \mu(B(x,r))\, d\mu(x)}{\log r} \,. \tag{1}$$

It always holds that $\dim_c \mu \le \dim_H \mu$. For $s \ge 0$ the s-energy of μ is defined by,

$$I_s \mu = \int \int \frac{d\mu(x)\, d\mu(y)}{|x-y|^s} \,.$$

It is not hard to verify that,

$$\dim_c \mu = \sup\{s \ge 0 \,:\, I_s\mu < \infty\} \,. \tag{2}$$

The Fourier transform of μ is defined by,

$$\widehat{\mu}(\xi) = \int e^{i\xi x}\, d\mu(x) \text{ for } \xi \in \mathbb{R} \,.$$

It is well known that for each $0 < s < 1$ there exists a constant $c(s) > 0$ such that,

$$I_s\mu = c_s \int |\xi|^{s-1} |\widehat{\mu}(\xi)|^2 \, d\xi \,. \tag{3}$$

3. Shmerkin's smoothing lemma

Denote by $\mathcal{D}(\mathbb{R})$ the class of measures in $\mathcal{P}(\mathbb{R})$ whose Fourier transform has at least power decay, i.e.

$$\mathcal{D}(\mathbb{R}) = \{\mu \in \mathcal{P}(\mathbb{R}) \,:\, \exists\, C, t > 0 \quad \text{s.t. } |\widehat{\mu}(\xi)| \leq C|\xi|^{-t} \quad \forall\, \xi \in \mathbb{R}\}\,.$$

Lemma 1. *Let $\nu \in \mathcal{D}(\mathbb{R})$ and $\mu \in \mathcal{P}(\mathbb{R})$ with $\dim_H \mu = 1$, then $\nu * \mu$ is absolutely continuous.*

Proof. Since $\nu \in \mathcal{D}(\mathbb{R})$ there exists $C > 0$ and $t \in (0, \frac{1}{2})$ with $|\widehat{\nu}(\xi)| \leq C|\xi|^{-t}$ for all $\xi \in \mathbb{R}$. For $n \geq 1$ let,

$$E_n = \{x \in \mathbb{R} \,:\, \mu(B(x,r)) \leq r^{1-t} \text{ for all } r \in (0, \frac{1}{n})\}\,.$$

From $\dim_H \mu = 1$ it follows $\mu(\cup_{n\geq 1} E_n) = 1$. For $n \geq 1$ with $\mu(E_n) > 0$ set $\mu_n = \frac{\mu|_{E_n}}{\mu(E_n)}$. It is not hard to show, directly from (1), that $\dim_c \mu_n \geq 1 - t$. From this and (2) it follows $I_{1-2t}\mu < \infty$. Now by (3),

$$\int |\widehat{\nu * \mu_n}(\xi)|^2 \, d\xi = \int |\widehat{\nu}(\xi)|^2 \cdot |\widehat{\mu_n}(\xi)|^2 \, d\xi$$

$$\int C^2 |\xi|^{-2t} \cdot |\widehat{\mu_n}(\xi)|^2 \, d\xi = \frac{C^2}{c_{1-2t}} \cdot I_{1-2t}\mu_n < \infty,$$

which shows that $\nu * \mu_n$ is absolutely continuous. Since this is true for every large enough $n \geq 1$ the lemma follows. $\square$

4. Proof of the main theorem

Proof of Theorem 1. Given $\lambda \in (\frac{1}{2}, 1)$ and $k \geq 2$ consider the IFS

$$\{f^w_{\lambda,k}(x) = \lambda^k x + \sum_{j=1}^{k-1} (-1)^{w_j} \lambda^j \,:\, w \in \{0,1\}^{k-1}\},$$

and let $\nu_{\lambda,k} \in \mathcal{P}(\mathbb{R})$ be with

$$\nu_{\lambda,k} = \sum_{w \in \{0,1\}^{k-1}} 2^{-k+1} \cdot f^w_{\lambda,k} \nu_{\lambda,k}\,.$$

Note that

$$\dim_s \nu_{\lambda,k} = \frac{\log 2^{k-1}}{\log \lambda^k} = (1 - \frac{1}{k}) \dim_s \nu_\lambda,$$

where $\dim_s$ is the similarity dimension. Also observe that

$$\nu_\lambda = \nu_{\lambda^k} * \nu_{\lambda,k}, \tag{4}$$

which follows from the fact that ν_{λ^k} is the distribution of $\sum_j \pm\lambda^{kj}$ and $\nu_{\lambda,k}$ is the distribution of $\sum_{j \nmid k} \pm\lambda^j$.

From Hochman's theorem, on parametric families of measures, it follows that there exists a set $E_k \subset (\frac{1}{2}, 1)$, with $\dim_H E_k = 0$, such that $\dim_H \nu_{\lambda,k} = \min\{1, \dim_s \nu_{\lambda,k}\}$ for all $\lambda \in (\frac{1}{2}, 1) \setminus E_k$.

By the Erdos-Kahane argument there exists a set $F \subset (0,1)$, with $\dim_H F = 0$, such that $\nu_\lambda \in \mathcal{D}(\mathbb{R})$ for all $\lambda \in (0,1) \setminus F$. Set

$$E = \cup_{k\geq 1}(E_k \cup \{\lambda \in (\frac{1}{2}, 1) \;:\; \lambda^k \in F\}),$$

then $\dim_H E = 0$. Fix $\lambda \in (\frac{1}{2}, 1) \setminus E$. Since $\dim_s \nu_\lambda > 1$ there exists $k \geq 1$ with $\dim_s \nu_{\lambda,k} > 1$. From $\lambda \notin E_k$ it follows $\dim_H \nu_{\lambda,k} = 1$. From $\lambda^k \notin F$ it follows $\nu_{\lambda^k} \in \mathcal{D}(\mathbb{R})$. By (4) and Lemma 1 it now follows that ν_λ is absolutely continuous, which completes the proof of the theorem. □

5. Related results

By using similar ideas it is possible to obtain the following results. For $u, x \in \mathbb{R}$ write $T_u(x) = ux$.

Theorem 2 (P.Shmerkin, [1])**.** *There exists a set $E \subset (0, \frac{1}{2})$, with $\dim_H E = 0$, such that the following holds. Let $a, b \in (0, \frac{1}{2})$ be such that $a \notin E$, $\frac{\log a}{\log b} \notin \mathbb{Q}$, and $\dim_H \nu_a + \dim_H \nu_b > 1$. Then $\mu^u_{a,b} := \nu_a * T_u \nu_b$ is absolutely continuous for all $u \in \mathbb{R} \setminus \{0\}$.*

Theorem 3 (P.Shmerkin and B.Solomyak, [2])**.** *Let $a, b \in (0, \frac{1}{2})$ be with $\dim_H \nu_a + \dim_H \nu_b > 1$. Then there exists a set $B \subset \mathbb{R}$, with $\dim_H B = 0$, such that $\mu^u_{a,b}$ is absolutely continuous for all $u \in \mathbb{R} \setminus B$.*

6. A counter example

The following theorem shows that it is not possible to remove the assumption $a \notin E$ from Theorem 2.

Theorem 4 (F.Nazarov, Y.Peres and P.Shmerkin, [3])**.** *Let $a, b \in (0, \frac{1}{2})$ be such that $\frac{\log a}{\log b} \notin \mathbb{Q}$ and a^{-1}, b^{-1} are both Pisot numbers. Then there exists a dense G_δ subset B of $(0, \infty)$, such that $\mu^u_{a,b}$ is singular for all $u \in B$.*

Remark 1. *Note that in the last theorem it is possible to take $a = \frac{1}{3}$ and $b = \frac{1}{4}$, in which case $\dim_H \nu_a + \dim_H \nu_b > 1$. Also, it is proven in the same paper that if $\frac{\log a}{\log b} \notin \mathbb{Q}$ then*

$$\dim_H \mu^u_{a,b} = \min\{1, \dim_H \nu_a + \dim_H \nu_b\} \textit{ for all } u \in (0, \infty)\,.$$

Proof. There exists a constant $c > 0$ such that,

$$|\widehat{\nu_a}(\pi a^{-n})|, |\widehat{\nu_b}(\pi b^{-n})| \geq c \text{ for all } n \in \mathbb{N}\,.$$

Since ν_b is compactly supported the function $\widehat{\nu_b}$ is K-Lipschitz for some $K > 1$. Write $\epsilon = \frac{c}{2\pi K}$, and for every $N \in \mathbb{N}$ let

$$V_N = \{u \in (0, \infty) \;:\; |ua^{-n} - b^{-m}| < \epsilon \text{ for some } n, m \geq N\}\,.$$

Clearly V_N is open in $(0, \infty)$. It is also dense in $(0, \infty)$, which follows from $\frac{\log a}{\log b} \notin \mathbb{Q}$. Set $B = \cap_{N\in\mathbb{N}} V_N$, then B is a dense G_δ subset of $(0, \infty)$.

Fix $u \in B$ and let $N \in \mathbb{N}$. Since $u \in V_N$ there exist $n, m \geq N$ with $|ua^{-n} - b^{-m}| < \epsilon$. We now have,

$$\begin{aligned}|\widehat{\mu^u_{a,b}}(\pi a^{-n})| &= |\widehat{\nu_a}(\pi a^{-n})| \cdot |\widehat{\nu_b}(\pi u a^{-n})| \\ &\geq c \cdot (|\widehat{\nu_b}(\pi b^{-m})| - |\widehat{\nu_b}(\pi b^{-m}) - \widehat{\nu_b}(\pi u a^{-n})|) \\ &\geq c \cdot (c - K \cdot |\pi b^{-m} - \pi u a^{-n}|) \geq c^2 - cK\pi\epsilon = \frac{c^2}{2} > 0\,.\end{aligned}$$

This shows that $\widehat{\mu^u_{a,b}}(\xi)$ does not tend to 0 as $\xi \to \infty$. Hence, by the Riemann–Lebesgue lemma, $\mu^u_{a,b}$ is not absolutely continuous. From this and the law of pure types the theorem follows. □

References

[1] Pablo Shmerkin. On the exceptional set for absolute continuity of Bernoulli convolutions. Geom. Funct. Anal., 24(3):946958, 2014.

[2] Pablo Shmerkin and Boris Solomyak. Absolute continuity of self-similar measures, their projections and convolutions. Trans. Amer. Math. Soc., 368:51255151, 2016.

[3] Fedor Nazarov, Yuval Peres, and Pablo Shmerkin. Convolutions of Cantor measures without resonance. Israel J. Math., 187:93116, 2012.

Varjú's theorem on smoothness of BC for algebraic parameters

Weikun He

This talk is based on the recent paper [3] due to Varjú. In this talk, log denotes the logarithm of base 2. Let $0 < \lambda < 1$. Denote by μ_λ the Bernoulli convolution with parameter λ. We have seen in a previous talk that for almost all $\lambda \in (1/2, 1)$, the BC μ_λ is absolutely continuous. However, showing absolute continuity for explicit parameter λ is a different problem. Here we consider algebraic parameters. For an algebraic number λ, its Mahler measure M_λ is defined to be the Mahler measure of its minimal polynomial. Previously, Garsia showed that μ_λ is absolutely continuous if λ^{-1} is an algebraic integer of Mahler measure 2. The main result of this talk is the following.

Theorem 1. *For any $\epsilon > 0$ there is $c > 0$ such that the following is true. Let $0 < \lambda < 1$ be an algebraic number. If*

$$1 - c \min(\log M_\lambda, (\log M_\lambda)^{-1-\epsilon}) < \lambda < 1,$$

then μ_λ is absolutely continuous with respect to the Lebesgue measure.

Note that in [3], Varjú showed Theorem 1 for biased Bernoulli convolution $\mu_{\lambda,p}$ (the constant c then depends on ϵ and p). Moreover, if we restrict λ to the set of algebraic numbers which are not roots of any polynomial with coefficients in $\{-1, 0, 1\}$, the condition on λ can be replaced by an entirely explicit bound

$$1 - 10^{-37}(\log(M_\lambda + 1))^{-1}(\log\log(M_\lambda + 2)^{-3}) < \lambda < 1,$$

Thus, Theorem 1 gives new explicit examples of λ for which μ_λ is absolutely continuous.

We outline the proof of Theorem 1. The main tool used in the proof is an averaged entropy at certain scale. Let X be a real valued random variable. Let $r > 0$ a real number. We define

$$H(X;r) = \int_0^1 H\Big(\Big\lfloor \frac{X}{r} + t \Big\rfloor\Big)\,\mathrm{d}t$$

where $H(\,\cdot\,)$ denotes the Shannon entropy. Moreover, for $0 < r < r'$, define

$$H(X;r \mid r') = H(X;r) - H(X;r').$$

This notion of entropy enjoys several nice properties among which is the following Garsia's absolute continuity criterion. A random variable is absolutely continuous with respect to the Lebesgue measure with class $L \log L$ density if and only if

$$\limsup_{r\to 0^+}\ \log(1/r) - H(X;r) < +\infty.$$

In light of this, the proof of Theorem 1 reduces to the proof of

$$1 - H(\mu_\lambda; 2^{-n} \mid 2^{-n+1}) \le \frac{1}{n^2} \tag{1}$$

for $n \ge 1$ sufficiently large under the assumption of Theorem 1. For an interval $I \subset (0,1]$, we write μ_λ^I for the distribution of the random variable

$$\sum_{n:\lambda^n \in I} \xi_n \lambda^n$$

where (ξ_n) is a sequence of independent and identically distributed Bernoulli random variables taking value in $\{-1,1\}$ with equal probability. With this notation, we have $\mu_\lambda = \mu_\lambda^{I_1} * \cdots * \mu_\lambda^{I_m} * \mu_\lambda^{(0,1]\setminus \cup_{i=1}^m I_i}$ whenever $I_1, \ldots, I_m$ are disjoint intervals.

To achieve (1), we start with the building blocks $\mu_\lambda^{(\lambda^l,1]}$ whose entropy at small scale can be bounded below in terms of the Garsia entropy h_λ which in turn is greater than $0.44\min(M_\lambda,1)$ by a result in [1]. Now using scale invariance and iterative convolution between these building blocks, we expect the entropy between certain scale 2^{-n} and 2^{-n+1} to gradually increase towards 1. The following two theorems quantify this growth. The first one is used at high entropy regime ($H(\mu;2^{-n},2^{-n+1})$ greater than 1 minus a small constant.)

Theorem 2. *Let μ and ν be compactly supported probability on $\mathbb{R}$. Let $\alpha \in (0,1/2)$ and $r > 0$ be real numbers. Assume that for all $s \in (\alpha^3 r, \alpha^{-3} r)$,*

$$1 - H(\mu; s \mid 2s) \le \alpha \text{ and } 1 - H(\nu; s \mid 2s) \le \alpha.$$

Then

$$1 - H(\mu * \nu; r \mid 2r) \le 10^8(-\log \alpha)^3 \alpha^2$$

The next theorem is used at low entropy regime ($H(\mu;2^{-n},2^{-n+1})$ is away from 0 but the previous theorem is not yet valid).

Theorem 3. *Given $\alpha \in (0, 1/2)$, there exists $c = c(\alpha) > 0$ such that the following is true. Let μ and ν be compactly supported probability on $\mathbb{R}$. For any negative integers $\sigma_2 < \sigma_1 < 0$ and any real number $\beta \in (0, 1/2]$, if*

$$\#\{\sigma \in \mathbb{Z} \cap [\sigma_2, \sigma_1] : 1 - H(\mu; 2^{\sigma} \mid 2^{\sigma+1}) < \alpha\} < c\beta(\sigma_1 - \sigma_2)$$

and

$$H(\nu; 2^{\sigma_2} \mid 2^{\sigma_1}) > \beta(\sigma_1 - \sigma_2),$$

then

$$H(\mu * \nu) \geq H(\mu; 2^{\sigma_2} \mid 2^{\sigma_1}) + \frac{c\beta}{-\log \beta}(\sigma_1 - \sigma_2) - 3.$$

References

[1] E. Breuillard and P. Varjú, *Entropy of Bernoulli convolutions and uniform exponential growth for linear groups*, Preprint, arXiv:1610.09154, 2015.

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

[3] P. Varjú, *Absolute continuity of Bernoulli convolutions for algebraic parameters*, Preprint, arXiv:1602.00261, 2016.

Breuillard-Varju's Inequality between Entropy and Mahler Measure

Or Landesberg

A *Bernoulli convolution* with parameter $\lambda \in (0, 1)$ is the distribution μ_λ of the following random power series $\sum_{k=0}^{\infty} \pm\lambda^k$ where the signs $\pm$ are chosen independently with equal probability. Denote the distribution of the finite sum $\sum_{0 \leq k \leq n-1} \pm\lambda^k$ by $\mu_\lambda^{(n)}$. The *random walk entropy* h_λ of μ_λ is defined to be:

$$h_\lambda := \lim_{n \to \infty} \frac{H(\mu_\lambda^{(n)})}{n}$$

where $H(\mu_\lambda^{(n)})$ is the Shannon entropy of the finitely supported measure $\mu_\lambda^{(n)}$.

Let $\pi_\lambda(x) = a_r \cdot \Pi_{i=0}^{i=r}(x - \lambda_i) = a_r x^r + a_{r-1}x^{r-1} + ... + a_0$ be the minimal polynomial in $\mathbb{Z}[x]$ of an algebraic number $\lambda \in \overline{\mathbb{Q}}$, with $\lambda_1, ..., \lambda_r$ its Galois conjugates (including $\lambda_1 = \lambda$). The *Mahler measure* of λ is defined to be $M_\lambda := |a_r| \cdot \Pi_{|\lambda_i > 1}|\lambda_i|$.

The main theorem presented in this talk was the following due to Emmanuel Breuillard and Peter Varju [1]:

Theorem 1. *There exists a positive constant $c > 0$ such that given any algebraic number λ:*

$$c \cdot \min(1, \log_2 \lambda) \leq h_\lambda \leq \min(1, \log_2 \lambda).$$

This constant can be taken to be $c = 0.44$.

A special case of Hochman's theorem on Bernoulli convolutions [3] connects the random walk entropy h_λ with algebraic parameter $\lambda \in (\frac{1}{2}, 1)$ to the Hausdorff dimension of the measure μ_λ:

$$\dim \mu_\lambda = \min\left(1, \frac{h_\lambda}{\log_2 \lambda^{-1}}\right)$$

It is a famous conjecture by Lehmer that the Mahler measure of all algebraic numbers is uniformly bounded away from 1 whenever λ is not 0 or a root of unity.

Theorem 2. *If the Lehmer conjecture holds, then there exists an $\varepsilon > 0$ such that for every real algebraic $1 - \varepsilon < \lambda < 1$ the dimension of the Bernoulli convolution $\dim \mu_\lambda = 1$.*

An outline of the proof of theorem 1 was given, emphasizing the role of the Gaussian averaged entropy of a random variable.

References

[1] Emmanuel Breuillard and Pter Varjú, *Entropy of Bernoulli convolutions and uniform exponential growth for linear groups*, Preprint, arXiv:1510.04043v2 (2016).

[2] Péter Varjú, *Recent progress on Bernoulli convolutions*, Preprint, arXiv:1608.04210 (2016).

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

Bernoulli convolutions with transcendental parameters, work of E. Breuillard and P. Varjú

Nicolas de Saxcé

Assume $(\xi_n)_{n\geq 0}$ is a sequence of ± 1 valued unbiased coin tosses. Given $\lambda \in (0, 1)$, we study the law μ_λ of the random variable $\sum_{n\geq 0} \xi_n \lambda^n$. We present the proof of the following result.

Theorem 1 (Breuillard-Varjú)**.** *The set $\{\lambda \in (1/2, 1) \mid \dim \mu_\lambda < 1\}$ is contained in the closure, for the usual topology on $\mathbb{R}$, of the set $\{\lambda \in \overline{\mathbb{Q}} \cap (1/2, 1) \mid \dim \mu_\lambda < 1\}$, where $\overline{\mathbb{Q}}$ denotes the algebraic closure of $\mathbb{Q}$ in $\mathcal{C}$.*

1. A question of Hochman

For $n \geq 1$, let $\mathcal{P}_n$ be the set of polynomials of degree less than n and with coefficients in $\{-1, 0, 1\}$, and let

$$E_n = \{\alpha \in \mathcal{C} \mid \exists P \in \mathcal{P}_n \,:\, P(\alpha) = 0\}.$$

In 2014, Hochman asked whether the following assertion was true:

$$\exists A \geq 0 : \ \forall \alpha \neq \beta \in \mathcal{P}_n, \ |\alpha - \beta| \geq e^{-An}. \tag{1}$$

As observed by Hochman, such a separation between the elements of E_n has a nice consequence on the dimension of Bernoulli convolutions.

Theorem 2 (Hochman)**.** *If* (1) *holds, then for any transcendental parameter* $\lambda \in (1/2, 1)$*,* $\dim \mu_\lambda = 1$*.*

We now present a short proof of this theorem, based on the ideas of Breuillard and Varjú. The following lemma will be crucial for the proof.

Lemma 1. *Let* $\mathcal{P} = \bigcap_n \mathcal{P}_n$*. For every* $\varepsilon > 0$*, there exists* $k = k(\varepsilon) \geq 0$ *such that every* $P \in \mathcal{P}$ *satisfying* $P(0) \neq 0$ *has at most* k *roots (with multiplicity) inside the disk* $B_{\mathbb{C}}(0, 1-\varepsilon)$*.*

Proof of Theorem 2, after Breuillard-Varjú. Assume (1) holds and let $\lambda \in (1/2, 1)$ such that $\dim \mu_\lambda < 1$. By Hochman's result on self-similar sets, this implies super-exponential decay of overlaps:

$$\forall C \geq 0,\ \exists N :\ \forall n \geq N,\ \exists P_n \in \mathcal{P}_n :\ |P_n(\lambda)| \leq e^{-Cn}.$$

Take $\varepsilon = \frac{1-\lambda}{2}$, let k be the number given by Lemma 1, and let $C = (A+1)k + \log \frac{1}{\varepsilon}$. Writing $P_n = X^r \prod_{i=1}^{n-r} X - \alpha_i$, with $|\alpha_i| > 1 - \varepsilon$ if $i > k$, we have

$$e^{-Cn} \geq |P_n(\lambda)| = \lambda^r \varepsilon \prod_{i=1}^{n-r} |\lambda - \alpha_i| \geq \lambda^r \varepsilon^{n-r} \prod_{i=1}^{k} |\lambda - \alpha_i|,$$

so there exists $i \in \{1, \ldots, k\}$ such that

$$|\lambda - \alpha_i| \leq \varepsilon^{-n/k} e^{-Cn/k} \leq e^{-n(A+1)}.$$

In short, we have shown:

$$\forall n \geq N,\ \exists \alpha \in E_n :\ |\alpha - \lambda| \leq e^{-n(A+1)}.$$

To obtain a contradiction, proceed as follows:
Let $n_0 = N$. Take $\alpha \in E_{n_0}$ such that $|\lambda - \alpha| := e^{-n_0 B} \leq e^{-n_0(A+1)}$.
Let $n_1 = \frac{n_0 B + 1}{A+1}$. Let $\beta \in E_{n_1}$ be such that $|\lambda - \beta| \leq e^{-n_1(A+1)}$. In particular, we must have $\beta \neq \alpha$. Then

$$|\alpha - \beta| \leq |\alpha - \lambda| + |\lambda - \beta| \leq e^{-n_0 B} + e^{-n_1(A+1)} \leq 2e^{-n_1(A+1)} < e^{-n_1 A}$$

contradicts (1), because $\alpha, \beta \in E_{n_1}$. □

2. A result of Mahler

The best separation result available for the roots of polynomials in $\mathcal{P}_n$ is due to Mahler.

Theorem 3 (Mahler)**.** *If* $\alpha \neq \beta$ *are elements in* E_n*, then* $|\alpha - \beta| \geq 2n^{-4n}$*.*

Exercise 1. *Using this separation result and mimicking the proof presented in the previous section, show that if the overlaps decay faster than* $e^{-Cn \log n}$ *then* $\lambda \in \overline{\mathbb{Q}}$*.*

A slightly weaker version of the following result of Breuillard and Varjú can be proved using Mahler's separation bound. From now on, if I is some subset of $\mathbb{R}_+$, we let μ_λ^I denote the law of the random variable $\sum_{n \geq 0; \lambda^n \in I} \xi_n \lambda^n$. Moreover, $H(\mu; r)$ denotes the entropy at scale r of the measure μ. Finally, for $\lambda \in (1/2, 1)$, h_λ denotes the Garsia entropy of λ.

Theorem 4 (Breuillard-Varjú). *Fix $\lambda \in (1/2, 1)$. There exists $c > 0$ such that for all n large enough, for all $r \in (0, n^{-3n})$, the following holds.*
Assume $H(\mu_\lambda^{(\lambda^n,1]}; r) < n$. Then there exists $\eta \in E_n$ such that $|\eta - \lambda| \leq r^c$ and $h_\eta \leq \frac{1}{n} H(\mu_\lambda^{(\lambda^n,1]};) < 1$.

Sketch of proof. Consider the family $\mathcal{A} = \{P \in \mathcal{P}_n \mid |P(\lambda)| \leq r\}$. Using some effective version Euclidean algorithm, write the greatest common divisor D of all polynomials in $\mathcal{A}$:

$$D = P_1 Q_1 + \cdots + P_\ell Q_\ell,$$

with $P_i \in \mathcal{A}$, $\ell \leq n$, $h(Q_i) \leq 2^n (2n)!$ and $\deg Q_i \leq n$. By definition of $\mathcal{A}$, this shows that

$$|D(\lambda)| \leq n^2 2^n (2n)! r \leq r^{1/5},$$

which implies that there exists a root η of D such that $|\eta - \lambda| \leq r^c$, for some small constant $c > 0$ depending only on λ. (The detailed argument uses Lemma 1 and is similar to the one presented in the first section.)

For $a \in \mathbb{N}$, let $\Omega_a = \{(\omega_0, \ldots, \omega_{n-1}) \mid \sum \omega_i \lambda^i \in [ar, (a+1)r)\}$, so that

$$H(\mu_\lambda^{(\lambda^n,1]}; r) = \sum \frac{|\Omega_a|}{2^n} \log \frac{2^n}{|\Omega_a|}.$$

If $\omega \neq \omega' \in \Omega_a$, then $P = \sum \frac{\omega_i - \omega_i'}{2} x^j$ is in $\mathcal{A}$, so $P(\eta) = 0$. Therefore $\sum \omega_i \eta^i = \sum \omega_i' \eta^i$, and

$$H(\sum_i = 0^{n-1} \xi_i \eta^i) \leq \sum \frac{|\Omega_a|}{2^n} \log \frac{2^n}{|\Omega|} = H(\mu_\lambda^{(\lambda^n,1]}; r).$$

Hence, $h_\lambda = \inf_m \frac{1}{m} H(\sum^{m-1} \xi_i \eta^i) \leq \frac{1}{n} H(\mu_\lambda^{(\lambda^n,1]}; r)$. □

3. Increasing entropy using self-similarity

We now want to sketch the proof of Theorem 1. The interested reader is referred to the original paper of Breuillard and Varjú for the detailed proof. The idea is to use Varjú's inverse theorem for entropy and the self-similarity property of μ_λ to increase the lower bound on entropy proved in the previous section.

Let $\lambda \in (1/2, 1)$ be such that $\dim \mu_\lambda < 1$ and assume for a contradiction that there the set of $\eta \in \overline{\mathbb{Q}}$ satisfying $\dim \mu_\eta < 1$ is bounded away from λ by a small number τ. Fix $\varepsilon > 0$ such that $\dim \mu_\lambda < 1 - 4\varepsilon$.
First step. Choose n_0 such that:

- $\forall r \leq \lambda^{n_0}$, $\frac{H(\mu_\lambda; r)}{\log 1/r} \leq \dim \mu_\lambda + \varepsilon$ (exact-dimensionality)
- $\forall n \geq n_0$, $H(\mu_\lambda^{(\lambda^n,1]}; \lambda^{10n} | \lambda^n) \leq n\varepsilon \log 1/\lambda$ (Hochman)
- $n_0^{-3n_0/c} < \tau$ (separation from exceptional algebraic parameters).

We claim that

$$H(\mu_\lambda^{(\lambda^{n_0},1]}; n_0^{-3n_0} | \lambda^{10n_0}) \geq \varepsilon n_0 \log 1/\lambda.$$

Indeed, otherwise, we would have

$$\begin{aligned} H(\mu_\lambda^{(\lambda^{n_0},1]}; n_0^{-3n_0}) &\leq \varepsilon n_0 \log 1/\lambda + H(\mu_\lambda^{(\lambda^{n_0},1]}; \lambda^{10n_0}|\lambda^{n_0}) + H(\mu_\lambda^{(\lambda^{n_0},1]}; \lambda^{n_0}) \\ &\leq n_0(\log 1/\lambda)(\dim \mu_\lambda + 3\varepsilon) \\ &< n_0. \end{aligned}$$

But by Theorem 4, there would exist $\eta \in \overline{\mathbb{Q}}$ such that $|\eta - \lambda| < n_0^{-3n_0/c}$ and $\dim \mu_\eta \leq \frac{h_\eta}{\log 1/\eta} \leq \dim \mu_\lambda + 4\varepsilon < 1$, contradicting our absurd assumption on λ.

Assuming n_i has been defined, define K_i and n_{i+1} so that one has

$$\begin{cases} \lambda^{K_i n_i} = n_i^{-3n_i} \\ n_{i+1} = K_i n_i \end{cases}$$

This way one obtains an increasing sequence $n_0, n_1, \ldots$ of integers such that for each i,

$$H(\mu_\lambda^{(\lambda^{n_i},1]}; n_i^{-3n_i}|\lambda^{10n_i}) \geq \varepsilon n_i \log 1/\lambda.$$

Second step. Using self-similarity and Varjú's inverse theorem for entropy, we deduce from the above lower bound that for some $c > 0$ depending only on λ, for all r small enough and all $p \in \{2, \ldots, n_i\}$,

$$H(\mu_\lambda^{(\lambda^{n_i} r, \lambda^{K_i n_i} r]}; r|\lambda^p r) \geq \frac{cp}{(\log K_i)(\log^{(2)} K_i)}.$$

Third step. Since the intervals $(\lambda^{n_i} r, \lambda^{K_i n_i} r]$ are disjoint, we may apply again Varjú's theorem and get

$$H(\mu_\lambda^{(\lambda^{n_0} r, \lambda^{-K_N n_N} r]}; r|\lambda^p r) \geq cp \sum_{i \leq N} \frac{1}{(\log K_i)(\log^{(2)} K_i)^2}.$$

Conclusion. To conclude, it suffices to show that the sum on the right-hand side diverges. For that, we check by induction that $\log K_i \leq \sqrt{i + i_0}$, where $i_0 = (\log K_0)^2$. This is clear for $i = 0$, and then, write

$$\begin{aligned} \log K_{i+1} &= (\log K_i) + \log(1 + C\frac{\log K_i}{K_i}) \\ &\leq (\log K_i) + C\frac{\log K_i}{K_i} \\ &= (\log K_i)(1 + \frac{C}{K_i}) \\ &\leq \sqrt{i + i_0 + 1}. \end{aligned}$$

(In the last line, we used the induction hypothesis and the fact that $t \mapsto (\log t)(1 + \frac{C}{t})$ is an increasing function of t for t large enough.)

4. The refined version of the theorem

In the paper by Breuillard and Varjú, the application of Varjú's inverse theorem is done more carefully, in order to have a quantitative result about the approximation rate by elements of E_n of the parameter λ satisfying $\dim \mu_\lambda < 1$. Their result is as follows.

Theorem 5 (Breuillard-Varjú). *Let $\lambda \in (1/2, 1]$ be such that $\dim \mu_\lambda < 1$. Then, for all $\varepsilon > 0$, there exists $A \geq 0$ such that for all d_0 sufficiently large, there exists $d \in [d_0, exp^{(5)}(\log^{(5)}(d_0) + A)]$ and $\eta \in E_{d,\dim_{\mu_\lambda}+\varepsilon}$ such that*

$$|\lambda - \eta| \leq \exp(-d^{\log^{(3)} d}).$$

As a corollary, we find the first explicit examples of transcendental parameters λ for which $\dim \mu_\lambda = 1$.

Corollary 1. *Let $\lambda \in (1/2, 1)$ be a number such that for all $n \gg 1$ and all $P \in P_n$, $|P(\lambda)| > \exp(-d^{\log^{(3)} d})$. Then $\dim \mu_\lambda = 1$. In particular, for $\lambda \in \{\ln 2, e^{-1/2}, \pi/4\}$, one has $\dim \mu_\lambda = 1$.*

Orponen's Distance Set Theorem

Demi Allen

Given a planar set $K \subset \mathbb{R}^2$ we consider its *distance set*

$$D(K) := \{|x - y| : x, y \in K\}.$$

Specifically, we are interested in how the sizes of K and $D(K)$ are related. We begin by surveying some classical results in this area. When K is a finite set, this problem is the essence of a question asked by Erdős.

Question (Erdős). *What is the minimum number of different distances, $f(n)$, determined by n distinct points in the plane?*

Erdős himself gave the following bounds for the number $f(n)$.

Theorem (Erdős [2], 1946). *The minimum number $f(n)$ of distances determined by n points of a plane satisfies the inequalities*

$$\left(n - \frac{3}{4}\right)^{\frac{1}{2}} - \frac{1}{2} \leq f(n) \leq \frac{cn}{\sqrt{\log n}},$$

where c is some universal constant.

A conjecture attributed to Erdős is that the lower bound for $f(n)$ given above should match the upper bound in the following sense.

Erdős Distance Conjecture. *The minimum number $f(n)$ of distances determined by n points of a plane satisfies*

$$f(n) \gtrsim \frac{n}{\sqrt{\log n}},$$

where we write $A \lesssim B$ to mean that there exists some universal constant $c > 0$ such that $A < cB$.

Some remarkable progress has been made recently by Guth and Katz towards proving this conjecture.

Theorem (Guth – Katz [5], 2015). *The quantity $f(n)$ satisfies*

$$f(n) \gtrsim \frac{n}{\log n}.$$

In considerations of the "continuous" version of the question posed by Erdős, notions of measure and dimension make an appearance. An early result relating to the "continuous" problem is the following theorem of Steinhaus.

Theorem (Steinhaus [10], 1920). *If $K \subset \mathbb{R}^2$ is a planar set with positive 2-dimensional Lebesgue measure, then the distance set $D(K)$ contains an interval $[0, \varepsilon)$ for some $\varepsilon > 0$.*

In the 1980s, Falconer proved a result relating the Hausdorff dimension of a set $K \subset \mathbb{R}^2$ with that of its distance set $D(K)$.

Theorem (Falconer [3], 1985). *If $K \subset \mathbb{R}^2$ is any set, then*

$$\dim_H D(K) \geq \dim_H K - 1.$$

The continuous analogue of the Erdős Distance Conjecture is attributed to Falconer.

Falconer Distance Conjecture. *If $K \subset \mathbb{R}^2$ is a Borel set with $\dim_H K > 1$, then $D(K)$ has positive length, i.e. $\mathcal{H}^1(D(K)) > 0$ (where, for $s \geq 0$, $\mathcal{H}^s$ is the usual Hausdorff s-measure).*

The current best known results towards the Falconer Distance Conjecture for general sets K are due to Wolff and Bourgain.

Theorem (Wolff [11], 1999). *If $K \subset \mathbb{R}^2$ is Borel with $\dim_H K > \frac{4}{3}$, then $D(K)$ has positive length.*

Theorem (Bourgain [1], 2003). *If $K \subset \mathbb{R}^2$ is Borel with $\dim_H K \geq 1$, then $\dim_H D(K) \geq \frac{1}{2} + \varepsilon$ for some (small) absolute constant $\varepsilon > 0$.*

While these theorems of Wolff and Bourgain represent the best known progress towards proving Falconer's Distance Conjecture for general planar sets K, in recent years there has been an increased interest in proving Falconer's Distance Conjecture or improving upon the results of Wolff and Bourgain for special classes of sets. For example, the problem for self-similar sets has been studied by Orponen [6] (2012), and self-affine sets have been considered by Ferguson, Fraser and Sahlsten [4] (2015). Very recently, the class of Ahlfors–David regular sets has also been studied by Orponen [7] (2017) and Shmerkin [8, 9] (2017).

Definition. *A Borel measure μ on $\mathbb{R}^d$ is said to be (s, A)*-Ahlfors–David regular *if*

$$\frac{r^s}{A} \leq \mu(B(x, r)) \leq Ar^s$$

for all $x \in \operatorname{spt}\mu$ and $0 < r \leq \operatorname{diam}(\operatorname{spt}\mu)$ for some constant $A \geq 1$. An $\mathcal{H}^s$-measurable set $K \subset \mathbb{R}^d$ is said to be (s, A)-Ahlfors–David regular if $0 < \mathcal{H}^s(K) < \infty$ and the restriction $\mathcal{H}^s|_K$ of $\mathcal{H}^s$ to K is (s, A)-Ahlfors–David regular.

For this class of sets, Orponen proved the following result regarding the upper box dimension of the corresponding distance sets.

Theorem (Orponen [7], 2017)**.** *Assume that $\emptyset \neq K \subset \mathbb{R}^2$ is a bounded $\mathcal{H}^s$-measurable (s, A)-Ahlfors–David regular set with $s \geq 1$. Then*

$$\overline{\dim}_B D(K) = 1,$$

where $\overline{\dim}_B$ denotes the upper box dimension.

We devote a large part of this talk to discussing the proof of this theorem, which relies on a careful covering argument and also employs several properties of entropy. In particular, the proof uses a projection theorem for entropy and a multi-scale decomposition of entropy to bound the entropy of projections.

References

[1] J. Bourgain, *On the Erdős-Volkmann and Katz-Tao ring conjectures*, Geom. Funct. Anal. **13** (2003), no. 2, 334–365.

[2] P. Erdős, *On sets of distances of n points*, Amer. Math. Monthly **53** (1946), 248–250.

[3] K. J. Falconer, *On the Hausdorff dimensions of distance sets*, Mathematika **32** (1985), no. 2, 206–212.

[4] A. Ferguson, J. M. Fraser, T. Sahlsten, *Scaling scenery of (m,n) invariant measures*, Adv. Math. **268** (2015), 564–602.

[5] L. Guth, N. H. Katz, *On the Erdős distinct distances problem in the plane*, Ann. of Math. (2) **181** (2015), no. 1, 155–190.

[6] T. Orponen, *On the distance sets of self-similar sets*, Nonlinearity **25** (2012), no. 6, 1919–1929.

[7] T. Orponen, *On the distance sets of Ahlfors–David regular sets*, Adv. Math. **307** (2017), 1029–1045.

[8] P. Shmerkin, *On distance sets, box-counting and Ahlfors regular sets*, Discrete Anal. 2017, Paper No. 9, 22 pp.

[9] P. Shmerkin, *On the Hausdorff dimension of pinned distance sets*, preprint, (2017), arXiv:1706.00131

[10] H. Steinhaus, *Sur les distances des points dans les ensembles de mesure positive*, Fund. Math. **1** (1920), 93–104.

[11] T. Wolff, *Decay of circular means of Fourier transforms of measures*, Internat. Math. Res. Notices (1999), no. 10, 547–567.

Balog-Szemerédi-Gowers Theorem

Oleg Pikhurko

The Balog-Szemerédi-Gowers Theorem [2, 3] is a very powerful tool in additive combinatorics which, roughly speaking, states that any two sets A, B in an Abelian group Γ with a large fraction of sums $a+b$, $(a,b) \in A \times B$, concentrated on few values necessarily have large subsets $A' \subseteq A$ and $B' \subseteq B$ with

$$A' + B' := \{a + b \mid a \in A',\ b \in B'\},$$

the set of **all** possible sums, having small size. It was first proved by Balog and Szemerédi [2] using Szemerédi's Regularity Lemma for graphs, with the quantitative dependences between the parameters being rather bad. Later, Gowers [3] found a proof that avoided regularity and gave polynomial dependence between the parameters. A detailed discussion of the Balog-Szemerédi-Gowers Theorem can be found in the book of Tao and Vu [4].

One version of the Balog-Szemerédi-Gowers Theorem is as follows (see [4, Theorem 2.29]). Given $G \subseteq A \times B$, define the *partial sumset*

$$A \overset{G}{+} B := \{a + b \mid (a, b) \in G\}.$$

Theorem 1 (Balog-Szemerédi-Gowers Theorem (Symmetric Version))**.** *Let A, B be subsets of an Abelian group Γ. Let $G \subseteq A \times B$ and reals $K \geq 1$ and $K' > 0$ satisfy:*

$$\begin{aligned} |G| &\geq \frac{|A| \cdot |B|}{K}, \\ \left| A \overset{G}{+} B \right| &\leq K' \, |A|^{1/2} \, |B|^{1/2}. \end{aligned}$$

Then there are $A' \subseteq A$ and $B' \subseteq B$ such that

$$\begin{aligned} |A'| &\geq \frac{|A|}{4\sqrt{2}\,K}, \\ |B'| &\geq \frac{|B|}{4K}, \\ |A' + B'| &\leq 2^{12}\, K^5\, (K')^3\, |A|^{1/2}\, |B|^{1/2}. \end{aligned}$$

In fact, there are other properties that can be attained in the conclusion of Theorem 1. These are discussed by Balog [1] who concentrates on the case when $|A|, |B|, \left|A \overset{G}{+} B\right| \leq N$ and $|G| \geq N^2/K$, and shows that one can attain $|A' - B'| \leq O(K^7 N)$, $|A' - A'| \leq O(K^5 N)$, $|A' + A'| \leq O(K^7 N)$, and $|(A' \times B') \cap G| \geq \Omega(N^2/K^4)$.

The above theorem is particularly useful if the sizes of A and B are within constant factor of each other, since then it can be combined with the Freiman-Ruzsa Theorem to derive a very strong structural information about the obtained sets A' and B'. If $|A|$ is significantly larger than $|B|$, then the direct application of Theorem 1 to the sets A and B is not very useful as then

$$KK' \geq \left| A \overset{G}{+} B \right| \frac{|A|^{1/2}\,|B|^{1/2}}{|G|} \geq \frac{|A|^{1/2}}{|B|^{1/2}}$$

has to be large. There is an "asymmetric" version of the Balog-Szemerédi-Gowers Theorem designed for such cases, see Theorem 2.35 in [4], which can be derived with some work from Theorem 1. For its statement and proof, see Section 2.6 in [4].

References

[1] A. Balog, *Many additive quadruples*, Additive combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, pp. 39–49.

[2] A. Balog and E. Szemerédi, *A statistical theorem on set addition*, Combinatorica **14** (1994), 263–268.

[3] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four*, Geom. Func. Analysis **8** (1998), 529–551.

[4] T. Tao and V. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge Univ. Press, 2006.

Bourgain's sum-product and projection theorems. Part I

Giorgis Petridis

The topic of this expository talk is a sum-product result of Bourgain, where the "largeness" of a finite set is measured by covering number. Applications of Borgain's theorem are discussed in the subsequent talk.

The classical sum-product phenomenon relies on "largeness" to measure the size of a set. The best known of its many manifestations is that every finite set must either have a large number of pairwise sums or a large number of pairwise products. Let $A \subset \mathbb{R}$ and denote by

$$A + A = \{a_1 + a_2 : a_1, a_2 \in A\}$$

its sum set and by AA the similarly defined product set. Konyagin and Shkredov [6], building on work of Solymosi [7], prove that at least one of $|A + A|$ and $|AA|$ must be much larger than $|A|$: there exist positive constants c and ε, with ε approximately equal to $1/5,000$, such that for all finite sets $A \subset \mathbb{R}$ we have

$$\max\{|A + A|, |AA|\} \geq c|A|^{4/3+\varepsilon}.$$

The conjectured exponent is 2. A different manifestation is to prove that sets like

$$AA + AA = \{a_1a_2 + a_3a_4 : a_1, \ldots, a_4 \in A\}$$

have cardinality much larger than $|A|$. The best known result is due to Iosevich, Roche-Newton, and Rudnev, who show (up to logarithms) that $|AA + AA| \geq c|A|^{19/12}$ [4].

Bourgain, inspired by questions in geometric measure theory, considered a sum-product question for finite sets

$$A \subset \delta\mathbb{Z} = \{\delta n : n \in \mathbb{Z}\},$$

for some absolute $\delta > 0$. He uses the δ-covering number to measure how "large" sets like $A, A + A, AA, \ldots$ are. Given a finite set $S \subseteq \mathbb{R}$, he denotes by $N(A, \delta)$ the minimum number of intervals in $\delta\mathbb{Z}$ that S intersects. While it is true that for $A \subset \delta\mathbb{Z}$, we have $N(A, \delta) = |A|$, and even $N(A + A, \delta) = |A + A|$, it is not true that $N(AA, \delta) = |AA|$. We do however have $|AA| > N(AA, \delta)$. In this sense Bourgain proved a strong sum-product theorem by establishing something like:

$$\max\{N(A + A, \delta), N(AA, \delta)\} \geq cN(A, \delta)^{1+\varepsilon}.$$

Before making the above heuristic statement precise, we note that taking $A = \delta\mathbb{Z} \cap [1, x)$ (for any $x > \delta$) shows that the above is false, because all three covering numbers are of the order of magnitude of $x\delta^{-1}$. This means that some care must be taken to make sure that the sets considered are not similar to the intersection of $\delta\mathbb{Z}$ and an interval. Here is Bourgain's theorem from [2]. $\ell(\cdot)$ denotes Lebesgue measure.

Theorem 1 (Bourgain)**.** *Let $\alpha > 0$ and κ be positive absolute constants. There exist ε_0 and ε_1 with the following property.*

For all $\delta > 0$ and all $A \subset \delta Z$ of cardinality $|A| \geq \delta^{\alpha}$ such that

$$|A \cap I| \leq \ell(I)^{\kappa}|A|$$

for all intervals I of length $\delta < \ell(I) < \delta^{\varepsilon_0}$, we have

$$\max\{N(A + A, \delta), N(AA, \delta)\} \geq \delta^{-\varepsilon_1}N(A, \delta).$$

The theorem is deduced from the following more general result from [2], which Bourgain also applies to projection theorems for fractal sets.

Theorem 2 (Bourgain)**.** *Let $\alpha > 0$ and κ be positive absolute constants. There exist ε_0 and ε_1 with the following property.*

For all $\delta > 0$, and all probability measures μ supported on $[0, 2]$ with the property

$$\mu(I) \leq \ell(I)^{\kappa}$$

for all intervals I of length $\delta < \ell(I) < \delta^{\varepsilon_0}$, and all $A \subset \delta Z$ of cardinality $|A| \geq \delta^{\alpha}$ such that

$$|A \cap I| \leq \ell(I)^{\kappa}|A|$$

for all intervals I of length $\delta < \ell(I) < \delta^{\varepsilon_0}$, there exists $x \in \mathrm{support}(\mu)$ such that

$$N(A + xA, \delta) \geq \delta^{-\varepsilon_1}N(A, \delta).$$

The deduction of Theorem 1 from Theorem 2 relies on methods from additive combinatorics presented in the seventh talk of this study group and is reminiscent of arguments of Gowers in [3] and Katz and Tao in [5]. A key step is to show that if both $N(A + A, \delta)$ and $N(AA, \delta)$ are comparable to $N(A, \delta)$, then $N(AA + AA, \delta)$ is also comparable to $N(A, \delta)$. This leads to a contradiction by a robust generalisation of the result of Iosevich, Roche-Newton, and Rudnev [4].

The proof of Theorem 2 is very intricate. It also uses techniques from additive combinatorics, but the key ingredient is a tree-structure theorem, which in a way is an inverse result for an example discussed in the eight talk of this workshop. Bourgain proves that if $N(A+A,\delta)$ is comparable to $N(A,\delta)$, then A contains a fairly large subset with a tree-structure similar to that of the sets discussed in the eighth lecture on Hochman's inverse theorem for entropy. This inverse result was used by Shmerkin in the work discussed in the final two talks of this workshop.

References

[1] J. Bourgain, *On the Erdős-Volkmann and Katz-Tao ring conjectures*, Geom. Funct. Anal. **13** (2003), 334–365.

[2] J. Bourgain, *The discretized sum-product and projection theorems*, J. Anal. Math. **112** (2010), 193–1236.

[3] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length four.*, Geom. Funct. Anal. **8** (1998), 529–551.

[4] A. Iosevich, O. Roche-Newton, M Rudnev, *On discrete values of bilinear forms*, preprint.

[5] N. Katz, T. Tao, *Some connections between the Falconer and Furstenburg conjectures*, New York J. Math. **7** (2001), 148–187.

[6] S. V. Konyagin, I. D. Shkredov, *New results on sum-products in* $\mathbb{R}$, Tr. Mat. Inst. Steklova **294** (2016), 87–98.

[7] J. Solymosi, *Bounding multiplicative energy by the sumset*, Adv. Math. **222** (2009), 402–4088.

Bourgain's sum-product and projection theorems. Part II.

Andras Mathe

Given a Borel set in the plane of a certain (Hausdorff) dimension, what can we say about the (Hausdorff) dimension of its orthogonal projections to various lines in the plane? Answers to questions of this type are called projection theorems.

This talk presented Bourgain's projection theorem and its background, and indicated how the proof follows from his sum-product theorems [1]. The sum-product theorems were the subject of the previous talk.

Let π_θ denote the orthogonal projection from the plane to the line in direction θ. Hausdorff dimension is denoted by $\dim_H$. The classical projection theorem is the following.

Theorem 1 (Marstrand)**.** *Let $E \subset \mathbb{R}^2$ be a Borel set with $\dim_H E = \alpha$. Then for almost every direction θ,*

$$\dim_H \pi_\theta(E) = \min(1, \alpha).$$

Moreover, if $\alpha > 1$ then $\pi_\theta(E)$ has positive Lebesgue measure for almost all θ.

Theorem 2 (Falconer)**.** *Let $E \subset \mathbb{R}^2$ be Borel with $\dim_H E = 1 + t$. Then the dimension of the "exceptional set of directions" is*

$$\dim_H\{\theta : \dim_H \pi_\theta(E) < 1\} \leq 1 - t.$$

Upper box dimension of product sets satisfies the inequality

$$\overline{\dim}_B(X \times Y) \le \overline{\dim}_B X + \overline{\dim}_B Y.$$

From this we immediately obtain

Observation 3. *Let $E \subset \mathbb{R}^2$ have upper box dimension α. Then there is at most one direction in which the projection has upper box dimension less than $\alpha/2$.*

One could also replace box dimension with Hausdorff dimension in the previous observation: then the exceptional set of directions will be of zero Hausdorff dimension.

In the other direction, we have the following construction.

Theorem 4 (Kaufman, Mattila)**.** *Let $0 \le \gamma \le 1$, $\gamma \le \alpha \le 2-\gamma$. There is a Borel set $E \subset \mathbb{R}^2$ such that $\dim_H E = \alpha$ and*

$$\dim_H\{\theta \,:\, \dim_H \pi_\theta(E) < (\alpha+\gamma)/2\} = \gamma.$$

It is not known whether this result is sharp in general. It is sharp if $\gamma = 0$ or if $\alpha = 1+t$, $\gamma = 1-t$, or if $\alpha = \gamma$.

Now we state Bourgain's projection theorem.

Theorem 5 (Bourgain [1])**.** *For every $0 < \alpha < 2$ and $\gamma > 0$ there is $\varepsilon > 0$ with the following property. Let $E \subset \mathbb{R}^2$ be Borel with $\dim_H E \ge \alpha$. Then*

$$\dim_H\{\theta \,:\, \dim_H \pi_\theta(E) < \alpha/2 + \varepsilon\} \le \gamma.$$

The proof of this theorem relies on the following discretised sum-product theorem.

Theorem 6 (Bourgain [1])**.** *Given $0 < \sigma < 1$ and $\gamma > 0$, there exist $\varepsilon > 0$ and $\varepsilon_0 > 0$ such that the following holds for $\delta > 0$ sufficiently small.*

Let μ be a probability measure on $[0,1]$ satisfying

$$\mu([x,y]) \le |y-x|^\gamma$$

whenever $\delta < |y-x| < \delta^{\varepsilon_0}$. Let $A \subset [1,2]$ be a discrete set consisting of δ-separated points satisfying

$$|A| \ge \delta^{-\sigma}$$

such that also

$$|A \cap [x,y]| \le |y-x|^\gamma |A|$$

whenever $\delta < |y-x| < \delta^{\varepsilon_0}$. Then there exists $x \in \operatorname{supp}\mu$ such that

$$N(A + xA,\, \delta) > \delta^{-\varepsilon}|A|$$

where $N(X,\, \delta)$ denotes the minimum number of δ-intervals needed to cover the set X.

In the non-discrete world this theorem roughly says the following: For every $0 < \sigma < 1$ and $\gamma > 0$ there is $\varepsilon > 0$ such that whenever $A \subset \mathbb{R}$ has dimension at least σ, and $S \subset \mathbb{R}$ has dimension at least γ, then there is $x \in S$ such that $A + xA$ has dimension at least $\sigma + \varepsilon$. However, this is vague (we did not specify what we mean by dimension), and the discrete version is actually much more powerful than this analogue suggests.

Notice that $A + xA$ is essentially the projection of $A \times A$ to the line of slope x.

Theorem 7 (Bourgain [1]). *Given $0 < \alpha < 2$, $\alpha' > 0$ and $\gamma > 0$, there exist $\varepsilon_0 > 0$ and $\varepsilon > 0$ such that the following holds.*

Let μ be a probability measure on S^1 (set of directions in the plane) such that the μ-measure of every interval (arc) of length ℓ is at most $C\ell^\gamma$. [This implies that the support of μ has Hausdorff dimension at least γ.]

Let $\delta > 0$ be chosen sufficiently small and let $E \subset [1,2] \times [1,2]$ be a δ-separated set satisfying

$$|E| = \delta^{-\alpha}$$

and

$$|E \cap B(x,r)| \leq r^{\alpha'} |E|$$

for every disc $B(x,r)$ of radius r with $\delta < r < \delta^{\varepsilon_0}$.

Then there exists $\theta \in \operatorname{supp}\mu$ such that

$$N(\pi_\theta(E), \delta) \geq \delta^{-(\alpha/2+\varepsilon)}.$$

Theorem 7 is proved from Theorem 6 using the Balog–Szemerédi–Gowers theorem (and sumset inequalities). This proof may be described (slightly incorrectly) in the following way. Assume there is no such θ. Then take two directions in the support of μ; these guarantee two small projections. That is, E (after applying an affine map and modulo moving points by distances at most δ) is contained in a product $B \times B$ where B is a δ-separated set and $|E| \approx |B|^2$. In this setting, the Balog–Szemerédi–Gowers theorem can be effectively applied to $E \subset B \times B$ to yield a product set $A \times A$. It can be shown that $A \times A$ can be covered by a few translates of E and since E has small projections, so does $A \times A$. However, this means that A contradicts Theorem 6.

Theorem 5 (Bourgain's projection theorem) follows from Theorem 7 using standard techniques and some ideas that are perhaps less standard and not explained in [1]. See [3] for details.

References

[1] J. Bourgain. The discretized sum-product and projection theorems. *J. Anal. Math.* 112 (2010), 193–236.

[2] K.J. Falconer. Hausdorff dimension and the exceptional set of projections. *Mathematika* 29 (1982), no. 1, 109–115.

[3] W. He. *Sums, products and projections of discretized sets.* PhD thesis (2017), Université Paris-Sud, Orsay.

[4] R. Kaufman, P. Mattila. Hausdorff dimension and exceptional sets of linear transformations, *Ann. Acad. Sci. Fenn. A Math.* 1 (1975), 387–392.

[5] J.M. Marstrand. Some fundamental geometrical properties of plane sets of fractional dimensions. *Proc. London Math. Soc.* (3) 4 (1954), 257–302.

Shmerkin's theorem on L^p dimensions, and applications. Part I

Alex Iosevich

The purpose of this talk is to present the first half of Pablo Shmerkin's paper, entitled "On the Furstenberg intersection conjecture, self-similar measures and the L^q norms of convolutions". This paper combines analytic, geometric and combinatorial methods to prove several open conjectures involving intersection of sets, Bernoulli convolutions and others. More precisely, the author studies measures possessing a self-similar structure, which he calls dynamically driven self-similar measures, and contain some classical self-similar measures such as Bernoulli convolutions as special cases. The main result of the paper gives an expression for the L^q dimensions of such dynamically driven self-similar measures. As an application, the the celebrated Furstenberg intersection conjecture established. It says that If A,B are closed subsets of the circle $[0,1)$, invariant under T_p,T_q respectively, with p and q multiplicatively independent, then

$$dim_{\mathcal{H}}(A \cap B) \leq \max\{dim_{\mathcal{H}}(A) + dim_{\mathcal{H}}(B) - 1, 0\},$$

where

$$T_p(x) = px \ mod\ 1.$$

In this talk we shall present the notion of the L^q-dimensional and explain how it be used to study the Furstenberg intersection conjecture and related problems about the Bernoulli convolutions. The key is to relate the notion of the L^q-dimension to the Frostman exponents governing the size of small balls and the sizes of fibers. Let μ be a Borel probability measure on $[0,1]$. Consider the family of intervals of length 2^{-m} $\{j2^{-m},(j+1)2^{-m}\}$, $j \in \mathbb{Z}$, denoted by $\mathcal{D}_m$. The L^q dimension of the measure μ is defined to be

$$\liminf_{m \to \infty} -\frac{\log \sum_{I \in \mathcal{D}_m} \mu^q(I)}{m(q-1)}.$$

Once can check that, roughly speaking, that under the assumption that under the assumption that the L^q dimension is s, it follows that $\mu(B(x,r)) \leq Cr^{(1-1/q)s}$, thus linking the notion of the L^q-dimension and the usual Hausdorff dimension. One can also check that under the same assumption, the upper box dimension of the fibers under Lipschitz maps cannot exceed $s-\alpha$, where α is the Frostman exponent of μ under the said Lipschitz map. These elegant observation lie at the core of the intricate web spun by the author which allows him to knock off the intersection conjecture mentioned above along with several related results.

We shall describe the set of dynamically self-driven models and their L^q-dimension, building up to the main result of the Shmerkin paper (Theorem 11.1) from which everything else is derived. These measures are modeled after p-Cantor sets, that is sets whose base p expansion digits lie in a given subset of $\{0,1,\dots,p-1$, p prime. In the process we shall explain how the basic ideas described above are combined with the notion of dynamically self-driven measures to set up the main mechanism of the paper.

References

[1] P. Shmerkin, *On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions*, Preprint 2016.

L^q-spectrum of homogeneous self-similar measures and inverse theorem for the decay of L^q-norms under convolutions. After P. Shmerkin

Julien Barral

Let $\lambda \in (0,1)$, $b \in \mathbb{N}_{\geq 2}$, $\mathcal{A} = \{0, \dots, b-1\}$, $p = (p_i)_{i \in \mathcal{A}}$ a probability vector, and $(t_i)_{i \in \mathcal{A}} \in \mathbb{R}^b$. For $a \in \mathbb{R}_+^*$, define $S_a : x \in \mathbb{R} \mapsto ax$. For $i \in \mathcal{A}$ define $\varphi_i : x \in \mathbb{R} \mapsto S_\lambda(x) + t_i$. For $n \in \mathbb{N}$ and $I = i_1 \cdots i_n \in \mathcal{A}^n$ define $\varphi_I = \varphi_{i_1} \circ \cdots \circ \varphi_{i_n}$ and $p_I = p_{i_1} \cdots p_{i_n}$. Also define

$$\mu_n = *_{i=0}^{n-1} S_{\lambda^i *}\Big(\sum_{j \in \mathcal{A}} p_j \delta_{t_j}\Big) = \sum_{I \in \mathcal{A}^n} p_I \delta_{\varphi_I(0)}. \tag{1}$$

For $m \in \mathbb{N}$, define $\mathcal{D}_m = \{[k2^{-m}, (k+1)2^{-m}) : k \in \mathbb{Z}\}$.

Denote by μ the unique homogeneous self-similar probability measure associated with p and the IFS $\{\varphi_i : i \in \mathcal{A}\}$, i.e. the unique Borel probability measure on $\mathbb{R}$ such that

$$\mu = \sum_{i \in \mathcal{A}} p_i \mu \circ \varphi_i^{-1}.$$

Recall that μ is supported by the unique non-empty compact set $K \subset \mathbb{R}$ such that $K = \bigcup_{i \in \mathcal{A}} \varphi_i(K)$. Without loss of generality we assume that $K \subset [0,1]$. The measure μ can also be written in the following form:

$$\mu = *_{i=0}^{\infty} S_{\lambda^i *}\Big(\sum_{j \in \mathcal{A}} p_j \delta_{t_j}\Big) = \mu_n * (S_{\lambda^n *}\mu). \tag{2}$$

Define the L^q-spectrum of μ as the concave mapping

$$\tau_\mu : q \in \mathbb{R}_+ \mapsto \lim_{m \to \infty} -\frac{1}{m} \log_2 \sum_{I \in \mathcal{D}_m} \mu(I)^q \tag{3}$$

(that the limit does exist was proved in [4] and follows from a submultiplicative property). The convolution structure (2) can be exploited to prove the following deep theorem.

Theorem 1 (P. Shmerkin [5]). *Either $\Delta_n = \min\{|\varphi_I(0) - \varphi_J(0)| : I \neq J \in \mathcal{A}^n\}$ converges super-exponentially to 0 as $n \to \infty$, or*

$$\forall\, q \geq 1,\ \tau_\mu(q) = \tau(q) := \min\left(q - 1, \frac{\log \sum_{j \in \mathcal{A}} p_j^q}{\log \lambda}\right).$$

Remark 2. *(1) Theorem 1 is a special case of the main result of* [5] *on dynamically driven self-similar measures.*

(2) Under the open set condition, Theorem 1 is known and it holds in any dimension [3].

(3) The measure μ is exact dimensional [1]*, hence its entropy dimension $H(\mu)$ is well defined and equals* $\dim(\mu)$*. On the other hand, it is easily seen that $H(\mu) \leq \tau'(1+)$ always hold, and it is always true that* $\dim(\mu) \geq \tau_\mu'(1+)$*. Consequently, Theorem 1 implies M. Hochman's result on the dimension of homogeneous self-similar measures* [2]*.*

For $n \in \mathbb{N}$, let $\pi_n : x \in \mathcal{A}^{\mathbb{N}} \mapsto \varphi_{x_1 \cdots x_n}(0)$; then define $\pi = \lim_{n\to\infty} \pi_n$. Denoting by ρ the Bernoulli product measure $\left(\sum_{i\in\mathcal{A}} p_i \delta_i\right)^{\otimes \mathbb{N}}$, we have $\mu = \pi_* \rho$, $\mu_n = \pi_{n*}\rho$, and $\|\pi - \pi_n\|_\infty = O(\lambda^n)$, from which it follows that for all $q > 1$, there exists $C_q > 0$ such that

$$C_q^{-1} \sum_{I \in \mathcal{D}_{m(n)}} \mu(I)^q \leq \sum_{I \in \mathcal{D}_{m(n)}} \mu_n(I)^q \leq C_q \sum_{I \in \mathcal{D}_{m(n)}} \mu(I)^q, \quad \forall n \in \mathbb{N}, \tag{4}$$

where $2^{-m(n)} \leq \lambda^{-n} < 2^{-m(n)+1}$. Thus, one can focus on $S_n(q) = \sum_{I\in\mathcal{D}_{m(n)}} \mu_n(I)^q$ to get Theorem 1. In particular, since it is easy to see using the subadditivity of $x \mapsto x^q$ that $-\frac{1}{m(n)} \log_2 S_n(q) \leq \frac{\log \sum_{j\in\mathcal{A}} p_j^q}{\log \lambda}$ for all $q > 1$, while on the other hand the upper bound $q-1$ holds for any probability measure, we get $\tau_\mu(q) \leq \tau(q)$. When Δ_n does not converge super-exponentially to 0, the opposite inequality follows from the following remarkable fact.

Theorem 3 ([5])**.** *Let $q > 1$. Suppose that $\tau_\mu(q) < q-1$ and $\tau_\mu'(q)$ exists. Then, for all $R \in \mathbb{N}$, $\lim_{n\to\infty} \sum_{I\in\mathcal{D}_{Rm(n)}} \mu_n(I)^q = \tau_\mu(q)$.*

Suppose that Δ_n does not converge super-exponentially to 0. Then, there exists $R \in \mathbb{N}$ such that for infinitely many n, for all $q > 1$, $\sum_{I\in\mathcal{D}_{Rm(n)}} \mu_n(I)^q = \sum_{J\in\mathcal{A}^n} p_J^q = \left(\sum_{i\in\mathcal{A}} p_i^q\right)^n$. Consequently, due to Theorem 3 and (4), for any q of the dense subset of $(1,\infty)$ over which τ_μ is differentiable, the equality $\tau_\mu(q) = \tau(q)$ holds, and it extends to $[1,\infty)$ by continuity.

Theorem 3 is a consequence of (2), (4), and a flattening theorem for L^q-norms of discretized version of μ. Before stating this result we need some new definitions.

If $m \in \mathbb{N}$, a 2^{-m}-measure is a probability measure supported on $2^{-m}\mathbb{Z} \cap [0,1]$. For any compactly supported Radon measure ν on $\mathbb{R}$, define $\nu^{(m)} = \sum_{k\in\mathbb{Z}} \nu([k2^{-m},(k+1)2^{-m})\delta_{k2^{-m}}$. If $q \geq 1$, the L^q norm of any finitely supported measure ρ is defined by $\|\rho\|_q^q = \sum_{y\in\mathrm{supp}(\rho)} \rho(\{y\})^q$. The Young inequality $\|\rho * \nu^{(m)}\|_q \leq \|\rho\|_1 \|\nu^{(m)}\|_q$ holds.

Remind that due to (3): $\forall\, \epsilon' > 0$, for m large enough, $\|\mu^{(m)}\|_q^q \geq 2^{-(\tau_\mu(q)+\epsilon')m}$.

Theorem 4 (Flattening property for L^q-norms of $\mu^{(m)}$ under convolution [5])**.** *Let $\sigma > 0$ and $q > 1$ such that $\tau'(q)$ exists and $\tau(q) < q-1$. Then, there exists $\epsilon = \epsilon(\sigma,q) > 0$ such that for m large enough, if ν is a 2^{-m}-measure and $\|\nu\|_q^q \leq 2^{-\sigma(q-1)m}$, then $\|\nu * \mu^{(m)}\|_q^q \leq 2^{-(\tau_\mu(q)+\epsilon)m} \leq 2^{-\epsilon m/2} \|\mu^{(m)}\|_q^q$.*

Theorem 4 is proved by contradiction. The proof combines fine large deviations estimates associated with μ at "temperature" $1/q$ when $q > 1$, $\tau_\mu'(q)$ exists, and

$q\tau'_\mu(q) - \tau_\mu(q) < 1$ (which holds when $\tau_\mu(q) < q - 1$), and an inverse theorem for the decay of L^q-norms under convolutions (Theorem 5) that we state after introducing new definitions.

If $A \subset \mathbb{R}$ and $s \in \mathbb{N}$, define $\mathcal{D}_s(A) = \{I \in \mathcal{D}_s : I \cap A \neq \emptyset\}$ and $\mathcal{N}_s(A) = |\mathcal{D}_s(A)|$.

If $x \in \mathbb{R}$, $\mathcal{D}_s(x)$ stands for the unique $I \in \mathcal{D}_s$ such that $x \in I$. If $a > 0$ and I is an interval, aI stands for the interval with the same center as I and length $a|I|$.

If $m \in \mathbb{N}$, a 2^{-m}-set is a subset of $2^{-m}\mathbb{Z} \cap [0, 1]$.

Let $D, \ell \in \mathbb{N}$ and set $m = D\ell$. Given $R = (R_s)_{0 \leq s \leq \ell - 1} \in [1, 2^D]^\ell$, say that a 2^{-m}-set A is (D, ℓ, R)-uniform if $\mathcal{N}_{(s+1)D}(A \cap I) = R_s$ for all $0 \leq s \leq \ell - 1$ and $I \in \mathcal{D}_{sD}(A)$, i.e. A has the structure of a spherical tree of height ℓ, with branching number R_s at generation sD, $0 \leq s \leq \ell - 1$.

The following result gives a very precise structural description of two 2^{-m}-measures ρ and ν whenever $\|\rho * \nu\|_q^q \geq 2^{-m\epsilon q}\|\rho\|_q^q$. In particular, it tells that along an arithmetic sequence of scales, either a large proportion of ν looks atomic, or ρ is distributed rather uniformly when restricted to a subset which carries a large proportion of the L^q norm of ρ. This has a flavour similar to that of the inverse theorem for the growth of entropy established in [2].

Theorem 5 ([5])**.** *Let $q > 1$, $\delta > 0$ and $D_0 \in \mathbb{N}$. There are $\epsilon > 0$ and $D \geq D_0$, such that if ℓ is large enough, $m = \ell D$ and ρ and ν are 2^{-m}-measures such that $\|\rho * \nu\|_q \geq 2^{-m\epsilon}\|\rho\|_q$, the following holds: after translating ρ and ν by appropriate numbers of the form $k2^{-m}$, there exist $A \subset \operatorname{supp}(\rho)$ and $B \subset \operatorname{supp}(\nu)$ such that*

(i) $\|\rho_{|A}\|_q \geq 2^{-m\delta}\|\rho\|_q$ *and* $\nu(B) \geq 2^{-m\delta}$*;*

(ii) $\rho(\{y\}) \leq 2\rho(\{x\})$ *for all* $x, y \in A$ *and* $\nu(\{y\}) \leq 2\nu(\{x\})$ *for all* $x, y \in B$*;*

(iii) $x \in \frac{1}{2}\mathcal{D}_{sD}(x)$ *for all* $x \in A \cup B$ *and* $0 \leq s \leq \ell - 1$*;*

(iv) There exists R^A and R^B such that A and B are respectively (D, ℓ, R^A) and (D, ℓ, R^B)-uniform.

(v) For all $0 \leq s \leq \ell - 1$, either $R^B_s = 1$ or $R^A_s \geq 2^{(1-\delta)}$.

(vi) $-\dfrac{\log(\|\nu\|_q^q)}{q-1} - \delta m \leq D \cdot |\{0 \leq s \leq \ell - 1 : R^A_s \geq 2^{(1-\delta)}\}| \leq -\dfrac{\log(\|\rho\|_q^q)}{q-1} + \delta m.$

To prove Theorem 5, P. Shmerkin first exploits the inequality $\|\rho * \nu\|_q \geq 2^{-m\epsilon}\|\rho\|_q$ to construct sets A_1 and B_1 such that (i) and (ii) hold, and the additive energy of A_1 and A_2 fulfills the assumption of a slightly simplified version of the asymmetric Balog-Szemeredi-Gowers theorem. He also establishes a beautiful refinement of Bourgain's structural result for small doubling sets. This result is then applied to the small doubling set produced by asymmetric B-S-G theorem to get the desired sets A and B, after a series of delicate manipulations.

References

[1] D.-J. Feng, H. Hu, *Dimension theory of iterated function systems*, Commun. Pure Appl. Math. **62** (2009), 1435–1500.

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

[3] K.-S. Lau, S.-M. Ngai, *Multifractal measures and a weak separation condition*, Adv. Math. **141** (1999), 45–96.

[4] Y. Peres, B. Solomyak, *Existence of L^q dimensions and entropy for self-imilar measure*, Indiana Univ. Math. J. **49** (2000), 1603–1621.
[5] P. Shmerkin, *On Furstenberg's intersection conjecture, self-similar measures, and the L^q norms of convolutions*, arXiv:1609.07802.

Reporter: Yuki Takahashi