Spacings and pair correlations for finite Bernoulli convolutions

Boris Solomyak (joint work with Itai Benjamini)

Strobl, Fractals and Tilings, July 6-10, 2009

Nearest-neighbor spacings

Suppose a sequence $\{\theta_n\}$ is equidistributed in $[0, 1)$, that is, for every $[a, b] \subset [0, 1)$,

$$\lim_{N\to\infty} N^{-1}\#\{n \le N :\ \theta_n \in [a, b]\} = b - a.$$

We order the first N elements of the sequence:

$$\theta_{1,N} \le \theta_{2,N} \le \cdots \le \theta_{N,N},$$

and then consider the normalized spacings

$$\delta_n^{(N)} := N(\theta_{n+1,N} - \theta_{n,N}).$$

More generally, for any fixed $\ell \geq 1$,

$$\delta_{\ell,n}^{(N)} := N(\theta_{n+\ell,N} - \theta_{n,N}).$$

A function $P_\ell(s)$ is called the **limiting distribution function** of $\{\delta_{\ell,n}^{(N)}\}$ as $N \to \infty$ if, for any interval $[a, b] \subset (0, \infty)$,

$$\lim_{N\to\infty} \frac{1}{N} \# \left\{ n \leq N : \ \delta_{n,\ell}^{(N)} \in [a, b] \right\} = \int_a^b P_\ell(s)\, ds.$$

If θ_n are i.i.d. uniform on $[0, 1]$, we get the **Poisson model**, with

$$P_\ell(s) = \frac{s^{(\ell-1)}}{(\ell-1)!} e^{-s}, \quad \text{for all } \ell \geq 1.$$

Correlation functions

The pair correlation function for the sequence $\{\theta_n\}$ is defined by

$$R_2(s, \{\theta_n\}, N) = \frac{1}{N} \# \Big\{ (n, m) : \; n \neq m, \; n, m \leq N, \; |\theta_{n,N} - \theta_{m,N}| \leq s \Big\}.$$

In the Poisson model we have $R_2(s, \{\theta_n\}, N) \to 2s$, as $N \to \infty$. Similarly are defined higher level correlations. It is well-known that all local spacing measures are determined by the correlation functions (of all levels).

Spacings of αn^d mod 1

- The spacing distributions of αn^d mod 1, for $\alpha \notin \mathbb{Q}$ and $d \geq 2$, have been studied in a number of papers (Rudnick, Sarnak, Zaharescu, Boca) since 1990's, motivated in part by Quantum Chaos.
- It is conjectured that for α badly approximable by rationals the spacing distribution is Poissonian.
- For a.e. α, the pair correlations converge to $2s$ as $N \to \infty$.
- For some irrational α (well-approximable by certain rationals) the behavior of spacings is not Poissonian.

For $\lambda \in (\frac{1}{2}, 1)$ let

$$A_N(\lambda) = \left\{ (1-\lambda) \sum_{n=0}^{N-1} a_n \lambda^n : \; a_n \in \{0,1\} \right\}.$$

We are interested in the randomness properties of these sets as $N \to \infty$. Note that a point in $A_N(\lambda)$ may have more than one representation. This will happen whenever λ is a zero of a polynomial with coefficients $\{-1, 0, 1\}$. In this case we take the point "with multiplicity," so that our sequence always has size 2^N. Consider the uniform measure ν_λ^N on $A_N(\lambda)$; this is a discrete measure, with a point in $A_N(\lambda)$ having mass 2^{-N} times the number of representations.

Bernoulli convolutions

The measures ν_λ^N converge vaguely to the **infinite Bernoulli convolution measure** ν_λ, defined as the distribution of the random series $\sum_{n=0}^\infty \omega_n(1-\lambda)\lambda^n$, where ω_n are i.i.d. Bernoulli random variables taking the values $0, 1$ with equal probability.

- ν_λ is continuous and has "pure type" for all λ
- ν_λ is singular for all $\lambda < \frac{1}{2}$ and a.c. for a.e. $\lambda \in (\frac{1}{2}, 1)$
- PV reciprocal λ: the only known singular ν_λ for $\lambda \in (\frac{1}{2}, 1)$.
- Garsia numbers: the only explicitly known λ with a.c. ν_λ.

Rescaling the sequence

Let $\{\xi_{j,N}\}_{j=1}^{2^N}$ be the ordering of $A_N(\lambda)$, counting with multiplicity, so that

$$0 = \xi_{1,N} \le \xi_{2,N} \le \cdots \le \xi_{2^N,N} = 1 - \lambda^N.$$

Spacing statistics provide a measure of randomness for the sequences $\{\xi_{n,N}\}$ approximating ν_λ. Numerical evidence seems to indicate that, for a typical λ, there exist level spacing distributions. However, we first need to rescale the sequence. Let

$$\widetilde{\xi}_{n,N} = F_\lambda(\xi_{n,N}), \quad n = 1, \ldots, 2^N, \quad \text{where } F_\lambda(\xi) = \nu_\lambda(-\infty, \xi).$$

These sequences are uniformly distributed modulo one.

Conjecture 1. *For almost every $\lambda \in (\frac{1}{2}, 1)$ the rescaled sequences $\widetilde{\xi}_{n,N}$ are distributed according to the Poisson model.*

- A key difficulty is that very few measures ν_λ are known explicitly.
- We have numerical evidence in support of this conjecture. In simulations, we used for rescaling the CDF of one of the few explicitly known cases: $\lambda = 2^{-1/2} \approx 0.7017$, for the whole range of parameters λ. This doesn't affect the histograms too much, mostly just by "smoothing" them a little.

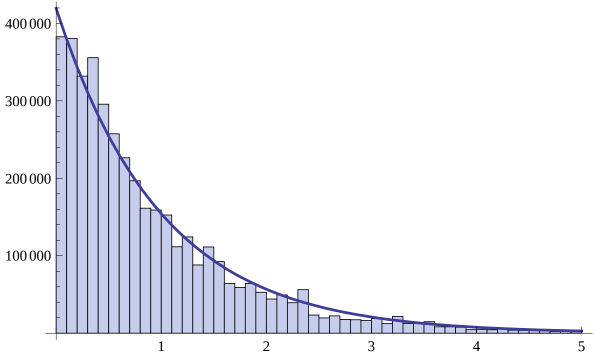

Figure: non-rescaled, $\lambda = 0.6429$, $\ell = 1$, $N = 22$

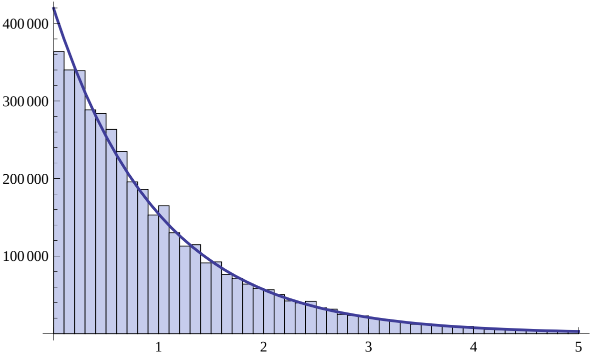

Figure: rescaled, $\lambda = 0.6429$, $\ell = 1$, $N = 22$

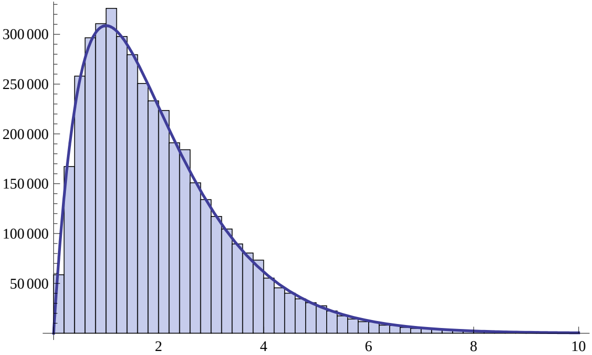

Figure: rescaled, $\lambda = 0.6429$, $\ell = 2$, $N = 22$

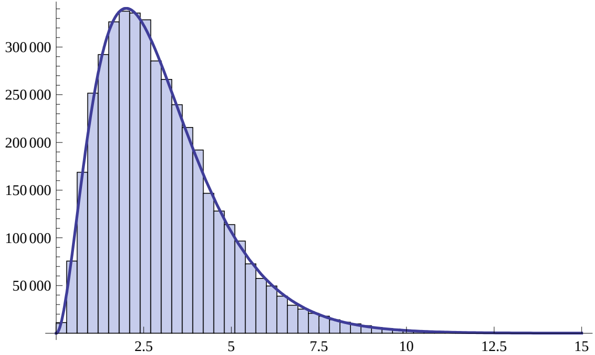

Figure: rescaled, $\lambda = 0.6429$, $\ell = 3$, $N = 22$

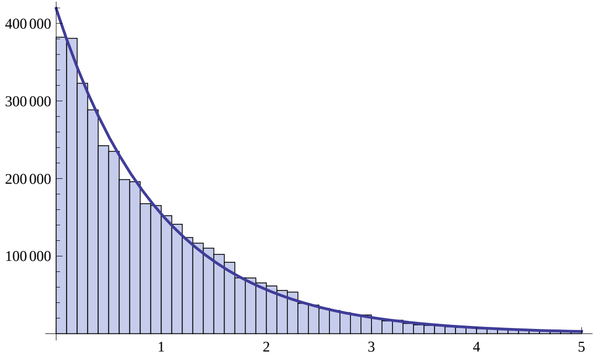

Figure: rescaled, $\lambda = 0.7088$, $\ell = 1$, $N = 22$

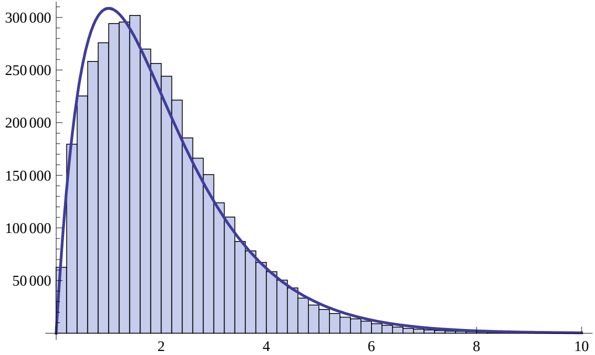

Figure: rescaled, $\lambda = 0.7088$, $\ell = 2$, $N = 22$

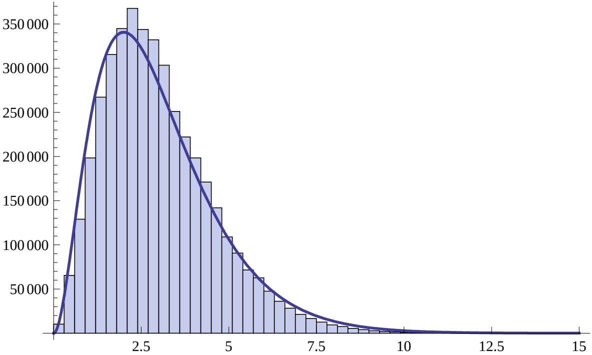

Figure: rescaled, $\lambda = 0.7088$, $\ell = 3$, $N = 22$

We mostly focus on the pair correlation function for the non-rescaled sequences

$$R_2(s,\lambda,2^N) := \frac{1}{2^N}\#\Big\{(x,y):\ x\ne y,\ x,y\in A_N(\lambda),\ |x-y|\le \frac{s}{2^N}\Big\}$$

which provide information about the variance of average spacings.

Conjecture 2. *For almost every $\lambda\in(\frac12,1)$, there exist $c,C>0$ such that*

$$cs \le R_2(s,\lambda,2^N)\le Cs \ \ \textit{for all } N \textit{ and } s>0.$$

This would mean that our sequences exhibit **no attraction** and **no repulsion**. Our main results make a step in this direction, but we have to do averaging over the parameter λ.

Results I

Theorem

Let $J = [\lambda_0, \lambda_1] \subset (\frac{1}{2}, 0.668)$. *Then* $\exists\, C_1 = C_1(J)$ *such that*

$$\forall\ N \geq 1,\ \forall\ s > 0,\ \int_J R_2(s, \lambda, 2^N)\, d\lambda \leq C_1 s,$$

Remarks.

- The appearance of the number 0.668 is due to "transversality".
- By Fatou's Lemma, $\int_J \liminf_{N\to\infty} R_2(s, \lambda, 2^N)\, d\lambda \leq C_1 s$ for $s > 0$. This shows that, on average, our spacing distributions exhibit no attraction, at least in the liminf sense.
- By Borel-Cantelli, for any $\varepsilon > 0$, for a.e. $\lambda \in (\frac{1}{2}, 0.668)$,

$$\forall\, s > 0,\ R_2(s, \lambda, 2^N) \leq sN^{1+\varepsilon} \quad \text{for all } N \text{ sufficiently large.}$$

Results II

Theorem

For any $\lambda_0 \in (\frac{1}{2}, 1)$, $\varepsilon \in (0, 1 - \lambda_0)$, *and* $L > 0$, *there exists* $C_2 = C_2(\lambda_0, \varepsilon) > 0$ *such that*

$$\forall\ N \geq 1,\ \forall\ s \in (0, L), \quad \int_{\lambda_0 - \varepsilon}^{\lambda_0 + \varepsilon} R_2(s, \lambda, 2^N)\, d\lambda \geq C_2 s.$$

Non-Poissonian behavior I

Attraction. Let $\theta = 1/\lambda$ be a PV number, that is, an algebraic integer > 1 whose conjugates are all less than one in modulus.

- Such ν_λ, with $\lambda \in (\frac{1}{2}, 1)$, are the only singular infinite Bernoulli convolutions known [Erdős 1939].
- Spacings between distinct points in $A_N(\lambda)$ are all bounded below by $C \cdot \lambda^N$, which implies that there are "massive" coincidences, since $2^N - C \cdot \lambda^{-N}$ spacings are zeros.
- This is an extreme case of attraction. Something similar happens for all zeros of $\{0, \pm 1\}$ polynomials.

Results III

Theorem

(i) *For any* λ_0 *which is a zero of a* $\{0, \pm 1\}$ *polynomial, there exists* $\rho = \rho(\lambda_0) < 2$ *and* C_3 *such that*

$$R_2(0, \lambda_0, 2^N) \geq C_3^{-1}(2/\rho)^N.$$

(ii) *For any interval* $J \subset (\frac{1}{2}, 1)$ *and* $\varepsilon > 0$*, there is an uncountable set* $\mathcal{E}_J \subset J$*, such that for* $\lambda \in \mathcal{E}_J$

$$\forall\, s > 0,\ R_2(s, \lambda, 2^N) \geq 2^{N^{1-\varepsilon}} \quad \textit{for infinitely many } N.$$

Remarks.

- The points in $\mathcal{E}_J$ are extremely well approximable by zeros of $\{0, \pm 1\}$ polynomials.
- Our construction yields $\mathcal{E}_J$ of zero Hausdorff dimension.

Non-Poissonian behavior II

Repulsion. Let $\theta = 1/\lambda$ be a **Garsia** number, that is, a zero of a monic polynomial with integer coefficients and constant term ± 2, whose conjugates are all greater than one in modulus.

- By Garsia's Lemma, the spacings between points in $A_N(\lambda)$ are bounded below by $C \cdot 2^{-N}$, and there are no coincidences. This is an extreme case of repulsion.
- Such ν_λ, with $\lambda \in (\frac{1}{2}, 1)$, are the only **explicitly known a.c.** known [Garsia 1960], and moreover, they have bounded density.
- Garsia reciprocals include $2^{-1/k}$ for $k \geq 2$, but also many other numbers (e.g. the reciprocal of the positive root of $x^3 - 2x - 2 = 0$, which is about .5652, or the reciprocal of the positive root of $x^3 - x^2 - 2 = 0$, which is about .5898).

Additional Remarks

- Some spacings have very high multiplicity (producing "spikes" in the distribution of non-rescaled spacings). This may be explained by the fact that every spacing g is a zero of a $\{0, \pm 1\}$ polynomial. For every such polynomial with k zero coefficients, the same difference between points in $A_N(\lambda)$ will be achieved 2^k times. If g is a very small number, it is likely to be a spacing, and if $N - k$ is small, we will see a "spike."
- The smallest spacing between points of $A_N(\lambda)$, for a.e. λ in a transversality interval, is bounded below by $3^{-N}N^{-1-\varepsilon}$, for infinitely many N, for any $\varepsilon > 0$, which is much larger than $(2^{-N})^2$ which we would get if the points were sampled randomly and independently from the uniform distribution.