How random is a uniformly distributed sequence?

by Niclas Technau

April 30, 2020

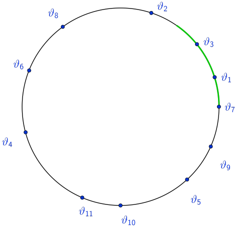
TEL AVIV UNIVERSITY אוניברסיטת תל אביב

Uniform Distribution mod one

$\vartheta = (\vartheta_n)_{n\geq 1}$ in $[0,1)$is called *uniformly distributed mod* 1 if

$$\frac{\#\{1 \leq n \leq N : \vartheta_n \in I\}}{N} \underset{N\to\infty}{\longrightarrow} \mathrm{meas}(I)$$

for each sub-interval $I \subseteq [0,1]$.



Sketch of a uniformly distributed sequence in $[0,1) \cong \mathbb{R}/\mathbb{Z}$.

Abbreviation: ϑ u.d.

Examples I

Lemma 1 (Weyl's Criterion [3]).

A sequence sϑ is u.d. if and only if

$$\frac{1}{N}\sum_{n\leq N} e^{2\pi i h \vartheta_n} \underset{N\to\infty}{\longrightarrow} 0$$

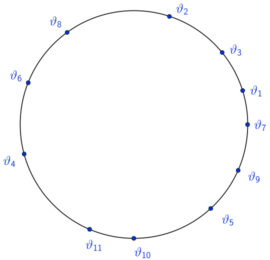
for each $h \in \mathbb{Z}\setminus\{0\}$.

Let $\alpha \in (0,\infty)$.

1. The Kronecker sequence $(\langle \alpha n\rangle)_{n\geq 1}$ is u.d. iff $\alpha \notin \mathbb{Q}$; Bohl, Sierpinski, Weyl independently (1909–1910).
2. More generally, $(\langle \alpha n^d\rangle)_{n\geq 1}$, $d\in\mathbb{N}$, is u.d. iff $\alpha \notin \mathbb{Q}$; Weyl (1916).
3. The sequence $(\langle n^\alpha\rangle)_{n\geq 1}$ is u.d. iff $\alpha\notin\mathbb{N}$; Fejér and Csillag (≈1930).
4. If $\vartheta_n : [0,1) \to [0,1)$ are $[0,1]$-uniformly distributed, independent random variables, then $(\vartheta_n(\alpha))_{n\geq 1}$ is almost surely u.d. .

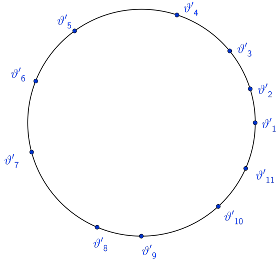
Gap Distributions I

(a) Let ϑ_n' denote the n^{th} (largest) element of $\{\vartheta_n : n \leq N\}$. Note that the average gap of ϑ_n' is $1/N$.



Unordered points on the unit circle.

Gap Distributions II



Ordered points on the unit circle.

(b) Put $\delta_{n,N} = N(\vartheta'_{n+1} - \vartheta'_n)$, where $n = 1, \ldots, N$ (and $\vartheta'_{n+1} := \vartheta_n + 1$).

Gap Distributions III

Definition 1.

If P exists such that

$$\frac{\#\{n \leq N : \delta_{n,N} \in I\}}{N} \underset{N \to \infty}{\longrightarrow} \int_I P(s)\, \mathrm{d}s$$

for each interval $I \subseteq \mathbb{R}_{\geq 0}$, then P is called the *gap distribution* of ϑ. If $P(s) = e^{-s}$, then ϑ has *Poissonian gap distribution*.

Remark.

There are other interesting statistics. E.g. the size of the minimal gap $\min_{n \leq N} \delta_{n,N}$, as $N \to \infty$.

The Berry–Tabor conjecture

Conjecture 1 (Berry and Tabor [4]).

A "generic" quantum system whose underlying classical dynamics is completely integrable, the gaps between the energy levels have a Poissonian limiting distribution.

Remark.

The energy levels of a boxed 2-dimensional oscillator are decoded by $\big(\langle \alpha n^2 \rangle\big)_{n\geq 1}$. For further reading: Rudnick [7].

Examples II

1. If $\vartheta_n : [0,1) \to [0,1)$ are $[0,1]$-uniformly distributed, independent random variables, then $(\vartheta_n(\alpha))_{n\geq 1}$ has a.s. Poissonian gap distribution.
2. Quadratic residues mod q, as the number of prime factors $\omega(q) \to \infty$, have a Poissonian gap distribution; Rudnick and Kurlberg (1999).
3. (Not-example) $(\langle \sqrt{n} \rangle)_{n\geq 1}$ has non-Poissonian gap distribution; Elkies and McMullen (2002).
4. Almost surely $(\langle \alpha 2^n \rangle)_{n\geq 1}$ has Poissonian gap distribution; Rudnick and Zaharescu (1999); see [5], [10], [6].

k-point Correlation Function I

Definition 2.

For $f \in C_c^\infty(\mathbb{R}^{k-1})$ or f being an indicator function of a parallelepiped, put

$$F_M(\mathbf{t}) := \sum_{\mathbf{u}\in\mathbb{Z}^{k-1}} f\left(M(\mathbf{t}+\mathbf{u})\right).$$

If, for any such f, the *k-point correlation function*

$$R_k(\vartheta, N, f) := \frac{1}{N} \sum_{\substack{n_1,\ldots,n_k \le N \\ n_i \text{distinct}}} F_N\left(\vartheta_{n_1} - \vartheta_{n_2}, \vartheta_{n_2} - \vartheta_{n_3}, \ldots \vartheta_{n_{k-1}} - \vartheta_{n_k}\right)$$

converges to $\int_{\mathbb{R}^{k-1}} f(\mathbf{t})\,\mathrm{d}\mathbf{t}$, as $N \to \infty$, then ϑ has *Poissonian k-point correlations*.

k-point Correlation Function II

Remark.

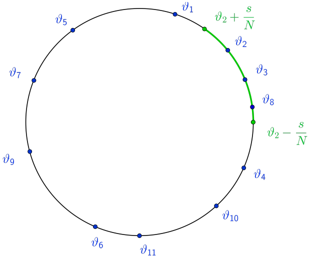
(a) If $k = 2$ (resp. $k = 3$) we speak of the pair (resp. triple) correlation function.
(b) If $\vartheta(\alpha)$ is a function, we simply set $R_k(\alpha, N, f) = R_k(\vartheta(\alpha), N, f)$.
(c) Poissonian pair correlation imply uniform distribution. The converse fails: $(\langle \alpha n \rangle)_{n \geq 1}$ has for *no* $\alpha \in (0, 1)$ the Poissonian pair correlation.

Pair Correlation Function

Let $k = 2$, $f = 1_{[0,s]}$. Then $F_N(t) = 1_{[0,s]}(N\|t\|)$ which equals $1_{[0,s/N]}(\|t\|)$. So,

$$R_2(\vartheta, N, f) = \frac{1}{N} \sum_{n \neq m \leq N} 1_{[0,s/N]}(\|\vartheta_n - \vartheta_m\|).$$

What does it mean that this function is $\stackrel{!}{\approx} \int_{\mathbb{R}} 1_{[0,s]}(\|t\|)\,\mathrm{d}t = 2s$?



Heuristic for the pair correlation function. Each green neighborhood contains $\approx 2s$ points on average.

Metric Theory of the Pair Correlation Function I

Theorem 1 (Marklof, Vinogradov, El-Baz [2]).

The sequence $\left(\langle\sqrt{n}\rangle\right)_{n\geq 1}$ *has Poissonian pair correlation.*

Problem 1.

Has $\left(\langle\sqrt{2}n^2\rangle\right)_{n\geq 1}$ *Poissonian pair correlation?*

Definition 3.

A strictly increasing sequence $\left(a_n\right)_{n\geq 1} \subseteq \mathbb{Z}_{\geq 1}$ is said to have the metric Poissonian pair correlation property if for (Lebesgue) almost every $\alpha \in (0,1)$ the sequence $(\langle \alpha a_n \rangle)_{n\geq 1}$ has Poissonian pair correlations.

For $I \subset \mathbb{Z}$, define

$$E(I) := \sum_{\substack{a+b=c+d \\ a,b,c,d \in I}} 1, \quad \text{and observe} \quad (\# I)^2 \leq E(I) \leq (\# I)^3 .$$

Metric Theory of the Pair Correlation Function II

Theorem 2 (T. F. Bloom, A. Walker [8]).

There is a universal constant $C > 0$ such that if the cut-offs $A_N := \{a_n : n \leq N\}$ satisfy

$$E(A_N) \ll \frac{N^3}{(\log N)^C}, \quad (1)$$

then $(a_n)_{n\geq 1}$ has the metric Poissonian pair correlations.

Remark.

(1) is satisfied, e.g., for lacunary sequences, or $\left(n^d\right)_{n\geq 1}$ with $d \in \mathbb{Z}_{\geq 2}$.

Metric Theory of the Pair Correlation Function III

Theorem 3 (A. Walker [9]).

The prime numbers $(p_n)_{n\geq 1}$ does not have metric Poissonian pair correlation.

Remark.

Does the metric Poissonian pair correlation property follows a zero–one law? Suppose, say,

$$\sum_{N\geq 1} \frac{E\left(A_N\right)}{N^3} < \infty, \qquad \frac{E\left(A_N\right)}{N^3} \quad \mathrm{well-behaved},$$

if and only if $(a_n)_{n\geq 1}$ has the metric Poissonian pair correlation property?

Metric Theory for the Pair Correlation Function IV

This fails on logarithmic scales:

Theorem 4 (Aistleitner, Lachmann, N.T. [1]).

There is a metric Poissonian sequence $(a_n)_{n\geq 1}$ such that

$$E\left(A_N\right) \gg \frac{N^3}{(\log N)^{\frac{3}{4}+\varepsilon}}.$$

The Borel–Cantelli lemma

Lemma 2 (Borel–Cantelli).

Suppose Lebesgue measurable sets $\Omega_n \subseteq [0,1)$ *satisfy*

$$\sum_{n\geq 1} \lambda(\Omega_n) < \infty.$$

Then, $\{\alpha \in [0,1) : \alpha \in \Omega_n \,\text{for}\, \infty \,\text{many}\, n\}$ *has measure zero.*

A blue-print for dilated squares I

An L^2-approach: Fix s, and let $R_2\left(\alpha, N, 1_{[-s,s]}\right) = \frac{1}{N}\sum_{n\neq m\leq N} 1_{[0,s/N]}\left(\left\|\alpha(n^2-m^2)\right\|\right)$ denote the pair correlation function of $(\langle \alpha n^2\rangle)_{n\geq 1}$.

1. First show that $R_2\left(\alpha, N_m, 1_{[-s,s]}\right) \underset{m\to\infty}{\longrightarrow} 2s$ a.s. on a (polynomially-thin) sub-sequence $(N_m)_{m\geq 1}$.
2. Then deduce convergence on the full sequence, i.e. $R_2\left(\alpha, N, 1_{[-s,s]}\right) \underset{N\to\infty}{\longrightarrow} 2s$ a.s., via a sandwiching argument.

A blue-print for dilated squares II: Sub-sequence step

- For any integer $d \in \mathbb{Z}$, we have $\int_{[0,1]} 1_{[0,\delta]}(\|d\alpha\|)\,\mathrm{d}\alpha = 2\delta$. So,

$$\mathbb{E}(R) := \int_{[0,1]} R\left(\alpha, N_m, 1_{[0,s]}\right)\,\mathrm{d}\alpha = \frac{1}{N}N(N-1)\frac{2s}{N} = 2s + O_s(N^{-1}),$$

- By Chebycheff,

$$\mathrm{meas}\left(\left\{\alpha \in [0,1) : \left|R\left(\alpha, N_m, 1_{[0,s]}\right) - \mathbb{E}(R)\right| > k\right\}\right) \leq \frac{\mathrm{Var}(R)}{k^2},$$

for each $k > 0$, where

$$\mathrm{Var}(R) := \int_{[0,1]} \left(R\left(\alpha, N_m, 1_{[0,s]}\right) - \mathbb{E}(R)\right)^2\,\mathrm{d}\alpha.$$

- Show that $\mathrm{Var}(R) \ll N^{-1+\varepsilon}$. (More details soon.)
- Apply Lemma 2 with $k = N_m^{-1/4}$ and $N_m = m^2$.

A blue-print for dilated squares III: Variance Estimate

Recall

$$R_2\left(\alpha, N, 1_{[-s,s]}\right) = \frac{1}{N}\sum_{n\neq m\leq N} 1_{[0,s/N]}\left(\left\|\alpha(n^2-m^2)\right\|\right)$$
$$\approx \sum_{n\neq m\leq N}\ \sum_{0\leq|k|\leq N^{1+\varepsilon}} c_k e^{2\pi i k(n^2-m^2)} + \mathrm{Error}.$$

- Indeed,

$$\mathrm{Var}\left(R\right) \ll \frac{1}{N^4}\#\left\{k_1(n_1^2-m_1^2)+k_2(n_2^2-m_2^2)=0:\ n_i, m_i, k_i \text{ as above}\right\}$$
$$\ll \frac{1}{N^4}\sum_d \#\left\{k_2(n_2^2-m_2^2)=d:\ n_2, m_2, k_2 \text{ as above}\right\}$$

where the summation runs over $\ll N^3$ many values of $d \ll N^3$.Using that $\#\{y\in\mathbb{Z}:\ y\mid d\}\ll d^\varepsilon$ produces

$$\mathrm{Var}\left(R\right) \ll N^{-1+\varepsilon}.$$

Main result

Theorem 5 (N. T. & Nadav Yesha).

The k-point correlation function of $(\langle n^\alpha \rangle)_{n\geq 1}$ is Poissonian for almost every $\alpha \gg k^2$. In particular, $(\langle n^\alpha \rangle)_{n\geq 1}$ has Poissonian pair correlations for almost every $\alpha > 7$.

Conjecture 2.

The gap distribution of $(\langle n^\alpha \rangle)_{n\geq 1}$ is Poissonian if $\alpha \in \mathbb{R}_{>0} \setminus \left(\mathbb{N} \cup \left\{\frac{1}{2}\right\}\right)$.

The Rudnick–Sarnak Conjecture

Conjecture (Z. Rudnick and P. Sarnak [11]).

If $\alpha \in (0,1)$ is Diophantine, i.e. $\|n\alpha\| \gg_\varepsilon n^{-(1+\varepsilon)}$ for all $\varepsilon > 0$, and $d \geq 2$ is an integer, then

$$\left(\langle \alpha n^d \rangle\right)_{n \geq 1}$$

has Poissonian gap distribution.

Remark.

For $d = 2$, see also work of Heath-Brown.

The Rudnick–Sarnak Obstruction

Recall $F_M(\mathbf{t}) = \sum_{\mathbf{u}\in\mathbb{Z}^2} f(M(\mathbf{t}+\mathbf{u}))$ for $f \in C_c^\infty(\mathbb{R}^2)$. Put

$$R_3(\alpha, N, L, f) := \frac{1}{N^3} \sum_{\substack{x_1, x_2, x_3 \leq N \\ x_i \text{ distinct}}} F_{N/L}(\alpha(x_1^2 - x_2^2), \alpha(x_2^2 - x_3^2)).$$

Now we expect $R_3(\alpha, N, L, f) \sim L^2 \int_{\mathbb{R}^2} f(\mathbf{t})\,\mathrm{d}\mathbf{t}$.

Proposition 1 (Z. Rudnick and P. Sarnak [11]).

If $\varepsilon > 0$ and $1 \leq L \leq N^{\frac{1}{3}-\varepsilon}$, then there exists $f \in C_c^\infty(\mathbb{R}^2)$ such that

$$\frac{1}{L^2}\int_{[0,1]} (R_3(\alpha, N, L, f))^2\,\mathrm{d}\alpha \underset{N\to\infty}{\longrightarrow} \infty.$$

Why?

The Rudnick–Sarnak Obstruction

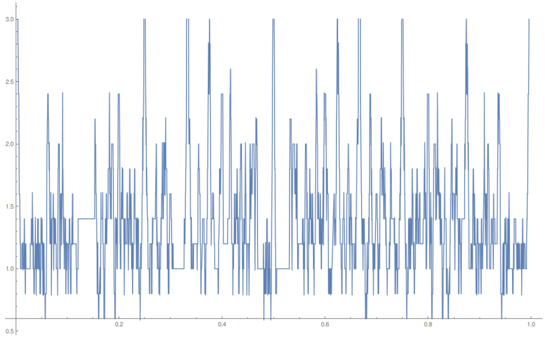


Figure: The pair correlation function plotted with Mathematica, for $N = 10$ and $f = 1_{[0,1]}$. Observe the large peaks at rationals with small denominator; in particular at $\frac{0}{1}$ and $\frac{1}{1}$!

Beyond the Obstruction

Theorem 6 (N.T., Aled Walker).

If $\varepsilon > 0$ and $N^{\frac{1}{4}+\varepsilon} < L < N^{1-\varepsilon}$, then

$$R_3(\alpha, N, L, f) = (1 + o_{\alpha,\varepsilon,f}(1))L^2 \int_{\mathbb{R}^2} f(\mathbf{t}) \, \mathrm{d}\mathbf{t}$$

for almost all $\alpha \in [0,1]$ for each $f \in C_c^\infty(\mathbb{R}^2)$, uniformly in L.

Remark.

Note: $\frac{1}{4} < \frac{1}{3}$. In the regime $N^{\frac{1}{2}+\varepsilon} \leq L \leq N$ one can use metric discrepancy theory.

The End

Thank you very much for your attention!

References

C. Aistleitner, T. Lachmann, N. Technau: *There is no Khintchine threshold for metric pair correlations.*
Mathematika, 65(4):929–949, 2019.

D. El-Baz, J. Marklof, I. Vinogradov: *The two-point correlation function of the fractional parts of $\sqrt{n}$ is Poisson.*
Proceedings of the American Mathematical Society, 143(7):2815–2828, 2015.

L. Kuipers, H. Niederreiter: *Uniform distribution of sequences.*
Courier Corporation, 2012.

M.V. Berry, M. Tabor: *Level clustering in the regular spectrum.*
Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences, 356(1686):375–394, 1977.

N. Elkies, C. McMullen: *Gaps in $\sqrt{n}$ mod 1 and ergodic theory.*
Duke Mathematical Journal, 123(1):95–139, 2004.

P. Kurlberg, Z. Rudnick: *The distribution of spacings between quadratic residues.*
Duke mathematical journal, 100(2):211–242, 1999.

Rudnick, Z.: *Quantum chaos?*
Notices of the AMS, 55(1):32–34, 2008.

T.F. Bloom, A. Walker: *GCD sums and sum-product estimates.*
Israel Journal of Mathematics, 235(1):1–11, 2020.

Walker, A.: *The primes are not metric Poissonian.*
Mathematika, 64(1):230–236, 2018.

Z. Rudnick, A. Zaharescu: *The distribution of spacings between fractional parts of lacunary sequences.*
Forum Mathematicum, 14, 691–712. Walter de Gruyter GmbH & Co. KG, 2002.

Z. Rudnick, P. Sarnak: *The pair correlation function of fractional parts of polynomials.*