The metric theory of the pair correlation function of real-valued lacunary sequences

Zeév Rudnick and Niclas Technau

Abstract Let $\{a(x)\}_{x=1}^{\infty}$ be a positive, real-valued, lacunary sequence. This note shows that the pair correlation function of the fractional parts of the dilations $\alpha a(x)$ is Poissonian for Lebesgue almost every $\alpha \in \mathbb{R}$. By using harmonic analysis, our result—irrespective of the choice of the real-valued sequence $\{a(x)\}_{x=1}^{\infty}$—can essentially be reduced to showing that the number of solutions to the Diophantine inequality

$$\left|n_1\big(a(x_1)-a(y_1)\big)-n_2\big(a(x_2)-a(y_2)\big)\right|<1$$

in integer six-tuples $(n_1,n_2,x_1,x_2,y_1,y_2)$ located in the box $[-N,N]^6$ with the "excluded diagonals"; that is,

$$x_1 \neq y_1, \qquad x_2 \neq y_2, \qquad (n_1,n_2) \neq (0,0),$$

is at most $N^{4-\delta}$ for some fixed $\delta>0$, for all sufficiently large N.

1. Introduction

A sequence of points $\{\theta_n\}_{n=1}^{\infty}$ is uniformly distributed modulo one if given any fixed interval I in the unit circle $\mathbb{R}/\mathbb{Z}$, the proportion of fractional parts θ_n mod 1 which lie in I tends to the length of the interval I; that is,

$$\#\{n \le N : \theta_n \bmod 1 \in I\} \sim \operatorname{length}(I)\cdot N, \quad N \to \infty.$$

We study the *pair correlation function* R_2, defined for every fixed interval $I \subset \mathbb{R}$ by the property that

$$\lim_{N\to\infty}\frac{1}{N}\#\left\{1 \le m \neq n \le N : |\theta_n-\theta_m| \in \frac{1}{N}I\right\} = \int_I R_2(x)\,dx,$$

assuming that the limit exists. For a random sequence of N elements—that is, N uniform independent random variables in $[0,1)$ (the Poisson model)—the limiting pair correlation function is almost surely $R_2(x) \equiv 1$.

There are very few positive results on the pair correlation function available for specific sequences, a notable exception being the fractional parts of $\sqrt{n}$ [6]; a more

Illinois Journal of Mathematics, Vol. 64, No. 4 (2020), 583–594
DOI 10.1215/00192082-8720506,
Received February 4, 2020. Received in final form June 10, 2020.
First published online September 16, 2020.
2010 Mathematics Subject Classification: Primary 11J54; Secondary 11J71.
This result is part of a project that received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 786758).

tractable problem is to randomize (a "metric" theory, in the terminology of uniform distribution theory) by looking at random multiples[1] $\theta_n = \alpha a(n) \bmod 1$, for almost all α. There is a well-developed metric theory of the pair correlation function for *integer* valued sequences $\{a(n)\}_{n=1}^{\infty}$, initiated in [11], where polynomial sequences such as $a(n) = n^d$, $d \geq 2$, are studied, with several developments in the last few years; see, e.g., [2, 3, 5, 8–10, 12, 13]. In this note, we study the case of real-valued lacunary sequences: Let $a(x) > 0$ be a lacunary sequence of positive reals—that is, there is some $C > 1$ so that for all integers $x \geq 1$,

$$a(x+1) \geq Ca(x).$$

For instance, we can take $a(x) = e^x$. It is known that for almost all α, the sequence $\alpha a(x) \bmod 1$ is uniformly distributed mod one [7, Chapter 1, Corollary 4.3]. Here and throughout this note, "almost all" is meant with respect to the Lebesgue measure on $\mathbb{R}$.

THEOREM 1.1
Assume that $\{a(x)\}_{x=1}^{\infty}$ is a lacunary sequence of positive reals. Then the pair correlation function of the sequence $\{\alpha a(x)\}_{x=1}^{\infty}$ is Poissonian for almost all α.

When $a(x)$ takes *integer* values, it was shown in [12] that for almost all α, the pair correlation function is Poissonian. The case of pair correlation of sequences of *rationals* $x_n = a_n/b_n$ with a_n integer-valued and lacunary and b_n integer-valued and (roughly speaking) sufficiently small (e.g., $a_n/b_n = 2014^n/[\log\log n]$) was treated in [5]. Here we treat any real-valued sequences.

We will reduce the problem to giving a bound for the number of lattice points satisfying a Diophantine inequality: For $M = N^{1+\varepsilon}$ and $K = N^{\varepsilon}$, let $\mathcal{S}(N)$ be the set of integer six-tuples with

$$1 \leq y_i \neq x_i \leq N, \qquad 1 \leq |n_i| \leq M \quad (i = 1, 2),$$

satisfying

$$\big|n_1\big(a(x_1) - a(y_1)\big) - n_2\big(a(x_2) - a(y_2)\big)\big| < K.$$

Assume that

$$\text{(A)} \qquad \#\mathcal{S}(N) \ll N^{4-\delta}.$$

THEOREM 1.2
Let $\{a(x)\}_{x=1}^{\infty}$ be a sequence of distinct positive reals. Assume that (A) holds for some $\delta > 0$. Then the pair correlation function of $\alpha a(x)$ is Poissonian for Lebesgue almost all $\alpha \in \mathbb{R}$.

1. A different notion of randomizing has recently been investigated in [1], which studies the pair correlation of the sequence $\alpha^n \bmod 1$ with α random.

In the case of integer-valued sequences, the almost sure convergence of the pair correlation function to the Poisson limit (metric Poisson pair correlation) follows [11, 12] from a similar bound for the equation

$$n_1\big(a(x_1)-a(y_1)\big)-n_2\big(a(x_2)-a(y_2)\big)=0.$$

See [2, 4] for a streamlined criterion for metric Poisson pair correlation in terms of the *additive energy* of the sequence.

In Section 4, we verify that (A) holds for lacunary sequences; hence, we obtain Theorem 1.1.

2. The pair correlation function

To study the pair correlation function, we use a smooth count (cf. [11]): For $f \in C_c^\infty(\mathbb{R})$ or f being an indicator function of a compact interval, set

$$F_N(x)=\sum_{j\in\mathbb{Z}} f\big(N(x+j)\big),$$

which is periodic and localized on scale $1/N$. For a sequence $\{\theta_n\}_{n=1}^\infty \subset \mathbb{R}/\mathbb{Z}$, we define its pair correlation function by

$$R_2(f,N)\big(\{\theta_n\}_{n=1}^\infty\big)=\frac{1}{N}\sum_{1\le m\ne n\le N} F_N(\theta_n-\theta_m). \tag{2.1}$$

In particular, for a fixed sequence $\{a_n\}_{n=1}^\infty$, we take $\theta_n=\alpha a_n \bmod 1$, and abbreviate the pair correlation function $R_2(f,N)(\{\theta_n\}_{n=1}^\infty)$, having fixed f, by $R_2(f,N)(\alpha)=R_2(\alpha)$.

It suffices to restrict α to lie in a fixed finite interval and to consider a smooth average: Let $\rho \in C_c^\infty(\mathbb{R})$, $\rho \ge 0$, be a smooth, compactly supported, non-negative weight function, normalized to give a probability density: $\int_{\mathbb{R}} \rho(\alpha)\,d\alpha = 1$. We define a smooth average

$$\langle X\rangle=\int_{\mathbb{R}} X(\alpha)\rho(\alpha)\,\mathrm{d}\alpha. \tag{2.2}$$

2.1. The expected value

We verify later (see Lemma 2.2) that (A) implies that there exists $\delta>0$ such that for $M=N^{1+\varepsilon}$ and $K \ll N^\varepsilon$ we have that

$$\#\big\{1\le n\le M, 1\le x\ne y\le N : n\big|a(x)-a(y)\big|<K\big\} \ll N^{2-\delta}. \tag{B}$$

Assuming for now this statement is true, we know $\langle R_2(f,N)\rangle$ asymptotically, as follows.

LEMMA 2.1

If (B) holds, then the expected value of $R_2(f,N)(\alpha)$ is

$$\big\langle R_2(f,N)\big\rangle=\int_{-\infty}^{\infty} f(x)\,dx+O(N^{-\delta}).$$

Proof
Let $f \in C_c^\infty(\mathbb{R})$. By using Poisson summation, we have the expansion

$$F_N(x) = \sum_{j\in\mathbb{Z}} f\big(N(x+j)\big) = \frac{1}{N}\sum_{n\in\mathbb{Z}} \widehat{f}\Big(\frac{n}{N}\Big)e(nx)$$

with $e(z) := e^{2\pi i z}$, which gives

$$R_2(\alpha) = \frac{1}{N^2}\sum_{n\in\mathbb{Z}} \widehat{f}\Big(\frac{n}{N}\Big)S_{n,N}(\alpha), \tag{2.3}$$

where

$$S_{n,N}(\alpha) = \sum_{1\le x\ne y\le N} e\big(\alpha n\big(a(x)-a(y)\big)\big).$$

Therefore, the expected value is

$$\langle R_2\rangle = \int_{-\infty}^{\infty} R_2(\alpha)\rho(\alpha)\,d\alpha = \frac{1}{N^2}\sum_{n\in\mathbb{Z}} \widehat{f}\Big(\frac{n}{N}\Big)\sum_{1\le x\ne y\le N} \widehat{\rho}\big(n\big(a(x)-a(y)\big)\big).$$

The zero mode $n = 0$ gives a contribution of

$$\frac{1}{N^2}\widehat{f}(0)N(N-1) = \int_{-\infty}^{\infty} f(x)\,dx\big(1+O(1/N)\big).$$

We split the sum over non-zero modes into two terms: Those with $1 \le |n| \le M = N^{1+\varepsilon}$, and those with $|n| > M$. To treat the contribution of modes with $|n| > M = N^{1+\varepsilon}$, we use $|\widehat{f}(x)| \ll x^{-A}$ and $|\widehat{\rho}| \ll 1$ to bound that term by

$$\frac{1}{N^2}\sum_{|n|>M}\Big(\frac{n}{N}\Big)^{-A}\sum_{1\le x\ne y\le N} 1 = \frac{N^A}{M^{A-1}} \ll \frac{1}{N^{1-\varepsilon}}$$

on choosing $A = 2/\varepsilon$.

To bound the contribution of modes with $1 \le |n| \le M$, we separate into a contribution of terms with $|n(a(x)-a(y))| < K$ and the rest.

We use $|\widehat{\rho}|, |\widehat{f}| \ll 1$ to obtain that the contribution of terms with $|n(a(x) - a(y))| < K$ is

$$\ll \frac{1}{N^2}\#\big\{1 \le n < N^{1+\varepsilon}, 1 \le y \ne x \le N : n\big(a(x)-a(y)\big) < N^{\varepsilon}\big\}.$$

By (B), this is $\ll N^{-\delta}$.

The contribution of terms with $|n(a(x)-a(y))| > K$ is bounded using

$$\big|\widehat{\rho}\big(n\big(a(x)-a(y)\big)\big)\big| \ll \big|n\big(a(x)-a(y)\big)\big|^{-A} \le K^{-A} = N^{-2}$$

and $|\widehat{f}| \ll 1$ by

$$\frac{1}{N^2}\sum_{\substack{1\le |n|\le M\\ 1\le x\ne y\le N\\ |n(a(x)-a(y))|>K}} \Big|\widehat{f}\Big(\frac{n}{N}\Big)\Big|\widehat{\rho}\big(n\big(a(x)-a(y)\big)\big)| \ll \frac{1}{N^2}\sum_{\substack{1\le |n|\le M\\ 1\le x\ne y\le N}} \frac{1}{N^2} \le \frac{M}{N^2},$$

which is $\ll N^{-1+\varepsilon}$. ◻

2.2. The condition (B)

LEMMA 2.2

If (A) holds for $\delta > 0$ and $K = N^{\varepsilon}$, then (B) is satisfied with $\delta' = \delta/2$ in place of δ and $K' = 2N^{\varepsilon}$ in place of K.

Proof

Let $1_{[0,K']}$ be the indicator function of $[0, K']$, and let $N \geq 1$ be large. With n, x, y and n_i, x_i, y_i constrained as in (B), we note that

$$\Big(\sum_{n,x,y} 1_{[0,K']}\big(n\big|a(x)-a(y)\big|\big)\Big)^2$$

equals

$$\begin{aligned}\sum_{n_i,x_i,y_i} & 1_{[0,K']}\big(\big|n_1\big(a(x_1)-a(y_1)\big)\big|\big)1_{[0,K']}\big(\big|n_2\big(a(x_2)-a(y_2)\big)\big|\big) \\ &\leq \sum_{n_i,x_i,y_i} 1_{[0,K]}\big(\big|n_1\big(a(x_1)-a(y_1)\big)-n_2\big(a(x_2)-a(y_2)\big)\big|\big) = \#S(N).\end{aligned}$$

Since $\#\mathcal{S}(N) \ll N^{4-\delta}$, combining these considerations and taking a square root completes the proof. □

2.3. The variance

PROPOSITION 2.3

Assume that $a(x)$ is a sequence of real numbers such that (A) holds. Then

$$\Big\langle\Big|R_2(f,N)-\int_{-\infty}^{\infty} f(x)\,dx\Big|^2\Big\rangle \ll N^{-\delta}.$$

Proof

By Cauchy–Schwarz,

$$\begin{aligned}\Big\langle\Big|R_2(f,N)-\int_{-\infty}^{\infty} f(x)\,dx\Big|^2\Big\rangle &\leq 2\big\langle\big|R_2(f,N)-\langle R_2\rangle\big|^2\big\rangle \\ &\quad + 2\Big\langle\Big|\langle R_2\rangle-\int_{-\infty}^{\infty} f(x)\,dx\Big|^2\Big\rangle.\end{aligned}$$

By Lemma 2.1,

$$\Big\langle\Big|\langle R_2\rangle-\int_{-\infty}^{\infty} f(x)\,dx\Big|^2\Big\rangle \ll N^{-2\delta}.$$

We now show that

$$\operatorname{Var} R_2 = \big\langle\big|R_2(f,N)-\langle R_2\rangle\big|^2\big\rangle \ll N^{-\delta} \tag{2.4}$$

which will prove Proposition 2.3.

To prove (2.4), it suffices to show by (A) that

$$\operatorname{Var} R_2 \ll_f \frac{\#\mathcal{S}(N)}{N^4}. \tag{2.5}$$

By using the expansion (2.3), the variance can be written as

$$\operatorname{Var}(R_2) = \frac{1}{N^4} \sum_{(n_1,n_2)\in\mathbb{Z}^2\setminus\{0\}} \widehat{f}\Big(\frac{n_1}{N}\Big)\widehat{f}\Big(\frac{n_2}{N}\Big) w(n_1,n_2,N), \tag{2.6}$$

where for integers n_1, n_2, we let

$$w(n_1,n_2,N) = \sum_{\substack{1\le x_1\ne x_3\le N,\\ 1\le x_2\ne x_4\le N}} \widehat{\rho}\big(n_1\big(a(x_3)-a(x_1)\big) - n_2\big(a(x_4)-a(x_2)\big)\big)$$

and ρ as in (2.2).

Due to the rapid decay of $\widehat{f}$, the contribution from the range in which $|n_1|$ or $|n_2|$ exceeds $M = N^{1+\varepsilon}$ is negligible, as we will argue now. We detail only the case $\max\{|n_1|,|n_2|\} = n_1 \ge M$ since the other case can be done similarly. We observe the trivial bound $|w(n_1,n_2,N)| \ll N^4$. Moreover,

$$n_1 = n_1^{\varepsilon/2} n_1^{1-\varepsilon/2} \ge n_1^{\varepsilon/2} N^{1+\varepsilon/2-\varepsilon^2/2},$$

which, since ε is small, yields $n_1 > n_1^{\varepsilon/2} N^{1+\varepsilon/3}$. Hence, the contribution to the right-hand side of (2.6) arising from the terms with $\max\{|n_1|,|n_2|\} = n_1 \ge M = N^{1+\varepsilon}$ and $n_2 \ne 0$ is

$$\ll \frac{1}{N^4} \sum_{\substack{n_1,n_2\in\mathbb{Z}\setminus\{0\}\\ |n_1|>N^{1+\varepsilon}}} \Big(\frac{n_1}{N}\Big)^{-18/\varepsilon} \sum_{n_2\ne 0} \widehat{f}\Big(\frac{n_2}{N}\Big) N^4 \ll \frac{1}{N^4}.$$

Moreover, the terms satisfying $\max\{|n_1|,|n_2|\} = n_1 \ge N^{1+\varepsilon}$, and $n_2 = 0$ are in absolute value

$$\ll \frac{1}{N^4} \sum_{|n_1|\ge N^{1+\varepsilon}} \widehat{f}\Big(\frac{n_1}{N}\Big) N^4 \ll \frac{1}{N^4}.$$

So, the upshot is that on the right-hand side of (2.6), the sum over all (n_1,n_2) with $\max(|n_1|,|n_2|) > N^{1+\varepsilon}$ contributes $\ll N^{-4}$. By the rapid decay of $\widehat{\rho}$, we can dispose of the regime where $|n_1(a(x_3)-a(x_1)) - n_2(a(x_4)-a(x_2))| \ge N^{\varepsilon}$.

By bounding $\hat{\rho}$ trivially, we find that

$$\operatorname{Var} R_2 \ll \frac{\#\mathcal{S}(N)}{N^4} + O\Big(\frac{1}{N^4}\Big).$$

Since $\#\mathcal{S}(N) \ge N^3$, we obtain (2.5). □

3. Almost everywhere convergence: Proof of Theorem 1.2

We now deduce almost everywhere convergence from a polynomial variance bound.

3.1. Preparations

We need a general property of the pair correlation function. Recall that for any sequence of points $\{\theta_n\}_{n=1}^{\infty} \subset \mathbb{R}/\mathbb{Z}$, we defined

$$R_2(f, N) = \frac{1}{N} \sum_{1 \le j \ne k \le N} F_N(\theta_j - \theta_k)$$

with $F_N(x) = \sum_{j \in \mathbb{Z}} f(N(x - j))$.

LEMMA 3.1

Suppose there is a strictly increasing sequence $\{N_m\}_{m=1}^{\infty} \subseteq \mathbb{Z}_{\ge 1}$, with

$$\lim_{m\to\infty} \frac{N_{m+1}}{N_m} = 1$$

so that for all $f \in C_c^{\infty}(\mathbb{R})$,

$$\lim_{m\to\infty} R_2(f, N_m) = \int_{-\infty}^{\infty} f(x)\, dx. \tag{3.1}$$

Then we can pass from the sub-sequence to the set of all integers:

$$\lim_{N\to\infty} R_2(f, N) = \int_{-\infty}^{\infty} f(x)\, dx \tag{3.2}$$

for all $f \in C_c^{\infty}(\mathbb{R})$.

Proof

We will first deduce that (3.1) holds for the indicator functions

$$I_s(x) = \begin{cases} 1, & |x| < s/2, \\ 0, & \text{otherwise,} \end{cases}$$

by approximating with smooth functions, and show that (3.2) holds for the functions I_s, and then deduce by approximating a general even smooth $f \in C_c^{\infty}(\mathbb{R})$ by linear combinations of I_s that (3.2) holds for all such f. Note that for odd smooth $f \in C_c^{\infty}(\mathbb{R})$, we have $F_N(-x) = -F_N(x)$ which entails $R_2(f, N) = 0$, so the pair correlation function $R_2(f, N)$ converges trivially to the right limit.

From the definition of $R_2(I_s, N)$, we have a monotonicity property: Let $0 < \varepsilon < 1$. If $(1-\varepsilon)N' < N < N'$ and $N'' < N < (1+\varepsilon)N''$, then

$$(1-\varepsilon) R_2(I_{(1-\varepsilon)s}, N'') \le R_2(I_s, N) \le \frac{1}{1-\varepsilon} R_2(I_{s/(1-\varepsilon)}, N'). \tag{3.3}$$

Indeed, using positivity of I_s (hence, of F_N),

$$N \cdot R_2(I_s, N) = \sum_{1 \le j \ne k \le N} F_N(\theta_j - \theta_k) \le \sum_{1 \le j \ne k \le N'} F_N(\theta_j - \theta_k).$$

Now if $1 > N/N' \ge 1 - \varepsilon > 0$, then since I_s is even and decreasing on $[0, \infty)$, we have

$$I_s(Ny) = I_s\Big(N'y \cdot \frac{N}{N'}\Big) \le I_s\big(N'y(1-\varepsilon)\big) = I_{s/(1-\varepsilon)}(N'y).$$

So

$$F_N(x) = \sum_{j\in\mathbb{Z}} I_s\big(N\cdot(x-j)\big) \le \sum_{j\in\mathbb{Z}} I_{s/(1-\varepsilon)}\big(N'\cdot(x-j)\big) = \tilde{F}_{N'}(x),$$

where $\tilde{F}_{N'}(y) = \sum_{j\in\mathbb{Z}} I_{s/(1-\varepsilon)}(N'(y-j))$. Hence,

$$R_2(I_s, N) \le \frac{N'}{N} R_2(I_{s/(1-\varepsilon)}, N') \le \frac{1}{1-\varepsilon} R_2(I_{s/(1-\varepsilon)}, N'),$$

which proves the upper bound in (3.3). The lower bound of (3.3) follows from switching the roles of N and N'' and inserting in the upper bound.

Next, fix $\varepsilon \in (0,1)$ small, let $N \gg 1$, and take $m \gg 1$ so that

$$N_m < N_{m+1} < (1+\varepsilon)N_m,$$

and so if $N_m \le N < N_{m+1}$, then

$$(1-\varepsilon)N_{m+1} < N < N_{m+1}, \qquad N_m \le N < (1+\varepsilon)N_m.$$

Then for all $s > 0$, we have

$$(1-\varepsilon)R_2(I_{(1-\varepsilon)s}, N_m) \le R_2(I_s, N) \le \frac{1}{1-\varepsilon} R_2(I_{s/(1-\varepsilon)}, N_{m+1}).$$

Taking $m \to \infty$, we find by (3.1)

$$\limsup_{N\to\infty} R_2(I_s, N) \le \frac{1}{1-\varepsilon}\int_{-\infty}^{\infty} I_{s/(1-\varepsilon)}\,dx = \frac{s}{(1-\varepsilon)^2}$$

and

$$\liminf_{N\to\infty} R_2(I_s, N) \ge (1-\varepsilon)\int_{-\infty}^{\infty} I_{(1-\varepsilon)s}\,dx = (1-\varepsilon)^2 s.$$

Since $\varepsilon > 0$ is arbitrary, we finally obtain

$$\lim_{N\to\infty} R_2(I_s, N) = s = \int_{-\infty}^{\infty} I_s(x)\,dx$$

so that (3.2) holds for all indicator functions I_s. Therefore, (3.2) holds for all test functions $f \in C_c^\infty(\mathbb{R})$. □

3.2. *Proof of Theorem 1.2*

Proof

It suffices to show that for almost every α in a fixed compact interval I, we have

$$\lim_{N\to\infty} R_2(f, N)(\alpha) = \int_{-\infty}^{\infty} f(x)\,dx \tag{3.4}$$

for all $f \in C_c^\infty(\mathbb{R})$.

Let $\rho \in C_c^\infty(\mathbb{R})$ be a non-negative function majorizing the indicator function of the interval I: $\mathbf{1}_I \le \rho$. Then from the variance bound of Proposition 2.3, we find that for some $\delta > 0$, for all $f \in C_c^\infty(\mathbb{R})$,

$$\int_I \left| R_2(f, N_m)(\alpha) - \int_{-\infty}^{\infty} f(x)\,dx \right|^2 \rho(\alpha)\,d\alpha \ll_f N^{-\delta}.$$

Hence, for the sequence

$$N_m = \lfloor m^{2/\delta} \rfloor,$$

we have that for almost all $\alpha \in I$,

$$\lim_{m\to\infty} R_2(f, N_m)(\alpha) = \int_{-\infty}^{\infty} f(x)\,dx \tag{3.5}$$

for all f. Indeed, for each fixed f set,

$$X_m(\alpha) = \Big| R_2(f, N_m)(\alpha) - \int_{-\infty}^{\infty} f(x)\,dx \Big|^2.$$

Then

$$\int_I X_m(\alpha)\,d\alpha \leq \int_{-\infty}^{\infty} X_m(\alpha)\rho(\alpha)\,d\alpha \ll \frac{1}{N_m^{\delta}} \ll \frac{1}{m^2}.$$

Therefore,

$$\int_I \Big(\sum_{m\geq 1} X_m(\alpha)\Big)\,d\alpha \leq \sum_{m\geq 1} \int_{-\infty}^{\infty} X_m(\alpha)\,d\alpha \ll_f \sum_{m\geq 1} \frac{1}{m^2} < \infty,$$

so that $\sum_{m\geq 1} X_m(\alpha)$ converges for almost all $\alpha \in I$. Thus,

$$\lim_{m\to\infty} X_m(\alpha) = 0$$

for almost all α—i.e., (3.5) holds for our specific f for almost all $\alpha \in I$.

By a diagonalization argument (see [11]) that involves selecting a countable and dense collection of functions f, there is a set of α whose complement has measure zero so that (3.5) holds for all f. Since $N_{m+1}/N_m \to 1$, we can use Lemma 3.1 to deduce (3.4) holds, proving Theorem 1.2. □

4. Lacunary sequences

From now on, we assume that $\{a(x)\}_{x=1}^{\infty}$ is a lacunary sequence of (strictly) positive reals—that is, there is some $C > 1$ so that

$$a(x+1) \geq C a(x)$$

for all integers $x \geq 1$. Consequently, we have for all $x \geq y \geq 1$ that

$$a(x) \geq C^{x-y} a(y).$$

We will show that (A) holds, hence proving that the pair correlation function of $\{\alpha a(x) \bmod 1\}_{x=1}^{\infty}$ is Poissonian for almost all α (that is, Theorem 1.1).

4.1. The condition (A)

PROPOSITION 4.1

Denote by $\mathcal{S}(N)$ the set of integer six-tuples with $1 \leq y_i \neq x_i \leq N, 1 \leq |n_i| \leq M$ for $i = 1, 2$ that satisfy

$$\big| n_1\big(a(x_1) - a(y_1)\big) - n_2\big(a(x_2) - a(y_2)\big)\big| < K.$$

Assume that $\{a(x)\}_{x=1}^{\infty}$ is a lacunary sequence of positive real numbers. Further, suppose $K = o(M)$ and $N^{\gamma} \ll M \ll N^{\Gamma}$ for some $0 < \gamma < \Gamma < 2$. Then

$$\#\mathcal{S}(N) \ll MN^2(\log M)^2.$$

Note that Theorem 1.2 and the Proposition 4.1 together imply Theorem 1.1.

Proof

The proof is a modification of [12, Proposition 2]: We are given the inequality

$$\left|n_1\big(a(x_1)-a(y_1)\big)-n_2\big(a(x_2)-a(y_2)\big)\right| < K. \tag{4.1}$$

We may assume that $n_i > 0$, and $1 \le y_i < x_i \le N$, $i = 1,2$, and that $x_1 \ge x_2$. In particular, we may then assume that $x_1 \ge 4\log_C M \gg \log N$ because the number of such tuples with $x_1 \ll \log N$ is at most $O(M^2(\log N)^4)$, which is admissible (that is, $o(MN^2(\log N)^2)$) if $M = O(N^{\Gamma})$ for $\Gamma < 2$.

We fix n_1, x_1, y_1, and first show that (recall $x_1 \ge x_2$)

$$x_1 - x_2 \le 2\log_C M. \tag{4.2}$$

Indeed, we have a lower bound,

$$n_1\big(a(x_1)-a(y_1)\big) \ge 1\cdot\big(a(x_1)-a(x_1-1)\big) \ge \Big(1-\frac{1}{C}\Big)a(x_1)$$

(since $y \le x_1 - 1$), and an upper bound,

$$n_2\big(a(x_2)-a(y_2)\big) \le Ma(x_2) = a(x_1)M\frac{a(x_2)}{a(x_2+(x_1-x_2))} \le \frac{M}{C^{x_1-x_2}}a(x_1)$$

since $a(x+h) \ge C^h a(x)$ for $h \ge 1$. Hence,

$$n_1\big(a(x_1)-a(y_1)\big)-n_2\big(a(x_2)-a(y_2)\big) \ge \Big(1-\frac{1}{C}\Big)a(x_1) - \frac{M}{C^{x_1-x_2}}a(x_1).$$

Assuming that $x_1 - x_2 > 2\log_C M$ gives in particular

$$1-\frac{1}{C}-\frac{M}{C^{x_1-x_2}} > 1-\frac{1}{C}-\frac{1}{M} > \frac{1}{2}\Big(1-\frac{1}{C}\Big) > 0$$

for sufficiently large N. The condition (4.1) now forces

$$\frac{1}{2}\Big(1-\frac{1}{C}\Big) < \Big(1-\frac{1}{C}\Big) - \frac{M}{C^{x_1-x_2}} \le \frac{K}{a(x_1)} \ll \frac{K}{C^{x_1}}$$

which forces $x_1 \ll \log_C K \le \varepsilon \log M$, which we assumed was not the case. Thus. we may assume that $x_1 - x_2 \le 2\log_C M$, which forces $x_2 \ge 2\log_C M$ since $x_1 > 4\log_C M$.

Now fix x_2 as well; then n_2 will be determined by y_2 because

$$n_2 = \frac{n_1(a(x_1)-a(y_1))}{a(x_2)-a(y_2)} + O\Big(\frac{K}{a(x_2)-a(y_2)}\Big)$$

and since $a(y)$ is lacunary, $K/(a(x_2)-a(y_2)) = o(1)$ if $y_2 \ge \log_C N$ because

$$a(x_2)-a(y_2) \ge a(x_2)-a(x_2-1) \ge a(x_2)\Big(1-\frac{1}{C}\Big) \gg C^{x_2} \ge M^2$$

since $x_2 > 2\log_C M$.

So we will be done if we show that there are $\ll \log N$ many choices of y_2 such that $x_2 - y_2 > 2\log_C M$. Indeed, if there are two pairs (y_2, n_2) and (y_2', n_2') for which (4.1) holds (recall all other variables are now fixed), with $x_2 - y_2 > 2\log_C M$, $x_2 - y_2' > 2\log_C M$, then since

$$a(y_2) \le \frac{a(x_2)}{C^{x_2-y_2}} \le \frac{a(x_2)}{M^2},$$

we find that (4.1) implies

$$\begin{aligned} n_1\big(a(x_1) - a(y_1)\big) &= n_2\big(a(x_2) - a(y_2)\big) + O(K) \\ &= n_2 a(x_2)\Big(1 + \frac{a(y_2)}{a(x_2)} + O\Big(\frac{K}{n_2 a(x_2)}\Big)\Big) \\ &= n_2 a(x_2)\Big(1 + O\Big(\frac{K}{M^2}\Big)\Big) \end{aligned}$$

since $n_2 a(x_2) \ge a(x_2) \ge C^{x_2} \gg M^2$ if $x_2 \ge 2\log_C M$, and $a(x_2)/a(y_2) \ge C^{x_2-y_2} \gg M^2$. If n_2', y_2' is another such pair, then we also find

$$n_1\big(a(x_1) - a(y_1)\big) = n_2' a(x_2)\Big(1 + O\Big(\frac{K}{M^2}\Big)\Big)$$

so that

$$n_2 a(x_2)\Big(1 + O\Big(\frac{K}{M^2}\Big)\Big) = n_2' a(x_2)\Big(1 + O\Big(\frac{K}{M^2}\Big)\Big)$$

which gives

$$n_2' = n_2\Big(1 + O\Big(\frac{K}{M^2}\Big)\Big) = n_2 + O\Big(\frac{K}{M}\Big)$$

since $n_2 \le M$. Thus, for $M \gg N^\gamma$, while $K \ll N^\varepsilon = o(M)$, we obtain $n_2' = n_2$. Substituting this information in (4.1), and recalling that n_1, x_1, y_1 are fixed and n_2, x_2 are determined now (up to $\ll \log N$ many choices), we infer that there are $\ll \log N$ many choices for y_2. This completes the proof. □

Acknowledgments. We thank the referee for a careful reading of the manuscript, and useful suggestions.

References

[1] C. Aistleitner and S. Baker, *On the pair correlations of powers of real numbers*, to appear in Israel J. Math, preprint, arXiv:1910.01437 [math.NT].

[2] C. Aistleitner, G. Larcher, and M. Lewko, *Additive energy and the Hausdorff dimension of the exceptional set in metric pair correlation problems*, with an appendix by J. Bourgain, Israel J. Math. **222** (2017), no. 1, 463–485. MR 3736514. DOI 10.1007/s11856-017-1597-5.

[3] V. Becher, O. Carton, and I. Cunningham, *Low discrepancy sequences failing Poissonian pair correlations*, Arch. Math. (Basel) **113** (2019), no. 2, 169–178. MR 3977510. DOI 10.1007/s00013-019-01336-3.

[4] T. F. Bloom, S. Chow, A. Gafni, and A. Walker, *Additive energy and the metric Poissonian property*, Mathematika **64** (2018), no. 3, 679–700. MR 3826482. DOI 10.1112/s0025579318000207.

[5] S. Chaubey, M. Lanius, and A. Zaharescu, *Pair correlation of fractional parts derived from rational valued sequences*, J. Number Theory **151** (2015), 147–158. MR 3314206. DOI 10.1016/j.jnt.2014.12.014.

[6] D. El-Baz, J. Marklof, and I. Vinogradov, *The two-point correlation function of the fractional parts of* $\sqrt{n}$ *is Poisson*, Proc. Amer. Math. Soc. **143**, Amer. Math. Soc., Providence, RI, 2015, 2815–2828. MR 3336607. DOI 10.1090/S0002-9939-2015-12489-6.

[7] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, New York, 1974. MR 0419394.

[8] T. Lachmann and N. Technau, *On exceptional sets in the metric Poissonian pair correlations problem*, Monatsh. Math. **189** (2019), no. 1, 137–156. MR 3948291. DOI 10.1007/s00605-018-1199-2.

[9] G. Larcher and W. Stockinger, *Some negative results related to Poissonian pair correlation problems*, Discrete Math. **343** (2020), art. 111656. MR 4040039. DOI 10.1016/j.disc.2019.111656.

[10] I. Pirsic and W. Stockinger, *The Champernowne constant is not Poissonian*, Funct. Approx. Comment. Math. **60** (2019), no. 2, 253–262. MR 3964264. DOI 10.7169/facm/1749.

[11] Z. Rudnick and P. Sarnak, *The pair correlation function of fractional parts of polynomials*, Comm. Math. Phys. **194** (1998), 61-70 . MR 1628282. DOI 10.1007/s002200050348.

[12] Z. Rudnick and A. Zaharescu, *A metric result on the pair correlation of fractional parts of sequences*, Acta Arith. **89** (1999), 283–293. MR 1691857. DOI 10.4064/aa-89-3-283-293.

[13] A. Walker, *The primes are not metric Poissonian*, Mathematika 64 (2018), no. 1, 230–236. MR 3778222. DOI 10.1112/S002557931700050X.

Rudnick: School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel; rudnick@tauex.tau.ac.il

Technau: School of Mathematical Sciences, Tel Aviv University, Tel Aviv, Israel; niclast@mail.tau.ac.il