A METRIC THEORY OF MINIMAL GAPS

ZE’EV RUDNICK

Abstract. We study the minimal gap statistic for fractional parts of sequences of the form $A^\alpha = \{\alpha a(n)\}$ where $A = \{a(n)\}$ is a sequence of distinct integers. Assuming that the additive energy of the sequence is close to its minimal possible value, we show that for almost all $\alpha$, the minimal gap $\delta_{\min}^\alpha(N) = \min\{\alpha a(m) - \alpha a(n) \mod 1 : 1 \leq m \neq n \leq N\}$ is close to that of a random sequence.

We start with a sequence of points $X = \{x_n : n = 1, 2, \ldots\} \subset \mathbb{R}/\mathbb{Z}$ in the unit interval/circle, which we assume is asymptotically uniformly distributed: For any subinterval $I \subset \mathbb{R}/\mathbb{Z}$, we have

(1) $\lim_{N \to \infty} \frac{1}{N} \#\{n \leq N : x_n \in I\} = |I|.$

In particular, the mean spacing between the points lying in any subinterval is $1/N$.

Our goal is to understand the minimal gap

$\delta_{\min}(X, N) = \min(|x_n - x_m| : n, m \leq N, n \neq m)$

(with a suitable modification for wrapping around).

For random points, namely $N$ independent uniform points in the unit interval (Poisson process), the minimal gap is almost surely of size $1/N^2$ [9]. In this note we study the metric theory of the minimal gap statistic for a class of deterministic sequences of fractional parts, such as fractional parts of polynomials. The case of quadratic polynomials $x_n = \alpha n^2$ has its roots in the recent paper [3], which studies the more complicated case of the minimal gap statistic for the sequence of eigenvalues of the Laplacian on a rectangular billiard, namely the points $\{\alpha m^2 + n^2 : m, n \geq 1\}$ on the real line.

We fix a sequence $A = \{a(n) : n = 1, 2, \ldots\} \subset \mathbb{Z}$ of distinct integers ($a(n) \neq a(m)$ if $m \neq m$), and study the minimal gap statistic of fractional parts of the set

$A^\alpha = \{\alpha a(n) \mod 1 : n = 1, 2, \ldots\} \subset \mathbb{R}/\mathbb{Z}$.

(it is an old result of Weyl that $A^\alpha$ satisfies (1) for almost all $\alpha$). We want to know under which conditions we can show that for almost all $\alpha$, the minimal gap statistics

$\delta_{\min}^\alpha(N) = \delta_{\min}(A^\alpha, N)$

follows that of the random case, that is of size about $1/N^2$ for almost all $\alpha$. It is easy to see that we cannot have much smaller minimal gaps:

**Theorem 1.** Assume that $A$ consists of distinct integers. Then for all $\eta > 0$, for almost all $\alpha$,

$\delta_{\min}^\alpha(N) > \frac{1}{N^{2+\eta}}, \quad \forall N > N_0(\alpha).$

Date: August 11, 2018.
To make the minimal gap small, we give a criterion in terms of the “additive energy” $E(A, N)$ of the sequence:

$$E(A, N) := \# \{(n_1, n_2, n_3, n_4) \in [1, N]^4 : a(n_1) + a(n_2) = a(n_3) + a(n_4)\}.$$  

Note that $N^2 \leq E(A, N) \leq N^3$. The result is

**Theorem 2.** Assume that $A$ consists of distinct integers, and that the additive energy satisfies

$$E(A, N) \ll N^{2+o(1)}, \ \forall N \gg 1.$$  

Then for all $\eta > 0$, for almost all $\alpha$,

$$\delta_{\min}(N) < \frac{1}{N^{2-\eta}}, \ \forall N > N_0(\alpha).$$  

Examples: For $a(n) = n^d$, $d \geq 2$, it is shown in [11] that the $E(A, N) \ll N^{2+o(1)}$. For lacunary sequences, we have $E(A, N) \ll N^2$ [12]. Hence Theorem 2 applies to these sequences.

Relaxing the required bound on the additive energy will give a weaker result on the minimal spacing, basically that $\delta_{\min}(N) < E(A, N)/N^{\frac{3-\eta'}{4}}$ for some $\eta' > 0$. A notable case where the additive energy is bigger is that of $A = \mathcal{P}$ being the sequence of primes, where $E(\mathcal{P}, N) \approx N^3/\log N$. In this case, we cannot have gaps much larger than the average gap: A simple argument shows that given any $\varepsilon > 0$, for almost all $\alpha$, we have $\delta_{\min}(\mathcal{P}^\alpha, N) \gg 1/(N/(N(\log N)^{2+\varepsilon}))$, see § 8.

**Acknowledgements:** We thank Andrew Granville and Niclas Technau for their comments. The work was supported by the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement n° 320755.

1. A bilinear statistic

To study the minimal gap, we introduce a statistics counting all possible gaps: We start with a smooth, compactly supported window function $f \in \mathcal{C}_c^\infty([-1/2, 1/2])$, which is non-negative: $f \geq 0$, and of unit mass $\int f(x)dx = 1$, and define

$$F_M(x) = \sum_{j \in \mathbb{Z}} f(M(x + j))$$

which is localized on the scale of $1/M$, and periodic: $F_M(x + 1) = F_M(x)$. We then set

$$D_A(N, M)(\alpha) = \sum_{1 \leq m, n \leq N} F_M(\alpha a(n) - \alpha a(m)).$$

The expected value of $D_A(N, M)$ is easily seen to equal

$$\int_0^1 D_A(N, M)(\alpha)d\alpha = \frac{N(N-1)}{M} \sim \frac{N^2}{M}.$$  

This already suffices to show that minimal gaps cannot typically be small (Theorem 1), see § 2.

We will bound the variance of $D_A(N, M)$, from which Theorem 2 will follow:
**Proposition 3.**

\[
\text{Var } D_A(N, M) \ll \frac{1}{M} N^\epsilon E(A, N). 
\]

The statistic \( D_A(N, M)(\alpha) \) is related to the pair correlation function of the sequence \( A^\alpha \), which in our notation is \( D_A(N, N)(\alpha)/N \). Pair correlation measures gaps on the scale of the mean spacing, assumed here to be \( 1/N \), corresponding to \( M = N \); here we are looking at much smaller scales of \( M \) close to \( N^2 \).

The metric theory of the pair correlation function of fractional parts was initiated in [11], where the sequences \( a(n) = an^d \) were shown to almost surely have Poissonian pair correlation for \( d \geq 2 \) (see [10, 7] for different proofs of the quadratic case \( d = 2 \)). The problem has since been studied in several other cases and has recently been revived in an abstract setting [1, 14, 8, 4]. In particular, a convenient criterion for almost sure Poisson pair correlation has been formulated by Aistleitner, Larcher and Lewko [1] in terms of the additive energy \( E(A, N) \) of the sequence. The proof of Proposition 3 is close to that of the analogous statement for the pair correlation function in [1], which in turn is based on [11 12].

**Proof.** The Fourier expansion of \( F_M(x) \) is

\[
F_M(x) = \sum_{k \in \mathbb{Z}} \frac{1}{M} \hat{f}(\frac{k}{M}) e(kx) 
\]

where \( \hat{f}(y) = \int_{-\infty}^{\infty} f(x)e^{-2\pi ixy}dx \). Inserting into the definition of \( D(N, M) \) gives

\[
D(N, M)(\alpha) = \sum_{k \in \mathbb{Z}} \frac{1}{M} \hat{f}(\frac{k}{M}) \sum_{1 \leq m, n \leq N} \sum_{m \neq n} e(k\alpha(a(m) - a(n))). 
\]

Integrating over \( \alpha \) gives the expected value (we assume \( a(m) \neq a(n) \) if \( n \neq m \))

\[
\int_0^1 D(N, M)(\alpha)d\alpha = \frac{1}{M} \hat{f}(0) N(N - 1) = \frac{N(N - 1)}{M}. 
\]

The variance is the second moment of the sum over nonzero frequencies:

\[
\text{Var } D(N, M) = \int_0^1 \left| \sum_{0 \neq k \in \mathbb{Z}} \frac{1}{M} \hat{f}(\frac{k}{M}) \sum_{1 \leq m, n \leq N} \sum_{m \neq n} e(k\alpha(a(m) - a(n))) \right|^2 d\alpha. 
\]

Squaring out and integrating gives

\[
\text{Var } D(N, M) = \sum_{k_1, k_2 \neq 0} \frac{1}{M^2} \hat{f}(\frac{k_1}{M}) \hat{f}(\frac{k_2}{M}) \times 
\]

\[
\# \{m_1 \neq n_1, m_2 \neq n_2 : k_1(a(m_1) - a(n_1)) = k_2(a(m_2) - a(n_2)) \}. 
\]

We now follow [11] Lemma 3] to convert this to “GCD sums”. Let

\[
R(v) = \# \{1 \leq m \neq n \leq N : a(m) - a(n) = v \}. 
\]

Then

\[
\text{Var } D(N, M) = \sum_{v_1, v_2 \neq 0} \sum_{k_1, k_2 \neq 0} c(k_1)c(k_2)R(v_1)R(v_2)\delta(k_1v_1 = k_2v_2) 
\]

where we set

\[
c(k) := \frac{1}{M} \hat{f}(\frac{k}{M}). 
\]
Lemma 4. Let $f \in C_c^\infty(\mathbb{R})$. For any nonzero integers $v_1, v_2 \neq 0$,

$$\sum_{k_1, k_2 \neq 0} c(k_1)c(k_2)\delta(k_1v_1 = k_2v_2) \ll_f \frac{1}{M} \frac{\gcd(v_1, v_2)}{\sqrt{|v_1v_2|}}. \quad (5)$$

Proof. For $k_1, k_2 \neq 0$, we have $k_1v_1 = k_2v_2$ if and only if

$$(k_1, k_2) = \ell \left(\frac{v_2}{\gcd(v_1, v_2)}, \frac{v_1}{\gcd(v_1, v_2)}\right)$$

for some nonzero integer $0 \neq \ell \in \mathbb{Z}$. Abbreviating $a_i = v_i/(M \gcd(v_1, v_2))$, we find

$$\sum_{k_1, k_2 \neq 0} c(k_1)c(k_2)\delta(k_1v_1 = k_2v_2) = \sum_{0 \neq \ell \in \mathbb{Z}} c(\ell \frac{v_2}{\gcd(v_1, v_2)})c(\ell \frac{v_1}{\gcd(v_1, v_2)})$$

$$= \frac{1}{M} \sum_{0 \neq \ell \in \mathbb{Z}} \hat{f}(a_2\ell)\hat{f}(a_1\ell)$$

so that it suffices to show that

$$\sum_{0 \neq \ell \in \mathbb{Z}} \hat{f}(a_1\ell)\hat{f}(a_2\ell) \ll_f \frac{1}{\sqrt{|a_1a_2|}} = \frac{M \gcd(v_1, v_2)}{\sqrt{|v_1v_2|}}. \quad (6)$$

Applying Cauchy-Schwarz we get

$$\sum_{0 \neq \ell \in \mathbb{Z}} \hat{f}(a_1\ell)\hat{f}(a_2\ell) \leq \left(\sum_{\ell \neq 0} \hat{f}(a_1\ell)^2\right)^{1/2} \left(\sum_{\ell \neq 0} \hat{f}(a_2\ell)^2\right)^{1/2}.$$

We will obtain (6) if we show that for any $a > 0$,

$$\sum_{\ell \neq 0} \hat{f}(a\ell)^2 \ll_f \frac{1}{a}.$$

Indeed, if $0 < a \ll 1$ then we get a Riemann sum for $(\hat{f})^2$:

$$\sum_{\ell \neq 0} \hat{f}(a\ell)^2 \sim \frac{1}{a} \int_{-\infty}^{\infty} \hat{f}(y)^2dy = \frac{1}{a} \int_{-\infty}^{\infty} f(x)^2dx.$$

If $a \gg 1$ then use the decay rate of the Fourier transform: For $y \neq 0$,

$$|\hat{f}(y)| \leq \frac{1}{2\pi|y|} \int_{-\infty}^{\infty} |f'(x)|dx,$$

to obtain

$$\sum_{\ell \neq 0} \hat{f}(a\ell)^2 \ll \sum_{\ell \neq 0} \left(\frac{\int_{-\infty}^{\infty} |f'(x)|dx}{|a\ell|}\right)^2 \ll_f \frac{1}{a^2}$$

which for $a \gg 1$ is $\ll 1/a$. \hfill \Box

Hence

$$\text{Var } D(N, M) \ll \frac{1}{M} \sum_{v_1, v_2 \neq 0} R(v_1)R(v_2) \frac{\gcd(v_1, v_2)}{\sqrt{|v_1v_2|}}.$$

According to the GCD bounds of [3],

$$\sum_{v_1, v_2 \neq 0} R(v_1)R(v_2) \frac{\gcd(v_1, v_2)}{\sqrt{|v_1v_2|}} \ll \exp\left(\frac{10 \log N}{\log \log N}\right) \sum_v R(v)^2$$

(see [3] for an essentially optimal refinement). Now

$$\sum R(v)^2 = \#\{m_i, n_j \leq N, m_1 \neq n_1, m_2 \neq m_2 : a(m_1) - a(n_1) = a(m_2) - a(n_2)\}$$
is at most the additive energy $E(\mathcal{A}, N)$. Thus

$$\text{Var} D(N, M) \ll \frac{1}{M} N^\epsilon E(\mathcal{A}, N)$$

as claimed. □

**Corollary 5.** Assume that the additive energy satisfies $E(\mathcal{A}, N) < N^{2+\epsilon(1)}$. If $M < N^{2-\eta}$, then for almost all $\alpha$,

$$D_{\mathcal{A}}(N, M)(\alpha) \sim \frac{N^2}{M}.$$  

**Proof.** Take $N_k = \lfloor k^{\frac{4}{\eta}} \rfloor$, so that $\sum_k N_k^{-\eta/2} < \infty$, and pick any $M_k < N^{2-\eta}_k$, then we find from Proposition 3

$$\sum_k \int_0^1 \left| \frac{D(N_k, M_k)(\alpha)}{N_k(N_k-1)/M_k} - 1 \right|^2 d\alpha = \sum_k \frac{\text{Var} D(N_k, M_k)}{(N_k(N_k-1)/M_k)^2} \lesssim \sum_k N_k^{o(1)} \frac{E(N_k)M_k}{N_k^4} \ll \sum_k \frac{1}{N_k^{\eta/2}} < \infty$$

and so for almost all $\alpha$

$$D(N_k, M_k)(\alpha) \sim \frac{N_k^2}{M_k}, \quad \forall k > k_0(\alpha).$$

Apriori the set depends on the test function $f$, but that can be taken care of by a standard diagonalization procedure; for our purposes we only need one test function.

Given $N \gg 1$, there is a unique value of $k$ so that $N_k \leq N < N_{k+1}$. Note that $N/N_k = 1 + O(N^{-\eta/4})$. Since $M < N^{2-\eta} \sim N^{2-\eta}_k < N^{2-\eta}_{k+1}$, we know (7) that almost surely $D(N_k, M)/(N_k^2/M) \to 1$.

Note that

$$D(N_k, M) \leq D(N, M) \leq D(N_{k+1}, M).$$

This is because the sums $D(N, M)$ consist of non-negative terms, and hence

$$D(N, M) = \sum_{1 \leq m \neq n \leq N} F_M(\alpha(a(m) - a(n))) \geq \sum_{1 \leq m \neq n \leq N_k} F_M(\alpha(a(m) - a(n))) = D(N_k, M)$$

(we dropped all pairs $(m, n)$ where $\max(m, n) > N_k$).

Since $N/N_k = 1 + O(N^{-\eta/4})$, we have

$$\frac{D(N_k, M)}{N_k^2/M} \leq \frac{D(N, M)}{N^2/M(1 + O(N^{-\eta/4}))} = \frac{D(N, M)}{N^2/M}(1 + O(N^{-\eta/4}))$$

and likewise

$$\frac{D(N, M)}{N_k^2/M} \leq \frac{D(N_{k+1}, M)}{N_{k+1}^2/M}(1 + O(N^{-\eta/4})).$$

Since we know that almost surely $D(N_k, M)/(N_k^2/M) \to 1$, we deduce that almost surely also $D(N, M)/(N^2/M) \to 1$. □

**Corollary 6.** Theorem 2 holds.
Proof. Fix \( \eta > 0 \), and let \( M = \frac{1}{4} N^{2-\eta} \). Since \( D(N, M)(\alpha) \sim \frac{N^2}{2M} > 1 \) by Corollary 5, we have a gap of size at most \( 1/(2M) = 1/N^{2-\eta} \), that is \( \delta_{\min}^\alpha(N) < 1/N^{2-\eta} \) almost surely. \( \square \)

2. Lower bounds: Proof of Theorem 1

We take any sequence of integers \( A = \{a(n)\} \) with distinct elements. We want to show that for any \( \eta > 0 \), almost surely,

\[ \delta^\alpha(N) > 1/N^{2+\eta}, \quad \forall N > N_0(\alpha). \]

Let \( N_k = [k]^{2/\eta} \). We claim that it suffices to show that for almost all \( \alpha \),

\[ \delta_{\min}^\alpha(N_k) > 2/N^{2+\eta}, \quad \forall k > k_0(\alpha). \]

Indeed, note that if \( N_k \leq N < N_{k+1} \) then \( \delta_{\min}^\alpha(N_k) \geq \delta_{\min}^\alpha(N_{k+1}) \). Since \( N_{k+1} \sim N \), by (8) we have for almost all \( \alpha \)

\[ \delta_{\min}^\alpha(N_k) > 2/N^{2+\eta} > 1/N^{2+\eta} \]

for \( N > N_0(\alpha) \).

To prove (8), it suffices, by the Borel-Cantelli lemma, to show that

\[ \sum_k \text{Prob} \left( \delta_{\min}^\alpha(N_k) \leq 2/N^{2+\eta} \right) < \infty. \]

In the definition of \( D(N, M) = D_f(N, M) \), choose \( f \) so that \( f(x) \geq 1 \) if \( |x| \leq 1/4 \) (and in addition, \( f \geq 0 \) is non-negative, \( \int_{-\infty}^{\infty} f(x) dx = 1 \), \( f \) is smooth and supported in \([-1/2, 1/2]\)). Now note that for such \( f \), if \( D_f(N, M) < 1 \) then \( \delta_{\min}^\alpha(N) > 1/(4M) \). This is because \( D_f(N, M) \) is a sum of non-negative terms, and if there is one gap of size \( \leq 1/(4M) \) then the corresponding term \( F_M(\alpha(a(m) - a(n))) = \sum_j f(M(\alpha(a(m) - a(n)) + j) \geq 1 \) by the choice of \( f \), so that \( D_f(N, M) \geq 1 \). Thus we find that

\[ \delta_{\min}^\alpha(N) \leq \frac{1}{4M} \implies D_f(N, M) \geq 1 \]

and hence

\[ \text{Prob} \left( \delta_{\min}^\alpha(N) \leq \frac{1}{4M} \right) \leq \text{Prob} \left( D_f(N, M) \geq 1 \right). \]

Now since \( D_f \geq 0 \),

\[ \text{Prob} \left( D_f(N, M) \geq 1 \right) \leq \int_0^1 D_f(N, M)(\alpha) d\alpha \]

so that by (2) for \( M_k = \frac{1}{8} N_k^{2+\eta} \),

\[ \int_0^1 D(N_k, M_k)(\alpha) d\alpha \sim 8N_k^{-\eta} \ll \frac{1}{k^2} \]

which together with (10) proves (9), hence (8). This proves Theorem 1.
3. Minimal gaps for the primes

Let \(a(n) = p_n\), the \(n\)-th prime. By Khinchin’s theorem, for all \(\varepsilon > 0\) there is a set of full measure of \(\alpha\)’s so that \(||\alpha| | \geq 1/(q(\log q)^{1+\varepsilon})\) for any integer \(q \geq q_0(\alpha)\).

In particular, for such \(\alpha\), the gap between fractional parts of \(\alpha p \mod 1\) are

\[
||\alpha(p_m - p_n)|| \gg \frac{1}{|p_m - p_n|(|\log |p_m - p_n||)^{1+\varepsilon}} \geq \frac{1}{N(\log N)^{2+\varepsilon}},
\]

since \(|p_m - p_n| \leq p N \sim N \log N\) for \(m < n \leq N\). Hence for such \(\alpha\), the minimal gap satisfies \(\delta_{\alpha_{\text{min}}}(N) > 1/N(\log N)^{2+\varepsilon}\).

A similar argument applies to other dense cases, such as the sequence of square-free integers. An extreme case is that when \(A = N\) is the sequence of all natural numbers. The argument above gives the minimal gap here is, for almost all \(\alpha\), at least \(\delta_{\alpha_{\text{min}}}(N) \gg 1/(N(\log N)^{1+\varepsilon})\). Note that in this case the “three-gap” theorem shows that there are at most three distinct gaps between the fractional parts \(\{\alpha n \mod 1 : n \leq N\}\). Concerning other “dense” sequences, it is known that for any sequence of integers \(A \subset [1, M]\), the fractional parts \(\alpha a(m) \mod 1\) have at most \(O(\sqrt{M})\) distinct gaps [13, 2].

REFERENCES

[1] 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., to appear.
[2] A. Balog, A. Granville and J. Solymosi. Gaps between fractional parts, and additive combinatorics. Quarterly Journal of Mathematics, 68 (2017), 1–12.
[3] V. Blomer, J. Bourgain, M. Radziwiłł and Z. Rudnick. Small gaps in the spectrum of the rectangular billiard. arXiv:1604.02413 [math.AP]. To appear in Annales Scientifiques de l’École Normale Supérieure.
[4] T. F. Bloom, S. Chow, A. Gafni and A. Walker. Additive energy and the metric Poissonian property. arXiv:1709.02654 [math.NT]
[5] A. Bondarenko and K. Seip. GCD sums and complete sets of square-free numbers. Bull. Lond. Math. Soc., 47(1):29–41, 2015.
[6] T. Dyer and G. Harman. Sums involving common divisors, J. London Math. Soc. 34 (1986) 1–11.
[7] D.R. Heath-Brown. Pair correlation for fractional parts of \(\alpha n^2\). Math. Proc. Cambridge Philos. Soc. 140 (2010), no. 3, 385–407.
[8] T. Lachmann and N. Technau. On Exceptional Sets in the Metric Poissonian Pair Correlations Problem. arXiv:1708.08899
[9] P. Lévy. Sur la division d’un segment par des points choisis au hasard, C.R. Acad. Sci. Paris 208 (1939), 147–149.
[10] J. Marklof and A. Strömbergsson. Equidistribution of Kronecker sequences along closed horocycles. Geom. Funct. Anal. 13 (2003), 1239–1280.
[11] Z. Rudnick and P. Sarnak. The pair correlation function of fractional parts of polynomials. Comm. Math. Phys., 194(1):61–70, 1998.
[12] Z. Rudnick and A. Zaharescu. A metric result on the pair correlation of fractional parts of sequences. Acta Arith., 89(3):283–293, 1999.
[13] M. Vâjâitu and A. Zaharescu. Distinct gaps between fractional parts of sequences. Proc. Amer. Math. Soc. 130 (2002), 3447–3452.
[14] A. Walker. The primes are not metric Poissonian. arXiv:1702.07365

RAYMOND AND BEVERLY SACKLER SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, TEL AVIV 69978, ISRAEL
E-mail address: rudnick@post.tau.ac.il