Thomas Lachmann, B.Sc. M.Sc.

Pair Correlations, Additive Combinatorics, and Metric Diophantine Approximation

DOCTORAL THESIS

to achieve the university degree of

Doktor der technischen Wissenschaften

submitted to

Graz University of Technology

Supervisor

Assoc.Prof. Dipl.-Ing. Dr.techn. Christoph Aistleitner

Institute of Analysis and Number Theory

Graz, February 2019

Contents

Acknowledgements

On my way, starting with my bachelor's studies towards finishing this thesis, I have learnt about numerous interesting subjects in mathematics and I have met extraordinary people. Both in terms of tutors and fellow students during my bachelor's and master's studies in Würzburg as well as my dear colleagues here at the TU Graz. Some will be mentioned multiple times in the following lines for the different roles they have played in this phase of my life.

My first gratitude is directed to my supervisor Christoph Aistleitner for giving me the opportunity to work with him and opening the possibility to write this thesis, ignoring anybody's fear I could die to the Austrian winter. Throughout the last three years he has taught me a considerable amount of new techniques and views on various topics as well as sharing a lot of ideas with me. He was able to make my whole time in Graz a wonderful, enjoyable and humorous experience whilst encouraging me to pursue my own ideas and goals, too. He was always ready to let me present my own thoughts and to give feedback, may it have been on mathematical thinking or on my cooking skills.

Furthermore, I want to thank a professor from Würzburg I view as my first mentor in the mathematical way of thinking, Jörn Steuding. He gave several excellent lectures centred on topics of number theory, which led me to developing a major interest in these subjects. Concluding in him supervising my master thesis about a problem in the branch of algebraic number theory. I also want to thank him for encouraging me to go to Graz where I was able to develop myself even further.

During my time in Graz we have established a small seminar consisting of several colleagues working or having worked in Graz at some point, the "Fufu-Seminar". The name originates from the typical dish we ate at the end of every successful seminar, whereas "successful" is to be understood subjectively. In this seminar we have worked through and discussed the results of a large diversity of different topics in number theory and were inspired to further advance some ideas. An example for that is the last chapter of this thesis. Hereby, I want to thank all the other current members of the Fufu-Seminar, namely, "Archduke" Christoph Aistleitner, "Duke" Marc Alexandre Munsch, "Marquis" Antoine André Marnat, "Earl" Agamemnon Zafeiropoulos and "Sir" Niclas Alexander Rüdiger Technau.

Next, I want to say how grateful I am for my family, my mother Sieglinde, and my siblings Marco, Manuela and Christian, for supporting me all this time and motivating me to always move on no matter what happened in life. Additionally to that I especially want to thank my girlfriend, Lisa Mündlein, for the ten best years of my life and for enduring a distance relationship over the last three years while I was studying and writing this thesis.

Extra thanks go to Chacha (Ade Irma Suriajaya), an Indonesian mathematician, currently working in Japan, for always motivating me with her energetic spirit and inspiring me to further pursue my goal of learning the Japanese language.

I also want to thank the two women without whom everyday life could never run so smoothly at our institute, Irene Pfeifer-Wilfinger and Hermine Panzenböck. Both of them are always there to answer all those terrible questions about bureaucracy, help for every form that has to be filled and look over all the different events that come up at a university. Another important person who always makes sure that everything runs smoothly, especially all the office computers, is Jochen Resch. He is a real jack of all trades and able to support a multitude of institutes with his wide range of talents.

Lastly, I want to thank my closest fellow students and colleagues whereas the chance is high that I forget some and hope nobody takes offence for not being mentioned. So, I thank my friends I got to know in Würzburg: Philipp Muth, Lena Ries, David Pfrang, Marc Technau, Niclas Technau, Torsten Ballin, Lena Lorang, Fabian Hoppe, Simon Schnürch, Sophia Granow, who all studied alongside me and worked together through all of our studies. Fellow PhD students from Graz: Niclas Technau, Damir Ferizović, Gigi (Kwok Chi Chim), Besmira Dervishi, Nicola Del Giudice; here I want to especially mention Stefan Lendl and Gundelinde (Gundi) Wiegel for not only being awesome friends and colleagues but also giving me the opportunity to fruitfully work together with them on interesting topics some steps away from number theory and kindle a new passion for certain branches of mathematics I did not work on before.

Chapter 1

Preface

All the results of this thesis are primarily based in the mathematical field of Metric Number Theory. This field of mathematics builds a connection between analytical number theory and probabilistic methods. The main interest is to metrically characterise certain sets of numbers. In other words, one looks for the size of the set of numbers which fulfil a certain condition, whereas this size is usually considered with respect to the Lebesgue measure. For the whole thesis the Lebesgue measure will be denoted by λ. Deep results of Metric Number Theory are in the areas of Diophantine approximation, the theory of uniform distribution of sequences, as well as continued fractions. These topics together with a so far less known notion — the pair correlations of sequences — are unified in this thesis. While techniques of all the aforementioned areas play a crucial role, the main part of this work is centred around pair correlations. We will start with basic definitions and some results which will be helpful later on. Moreover we summarise the most important results of each chapter. For more precise introductions we refer to the single chapters.

1.1 Equidistribution and Pair Correlations

The consideration of sequences of real numbers in terms of the distribution of their fractional parts can be said to have its genesis as fully acknowledged branch of mathematics with the seminal paper *Über die Gleichverteilung von Zahlen mod. Eins* by Hermann Weyl in 1916 [46].

Let $(x_n)_{n\geq 1}$ be a sequence of real numbers and $\langle x \rangle$ denotes the fractional part of x. If for all $a, b \in [0,1]$, $a \leq b$, it holds that

$$\lim_{N\to\infty} \frac{1}{N} \#\Big\{1 \leq n \leq N : \ \langle x_n \rangle \in [a,b]\Big\} = b - a,$$

the sequence is said to be *equidistributed* or *uniformly distributed modulo one* ("Gleichverteilt mod. Eins").

As we will see later this notion is closely related to a sequence having *Poissonian pair correlations* (PPC for short). The *pair correlation function* of the aforementioned sequence $(x_n)_{n\geq 1}$ up to the N-th member is defined by

$$R_2(s, (x_n)_n, N) = \frac{1}{N} \# \left\{1 \leq m, n \leq N, \ m \neq n : \ \|x_m - x_n\| \leq \frac{s}{N}\right\}$$

whereas we shorten this notation to

$$R(s,N) := R_2(s,(x_n)_n, N)$$

if it is clear which sequence is considered. The sequence is said to have *Poissonian pair correlations* if

$$R(s) := \lim_{N\to\infty} R(s,N) = \lim_{N\to\infty} R_2(s,(x_n)_n, N)$$

exists for every $s \geq 0$ and it holds that

$$R(s) = 2s.$$

Then, it can also be said that the distribution of its pair correlation statistics is asymptotically Poissonian.

The concept of equidistribution will only appear in Chapter 2 where we show that the property of having Poissonian pair correlations is stronger than being equidistributed:

Theorem 2.1. *Let $(x_n)_{n\geq 1}$ be a sequence of real numbers in $[0,1]$, and assume that the distribution of its pair correlation statistic is asymptotically Poissonian. Then the sequence is equidistributed.*

1.2 Additive Energy

In Chapter 3 and Chapter 4 we will discuss, given a sequence $(\alpha a_n)_n$, with $\alpha \in [0,1]$ and $a_n \in \mathbb{N}$ strictly increasing, the dependence of having Poissonian pair correlations on the *additive energy* of the sequence $(a_n)_n$. We will also be interested in the set of α such that $(\alpha a_n)_n$ has no PPC and call this set $\mathrm{NPPC}((a_n)_n)$. If $\lambda(\mathrm{NPPC}((a_n)_n)) = 0$, we say that $(a_n)_n$ has *metric PPC.*

For a finite set I the *additive energy* $E(I)$ is defined as

$$E(I) := \#\left\{(a,b,c,d) \in I^4 \colon a+b=c+d\right\} = \sum_{\substack{a+b=c+d\\ a,b,c,d\in I}} 1.$$

Trivial bounds for the additive energy of I are given by

$$|I|^2 \leq E(I) \leq |I|^3.$$

Let $A_N := \{a_n\}_{1\leq n\leq N}$ be the set of the first $N \in \mathbb{N}$ members of the sequence $(a_n)_n$, then the additive energy of A_N can be written as

$$E(A_N) = \#\{n_1,n_2,n_3,n_4 \leq N \colon a_{n_1} + a_{n_2} = a_{n_3} + a_{n_4}\}.$$

In Chapter 3 we show the following two results.

Theorem 3.1. *If $E(A_N) = \Omega\left(N^3\right)$, then $\mathrm{NPPC}((a_n)_n)$ has full Lebesgue measure.*

Theorem 3.2. *Let $f : \mathbb{R}_{>0} \to \mathbb{R}_{>2}$ be a function increasing monotonically to ∞, and satisfying $f(x) = \mathcal{O}\big(x^{1/3} (\log x)^{-7/3}\big)$. Then, there is a strictly increasing sequence $(a_n)_n$ of positive integers with $E(A_N) = \Theta\big(N^3/f(N)\big)$ such that if*

$$\sum_{n \geq 1} \frac{1}{nf(n)}$$

diverges, then for Lebesgue almost all $\alpha \in [0,1]$

$$\limsup_{N \to \infty} R(s, \alpha, N) = \infty$$

holds for any $s > 0$.

As one will see in the chapter, the second theorem also yields information about the Hausdorff dimension of NPPC $((a_n)_n)$, but we do not write it here in its full form since it is not part of the original motivation to consider this problem.

Investigation of such problems is motivated by questions asked after a result of Aistleitner, Larcher, and Lewko in [4]. They show:

Theorem (Aistleitner, Larcher, Lewko, [4])**.** *If for a sequence $(a_n)_n$ we have*

$$E(A_n) = \mathcal{O}\left(N^{3-\varepsilon}\right),$$

for some $\varepsilon > 0$, then $(a_n)_n$ has metric PPC.

The natural questions that arose after this result are the following.

Question 1. *Is it possible for $(a_n)_n$ with $E(A_N) = \Omega\left(N^3\right)$ to have metric PPC?*

Question 2. *Do all $(a_n)_n$ with $E(A_N) = o\left(N^3\right)$ have metric PPC?*

Both these questions are answered in the negative by Jean Bourgain whose results are in [4] as well. The aforementioned theorems 3.1 and 3.2 are improving his answers in a quantitative way. A certain time after Theorem 3.1 was originally proved, Larcher was able to show, under the same assumption, that NPPC $((a_n)_n)$ is the whole interval, i.e. there is **no** α such that $(\alpha a_n)_n$ has PPC (see [29]).

Theorem 3.2 has been proven by constructing a sequence built up of arithmetic and geometric progressions while making sure that the desired properties stay intact. Sticking with the same idea of combining arithmetic and geometric progression but modifying their lengths and the used moduli of the arithmetic progressions, we are able to prove:

Theorem 4.1. *For every $\varepsilon > 0$ there exists a strictly increasing sequence $(a_n)_n$ of positive integers which has the metric pair correlation property, and whose additive energy satisfies*

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

Together with Theorem 3.2 this gives a negative answer to a fundamental question posed by Bloom, Chow, Gafni, and Walker in [9]. They ask if there is a "switching" threshold. This means in terms of additive energy, for a sequence to always having PPC below and never having it above that threshold. In this range the conjectured threshold would have been essentially the one given by Theorem 3.2. These results show that there is a sort of continuum where it cannot only be based on the order of the additive energy.

1.3 Diophantine Approximation and Continued Fractions

Diophantine approximation was first studied in the 18th century. One is interested in how good any real number could be approximated by rationals in a certain sense. Dirichlet's well-known result about upper bounds for these approximation was one of the starting points for a later popular field of studies in a lot of different branches of math, asking questions of a similar type, such as in analytic and algebraic number theory, fractal geometry, ergodic theory as well as having practical applications.

Dirichlet's approximation theorem. *For any $\alpha \in [0,1] \setminus \mathbb{Q}$ and $N \in \mathbb{N}$ there exists a pair $(p,q) \in \mathbb{Z}^2$ with $1 \leq q \leq N$ such that*

$$|q\alpha - p| < \frac{1}{N}.$$

Dirichlet proved this as an easy consequence of the pigeonhole principle which also became quite popular thanks to this result.

A result that follows immediately by Dirichlet's theorem is that there are infinitely many coprime $p, q \in \mathbb{N}$ such that

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}.$$

About a century later, it was shown in a breakthrough paper that the 2 in the exponent is optimal for algebraic α and cannot be bigger if one wants to find infinitely many solutions. The corresponding theorem is called Roth's theorem, sometimes also Thue-Siegel-Roth theorem.

In the first half of the 20th century Khintchine proved a generalization of Dirichlet's approximation theorem which shows a certain similarity to Theorem 3.2.

Khintchine's theorem. *Let $\psi\colon \mathbb{N} \to \mathbb{R}$ be a non-negative, non-increasing function. Then for almost all $\alpha \in [0,1]$ there are infinitely many $p, q \in \mathbb{N}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \frac{\psi(q)}{q}$$

if and only if

$$\sum_{n \geq 1} \psi(n)$$

diverges.

Choosing $\psi(q) = \frac{1}{q}$ we obtain Dirichlet's approximation theorem again.

It is noteworthy to mention that it is important that ψ is non-increasing and not just converging to 0 as $n \to \infty$. Otherwise there is a counterexample, which can be found in [21]. There were a lot of attempts to loosen the condition of ψ being non-increasing. These attempts led to a very famous conjecture. Let φ be Euler's totient function.

Duffin–Schaeffer conjecture. *Let $\psi\colon \mathbb{N} \to \mathbb{R}$ be a non-negative function. Then for almost all $\alpha \in [0,1]$ there are infinitely many coprime $p, q \in \mathbb{N}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \frac{\psi(q)}{q}$$

if and only if

$$\sum_{n\geq 1} \psi(n) \frac{\varphi(n)}{n}$$

diverges.

This problem remains unsolved until today. The main result of Chapter 5 is a slightly weaker version of the Duffin–Schaeffer conjecture.

Theorem 5.1. *Let $\psi\colon \mathbb{N} \to \mathbb{R}$ non-negative function. Then for almost all $\alpha \in [0,1]$ there are infinitely many coprime $p, q \in \mathbb{N}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \frac{\psi(q)}{q}$$

if there is some $\varepsilon > 0$ such that

$$\sum_{n\geq 1} \psi(n) \frac{\varphi(n)}{n(\log n)^{\varepsilon}}$$

diverges.

For proving this theorem a result due to Gallagher is utterly important:

Gallagher's zero-one law. *Let $\psi\colon \mathbb{N} \to \mathbb{R}$ non-negative function. Then for either almost all $\alpha \in [0,1]$ or almost no $\alpha \in [0,1]$ there are infinitely many coprime $p, q \in \mathbb{N}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \psi(q).$$

On the other hand it is well-known that good approximations of any real number α can be obtained by studying the convergents in its continued fraction expansion. Every $\alpha \in [0,1]$ can be written as a continued fraction, namely

$$\alpha = \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_2 + \cfrac{1}{\ddots}}},$$

with $(\alpha_n)_n \subseteq \mathbb{N}$, whereas $(\alpha_n)_n$ can be either finite or infinite. This fraction is finite if and only if $\alpha \in \mathbb{Q}$. Usually, the continued fraction for $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ is denoted by

$$\alpha = [\alpha_1, \alpha_2, \ldots]$$

and for $\alpha \in \mathbb{Q}$ by

$$\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_n].$$

For $\alpha = [\alpha_1, \alpha_2, \ldots]$ and $n \in \mathbb{N}$ we call $[\alpha_1, \alpha_2, \ldots, \alpha_n] = \frac{p_n}{q_n}$ the n-th convergent with $p_n, q_n \in \mathbb{N}$. These convergents are good approximations to α. The following results on continued fractions will be important in Chapter 4.

Legendre's theorem. *If a/b is a fraction with*

$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{2b^2},$$

then a/b is a convergent to α.

Borel–Bernstein theorem. *Let $B \coloneqq (b_n)_n$ be a sequence of (strictly) positive real numbers, and consider the series*

$$\sum_{n\geq 1} \frac{1}{b_n}. \tag{$\star$}$$

If $V_B \subset [0,1]$ denotes the set of numbers $\alpha = [\alpha_1, \alpha_2, \ldots]$ satisfying $\alpha_n \leq b_n$ for all sufficiently large $n \geq 1$, then

$$\lambda\left(V_B\right) = \begin{cases} 1 & \text{if } (\star) \text{ converges,} \\ 0 & \text{if } (\star) \text{ diverges.} \end{cases}$$

1.4 Results from Probability Theory and Number Theory

An often used result in probability theory is the Borel–Cantelli lemma. We want to state it in a version in terms of Lebesgue measurable sets.

Borel–Cantelli lemma. *Let $(A_n)_n$ be a sequence of Lebesgue measurable sets in $[0,1]$ satisfying*

$$\sum_{n\geq 1} \lambda\left(A_n\right) = \infty.$$

Then

$$\lambda\left(\limsup_{n\to\infty} A_n\right) = 1.$$

We will need a version of this lemma due to Erdős, and Rényi which will play a crucial role in proofs of Chapter 3 and Chapter 5.

Lemma 3.4, Lemma 5.1 (Erdős–Rényi, [21])**.** *Let $(A_n)_n$ be a sequence of Lebesgue measurable sets in $[0,1]$ satisfying*

$$\sum_{n\geq 1} \lambda\left(A_n\right) = \infty.$$

Then,

$$\lambda\left(\limsup_{n\to\infty} A_n\right) \geq \limsup_{N\to\infty} \frac{\left(\sum_{n\leq N} \lambda\left(A_n\right)\right)^2}{\sum_{m,n\leq N} \lambda\left(A_n \cap A_m\right)}.$$

Furthermore we need one of Mertens' theorems giving information on the distribution of primes. More precisely, Mertens' second theorem will be applied in Chapter 5 several times.

Mertens' second theorem. *It holds*

$$\lim_{n\to\infty}\left(\sum_{p\le n}\frac{1}{p}-\log\log n-M\right)=0,$$

where $p \le n$ means all primes smaller than n and M is the Meissel-Mertens constant*. In particular it holds that*

$$\sum_{p\le n}\frac{1}{p}\ll\log\log n.$$

The above list of results used is not exhaustive and mentions only the most important results used in the field of metric number theory and this work. For example in Chapter 2 some knowledge about quadratic forms, basic linear algebra, the Dirchlet and Fejér Kernel as well as in Chapter 3 and Chapter 5 some knowledge about GCD sums and Fourier analysis will be used.

To complete the preface it follows a summary of all the original papers containing the results presented.

- Chapter 2: Published in *Journal of Number Theory* ([3]), joint work with Christoph Aistleitner, and Florian Pausinger.
- Chapter 3: Published in *Monatshefte für Mathematik* ([28]), joint work with Niclas Technau.
- Chapter 4: Submitted, joint work with Christoph Aistleitner, and Niclas Technau. Available at arXiv:1802.02659.
- Chapter 5: Submitted, joint work with Christoph Aistleitner, Marc Munsch, Niclas Technau, and Agamemnon Zafeiropoulos. Available at arXiv:1803.05703.

Chapter 2

Pair Correlations and Equidistribution

Let $(x_n)_n$ be a sequence of real numbers. We say that this sequence is *equidistributed* or *uniformly distributed modulo one* if asymptotically the relative number of fractional parts of elements of the sequence falling into a certain subinterval is proportional to the length of this subinterval. More precisely, we require that

$$\lim_{N\to\infty} \frac{1}{N} \#\Big\{1 \le n \le N: \ \langle x_n \rangle \in [a,b]\Big\} = b - a$$

for all $0 \le a \le b \le 1$, where $\langle \cdot \rangle$ denotes the fractional part. This notion was introduced in the early twentieth century, and received widespread attention after the publication of Hermann Weyl's seminal paper *Über die Gleichverteilung von Zahlen mod. Eins* in 1916 [46]. Among the most prominent results in the field are the facts that the sequences $(n\alpha)_n$ and $(n^2\alpha)_n$ are equidistributed whenever $\alpha \notin \mathbb{Q}$, and the fact that for any distinct integers $n_1, n_2, \ldots$ the sequence $(n_k\alpha)_{k\ge 1}$ is equidistributed for almost all α. All of these results were already known to Weyl, and can be established relatively easily using the famous *Weyl criterion*, which links equidistribution theory with the theory of exponential sums. For more background on uniform distribution theory, see the monographs [15, 25]. We note that when $(X_n)_n$ is a sequence of independent, identically distributed (i.i.d.) random variables having uniform distribution on $[0,1]$, then by the strong law of large numbers this sequence is almost surely equidistributed. Consequently, in a very vague sense equidistribution can be seen as an indication of "pseudorandom" behaviour of a deterministic sequence.

The investigation of pair correlations can also be traced back to the beginning of the twentieth century, when such quantities appeared in the context of statistical mechanics. In our setting, when $(x_n)_n$ are real numbers in the unit interval, we define a function $R: \mathbb{R}_{\ge 0} \times \mathbb{N} \mapsto \mathbb{R}_{\ge 0} \cup \{\infty\}$ by

$$R(s,N) = \frac{1}{N} \#\left\{1 \le m,n \le N, \ m \ne n: \ \|x_m - x_n\| \le \frac{s}{N}\right\}, \tag{2.1}$$

and set

$$R(s) = \lim_{N\to\infty} R(s,N),$$

provided that such a limit exists; here $s \geq 0$ is a real number, and $\|\cdot\|$ denotes the distance to the nearest integer. The function $R(s, N)$ counts the number of pairs (x_m, x_n), $1 \leq m, n \leq N$, $m \neq n$, of points which are within distance at most s/N of each other (in the sense of the distance on the torus). If $R(s) = 2s$ for all $s \geq 0$, then we say that the asymptotic distribution of the pair correlations of the sequence is *Poissonian*. Again, one can show that an i.i.d. random sequence (sampled from the uniform distribution on $[0, 1]$) has this property, almost surely. Questions concerning the distribution of pair correlations of sequences such as $(\langle n\alpha \rangle)_n$ or $(\langle n^2 \alpha \rangle)_n$ are linked with statistical properties of the distribution of the energy spectra of certain quantum systems, and thus play a role for the Berry–Tabor conjecture [8]. See [32, 37] for more information on this connection. It turns out that unlike the situation in the case of equidistribution, there is *no* value of α for which the sequence $(\langle n\alpha \rangle)_n$ has Poissonian pair correlations, and the question whether the pair correlations of the sequence $(\langle n^2 \alpha \rangle)_n$ are Poissonian or not depends on delicate number-theoretic properties of α, in particular on properties concerning Diophantine approximation and the continued fraction expansion of α. Here many problems are still open, see [24, 38]. Furthermore, for $(n_k)_{k \geq 1}$ being distinct integers the question whether $(\langle n_k \alpha \rangle)_{k \geq 1}$ has Poissonian pair correlations for almost all α or not depends on certain number-theoretic properties of $(n_k)_{k \geq 1}$, in particular on the number of possible ways to represent integers as a difference of elements of this sequence; see [4].

It is remarkable that (to the best of our knowledge) the relation between these two notions (being equidistributed, and having Poissonian pair correlations) has never been clarified, a fact which came to our attention by a question asked by Arias de Reyna [5] in a slightly different, but related context (we will repeat this question at the end of the present section). As a starting observation, we note that in a probabilistic sense Poissonian pair correlations actually require uniform distribution. More precisely, assume that $(X_n)_n$ are i.i.d. random variables, which for simplicity we assume to have a density g on $[0, 1]$. Then we have

$$\begin{aligned}
& \mathbb{E}\left(\frac{1}{N} \# \left\{ 1 \leq m, n \leq N,\ m \neq n :\ \|X_m - X_n\| \leq \frac{s}{N} \right\} \right) \\
\approx \ & \frac{1}{N} N^2 \int_0^1 g(x) \underbrace{\int_{x - \frac{s}{N}}^{x + \frac{s}{N}} g(y)\ dy}_{\approx \frac{2s}{N} g(x)}\ dx \\
\approx \ & 2s \int_0^1 g(x)^2\ dx,
\end{aligned}$$

which can be turned into a rigorous argument to show that almost surely

$$R(s) = 2s \int_0^1 g(x)^2\ dx, \qquad s > 0, \tag{2.2}$$

in this case (and $R(s) = \infty$ for all $s > 0$, almost surely, in the case when the distribution of the X_n is not absolutely continuous with respect to the Lebesgue measure). Now we clearly have $\int_0^1 g(x)^2\ dx = 1$ if and only if $g \equiv 1$, which means that g is the density of the uniform distribution. Thus Poissonian pair correlations require uniform distribution in a probabilistic sense; however, it is a priori by no means clear that a similar relation also holds for the case of deterministic sequences $(x_n)_n$. Our Theorem 2.1 below shows that this actually is the case.

Theorem 2.1. *Let $(x_n)_n$ be a sequence of real numbers in $[0,1]$, and assume that the distribution of its pair correlation statistic is asymptotically Poissonian. Then the sequence is equidistributed.*

There also is a quantitative "density" version of the theorem, which has a resemblance of (2.2), and which we state as Theorem 2.2 below.

Theorem 2.2. *Let $(x_n)_n$ be a sequence of real numbers in $[0,1]$. Assume that it has an asymptotic distribution function G on $[0,1]$, i.e., that there is a function G such that*

$$G(x) = \lim_{N\to\infty} \frac{1}{N} \#\Big\{1 \le n \le N: \ x_n \in [0,x]\Big\}, \qquad x \in [0,1].$$

Assume also that there is a function $F: \ [0,\infty) \mapsto [0,\infty]$ such that

$$R(s) = \lim_{N\to\infty} \frac{1}{N} \#\left\{1 \le m,n \le N, \ m \neq n: \ \|x_m - x_n\| \le \frac{s}{N}\right\}, \qquad s > 0.$$

Then the following hold.

- *If G is not absolutely continuous, then $R(s) = \infty$ for all $s > 0$.*
- *If G is absolutely continuous, then, writing g for the density function of the corresponding measure, we have*
$$\limsup_{s\to\infty} \frac{R(s)}{2s} \ge \int_0^1 g(x)^2 \ dx. \tag{2.3}$$

We believe that Theorem 2.1 is quite remarkable; actually, we initially set out to prove its opposite, namely that a sequence which has Poissonian pair correlations does not have to be equidistributed. This seemed natural to us since equidistribution is controlled by the "large-scale" behaviour, while pair correlations are determined by "fine-scale" behaviour. Only after some time we realised why it is not possible to construct a non-equidistributed sequence which has Poissonian pair correlations; roughly speaking, the reason is that regions where too many points are situated contribute to the pair correlation function proportional to the square of the local density, and regions with fewer elements cannot compensate this larger contribution — this is exactly what (2.2) and (2.3) also tell us.

As noted above, there is a characterisation of equidistribution in terms of exponential sums called Weyl's criterion. In a similar way one could characterise the asymptotic pair correlation function by exponential sums, and then assuming Poissonian pair correlations one could try to control the exponential sums in Weyl's criterion. However, we have not been able to do this; the problem is of course that the pair correlations are determined by "fine" properties at the scale of $1/N$, while equidistribution is a "global" property on full scale — in other words, the trigonometric functions which determine the distribution of the pair correlations have frequencies of order N, while equidistribution is determined by trigonometric functions with constant frequencies. Instead of following such an approach, our proof of Theorem 2.1 is based on estimating the pair correlation function by a certain quadratic form, which is attached to a so-called *circulant matrix*. We can calculate the eigenvalues and eigenvectors of this matrix, and after averaging over different values of s reduce the problem to the fact that the Fejér kernel is a non-negative kernel.

Finally we return to the question of Arias de Reyna, which was mentioned above. Elliott and Hlawka independently proved that the imaginary parts $(\gamma_n)_n$ of the non-trivial zeros of the Riemann zeta function are equidistributed. However, the proof of this result is simplified by the fact that the zeros of the zeta function are relatively dense; more precisely, the number of zeros up to height T is roughly $\frac{T}{2\pi} \log\left(\frac{T}{2\pi}\right) - \frac{T}{2\pi}$. Thus to get a statement about the pseudorandomness of these zeros it is more interesting to consider the sequence of imaginary parts of zeros after normalising them to have average distance 1; that is, instead of investigating the equidistribution of the sequence $(\gamma_n)_n$ itself one asks for the equidistribution of the normalised sequence $(x_n)_n = \left(\frac{\gamma_n}{2\pi} \log\left(\frac{\gamma_n}{2\pi e}\right)\right)_n$. This seems to be a very difficult problem; see [5] for the current state of research in this direction. On the other hand, the famous *Montgomery pair correlations conjecture* predicts a certain asymptotic distribution R for the pair correlations between elements of this normalised sequence $(x_n)_n$. For the statement of this conjecture see [34]; we only mention that the distribution is not the same as in the case of a random sequence, but coincides with a distribution that also appears as the correlation function of eigenvalues of random Hermitian matrices and shows a certain "repulsion" phenomenon. Arias de Reyna asked whether Montgomery's pair correlation conjecture is compatible with equidistribution of the normalised zeros.

Note that the setting of this question is different from our setting; while in our setting the whole sequence is contained in $[0,1]$ and the average spacing of the first N points is $1/N$, in the setting of Arias de Reyna's question the equidistribution property is requested for the reduction of the sequence $(x_n)_n$ modulo one, while the pair correlations are calculated for the increasing sequence $(x_n)_n$ itself, for which the average spacing between consecutive elements is 1 in the limit. Thus the results from the present paper cannot be applied to this setting. A general form of Arias de Reyna's question is: Let (x_n) be an increasing sequence with average spacing 1, that is, $x_n/n \to 1$. Assume that $(x_n)_n$ asymptotically has the pair correlation distribution R from Montgomery's conjecture. Is it possible that $(x_n)_n$ is equidistributed? Is it possible that $(x_n)_n$ is *not* equidistributed? It is known that there exists a random process whose pair correlation function is R (see [26, 27]), and one should be able to show that such a random process (or a further randomization of it) is equidistributed almost surely. So the answer to the first question is "yes". The answer to the second question should be "yes" as well, but we have not been able to construct an example.

Remark added during the revision stage: Simultaneously and independently, Grepstad and Larcher [20] also gave a proof of the result stated as Theorem 2.1 in this paper. They used a more elementary argument, and also obtained quantitative results which relate the speed of convergence of the distribution of pair correlations to the limit function, and the speed of convergence towards equidistribution of the empirical distribution of the point set (the so-called *discrepancy*). Shortly after, Steinerberger [41] provided a more general framework for these results, and established a class of tests which allow to deduce equidistribution from the convergence of normalised sums $\sum_{m,n} f(x_m - x_n)$, where f are for example Jacobi θ-functions or Gaussians. His results also apply to sequences on general compact manifolds, where the role of the test functions f is played by the corresponding heat kernel.

2.1 Preliminaries

Throughout this section, we will use the following notation. Assume that $x_1, \dots, x_N$ are given. Let $R(s,N)$ be defined as in (2.1). We partition the unit interval $[0,1)$ into subintervals $I_1, \dots, I_M$, where $I_m = [m/M, (m+1)/M)$, and we set

$$y_m = \#\Big\{1 \le n \le N : \ x_n \in I_m\Big\}.$$

Then trivially we have

$$\sum_{m=1}^{M} y_m = N.$$

For notational convenience, we assume that the sequence $(y_m)_{1 \le m \le M}$ and the partition $I_1, \dots, I_M$ are extended periodically; in other words, we set

$$y_m = y_{(m \bmod M)}, \qquad \text{and} \qquad I_m = I_{(m \bmod M)}, \qquad m \in \mathbb{Z}.$$

Let $s \ge 1$ be an integer. We set

$$H_{N,M}(s) = \sum_{m=1}^{M} \sum_{-s+1 \le \ell \le s-1} y_m y_{m+\ell}.$$

Then by construction we have

$$\begin{aligned} H_{N,M}(s) &= \sum_{m=1}^{M} \sum_{n \in \{1,\dots,N\}:\ x_n \in I_m} \#\left\{1 \le k \le N : \ x_k \in \bigcup_{\ell=-s+1}^{s-1} I_{m+\ell}\right\} \\ &\le \sum_{n=1}^{N} \#\left\{1 \le k \le N : \ \|x_k - x_n\| \le \frac{s}{M}\right\} \\ &= \left(\sum_{n=1}^{N} \#\left\{1 \le k \le N,\ n \ne k : \ \|x_k - x_n\| \le \frac{s}{M}\right\}\right) + N \\ &= N R(s,N) \left(\frac{sN}{M}\right) + N. \end{aligned} \tag{2.4}$$

Thus a lower bound for $H_{N,M}$ implies a lower bound for $R(s,N)$.

We have the following lemma.

Lemma 2.1. *Let $y_1, \dots, y_M$ be non-negative real numbers whose sum is N, assume that $(y_m)_{1 \le m \le M}$ is extended periodically as above, and let $H_{N,M}$ be defined as above. Let $S \ge 1$ be an integer for which $2S < M$. Then*

$$\frac{1}{S} \sum_{s=1}^{S} H_{N,M}(s) \ge \frac{SN^2}{M}.$$

Proof. The sum

$$\sum_{m=1}^{M} \sum_{-s+1 \le \ell \le s-1} y_m y_{m+\ell}$$

in the definition of $H_{N,M}$ is a quadratic form which is attached to the matrix

$$A^{(s)} = \left(a_{ij}^{(s)}\right)_{1 \le i,j \le M} = \begin{cases} 1 & \text{if } \operatorname{dist}(i-j) \le s-1, \\ 0 & \text{otherwise,} \end{cases}$$

where dist is the periodic distance such that $\operatorname{dist}(i-j) \le s-1$ whenever

$$i - j \in (-\infty, -M+s-1] \cup [-s+1, s-1] \cup [M-s+1, \infty).$$

Thus $A^{(s)}$ is a band matrix which also has non-zero entries in its right upper and left lower corner. This matrix $A^{(s)}$ is symmetric, and it is of a form which is called *circulant*. Generally, a circulant matrix is a matrix of the form

$$\begin{pmatrix} c_0 & c_1 & c_2 & \dots & c_{M-1} \\ c_{M-1} & c_0 & c_1 & \dots & c_{M-2} \\ c_{M-2} & c_{M-1} & c_0 & \ddots & c_{M-3} \\ \vdots & & \ddots & \ddots & \vdots \\ c_1 & & \dots & c_{M-1} & c_0 \end{pmatrix},$$

where each row is obtained by a cyclic shift of the previous row. We recall some properties of such matrices; for a reference see for example [19, Chapter 3]. The eigenvectors of such a matrix are

$$v_m = \left(1, \omega^m, \omega^{2m}, \dots, \omega^{(M-1)m}\right), \qquad m = 0, \dots, M-1, \tag{2.5}$$

where $\omega = e^{\frac{2\pi i}{M}}$. Note that these eigenvectors are pairwise orthogonal, and that they are independent of the coefficients of the matrix (they just depend on the fact that the matrix is circulant). The eigenvalue λ_m to the eigenvector v_m is given by

$$\lambda_m = \sum_{\ell=0}^{M-1} c_\ell \omega^m. \tag{2.6}$$

We have already noted that our matrix $A^{(s)}$ is symmetric, which implies that all its eigenvalues are real. Furthermore, if we use the formula (2.6) to calculate the eigenvalues of $A^{(s)}$, we obtain

$$\lambda_m^{(s)} = \sum_{\ell=-s+1}^{s-1} \omega^m = \frac{\sin\left(\frac{(2s-1)\pi m}{M}\right)}{\sin\left(\frac{\pi m}{M}\right)}, \qquad m \ne 0,$$

which is the $(s-1)$-th order Dirichlet kernel D_{s-1} (with period 1 rather than the more common period 2π), evaluated at position m/M. Note that the largest eigenvalue is $\lambda_0^{(s)} = 2s-1$.

Since the eigenvectors of $A^{(s)}$ form an orthogonal basis, we can express our vector $(y_1, \dots, y_M)$ in this basis. We write

$$(y_1, \dots, y_M) = \sum_{m=0}^{M-1} \varepsilon_m v_m$$

for appropriate coefficients $(\varepsilon_m)_{1 \le m \le M}$. Note that we have $y_1 + \dots + y_M = N$, which can be rewritten as $(y_1, \dots, y_M) v_0 = N$; thus we must have $\varepsilon_0 = N/M$ (since the eigenvectors are orthogonal). Furthermore, we have

$$\begin{aligned} H_{N,M}(s) &= \left(\sum_{m=0}^{M-1} \varepsilon_m v_m \right)^T A^{(s)} \left(\sum_{m=0}^{M-1} \varepsilon_m v_m \right) \\ &= \sum_{m=0}^{M-1} \lambda_m^{(s)} \varepsilon_m^2 \|v_m\|_2^2 \\ &= M \sum_{m=0}^{M-1} \lambda_m^{(s)} \varepsilon_m^2, \end{aligned} \tag{2.7}$$

again by orthogonality. However, from this we cannot deduce that $H_{N,M}(s) \ge M \lambda_0^{(s)} \varepsilon_0^2 = (2s-1)N^2/M$, since (in general) some of the eigenvalues are negative.[1] To solve this problem we will make a transition from the Dirichlet kernel to the Fejér kernel, which is non-negative.

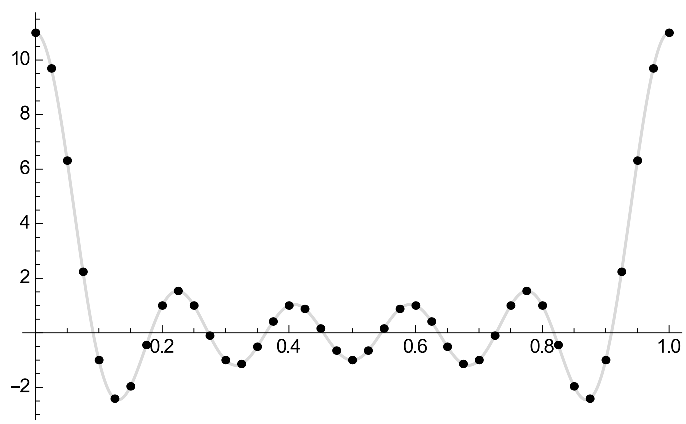

Figure 2.1: Fifth order Dirichlet kernel D_5 (gray line), evaluated at positions $m/40$, $0 \le m \le 39$ (black dots). The m-th eigenvalue $\lambda_m^{(6)}$ of $A^{(6)}$ equals $D_5(m/40)$; note that some of the eigenvalues are negative.

We repeat that the eigenvectors of $A^{(s)}$ depend on M, but not on s. Let $S \ge 1$ be an integer and consider

$$A^{(\Sigma)} = \frac{1}{S} \sum_{s=1}^{S} A^{(s)},$$

[1] However, if $s = 1$ then the matrix $A^{(s)}$ is the unit matrix, all eigenvalues are 1, and we have $H_{N,M}(s) \ge N^2/M$; this fact will be used in the proof of Theorem 2.2.

where we assume that $2S < M$ (to retain the structure of the matrix). Then clearly the eigenvectors of this matrix are also given by $v_0, \dots, v_{M-1}$, and the corresponding eigenvalues are

$$\lambda_m^{(\Sigma)} = \frac{1}{S} \sum_{s=1}^{S} \lambda_m^{(s)} = \frac{1}{S} \sum_{s=1}^{S} \sum_{\ell=-s+1}^{s-1} \omega^m, \qquad 0 \leq m \leq M-1.$$

Now $\lambda_m^{(\Sigma)}$ can be identified as the Fejér kernel of order $S-1$ (with period 1 instead of 2π), evaluated at position m/M. It is well-known that the Fejér kernel is non-negative, so we have

$$\lambda_m^{(\Sigma)} \geq 0, \qquad m = 0, \dots, M-1, \tag{2.8}$$

and we also have

$$\lambda_0^{(\Sigma)} = \frac{1}{S} \sum_{s=1}^{S} (2s-1) = S.$$

Now using again the considerations which led to (2.7) we can show that

$$\frac{1}{S} \sum_{s=1}^{S} H_{N,M}(s) \geq M \sum_{m=0}^{M-1} \lambda_m^{(\Sigma)} \varepsilon_m^2 \geq M \lambda_0^{(\Sigma)} \varepsilon_0^2 = SN^2/M,$$

where (2.8) played a crucial role. This proves the lemma. □

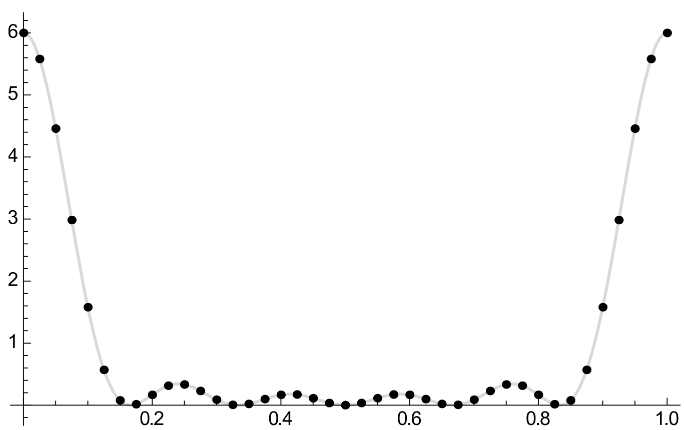

Figure 2.2: Fifth order Fejér kernel σ_5 (gray line), evaluated at positions $m/40$, $0 \leq m \leq 39$ (black dots). The m-th eigenvalue of $A^{(\Sigma)}$ for $S = 6$ equals $\sigma_5(m/40)$; note that all eigenvalues are non-negative.

2.2 Proof of Theorem 2.1

Let $(x_n)_n$ be a sequence of real numbers in $[0,1]$, and assume that it is *not* equidistributed. Thus there exists an $a \in (0,1)$ for which

$$\frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_{[0,a)}(x_n) \not\to a \qquad \text{as } N \to \infty$$

(here, and in the sequel, $\mathbb{1}_B$ denotes the indicator function of a set B). However, for this value of a by the Bolzano–Weierstraß theorem there exists a subsequence $(N_r)_{r\geq 1}$ of $\mathbb{N}$ for which a limit exists; that is, there exists a number $b \neq a$ such that

$$\lim_{r\to\infty} \frac{1}{N_r} \sum_{n=1}^{N_r} \mathbb{1}_{[0,a)}(x_n) = b. \tag{2.9}$$

Let $\varepsilon > 0$ be given, and assume that ε is "small". Choose an integer S (which is "large"). Let $r \geq 1$ be given, let N_r be from the subsequence in the previous paragraph, and consider the points $x_1, \ldots, x_{N_r}$. Let $\mathcal{E}$ denote the union of the sets

$$\left[0, \frac{2S}{N_r}\right] \cup \left[a - \frac{2S}{N_r}, a + \frac{2S}{N_r}\right] \cup \left[1 - \frac{2S}{N_r}, 1\right].$$

Furthermore, we set $B_1 = [0,a] \backslash \mathcal{E}$ and $B_2 = [a,1] \backslash \mathcal{E}$.

First consider the case that $\#\{1 \leq n \leq N_r : \ x_n \in \mathcal{E}\} \geq \varepsilon N_r$. Then by the pigeon hole principle there exists an interval of length at most $1/N_r$ in $\mathcal{E}$ which contains at least $\varepsilon N_r/(8S)$ elements of $\{x_1, \ldots, x_{N_r}\}$. All of these points are within distance $1/N_r$ of each other, which implies that

$$N_r F_{N_r}(1) \geq \left(\frac{\varepsilon N_r}{8S}\right)^2 - N_r.$$

If this inequality holds for infinitely many r, then

$$\limsup_{r\to\infty} F_{N_r}(1) = \infty,$$

which implies that the pair correlations distribution cannot be asymptotically Poissonian.

Thus we may assume that $\#\{1 \leq n \leq N : \ x_n \in \mathcal{E}\} < \varepsilon N_r$ for all elements of the subsequence $(N_r)_{r\geq 1}$. Then $[0,1]\backslash\mathcal{E} = B_1 \cup B_2$ contains at least $(1-\varepsilon)N_r$ elements of $\{x_1, \ldots, x_{N_r}\}$. Consequently, if r is sufficiently large, by (2.9) we have

$$\#\{1 \leq n \leq N_r : x_n \in B_1\} \geq (b - 2\varepsilon)N_r,$$

and

$$\#\{1 \leq n \leq N_r : x_n \in B_2\} \geq ((1-b) - 2\varepsilon)N_r.$$

We assume that r is so large that we can find positive integers M_1, M_2 for which $a/M_1 \approx (1-a)/M_2 \approx \frac{1}{N_r}$; more precisely, we demand that

$$\frac{a}{M_1} \in \left[\frac{1-\varepsilon}{N_r}, \frac{1}{N_r}\right], \qquad \frac{1-a}{M_2} \in \left[\frac{1-\varepsilon}{N_r}, \frac{1}{N_r}\right]. \tag{2.10}$$

We partition $[0,a)$ and $[a,1)$ into M_1 and M_2 disjoint subintervals of equal length, respectively, and write $y_1, \ldots, y_{M_1}$ and $z_1, \ldots, z_{M_2}$, respectively, for the number of elements of

$\{x_1,\dots,x_{N_r}\}\cap([0,1]\backslash\mathcal{E})$ contained in each of these subintervals (we assume that the subintervals are sorted in the "natural" order from left to right). Next, for $s\in\{1,\dots,S\}$ we define

$$H^*_{M_1}(s)=\sum_{m=1}^{M_1}\sum_{-s+1\le\ell\le s-1} y_m y_{m+\ell}$$

and

$$H^*_{M_2}(s)=\sum_{m=1}^{M_2}\sum_{-s+1\le\ell\le s-1} z_m z_{m+\ell}.$$

By construction, $\sum_{m=1}^{M_1} y_m\ge(b-2\varepsilon)N_r$ and $\sum_{m=1}^{M_2} z_m\ge((1-b)-2\varepsilon)N_r$. Also by construction the cyclic extension is not necessary here, provided that r is sufficiently large; by excluding all the points in $\mathcal{E}$ we have $y_1=\dots=y_S=0$ and $y_{M_1-S+1}=\dots=y_{M_1}=0$, and the same holds for the z_m's.

Then by Lemma 2.1 and by our choice of M_1,M_2 we have

$$\frac{1}{S}\sum_{s=1}^{S} H^*_{M_1}(s)\ge\frac{S((b-2\varepsilon)N_r)^2}{M_1}\ge\frac{S(b-2\varepsilon)^2N_r(1-\varepsilon)}{a},\tag{2.11}$$

and

$$\frac{1}{S}\sum_{s=1}^{S} H^*_{M_2}(s)\ge\frac{S((1-b-2\varepsilon)N_r)^2}{M_2}\ge\frac{S(1-b-2\varepsilon)^2N_r(1-\varepsilon)}{1-a}.\tag{2.12}$$

As in the calculation leading to (2.4) we can obtain a lower bound for the pair correlation function F_{N_r} from the lower bounds for $H^*_{M_1}$ and $H^*_{M_2}$. More precisely, we obtain

$$N_rF_{N_r}(s)+N_r\ge H^*_{M_1}(s)+H^*_{M_2}(s),$$

and accordingly, by (2.11) and (2.12), we have

$$\begin{aligned}&\frac{1}{S}\sum_{s=1}^{S}\left(N_rF_{N_r}(s)+N_r\right)\\ \ge\;&(1-\varepsilon)SN_r\left(\frac{(b-\varepsilon)^2}{a}+\frac{(1-b-\varepsilon)^2}{1-a}\right).\end{aligned}\tag{2.13}$$

Now note that for $0\le a,b\le 1$ we can only have $b^2/a+(1-b)^2/(1-a)=1$ if $a=b$; however, this is ruled out by assumption. For all other pairs (a,b) we have $b^2/a+(1-b)^2/(1-a)>1$, and thus (2.13) implies that

$$\frac{1}{S}\sum_{s=1}^{S}N_rF_{N_r}(s)\ge N_r\left(S(1+2c_\varepsilon)-1\right)$$

for a positive constant c_ε depending only on ε, provided that ε is sufficiently small. This implies

$$\frac{1}{S}\sum_{s=1}^{S}F_{N_r}(s)\ge S(1+2c_\varepsilon)-1\ge S(1+c_\varepsilon)\left(1+\frac{1}{S}\right),\tag{2.14}$$

where the last inequality holds under the assumption that S is sufficiently large. Consequently there exists an $s \in \{1, \dots, S\}$ such that

$$F_{N_r}(s) \geq (1 + c_\varepsilon)2s, \tag{2.15}$$

since otherwise (2.14) is impossible.

For every sufficiently large N_r in the subsequence in (2.9) such an $s \in \{1, \dots, S\}$ exists; accordingly, there is an s such that for infinitely many r we have (2.15). Thus for this s we have

$$\limsup_{r \to \infty} \frac{F_{N_r}(s)}{2s} \geq (1 + c_\varepsilon) > 1,$$

which proves the theorem.

2.3 Proof of Theorem 2.2

First assume that the measure μ_G defined by the asymptotic distribution function G is not absolutely continuous with respect to the Lebesgue measure. A function which is not absolutely continuous is not Lipschitz continuous as well. Thus there is an $\varepsilon > 0$ such that for every (small) $\delta > 0$ there exists an interval $I = [a, b] \subset [0, 1]$ such that

$$\lambda(I) \leq \delta, \qquad \text{but} \qquad \mu_G(I) \geq \varepsilon,$$

where λ denotes the Lebesgue measure (that is, the length) of I. The interval $I_1 = [a - \delta/2, b + \delta/2]$ must contain at least $\varepsilon N/2$ elements of $(x_n)_{1 \leq n \leq N}$ for sufficiently large N. Set $I_2 = [a - \delta, b + \delta]$, and set $M = \lceil \lambda(I) N \delta^{-1/2} \rceil$, split I_2 into M subintervals, and denote the number of elements of the set $\{x_1, \dots, x_N\} \cap I_1$ contained in each of these subintervals by $y_1, \dots, y_M$, respectively. The sets I_1 and I_2 are constructed in a way to remove the influence of the cyclic "overlap" in Lemma 2.1; in fact, by construction $y_m = 0$ for all m being close to either 0 or M. Let $\hat{N} = y_1 + \dots + y_M$, and define $H_{\hat{N},M}(1) = y_1^2 + \dots + y_M^2$. Note that $\lambda(I_2) \leq 2\delta$ and that $\hat{N} \geq \{x_1, \dots, x_N\} \cap I_1 \geq \varepsilon N/2$ for sufficiently large N. Applying Lemma 2.1 and using a rescaled version of (2.4) we have

$$\begin{aligned} NR(s,N)\left(4\delta^{1/2}\right) + N &\geq NR(s,N)\left(\frac{2N\lambda(I_2)}{M}\right) + N \\ &\geq H_{\hat{N},M}(1) \\ &\geq \frac{(\varepsilon N/2)^2}{M} \\ &\geq \frac{\varepsilon^2 N}{5\delta^{1/2}} \end{aligned} \tag{2.16}$$

for sufficiently large N. Recall that by definition $R(s, N)$ is monotonic increasing as a function of s, and thus by (2.16) we have

$$R(s,N)\,(s) \geq \frac{\varepsilon^2}{5\delta^{1/2}} - 1 \qquad \text{for all } s \geq 4\delta^{1/2},$$

for all sufficiently large N. Since ε is fixed and δ can be chosen arbitrarily small, this proves the theorem when μ_G is not absolutely continuous.

Now assume that the measure μ_G defined by G is absolutely continuous with respect to the Lebesgue measure, and thus has a density g. In the upcoming sequel we will think of $([0,1],\mathcal{B}([0,1]),\lambda)$ as a probability space, and write $\mathbb{E}$ for the expected value (of a measurable real function) with respect to this space. Let $R \geq 1$ be an integer. We split the unit interval into 2^R intervals of equal lengths. Let $\mathcal{F}_R$ denote the σ-field generated by these intervals. Assume for simplicity that g is bounded on $[0,1]$. Then for any of the intervals I which generate $\mathcal{F}_R$ the number of elements of $(x_n)_{1 \leq n \leq N}$ which are contained in I is asymptotically proportional to $\mu_G(I)N$, and we can use arguments similar to those in the proof of Theorem 2.1 to prove that for given $\varepsilon > 0$ there exist infinitely many values of s such that for each of these values we have

$$\begin{aligned} \frac{F(s)}{2s} &\geq \sum_{\text{intervals } I \text{ generating } \mathcal{F}_R} \frac{(\mu_G(I))^2}{2^R} - \varepsilon \\ &= \mathbb{E}\left(\left(\mathbb{E}(g|\mathcal{F}_R)\right)^2\right) - \varepsilon, \end{aligned}$$

where $\mathbb{E}(g|\mathcal{F}_k)$ denotes the conditional expectation of g under the σ-field $\mathcal{F}_R$.[2] Note that a direct generalization of the proof of Theorem 2.1 only guarantees the existence of *one* such integer s; however, we can use the fact that F is monotonically increasing to show that there actually must be infinitely many such values of s. The family $(\mathcal{F}_R)_{R\geq 1}$ forms a filtration whose limit is $\mathcal{B}([0,1])$, in the sense that $\mathcal{B}([0,1])$ is the sigma-field generated by $\bigcup_{R\geq 1}\mathcal{F}_R$. Thus by the convergence theorem for conditional expectations (also known as *Lévy's zero-one law*) we have

$$\lim_{R\to\infty} \mathbb{E}\left(\left(\mathbb{E}(g|\mathcal{F}_R)\right)^2\right) = \mathbb{E}\left(g^2\right) = \int_0^1 g(x)^2 \, dx,$$

which proves the theorem in the case when g is bounded. Finally, if g is not bounded then we can apply the argument above to a truncated version g_{trunc} of g and show that in this case

$$\limsup_{s\to\infty} \frac{R(s)}{2s} \geq \int_0^1 g_{\text{trunc}}(x)^2 \, dx.$$

By raising the level where g is truncated this square-integral can be made arbitrarily close to $\int_0^1 g(x)^2 \, dx$, or arbitrarily large in case we have $\int_0^1 g(x)^2 \, dx = \infty$. This proves the theorem.

[2] In the proof of Theorem 2.1, the role of the second moment of the conditional expectation function is played by the expression $b^2/a + (1-b)^2/(1-a)$, which appears in line (2.13). Instead of splitting $[0,1]$ into two subintervals and applying Lemma 2.1 to these two pieces, as in the proof of Theorem 2.1, we now split $[0,1]$ into 2^R subintervals and apply the lemma 2^R times. For each of these subintervals I the number of points contained is roughly $\mu_G(I)N$, and the number M which appears in Lemma 2.1 is chosen as $2^{-R}N$.

Chapter 3

On Exceptional Sets in the Metric Poissonian Pair Correlations Problem

3.1 Introduction

The theory of uniform distribution modulo 1 dates back, at least, to the seminal paper [46] of Weyl who showed that for any fixed $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ and integer $d \geq 1$ the sequences $\left(\langle \alpha n^d \rangle\right)_n$ are uniformly distributed modulo 1 where $\langle x \rangle$ denotes the fractional part of $x \in \mathbb{R}$. However, in recent years various authors [4, 9, 11, 24, 31, 33, 35, 37, 38, 39, 43, 45] have been investigating a more subtle distribution property of such sequences — namely, whether the asymptotic distribution of the pair correlations has a property which is called Poissonian, and defined as follows:

Definition. Let $\|\cdot\|$ denote the distance to the nearest integer. A sequence $(\theta_n)_n$ in $[0,1]$ is said to have *(asymptotically) Poissonian pair correlations*, for shorthand notation PPC, if for each $s \geq 0$ the pair correlation function[1]

$$R_2\left(s, (\theta_n)_n, N\right) \coloneqq \frac{1}{N} \# \left\{ 1 \leq i \neq j \leq N : \|\theta_i - \theta_j\| \leq \frac{s}{N} \right\} \tag{3.1}$$

tends to $2s$ as $N \to \infty$. Moreover, let $(a_n)_n$ denote a strictly increasing sequence of positive integers. If no confusion can arise, we write

$$R(s, \alpha, N) \coloneqq R_2\left(s, (\alpha a_n)_n, N\right)$$

and say that a sequence $(a_n)_n$ has *metric Poissonian pair correlations* if $(\alpha a_n)_n$ has PPC for almost all $\alpha \in [0,1]$ where almost all throughout this article is meant with respect to the Lebesgue measure.

As seen in Chapter 2 seen that if a sequence $(\theta_n)_n$ has PPC, then it is uniformly distributed modulo 1. Yet, the sequences $\left(\langle \alpha n^d \rangle\right)_n$ do *not* have PPC for *any* $\alpha \in \mathbb{R}$ if $d = 1$. For $d \geq 2$, Rudnick and Sarnak [37] proved that $\left(n^d\right)_n$ has metric Poissonian pair correlations (metric PPC). A result of Aistleitner, Larcher, and Lewko [4], who used a Fourier analytic

[1] The subscript 2 in R_2 indicates that relations of second order, i.e. pair correlations, are counted.

approach combined with a bound on GCD sums of Bondarenko and Seip [12], uncovered the connection of the metric PPC property of $(a_n)_n$ with its combinatoric properties. For stating it, we introduce some notation. Let $(a_n)_n$ denote throughout this article a strictly increasing sequence of positive integers, and abbreviate the set of the first N elements of $(a_n)_n$ by A_N. Moreover, define the additive energy $E(I)$ of a finite set of integers I via

$$E(I) := \#\left\{(a,b,c,d) \in I^4 : \, a+b=c+d\right\},$$

and note that $(\#I)^2 \le E(I) \le (\#I)^3$ where $\#S$ denotes the cardinality of a set S. In the following, let $\mathcal{O}$ and o denote the Landau symbols/O-notation, and $\ll$ or $\gg$ the Vinogradov symbols. The dependence of an implied constant in one of these symbols will be indicated by mentioning this parameter in a subscript.
Now, a main finding of [4] can be stated as the implication that if the truncations A_N satisfy

$$E(A_N) = \mathcal{O}\left(N^{3-\varepsilon}\right) \tag{3.2}$$

for some fixed $\varepsilon > 0$, then $(a_n)_n$ has metric PPC. Roughly speaking, a set I has large additive energy if and only if it contains a "large" arithmetic progression like structure. Indeed, if $(a_n)_n$ is a geometric progression or of the form $\left(n^d\right)_n$ for $d \ge 2$, then (3.2) is satisfied.

Recently, Bloom, Chow, Gafni, Walker relaxed — provided that, roughly speaking, the density of the sequence does not decay faster than $1/(\log N)^2$ — the power saving bound (3.2) for detecting the metric PPC property of $(a_n)_n$ significantly:

Theorem A (Bloom, Chow, Gafni, Walker [9]). *If there exists an $\varepsilon > 0$ such that*

$$E(A_N) \ll_\varepsilon \frac{N^3}{(\log N)^{2+\varepsilon}} \quad \text{and} \quad \frac{1}{N}\#\left(A_N \cap \{1,\dots,N\}\right) \gg_\varepsilon \frac{1}{(\log N)^{2+2\varepsilon}}$$

hold, then $(a_n)_n$ has metric PPC.

In accordance with probabilistic considerations, cf. [9, Thm. 1.5], the above result could be seen as a sign of Khintchine-type law underpinning the characterization of the metric PPC property of $(a_n)_n$. Indeed, the following basic question about the nature of the connection between additive energy and the metric PPC property was raised in [9]:

Fundamental Question (Bloom, Chow, Gafni, Walker [9]). *Is it true that if $E(A_N) \sim N^3\psi(N)$ for some weakly decreasing function $\psi : \mathbb{N} \to [0,1]$, then $(a_n)_n$ has metric PPC if and only if $\sum_{N\ge 1}\psi(N)/N$ converges?*

Remark. This question will be answered in the negative in the following chapter.

Regarding the optimal bound for $E(A_N)$ to ensure the metric PPC property of $(a_n)_n$, the following two questions were raised in [4]. For stating those, we use the convention that $f = \Omega(g)$ means for $f,g : \mathbb{N} \to \mathbb{R}$ there is a constant $c > 0$ such that $g(n) > cf(n)$ holds for infinitely many n.

Question 1. *Is it possible for $(a_n)_n$ with $E(A_N) = \Omega\left(N^3\right)$ to have metric PPC?*

Question 2. *Do all $(a_n)_n$ with $E(A_N) = o\left(N^3\right)$ have metric PPC?*

Both questions were answered in the negative by Bourgain whose proofs can be found in [4] as an appendix, without giving an estimate on the measure of the set that was used to answer Question 1, and without a quantitative bound on $E(A_N)$ appearing in the negation of Question 2. However, a quantitative analysis, as noted in [45], shows that the sequence Bourgain constructed for Question 2 satisfies

$$E(A_N) = \mathcal{O}_\varepsilon\left(\frac{N^3}{(\log\log N)^{\frac{1}{4}+\varepsilon}}\right) \tag{3.3}$$

for any fixed $\varepsilon > 0$. Moreover, Nair posed the problem[2] whether the sequence of prime numbers $(p_n)_n$, ordered by increasing value, has metric PPC. Recently, Walker [45] answered this question in the negative by showing that there is a constant $c > 0$ satisfying that for almost every $\alpha \in [0,1]$ the inequality $R(s,\alpha,N) > c$ holds for infinitely many N. Thereby he gave a significantly better bound than (3.3) for the additive energy $E(A_n)$ for a sequence $(a_n)_n$ not having metric PPC — since the additive energy of the truncations of $(p_n)_n$ is in $\Theta\big((\log N)^{-1} N^3\big)$ where $f = \Theta(g)$, for functions f, g, means that both $f = \mathcal{O}(g)$ and $g = \mathcal{O}(f)$ holds.

For a given sequence $(a_n)_n$, we denote by $\mathrm{NPPC}((a_n)_n)$ the "exceptional" set of all $\alpha \in (0,1)$ such that $(\langle \alpha a_n \rangle)_n$ does not have PPC.

Theorem B (Bourgain, [4]). *Assume $E(A_N) = \Omega(N^3)$, then $\mathrm{NPPC}((a_n)_n)$ has positive Lebesgue measure.*

We prove the following sharpening.

Theorem 3.1. *If $E(A_N) = \Omega(N^3)$, then $\mathrm{NPPC}((a_n)_n)$ has full Lebesgue measure.*

Some remarks are in order.

Remark. (a) $(a_n)_n$ is called quasi-arithmetic of degree one, cf. [1, Def. 1], if infinitely often at least a constant proportion of elements of A_N is contained in some arithmetic progression of length $\ll N$. Any such sequence obviously satisfies $E(A_N) = \Omega(N^3)$. Theorem 3.1 improves upon a recent result of Aichinger, Aistleitner, and Larcher [1, Thm. 3] who showed that $\mathrm{NPPC}((a_n)_n)$ has full Lebesgue measure, if $(a_n)_n$ is quasi-arithmetic of degree one.
(b) Recently, Larcher [29, Thm. 1] sharpened this result to $\mathrm{NPPC}((a_n)_n) = (0,1)$, and subsequently Larcher and Stockinger [30, Thm. 1] extended this to quasi-arithmetic sequences of any degree $d \geq 1$ — which due to Freiman's theorem (cf. [29, text above Def. 2]) implies that if $E(A_N) = \Omega(N^3)$, then $\mathrm{NPPC}((a_n)_n) = (0,1)$.

For stating our second main theorem, we denote by $\mathbb{R}_{>x}$ the set of real numbers exceeding a given $x \in \mathbb{R}$.

Theorem 3.2. *Let $f : \mathbb{R}_{>0} \to \mathbb{R}_{>2}$ be a function increasing monotonically to ∞, and satisfying $f(x) = \mathcal{O}\big(x^{1/3}(\log x)^{-7/3}\big)$. Then, there is a strictly increasing sequence $(a_n)_n$ of positive integers with $E(A_N) = \Theta\big(N^3/f(N)\big)$ such that if*

$$\sum_{n\geq 1} \frac{1}{nf(n)} \tag{3.4}$$

[2] This problem was posed at the problem session of the ELAZ conference in 2016.

diverges, then for Lebesgue almost all $\alpha \in [0,1]$

$$\limsup_{N\to\infty} R\left(s,\alpha,N\right) = \infty \tag{3.5}$$

holds for any $s > 0$*; additionally, if (3.4) converges and* $\sup\left\{f\left(2x\right)/f\left(x\right) : x \geq x_0\right\}$ *is strictly less than* 2 *for some* $x_0 > 0$*, then* $\mathrm{NPPC}\left(\left(a_n\right)_n\right)$ *has Hausdorff dimension at least* $\left(1+\lambda\left(f\right)\right)^{-1}$ *where*

$$\lambda\left(f\right) \coloneqq \liminf_{x\to\infty} \frac{\log f\left(x\right)}{\log x}$$

denotes the lower order of infinity of f*.*

We record an immediate consequence of Theorem 3.2 by using the convention that the r-folded iterated logarithm is denoted by $\log_r\left(x\right)$, i.e.

$$\log_r\left(x\right) \coloneqq \log_{r-1}\left(\log\left(x\right)\right)$$

and $\log_1\left(x\right) \coloneqq \log\left(x\right)$.

Corollary 3.1. *Let* r *be a positive integer. Then, there is a strictly increasing sequence* $\left(a_n\right)_n$ *of positive integers with*

$$E\left(A_N\right) = \Theta\left(\frac{N^3}{\log\left(N\right)\log_2\left(N\right)\ldots\log_r\left(N\right)}\right)$$

such that $\mathrm{NPPC}\left(\left(a_n\right)_n\right)$ *has full Lebesgue measure. Moreover, for any* $\varepsilon > 0$ *there is a strictly increasing sequence* $\left(a_n\right)_n$ *of positive integers with*

$$E\left(A_N\right) = \Theta\left(\frac{\left(\log_r\left(N\right)\right)^{-\varepsilon} N^3}{\log\left(N\right)\log_2\left(N\right)\ldots\log_r\left(N\right)}\right)$$

such that $\mathrm{NPPC}\left(\left(a_n\right)_n\right)$ *has full Hausdorff dimension.*

The proof of Theorem 3.2 connects the metric PPC property to the notion of optimal regular systems from Diophantine approximation. It uses, among other things, a Khintchine-type theorem due to Beresnevich. Furthermore, despite leading to better bounds, the nature of the sequences underpinning Theorem 3.2 is much simpler than the nature of those sequences previously constructed by Bourgain [4] (who used, inter alia, large deviations inequalities from probability theory), or the sequence of prime numbers studied by Walker [45] (who relied on estimates, derived by the circle-method, on the exceptional set in Goldbach-like problems).

3.2 First Main Theorem

Let us give an outline of the proof of Theorem 3.1. For doing so, we begin by sketching the reasoning of Theorem B: As it turns out, except for a set of negligible measure, the counting function in (3.1) can be written as a function (of α) that admits a non-trivial estimate for its L^1-mean value. The mean value is infinitely often too small on sets whose measure is uniformly bounded from below. Thus, there exists a sequence of sets $(\Omega_r)_r$ of $\alpha \in [0,1]$ such that $R(s,\alpha,N)$ is too small for every $\alpha \in \Omega_r$ for having PPC and Theorem B follows.

Our reasoning for proving Theorem 3.1 is building upon this argument of Bourgain while we introduce new ideas to construct a sequence of sets $(\Omega_r)_r$ that are "pairwise quasi independent" - meaning that for every fixed t the relation

$$\lambda(\Omega_r \cap \Omega_t) \leq \lambda(\Omega_r)\lambda(\Omega_t) + o(1)$$

holds as $r \to \infty$ where λ is the Lebesgue measure. Roughly speaking, applying a suitable version of the Borel–Cantelli lemma, combined with a sufficiently careful treatment of the $o(1)$ term, will then yield Theorem 3.1. However, before proceeding with the details of the proof we collect in the next paragraph some tools from additive combinatorics that are needed.

3.2.1 Preliminaries

We start with a well-know result relating, in a quantitative manner, the additive energy of a set of integers with the existence of a (relatively) dense subset with small difference set where the difference set $B - B := \{b - b' : b, b' \in B\}$ for a set $B \subseteq \mathbb{R}$.

Lemma 3.1 (Balog–Szeméredi–Gowers lemma, [42, Thm 2.29])**.** *Let $A \subseteq \mathbb{Z}$ be a finite set of integers. For any $c > 0$ there exist $c_1, c_2 > 0$ depending only on c such that the following holds. If $E(A) \geq c\,(\#A)^3$, then there is a subset $B \subseteq A$ such that*

1. $\#B \geq c_1 \#A$,
2. $\#(B - B) \leq c_2 \#A$.

Moreover, we recall that for $\delta > 0$ and $d \in \mathbb{Z}$ the set

$$B(d,\delta) := \{\alpha \in [0,1] : \|d\alpha\| \leq \delta\}$$

is called Bohr set. The following two simple observations will be useful.

Lemma 3.2. *Let $B \subseteq \mathbb{Z}$ be a finite set of integers. Then,*

$$\lambda\left(\left\{\alpha \in [0,1] : \min_{d \in (B-B)\setminus\{0\}} \|d\alpha\| < \frac{\varepsilon}{\#(B-B)}\right\}\right) \leq 2\varepsilon$$

for every $\varepsilon \in (0,1)$.

Proof. By observing that the set under consideration is contained in

$$\bigcup_{\substack{m,n \in B \\ m \neq n}} B\left(m-n, \frac{\varepsilon}{\#(B-B)}\right),$$

and

$$\lambda\left(B\left(m-n, \frac{\varepsilon}{\#(B-B)}\right)\right) = \frac{2\varepsilon}{\#(B-B)},$$

the claim follows at once. □

Lemma 3.3. *Suppose A is a finite intersection of Bohr sets, and B is a finite union of Bohr sets. Then, $A \setminus B$ is the union of finitely many intervals.*

Furthermore, we shall use the Borel–Cantelli lemma in a version due to Erdős, and Rényi.

Lemma 3.4 (Erdős–Rényi, cf. [21, Lem. 2.3])**.** *Let $(A_n)_n$ be a sequence of Lebesgue measurable sets in $[0,1]$ satisfying*

$$\sum_{n\geq 1} \lambda\left(A_n\right) = \infty.$$

Then,

$$\lambda\left(\limsup_{n\to\infty} A_n\right) \geq \limsup_{N\to\infty} \frac{\left(\sum_{n\leq N} \lambda\left(A_n\right)\right)^2}{\sum_{m,n\leq N} \lambda\left(A_n \cap A_m\right)}.$$

Moreover, let us explain the main steps in the proof of Theorem 3.1. Let

$$\varepsilon := \varepsilon\left(j\right) := \frac{1}{10^j} c_1^2$$

where $c_1 > 0$ is a constant to be specified later-on, and j denotes a positive integer. In the first part of the argument, we show how a sequence — that is constructed in the second part of the argument — can be used to deduce Theorem 3.1. For every fixed j, we find a corresponding $s = s(j)$ and construct inductively a sequence $(\Omega_r)_r$ of exceptional values α with the following properties:

(i) For all $\alpha \in \Omega_r$, the pair correlation function admits the upper bound

$$R\left(s, \alpha, N\right) \leq 2\tilde{c}s \tag{3.6}$$

for some absolute constant $\tilde{c} \in (0,1)$, depending on $(a_n)_n$ only.

(ii) For all integers $r > t \geq 1$, the relation

$$\lambda\left(\Omega_r \cap \Omega_t\right) \leq \lambda\left(\Omega_r\right)\lambda\left(\Omega_t\right) + 2\varepsilon\lambda\left(\Omega_t\right) + \mathcal{O}\left(r^{-2}\right) \tag{3.7}$$

holds.

(iii) Each Ω_r is the union of finitely many intervals (hence measurable).

(iv) For all $r \geq 1$, the measure $\lambda\left(\Omega_r\right)$ is uniformly bounded from below by

$$\lambda\left(\Omega_r\right) \geq \frac{c_1^2}{8}. \tag{3.8}$$

3.2.2 Proof of Theorem 3.1

1. Suppose there is $(\Omega_r)_r$ satisfying (i)–(iv). Then, by using (3.7), we get

$$\begin{aligned}\sum_{r,t\le N}\lambda\left(\Omega_r\cap\Omega_t\right) &\le 2\sum_{2\le t\le N}\sum_{1\le r<t}\left(\lambda\left(\Omega_r\right)\lambda\left(\Omega_t\right)\right)+2\varepsilon N^2+\mathcal{O}\left(N\right)\\ &\le \left(\sum_{t\le N}\lambda\left(\Omega_t\right)\right)^2+2\varepsilon N^2+\mathcal{O}\left(N\right).\end{aligned}$$

By recalling that $\Omega_r=\Omega_r\left(\varepsilon\right)=\Omega_r\left(j\right)$, we let

$$\Omega(j):=\limsup_{r\to\infty}\Omega_r.$$

By using the inequality above in combination with Lemma 3.4 and the bound (3.8), we obtain that the set $\Omega(j)$ has measure at least

$$\limsup_{N\to\infty}\frac{\left(\sum_{r\le N}\lambda\left(\Omega_r\right)\right)^2}{\sum_{r,t\le N}\lambda\left(\Omega_r\cap\Omega_t\right)}\ge\limsup_{N\to\infty}\frac{1}{1+\frac{4\varepsilon N^2}{\left(\sum_{r\le N}\lambda(\Omega_r)\right)^2}}\ge\limsup_{N\to\infty}\frac{1}{1+\frac{256}{c_1^4}\varepsilon}=\frac{1}{1+\frac{256}{c_1^4}\varepsilon}.$$

Note that due to (3.6), for every $\alpha\in\Omega\left(j\right)$ the sequence $\left(\alpha a_n\right)_n$ does not have PPC. Now, letting $j\to\infty$ proves the assertion.

2. For constructing $(\Omega_r)_r$ with the required properties, let $c>0$ such that $E\left(A_N\right)>cN^3$ for infinitely many integers N. By choosing an appropriate subsequence $(N_i)_i$ and omitting the subscript i for ease of notation, we may suppose that $E\left(A_N\right)>cN^3$ holds for every N occurring in this proof. Moreover, let c_1,c_2 and B_N be as in Lemma 3.1, corresponding to the c just mentioned. Let

$$s=\frac{\varepsilon}{2c_2}.$$

Arguing inductively, while postponing the base step,[3] we may assume that there are sets $(\Omega_r)_{1\le r<R}$ given that satisfy the properties (i)–(iv) for all distinct integers $1\le r,t<R$. Let $N\ge R$. Since, due to Lemma 3.1,

$$\frac{s}{N}\le\frac{\varepsilon}{\#\left(B-B\right)},$$

Lemma 3.2 implies that the set $\Omega_{\varepsilon,N}$ of all $\alpha\in[0,1]$ satisfying $\|\left(r-t\right)\alpha\|<N^{-1}s$ for some distinct $r,t\in B_N$ has measure at most 2ε. Setting

$$\mathcal{D}_N:=\left\{(r,t)\in\left(A_N\times A_N\right)\setminus\left(B_N\times B_N\right):\ r\ne t\right\},$$

we get for $\alpha\notin\Omega_{\varepsilon,N}$ that

$$R\left(s,\alpha,N\right)=\frac{1}{N}\#\left\{(r,t)\in\mathcal{D}_N:\ \|\left(r-t\right)\alpha\|<N^{-1}s\right\}.$$

[3]The base step uses simplified versions of the arguments exploited in the induction step, and will therefore be postponed.

Let ℓ_R denote the length of the smallest subinterval of Ω_r for $1 \leq r < R$, and define $C(\Omega_r)$ to be the set of subintervals of Ω_r. Note that $\ell_R > 0$, and $\max_{1\leq r<R} \#C(\Omega_r) < \infty$. We divide $[0,1)$ into

$$P := \left\lfloor 1 + 2\ell_R^{-1} R^2 \max_{1\leq r<R} \#C(\Omega_r) \right\rfloor$$

parts $\mathcal{P}_i$ of equal lengths (where $\lfloor x \rfloor$ is the integer part of x), i.e.

$$\mathcal{P}_i := \left[\frac{i}{P}, \frac{i+1}{P}\right)$$

where $i = 0, \ldots, P-1$. Let $\mathbf{1}_X$ denote the characteristic function of a Borel set $X \subseteq [0,1]$. After writing

$$\frac{1}{N} \int_{\mathcal{P}_i} \#\left\{(r,t) \in \mathcal{D}_N : \|(r-t)\alpha\| \leq N^{-1}s\right\} \mathrm{d}\alpha = \frac{1}{N} \sum_{(r,t)\in\mathcal{D}_N} \int_{\mathcal{P}_i} \mathbf{1}_{[-\frac{s}{N},\frac{s}{N}]} (\|(r-t)\alpha\|) \,\mathrm{d}\alpha, \tag{3.9}$$

we split the sum into two parts: one part containing differences $|r-t| > R^k P$, and a second part containing differences $|r-t| \leq R^k P$ where

$$k := \left\lfloor \frac{1}{\log 2} \log \frac{8(4s+1)}{\left(c_1^2 - 2^{-1}c_1^4\right)s} \right\rfloor + 1.$$

The Cauchy–Schwarz inequality implies

$$\int_{\mathcal{P}_i} \mathbf{1}_{[0,\frac{s}{N}]} (\|(r-t)\alpha\|) \,\mathrm{d}\alpha \leq \sqrt{\frac{1}{P}\frac{2s}{N}}.$$

Since for any $x > 0$ there are at most $2xN$ choices of $(r,t) \in \mathcal{D}_N$ such that $|r-t| \leq x$, we obtain

$$\frac{1}{N} \sum_{\substack{(r,t)\in\mathcal{D}_N \\ |r-t|\leq PR^k}} \int_{\mathcal{P}_i} \mathbf{1}_{[0,\frac{s}{N}]} (\|(r-t)\alpha\|) \,\mathrm{d}\alpha \leq 2PR^k \sqrt{\frac{1}{P}\frac{2s}{N}}$$

which is $\leq P^{-1}R^{-k}$ if N is sufficiently large. Moreover, for any $|r-t| > PR^k$ we observe that

$$\int_{\mathcal{P}_i} \mathbf{1}_{[0,\frac{s}{N}]} (\|(r-t)\alpha\|) \,\mathrm{d}\alpha \leq \frac{2s}{N|r-t|} \left(\#\left\{j \leq |r-t| : j/|r-t| \in \mathcal{P}_i\right\} + 1\right) \leq \frac{2s}{PN} + \frac{4s}{PR^k N}.$$

Also note that $\#\mathcal{D}_N \leq N^2 - \left(\#B_N\right)^2 \leq \tilde{c}N^2$ where $\tilde{c} := 1 - c_1^2$. Therefore, the mean value (3.9) of the modified pair correlation counting function on the interval $\mathcal{P}_i$ admits the upper bound

$$\frac{1}{N}(\#\mathcal{D}_N)\left(\frac{2s}{PN} + \frac{4s}{PR^k N}\right) + \frac{1}{PR^k} \leq \frac{2\tilde{c}s}{P} + \frac{4s+1}{PR^k}.$$

Hence, it follows that the measure of the set $\Delta_N(i)$ of $\alpha \in \mathcal{P}_i$ with

$$\frac{1}{N}\#\left\{(r,t) \in \mathcal{D}_N : \|(r-t)\alpha\| \le N^{-1}s\right\} \le 2\left(1 - \frac{c_1^2}{2}\right)s \tag{3.10}$$

admits, by the choice of k, the lower bound

$$\lambda(\Delta_N(i)) \ge \frac{1}{P} - \frac{1}{P}\frac{2\tilde{c}s + (4s+1)R^{-k}}{2\left(1 - \frac{c_1^2}{2}\right)s} \ge \frac{1}{P}\left(\frac{c_1^2}{2} - \frac{c_1^2}{8}\right). \tag{3.11}$$

Note that $\Delta_N(i)$ is the union of finitely many intervals, due to Lemma 3.3. So, we may take $\Delta_N'(i) \subset \Delta_N(i)$ being a finite union of intervals such that $\lambda(\Delta_N'(i))$ equals the lower bound in (3.11). Let

$$\Omega_R \coloneqq \Omega_R(N) \coloneqq \Delta_N \setminus \Omega_{\varepsilon,N} \qquad \text{where} \qquad \Delta_N \coloneqq \bigcup_{i=0}^{P-1} \Delta_N'(i).$$

We are going to show now that Ω_R satisfies the properties (i)–(iv). Now, Ω_R satisfies property (iv) with $r = R$ since

$$\lambda(\Omega_R) \ge \lambda(\Delta_N) - \lambda(\Omega_{\varepsilon,N}) = \frac{c_1^2}{2} - \frac{c_1^2}{8} - 2\varepsilon \ge \frac{c_1^2}{8}.$$

Furthermore, Ω_R satisfies property (i) by construction and also property (iii) since all sets involved in the construction of Ω_R were a finite union of intervals. Let $1 \le r < R$, and I be a subinterval of Ω_r. Then,

$$\lambda(I \cap \Delta_N) = \sum_{i:\mathcal{P}_i \cap I \neq \emptyset} \lambda(\mathcal{P}_i \cap I \cap \Delta_N) \le \frac{2}{P} + \sum_{i:\mathcal{P}_i \subsetneq I} \lambda(\mathcal{P}_i \cap \Delta_N) \le \frac{2}{P} + \sum_{i:\mathcal{P}_i \subsetneq I} \lambda\left(\Delta_N'(i)\right).$$

By summing over all subintervals $I \in C(\Omega_r)$, we obtain that

$$\begin{aligned}\lambda(\Omega_r \cap \Delta_N) &\le \sum_{I \in C(\Omega_r)} \left(\frac{2}{P} + \sum_{i:\mathcal{P}_i \subsetneq I} \lambda\left(\Delta_N'(i)\right)\right) \le \frac{1}{R^2} + \sum_{I \in C(\Omega_r)} P\lambda(I)\frac{\lambda(\Delta_N)}{P} \\ &= \lambda(\Omega_r)\lambda(\Delta_N) + \frac{1}{R^2}.\end{aligned}$$

We deduce property (ii) from this estimate and Lemma 3.2 via

$$\begin{aligned}\lambda(\Omega_r \cap \Omega_R) &\le \lambda(\Omega_r \cap \Delta_N) \\ &\le \lambda(\Omega_r)(\lambda(\Delta_N) - \lambda(\Omega_{\varepsilon,N})) + R^{-2} + \lambda(\Omega_r)\lambda(\Omega_{\varepsilon,N}) \\ &\le \lambda(\Omega_r)\lambda(\Omega_R) + 2\varepsilon\lambda(\Omega_r) + R^{-2}.\end{aligned}$$

This concludes the induction step. The only part missing now is the base step of the induction. For realizing it, let N denote the smallest integer m with $E(A_m) > cm^3$. We replace $\mathcal{P}_i$ in (3.9) by $[0,1]$ to directly derive

$$\int_0^1 \frac{1}{N}\#\left\{(r,t) \in \mathcal{D}_N : \|(r-t)\alpha\| \le N^{-1}s\right\} \mathrm{d}\alpha \le 2\tilde{c}s,$$

and conclude that the set Ω_1' of $\alpha \in [0,1]$ satisfying (3.10) has a measure of at least $c_1^2/2$. Thus, $\Omega_1 \coloneqq \Omega_1' \setminus \Omega_{N,\varepsilon}$ has measure at least as large as the right hand side of (3.8). For property (3.7), there is nothing to check and that Ω_1 is a finite union of intervals follows from Lemma 3.3 by observing that

$$\Omega_1' = \bigcap_{d_1,\ldots,d_{\lfloor N2\tilde{c}s\rfloor}} \left(B\left(d_1, N^{-1}s\right)^C \cup \ldots \cup B\left(d_{\lfloor N2\tilde{c}s\rfloor}, N^{-1}s\right)^C \right)$$

where the intersection runs through any set of $\lfloor N2\tilde{c}s\rfloor$-tuples of differences $d_i = r_i - t_i \neq 0$ of components of $(r_i,t_i) \in \mathcal{D}_N$ for $i = 1,\ldots,\lfloor N2\tilde{c}s\rfloor$.

Thus, the proof is complete.

3.3 Second Main Theorem

The sequences $(a_n)_n$ enunciated in Theorem 3.2 are constructed in two steps. In the first step, we concatenate (finite) blocks, with suitable lengths, of arithmetic progressions to form a set P_A. In the second step, we concatenate (finite) blocks, with suitable lengths, of geometric progressions to form a set P_G and then define a_n to be the n-th smallest element of $P_A \cup P_G$. On the one hand, the arithmetic progression like part P_A serves to ensure, due to considerations from metric Diophantine approximation, the divergence property (3.5) on a set with full measure or controllable Hausdorff dimension; on the other hand, the geometric progression like part P_G lowers the additive energy, as much as it can. For doing so, a geometric block will appear exactly before and after an arithmetic block, and have much more elements.

For writing the construction precisely down, we introduce some notation. Suppose throughout this section that f is as in Theorem 3.2. We set $P_A^{(1)}$ to be the empty set while $P_G^{(1)} \coloneqq \{1,2\}$. Suppose $P_A^{(j-1)}, P_G^{(j-1)}$ for $j \geq 2$ are already constructed. Let $C_j = 2\max\{P_G^{(j-1)}\}$. Then

$$P_A^{(j)} \coloneqq \left\{ C_j + h :\, 1 \leq h \leq \left\lfloor \left(f(2^j)\right)^{-\beta} 2^j \right\rfloor \right\},$$

and $P_G^{(j)}$ is defined via

$$P_G^{(j)} \coloneqq \left\{ 2C_j + 2^i :\, 1 \leq i \leq \left\lfloor \left(f(2^j)\right)^{-\gamma} 2^j \left(1 - \left(f(2^j)\right)^{\gamma-\beta}\right) \right\rfloor \right\}$$

where $0 < \gamma < \beta < 3/4$ are parameters[4] to be chosen later-on. Letting

$$P_A \coloneqq \bigcup_{j\geq 1} P_A^{(j)}, \qquad P_G \coloneqq \bigcup_{j\geq 1} P_G^{(j)},$$

we denote by a_n the n-th smallest element in $P_A \cup P_G$. For $d \in \mathbb{Z}$ and finite sets of integers X, Y, we abbreviate the number of representations of d as a difference of an $x \in X$ and a $y \in Y$ by

$$\mathrm{rep}_{X,Y}(d) \coloneqq \#\{(x,y) \in X \times Y :\, x - y = d\};$$

[4] No particular importance should be attached to requiring $\beta < 3/4$, or using "dyadic steps lengths 2^j". Doing so is for simplifying the technical details only - eventually, it will turn out that $\beta = 2/3 = 2\gamma$ is the optimal choice of parameters in this approach. For proving this to the reader, we leave γ, β undetermined till the end of this section.

for later reference, we record here that the additive energy of a set X and the pair correlation counting function can be written as

$$E(X) = \sum_{d\in\mathbb{Z}} \left(\mathrm{rep}_{X,X}(d)\right)^2, \tag{3.12}$$

and

$$R(s,\alpha,N) = \frac{1}{N}\sum_{d\in\mathbb{Z}\setminus\{0\}} \mathrm{rep}_{A_N,A_N}(d)\mathbf{1}_{[0,\frac{s}{N}]}(\|\alpha d\|). \tag{3.13}$$

3.3.1 Preliminaries

For determining the order of magnitude of $E(A_N)$, the following considerations are useful. Since the cardinality $P_G^{(j)} \cup P_A^{(j)}$ has about exponential growth, it is reasonable to expect $E(A_N)$ to be of the same order of magnitude as the additive energy of the last block $P_G^{(J)} \cup P_A^{(J)}$ that is fully contained in A_N — note that $J = J(N)$; i.e. to expect the magnitude of $E\big(P_G^{(J)} \cup P_A^{(J)}\big)$ which is roughly $E\big(P_A^{(J)}\big)$. The next proposition verifies this heuristic.

Proposition 3.1. *Let $(a_n)_n$ be as in the beginning of Section 3, and f be as in one of the two assertions in Theorem 3.2. Then,* $E(A_N) = \Theta\big(N^3(f(N))^{-3(\beta-\gamma)}\big)$.

For the proof of Proposition 3.1, we need the following technical lemma.

Lemma 3.5. *Let $F_j := 2^j\big(f(2^j)\big)^{-\delta}$, for $j \geq 1$ and fixed $\delta \in (0,1)$, where f is as in Proposition 3.1. Then, $\sum_{i\leq j} F_i = \mathcal{O}(F_j)$ and*

$$\sum_{d\in\mathbb{Z}}\Bigg(\sum_{j,i\leq J}\mathrm{rep}_{P_G^{(j)},P_A^{(i)}}(d)\Bigg)^2 = \mathcal{O}\left(J^6 2^{2J}\right).$$

Proof. Suppose that $f(x) = \mathcal{O}\big(x^{1/3}(\log x)^{-7/3}\big)$ is such that (3.4) diverges. Because

$$\sum_{j\leq J+1}\frac{1}{f(2^j)} \geq \sum_{k\leq 2^J}\frac{1}{kf(k)}$$

diverges as $J\to\infty$ and $\big(f(2^j)/f(2^{j+1})\big)_j$ is non-decreasing, we conclude that

$$\lim_{j\to\infty}\big(f(2^j)/f(2^{j+1})\big) = 1.$$

Therefore, there is an i_0 such that the estimate

$$\big(f(2^i)\big)^{-1}f\big(2^{i+h}\big) < (3/2)^{\frac{h}{\delta}}$$

holds for any $i \geq i_0$ and $h\in\mathbb{N}$. Hence,

$$\frac{1}{F_j}\sum_{i\leq j}F_i \leq o(1) + \sum_{i_0\leq i\leq j}2^{i-j}\left(\frac{3}{2}\right)^{j-i} = \mathcal{O}(1).$$

If f is such that (3.4) converges and $f(2x) \leq (2-\varepsilon)f(x)$ for x large enough, then we obtain by a similar argument that $\sum_{i\leq j}F_i$ is in $\mathcal{O}(F_j)$. Furthermore, $\mathrm{rep}_{P_G^{(j)},P_A^{(i)}}(d) = \mathcal{O}(i)$, for every $j\geq 1$, and non-vanishing for $\mathcal{O}(2^{2j})$ values of d which implies the last claim. □

We can now prove the proposition.

Proof of Proposition 3.1. Let $N \geq 1$ be large and denote by $J = J(N) \geq 0$ the greatest integer j such that $P_G^{(j-1)} \subseteq A_N$. By exploiting (3.12),

$$E\big(A_N\big) \geq E\big(P_A^{(J-1)}\big) \gg \big(\#P_A^{(J-1)}\big)^3$$

which is seen to be $\gg f\big(N\big)\big)^{-3(\beta-\gamma)} N^3$. Hence, it remains to show that

$$E\big(A_N\big) = \mathcal{O}\big(\big(f\big(N\big)\big)^{-3(\beta-\gamma)} N^3\big).$$

Note that

$$E\big(A_N\big) \leq \sum_{d\in\mathbb{Z}} \big(\mathrm{rep}_{A_{T_J}, A_{T_J}}(d)\big)^2 \quad \text{where} \quad T_J := \# \bigcup_{j\leq J} \Big(P_A^{(j)} \cup P_G^{(j)}\Big).$$

Moreover, $\mathrm{rep}_{A_{T_J}, A_{T_J}}(d) = S_1(d) + S_2(d)$ where $S_2(d)$ denotes the mixed sum

$$\sum_{i,j\leq J} \big(\mathrm{rep}_{P_A^{(j)}, P_G^{(i)}}(d) + \mathrm{rep}_{P_G^{(i)}, P_A^{(j)}}(d)\big),$$

and $S_1(d)$ abbreviates

$$\sum_{i,j\leq J} \big(\mathrm{rep}_{P_G^{(i)}, P_G^{(j)}}(d) + \mathrm{rep}_{P_A^{(i)}, P_A^{(j)}}(d)\big).$$

Using that for any $a, b \in \mathbb{R}$ the inequality $(a+b)^2 \leq 2\big(a^2+b^2\big)$ holds, we obtain

$$E\big(A_N\big) = \mathcal{O}\Bigg(\sum_{d\in\mathbb{Z}} \big(S_1(d)\big)^2 + \sum_{d\in\mathbb{Z}} \big(S_2(d)\big)^2\Bigg).$$

Lemma 3.5 implies that $\sum_{d\in\mathbb{Z}} \big(S_2(d)\big)^2 = \mathcal{O}\big((\log N)^6 N^2\big)$ due to $J = \mathcal{O}(\log N)$. Furthermore letting $F_j = 2^j \big(f\big(2^j\big)\big)^{-\beta}$, we observe that $\mathrm{rep}_{P_A^{(i)}, P_A^{(j)}}(d)$ is non-vanishing for at most $4F_J$ values of d as $i, j \leq J$. Since $\mathrm{rep}_{P_A^{(i)}, P_A^{(j)}}(d) \leq F_{\min(i,j)}$ holds, we deduce that

$$\sum_{i,j\leq J} \mathrm{rep}_{P_A^{(i)}, P_A^{(j)}}(d) = \mathcal{O}\Bigg(\sum_{j\leq J}\sum_{i\leq j} F_i\Bigg) = \mathcal{O}\big(F_J\big).$$

Since $\mathrm{rep}_{P_G^{(i)}, P_G^{(j)}}(d) \leq 1$, as $i, j \leq J$, is non-zero for at most $\mathcal{O}\big(T_J^2\big) = \mathcal{O}\big(N^2\big)$ values of d, we obtain that

$$\sum_{d\in\mathbb{Z}} \big(S_1(d)\big)^2 = \mathcal{O}\big(F_J^3 + (\log N)^6 N^2\big) = \mathcal{O}\big(N^3 \big(f\big(N\big)\big)^{-3(\beta-\gamma)}\big)$$

Hence, $E\big(A_N\big) = \mathcal{O}\big(N^3\big(f\big(N\big)\big)^{-3(\beta-\gamma)}\big)$. □

For estimating the measure or the Hausdorff dimension of $\mathrm{NPPC}\left(\left(a_n\right)_n\right)$ from below, we recall the notion of an optimal regular system. This notion, roughly speaking, describes sequences of real numbers that are exceptionally well distributed in any subinterval, in a uniform sense, of a fixed interval.

Definition. Let J be a bounded real interval, and $S = (\alpha_i)_i$ a sequence of distinct real numbers. S is called an *optimal regular system* in J if there exist constants c_1, c_2, $c_3 > 0$ — depending on S and J only — such that for any interval $I \subseteq J$ there is an index $Q_0 = Q_0(S, I)$ such that for any $Q \geq Q_0$ there are indices

$$c_1 Q \leq i_1 < i_2 < \ldots < i_t \leq Q \tag{3.14}$$

satisfying $\alpha_{i_h} \in I$ for $h = 1, \ldots, t$, and

$$|\alpha_{i_h} - \alpha_{i_\ell}| \geq \frac{c_2}{Q} \tag{3.15}$$

for $1 \leq h \neq \ell \leq t$, and

$$c_3 \lambda(I) Q \leq t \leq \lambda(I) Q. \tag{3.16}$$

Moreover, we need the following result(s) due to Beresnevich which may be thought of as a far reaching generalization of the classical Khintchine theorem, and the Jarník–Besicovitch theorem in Diophantine approximation.

Theorem 3.3 ([13, Thm. 6.1, Thm. 6.2])**.** *Suppose $\psi : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is a continuous, non-increasing function, and $S = \left(\alpha_i\right)_i$ an optimal regular system in $(0, 1)$. Let $\mathcal{K}_S(\psi)$ denote the set of ξ in $(0, 1)$ such that $|\xi - \alpha_i| < \psi(i)$ holds for infinitely many i. If*

$$\sum_{n \geq 1} \psi(n) \tag{3.17}$$

diverges, then $\mathcal{K}_S(\psi)$ has full measure.
Conversely, if (3.17) converges, then $\mathcal{K}_S(\psi)$ has measure zero and the Hausdorff dimension equals the reciprocal of the lower order of $\frac{1}{\psi}$ at infinity.

For a rational $\alpha = \frac{p}{q}$, where $p, q \in \mathbb{Z}$, $q \neq 0$, we denote by $H(\alpha)$ its (naive) height, i.e. $H(\alpha) := \max\{|p|, |q|\}$. It is well-known that the set of rational numbers in $(0, 1)$ — first running through all rationals of height 1 ordered by increasing numerical value, then through all rationals with height 2 ordered by increasing numerical value, and so on — gives rise to an optimal regular system in $(0, 1)$. The following lemma says, roughly speaking, that this assertion remains true for the set of rationals in $(0, 1)$ whose denominators are members of a special sequence that is not too sparse in the natural numbers, and hand-tailored for our purposes. The proof can be given by modifying the proof of the classical case, compare [13, Prop. 5.3]; however, we shall give the details for making this article more self-contained.

Lemma 3.6. *Let $\vartheta : \mathbb{R}_{>0} \to \mathbb{R}_{>1}$ be monotonically increasing to infinity with $\vartheta(x) = \mathcal{O}\left(x^{1/4}\right)$ and $\vartheta\left(2^{j+1}\right)/\vartheta\left(2^j\right) \to 1$ as $j \to \infty$. For each $j \in \mathbb{N}$, we let*

$$B_j := \frac{2^j}{f(2^j)\sqrt{\vartheta(2^j)}}, \qquad b_j := \frac{2}{3} B_j.$$

Let $S = \left(\alpha_i\right)_i$ denote a sequence running through all rationals in $(0, 1)$ whose denominators are in $M := \bigcup_{j \geq 1} \left\{n \in \mathbb{N} : b_j \leq n \leq B_j\right\}$ such that $i \mapsto H\left(\alpha_i\right)$ is non-decreasing. Then, S is an optimal regular system in $(0, 1)$.

Proof. Let $X \geq 2$. There are strictly less than $2X^2$ rational numbers in $(0,1)$ with height bounded by X. We take $J = J(X)$ to be the largest integer $j \geq 1$ such that $B_j \leq X$. Then, for X large enough, there are at least, due to a basic property of Euler's totient function,

$$\sum_{j \leq J} \sum_{b_j \leq q \leq B_j} \varphi(q) \geq \sum_{j \leq J} \left(\frac{1}{3\pi^2} B_j^2 + \mathcal{O}(B_j \log B_j) \right) \geq \frac{1}{6\pi^2} \frac{2^{2J}}{f^2(2^J)\,\vartheta(2^J)} + \mathcal{O}\left(J2^J\right) > \left(\frac{X}{5\pi}\right)^2$$

distinct such rationals in $(0,1)$ with height not exceeding X. Hence, we obtain

$$\frac{\sqrt{i}}{2} \leq H(\alpha_i) \leq \sqrt{25\pi^2(i+1)} + 1$$

for i sufficiently large. Let $Q \in \mathbb{N}$, $I \subseteq [0,1]$ be a non-empty interval, and let F denote the set of $\xi \in I$ satisfying the inequality $\|q\xi\| < Q^{-1}$ with some $1 \leq q \leq \frac{1}{1000}Q$. Note that F has measure at most

$$\sum_{q \leq \frac{1}{1000}Q} \left(\frac{2}{qQ} q\lambda(I) + \frac{2}{qQ} \right) = \frac{1}{500}\lambda(I) + \mathcal{O}\left(\frac{\log Q}{Q}\right) < \frac{1}{400}\lambda(I)$$

for $Q \geq Q_0$ where $Q_0 = Q_0(S,I)$ is sufficiently large. Let $\left\{{}^{p_j}/_{q_j}\right\}_{1 \leq j \leq t}$ be the set of all rationals ${}^{p_j}/_{q_j} \in (0,1)$ with $q_j \in M$, $\frac{1}{1000}Q < q_j < Q$ that satisfy

$$\left| \frac{p_j}{q_j} - \frac{p_{j'}}{q_{j'}} \right| > \frac{2000}{Q^2}$$

whenever $1 \leq j \neq j' \leq t$. Observe that for J as above with $X = Q$ sufficiently large, it follows that

$$\{q \in M : b_J \leq q \leq B_J\} \subseteq \left\{ \left\lfloor \frac{Q}{1000} \right\rfloor, \left\lfloor \frac{Q}{1000} \right\rfloor + 1, \ldots, Q \right\}$$

holds and hence, there are at least

$$\frac{1}{3\pi^2} B_J^2 + \mathcal{O}(B_J \log B_J) > \frac{1}{400} Q^2$$

choices of ${}^{p_j}/_{q_j} \in (0,1)$ with $q_j \in M$ and $\frac{1}{1000}Q < q_j < Q$. Due to $\lambda(I \setminus F) > \frac{399}{400}\lambda(I)$, we conclude

$$t \geq 400 \frac{Q^2}{4000} \frac{399}{400} \lambda(I).$$

Thus, taking $c_1 := {}^1/_{1000}$, $c_2 := 2000$, and $c_3 := \frac{399}{4000}$ in (3.14), (3.15) and (3.16), respectively, S is shown to be an optimal regular system. $\square$

Now we can proceed to the proof of Theorem 3.2.

3.3.2 Proof of Theorem 3.2

We argue in two steps depending on whether or not the series (3.4) converges. Proposition 3.1 implies the announced Θ-bounds on the additive energy of A_N, in both cases.

(i) Suppose (3.4) diverges, and fix $s > 0$. Let $\vartheta : \mathbb{R}_{>0} \to \mathbb{R}_{>1}$ be monotonically increasing to infinity with $\vartheta(x) = \mathcal{O}\left(x^{1/4}\right)$ such that

$$\psi(n) := \frac{1}{nf(n)\vartheta(n)} \tag{3.18}$$

satisfies the divergence condition (3.17). Thus, $\vartheta\left(2^j\right)/\vartheta\left(2^{j-1}\right) \to 1$ as $j \to \infty$, and $S = (\alpha_i)_i$ from Lemma 3.6 is an optimal regular system. Furthermore, if $b_J \leq n \leq B_J$, for some integer J, then, by the properties of ϑ from Lemma 3.6 and the relation $\sum_{j \leq J} F_j = \mathcal{O}(F_J)$ from Lemma 3.5, we conclude that

$$\sum_{j \leq J-1} \sum_{b_j \leq m \leq B_j} \varphi(m) = \Theta\big(B_J^2\big)$$

implies that $\alpha_i = {}^{m}\!/\!{}_{n}$ entails $i \geq cn^2$ where $c = c(f, \vartheta) > 0$ is a constant. Therefore, $\psi(i) \leq c^{-1}n^{-2}\big(f\big(cn^2\big)\vartheta\big(cn^2\big)\big)^{-1}$. The growth assumption on f and $\vartheta(x) = \mathcal{O}\big(x^{1/4}\big)$ yields that if j is large enough, then $b_j \leq n \leq B_j$ implies $cn^2 > 2^j$ and hence we obtain $\psi(i) \leq c^{-1}n^{-2}\big(f\big(2^j\big)\vartheta\big(2^j\big)\big)^{-1}$. Combining these considerations, we infer that

$$n\psi(i) = \mathcal{O}\big(2^{-j}\left(\vartheta\big(2^j\big)\right)^{-1/2}\big).$$

Applying Theorem 3.3 with ψ as in (3.18), implies that $\mathcal{K}_S(\psi)$ has full Lebesgue measure. Therefore, for any $\alpha \in \mathcal{K}_S(\psi)$ we get

$$\|n\alpha\| \leq n\,|\alpha - \alpha_i| = \mathcal{O}\Big(2^{-j}\left(\vartheta\big(2^j\big)\right)^{-1/2}\Big) \tag{3.19}$$

for infinitely many i and $j = j(i)$. Now if $b_j \leq n \leq B_j$ for j sufficiently large and n, α as in (3.19), then it follows that by taking any integer $m \leq \left(f\left(2^j\right)\right)^{\gamma}\left(\vartheta\big(2^j\big)\right)^{1/3}$ that also the multiples

$$nm \leq 2^j\left(f\left(2^j\right)\right)^{\gamma-1}\left(\vartheta\left(2^j\right)\right)^{-1/6}$$

satisfy that $\mathbf{1}_{[0,s/T_j]}(\|\alpha(mn)\|) = 1$ where $T_j = \mathcal{O}\big(2^j\left(f\left(2^j\right)\right)^{-\gamma}\big)$ is as in the proof of Proposition 3.1. If additionally $\gamma - 1 \leq -\beta$ holds, then we obtain that

$$\mathrm{rep}_{A_{T_j},A_{T_j}}(mn) \geq 2^{j-1}\left(f\left(2^j\right)\right)^{-\beta}$$

holds for j sufficiently large. By (3.13), we obtain

$$R\big(s, \alpha, T_j\big) \geq C\left(f\left(2^j\right)\right)^{2\gamma-\beta}\left(\vartheta\big(2^j\big)\right)^{1/3}$$

for infinitely many j where $C > 0$ is some constant. For the optimal choice of the parameters $\beta, \gamma > 0$, we are therefore led to maximise $\beta - \gamma$ where $2\gamma - \beta \geq 0$ and $\gamma - 1 \leq -\beta$ have to be satisfied. The solution is given if equality in the first inequality occurs, leading to $\beta = 2/3$ and $\gamma = 1/3$. Hence, (3.5) follows for $\alpha \in \mathcal{K}_S(\psi)$.

(ii) Suppose the series (3.4) converges. We keep the same sequence as in step (i) while taking $\vartheta(x) = 1 + \log(x)$, as we may. The arguments of step (i) show that any $\alpha \in \mathcal{K}_S(\psi)$ satisfies (3.5); now the conclusion is that $\mathcal{K}_S(\psi)$ has Hausdorff dimension at least equal to the reciprocal of

$$\liminf_{x\to\infty} \frac{-\log(\psi(x))}{\log x} = 1 + \liminf_{x\to\infty} \frac{\log f(x)}{\log x}.$$

Thus, the proof is complete.

Concluding remarks It should be possible to relax the growth restriction

$$f(x) = \mathcal{O}\big(x^{1/3}(\log(x))^{-7/3}\big)$$

in Theorem 3.2 on the expense of some additional technical work; as the main objective in this section was to get as close as possible to the Khintchine-type threshold for making progress on the Fundamental Question, we have not expended much effort in possible relaxations.

We would like to mention an open problem related to this article. It asks about how much the PPC property is violated for a sequence that has not metric PPC.

Problem. *Under which conditions on $(a_n)_n$ is it true that* NPPC$((a_n)_n)$ *having full Lebesgue measure implies that the set of $\alpha \in [0,1]$ such that for all $s > 0$*

$$\limsup_{N\to\infty} R(s, \alpha, N) = \infty$$

holds also has full Lebesgue measure?

Chapter 4

There is no Khintchine Threshold for Metric Pair Correlations

4.1 Introduction

Let $x_1, \dots, x_N$ be numbers in the unit interval. The distribution of the pair correlations of these numbers is described by the function

$$R(s,N) = \frac{1}{N} \{1 \le i \ne j \le N : \ \|x_i - x_j\| \le s/N\}, \qquad s \ge 0, \tag{4.1}$$

where $\| \cdot \|$ denotes the distance to the nearest integer. If for an infinite sequence $(x_n)_n$ we have

$$R(s,N) \to 2s \tag{4.2}$$

for all $s \ge 0$, then we say that the distribution of pair correlations is (asymptotically) *Poissonian.* Note that a sequence of independent, identically distributed (i.i.d.) random points, picked from a uniform distribution on $[0,1]$, almost surely has Poissonian pair correlations. The term "Poissonian" comes from a similarity with the distribution of the spacings between points in a Poisson process, which, however, only becomes really meaningful when also considering higher correlations (triple, quadruple etc.) or so-called level spacings (which are in general much more difficult to handle than pair correlations).

The interest in such problems goes back to a paper of Berry and Tabor [8], where they gave a conjectural framework for the distribution of energy spectra of integrable quantum systems (see [32] for a survey). Their model led to strong mathematical interest in distributional properties of spacing of sequences such as $(n\alpha)_n$ mod 1 (corresponding to the "harmonic oscillator") and $(n^2\alpha)_n$ mod 1 (corresponding to the "boxed oscillator"). The case of $(n\alpha)_n$ is easier to analyse; one can use considerations based on continued fractions to show that the pair correlations of this sequence cannot be Poissonian for any α, since for some N the initial segment $(\alpha, 2\alpha, \dots, N\alpha)$ mod 1 is too regularly spaced. The case of $(n^2\alpha)_n$ is much harder and is far from being well-understood. It is conjectured that the pair correlations for this sequence should be Poissonian, unless α is very well approximable by rationals; however, there exist only some partial results in this direction (see for example [24, 38, 43]). From the metric perspective, the situation is easier: it is known that the pair correlations of $(n^2\alpha)_n$ mod 1 are Poissonian for almost all α, in the sense of Lebesgue measure. The same is true if $(n^2)_n$ is replaced by $(n^d)_n$ for some integer $d \geq 3$, or by an exponentially growing sequence $(a_n)_n$ of integers, see [37, 40]. We denote this property by saying that these sequences have the *metric pair correlation property.* In a recent paper [4], a connection was established between the question whether a sequence has the metric pair correlation property, and the asymptotic order of its so-called additive energy. Let $(a_n)_n$ be a sequence of distinct positive integers, let A_N denote its initial segment $a_1, \dots, a_N$, and denote by $E(A_N)$ the *additive energy* of A_N, which is defined as

$$E(A_N) = \#\{n_1, n_2, n_3, n_4 \leq N : \ a_{n_1} + a_{n_2} = a_{n_3} + a_{n_4}\}. \tag{4.3}$$

Trivially, the additive energy is always between N^2 and N^3. Throughout this chapter we will use the formulation "the order of the additive energy of a sequence", when more precisely speaking we mean the order (as a function of N) of the additive energy of the N first elements of the sequence.

The main results of [4] say that a sequence has the metric pair correlation property if its additive energy is of order at most $N^{3-\varepsilon}$ for some $\varepsilon > 0$, while it does not have the metric pair correlation property if the additive energy exceeds cN^3 for infinitely many N for some constant $c > 0$. This fits together very well with the examples from above, since sequences of the form $(n^d)_n$ for $d \geq 2$ and lacunary sequences are known to have very small additive energy, while the additive energy of the sequence $a_n = n, \ n \geq 1$, is of the maximal possible order.

So the general philosophy is that a sequence has the metric pair correlation property as soon as its additive energy is slightly below the maximal possible order. However, a precise threshold is not known. Some results in this direction are:

- The primes do not have the metric pair correlation property, as shown by Walker [45]. The additive energy of the sequence of primes is roughly of order $\frac{N^3}{\log N}$.
- There exists a sequence having additive energy of order $\frac{N^3}{\log N \log\log N}$ which does not have the metric pair correlation property, as seen in Chapter 3.
- For every $\varepsilon > 0$ there exists a sequence having additive energy of order $\frac{N^3}{\log N (\log\log N)^{1+\varepsilon}}$ which has the metric pair correlation property (unpublished, but not difficult to construct using methods from Chapter 3 or [9]).

These results indicate that there is a sort of transitional behaviour when the additive energy lies around the "critical" order of roughly $\frac{N^3}{\log N \log\log N}$. The methods used in Chapter 3 and [9] indicate a close connection between this sort of question and problems from metric Diophantine approximation, where the classical theorem of Khintchine gives a zero-one law in terms of the convergence/divergence of the series of measures of the target intervals (see for example [21]). It is tempting to speculate that a similar convergence/divergence criterion might also exist for the metric theory of pair correlations, where the crucial quantity is the additive energy of $(a_n)_n$. This idea was discussed in a recent paper of Bloom, Chow, Gafni, and Walker [9], where they noted that there “appears to be reasonable evidence to speculate a sharp Khintchine-type threshold, that is, to speculate that the metric Poissonian property should be completely determined by whether or not a certain sum of additive energies is convergent or divergent”. They raise the following problem, which they call the “Fundamental Question”:

> Is it true that if $E(A_N) \sim N^3\psi(N)$, for some weakly decreasing function $\psi : \mathbb{Z}_{\geq 1} \to [0,1]$, then $(a_n)_n$ is metric Poissonian if and only if
>
> $$\sum_{N \geq 1} \psi(N)/N \tag{4.4}$$
>
> converges?

In the present chapter, we show that the answer to the question above is negative, and that the metric pair correlation property cannot be fully characterised in terms of the additive energy alone. For this purpose, we construct a sequence $(a_n)_n$ whose additive energy is of order roughly $N^3/(\log N)^{3/4}$, and which *does* have the metric pair correlation property. More precisely, we prove the following theorem.

Theorem 4.1. *For every $\varepsilon > 0$ there exists a strictly increasing sequence $(a_n)_n$ of positive integers which has the metric pair correlation property, and whose additive energy satisfies*

$$E(A_N) \gg \frac{N^3}{(\log N)^{3/4+\varepsilon}}. \tag{4.5}$$

Note that the additive energy of the sequence from the conclusion of Theorem 4.1 is significantly larger than the putative threshold, which is rather around $N^3/\log N$. Furthermore, as the examples above showed, the additive energy of a sequence which does *not* have the metric pair correlation property can be of asymptotic order N^3, but it can also be of asymptotic order roughly $N^3/\log N$. Thus the metric theory of pair correlations cannot simply be reduced to a convergence/divergence criterion in terms of the additive energy alone. Instead, the picture is more complicated and looks as follows:

- If the additive energy is below a certain threshold, then the sequence does have the metric pair correlation property.
- If the additive energy is above a certain threshold (for infinitely many N), then the sequence cannot have the metric pair correlation property. (This threshold is different from the one in the point above.)
- Between these upper and lower thresholds there is a transition zone, where knowing the additive energy alone is not sufficient to determine the metric pair correlation behaviour of the sequence. Thus, in this range the metric pair correlation property is determined by some additional arithmetic properties of the sequence.

We note that while our result says that the metric pair correlation property cannot be characterised in terms of the additive energy alone, it leaves the problem of finding some other way of characterising the metric pair correlation property in terms of some arithmetic properties of $(a_n)_n$. It is likely that there is a zero-one law in the metric theory of pair correlations, but actually even this is not known. Also, our result still leaves questions concerning the quantitative connection between additive energy and the metric theory of pair correlations. For example, is it possible that a sequence having additive energy of order $N^3/(\log \log N)$ also has the metric pair correlation property? In the other direction, is it possible that the additive energy is of order $N^3/(\log N)^2$ and the sequence does not have the metric pair correlation property?[1] Closing the gaps in our knowledge in this field would be very desirable, as phenomena from both additive combinatorics and Diophantine approximation seem to be at work here.

4.2 Preliminaries

4.2.1 Construction of the Sequence

We will construct our sequence $(a_n)_n$ as the concatenation of successive "blocks". All these blocks are either finite arithmetic or finite geometric progressions. The geometric blocks will contain the majority of the numbers which constitute the final sequence, but they will not be responsible for making the additive energy of the final sequence large, since geometric progressions always have small additive energy. The contribution of the geometric blocks to the distribution of pair correlations will be "random", in accordance with the well-known heuristics that lacunary systems exhibit properties which are also shown by independent random systems (see for example [40] for this phenomenon in connection with pair correlations, and [2] for the wider context). In our context "random" behaviour means Poissonian behaviour of the pair correlations, so the geometric blocks are responsible that the final sequence is metric Poissonian. The arithmetic blocks contain only a minority of all the elements of the final sequence, while being responsible for making the additive energy large. The main task will be to show that while these arithmetic blocks boost the additive energy, their contribution to the distribution of pair correlations is asymptotically negligible. To control the contribution of arithmetic blocks we will use tools from metric Diophantine approximation.

The key point of the construction lies in the fact that the arithmetic blocks which are used in the construction have different prime numbers as their step sizes.[2] The fact that the step sizes of the arithmetic blocks are prime numbers will play a key role in two parts of the proof. On the one hand, using the theory of continued fractions at some point we will be led to counting the number of solutions of a certain equation; the assumption that the step size is a (large) prime will imply that we only have to count solutions which are a multiple of that prime, thus effectively reducing the number of solutions. This will allow us to control the contribution which comes from elements contained within one and the same arithmetic block. On the other hand, to control the contribution of the interaction of elements from two different arithmetic blocks, we will use a variance estimate which boils down to counting

[1] While the present chapter was being refereed as submitted paper, a paper of Bloom and Walker addressing this question appeared on the arXiv. They proved that there exists an (unspecified) constant $C > 1$ such that a sequence has the metric pair correlations property whenever its additive energy is of asymptotic order at most $N^3/(\log N)^C$, see [10].

[2] We will need to use a "recycling process" for the prime moduli, since there are not enough different primes of the appropriate size available to have a different step size for each arithmetic block.

the maximal number of solutions of a simple Diophantine equation. Again, the fact that the moduli are (different) primes will reduce the maximal number of solutions of the equation. We will add some further comments on the heuristic reasoning behind the proof after first formulating precisely the way in which our sequence is constructed.

Notation. We fix some $\varepsilon > 0$. Throughout the rest of this chapter we assume w.l.o.g. that ε is "small", say $\varepsilon < 1/100$. We will use Landau notation o, $\mathcal{O}$, and Vinogradov symbols $\ll, \gg$, with their usual meaning in analytic number theory (i.e. $f \ll g$ meaning that $|f|$ is bounded by a constant times $|g|$, for all possible arguments). The symbol $f \asymp g$ means that $f \ll g$ as well as $f \gg g$. If the implied constant depends on some parameter, we will indicate the dependence by a corresponding subscript. However, we will not indicate any dependence on ε, since throughout the proof ε is assumed to be fixed. We write λ for the Lebesgue measure on $\mathbb{R}$. Finally, we write $\lfloor \cdot \rfloor$ for the integer part of a real number.

In Lemma 4.1 below we construct the moduli of the arithmetic blocks.

Lemma 4.1. *There exist an index $j_0 \geq 1$ and a sequence $(m_j)_{j \geq j_0}$ of primes such that*

$$m_j \asymp j^{1/4}, \tag{4.6}$$

and such that $m_j \neq m_i$ whenever $j - 3 \log j < i < j$, for all $i, j \geq j_0$.

Before proving the lemma, we briefly explain its meaning. The numbers m_j will be the moduli of the arithmetic progressions in our construction. The first condition in the lemma says that these moduli are of asymptotic order roughly $j^{1/4}$. The second condition guarantees that the step size of the j-th arithmetic progression is different from the step size of the i-th arithmetic progression, whenever i is "close" to j. So arithmetic blocks whose indices are close by can never have the same step size, which will guarantee that there is no undesired interaction between such blocks (this will play a crucial role in the proof of Lemma 4.7 below). On the other hand, if i and j are not close to each other, then the corresponding arithmetic blocks are allowed to have the same step size — this is the "recycling process", which was mentioned in the preceding footnote, and which is necessary since the step sizes of the blocks grow more slowly than the indices of the blocks themselves. However, this will not cause any problems since the block sizes increase very quickly and any interaction between a block and some other block of much smaller cardinality will always be negligible.

Proof of Lemma 4.1. To define the value of m_j for all indices j in the range $16^d \leq j < 16^{d+1}$, $d \geq 0$, we note that the number of primes in the range

$$\left(2^d, 2^{d+1}\right) \tag{4.7}$$

is certainly at least d^2 for all sufficiently large d, by the prime number theorem. So assume that d is "large", and let $p_{d,1} < \ldots < p_{d,d^2}$ denote the first d^2 primes in the interval (4.7). We set

$$m_j := p_{d,r(j)}, \qquad 16^d \leq j < 16^{d+1}, \tag{4.8}$$

where $r(j)$ is the unique remainder when reducing j mod d^2. Then (4.6) holds since $2^d = (16^d)^{1/4} \asymp j^{\frac{1}{4}}$. Furthermore, it can easily be seen that the second assertion of the lemma holds as well for sufficiently large d, since (4.8) guarantees that m_j cannot equal m_i whenever $|i-j|$ is small. Observe here that d^2 is of order roughly $(\log j)^2$, and thus much larger than $3\log j$ for all sufficiently large j and d. $\square$

Let j_0 be as in Lemma 4.1. For $j \geq j_0$ we recursively define sets $P_G(j)$ and $P_A(j)$ by setting

$$P_G(j) := \left\{ 2^{(\max P_A(j-1))} + 3^{j^h} : h = 0, \dots, 2^j - 1 \right\}, \tag{4.9}$$

$$P_A(j) := \left\{ 2^{(\max P_G(j))} + m_j h : h = 0, \dots, \left\lfloor 2^j / j^{1/4+\varepsilon/3} \right\rfloor \right\}. \tag{4.10}$$

To make the construction well-defined we need to specify the initial value $\max P_A(j_0-1)$, which is necessary for (4.9) in the case $j = j_0$; it does not matter what we choose, but let us agree that this quantity should be read as 1, and that accordingly $P_G(j_0) := \left\{ 2 + 3^{j_0^h} : h = 0, \dots, 2^{j_0} - 1 \right\}$.

The set $P_G(j)$ is a (shifted) geometric progression for each j, while the set $P_A(j)$ is a (shifted) arithmetic progression for each j. The sets $P_G(j)$ and $P_A(j)$ are arranged in increasing order; more precisely, we have

$$P_G(j) < P_A(j) < P_G(j+1) \tag{4.11}$$

for all $j \geq j_0$, where the symbol "$<$" means that every element of the set on the right side exceeds every element of the set on the left side.

The exponential factors $2^{(\cdots)}$ which appear in the definitions of all the sets P_G and P_A are quite arbitrary; what matters is only that the smallest element of $P_A(j)$ is much larger than the largest elements of $P_G(j)$, and so on. Thus the respective sets in our construction are not only ordered as shown by (4.11), but there actually are huge gaps separating one item in this chain of inequalities from the next.

Finally, we specify the sequence $(a_n)_n$ by defining $a_n = a_n(\varepsilon)$ as the n-th (smallest) element of

$$\bigcup_{j \geq j_0} \big(P_G(j) \cup P_A(j) \big),$$

for all $n \geq 1$. So $(a_n)_n$ contains all the numbers which are contained in $P_G(j)$ or $P_A(j)$ for some j, sorted in increasing order. We claim that the additive energy of this sequence is as large as specified in (4.5), and that the pair correlations of $(a_n \alpha)_n$ mod 1 are Poissonian for almost all α.

4.2.2 The Heuristic behind the Construction

Before turning to the proof of Theorem 4.1, we want to explain the heuristic behind the construction of the sequence $(a_n)_n$. In particular, we want to show why our construction allows to go beyond the alleged "Khintchine threshold". Note that the distribution of the pair correlations of $(a_n\alpha)_n$ mod 1 depends not so much on the sequence $(a_n)_n$ itself, but rather on the set of differences $\{a_n - a_m\}_{m,n}$ (as does the additive energy). Thus it is this difference set that has to be controlled.

Obviously the difference set of a finite arithmetic progression has a very special structure; it is essentially an arithmetic progression itself, and the cardinality of the difference set of an arithmetic progression is small while the additive energy is large. More precisely, the positive part of the difference set of an arithmetic progression with step size d and length M is itself an arithmetic progression with step size d, and length $M-1$, and each of the elements of the difference sets has at least 1 and at most $M-1$ representations as a difference of elements of the original set. In our construction we combine arithmetic progressions with different prime moduli m_j. The number of such arithmetic progressions and their respective length is so large that they boost the additive energy of the total sequence; in contrast, we have to show that their contribution to the distribution of pair correlations is asymptotically negligible. In our setting, at the j-th building block we have constructed roughly $N \approx 2^j$ elements of our sequence $(a_n)_n$. Each arithmetic progression at this level consists of roughly $\approx N/(\log N)^{1/4+\varepsilon/3}$ elements. One can easily check that this leads to the required lower bound for the additive energy. The size of the prime moduli m_j is roughly $(\log N)^{1/4}$.

To make sure that the contribution which one of these arithmetic progressions makes to the pair correlations is asymptotically negligible, we have to show (roughly speaking) that

$$\frac{N}{(\log N)^{1/4+\varepsilon/3}} \cdot \#\left\{q \leq N/(\log N)^{1/4+\varepsilon/3} :\ \|m_j q\alpha\| \leq \frac{1}{N}\right\} = o(N) \tag{4.12}$$

for "typical" α in the sense of Lebesgue measure (where for simplicity we took $1/N$ rather than s/N for the length of the test interval). Here the factor $N/(\log N)^{1/4+\varepsilon/3}$ on the very left arises as the maximal number of representations which the number $m_j q$ has as a difference of two elements of the arithmetic progression with step size m_j, and the upper bound $q \leq N/(\log N)^{1/4+\varepsilon/3}$ which restricts the maximal size of q comes from the length of the arithmetic blocks. The estimate (4.12) is true as long as the cardinality of the set on the left is asymptotically negligible in comparison to $(\log N)^{1/4+\varepsilon/3}$.

Essentially, the cardinality of this set only exceeds $(\log N)^{1/4+\varepsilon/3}$ when there is a

$$q \in \{1, \dots, N/(\log N)^{1/2+2\varepsilon/3}\}$$

such that $\|m_j q\alpha\|$ is less than $1/(N(\log N)^{1/4+\varepsilon/3})$, so that for the next $(\log N)^{1/4+\varepsilon/3}$ multiples of q we also have $\|m_j q\alpha\| \leq 1/N$, and so that all these multiples are still smaller than $N/(\log N)^{1/4+\varepsilon}$. Accordingly, one has to check if for typical α one should expect that there is a q such that

$$\|m_j q\alpha\| \leq \frac{1}{N(\log N)^{1/4+\varepsilon/3}}, \qquad 1 \leq q \leq N/(\log N)^{1/2+2\varepsilon/3}.$$

Writing $N/(\log N)^{1/2+2\varepsilon/3} =: Q$, the inequality above essentially becomes

$$\|m_j q\alpha\| \leq \frac{1}{Q(\log Q)^{3/4+\varepsilon}}, \qquad 1 \leq q \leq Q. \tag{4.13}$$

By Khintchine's convergence/divergence criterion this inequality looks like it should have infinitely many solutions for "typical" alpha, since the expression on the right-hand side is not summable as a function of Q. However, one major aspect is missing. The right-hand side of (4.13) is so small that the solutions of this inequality can be explicitly characterised by continued fraction theory; all solutions necessarily come from best approximations to α. We will show that (for typical alpha) we may assume that the denominators of best approximations to alpha are not divisible by the prime m_j; thus the number $m_j q$ cannot be a best approximation denominator itself. Rather, it must be the multiple of m_j and of a best approximation denominator, and accordingly for q itself to satisfy (4.13) we must have

$$\|q\alpha\| = \frac{\|m_j q\alpha\|}{m_j} \leq \frac{1}{m_j Q(\log Q)^{3/4+\varepsilon}} \approx \frac{1}{Q(\log Q)^{1+\varepsilon}}, \qquad 1 \leq q \leq Q. \tag{4.14}$$

The right-hand side of (4.14) is summable as a function of Q, and thus by Khintchine's criterion we should only expect finitely many solutions for typical alpha. It turns out that this heuristic reasoning can be turned into an actual proof.

We emphasise again that the fact that m_j always is a prime played a crucial role in this reasoning, together with the fact that we may assume that the denominators of best approximations are not divisible by m_j (we will prove this fact in Lemma 4.2 below). Another crucial aspect is to show that two different arithmetic blocks do not "interact" in an undesired way; that is, we have to show that the difference sets of these respective progressions do not overlap too much. For this it will again be important that all the moduli are (different) primes, since then a fixed integer can only show up in the difference set of two arithmetic progressions if it is a product of the two primes which constitute the respective step sizes. This will be proved in the form of a variance bound in Lemma 4.7.

Finally, let us remark why it is not possible to obtain even larger additive energy with such a construction. Obviously, the additive energy is increased when the length of the arithmetic blocks is increased, so we might try to do that. Furthermore, as (4.14) shows, increasing the size of the prime moduli m_j would also improve the argument, so we might try to do that as well. So let us assume that the length of the arithmetic blocks is changed from roughly $N/(\log N)^{1/4+\varepsilon/3}$ to $N/(\log N)^{\beta}$ for some β, and that the size of the prime moduli m_j is changed from roughly $(\log N)^{1/4}$ to $(\log N)^{\gamma}$ for some γ. If we do so, then instead of (4.12) we will have to show that

$$\frac{N}{(\log N)^{\beta}} \# \left\{ q \leq N/(\log N)^{\beta} : \ \|m_j q\alpha\| \leq \frac{1}{N} \right\} = o(N) \tag{4.15}$$

for "typical" alpha, with m_j of size roughly $m_j \approx (\log N)^{\gamma}$. Now recall that Legendre's theorem from continued fraction theory allows us to characterise the solutions (a, b) to $|b\alpha - a| < 1/(2b)$. We want to use this for $b = m_j q$, and thus in our application b might be as large as $N(\log N)^{\gamma-\beta}$. The term $1/N$ in (4.15) is preassigned, since it comes from the definition of pair correlations. So in order to apply Legendre's theorem we have to make sure that $N(\log N)^{\gamma-\beta} \ll N$, which implies $\gamma \leq \beta$. This restricts the size of the prime moduli (in terms of the length of the arithmetic progressions). When we carry out the heuristic reasoning above with general parameters (β, γ) instead of $(1/4 + \varepsilon/3, 1/4)$, then instead of (4.14) we will arrive at

$$\|q\alpha\| \leq \frac{1}{Q(\log Q)^{3\beta+\gamma}}, \qquad 1 \leq q \leq Q. \tag{4.16}$$

The right-hand side is summable if $3\beta + \gamma > 1$. Since the additive energy is maximised by taking β as small as possible, and since we already know that we need to take $\gamma \leq \beta$, the minimal permissible value for β is restricted by the requirement $\beta > 1/4$. This is the choice of parameters which is made in our construction. One can also show that our choice of parameters is optimal with respect to the conditions imposed by the variance bound in Lemma 4.7, which also requires that $3\beta + \gamma > 1$. Thus some significant new ideas would be necessary to further increase the additive energy while preserving the metric pair correlation property.

4.2.3 A Useful Partition, and Organisation of the Chapter

The following partitioning underpins the remaining part of this chapter. For doing so, we need some notation from additive combinatorics: We write $X - Y$ for the difference set

$$X - Y := \{x - y : x \in X,\, y \in Y\}$$

of two sets $X, Y \subseteq \mathbb{Z}$. By $\#X$ we denote the cardinality of X. Furthermore, we write r_{X-Y} for the number of ways in which $d \in \mathbb{Z}$ can be represented as a difference of elements of $X, Y \subseteq \mathbb{Z}$, that is,

$$\mathrm{r}_{X-Y}(d) := \#\{(x, y) \in X \times Y : d = x - y\}. \tag{4.17}$$

If no confusion can arise, we will simply write $\mathrm{r}(d)$ for $\mathrm{r}_{X-Y}(d)$. Recall that trivially

$$\mathrm{r}_{X-Y}(d) \leq \min\{\#X, \#Y\}.$$

Moreover, let $X^+ := X \cap \mathbb{Z}_{\geq 1}$ denote the set of positive elements of a set $X \subseteq \mathbb{Z}$. Since $A_N - A_N$ is symmetric around the origin, we can confine attention to its positive part.

Assume that $d \geq 0$ is the difference of two elements of A_N, that is, $d = x - y$. We will classify these differences, according to the origin of x and y. More precisely, we will distinguish between the following cases.

- Case (GG): x and y both come from geometric blocks, that is, $x, y \in \bigcup_j P_G(j)$.
- Case (AG): x comes from an arithmetic, and y comes from a geometric block, that is, $x \in \bigcup_j P_A(j)$ and $y \in \bigcup_j P_G(j)$. Or, reciprocally, x comes from a geometric block and y comes from an arithmetic block.
- Case ($\mathrm{AA}_{\mathrm{diff}}$): x and y come from *different* arithmetic blocks, that is, $x \in P_A(j_1)$ for some j_1 and $y \in P_A(j_2)$ for some j_2, such that $j_1 \neq j_2$.
- Case ($\mathrm{AA}_{\mathrm{same}}$): x and y come from *the same* arithmetic block, that is, $x, y \in P_A(j)$ for some j.

We write $\mathcal{D}_N(GG)$ for the set of those d in the difference set $(A_N - A_N)^+$ which can be represented as Case (GG). In a similar way, we define $\mathcal{D}_N(AG), \mathcal{D}_N(AA_{\text{diff}})$, and $\mathcal{D}_N(AA_{\text{same}})$.

The function R which was defined in (4.1) can be decomposed in a similar way in the form

$$R = R(GG) + R(AG) + R(AA_{\text{diff}}) + R(AA_{\text{diff}}). \tag{4.18}$$

For this decomposition, we set

$$R(GG) := R(GG, \alpha, s, N) := \frac{2}{N} \sum_{d \in \mathcal{D}_N(GG)} \mathrm{r}(d)\, I_{s,N}(d\alpha), \quad I_{s,N}(x) := \begin{cases} 1 & \|x\| \le s/N \\ 0 & \text{otherwise}, \end{cases} \tag{4.19}$$

where $r(d)$ counts only the number of Case (GG) representations which $d \geq 1$ has in the form $d = x - y$ such that $x, y \in A_N$. Note that the factor 2 in (4.19), which is not present in (4.1), comes from the fact that we restricted ourselves to the positive part of the difference set $A_N - A_N$. Similarly, we define $R(AG)$, $R(AA_{\text{diff}})$ and $R(AA_{\text{same}})$, where the function $r(d)$ is instead restricted to representations of d as Case (AG), Case (AA_{diff}), and Case (AA_{same}), respectively.

By using the same methods as in [4], one can easily conclude that

$$R(GG, \alpha, s, N) \to 2s \tag{4.20}$$

as $N \to \infty$, for almost all $\alpha \in [0, 1]$ and each $s > 0$. This follows from the fact that geometric progressions have small additive energy, and the fact that the cardinality of the geometric blocks is dominant over the total cardinality of the arithmetic blocks which implies that $1/N$ really is the correct normalisation factor such that $R(GG)$ converges as desired for $N \to \infty$.

Thus it remains to show that all the remaining terms $R(AG), R(AA_{\text{diff}})$ and $R(AA_{\text{same}})$ vanish in the limit $N \to \infty$, for almost all α.

The outline of the next sections is as follows. First, in Section 4.3, we analyse the contribution of $R(AA_{\text{same}})$. Here Diophantine approximation determines the counting.[3] Then, in Section 4.4, we prove variance estimates to control $R(AG)$ and $R(AA_{\text{diff}})$. Once these steps are completed, in Section 4.5 we use the Borel–Cantelli lemma with a sandwiching argument to finish the proof of Theorem 4.1.

[3] The mechanism furnishing these estimates is of a somewhat combinatorial nature, and related to so-called *Bohr sets*. The combinatorial nature of these sets also plays a key role in a recent paper of Chow, cf. [14].

4.3 Analysing the Contribution of the Small Differences

Before proceeding further, we need to recall some notions and results about continued fractions. For a (possibly finite) sequence $(\alpha_i)_i$ of strictly positive integers, we denote by

$$\alpha := [\alpha_1, \alpha_2, \ldots] = \cfrac{1}{\alpha_1 + \cfrac{1}{\alpha_2 + \cfrac{1}{\ddots}}}$$

the associated (possibly finite) continued fraction in the unit interval $[0, 1]$. Moreover, let p_n/q_n denote the n-th convergent to α. Then, the following are well-known facts, cf. for instance [13, Ch.1].

1. Legendre's theorem: If a/b is a fraction with
$$\left|\alpha - \frac{a}{b}\right| < \frac{1}{2b^2},$$
then a/b is a convergent to α.

2. We have
$$\left|\alpha - \frac{p_n}{q_n}\right| \asymp \frac{1}{\alpha_n q_n^2}, \tag{4.21}$$
where the implied constants are independent of α.

3. Borel–Bernstein theorem: Let $B := (b_n)_n$ be a sequence of (strictly) positive real numbers, and consider the series
$$\sum_{n\geq 1} \frac{1}{b_n}. \tag{4.22}$$
If $V_B \subset [0, 1]$ denotes the set of numbers $\alpha = [\alpha_1, \alpha_2, \ldots]$ satisfying $\alpha_n \leq b_n$ for all sufficiently large $n \geq 1$, then
$$\lambda(V_B) = \begin{cases} 1 & \text{if (4.22) converges,} \\ 0 & \text{if (4.22) diverges.} \end{cases}$$

Lemma 4.2. *Let $(m_j)_{j\geq 1}$ be the sequence of primes from Lemma 4.1, which was used in* (4.10) *for the construction of our sequence. Then for almost all $\alpha \in [0, 1]$ there exist only finitely many pairs of indices (j, n) such that the prime m_j divides q_n, and such that additionally $q_n/m_j \in [2^j/j^2, 2^j]$, where q_n is the denominator of a convergent to α.*

Proof. Assume that the denominator q_n of a convergent is divisible by a prime m_j, i.e. there is a k such that $q_n = km_j$. When q_n is a convergent to α then $\|q_n\alpha\| \leq 1/q_n$, and thus $\|km_j\alpha\| \leq 1/(km_j)$. Thus to prove the lemma we have to show that almost all $\alpha \in [0, 1]$ are contained in at most finitely many sets of the form

$$S_{j,k} := \left\{x \in [0, 1] : \|km_j x\| \leq \frac{1}{km_j}\right\}, \qquad j = 1, 2, \ldots, \quad 2^j/j^2 \leq k \leq 2^j.$$

We have

$$\lambda(S_{j,k}) = \frac{2}{km_j}.$$

Furthermore, we have

$$\sum_{j=1}^{\infty} \sum_{2^j/j^2 \leq k \leq 2^j} \frac{2}{km_j} \ll \sum_{j=1}^{\infty} \frac{\log j}{m_j}.$$

Recall that to construct our sequence $(m_j)_{j\geq 1}$ in Lemma 4.1 we selected d^2 primes from the range $\left(2^d, 2^{d+1}\right)$, for each (sufficiently large) d. Thus

$$\sum_{j=1}^{\infty} \frac{\log j}{m_j} \ll \sum_{d} \frac{(\log\log d) d^2}{2^d} < \infty.$$

Thus, by the Borel–Cantelli lemma, almost all α are contained in only finitely many sets $S_{j,k}$. □

Lemma 4.3. *Let*

$$M_j := \left\{ q \leq 2^j/j^{1/4+\varepsilon/3} : \left\| m_j q \alpha \right\| \leq s/2^j \right\}.$$

Then for almost all $\alpha \in [0,1]$ we have

$$\# M_j \ll_s j^{1/4}.$$

Proof. During this proof we suppress the potential dependence of the symbols "$\ll$" and "$\gg$" on s.

Let B denote the sequence $(n^{1+\varepsilon/3})_n$, and suppose that $\alpha \in V_B$ is an irrational number (recall that the set V_B was defined in the statement of the Borel–Bernstein theorem, before the statement of Lemma 4.2). By the Borel–Bernstein theorem, V_B has full Lebesgue measure. In the sequel we will assume that $\alpha \in [0,1]$ is a fixed number which is contained in V_B, and for which the conclusion of Lemma 4.2 holds. Note that the set of such α's has full Lebesgue measure.

Let us note the following. Let q_m be the denominator of a convergent to α. Assume that

$$\|q_m \alpha\| \leq \frac{s}{2^j}. \tag{4.23}$$

Then, as noted above, we have

$$\|q_m \alpha\| \asymp \frac{1}{\alpha_m q_m}.$$

Since q_m grows at least exponentially in m, and since $\alpha \in V_B$ implies that $\alpha_m \ll m^{1+\varepsilon/3} \ll (\log q_m)^{1+\varepsilon/3}$, we thus see that (4.23) is only possible if

$$\frac{1}{(\log q_m)^{1+\varepsilon/3} q_m} \ll \frac{1}{2^j},$$

which in turn is only possible if

$$q_m \gg \frac{2^j}{j^{1+\varepsilon/3}}.$$

Now we argue in two steps.
(i) We first claim the following. If j is large enough, and if M_j is non-empty, then there exists a unique value of n such that q_n is the denominator of a convergent to α, such that $q_n \geq 2^j/j^2$, and such that

$$M_j \subseteq q_n\mathbb{Z}. \tag{4.24}$$

Indeed, if some q is contained in M_j, then for this q we have $\|m_j q\alpha\| \leq s/2^j < 1/(2m_j q)$, if j is sufficiently large. This is a consequence of our construction, where we have $m_j \asymp j^{1/4}$ and $q \leq 2^j/j^{1/4+\varepsilon/3}$. Let $p \in \mathbb{Z}$ be such that $\|m_j q\alpha\| = |m_j q\alpha - p|$. Then Legendre's theorem implies that there is some $n \geq 1$ with

$$\frac{p}{m_j q} = \frac{p_n}{q_n}. \tag{4.25}$$

As a consequence, since p_n and q_n are coprime, there is some integer $g \geq 1$ such that $m_j q = g q_n$ and $p = g p_n$. Then we have $\|m_j q\alpha\| = |g q_n \alpha - g p_n| = g|q_n\alpha - p_n|$, and from the reasoning following equation (4.23) we can deduce that $q_n \gg 2^j/j^{1+\varepsilon/3}$. Thus, provided that j is sufficiently large, q_n/m_j lies in the range $[2^j/j^2, 2^j]$, and then, by Lemma 4.2, we can assume that q_n is not divisible by m_j.

Since we have now figured out that we may assume that m_j does not divide q_n, we conclude that p and q can actually both be written in the form $p = h m_j p_n$ and $q = h m_j q_n$ for some integer $h \geq 1$. Observe that (4.21) implies

$$\frac{m_j h}{\alpha_n q_n} \asymp \|m_j q\alpha\| \leq \frac{s}{2^j}, \tag{4.26}$$

and thus

$$\alpha_n q_n \gg m_j h 2^j \gg j^{1/4} 2^j.$$

Thus the well-known recursion $q_{n+1} = \alpha_n q_n + q_{n-1}$ yields $q_{n+1} \geq \alpha_n q_n \gg j^{1/4} 2^j$ for sufficiently large j. However, M_j by definition is a subset of $\{1, \ldots, 2^j/j^{1/4+\varepsilon/3}\}$. This shows that q_{n+1} is already too large to be contained in M_j, and consequently M_j consists only of integer multiples of q_n.

(ii) Now we give an upper bound for the largest possible value of $h \geq 1$ such that $h m_j q_n \in M_j$. From (4.26) and the definition of M_j we deduce that

$$h \leq \frac{2^j}{q_n j^{1/4+\varepsilon/3}} \qquad \text{as well as} \qquad h \ll \frac{\alpha_n q_n}{m_j 2^j}.$$

As noted above we have $2^j/j^{1+\varepsilon/3} \ll q_n \ll 2^j/j^{1/4+\varepsilon/3}$. Thus

$$h \ll \max_{\frac{2^j}{j^{1+\varepsilon/3}} \ll x \leq \frac{2^j}{j^{1/4+\varepsilon/3}}} \min\left\{\frac{2^j}{x j^{1/4+\varepsilon/3}}, \frac{\alpha_n x}{m_j 2^j}\right\}$$

where the $x \in \mathbb{R}$ maximising the right hand side, under the given constraints, is determined via

$$\frac{2^j}{x j^{1/4+\varepsilon/3}} = \frac{\alpha_n x}{m_j 2^j} \quad \Leftrightarrow \quad x^2 = \frac{m_j 2^{2j}}{j^{1/4+\varepsilon/3}\alpha_n}.$$

Thus, using $\alpha_n \ll (\log q_n)^{1+\varepsilon/3} \ll j^{1+\varepsilon/3}$, we finally obtain

$$h^2 \ll \frac{\alpha_n}{j^{1/4+\varepsilon/3} m_j} \ll j^{1/2}.$$

Thus $\#M_j \ll j^{1/4}$, which proves the lemma. □

4.4 Analysing the Contribution of the Large Differences

The Fourier series expansion of the indicator functions $I_{s,N}(\alpha)$ is given by

$$I_{s,N}(\alpha) \sim \sum_{n\in\mathbb{Z}} c_n e(n\alpha) \qquad \text{where} \quad c_n := \begin{cases} \sin(2\pi n s/N)/(\pi n) & \text{if } n \neq 0, \\ 2s/N & \text{if } n = 0, \end{cases} \tag{4.27}$$

where we write $e(\alpha)$ for $\exp(2\pi i\alpha)$. The next lemma is of a technical nature, and is used in a decoupling argument for the variance bounds, which are derived in Section 4.4.1.

Lemma 4.4. *Define for integers $u, v > 0$ the quantity*

$$C(u,v) := \sum_{\substack{n_1,n_2\in\mathbb{Z}\setminus\{0\},\\ n_1u=n_2v}} c_{n_1}c_{n_2}. \tag{4.28}$$

Then

$$C(u,v) \ll \frac{\gcd(u,v)}{\max\{u,v\}}. \tag{4.29}$$

Moreover, for $u \neq 0$ we have

$$C(u,u) \ll_s N^{-1}. \tag{4.30}$$

Proof. Note that $n_1u = n_2v$ holds if and only if there is an integer $h \neq 0$ satisfying $n_1 = hu/\gcd(u,v)$ and $n_2 = hv/\gcd(u,v)$. Moreover, we observe that $|c_n| \leq \min\{2s/N, 1/|n|\}$ for $n \neq 0$. Combining these estimates with the Cauchy–Schwarz inequality yields

$$\begin{aligned}
|C(u,v)|^2 &\leq \sum_{h\in\mathbb{Z}\setminus\{0\}} c^2_{h\frac{u}{\gcd(u,v)}} \sum_{h\in\mathbb{Z}\setminus\{0\}} c^2_{h\frac{v}{\gcd(u,v)}} \\
&\leq \sum_{h\in\mathbb{Z}\setminus\{0\}} \frac{(\gcd(u,v))^2}{(uh)^2} \sum_{h\in\mathbb{Z}\setminus\{0\}} \frac{(\gcd(u,v))^2}{(vh)^2},
\end{aligned}$$

which implies (4.29).

Furthermore,

$$C(u,u) \ll \sum_{n\leq\frac{N}{2s}} \frac{4s^2}{N^2} + \sum_{n>\frac{N}{2s}} \frac{1}{n^2},$$

which implies (4.30). □

From orthogonality relations, combined with (4.27), we obtain

$$N^2 \operatorname{Var}\big(R(AG,\cdot,s,N)\big) = \int_0^1 \Bigg(\sum_{d \in \mathcal{D}_N(AG)} \mathrm{r}(d) \sum_{n \in \mathbb{Z}\backslash\{0\}} c_n e(dn\alpha) \Bigg)^2 \mathrm{d}\alpha = \sum_{u,v \in \mathcal{D}_N(AG)} \mathrm{r}(u)\, \mathrm{r}(v)\, C(u,v), \tag{4.31}$$

where $r(\cdot)$ is the representation function which counts representations as Case (AG). A perfect analogue holds when (AG) is replaced by ($\mathrm{AA}_{\mathrm{diff}}$) everywhere in the formula (including in the definition of the representation function r).

The main term on the right hand side, as we shall see, is the sum over the diagonal $(\mathrm{r}(u))^2 C(u,u)$. To prove this, the next lemma shows that the contribution from the off-diagonal terms is small. More precisely, $C(u,v)$ is extremely small for two elements $u \neq v$ of $\mathcal{D}_N(AG)$ or $\mathcal{D}_N(AA_{\mathrm{diff}})$.

Lemma 4.5. *We have*

$$\sum_{\substack{u,v \in \mathcal{D}_N(AG),\\ u \neq v}} \mathrm{r}(u)\, \mathrm{r}(v)\, C(u,v) \ll 1, \tag{4.32}$$

where the representation function r counts representations from Case (AG). The same estimate holds if (AG) is replaced by (AA_{diff}).

Proof. This is not a critical part in the whole argument, and it is sufficient to use very rough estimates. We only give a brief outline of the proof. Let u and v be elements of the difference set $\mathcal{D}_N(AG)$ such that $0 < u < v$. Recall that different building blocks of our sequence are separated by huge constants. For u and v this leaves only two possibilities:

- Either u is of much smaller order than v, say $u \ll v^{1/2}$. By (4.29) we have $C(u,v) \ll \gcd(u,v)/\max\{u,v\}$. Since $\gcd(u,v) \leq u$, we have $C(u,v) \ll u/v \ll v^{-1/2}$.
- The second possibility is that u and v are of very similar size, and that consequently $v-u$ is very small in comparison with v. In this case we may assume for example that $v-u \ll v^{1/2}$. Again using $C(u,v) \ll \gcd(u,v)/\max\{u,v\}$, and now observing that $\gcd(u,v) \leq v-u \ll v^{1/2}$, we obtain $C(u,v) \ll v^{-1/2}$.

So in both cases $C(u,v)$ is small in comparison with v. By construction the difference set $\mathcal{D}_N(AG)$ is an extremely sparse set, due to the very fast growth of our sequence. Thus after summing over u and v we can obtain (4.32). A similar argument works when instead of $\mathcal{D}_N(AG)$ we consider $\mathcal{D}_N(AA_{\mathrm{diff}})$. □

4.4.1 Variance Bounds

Now we have the tools at hand to derive the variance bounds for the auxiliary functions $R(AG,\cdot,s,N)$ and $R(AA_{\mathrm{diff}},\cdot,s,N)$ which were defined in (4.19).

Lemma 4.6. *For every fixed $s > 0$, we have*

$$\operatorname{Var}\big(R\,(AG, \cdot, s, N)\big) \ll_s N^{-1/2}. \tag{4.33}$$

Proof. Again this is not a crucial lemma, and it is sufficient to use very rough estimates. Note that trivially $\#\mathcal{D}_N\,(AG) \leq N^2$. Let $u \in \mathcal{D}_N\,(AG)$. Then, using again the fact that our sequence increases very quickly, we can easily show that the number of Case (AG) representations $r(u)$ which u has as the difference of two elements from A_N is very small. To give a quantitative statement, we could easily show that $r(u) \ll N^{1/4}$, uniformly in u (this is just a very rough estimate). Hence (4.30) implies

$$\sum_{u \in \mathcal{D}_N(AG)} \mathrm{r}\,(u)^2\,|C\,(u,u)| \ll_s (\#\mathcal{D}_N\,(AG))\,N^{1/2}N^{-1} \ll_s N^{3/2}.$$

Together with (4.31) and (4.32) this implies (4.33).

□

The contribution coming from numbers which arise as the difference between two numbers from different arithmetic blocks is a bit more difficult to control. To see this, note that when there are two arithmetic progressions with different step sizes m_{j_1} and m_{j_2}, then there are certain numbers which have many representations as a number from the first arithmetic progression, minus a number from the second arithmetic progression. To control the contribution from such numbers, we will make crucial use of the fact that in our construction the step sizes m_{j_1} and m_{j_2} are prime numbers.

Lemma 4.7. *For every fixed $s > 0$, we have*

$$\operatorname{Var}\big(R\,(AA_{\mathit{diff}}, \cdot, s, N)\big) \ll_s \frac{1}{(\log N)^{1+\varepsilon/2}}. \tag{4.34}$$

Proof. Let N be given. There is some J such that $a_N \in P_A(J) \cup P_G(J)$, and by construction for this value of J we have $J \asymp \log N$. By (4.31) and Lemma 4.5 we have

$$\begin{aligned}
\operatorname{Var}\big(R\,(AA_{\text{diff}}, \cdot, s, N)\big) &\leq \frac{1}{N^2} \sum_{u,v \in \mathcal{D}_N\left(AA_{\text{diff}}\right)} r(u)r(v)|C(u,v)| \\
&\ll \frac{1}{N^2}\left(1 + \sum_{1 \leq j_1 < j_2 \leq J}\ \sum_{u \in P_A(j_2) - P_A(j_1)} \mathrm{r}\,(u)^2\,|C\,(u,u)|\right),
\end{aligned}$$

where $r(u)$ counts the number of representation of u as the difference between an element of $P_A\,(j_2)$ and an element of $P_A\,(j_1)$. Here we used the fact that due to the huge constants which separate different blocks in our construction, for given u there is only one pair (j_1, j_2) such that $u \in P_A\,(j_2) - P_A\,(j_1)$, except maybe for finitely many (small) values of u.

Let $j_1 < j_2$ be fixed. First assume that $j_1 < J - 2\log J$. We note that the cardinality of the set $P_A(j_2) - P_A(j_1)$ is bounded by

$$\begin{aligned}\#\{P_A(j_2) - P_A(j_1)\} &\ll \max\{P_A(j_2) - P_A(j_1)\} - \min\{P_A(j_2) - P_A(j_1)\} \\ &\ll m_{j_2} \frac{N}{(\log N)^{1/4+\varepsilon/3}} \\ &\ll \frac{N}{(\log N)^{\varepsilon/3}}. \end{aligned} \tag{4.35}$$

Then by the trivial estimate $\mathrm{r}(u) \ll \#P_A(j_1) \ll N/(\log N)^{2\log 2 + 1/4}$, and since $2\log 2 + 1/4 > 16/10$, we have

$$\begin{aligned}\frac{1}{N^2} \sum_{\substack{j_1, j_2, \\ j_1 < J - 2\log J}} \sum_{u \in P_A(j_2) - P_A(j_1)} \mathrm{r}(u)^2 |C(u,u)| &\ll_s \frac{J^2}{N^2} \frac{N^3}{(\log N)^{16/5}} \frac{1}{N} \\ &\ll \frac{1}{(\log N)^{6/5}}. \end{aligned} \tag{4.36}$$

It remains to control the contribution from the range $J - 2\log J \le j_1 < j_2 \le J$. Here it plays a crucial role that for j_1, j_2 in this range, by construction there are two *different* primes m_{j_1} and m_{j_2} which form the step sizes of the arithmetic progression $P_A(j_1)$ and $P_A(j_2)$, respectively (cf. Lemma 4.1). Therefore, in such a situation $\mathrm{r}(u)$ is bounded by the number of solutions $(x, y) \in \mathbb{Z}^2$ to the linear Diophantine equation

$$\tilde{u} = m_{j_2} x - m_{j_1} y \qquad \text{where} \quad \tilde{u} := u - \min\{P_A(j_2)\} + \min\{P_A(j_1)\},$$

and (x, y) satisfies the additional restriction that $1 \le x, y \le N/(\log N)^{1/4+\varepsilon/3}$. Since m_{j_1} and m_{j_2} are prime numbers, the set of integer solutions to this equation admits the form

$$(x_0 + h m_{j_1}, y_0 - h m_{j_2}),$$

where $h \in \mathbb{Z}$ and (x_0, y_0) is some solution to the above equation. Moreover, the size of j_1 and j_2, together with (4.6), ensures that $m_{j_1} \asymp m_{j_2} \asymp (\log N)^{1/4}$. Hence,

$$\mathrm{r}(u) \ll \frac{N}{(\log N)^{1/2+\varepsilon/3}}. \tag{4.37}$$

Thus using (4.30), (4.35) and (4.37), and noting that $\log J \ll \log\log N \ll (\log N)^{\varepsilon/2}$, we obtain that

$$\begin{aligned}\frac{1}{N^2} \sum_{\substack{j_1, j_2, \\ J - 2\log J \le j_1 < j_2 \le J}} \sum_{u \in P_A(j_2) - P_A(j_1)} \mathrm{r}(u)^2 |C(u,u)| &\ll_s \frac{1}{N^2} \sum_{\substack{j_1, j_2, \\ J - 2\log J \le j_1 < j_2 \le J}} \frac{N^3}{(\log N)^{1+\varepsilon}} \frac{1}{N} \\ &\ll \frac{1}{(\log N)^{1+\varepsilon/2}}. \end{aligned}$$

Combining this with (4.36) yields (4.34). □

4.5 Proof of Theorem 4.1

Let N be given. There is a number J such that $a_N \in P_A(J) \cup P_G(J)$, and for this value of J we have $J \asymp \log N$ and $2^J \asymp N$. Then

$$E\big(A_N\big) \geq E\big(P_A(J-1)\big) \gg \frac{N^3}{(\log N)^{3/4+\varepsilon}},$$

where we used that the additive energy of an arithmetic progression is proportional to the third power of its cardinality, and that by construction $\# P_A(J-1) \gg 2^J / J^{1/4+\varepsilon/3} \gg N/(\log N)^{1/4+\varepsilon/3}$. Thus the additive energy of the sequence constructed in our example is indeed as large as claimed in the statement of the theorem.

It remains to show that $(a_n)_n$ has the metric pair correlation property. Recall that for almost all α the contribution coming from the geometric blocks gives the desired convergence $R(GG, \alpha, s, N) \to 2s$, for every fixed $s > 0$, cf. (4.20). It is a standard procedure to use the variance estimates and the results from the previous section to conclude that the contribution of the parts $R(AG)$ and $R(AA_{\text{diff}})$ tends to zero in the limit; thus we will only give a brief outline. Fix a rational $s > 0$. Define the sequence

$$N_m = \big\lfloor \exp\big(m^{\frac{1}{1+\varepsilon/2}}\big)\big\rfloor$$

and note that $N_{m+1}/N_m \to 1$. If N is such that $N_m \leq N < N_{m+1}$, then

$$NR\big(AG, \alpha, s, N\big) \leq N_{m+1} R\big(AG, \alpha, N_{m+1}/N_m s, N_{m+1}\big).$$

Denote by $E_{AG,s}(N_m)$ the "exceptional" set

$$\Big\{\alpha \in [0,1] : \big|R\big(AG, \alpha, N_m/N_{m+1}s, N\big) - \mu_{AG,s}(N_m)\big| \geq 1/\log\log N_m\Big\}$$

where $\mu_{AG,s}(N_m)$ is the expected value of $R\big(AG, \alpha, N_m/N_{m+1}s, N\big)$. Now, observe that $\mu_{AG,s}(N_m) \to 0$ as $m \to \infty$, since the indices of those elements of $(a_n)_n$ which come from an arithmetic block are contained in a set of zero density within the total index set. Combining Chebyshev's inequality with the variance estimates from Lemma 4.6 and Lemma 4.7, and applying the Borel–Cantelli lemma, we obtain

$$R\big(AG, \alpha, s, N\big) \underset{N \to \infty}{\longrightarrow} 0, \tag{4.38}$$

for all rational s and for Lebesgue almost all $\alpha \in [0,1]$. Exactly the same argument works if (AG) is replaced by (AA_{diff}).

Finally we have to show that $R(AA_{\text{same}}) \to 0$ for almost all α. Let $s > 0$ be fixed, and assume that s is rational. By the Borel–Bernstein theorem, almost no $\alpha \in [0,1]$ has infinitely many $d \geq N/(\log N)^{3/2}$ such that $I_{s,N}(d\alpha) = 1$. Hence it is sufficient to estimate the contribution of those differences d which are contained in $(P_A(j) - P_A(j))^+$ for a value of j which is close to J. More precisely, we can restrict j to the range $J - 2\log J \le j \le J$. By Lemma 4.3, for almost all $\alpha \in [0,1]$ we have

$$\begin{aligned} R\big(AA_{\text{same}}, \alpha, s, N\big) &\ll \frac{1}{N} \sum_{J-2\log J \le j \le J} \frac{2^j}{j^{1/4+\varepsilon/3}} \cdot \# \left\{ d \in (P_A(j) - P_A(j))^+ : \ \|d\alpha\| \le \frac{s}{N} \right\} \\ &\ll_s \frac{1}{N} \sum_{J-2\log J \le j \le J} \frac{2^j}{j^{\varepsilon/3}} \\ &\ll (\log N)^{-\varepsilon/6}, \end{aligned} \tag{4.39}$$

where we estimated $\log J \ll \log\log N \ll (\log N)^{\varepsilon/6}$.

Thus we have $R(GG) \to 2s$, and $R(AG) \to 0$, $R(AA_{\text{diff}}) \to 0$, $R(AA_{\text{same}}) \to 0$, for all rational $s > 0$, for almost all α. However, if this convergence holds for all rational $s > 0$ and almost all α, then by monotonicity it must also hold for all real $s > 0$ and almost all α. In view of the decomposition (4.18) this concludes the proof of the theorem.

Chapter 5

The Duffin–Schaeffer Conjecture with Extra Divergence

5.1 Introduction and Statement of Results

Let $\psi : \mathbb{N} \to \mathbb{R}$ be a non-negative function. For every non-negative integer n define a set $\mathcal{E}_n \subset \mathbb{R}/\mathbb{Z}$ by

$$\mathcal{E}_n := \bigcup_{\substack{1 \le a \le n, \\ (a,n)=1}} \left(\frac{a - \psi(n)}{n}, \frac{a + \psi(n)}{n} \right). \tag{5.1}$$

The Lebesgue measure of $\mathcal{E}_n$ is at most $2\psi(n)\varphi(n)/n$, where φ denotes Euler's totient function. Thus, writing $W(\psi)$ for the set of those $x \in [0, 1]$ which are contained in infinitely many sets $\mathcal{E}_n$, it follows directly from the first Borel–Cantelli lemma that $\lambda(W(\psi)) = 0$ whenever

$$\sum_{n=1}^{\infty} \frac{\psi(n)\varphi(n)}{n} < \infty. \tag{5.2}$$

Here λ denotes the Lebesgue measure. The corresponding divergence statement, which asserts that $\lambda(W(\psi)) = 1$ whenever the series in (5.2) is divergent, is known as the Duffin–Schaeffer conjecture [16] and is one of the most important open problems in metric number theory. It remains unsolved since 1941.

The Duffin–Schaeffer conjecture is known to be true under some additional arithmetic conditions or regularity conditions on the function ψ. See for example [22, 44]. In [23] Haynes, Pollington and Velani initiated a program to establish the Duffin–Schaeffer condition without assuming any regularity or number-theoretic properties of ψ, but instead assuming a slightly stronger divergence condition. In [23] they proved that there is a constant c such that $\lambda(W(\psi)) = 1$, provided that

$$\sum_{n=1}^{\infty} \frac{\psi(n)\varphi(n)}{n \, e^{\left(\frac{c \log n}{\log \log n}\right)}} = \infty$$

(throughout this chapter we will understand $\log x$ as $\max(1, \log x)$, so that all appearing logarithms are positive and well-defined). In [6] Beresnevich, Harman, Haynes and Velani used a beautiful averaging argument, which is also at the core of the argument in the present chapter, to show that it is sufficient to assume

$$\sum_{n=1}^{\infty} \frac{\psi(n)\varphi(n)}{n(\log n)^{\varepsilon \log\log\log n}} = \infty$$

for some $\varepsilon > 0$. In the present chapter we prove that the extra divergence factor can be reduced to $(\log n)^{\varepsilon}$ for a fixed $\varepsilon > 0$. In particular this solves Problem 2 posed in [23], where it was asked whether the extra divergence factor $\log n$ is sufficient.

Theorem 5.1. *Let $\psi\colon \mathbb{N} \to \mathbb{R}$ non-negative function. Then for almost all α there are infinitely many coprime $p, q \in \mathbb{N}$ such that*

$$\left|\alpha - \frac{p}{q}\right| < \frac{\psi(q)}{q}$$

if there is some $\varepsilon > 0$ such that

$$\sum_{n\geq 1} \psi(n) \frac{\varphi(n)}{n(\log n)^{\varepsilon}} \tag{5.3}$$

diverges.

We note that by the mass transference principle of Beresnevich and Velani [7] it is possible to deduce Hausdorff measure statements from results for Lebesgue measure, in the context of the Duffin–Schaeffer conjecture. Roughly speaking, the quantitative "extra divergence" result in Theorem 5.1 translates into a corresponding condition on the dimension function of a Hausdorff measure for the set where the Duffin–Schaeffer conjecture is true. For details we refer the reader to Section 4 of [23], where this connection is explained in detail.

5.2 Proof of Theorem 5.1

Throughout the proof, we assume that $\varepsilon > 0$ is fixed. We use Vinogradov notation "$\ll$", where the implied constant may depend on ε, but not on m, n, h or anything else.

As noted in [6], we may assume without loss of generality that for all n either $1/n \leq \psi(n) \leq 1/2$ or $\psi(n) = 0$. Furthermore, by Gallagher's zero-one law [17] the measure of $W(\psi)$ can only be either 0 or 1. Thus $\lambda(W(\psi)) > 0$ implies $\lambda(W(\psi)) = 1$.

We will use the following version of the second Borel–Cantelli lemma (see for example [21, Lemma 2.3]).

Lemma 5.1. *Let $(A_n)_n$ be a sequence of Lebesgue measurable sets in $[0,1]$ satisfying*

$$\sum_{n\geq 1} \lambda(A_n) = \infty.$$

Then,

$$\lambda\left(\limsup_{n\to\infty} A_n\right) \geq \limsup_{N\to\infty} \frac{\left(\sum_{n\leq N}\lambda\left(A_n\right)\right)^2}{\sum_{m,n\leq N}\lambda\left(A_n\cap A_m\right)}.$$

The following lemma of Pollington and Vaughan [36] allows to estimate the ratio between the measure of the overlap $\mathcal{E}_m \cap \mathcal{E}_n$ and the product of the measures of $\mathcal{E}_m$ and $\mathcal{E}_n$, and is a key ingredient in [6]. In the statement of the lemma and in the sequel, we write (m,n) for the greatest common divisor of two positive integers m, n.

Lemma 5.2. *For $m \neq n$, assume that $\lambda(\mathcal{E}_m)\lambda(\mathcal{E}_n) \neq 0$. Define*

$$P(m,n) = \frac{\lambda(\mathcal{E}_m \cap \mathcal{E}_n)}{\lambda(\mathcal{E}_m)\lambda(\mathcal{E}_n)}. \tag{5.4}$$

Then

$$P(m,n) \ll \prod_{\substack{p \mid \frac{mn}{(m,n)^2},\\ p>D(m,n)}} \left(1-\frac{1}{p}\right)^{-1}, \tag{5.5}$$

where the product is taken over all primes p in the specified range, and where

$$D(m,n) = \frac{\max(n\psi(m), m\psi(n))}{(m,n)}. \tag{5.6}$$

In view of Lemma 5.1 it is clear that controlling $P(m,n)$ is the key to proving $\lambda(W(\psi)) > 0$. Following [6], we divide the set of positive integers into blocks

$$2^{4^h} \leq n < 2^{4^{h+1}}, \qquad h \geq 1, \tag{5.7}$$

and we may assume without loss of generality that the divergence condition (5.3) still holds when the summation is restricted to those n which are contained in a block with h being even. As noted in [6], when m and n are contained in different blocks, then automatically $P(m,n) \ll 1$. Thus the real problem is that of controlling $P(m,n)$ when m and n are contained in the same block (5.7) for some h.

In the sequel, let m,n be fixed, and assume that

$$2^{4^h} \leq m < n < 2^{4^{h+1}}$$

for some h. As in [6], we will average the factors $P(m,n)$ over a range of downscaled versions of the sets $\mathcal{E}_m$ and $\mathcal{E}_n$. More precisely, for $k = 1, 2, \ldots$, let $\mathcal{E}_n^{(k)}$ be defined as $\mathcal{E}_n$, but with $\psi(n)/e^k$ in place of $\psi(n)$. Correspondingly, we define

$$P_k(m,n) = \frac{\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big)}{\lambda\big(\mathcal{E}_m^{(k)}\big)\lambda\big(\mathcal{E}_n^{(k)}\big)}$$

and

$$D_k(m,n) = \frac{\max(n\psi(m), m\psi(n))}{e^k (m,n)},$$

and note that for P_k we have the same estimate as in (5.5), only with D replaced by D_k. At the core of the argument in [6] is the observation that

$$\begin{aligned}\sum_{k=1}^{K} P_k(m,n) &\ll \sum_{k=1}^{K} \prod_{\substack{p \mid \frac{mn}{(m,n)^2},\\ p>e^k}} \left(1-\frac{1}{p}\right)^{-1} \\ &\ll \sum_{k=1}^{K} \frac{\log\log n}{k} \\ &\ll (\log K)(\log\log n), \end{aligned} \tag{5.8}$$

where the product in the first line is estimated using Mertens' second theorem. Thus when $K \gg (\log\log n)(\log\log\log n)$ we have $\sum_{k=1}^{K} P_k(m,n) \ll K$, and accordingly there is at least one value of k in this range for which $P_k(m,n) \ll 1$. This argument can be extended over a range of pairs (m,n) instead of assuming that m,n are fixed. Together with Lemma 5.1 and Gallagher's zero-one law this allows to deduce the desired result, provided that we are allowed to divide $\psi(n)$ by $e^K \leq e^{\varepsilon(\log\log n)(\log\log\log n)}$ for all n and still keep the divergence of the sum of measures.

In our proof we will roughly follow the same plan. However, instead of taking (5.5) for granted and then averaging over different reduction factors e^k, we will take the averaging procedure into the proof of the overlap estimate which leads to Lemma 5.2. To see where a possible improvement could come from, we note that to obtain the estimate in Lemma 5.2 it is necessary to give upper bounds for sums

$$\sum_{\substack{1\leq b\leq\theta,\\ (b,t)=1}} 1,$$

where we can think of $\theta \ll \log t$ as being the number D from (5.6), and of t as being the number $\frac{mn}{(m,n)^2}$ which appears in (5.5). It is necessary to relate this sum to $\theta\varphi(t)/t$. To obtain Lemma 5.2 one applies the classical sieve bound

$$\sum_{\substack{1\leq b\leq\theta,\\ (b,t)=1}} 1 \ll \theta \prod_{\substack{p\mid t,\\ p\leq\theta}} \left(1-\frac{1}{p}\right) = \theta\frac{\varphi(t)}{t} \prod_{\substack{p\mid t,\\ p>\theta}} \left(1-\frac{1}{p}\right)^{-1}, \tag{5.9}$$

and the product on the very right is the one which also appears in (5.5). This sieve bound gives optimal results for some constellations of parameters, but we can use the fact that we are averaging over different values of k (which determine θ) to save some factors. We exhibit two extremal cases showing this phenomenon. The factor $P(m,n)$ can only be large when the product on the right of (5.9) is large. However, this product can only be large if a very large proportion of small primes divides t. Assume on the contrary that *no* small prime divides t. Then the sieve inequality in (5.9) is actually an equality, since on both sides we have exactly θ, but the product on the very right is extremely small and cannot cause problems. As a second extremal case, assume that *all* small primes divide t. Then the product on the very right is very large, but the sieve bound is not sharp, since in the sum on the left the only number we count is the number 1 (no other small number is coprime to t). So there is a trade-off between the way how a large proportion of primes dividing t is able to increase the value of the product on the right of (5.9), but at the same time reduces the quality of the sieve bound. It seems that this should be a very subtle relationship, and in general this is indeed the case (cf. [18, Proposition 2.6], where this phenomenon is addressed). However, quite surprisingly, it turns out that in our particular situation it is possible to exploit this phenomenon using only some simple calculations.

Following [36, Paragraph 3], we write m and n in their prime factorization

$$m = \prod_p p^{u_p}, \qquad n = \prod_p p^{v_p},$$

and define

$$r = \prod_{\substack{p,\\ u_p = v_p}} p^{u_p}, \qquad s = \prod_{\substack{p,\\ u_p \neq v_p}} p^{\min(u_p, v_p)}, \qquad t = \prod_{\substack{p,\\ u_p \neq v_p}} p^{\max(u_p, v_p)}.$$

Furthermore, we set

$$\delta = \min\left(\frac{\psi(m)}{m}, \frac{\psi(n)}{n}\right), \qquad \Delta = \max\left(\frac{\psi(m)}{m}, \frac{\psi(n)}{n}\right).$$

Then for every k from the first displayed formula on page 196 of [36] we have the estimate

$$\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big) \ll \frac{\delta}{e^k}\varphi(s)\frac{\varphi(r)^2}{r}\int_1^{4\Delta r t e^{-k}} S_t(\theta)\ d\theta,$$

where we write

$$S_t(\theta) = \sum_{\substack{1 \leq b \leq \theta,\\ (b,t)=1}} \frac{1}{\theta}$$

and where we used that changing $\psi(m) \mapsto \psi(m)/e^k$ and $\psi(n) \mapsto \psi(n)/e^k$ also changes $\delta \mapsto \delta/e^k$ and $\Delta \mapsto \Delta/e^k$. Since

$$\lambda\big(\mathcal{E}_m^{(k)}\big)\lambda\big(\mathcal{E}_n^{(k)}\big) = \frac{\varphi(m)\varphi(n)\delta\Delta}{e^{2k}}$$

this implies

$$\begin{aligned} P_k(m,n) &\ll \frac{e^k \varphi(s)\varphi(r)^2 \int_1^{4\Delta r t e^{-k}} S_t(\theta)\ d\theta}{\Delta r \varphi(m)\varphi(n)} \\ &= \frac{\varphi(t)t}{\varphi(t)t}\ \frac{\varphi(s)\varphi(r)^2}{\varphi(m)\varphi(n)}\ \frac{\int_1^{4\Delta r t e^{-k}} S_t(\theta)\ d\theta}{\Delta r e^{-k}} \\ &= \frac{t}{\varphi(t)} \frac{\int_1^{4\Delta r t e^{-k}} S_t(\theta)\ d\theta}{\Delta r t e^{-k}}, \end{aligned}$$

where the last line follows from $\varphi(s)\varphi(r)^2\varphi(t) = \varphi(m)\varphi(n)$. We set $K = K(h) = \lfloor \varepsilon h \log 4 \rfloor$. Note that with this choice of K we have

$$e^K \ll (\log m)^\varepsilon, (\log n)^\varepsilon \ll e^K. \tag{5.10}$$

Summing over k, we deduce that

$$\sum_{k=1}^{K} P_k(m,n) \ll \sum_{k=1}^{K} \frac{t}{\varphi(t)} \frac{\int_1^{4\Delta rte^{-k}} S_t(\theta)\ d\theta}{\Delta rte^{-k}}. \tag{5.11}$$

As noted in [36] and [6], if $2\Delta rte^{-k} \le 1$ then $P_k(m,n) = 0$, since in this case $\mathcal{E}_m^{(k)}$ and $\mathcal{E}_n^{(k)}$ are disjoint (see the fourth displayed formula from below on p. 195 of [36]). Furthermore, again as noted in [36] and [6], if $4\Delta rte^{-k} \ge e^K \gg (\log n)^\varepsilon$ then $P_k(m,n) \ll 1$, which follows from Lemma 5.2 and Mertens' second theorem. Accordingly, for the contribution to (5.11) of those k for which $4\Delta rte^{-k} \notin [1, e^K)$ we have

$$\sum_{\substack{1\le k\le K,\\ 4\Delta rte^{-k} \notin [1,e^K)}} P_k(m,n) \ll K. \tag{5.12}$$

To estimate the contribution of the other values of k, we note that there exists a number $c \in [1, e)$ such that

$$\Big(\big\{4\Delta rte^{-k},\ k = 1,\dots,K\big\} \cap [1, e^K)\Big) \subset \{ce^j, j = 0,\dots,K-1\}.$$

Thus for the contribution of these k to (5.11) we have

$$\sum_{\substack{1\le k\le K,\\ 4\Delta rte^{-k} \in [1,e^K)}} P_k(m,n) \ \ll\ \frac{t}{\varphi(t)} \sum_{j=0}^{K-1} \frac{1}{e^j} \int_1^{ce^j} S_t(\theta)\ d\theta. \tag{5.13}$$

For the term on the right-hand side of (5.13) we have

$$\begin{aligned}
\sum_{j=0}^{K-1} \frac{1}{e^j} \int_1^{ce^j} S_t(\theta)\, d\theta &\ll \sum_{j=1}^{K} \frac{1}{e^j} \int_1^{e^j} S_t(\theta)\, d\theta \\
&= \sum_{j=1}^{K} \frac{1}{e^j} \sum_{\substack{1\le b\le e^j,\\ (b,t)=1}} \int_b^{e^j} \frac{d\theta}{\theta} \\
&= \sum_{j=1}^{K} \sum_{\substack{1\le b\le e^j,\\ (b,t)=1}} \frac{j-\log b}{e^j} \\
&= \sum_{\substack{1\le b\le e^K,\\ (b,t)=1}} \sum_{j=\lceil \log b\rceil}^{K} \frac{j-\log b}{e^j} \\
&\ll \sum_{\substack{1\le b\le e^K,\\ (b,t)=1}} \frac{1}{b} \underbrace{\sum_{i=1}^{\infty} \frac{i}{e^i}}_{\ll 1} \\
&\ll \sum_{\substack{1\le b\le e^K,\\ (b,t)=1}} \frac{1}{b}.
\end{aligned} \tag{5.14}$$

The sum in (5.14) can be estimated using a sieve with logarithmic weights. Following the lines of [18, Lemma 2.1], we have

$$\begin{aligned}
\sum_{\substack{1\le b\le e^K,\\ (b,t)=1}} \frac{1}{b} &= \sum_{\substack{1\le b\le e^K,\\ p\mid b \implies p\nmid t}} \frac{1}{b} \\
&\le \prod_{\substack{p\le e^K,\\ p\nmid t}} \left(1-\frac{1}{p}\right)^{-1} \\
&= \left(\prod_{p\le e^K} \left(1-\frac{1}{p}\right)^{-1}\right) \left(\prod_{\substack{p\le e^K,\\ p\mid t}} \left(1-\frac{1}{p}\right)\right).
\end{aligned} \tag{5.15}$$

For the first product in (5.15) by Mertens' theorem we have

$$\prod_{p\le e^K} \left(1-\frac{1}{p}\right)^{-1} \ll K.$$

For the second product we have

$$\prod_{\substack{p\le e^K,\\ p|t}} \left(1-\frac{1}{p}\right) = \frac{\varphi(t)}{t} \underbrace{\prod_{\substack{p>e^K,\\ p|t}} \left(1-\frac{1}{p}\right)^{-1}}_{\ll 1},$$

where Mertens' theorem and (5.10) were used to estimate the last product. Inserting these bounds into (5.14), and combining this with (5.12) and (5.13) we finally obtain

$$\sum_{k=1}^{K} P_k(m,n) \ll K. \tag{5.16}$$

By the definition of $P_k(m,n)$ we have

$$\begin{aligned}\sum_{k=1}^{K} P_k(m,n) &= \sum_{k=1}^{K} \frac{\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big)}{\lambda\big(\mathcal{E}_m^{(k)}\big)\lambda\big(\mathcal{E}_n^{(k)}\big)} \\ &= \sum_{k=1}^{K} \frac{e^{2k}\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big)}{\lambda\big(\mathcal{E}_m\big)\lambda\big(\mathcal{E}_n\big)},\end{aligned}$$

and consequently (5.16) implies that

$$\sum_{k=1}^{K} e^{2k}\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big) \ll K\lambda\big(\mathcal{E}_m\big)\lambda\big(\mathcal{E}_n\big).$$

Note that the implied constant is independent of m and n. Thus, summing over m and n yields

$$\sum_{k=1}^{K} \sum_{2^{4^h} \le m<n<2^{4^{h+1}}} e^{2k}\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big) \ll K \sum_{2^{4^h} \le m<n<2^{4^{h+1}}} \lambda\big(\mathcal{E}_m\big)\lambda\big(\mathcal{E}_n\big).$$

Accordingly, there is at least one choice of $k = k(h)$ in the range $\{1,\dots,K\}$ such that

$$\sum_{2^{4^h} \le m<n<2^{4^{h+1}}} e^{2k}\lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big) \ll \sum_{2^{4^h} \le m<n<2^{4^{h+1}}} \lambda\big(\mathcal{E}_m\big)\lambda\big(\mathcal{E}_n\big),$$

or, equivalently, such that

$$\sum_{2^{4^h} \le m<n<2^{4^{h+1}}} \lambda\big(\mathcal{E}_m^{(k)} \cap \mathcal{E}_n^{(k)}\big) \ll \sum_{2^{4^h} \le m<n<2^{4^{h+1}}} \lambda\big(\mathcal{E}_m^{(k)}\big)\lambda\big(\mathcal{E}_n^{(k)}\big), \tag{5.17}$$

where the implied constant does not depend on h. We replace the original function $\psi(n)$ by a function $\psi^*(n)$, where

$$\psi^*(n) = \begin{cases} 0 & \text{when } n \text{ is not in } \big[2^{4^h}, 2^{4^{h+1}}\big) \text{ for some even } h, \\ \psi(n)e^{-k(h)} & \text{when } n \text{ is in } \big[2^{4^h}, 2^{4^{h+1}}\big) \text{ for some even } h, \end{cases}$$

and write $\mathcal{E}_n^*,\ n \geq 1$, for the corresponding sets, which are defined like (5.1) but with ψ^* in place of ψ. By (5.10) we have

$$\psi^*(n) \gg \frac{\psi(n)}{(\log n)^\varepsilon}.$$

Thus the extra divergence condition in the assumptions of Theorem 5.1 guarantees that

$$\sum_{n=1}^{\infty} \lambda(\mathcal{E}_n^*) = \infty,$$

while (5.17) guarantees that

$$\sum_{1 \leq m,n \leq N} \lambda\big(\mathcal{E}_m^* \cap \mathcal{E}_n^*\big) \ll \sum_{1 \leq m,n \leq N} \lambda(\mathcal{E}_m^*)\lambda(\mathcal{E}_n^*)$$

(recall that $\lambda(\mathcal{E}_m^* \cap \mathcal{E}_n^*) \ll \lambda(\mathcal{E}_m^*)\lambda(\mathcal{E}_n^*)$ holds automatically when m and n are not contained in the same block for some h). Thus by Lemma 5.1 we have $\lambda(W(\psi^*)) > 0$, and since $\mathcal{E}_n^* \subset \mathcal{E}_n$ we also have $\lambda(W(\psi)) > 0$. By Gallagher's zero-one law, positive measure of $W(\psi)$ implies full measure. Thus $\lambda(W(\psi)) = 1$, which proves the theorem.

Bibliography

[1] I. Aichinger, C. Aistleitner and G. Larcher. On Quasi-Energy-Spectra, Pair Correlations of Sequences and Additive Combinatorics. In *Contemporary Computational Mathematics - A Celebration of the 80th Birthday of Ian Sloan*, pages 1–16. Springer, Cham, 2018.

[2] C. Aistleitner. Metric number theory, lacunary series and systems of dilated functions. In *Uniform distribution and quasi-Monte Carlo methods*, volume 15 of *Radon Ser. Comput. Appl. Math.*, pages 1–16. De Gruyter, Berlin, 2014.

[3] C. Aistleitner, T. Lachmann and F. Pausinger. Pair correlations and equidistribution. *J. Number Theory*, 182:206-220, 2018.

[4] 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 Jean Bourgain. *Israel J. Math.*, 222(1):463–485, 2017.

[5] J. Arias de Reyna. On the distribution (mod 1) of the normalized zeros of the Riemann zeta-function. *J. Number Theory*, 153:37–53, 2015.

[6] V. Beresnevich, G. Harman, A. Haynes, and S. Velani. The Duffin–Schaeffer conjecture with extra divergence II. *Math. Z.*, 275(1-2):127–133, 2013.

[7] V. Beresnevich and S. Velani. A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures. *Ann. of Math. (2)*, 164(3):971–992, 2006.

[8] M. Berry and M. Tabor. Level clustering in the regular spectrum. *Proc. Royal Soc. London A: Math., Phys. and Engin. Sci.*, 356(1686):375–394, 1977.

[9] T. F. Bloom, S. Chow, A. Gafni, and A. Walker. Additive energy and the metric Poissonian property. *Mathematika*, 64(3):679–700, 2018.

[10] T. F. Bloom and A. Walker. GCD sums and sum-product estimates. preprint available at *arXiv:1806.07849.*

[11] F. P. Boca and A. Zaharescu. Pair correlation of values of rational functions (mod p). *Duke Math. J.*, 105(2):267–307, 2000.

[12] A. Bondarenko and K. Seip. GCD sums and complete sets of square-free numbers. *Bull. Lond. Math. Soc.*, 47(1):29–41, 2015.

[13] Y. Bugeaud. *Approximation by algebraic numbers*, volume 160. Cambridge University Press, 2004.

[14] S. Chow. Bohr sets and multiplicative Diophantine approximation. *Duke Math. J.*, 167(9):1623–1642, 2018.

[15] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*, volume 1651 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1997.

[16] R. J. Duffin and A. C. Schaeffer. Khintchine's problem in metric Diophantine approximation. *Duke Math. J.*, 8:243–255, 1941.

[17] P. Gallagher. Approximation by reduced fractions. *J. Math. Soc. Japan*, 13:342–345, 1961.

[18] A. Granville, D. Koukoulopoulos, and K. Matomäki. When the sieve works. *Duke Math. J.*, 164(10):1935–1969, 2015.

[19] R. M. Gray. Toeplitz and circulant matrices: a review. *Found. Trends Commun. Inf. Theory*, 2(3):93, 2005.

[20] S. Grepstad and G. Larcher. On pair correlation and discrepancy. *Arch. Math.*, 109(2):143–149, 2017.

[21] G. Harman. *Metric number theory*, volume 18 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1998.

[22] G. Harman. Some cases of the Duffin and Schaeffer conjecture. *Quart. J. Math. Oxford Ser. (2)*, 41(164):395–404, 1990.

[23] A. K. Haynes, A. D. Pollington, and S. L. Velani. The Duffin-Schaeffer conjecture with extra divergence. *Math. Ann.*, 353(2):259–273, 2012.

[24] D. R. Heath-Brown. Pair correlation for fractional parts of αn^2. *Math. Proc. Cambridge Philos. Soc.*, 148(3):385–407, 2010.

[25] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974.

[26] T. Kuna, J. L. Lebowitz, and E. R. Speer. Realizability of point processes. *J. Stat. Phys.*, 129(3):417–439, 2007.

[27] T. Kuna, J. L. Lebowitz, and E. R. Speer. Necessary and sufficient conditions for realizability of point processes. *Ann. Appl. Probab.*, 21(4):1253–1281, 2011.

[28] T. Lachmann and N. Technau. On exceptional sets in the metric Poissonian pair correlations problem. *Monatsh. Math.*, 2018. *doi:10.1007/s00605-018-1199-2.*

[29] G. Larcher. Remark on a result of Bourgain on Poissonian pair correlation. preprint available at *arXiv:1711.08663.*

[30] G. Larcher and W. Stockinger. Pair correlation of sequences $(\{a_n\alpha\})_{n\in\mathbb{N}}$ with maximal additive energy. *Mathematical Proceedings of the Cambridge Philosophical Society*, 2018. *doi:10.1017/S030500411800066X*.

[31] G. Larcher and W. Stockinger. Some negative results related to Poissonian pair correlation problems. preprint available at *arXiv:1803.05236*.

[32] J. Marklof. The Berry-Tabor conjecture. In *European Congress of Mathematics, Vol. II (Barcelona, 2000)*, volume 202 of *Progr. Math.*, pages 421–427. Birkhäuser, Basel, 2001.

[33] J. Marklof and A. Strömbergsson. Equidistribution of Kronecker sequences along closed horocycles. *Geom. Funct. Anal.*, 13(6):1239–1280, 2003.

[34] H. L. Montgomery. The pair correlation of zeros of the zeta function. In *Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972)*, pages 181–193. Amer. Math. Soc., Providence, R.I., 1973.

[35] Í. Piršić and W. Stockinger. The Champernowne constant is not Poissonian. *Funct. Approx. Comment. Math.*, 2018. *doi:10.7169/facm/1749*.

[36] A. D. Pollington and R. C. Vaughan. The k-dimensional Duffin and Schaeffer conjecture. *Mathematika*, 37(2):190–200, 1990.

[37] Z. Rudnick and P. Sarnak. The pair correlation function of fractional parts of polynomials. *Comm. Math. Phys.*, 194(1):61–70, 1998.

[38] Z. Rudnick, P. Sarnak, and A. Zaharescu. The distribution of spacings between the fractional parts of $n^2\alpha$. *Invent. Math.*, 145(1):37–57, 2001.

[39] Z. Rudnick and A. Zaharescu. A metric result on the pair correlation of fractional parts of sequences. *Acta Arith.*, 89(3):283–293, 1999.

[40] Z. Rudnick and A. Zaharescu. The distribution of spacings between fractional parts of lacunary sequences. *Forum Math.*, 14(5):691–712, 2002.

[41] S. Steinerberger. Localized quantitative criteria for equidistribution. *Acta. Arith.*, 180(2):183–199, 2017.

[42] T. Tao and V. Vu. *Additive combinatorics*. Vol. 105, Cambridge University Press, 2006.

[43] J. L. Truelsen. Divisor problems and the pair correlation for the fractional parts of $n^2\alpha$. *Int. Math. Res. Not. IMRN*, (16):3144–3183, 2010.

[44] J. D. Vaaler. On the metric theory of Diophantine approximation. *Pacific J. Math.*, 76(2):527–539, 1978.

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

[46] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.*, 77(3):313–352, 1916.

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date

Signature