

# Long-Range Correlations of Sequences Modulo 1

CHRISTOPHER LUTSKO\*

July 21, 2020

## Abstract

In this paper we consider the fractional parts of a general sequence, for example the sequence  $\alpha\sqrt{n}$  or  $\alpha n^2$ . We give a general method, which allows one to show that long-range correlations (correlations where the support of the test function grows as we consider more points) are Poissonian. We show that these statements about convergence can be reduced to bounds on associated Weyl sums. In particular we apply this methodology to the aforementioned examples. In so doing, we recover a recent result of Technau-Walker (2020) for the triple correlation of  $\alpha n^2$  and generalize the result to higher moments. For both of the aforementioned sequences this is one of the only results which indicates the pseudo-random nature of the higher level ( $m \geq 3$ ) correlations.

MSC2020: 11K06; 11K60; 11L07; 37A44; 37A44

KEY WORDS AND PHRASES: Local Statistics; Sequences Modulo 1; Exponential Sums.

## 1 Introduction

Let  $\{a(n)\}_{n \in \mathbb{N}}$  be a sequence in  $\mathbb{R}$ . For a long time, mathematicians have studied the distribution of the fractional parts of such sequences. That is, the sequence  $\{x_n\}_{n \in \mathbb{N}}$  given by

$$x_n := a(n) \pmod{1}.$$

In general, it is fairly well-understood which sequences are uniformly distributed on the interval (see [KN74]). However, uniform distribution is a relatively crude measure of pseudo-randomness. Recently, there has been a lot of interest in the fine-scale local statistics of such sequences. Namely, a central question is to understand when the distribution of gaps between neighboring points in these sequences converges to the exponential function, as they do for uniformly distributed random variables on the interval.

This problem is fundamental from a mathematical point of view in understanding the random nature of deterministic sequences. Furthermore, it also has implications in physics. Since such sequences can be used to understand the energy levels of quantum systems, understanding their distribution has numerous implications. For example, the sequence  $a(n) = \alpha n^2$  for  $\alpha \neq 0$  corresponds to the eigenvalues of a boxed harmonic oscillator. In general, the well-known Berry-Tabor [BT77] states that for generic surfaces, if the Hamiltonian dynamics on the surface are integrable, then the gap distribution for the spectrum has a well-defined limit, given by the exponential distribution. There exist counter-examples to this statement which can be explained away (thus the word generic), however in general this relationship has been confirmed by experiment. But very little can be rigorously proved. For more information see the review [Mar00].

---

\*Rutgers University  
Email: chris@lutsko.com

## 1.1 Long-Range Correlations

In general the gap distribution is very hard to work with directly. This owes to the fact that neighboring points are not necessarily consecutive points in the sequence. To get around this problem, we instead consider the  $m$ -level correlations: Given a sequence  $\{x_n\}_{n \in \mathbb{N}} \subset [0, 1)$ , and a vector  $\mathbf{j} \in \mathbb{Z}^m$ , let  $\Delta(\mathbf{j}) \in \mathbb{R}^{m-1}$  denote the difference vector

$$\Delta(\mathbf{j}) := (x_{j_1} - x_{j_2}, x_{j_2} - x_{j_3}, \dots, x_{j_{m-1}} - x_{j_m}).$$

Then for  $f \in C_c^\infty(\mathbb{R}^{m-1})$  we define the  $m$ -level correlation to be

$$R^{(m)}(N, f) := \frac{1}{N} \sum_{\mathbf{j} \in \{1, \dots, N\}^m}^* f(N(\Delta(\mathbf{j}))), \quad (1.1)$$

where the notation  $\sum^*$  indicates that all entries of the vector  $\mathbf{j}$  are distinct. We say that the  $m$ -level correlation is *Poissonian* if

$$R^{(m)}(N, f) \rightarrow \int_{\mathbb{R}^{m-1}} f(\mathbf{x}) d\mathbf{x} \quad (1.2)$$

in the limit as  $N \rightarrow \infty$ . And we say the sequence is *Poissonian* if all the correlations for  $m \geq 2$  are Poissonian. By the method of moments, it can be shown that if a sequence is Poissonian then the gap distribution is exponential. Therefore it is common to study the correlations rather than the gap distribution itself.

That said, it remains a very difficult problem to show that a given sequence is Poissonian. While there are exceptions (e.g Rudnick-Zaharescu showed that almost every dilate of a lacunary sequence is Poissonian [RZ02]), there are very few results in the area. For example, if a sequence grows with a power law, then very little is known (especially when considering higher power correlations ( $m \geq 3$ )). Rather than study the correlations, in this paper we consider a coarser measure – the *long-range correlations*: fix  $\tau \in (0, 1)$

$$R^{(m)}(N, f, \tau) := \frac{1}{N} \sum_{\mathbf{j} \in \{1, \dots, N\}^m}^* f(N^\tau(\Delta(\mathbf{j}))). \quad (1.3)$$

Hence, in comparison with the standard correlations, we are increasing the support of the function  $f$  as  $N \rightarrow \infty$ . For uniformly distributed random variables on the interval one would expect:

$$R^{(m)}(N, f, \tau) = N^{(m-1)(1-\tau)} \left( \int_{\mathbb{R}^{m-1}} f(\mathbf{x}) d\mathbf{x} + o(1) \right). \quad (1.4)$$

In this paper, we give a general method to show that (for particular values of  $\tau > 0$ ) these long-range correlations converge to this limit, and thus coincide with the Poissonian limit. We show that these long-range correlations can be expressed in terms of the Poissonian limit (1.4) and an error which can be written in terms of Weyl sums. Then if we can appropriately bound these Weyl sums we can show that the error is small.

It should be noted that if the discrepancy of a sequence goes to 0 fast enough, then one can prove that the long-range correlations converge to the Poissonian limit for  $\tau < \frac{1}{2} - \epsilon$  for all  $\epsilon > 0$  (see [TW20] for details on this relation). The methodology in this paper, when applied to our examples, will improve on this 'naïve' bound.

## 1.2 Moments

The long-range correlations are the most natural object to work with from a mathematical point of view. However from an intuitive point of view one can also consider the following random variable:

$$W_{L,N} = W_{L,N}(Y) := \#\{n \leq N : x_n \in [Y, Y + L/N] \pmod{1}\}. \quad (1.5)$$

That is, we consider the number of points in a randomly placed interval of size  $\frac{L}{N}$ . Using a standard technique (see [TW20]) one can prove the following proposition

**Proposition 1.** *Fix a sequence  $\{x_n\}_{n \in \mathbb{N}}$  and a moment  $m \geq 2$ . Then for a given  $\tau < 1$ , suppose that for all  $f \in C_c^\infty(\mathbb{R}^{m-1})$  we have*

$$R^{(m)}(N, f, \tau) = N^{(m-1)(1-\tau)} \left( \int_{\mathbb{R}^{m-1}} f(\mathbf{x}) d\mathbf{x} + o(1) \right)$$

as  $N \rightarrow \infty$ . Then

$$\mathbf{E}(W_{N,L}^m) = L^m(1 + o(1)), \quad (1.6)$$

as  $N \rightarrow \infty$ , where  $L = N^{1-\tau}$ .

In words, this shows that, if we can show that the  $m$ -level long-range correlation converges to the Poissonian limit, then we can show that the  $m^{\text{th}}$  moment of  $W_{N,L}$  is (to leading order)  $L^m$ , as is the case for uniformly distributed random variables.

## 1.3 Specific Examples

In Section 2 we explain the general methodology for proving convergence as in (1.3) however, prior to this we give two applications to well-studied sequences.

### 1.3.1 The Sequence $a(n) = \alpha n^2$

If we consider  $a(n) = \alpha n^2$ , then the gaps between the sequence  $x_n$  are related to the gaps in the energy levels of the “boxed oscillator”, owing to this fact, and the sequence’s importance mathematically, this case has been extensively studied [RS98, RSZ01, HB10]. In particular Rudnick and Sarnak showed that there is a set of  $\alpha$  of full Lebesgue measure for which the pair correlation ( $m = 2$ ) is Poissonian. However very little is known about higher level correlations.

We say  $\alpha \in \mathbb{R}$  is Diophantine if, for every  $\epsilon > 0$  there is a  $c = c(\alpha) > 0$  such that

$$\left| \alpha - \frac{p}{q} \right| > \frac{c}{q^{2+\epsilon}} \quad (1.7)$$

for all integers  $p, q$ . Our first result, shows that for  $\alpha$  Diophantine and for  $\tau$  lying in a particular range, the long-range correlations for  $a(n) = \alpha n^2$  are Poissonian:

**Theorem 2.** *Let  $a(n) = \alpha n^2$  for  $\alpha$  Diophantine, fix  $m \geq 2$  and  $\epsilon > 0$ , then if  $0 < \tau \leq \frac{m}{2m-2} - \epsilon$  the long-range  $m$ -level correlation is Poissonian. Namely:*

$$R^{(m)}(N, f, \tau) = N^{(m-1)(1-\tau)} \left( \int_{\mathbb{R}^{m-1}} f(\mathbf{x}) d\mathbf{x} + o(1) \right), \quad (1.8)$$

as  $N \rightarrow \infty$ .

By Dirichlet’s approximation theorem, the set of Diophantine  $\alpha$  has full Lebesgue measure.

Theorem 2 has a number of implications. First of all, for  $m = 3$  this recovers a Theorem proved by Technau and Walker [TW20]. Namely that the long-range correlations are Poissonian for  $\tau < \frac{3}{4} - \epsilon$ . In addition, for  $m \geq 4$  this is the first result proving convergence of long-range correlations that goes beyond the discrepancy bound of  $\tau = \frac{1}{2} - \epsilon$ .

Interestingly, our result holds for all Diophantine  $\alpha$ . In [RS98], Rudnick and Sarnak give an example of a Diophantine  $\alpha$  for which the (not long-range) correlations are not Poissonian. Theorem 2 thus shows that for such examples, the non-Poissonian behavior happens on a very fine scale which is  $o(N^{-\tau})$ .

### 1.3.2 The Sequence $a(n) = \alpha\sqrt{n}$

A second sequence of particular interest is the sequence  $a(n) = \alpha\sqrt{n}$ . If we consider  $a(n) = \alpha n^\beta$  for  $0 < \beta < 1$ , then it is expected that for  $\beta \neq 1/2$ , the gap statistics converge to the exponential distribution. In contrast, in a surprising paper, Elkies and McMullen [EM04] used homogeneous dynamics methods to show that if  $\alpha^2 \in \mathbb{Q}$  and we consider  $a(n) = \alpha\sqrt{n}$ , then the gap distribution converges to an explicit distribution which is not the exponential. Moreover they conjectured that for  $\alpha^2 \notin \mathbb{Q}$  the gap distribution *is* exponential. Also surprisingly El-Baz-Marklof-Vinogradov [EBMV15] showed that for  $\alpha = 1$  the pair correlation *is* Poissonian. Therefore, in short,  $\beta = 1/2$  is a surprising, special case exhibiting unusual behavior.

**Theorem 3.** *Let  $a(n) = \alpha\sqrt{n}$  for any  $\alpha \in \mathbb{R}^*$ , fix any  $\epsilon > 0$ . Take  $\tau \leq \frac{3m}{6m-4} - \epsilon$ . Then*

$$R^{(m)}(N, f, \tau) = N^{(m-1)(1-\tau)} \left( \int_{\mathbb{R}} f(x) dx + o(1) \right), \quad (1.9)$$

as  $N \rightarrow \infty$ .

Note that Theorem 3 holds for any  $\alpha \in \mathbb{R}^*$ . Thus for  $m = 2$ , (1.9) agrees with [EBMV15], while for  $m \geq 3$ , (1.9) states that on this long-range scale all the correlations converge to the Poissonian limit, regardless of the value of  $\alpha$ . Therefore, this shows that the structure observed by Elkies-McMullen occurs on a very fine scale and isn't apparent on these long-range scales.

**Notation:** To avoid confusion, fix the following notation, let  $f(n) = \mathcal{O}(g(n))$  if  $\overline{\lim}_{n \rightarrow \infty} |f(n)/g(n)| < \infty$ , and  $f(n) = o(g(n))$  if  $\overline{\lim}_{n \rightarrow \infty} |f(n)/g(n)| = 0$ . Equivalently let  $f(n) \ll g(n)$  denote  $f(n) = \mathcal{O}(g(n))$ .

For  $x \in \mathbb{R}$  let  $\{x\}$  denote the fractional part of  $x$ , and we let  $\|x\|$  denote the distance to the nearest integer. Given a set  $\mathcal{A}$  we denote  $\mathcal{A}^* := \mathcal{A} \setminus \{0\}$ . Finally, as usual  $e(z) := e^{2\pi iz}$ .

## 2 Methodology

To achieve the convergence in (1.4) we first use discrete Fourier analysis, to show that the long-range  $m$ -level correlation can be written as a main term and a remainder which can be written as a product of Weyl sums. Then, for each example we use bounds on the associated Weyl sums to show that the remainder is sufficiently small. For  $\mathbf{y} \in \mathbb{R}^m$ , write

$$\boldsymbol{\beta}(\mathbf{y}) = (\beta_1(\mathbf{y}), \dots, \beta_{m-1}(\mathbf{y})) \quad (2.1)$$

where  $\beta_i(\mathbf{y}) := a(y_{i+1}) - a(y_i)$ . Thus

$$R^{(m)}(N, f, \tau) = \frac{1}{N} \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* \sum_{\mathbf{k} \in \mathbb{Z}^{m-1}} f(N^\tau(\boldsymbol{\beta}(\mathbf{y}) + \mathbf{k})). \quad (2.2)$$

Now apply Poisson summation to the sum over  $\mathbf{k}$  to give:

$$R^{(m)}(N, f, \tau) = \frac{1}{N} \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* \sum_{\mathbf{k} \in \mathbb{Z}^{m-1}} e(\mathbf{k} \cdot \boldsymbol{\beta}(\mathbf{y})) \int_{\mathbb{R}^{m-1}} f(N^\tau \mathbf{x}) e(\mathbf{x} \cdot \mathbf{k}) d\mathbf{x}. \quad (2.3)$$

Now we isolate the term  $\mathbf{k} = 0$  which gives us our main term:

$$R^{(m)}(N, f, \tau) = N^{(1-\tau)(m-1)} \left( \int f(\mathbf{x}) d\mathbf{x} + o(1) \right) + \mathcal{E} \quad (2.4)$$

where the  $o(1)$ -error comes from the fact that the  $y_i$  are taken to be distinct and

$$\mathcal{E} := \frac{1}{N} \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* \sum_{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^{m-1}} e(\mathbf{k} \cdot \boldsymbol{\beta}(\mathbf{y})) \int_{\mathbb{R}^{m-1}} f(N^\tau \mathbf{x}) e(\mathbf{x} \cdot \mathbf{k}) d\mathbf{x}.$$

Thus our goal in the remainder of the paper is to show that  $\mathcal{E} = o\left(N^{(1-\tau)(m-1)}\right)$ . To achieve this bound let  $M = N^{\tau+\epsilon'}$  for some  $\epsilon' > 0$  which we fix later, then we can use the fast decay of Fourier coefficients to show:

$$\begin{aligned} \mathcal{E} &= \frac{1}{N} \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* \sum_{\substack{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < M}} e(\mathbf{k} \cdot \boldsymbol{\beta}(\mathbf{y})) \int_{\mathbb{R}^{m-1}} f(N^\tau \mathbf{x}) e(\mathbf{x} \cdot \mathbf{k}) d\mathbf{x} + o(1) \\ &\ll \frac{1}{N^{1+\tau(m-1)}} \sum_{\substack{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* e(\mathbf{k} \cdot \boldsymbol{\beta}(\mathbf{y})) \right| + o(1). \end{aligned}$$

For notation set  $k_0 = k_m = 0$ , then we can change variables and write:

$$\begin{aligned} \mathcal{E} &\ll \frac{1}{N^{1+\tau(m-1)}} \sum_{\substack{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* \prod_{i=1}^m e(a(y_i)(k_{i-1} - k_i)) \right| + o\left(N^{(m-1)(1-\tau)}\right) \\ &= \frac{1}{N^{1+\tau(m-1)}} \sum_{\substack{\mathbf{0} \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* e(a(y_m)d(\mathbf{k})) \prod_{i=1}^{m-1} e(a(y_i)k_i) \right| + o\left(N^{(m-1)(1-\tau)}\right) \end{aligned}$$

where  $d(\mathbf{k}) = -k_1 - k_2 - \dots - k_{m-1}$ .

If we were considering the full correlations ( $\tau = 1$ ), then the diagonal terms  $y_i = y_j$  ( $i \neq j$ ) would be of the same size as the main term. However since we work with  $\tau < 1$ , the diagonal terms are in fact negligible. Thus we can add them back into the sum. To that end, we use the inclusion-exclusion principle to write

$$\sum_{\mathbf{y} \in \{1, \dots, N\}^m}^* = \sum_{\mathbf{y} \in \{1, \dots, N\}^m} - \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^{(2)}$$

where the (2) above sum signifies that at least 2 terms in the sum must be equal. By symmetry:

$$\begin{aligned}
& \frac{1}{N^{1+\tau(m-1)}} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^m}^{(2)} e(a(y_m)d(\mathbf{k})) \prod_{i=1}^{m-1} e(a(y_i)k_i) \right| \\
& \ll \frac{M^2}{N^{1+\tau(m-1)}} \sum_{\substack{\mathbf{k} \in \mathbb{Z}^{m-3} \\ |k_i| < 2M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^{m-2}} e(a(y_{m-2})d(\mathbf{k})) \prod_{i=1}^{m-1-2} e(a(y_i)k_i) \right| \\
& \ll \frac{M^2}{N^{1+\tau(m-1)}} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-3} \\ |k_i| < 2M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^{m-j}} e(a(y_{m-2})d(\mathbf{k})) \prod_{i=1}^{m-3} e(a(y_i)k_i) \right| + \frac{M^2}{N^{1+\tau(m-1)}} N^{m-2}
\end{aligned} \tag{2.5}$$

If we write:

$$\mathcal{E}_m := \frac{1}{N^{1+\tau(m-1)}} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} \left| \sum_{\mathbf{y} \in \{1, \dots, N\}^m} e(a(y_m)d(\mathbf{k})) \prod_{i=1}^{m-1} e(a(y_i)k_i) \right|$$

Then we have shown that, for  $\tau < 1$

$$\mathcal{E} \ll \mathcal{E}_m + \frac{M}{N^\tau} \mathcal{E}_{m-1} + o\left(N^{(m-1)(1-\tau)}\right). \tag{2.6}$$

Now we write

$$S(N, k) := \sum_{y=1}^N e(ka(y)).$$

In this case

$$\mathcal{E}_m \ll \frac{1}{N^{1+\tau(m-1)}} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}))| \prod_{i=1}^{m-1} |S(N, k_i)| \tag{2.7}$$

Thus we get the following theorem:

**Theorem 4.** *Let  $\tau < 1$ , then for  $m \geq 2$ :*

$$\frac{\mathcal{E}_m}{N^{(m-1)(1-\tau)}} \ll \frac{1}{N^m} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}))| \prod_{i=1}^{m-1} |S(N, k_i)|. \tag{2.8}$$

In what remains, we use Theorem 4 to show that (under the hypotheses stated in the main theorems of the introduction)  $\mathcal{E}_m = o(1)$ . Thus, since the range of  $\tau$  decreases with  $m$ , we can use an inductive argument to prove our main theorems.

*Remark.* Unfortunately Theorem 4, as it is stated, is not powerful enough to handle the case  $\tau = 1$  (i.e the standard correlations). The reason for this is that in that case the diagonal terms are of leading order. Therefore, to use the same approach would require a precise asymptotic for the Weyl sums  $S(N, k)$ . These kinds of precise asymptotics are rare in the field. Most likely, to handle the full correlations would require a multi-dimensional analysis as is used in [TW20].

### 3 Proof of Theorem 2

The proof of Theorem 2 is more-or-less a straightforward application of Theorem 4 and Weyl differencing (see e.g [Dav05, Lemma 3.1]). Namely, as was shown in [RS98, Corollary 5] for  $\alpha$  Diophantine, we have, for any  $\delta > 0$ :

$$\sum_{1 \leq k \leq M} |S(k, N)|^2 \ll M^{1+\delta} N^{1+\delta} \quad (3.1)$$

$$\sum_{1 \leq k \leq M} |S(k, N)| \ll M^{1+\delta} N^{1/2+\delta}. \quad (3.2)$$

Therefore if we consider the r.h.s in Theorem 4: since  $\mathbf{k} \neq 0$ , we can assume  $k_1 \neq 0$ :

$$\frac{1}{N^m} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}))| \prod_{i=1}^{m-1} |S(N, k_i)| \ll \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \sum_{k_1=1}^M |S(N, k_1)| |S(N, d(\mathbf{k}))|$$

where  $\mathbf{k}' = (y_2, \dots, y_{m-1})$ . First we handle the term  $d(\mathbf{k}) = 0$ . Note that setting  $d(\mathbf{k}) = 0$ , fixes the value of  $k_1 = d(\mathbf{k}') = -k_2 - k_3 - \dots - k_{m-1}$ . Moreover, we know that  $k_1 > 0$ , thus (applying the trivial bound  $S(N, 0) = N$ ),

$$\begin{aligned} \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \sum_{k_1=1}^M |S(N, k_1)| |S(N, 0)| &= \frac{1}{N^m} \sum_{\substack{0 \neq \mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}'))| |S(N, 0)| \prod_{i=2}^{m-1} |S(N, k_i)| \\ &= \frac{1}{N^{m-1}} \sum_{\substack{0 \neq \mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}'))| \prod_{i=2}^{m-1} |S(N, k_i)| \\ &= \frac{1}{N^{m-1}} \sum_{\substack{0 \neq \mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}'))| \prod_{i=2}^{m-1} |S(N, 0)| \end{aligned}$$

Thus we can use an inductive argument to bound the  $d(\mathbf{k}) = 0$  term (i.e this is the same as the right hand side in Theorem 4 for  $\mathcal{E}_{m-1}$ ). For the other term, we apply Cauchy Schwarz to the sum over  $k_1$  giving:

$$\frac{1}{N^m} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M \\ d(\mathbf{k}) \neq 0}} |S(N, d(\mathbf{k}))| \prod_{i=1}^{m-1} |S(N, k_i)| \ll \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \left( \sum_{k=1}^M |S(N, k)|^2 \sum_{k=1}^M |S(N, k)|^2 \right)^{1/2},$$

where we have used the fact that  $0 < |d(\mathbf{k})| \leq mM$  for any value of  $k_1$ . Now we can use (3.1) on the final bracket, and for each of the remaining  $k_i$ -sums we use (3.2) and the trivial bound  $S(N, 0) = N$ . Thus

$$\begin{aligned} \frac{1}{N^m} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}))| \prod_{i=1}^{m-1} |S(N, k_i)| &\ll \frac{1}{N^m} \left( N + M^{1+\delta} N^{1/2+\delta} \right)^{m-2} \left( M^{1+\delta} N^{1+\delta} \right) \\ &\ll \frac{1}{N^m} M^{(1+\delta)(m-1)} N^{m(1/2+\delta)} + \frac{M^{1+\delta} N^\delta}{N} \\ &\ll \frac{M^{m-1+\delta'}}{N^{m/2}} + \frac{M^{1+\delta} N^\delta}{N}. \end{aligned} \quad (3.3)$$

Recall, we want to show  $\frac{\mathcal{E}_m}{N^{(m-1)(1-\tau)}} = o(1)$ . Therefore, provided  $M = N^{\frac{m}{2m-2}-\epsilon}$ , for some  $\epsilon > \delta' \left( \frac{m}{2(m-1)(m-1+\delta')} \right)$ , then the l.h.s of (3.3) is  $o(1)$ . Inserting this bound into (2.8) and using (2.6) proves Theorem 2.  $\square$

## 4 Proof of Theorem 3

For the proof of Theorem 3, we have not found any suitable bounds on Weyl sums in the literature. Fortunately we can use the established theory of exponential sums to prove the following:

**Theorem 5.** *Let  $0 < |k| \leq CM$  for any constant  $C < \infty$ , then as  $N \rightarrow \infty$ :*

$$S(N, k) \ll k^{1/2} N^{1/4} + \frac{N^{3/4}}{k^{1/2}} \log(N). \quad (4.1)$$

With that, Theorem 3 follows almost immediately from (2.8). First, note that: for any fixed constant  $C > 0$

$$\sum_{0 < |k_i| < CM} |S(N, k)|^2 \ll M^2 N^{1/2} + N^{3/2} \log(N)^2 + NM \log(N) \quad (4.2)$$

$$\sum_{0 < |k_i| < CM} |S(N, k)| \ll M^{3/2} N^{1/4} + N^{3/4} M^{1/2} \log(N) \quad (4.3)$$

as  $N \rightarrow \infty$ .

Just as we did in Section 3, we note that if  $d(\mathbf{k}) = 0$  then  $k_1 = -k_2 - \dots - k_{m-1} = d(\mathbf{k}')$ . Thus

$$\begin{aligned} \frac{1}{N^m} \sum_{\substack{0 \neq \mathbf{k} \in \mathbb{Z}^{m-1} \\ |k_i| < 2M}} |S(N, d(\mathbf{k}))| \prod_{i=1}^{m-1} |S(N, k_i)| &\ll \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \sum_{k_1=1}^{2M} |S(N, k_1)| |S(N, d(\mathbf{k}))|, \\ &\ll \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \sum_{\substack{k_1=1 \\ d(\mathbf{k}) \neq 0}}^{2M} |S(N, k_1)| |S(N, d(\mathbf{k}))| \\ &\quad + \frac{1}{N^{m-1}} \sum_{\substack{0 \neq \mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| |S(N, d(\mathbf{k}'))|. \end{aligned}$$

Thus (as before), we can use an inductive argument on  $m$  to show that the latter term in the last line is sufficiently small. For the former term we apply Cauchy-Schwarz to get

$$\begin{aligned} \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \sum_{\substack{k_1=1 \\ d(\mathbf{k}) \neq 0}}^{2M} |S(N, k_1)| |S(N, d(\mathbf{k}))| &\ll \frac{1}{N^m} \sum_{\substack{\mathbf{k}' \in \mathbb{Z}^{m-2} \\ |k_i| < 2M}} \prod_{i=2}^{m-1} |S(N, k_i)| \sum_{k_1=1}^{mM} |S(N, k_1)|^2 \\ &\ll \frac{1}{N^m} \left( M^{3/2} N^{1/4} \right)^{m-2} \left( N^{1/2} M^2 \right) \\ &\ll \frac{1}{N^{\frac{3m}{4}}} M^{\frac{3m-2}{2}} \end{aligned}$$

(we have assumed here that  $\tau \geq \frac{1}{2}$ ). Now inserting  $M = N^\tau$  and recalling that we want  $\frac{\mathcal{E}_m}{N^{(m-1)(1-\tau)}} = o(1)$ , this implies:

$$\frac{3m-2}{2}\tau - \frac{3m}{4} < 0 \quad \implies \quad \tau < \frac{3m}{2(3m-2)}$$

as in Theorem 3. □

## 5 Exponential Sum Bounds

This section is devoted to the proof of Theorem 5, which is based on Van der Corput's B-Process – see [Mon94, Chapter 3] for an excellent reference. However, for our problem, the range of the sum  $[1, N]$  is very large. As a consequence we do not have good control on the derivatives of the exponent. Hence we need to prove the stationary phase integrals rather than cite a known reference. Then we will use some classical analytic tools to control the various terms which arise. For simplicity of notation we assume  $k > 0$ . Then to achieve the same bounds for negative  $k$  take complex conjugates.

### 5.1 Analytic Tools

For a real-valued function  $g(x)$  defined on  $[A, B]$  we define:

$$V(g) := \max_{(A,B)} |g(x)| + t.v(g(x)), \quad (5.1)$$

where  $t.v(g(x)) := \int_A^B |g'(x)| dx$ . For a proof of the following classical lemma see [Hux96, Lemma 5.1.4]

**Lemma 6** (Van der Corput's  $k^{\text{th}}$  Derivative Test). *Let  $f(x)$  be real and  $k$  times differentiable on  $(A, B)$  with  $f^{(r)}(x) \geq \mu > 0$  on  $(A, B)$ . Let  $g(x)$  be a real valued function. Then:*

$$\left| \int_A^B g(x)e(f(x))dx \right| \ll \frac{V(g)}{\mu^{1/r}}. \quad (5.2)$$

#### 5.1.1 Stationary Phase Integrals

If  $\frac{k\alpha}{2\sqrt{N}} \leq r \leq \frac{k\alpha}{2}$ , then the function  $h_{k,r}(x) = k\alpha\sqrt{x} - rx$  has a stationary point,  $\gamma_{k,r} = \left(\frac{k\alpha}{2r}\right)^2 \in (1, N)$  where  $h'_{k,r}(\gamma_{k,r}) = 0$ . In that case the following is an application of classical stationary phase estimates (see for example [Hux96, Lemma 5.5.2]):

**Proposition 7** (Stationary Phase Integral). *Let  $r, k \in \mathbb{Z}$  be integers such that,  $k \in \{1, 2, \dots, N\}$  and  $\frac{k\alpha}{2\sqrt{N}} \leq r < \frac{k\alpha}{2}$ . Then in the limit as  $N \rightarrow \infty$ :*

$$\int_1^N e(h_{k,r}(x)) dx = e\left(\frac{(k\alpha)^2}{4r} + \frac{1}{8}\right) \frac{\gamma_{k,r}^{3/4}}{k^{1/2}} + \mathcal{O}\left(\frac{1}{h'_{k,r}(1)} + \frac{1}{h'_{k,r}(N)}\right) + \mathcal{O}\left(\frac{N^{1/4}}{k^{3/2}} + \frac{N^{1/2}}{k^2}\right). \quad (5.3)$$

*Proof.* For this proof we will suppress the subscripts  $k, r$  whenever they appear. To prove (5.3) we fix two constants  $0 < c_1 < 1 < c_2 < \infty$  independent of  $N$  and fix  $u := c_1\gamma$  and  $v := c_2\gamma$ . Now we consider the three integrals on  $(1, u)$ ,  $(u, v)$ , and  $(v, N)$  separately.

**Step 1:** To address the integral on  $(1, u)$ , apply integration by parts, to give:

$$\int_1^u e(h(x))dx = \left[ \frac{e(h(x))}{2\pi i h'(x)} \right]_{x=1}^u + \frac{1}{2\pi i} \int_1^u \frac{h''(x)e(h(x))}{h'(x)^2} dx \quad (5.4)$$

Now if  $u < 1$ , then the term of the right is  $\mathcal{O}\left(\frac{1}{h'(1)}\right)$ . If not, then  $r < \frac{k\alpha\sqrt{c_1}}{2}$ . Thus, applying the first integral test gives:

$$\int_1^u e(h(x))dx = \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=1}^u + \mathcal{O}\left(V\left(\frac{h''(x)}{h'(x)^2}\right) (\min|h'(x)|)^{-1}\right) \quad (5.5)$$

Now note that on  $(1, u)$ , the derivative  $|h'(x)| \geq C\frac{k}{\gamma^{1/2}}$  where the constant  $C = \alpha\left(\frac{1}{\sqrt{c_1}} - 1\right)$ . Thus

$$\begin{aligned} \int_1^u e(h(x))dx &= \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=1}^u + \mathcal{O}\left(\frac{\gamma^{1/2}}{k} \left[\frac{h''(x)}{h'(x)^2}\right]_1^u\right) \\ &= \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=1}^u + \mathcal{O}\left(\frac{\gamma^{1/2}}{k} \left(\frac{k}{(k\alpha - r)^2} + \frac{1}{\gamma^{1/2}k}\right)\right) \\ &= \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=1}^u + \mathcal{O}\left(\frac{N^{1/2}}{k^2} + \frac{1}{k^2}\right). \end{aligned} \quad (5.6)$$

**Step 2:** The proof for  $(v, N)$  is similar: First we use the first derivative test (Lemma 6) to show:

$$\begin{aligned} \int_v^N e(h(x))dx &= \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=v}^N + \frac{1}{2\pi i} \int_v^q \frac{h''(x)e(h(x))}{h'(x)^2} dx \\ &= \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=v}^N + \mathcal{O}\left(\frac{V\left(\frac{h''(x)}{h'(x)^2}\right)}{\min|h'(x)|}\right) \end{aligned}$$

Moreover, evaluating  $V$  gives:

$$\frac{V\left(\frac{h''(x)}{h'(x)^2}\right)}{\min|h'(x)|} = \mathcal{O}\left(\frac{1}{k^2}\right). \quad (5.7)$$

Therefore:

$$\int_v^N e(h(x))dx = \left[\frac{e(h(x))}{2\pi i h'(x)}\right]_{x=v}^N + \mathcal{O}\left(\frac{1}{k^2}\right). \quad (5.8)$$

**Step 3:** Working in  $[u, v]$ , set  $T = k\alpha\gamma^{1/2}$  and  $M = \gamma$ . Then for  $j = 2, 3, 4$ , there exist constants  $0 < C_j < \infty$  such that

$$\left|h^{(j)}(x)\right| \leq C_j \frac{T}{M^j} \quad , \quad h''(x) \geq h''(v) \geq C_2^{-1} \frac{T}{M^2}$$

Thus we can apply a classical stationary phase integral [Hux96, Lemma 5.5.2] to say:

$$\begin{aligned} \int_u^v h(x)dx &= \frac{e(h(\gamma) + 1/8)}{\sqrt{h''(\gamma)}} + \frac{e(h(v))}{2\pi i h'(v)} - \frac{e(h(u))}{2\pi i h'(u)} + \mathcal{O}\left(\frac{M}{T^{3/2}} + \frac{M^4}{T^2} \left(\frac{1}{\gamma}\right)^3\right) \\ &= \frac{e(h(\gamma) + 1/8)}{\sqrt{h''(\gamma)}} + \frac{e(h(v))}{2\pi i h'(v)} - \frac{e(h(u))}{2\pi i h'(u)} + \mathcal{O}\left(\frac{\gamma}{k^{3/2}\gamma^{3/4}} + \frac{\gamma^3}{k^2\gamma^3}\right) \\ &= \frac{e(h(\gamma) + 1/8)}{\sqrt{h''(\gamma)}} + \frac{e(h(v))}{2\pi i h'(v)} - \frac{e(h(u))}{2\pi i h'(u)} + \mathcal{O}\left(\frac{\gamma^{1/4}}{k^{3/2}} + \frac{1}{k^2}\right). \end{aligned} \quad (5.9)$$

(5.3), now follows directly from this bound together with (5.6) and (5.8).  $\square$

## 5.2 Proof of Theorem 5

Our goal is to estimate

$$S = S(N, k) = \sum_{y \in \{1, \dots, N\}} e(k\alpha\sqrt{y}).$$

**Step 1:** First, apply the truncated Poisson summation formula [Hux96, Lemma 5.4.3]: For  $A := \frac{k\alpha}{2\sqrt{N}}$  and  $B := \frac{k\alpha}{2}$ :

$$S = \sum_{A - \frac{1}{4} < r < B + \frac{1}{4}} \int_1^N e(k\alpha\sqrt{x} - rx) dx + \mathcal{O}(\log(N)).$$

First note that if  $r = 0$  then  $\int_1^N e(k\alpha\sqrt{x}) dx \ll \frac{N^{1/2}}{k}$  by the first derivative test. Thus, if we set  $\tilde{A} = \min(1, \lceil A - 1/4 \rceil)$  and  $\tilde{B} := \lfloor B + 1/4 \rfloor$ , and apply Proposition 7 and Lemma 6 (2<sup>nd</sup> derivative test) then we have the bound:

$$S = \sum_{\tilde{A} \leq r \leq \tilde{B}} \left( e \left( \frac{(k\alpha)^2}{4r} + \frac{1}{8} \right) \frac{\gamma_{k,r}^{3/4}}{k^{1/2}} + \mathcal{O} \left( \min \left\{ \frac{1}{h'_{k,r}(1)}, \frac{N^{3/4}}{k^{1/2}} \right\} + \min \left\{ \frac{1}{h'_{k,r}(N)}, \frac{N^{3/4}}{k^{1/2}} \right\} \right) \right) + \mathcal{O} \left( \frac{N^{1/2}}{k} + \frac{N^{1/4}}{k^{1/2}} + \log(N) \right). \quad (5.10)$$

We have used the fact that there are order  $k$  terms in the sum over  $r$  to pull two of the error terms out of the sum (and the resulting error is  $\mathcal{O}(k^{1/2})$ ).

**Step 2:** The latter two terms in the sum can both be controlled in the same way:

$$\sum_{\tilde{A} \leq r \leq \tilde{B}} \min \left\{ \frac{1}{h'_{k,r}(1)}, \frac{N^{3/4}}{k^{1/2}} \right\} \ll \left| \sum_{\tilde{A} \leq r \leq \tilde{B}} \min \left\{ \frac{1}{\frac{k\alpha}{2} - r}, \frac{N^{3/4}}{k^{1/2}} \right\} \right| \ll \log(N) + \min \left\{ \frac{1}{\|\frac{k\alpha}{2}\|}, \frac{N^{3/4}}{k^{1/2}} \right\}. \quad (5.11)$$

The same bound holds for the term involving  $h'_{k,r}(N)$ . Thus:

$$S = \sum_{\tilde{A} \leq r \leq \tilde{B}} e \left( \frac{(k\alpha)^2}{4r} + \frac{1}{8} \right) \frac{\gamma_{k,r}^{3/4}}{k^{1/2}} + \mathcal{O} \left( \frac{N^{3/4}}{k^{1/2}} + \frac{N^{1/2}}{k} + \log(N) \right). \quad (5.12)$$

**Step 3:** Inserting the definition of  $\gamma_{k,r}$ , the remaining term can be written:

$$\begin{aligned} E &:= k \sum_{\tilde{A} < r < \tilde{B}} e \left( \frac{(k\alpha)^2}{4r} \right) \frac{1}{r^{3/2}}, \\ &\leq k \sum_{\tilde{A} < r < \tilde{B}} \frac{1}{r^{3/2}} \\ &\ll k \frac{1}{\tilde{A}^{1/2}} \ll N^{1/4} k^{1/2}. \end{aligned} \quad (5.13)$$

□

## Acknowledgements

For part of this work, the author was supported by EPSRC Studentship EP/N509619/1 1793795. Furthermore, the author would like to thank Jens Marklof for several useful discussions and Niclas Technau for comments on an early preprint of the paper.

## References

- [BT77] M. Berry and M. Tabor. Level clustering in the regular spectrum. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 356(1686):375–394, 1977.
- [Dav05] H. Davenport. *Analytic methods for Diophantine equations and Diophantine inequalities*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, second edition, 2005. With a foreword by R. C. Vaughan, D. R. Heath-Brown and D. E. Freeman, Edited and prepared for publication by T. D. Browning.
- [EBMV15] D. El-Baz, J. Marklof, and I. Vinogradov. The two-point correlation function of the fractional parts of  $\sqrt{n}$  is Poisson. *Proceedings of the American Mathematical Society*, 143(7):2815–2828, 2015.
- [EM04] N. Elkies and C. McMullen. Gaps in  $\sqrt{n} \bmod 1$  and ergodic theory. *Duke Mathematical Journal*, 123(1):95–139, 2004.
- [HB10] D. R. Heath-Brown. Pair correlation for fractional parts of  $\alpha n^2$ . *Mathematical Proceedings of the Cambridge Philosophical Society*, 148(3):385–407, 2010.
- [Hux96] M. N. Huxley. *Area, lattice points, and exponential sums*, volume 13 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1996. Oxford Science Publications.
- [KN74] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences*. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974. Pure and Applied Mathematics.
- [Mar00] J. Marklof. The Berry-Tabor conjecture. In *Proceedings of the 3<sup>rd</sup> European Congress of Mathematics*, volume 202, pages 421–427, Barcelona, 2000. Birkhäuser, Basel.
- [Mon94] H. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics*. Published for the CBMS, Washington, DC; by the AMS, Providence, RI, 1994.
- [RS98] Z. Rudnick and P. Sarnak. The pair correlation function of fractional parts of polynomials. *Communications in Mathematical Physics*, 194(1):61–70, 1998.
- [RSZ01] Z. Rudnick, P. Sarnak, and A. Zaharescu. The distribution of spacings between the fractional parts of  $n^2\alpha$ . *Inventiones Mathematicae*, 145(1):37–57, 2001.
- [RZ02] Z. Rudnick and A. Zaharescu. The distribution of spacings between fractional parts of lacunary sequences. *Forum Mathematicum*, 14(5):691–712, 2002.
- [TW20] N. Technau and A. Walker. the triple correlations of fractional parts of  $\alpha n^2$ . *arXiv:2005.01490 [math.NT]*, 2020.