On the discrepancy of permuted lacunary series

C. Aistleitner,[*] I. Berkes[†] and R. Tichy[‡]

Abstract

It is known that for any smooth periodic function f the sequence $(f(2^k x))_{k \geq 1}$ behaves like a sequence of i.i.d. random variables, for example, it satisfies the central limit theorem and the law of the iterated logarithm. Recently Fukuyama showed that a permutation of $(f(2^k x))_{k \geq 1}$ can ruin the validity of the law of the iterated logarithm, a very surprising result. In this paper we present an optimal condition on $(n_k)_{k \geq 1}$, formulated in terms of the number of solutions of certain Diophantine equations, which ensures the validity of the law of the iterated logarithm for any permutation of the sequence $(f(n_k x))_{k \geq 1}$. A similar result is proved for the discrepancy of the sequence $(\{n_k x\})_{k \geq 1}$, where $\{\cdot\}$ denotes fractional part.

1 Introduction

Given a sequence $(x_1, \dots, x_N)$ of real numbers, the value

$$D_N = D_N(x_1, \dots, x_N) = \sup_{0 \leq a < b < 1} \left| \frac{\sum_{k=1}^N \mathbb{1}_{[a,b)}(\{x_k\})}{N} - (b-a) \right|$$

is called the discrepancy of the sequence. Here $\mathbb{1}_{[a,b)}$ denotes the indicator function of the interval $[a,b)$ and $\{\cdot\}$ denotes fractional part. An infinite sequence $(x_n)_{n \geq 1}$ is called uniformly distributed mod 1 if $D_N(x_1, \dots x_N) \to 0$ as $N \to \infty$. Weyl [24]

[*]Graz University of Technology, Department for Analysis and Computational Number Theory, Steyrergasse 30, 8010 Graz, Austria e-mail: aistleitner@math.tugraz.at. Research supported by FWF grant S9603-N23.

[†]Graz University of Technology, Institute of Statistics, Münzgrabenstraße 11, 8010 Graz, Austria. e-mail: berkes@tugraz.at. Research supported by FWF grant S9603-N23 and OTKA grants K 67961 and K 81928.

[‡]Graz University of Technology, Department for Analysis and Computational Number Theory, Steyrergasse 30, 8010 Graz, Austria. e-mail: tichy@tugraz.at. Research supported by FWF grant S9603-N23.

AMS 2010 Subject classification. 42C15, 42A55, 42A61, 11K38, 60G50, 60F15

Key words and phrases: lacunary series, permutation-invariance, Diophantine equations, law of the iterated logarithm

proved that for any increasing sequence $(n_k)_{k\geq 1}$ of integers, $(n_k x)_{k\geq 1}$ is uniformly distributed mod 1 for almost all $x \in \mathbb{R}$ in the sense of the Lebesgue measure. Computing the discrepancy of this sequence is a difficult problem and the precise asymptotics is known only in a few cases. Philipp [19] proved that if $(n_k)_{k\geq 1}$ satisfies the Hadamard gap condition

$$n_{k+1}/n_k \geq q > 1 \qquad (k = 1, 2, \ldots), \tag{1}$$

then the discrepancy of $(\{n_k x\})_{k\geq 1}$ obeys the law of the iterated logarithm, i.e.

$$\frac{1}{4\sqrt{2}} \leq \limsup_{N\to\infty} \frac{ND_N(n_k x)}{\sqrt{2N \log\log N}} \leq D(q) \quad \text{a.e.}, \tag{2}$$

where $D(q)$ is a number depending on q. (For the simplicity of the notation, we will write $D_N(x_k)$ instead of $D_N(x_1, \ldots x_N)$.) Note that for the discrepancy of an i.i.d. nondegenerate sequence $(X_k)_{k\geq 1}$ we have

$$\limsup_{N\to\infty} \frac{ND_N(X_k)}{\sqrt{2N \log\log N}} = 1/2 \quad \text{a.s.} \tag{3}$$

by the Chung-Smirnov law of the iterated logarithm (see e.g. [22, p. 504]). A comparison of (2) and (3) shows that for Hadamard lacunary $(n_k)_{k\geq 1}$, the sequence $(\{n_k x\})_{k\geq 1}$ of functions on $(0,1)$ behaves like a sequence of i.i.d. random variables. (Note, however, that for certain values of x the distribution of $(\{n_k x\})_{k\geq 1}$ can differ significantly from the uniform distribution even under (1), see e.g. [18].) Interestingly, the analogy between lacunary sequences and sequences of i.i.d. random variables is not complete. Fukuyama [8] determined the limsup in (2) for the sequences $n_k = \theta^k$, $\theta > 1$; his results show that the limsup is different from 1/2 for any integer $\theta \geq 2$. (On the other hand, in [8] it is shown that if θ^r is irrational for $r = 1, 2, \ldots$, then the limsup in (2) equals to the i.i.d. value 1/2.) Aistleitner [2] constructed a Hadamard lacunary sequence $(n_k)_{k\geq 1}$ such that the limsup in (2) is not a constant a.e. and Fukuyama and Miyamoto [11] showed that this actually happens for $n_k = 2^k - 1$. Even more surprisingly, Fukuyama [9] showed that even if the limsup in (2) is a constant (e.g., for $n_k = 2^k$), the value of the limsup can change by permuting the sequence $(n_k)_{k\geq 1}$, a phenomenon radically different from i.i.d. behavior, which is clearly permutation-invariant.

The previous results show that the behavior of $(\{n_k x\})_{k\geq 1}$ is very delicate, exhibiting both probabilistic and number-theoretic phenomena. It is natural to ask for which $(n_k)_{k\geq 1}$ the behavior of $(\{n_k x\})_{k\geq 1}$ follows exactly i.i.d. behavior, for example, when is the limsup in (2) equal to 1/2 a.e and under what conditions is the value of the limsup permutation-invariant. A near optimal number-theoretic condition for $\limsup = 1/2$ a.e. was given by Aistleitner [3] and the purpose of the present paper is to give an optimal condition for the permutation-invariance of the LIL.

Let f be a measurable function satisfying

$$f(x+1) = f(x), \quad \int_0^1 f(x)dx = 0, \quad \mathrm{Var}_{[0,1]}(f) < \infty. \tag{4}$$

A profound study of the behavior of the sequence $(f(n_k x))_{k\geq 1}$ for Hadamard lacunary $(n_k)_{k\geq 1}$ was given in Gaposhkin [12, 13]. By a classical theorem of Kac [15], under (4) the sequence $(f(2^k x))_{k\geq 1}$ satisfies the CLT and Erdős and Fortet showed (see [16, p. 646]) that the CLT generally fails for $(f((2^k-1)x))_{k\geq 1}$ (see also [6]). Gaposhkin showed that $(f(n_k x))_{k\geq 1}$ satisfies the central limit theorem provided n_{k+1}/n_k is an integer for all $k \geq 1$ or if $n_{k+1}/n_k \to \alpha$, where α^r is irrational for $r = 1, 2, \ldots$. More generally, he showed that $(f(n_k x))_{k\geq 1}$ satisfies the central limit theorem provided that for any nonzero integers a, b, c the number of solutions of the Diophantine equation

$$a n_k + b n_l = c, \qquad 1 \leq k, l \leq N \tag{5}$$

is bounded by a constant $K(a,b)$, independent of c. Aistleitner and Berkes [4] showed that the CLT remains valid if for any nonzero integers a, b, c the number of solutions of (5) is $o(N)$, uniformly in c, and this condition is best possible. Aistleitner [3] also proved that replacing $o(N)$ by $O(N/(\log N)^{1+\varepsilon})$ in the previous theorem, the limsup in (2) equals 1/2. As we will see, a two-term Diophantine condition will also give the precise condition for the permutation-invariance of the LIL for $D_N(n_k x)$.

2 Results

In what follows, we write $\|f\|$ for the $L^2(0,1)$ norm of a function f.

Theorem 1 *Let $(n_k)_{k\geq 1}$ be a sequence of positive integers satisfying (1), such that for any fixed integers $a \neq 0$, $b \neq 0$, c the number of solutions of the Diophantine equation (5) is bounded by a constant $K(a,b)$ independent of c, where for $c = 0$ we require also $k \neq l$. Let f be a function satisfying (4). Then for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$ we have*

$$\limsup_{N\to\infty} \frac{\sum_{k=1}^N f(n_{\sigma(k)} x)}{\sqrt{2N \log\log N}} = \|f\| \quad \text{a.e.}$$

As a consequence of Theorem 1 we obtain the following metric discrepancy result.

Theorem 2 *Let $(n_k)_{k\geq 1}$ be a sequence of positive integers satisfying (1), such that for any fixed integers $a \neq 0$, $b \neq 0$, c the number of solutions of the Diophantine equation (5) is bounded by a constant $K(a,b)$ independent of c, where for $c = 0$ we require also $k \neq l$. Then for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$ we have*

$$\limsup_{N\to\infty} \frac{N D_N(n_{\sigma(k)} x)}{\sqrt{2N \log\log N}} = \frac{1}{2} \quad \text{a.e.} \tag{6}$$

At the end of our paper we will show that if there exist integers $a \neq 0, b \neq 0$ and c such that the Diophantine equation (5) has infinitely many solutions $(k, l), k \neq l$, then the conclusion of Theorem 1 fails to hold for appropriate f. In fact, we can even obtain a non-constant limsup, which perfectly matches the results in [1, 2]. This shows that the Diophantine condition in Theorem 1 is essentially optimal.

We stress that in Theorems 1 and 2 we bounded the number of solutions of (5) also for $c = 0$ and thus $n_k = 2^k$ does not satisfy this condition. In fact, the conclusion of both theorems is false for $n_k = 2^k$: with the identity permutation $\sigma(k) = k$ the limsup in Theorem 1 equals

$$\left(\|f\|^2 + 2 \sum_{k=1}^{\infty} \int_0^1 f(x) f(2^k x)\, dx \right)^{1/2}$$

(see [14, 17]) and the limsup in Theorem 2 is $\sqrt{42}/9$ by the theorem of Fukuyama [8]. As we mentioned in the Introduction, for the CLT with a limit distribution of unspecified variance, it suffices to bound the number of solutions of (5) for coefficients a, b, c all different from 0.

As the proof of Theorem 1 will show, in the case when f is a trigonometric polynomial of degree d, it suffices to assume the Diophantine condition only with coefficients a, b satisfying $1 \le |a| \le d, 1 \le |b| \le d$. In particular, in the trigonometric case $f(x) = \cos 2\pi x$ it suffices to allow only coefficients ± 1, when the Diophantine condition in Theorem 1 is satisfied for any Hadamard lacunary sequence $(n_k)_{k \ge 1}$ (see e.g. Zygmund [25, pp. 203-204]). Thus we obtain the following corollary of Theorem 1, which is a permutation invariant version of the Erdős-Gál LIL in [7].

Theorem 3 *Let $(n_k)_{k \ge 1}$ be a sequence of positive integers satisfying the Hadamard gap condition (1), and let $\sigma : \mathbb{N} \to \mathbb{N}$ be a permutation of the set of positive integers. Then*

$$\limsup_{N \to \infty} \frac{\sum_{k=1}^N \cos 2\pi n_{\sigma(k)} x}{\sqrt{2N \log\log N}} = \frac{1}{\sqrt{2}} \qquad \text{a.e.} \tag{7}$$

If we assume the slightly stronger gap condition

$$n_{k+1}/n_k \to \infty \tag{8}$$

instead of Hadamard's gap condition (1), then the behavior of $f(n_k x)$ is permutation-invariant, regardless the number theoretic structure of $(n_k)_{k \ge 1}$. In fact, any sequence $(n_k)_{k \ge 1}$ satisfying the gap condition (8) satisfies the Diophantine condition in Theorem 1 automatically. This follows from the fact that for arbitrary fixed nonzero integers a, b the set-theoretic union of the sequences $(a n_k)_{k \ge 1}$ and $(b n_k)_{k \ge 1}$, arranged in increasing order, satisfies the Hadamard gap condition (1) and consequently the Diophantine condition in Theorem 1. Thus Theorem 1 implies the following

Theorem 4 *Let $(n_k)_{k \ge 1}$ be a sequence of positive integers satisfying the gap condition (8). Then for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the integers and for any function f satisfying (4) we have*

$$\limsup_{N \to \infty} \frac{\sum_{k=1}^N f(n_{\sigma(k)} x)}{\sqrt{2N \log\log N}} = \|f\| \quad \text{a.e.} \tag{9}$$

Moreover, for any permutation σ of $\mathbb{N}$ we have

$$\limsup_{N \to \infty} \frac{N D_N(n_{\sigma(k)} x)}{\sqrt{2N \log\log N}} = \frac{1}{2} \quad \text{a.e.} \tag{10}$$

3 Proofs

To deduce Theorem 2 from Theorem 1, we write $\mathbf{I}_{[a,b)}$ for the indicator of the interval $[a, b)$, centered at expectation and extended with period 1, i.e.

$$\mathbf{I}_{[a,b)}(x) = \mathbb{1}_{[a,b)}(\langle x \rangle) - (b - a),$$

where $\langle \cdot \rangle$ denotes the fractional part. For lacunary sequences (and also for permutations of lacunary sequences)

$$\limsup_{N\to\infty} \sup_{0\le a<b\le 1} \frac{\left|\sum_{k=1}^N \mathbf{I}_{[a,b)}(n_k x)\right|}{\sqrt{2N \log\log N}} = \sup_{0\le a<b\le 1} \limsup_{N\to\infty} \frac{\left|\sum_{k=1}^N \mathbf{I}_{[a,b)}(n_k x)\right|}{\sqrt{2N \log\log N}} \quad \text{a.e.}$$

(see [10, Theorem 1]). Thus Theorem 1 yields

$$\begin{aligned} &\limsup_{N\to\infty} \frac{N D_N(n_{\sigma(k)} x)}{\sqrt{2N \log\log N}} \\ = &\sup_{0\le a<b\le 1} \limsup_{N\to\infty} \frac{\left|\sum_{k=1}^N \mathbf{I}_{[a,b)}(n_{\sigma(k)} x)\right|}{\sqrt{2N \log\log N}} = \sup_{0\le a<b\le 1} \left\|\mathbf{I}_{[a,b)}\right\| = \frac{1}{2} \end{aligned}$$

almost everywhere.

We now turn to the proof of Theorem 1. In the sequel we will assume that the function $f(x)$, the sequence $(n_k)_{k\ge1}$ and the permutation σ are fixed. We assume that sequence $(n_k)_{k\ge1}$ is a lacunary sequence satisfying the Diophantine condition in Theorem 1. The method which we use for the proof of Theorem 1 is a multidimensional version of the method which was used in to prove the permutation-invariant CLT in our paper [5]. We recommend [5] as an introduction to the methods which are used in the present paper.

We begin with some auxiliary results.

Lemma 1 *Let $A_1, A_2, \dots$ be arbitrary events, satisfying*

$$\sum_{m=1}^{\infty} \mathbb{P}(A_m) = \infty$$

and

$$\liminf_{N\to\infty} \frac{\sum_{n=1}^N \sum_{m=1}^N \mathbb{P}(A_m A_n)}{\left(\sum_{m=1}^N \mathbb{P}(A_m)\right)^2} = 1.$$

Then

$$\mathbb{P}\left(\prod_{n=1}^{\infty} \sum_{m=n}^{\infty} A_m\right) = 1,$$

i.e. with probability 1 infinitely many of the events A_m occur.

A proof of this lemma can be found in [20].

Lemma 2 *Let P_1, P_2 be probability measures on $\mathbb{R}^2$, and write p_1, p_2 for the corresponding characteristic functions. Then for all $T_1, T_2, \delta_1, \delta_2, x, y > 0$*

$$\begin{aligned} & |P_1^*([-x,x]\times[-y,y]) - P_2^*([-x,x]\times[-y,y])| \\ \leq\ & xy \int_{(s,t)\in[-T_1,T_1]\times[-T_2,T_2]} |p_1(s,t)-p_2(s,t)|\ d(s,t) \\ & + xy\left(\delta_1^{-1}\delta_2\ \exp\left(-T_1^2\delta_1^2/2\right) + \delta_1\delta_2^{-1}\ \exp\left(-T_2^2\delta_2^2/2\right)\right), \end{aligned}$$

where

$$P_1^* = P_1 \star H, \quad P_2^* = P_2 \star H,$$

and H is a two-dimensional normal distribution with density

$$(2\pi\delta_1\delta_2)^{-1}\ e^{-\delta_1^2 u^2/2 - \delta_2^2 v^2/2}.$$

Proof: Assume that $T_1, T_2, \delta_1, \delta_2 > 0$ are fixed. Letting

$$h(s,t) = e^{-\delta_1^2 s^2/2 - \delta_2^2 t^2/2}$$

denote the characteristic function of H, we have $p_1^* = p_2 h$ and $p_2^* = p_2 h$. Writing γ_1 and γ_2 for the densities of P_1^* and P_2^*, respectively, we have

$$\begin{aligned} |\gamma_1(u,v)-\gamma_2(u,v)| \ &\leq\ (2\pi)^{-2}\left|\int_{\mathbb{R}^2} e^{-isu-itv}\left(p_1^*(s,t)-p_2^*(s,t)\right)\ d(s,t)\right| \\ &\leq\ (2\pi)^{-2}\int_{\mathbb{R}^2} |p_1(s,t)-p_2(s,t)|\,|h(s,t)|\ d(s,t) \\ &\leq\ (2\pi)^{-2}\int_{(s,t)\in[-T_1,T_1]\times[-T_2,T_2]} |p_1(s,t)-p_2(s,t)|\ d(s,t) \\ &\quad +(2\pi)^{-2}\ 2\int_{(s,t)\notin[-T_1,T_1]\times[-T_2,T_2]} |h(s,t)|\ d(s,t). \end{aligned}$$

Therefore

$$\begin{aligned} & |P_1^*([-x,x]\times[-y,y]) - P_2^*([-x,x]\times[-y,y])| \\ \leq\ & \int_{[-x,x]\times[-y,y]} |\gamma_1(u,v)-\gamma_2(u,v)|\ d(u,v) \\ \leq\ & 4xy(2\pi)^{-2}\int_{(s,t)\in[-T_1,T_1]\times[-T_2,T_2]} |p_1(s,t)-p_2(s,t)|\ d(s,t) \\ & +4xy(2\pi)^{-2}\ 2\int_{(s,t)\notin[-T_1,T_1]\times[-T_2,T_2]} |h(s,t)|\ d(s,t). \end{aligned}$$

Now

$$\int_{(s,t)\notin[-T_1,T_1]\times[-T_2,T_2]} |h(s,t)|\ d(s,t)$$

$$
\begin{aligned}
&= \int_{(s,t)\notin[-T_1,T_1]\times[-T_2,T_2]} e^{-\delta_1^2 s^2/2-\delta_2^2 t^2/2}\, d(s,t) \\
&\leq (2\pi)^{-1}\delta_1\delta_2 - \left((2\pi)^{-1/2}\delta_1 - 4\delta_1^{-1}\exp\left(-T_1^2\delta_1^2/2\right)\right)\times \\
&\qquad \times\left((2\pi)^{-1/2}\delta_2 - 4\delta_2^{-1}\exp\left(-T_2^2\delta_2^2/2\right)\right) \\
&\leq \delta_1^{-1}\delta_2\ \exp\left(-T_1^2\delta_1^2/2\right) + \delta_1\delta_2^{-1}\ \exp\left(-T_2^2\delta_2^2/2\right)
\end{aligned}
$$

proves the lemma. $\square$

The following lemma is a one-dimensional version of Lemma 2, which can be shown in the same way (a proof is contained in [5]).

Lemma 3 ([5, Lemma 4.2]) *Let P_1, P_2 be probability measures on $\mathbb{R}$, and write p_1, p_2 for the corresponding characteristic functions. Let*

$$P_1^* = P_1 \star H, \quad P_2^* = P_2 \star H,$$

where H is a normal distribution with mean zero and standard deviation δ. Then for all $T > 0$

$$|P_1^*([-x,x]) - P_2^*([-x,x])| \leq x\int_{s\in[-T,T]} |p_1(s)-p_2(s)|\ ds + 4x\delta^{-1}e^{-T^2\delta^2/2}.$$

Lemma 4 *Let P_1, P_2 be probability measures on $\mathbb{R}$, and write p_1, p_2 for the corresponding characteristic functions. Then for any $S, T, \delta > 0$*

$$
\begin{aligned}
&|P_1([-x-S,x+S]) - P_2([-x+S,x-S])| \\
\leq\ & x\int_{s\in[-T,T]} |p_1(s)-p_2(s)|\ ds + 4x\delta^{-1}e^{-T^2\delta^2/2} + 4e^{-S^2/(2\delta^2)}.
\end{aligned}
$$

Proof: Let S, T, δ be fixed. Let H be a normal distribution with mean zero and variance δ^2, and set $P_1^* = P_1 \star H$ and $P_2^* = P_2 \star H$. Then

$$
\begin{aligned}
&|P_1([-x-S,x+S]) - P_2([-x+S,x-S])| \\
\leq\ & |P_1^*([-x,x]) - P_2^*([-x,x])| + 2H\left(\mathbb{R}\backslash[-S,S]\right).
\end{aligned}
$$

Now

$$
\begin{aligned}
2H(\mathbb{R}\backslash[-S,S]) &\leq \frac{2}{\delta\sqrt{2\pi}}\int_{s\notin[-S,S]} e^{-s^2/(2\delta^2)}\ ds \\
&\leq 4e^{-S^2/(2\delta^2)},
\end{aligned}
$$

and thus by Lemma 3

$$
\begin{aligned}
&|P_1([-x-S,x+S]) - P_2([-x+S,x-S])| \\
\leq\ & x\int_{s\in[-T,T]} w(s)\ ds + 4e^{-S^2/(2\delta^2)} + 4x\delta^{-1}e^{-T^2\delta^2/2},
\end{aligned}
$$

which proves Lemma 4. $\square$

Let now $\theta > 1$, $\varepsilon > 0$ and $d \geq 1$ be arbitrary, but fixed. To simplify notations we assume in the sequel that f is even. We will also assume that $\|f\| > 0$, since otherwise Theorem 1 is trivial. We write

$$f(x) \sim \sum_{j=1}^{\infty} a_j \cos 2\pi jx = p(x) + r(x),$$

where

$$p(x) = \sum_{j=1}^{d} a_j \cos 2\pi jx, \quad r(x) = \sum_{j=d+1}^{\infty} a_j \cos 2\pi j n_k x.$$

Lemma 5 ([3, Lemma 3.1])

$$\limsup_{N\to\infty} \frac{\sum_{k=1}^{N} r(n_{\sigma(k)}x)}{\sqrt{2N \log\log N}} \leq Cd^{-1/4} \quad \text{a.e.}$$

for some number C which is independent of d and σ.

This lemma is valid for general lacunary sequences without any condition on the number of solutions of Diophantine equations. It is easy to see that the proof of [3, Lemma 3.1] is valid not only for strictly increasing lacunary sequences, but also for permutations of lacunary sequences, since it is based on a result of Philipp [19], which also has this property. We also recall that $\mathrm{Var}_{[0,1]} f \leq 1$ implies $|a_j| \leq j^{-1},\ j \geq 1$ (cf. Zygmund [25, p. 48]), and

$$\|r\|^2 \leq \sum_{j=d+1}^{\infty} |a_j|^2 \leq \sum_{j=d+1}^{\infty} j^{-2} \leq d^{-1}.$$

Lemma 6

$$\limsup_{N\to\infty} \frac{\sum_{k=1}^{N} p(n_{\sigma(k)}x)}{\sqrt{2N \log\log N}} = \|p\| \quad \text{a.e.}$$

In case $\|p\| = 0$ Lemma 6 is trivial. Therefore, to simplify formulas, we will assume in the sequel, without loss of generality, that $\|p\| = 1$, and prove

$$\limsup_{N\to\infty} \frac{\sum_{k=1}^{N} p(n_{\sigma(k)}x)}{\sqrt{2N \log\log N}} = 1 \quad \text{a.e.}$$

Since a finite number of elements of $(n_k)_{k\geq 1}$ does not influence the asymptotic behavior of $(n_{\sigma(k)}x)_{k\geq 1}$ we can also assume that

$$an_k + bn_l = 0 \tag{11}$$

does not have any solution $(k,l),\ k \neq l$ for $1 \leq |a| \leq d, 1 \leq |b| \leq d$.

Since d can be chosen arbitrarily large, Lemma 5 and and Lemma 6 together imply Theorem 1. Therefore it remains to prove Lemma 6. The proof of this lemma is crucial, and will be given in two parts below. The main ingredient is Lemma 7, which is formulated and proven below. For the proof we use ideas of Révész [21].

We define

$$\begin{aligned} \mu_k &= n_{\sigma(k)}, \qquad k \geq 1 \\ \Delta_m^* &= \left\{k \geq 1 : \ \theta^m \leq k < \theta^{m+1}\right\}, \qquad m \geq 1. \end{aligned}$$

We rearrange the sequence $(\mu_k)_{k\geq 1}$ in such a way that it is increasing within the blocks Δ_m^* and call this new sequence $(\nu_k)_{k\geq 1}$. In other words, $(\nu_k)_{k\geq 1}$ consists of the same elements as $(\mu_k)_{k\geq 1}$ (and $(n_{\sigma(k)})_{k\geq 1}$), but satisfies

$$\nu_k < \nu_l \quad \text{if} \quad k < l \quad \text{and} \quad k,l \in \Delta_m^* \quad \text{for some} \quad m \geq 1.$$

Moreover, we define

$$\overline{\Delta}_m = \left\{ k \in \Delta_m^* : \ \nexists l \in \bigcup_{n=1}^{m-\log_\theta m} \Delta_n^* : \ \frac{\nu_k}{\nu_l} \in \left[\frac{1}{2d}, 2d\right] \right\}, \quad m \geq 1, \tag{12}$$

$$\begin{aligned} \Delta_m &= \Big\{ k \in \overline{\Delta}_m : \ \big(k \mod \big(\lceil \sqrt{m} + \log_q(2d) \rceil\big)\big) \\ &\qquad \notin \left\{0, \dots, \lceil \log_q(2d) \rceil\right\} \Big\}, \quad m \geq 1, \\ \Delta_m^{(h)} &= \left\{ k \in \Delta_m : \ \frac{k}{\lceil \sqrt{m} + \log_q(2d) \rceil} \in [h, h+1) \right\}, \quad h \geq 0, m \geq 1, \\ \eta_m &= \frac{\sum_{k \in \Delta_m} p(\nu_k x)}{\sqrt{|\Delta_m|}}, \quad m \geq 1, \\ \alpha_m(s) &= \prod_{h \geq 0} \left(1 + \frac{is \sum_{k \in \Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j \nu_k x)}{\sqrt{|\Delta_m|}} \right), \quad m \geq 1, \\ \beta_m &= \sum_{h \geq 0} \sum_{k_1, k_2 \in \Delta_m^{(h)}} \sum_{j_1, j_2 = 1}^d \frac{a_{j_1} a_{j_2}}{2} \cos(2\pi (j_1 \nu_{k_1} + j_2 \nu_{k_2}) x) \\ &\quad + \sum_{h \geq 0} \underbrace{\sum_{k_1, k_2 \in \Delta_m^{(h)}} \sum_{j_1, j_2 = 1}^d}_{(k_1, j_1) \neq (k_2, j_2)} \frac{a_{j_1} a_{j_2}}{2} \cos(2\pi (j_1 \nu_{k_1} - j_2 \nu_{k_2}) x), \quad m \geq 1, \\ \varphi_{m,n}(s,t) &= \mathbb{E}\left(e^{is\eta_m + it\eta_n}\right), \quad m, n \geq 1, \quad s, t \in \mathbb{R}. \end{aligned}$$

Here $|\Delta_m|$ denotes the number of elements of the set Δ_m. Throughout the paper $\log x$ will be understood as $\max\{1, \log x\}$. $\sum_{k=m}^n$ means $\sum_{m \leq k \leq n}$, if m, n are not integers.

We observe that for $m \geq 1$

$$
\begin{aligned}
|\overline{\Delta}_m| &\geq |\Delta_m^*| - (2\log_q 2d) \sum_{n=1}^{m-\log_\theta m} |\Delta_n^*| \\
&\geq |\Delta_m^*| \left(1 - \frac{(2\log_q 2d)\theta^{m+1-\log_\theta m}}{|\Delta_m^*|}\right) \\
&\geq |\Delta_m^*| \left(1 - \frac{(2\log_q 2d)\theta}{(\theta-1)m}\right)
\end{aligned}
$$

and therefore

$$
\begin{aligned}
|\Delta_m| &\geq |\overline{\Delta}_m| - \left(\left(1 + \lceil \log_q(2d) \rceil\right)\left(\frac{|\overline{\Delta}_m|}{\lceil \sqrt{m}\ \rceil} + 1\right)\right) \\
&\geq |\Delta_m^*| \left(1 - \frac{(2\log_q 2d)\theta}{(\theta-1)m}\right)\left(1 - \frac{2\left(1 + \lceil \log_q(2d) \rceil\right)}{\lceil \sqrt{m}\ \rceil}\right). \qquad (13)
\end{aligned}
$$

Also, it is clear that

$$
\theta^m(\theta-1) \leq |\Delta_m^*| \leq \theta^m(\theta-1) + 1, \qquad (14)
$$

and

$$
\sum_{h\geq 0} \left|\Delta_m^{(h)}\right|^3 \leq \left(\max_{h\geq 0}\left|\Delta_m^{(h)}\right|\right)^2 \sum_{h\geq 0}\left|\Delta_m^{(h)}\right| \leq m|\Delta_m|. \qquad (15)
$$

By construction

$$
\sum_{h\geq 0} \sum_{k\in\Delta_m^{(h)}} p(\nu_k x) = \sum_{k\in\Delta_m} p(\nu_k x), \qquad m \geq 1,
$$

and since we have assumed $\|p\| = 1$, we also have

$$
|a_j| \leq \sqrt{2}, \qquad 1 \leq j \leq d, \qquad (16)
$$

and

$$
\sum_{k\in\Delta_m} \sum_{j=1}^{d} \frac{a_j^2}{2} = |\Delta_m|, \qquad m \geq 1. \qquad (17)
$$

Finally, we have for $m \geq 1$

$$
|\beta_m| \leq \sum_{h\geq 0} \sum_{k_1,k_2\in\Delta_m^{(h)}} \sum_{j_1,j_2=1}^{d} 1 \leq d^2 \left(\max_{h\geq 0}|\Delta_m^{(h)}|\right) \sum_{h\geq 0}\left|\Delta_m^{(h)}\right| \leq d^2\sqrt{m}|\Delta_m|. \qquad (18)
$$

Lemma 7 *Let $m, n \geq 1$, and assume that*

$$
m \leq n - \lceil \log_\theta n \rceil. \qquad (19)
$$

Then for sufficiently large m, n we have

$$\left\| \varphi_{m,n}(s,t) - e^{-(s^2+t^2)/2} \right\| \le \frac{1}{m^4+n^4}$$

provided

$$|s| \le m^{1/8} \quad |t| \le n^{1/8}.$$

Proof: Using

$$e^{ix} = (1+ix)e^{-x^2/2+w(x)}, \qquad |w(x)| \le |x|^3, \tag{20}$$

we have

$$\begin{aligned}
e^{is\eta_m} &= \prod_{h\ge 0} \exp\left(\frac{is\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j\nu_k x)}{\sqrt{|\Delta_m|}} \right) \\
&= \alpha_m(s)\ \exp\left(\sum_{h\ge 0} \frac{-s^2\left(\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j\nu_k x)\right)^2}{2|\Delta_m|} \right) \times \\
&\quad \times \exp\left(\sum_{h\ge 0} w\left(\frac{s\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j\nu_k x)}{\sqrt{|\Delta_m|}} \right) \right).
\end{aligned}$$

Since

$$\begin{aligned}
& \sum_{h\ge 0} \left(\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j\nu_k x) \right)^2 \\
= \;& \sum_{h\ge 0} \sum_{k_1,k_2\in\Delta_m^{(h)}} \sum_{j_1,j_2=1}^d \frac{a_{j_1}a_{j_2}}{2} \left(\cos(2\pi(j_1\nu_{k_1}+j_2\nu_{k_2})x) + \cos(2\pi(j_1\nu_{k_1}-j_2\nu_{k_2})x)\right) \\
= \;& \sum_{h\ge 0} \sum_{k_1,k_2\in\Delta_m} \sum_{j_1,j_2=1}^d \frac{a_{j_1}a_{j_2}}{2} \cos(2\pi(j_1\nu_{k_1}+j_2\nu_{k_2})x) \\
& |\Delta_m| + \sum_{h\ge 0} \underbrace{\sum_{k_1,k_2\in\Delta_m} \sum_{j_1,j_2=1}^d}_{(k_1,j_1)\ne(k_2,j_2)} \frac{a_{j_1}a_{j_2}}{2} \cos(2\pi(j_1\nu_{k_1}-j_2\nu_{k_2})x) \\
= \;& |\Delta_m| + \beta_m,
\end{aligned}$$

where we used (17), we can write

$$e^{is\eta_m} = \alpha_m(s)\ \exp\left(-\frac{s^2}{2}\left(1+\frac{\beta_m}{|\Delta_m|}\right) + w_m(s) \right), \tag{21}$$

where

$$w_m(s) = \sum_{h\ge 0} w\left(\frac{s\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j\nu_k x)}{\sqrt{|\Delta_m|}} \right),$$

and by (15), (16), (20),

$$
\begin{aligned}
|w_m(s)| &\leq \left| \sum_{h\geq 0} \frac{s^3 d^3 2^{3/2} \left|\Delta_m^{(h)}\right|^3}{|\Delta_m|^{3/2}} \right| \\
&\leq 3\,|s|^3\, d^3 m |\Delta_m|^{-1/2}.
\end{aligned} \tag{22}
$$

Note further that

$$
\begin{aligned}
|\alpha_m(s)| &\leq \prod_{h\geq 0} \left(1 + \frac{s^2 \left(\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j \nu_k x) \right)^2}{|\Delta_m|} \right)^{1/2} \\
&\leq \exp\left(\sum_{h\geq 0} \frac{s^2 \left(\sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j \nu_k x) \right)^2}{2|\Delta_m|} \right) \\
&= \exp\left(\frac{s^2}{2} \left(1 + \frac{\beta_m}{|\Delta_m|} \right) \right).
\end{aligned} \tag{23}
$$

Finally we observe, that for $m \leq n - \lceil \log_\theta n \rceil$

$$
\begin{aligned}
& \mathbb{E}\left(\alpha_m(s)\alpha_n(t)\right) \\
= & \int_0^1 \prod_{h\geq 0} \left(1 + \frac{is \sum_{k\in\Delta_m^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j \nu_k x)}{\sqrt{|\Delta_m|}} \right) \times \\
& \times \prod_{h\geq 0} \left(1 + \frac{it \sum_{k\in\Delta_n^{(h)}} \sum_{j=1}^d a_j \cos(2\pi j \nu_k x)}{\sqrt{|\Delta_n|}} \right) dx \\
= & \; 0,
\end{aligned} \tag{24}
$$

since by the construction of the sets Δ_m, Δ_n and $\Delta_m^{(h)}$, $\Delta_n^{(h)}$ we have for any $1 \leq j_1, j_2 \leq d$

$$
\begin{aligned}
\frac{j_1 \nu_{k_1}}{j_2 \nu_{k_2}} \not\in \left[\frac{1}{2}, 2\right] \qquad & \text{if} \quad k_1 \in \Delta_m^{(h_1)},\ k_2 \in \Delta_m^{(h_2)} \quad \text{for some} \quad h_1 \neq h_2 \\
& \text{or} \quad k_1 \in \Delta_n^{(h_1)},\ k_2 \in \Delta_n^{(h_2)} \quad \text{for some} \quad h_1 \neq h_2,
\end{aligned}
$$

and also

$$
\frac{j_1 \nu_{k_1}}{j_2 \nu_{k_2}} \not\in \left[\frac{1}{2}, 2\right] \quad \text{if} \quad k_1 \in \Delta_m^{(h_1)},\ k_2 \in \Delta_n^{(h_2)} \quad \text{for some} \quad h_1, h_2 \geq 0.
$$

In fact, if e.g. $k_1 \in \Delta_m^{(h_1)}$, $k_2 \in \Delta_m^{(h_2)}$ for $h_1 < h_2$, then necessarily

$$
\frac{j_1 n_{k_1}}{j_2 n_{k_2}} \leq \frac{d n_{k_1}}{n_{k_1 + \lceil \log_q(2d) \rceil}} < d q^{\log_q(2d)} \leq 1/2,
$$

and for $k_1 \in \Delta_m$, $k_2 \in \Delta_n$ we have

$$\frac{j_1 n_{k_1}}{j_2 n_{k_2}} \notin \left[\frac{1}{2}, 2\right] \qquad \text{since} \qquad \frac{n_{k_1}}{n_{k_2}} \notin \left[\frac{1}{2d}, 2d\right]$$

by (12) and (19). Thus by (21), (23), (24)

$$\begin{aligned}
& \left|\varphi_{m,n}(s,t) - e^{-s^2/2-t^2/2}\right| \\
= & \left|\mathbb{E}\left(\alpha_m(s)\alpha_n(t)\exp\left(-\frac{s^2}{2}\left(1+\frac{\beta_m}{|\Delta_m|}\right)+w_m(s)\right)\times\right.\right. \\
& \left.\times\exp\left(-\frac{t^2}{2}\left(1+\frac{\beta_n}{|\Delta_n|}\right)+w_n(t)\right)\right) - e^{-s^2/2-t^2/2}\Bigg| \\
= & \left|\mathbb{E}\left(\alpha_m(s)\alpha_n(t)\left(\exp\left(-\frac{s^2}{2}\left(1+\frac{\beta_m}{|\Delta_m|}\right)+w_m(s)\right)\times\right.\right.\right. \\
& \left.\left.\left.\times\exp\left(-\frac{t^2}{2}\left(1+\frac{\beta_n}{|\Delta_n|}\right)+w_n(t)\right) - e^{-s^2/2-t^2/2}\right)\right)\right| \\
\leq & \ \mathbb{E}\left(|\alpha_m(s)\alpha_n(t)|\left|\exp\left(-\frac{s^2}{2}\left(1+\frac{\beta_m}{|\Delta_m|}\right)+w_m(s)\right)\times\right.\right. \\
& \left.\left.\times\exp\left(-\frac{t^2}{2}\left(1+\frac{\beta_n}{|\Delta_n|}\right)+w_n(t)\right) - e^{-s^2/2-t^2/2}\right|\right) \\
\leq & \ \mathbb{E}\left|e^{w_m(s)+w_n(t)} - \exp\left(\frac{s^2\beta_m}{2|\Delta_m|}+\frac{t^2\beta_n}{2|\Delta_n|}\right)\right| \\
\leq & \ \mathbb{E}\left|e^{w_m(s)+w_n(t)}-1\right| + \mathbb{E}\left|e\left(\frac{s^2\beta_m}{2|\Delta_m|}+\frac{t^2\beta_n}{2|\Delta_n|}\right)-1\right|.
\end{aligned} \tag{25}$$

By (22) we have

$$\mathbb{E}\left|e^{w_m(s)+w_n(t)}-1\right| \leq e\left(3|s|^3 d^3 m|\Delta_m|^{-1/2} + 3|t|^3 d^3 n|\Delta_n|^{-1/2}\right) - 1. \tag{26}$$

The function β_m is a sum of at most $2\sqrt{m}|\Delta_m|$ trigonometric functions. The coefficients of these functions are bounded by some constant C^* by the Diophantine condition in Theorem 1. Using (18), this implies

$$\begin{aligned}
\|\beta_m\|^2 & \leq 2C^*\sqrt{m}|\Delta_m|, \\
\mathbb{P}\left(|\beta_m| > |\Delta_m|^{2/3}\right) & \leq \frac{2C^*\sqrt{m}}{|\Delta_m|^{-1/3}}, \\
\mathbb{E}\ \exp\left(\frac{s^2\beta_m}{|\Delta_m|}\right) & \leq \exp\left(s^2|\Delta_m|^{-1/3}\right) + \exp\left(s^2 d^2\sqrt{m}\right)\frac{2C^* m}{|\Delta_m|^{-1/3}}
\end{aligned} \tag{27}$$

and therefore

$$\begin{aligned}
& \mathbb{E}\left|\exp\left(\frac{s^2\beta_m}{2|\Delta_m|}+\frac{t^2\beta_n}{2|\Delta_n|}\right)-1\right| \\
\leq & \left(\mathbb{E}\exp\left(\frac{s^2\beta_m}{|\Delta_m|}\right)\right)^{1/2}\left(\mathbb{E}\exp\left(\frac{t^2\beta_n}{|\Delta_n|}\right)\right)^{1/2} - 1
\end{aligned} \tag{28}$$

$$\leq \quad \left(\left(\exp\left(s^2 |\Delta_m|^{-1/3}\right) + \exp\left(s^2 d^2 \sqrt{m}\right) \frac{2C^* m}{|\Delta_m|^{-1/3}} \right) \times \right.$$
$$\left. \times \left(\exp\left(t^2 |\Delta_n|^{-1/3}\right) + \exp\left(t^2 d^2 \sqrt{n}\right) \frac{2C^* n}{|\Delta_n|^{-1/3}} \right) \right)^{1/2} - 1 \qquad (29)$$

Now (25), (26), (29) and some elementary calculations show that for sufficiently large m, n, under the additional condition

$$|s| \leq m^{1/8}, \quad |t| \leq n^{1/8}$$

we have

$$\left| \varphi_{m,n}(s,t) - e^{-s^2/2 - t^2/2} \right| \leq \frac{1}{m^4 + n^4},$$

which proves the lemma. $\Box$

Lemma 8 *For sufficiently large m we have*

$$\left| \mathbb{E} e^{is\eta_m} - e^{-s^2/2} \right| \leq m^{-4},$$

for all $s \in [-m^{1/8}, m^{1/8}]$.

Proof: This lemma is an one-dimensional version of Lemma 7 and can be shown in exactly the same way.

Lemma 9 *Let B be a finite set of positive integers. Then if $|B|$ is sufficiently large, we can divide B into two disjoint sets B_1, B_2, such that*

$$|B_2| \leq C_1 |B| / \sqrt{\log |B|}$$

for some constant C_1, and

$$\left| \mathbb{E} \exp\left(is |B_1|^{-1/2} \sum_{k \in B_1} p(n_k x) \right) - e^{-s^2/2} \right| \leq (\log |B_1|)^{-4},$$

for $|s| \leq (\log |B|)^{1/8}$.

Proof: This lemma can be shown in the same way as the previous two lemmas (or exactly in the same way as [5, Lemma 4.3]).

Lemma 10 ([23, Lemma 2]) *Let B be a finite set of positive integers. Then for any $\lambda > 0$ satisfying*

$$4\lambda |B|^{1/3} < 1$$

we have

$$\int_0^1 \exp\left(\sum_{k \in B} p(n_{\sigma(k)} x) \right) dx \leq C_2 e^{C_3 \lambda^2 |B|},$$

where C_2, C_3 are positive constants.

Next we prove some Berry-Esseen type lemmas needed for our proof. We redefine the random variables $\eta_1, \eta_2, \dots$ on a larger probability space $(\Omega, \mathcal{A}, \hat{P})$ (we write $\hat{\eta}_1, \hat{\eta}_2, \dots$ for the redefined r.v.'s), such that their finite dimensional distributions remain unchanged, and such that on the new probability space there exists a sequence $\hat{h}_1, \hat{h}_2, \dots$ of i.i.d. random variables satisfying

- $\hat{h}_m \sim \mathcal{N}(0, \tau_m), \quad \text{where} \quad \tau_m = \dfrac{1}{\sqrt{8} \log\log\log \theta^m}, \quad m \geq 1$
- $\hat{h}_m \quad \text{and} \quad \eta_m \quad \text{are independent}, \quad m \geq 1,$
- the two-dimensional random variables $(\hat{h}_m, \hat{h}_n)$ and $(\hat{\eta}_m, \hat{\eta}_n)$ are independent, $m \neq n, \ m, n \geq 1$

Lemma 11 *Define*

$$z_m = \sqrt{(2-\varepsilon)(\theta/(\theta-1)) \log\log \theta^m}$$

and

$$A_m = \left\{ \omega \in \Omega : \ \hat{\eta}_m(\omega) + \hat{h}_m(\omega) > z_m \right\}, \quad m \geq 1.$$

Then

$$\begin{aligned} & \left| \hat{P}(A_m A_n) - R\left((1+\tau_m)^{-1} z_m\right) R\left((1+\tau_n)^{-1} z_n\right) \right| \\ \leq \ & (\log m)^2 (\log n)^2 \left(m^{-4} + n^{-4}\right) \end{aligned}$$

for sufficiently large m, n, provided $m \leq n - \lceil \log n \rceil$. Here

$$R(u) = 1 - (2\pi)^{-1/2} \int_{-u}^{u} e^{-s^2/2} ds, \quad u \geq 0.$$

Proof: We define two measures P_1, P_2 on $\mathbb{R}^2$: P_1 is the measure induced by $(\hat{\eta}_m, \hat{\eta}_n)$, and P_2 is a two-dimensional standard normal distribution. We apply Lemma 2 with $x = z_1, y = z_2, \sigma_1 = \tau_m, \sigma_2 = \tau_n$ and

$$T_1 = 8\sqrt{\log\log \theta^m} \log\log\log \theta^m \qquad T_2 = 8\sqrt{\log\log \theta^n} \log\log\log \theta^n.$$

Then we get, using the notations from Lemma 2,

$$\begin{aligned} & \left| P_1^*([-x,x] \times [-y,y]) - P_2^*([-x,x] \times [-y,y]) \right| \\ \leq \ & +xy \ 4 T_1 T_2 \frac{1}{m^4 + n^4} \\ & +xy \left(\tau_m^{-1} \tau_n \ e\left(-T_1^2 \tau_1^2/2\right) + \tau_m \tau_n^{-1} \ e\left(-T_2^2 \tau_2^2/2\right) \right) \\ \leq \ & (\log m)^2 (\log n)^2 \left(\frac{1}{m^4} + \frac{1}{n^4} \right) \end{aligned}$$

for sufficiently large m, n (we emphasize that $T_1 \leq m^{1/8}, \ T_2 \leq n^{1/8}$ for sufficiently large m, n, and therefore we can use Lemma 7). Since by construction $(\hat{\eta}_m, \hat{\eta}_n)$ and $(\hat{h}_m, \hat{h}_n)$ are independent,

$$\hat{P}(A_m A_n) = 1 - (P_1 \star H)\left([-z_m, z_m] \times [-z_n, z_n]\right) = 1 - P_1^*([-z_m, z_m] \times [-z_n, z_n]),$$

and since the random variables $\hat{h}_m$ have distribution $\mathcal{N}(0, \tau_m)$, we have

$$1 - P_2^* \left([-z_m, z_m] \times [-z_n, z_n]\right) = R\left((1+\tau_m)^{-1} z_m)\right) R\left((1+\tau_n)^{-1} z_m\right).$$

Summarizing our estimates, we have

$$\left|\hat{P}(A_m A_n) - R\left((1+\tau_m)^{-1} z_m\right) R\left((1+\tau_n)^{-1} z_n\right)\right| \quad \leq \quad (\log m)^2 (\log n)^2 \left(\frac{1}{m^4} + \frac{1}{n^4}\right)$$

for sufficiently large m, n. $\square$

Lemma 12 *For sufficiently large m*

$$\left|\hat{P}(A_m) - R\left((1+\tau_m)^{-1} z_m\right)\right| \leq \frac{(\log m)^2}{m^4}.$$

This can be shown like Lemma 11, using Lemma 3 instead of Lemma 2.

Lemma 13 *Let*

$$\overline{A}_m = \left\{ x \in (0,1) : \sum_{k=1}^{\theta^m} p(\nu_k x) > \sqrt{(2+\varepsilon) \log\log \theta^n} + 3 \frac{\sqrt{\log\log \theta^m}}{\log\log\log \theta^m} \right\}, \quad m \geq 1.$$

Then for sufficiently large m

$$\mathbb{P}(\overline{A}_m) \leq R\left(\sqrt{(2+\varepsilon)\log\log\theta^m}\right) + 2\frac{(\log m)^2}{m^4}.$$

Proof: This is a consequence of Lemma 4 and Lemma 9. In fact, let

$$B = \{1 \leq k \leq \theta^m\}.$$

Then by Lemma 9 there exist sets B_1, B_2 such that such that

$$|B_2| \leq C_1 |B| / \sqrt{\log |B|}$$

and

$$\left|\mathbb{E} \exp\left(\frac{is \sum_{k \in B_1} p(n_k x)}{|B_1|^{1/2}}\right) - e^{-s^2/2}\right| \leq \frac{1}{(\log |B_1|)^4},$$

for $|s| \leq (\log |B|)^{1/8}$. We apply Lemma 4 with

$$\begin{aligned}
T &= 8\sqrt{\log\log\theta^m}\,\log\log\log\theta^m \\
S &= \sqrt{\log\log\theta^m}(\log\log\log\theta^m)^{-1} \\
\sigma &= \tau_m \\
x &= \sqrt{(2+\varepsilon)\log\log\theta^m} + \sqrt{\log\log\theta^m}\,(\log\log\log\theta^m)^{-1}
\end{aligned}$$

and get

$$\mathbb{P}\left\{ x \in (0,1) : \sum_{k=1}^{\theta^m} p(\nu_k x) > \sqrt{(2+\varepsilon)\log\log\theta^m} + 2\frac{\sqrt{\log\log\theta^m}}{\log\log\log\theta^m} \right\}$$

$$\begin{aligned}
&\leq 1 - \left(\frac{1}{\sqrt{2\pi}} \int_{-\sqrt{(2+\varepsilon)\log\log\theta^m}}^{\sqrt{(2+\varepsilon)\log\log\theta^m}} e^{s^2/2}\ ds - 2xT \frac{1}{(\log|B_1|)^4} \right. \\
&\qquad \left. -4x\tau_m^{-1} \exp\left(-T^2\tau_m^2/2\right) - 2\exp\left(-S^2/(2\tau_m^2)\right) \right) \\
&\leq R\left(\sqrt{(2+\varepsilon)\log\log\theta^m}\right) + (\log m)^2 m^{-4}
\end{aligned}$$

for sufficiently large m. By Lemma 10

$$\mathbb{P}\left(\left| \sum_{k\in B_2} p(n_k x) \right| > S \right) \leq m^{-4}$$

for sufficiently large m, and the proof of the lemma is complete. $\square$

We are ready now to prove the upper bound in the LIL. We show

Lemma 14

$$\limsup_{N\to\infty} \frac{\left|\sum_{k=1}^N p(n_{\sigma(k)}x)\right|}{\sqrt{2N\log\log N}} \leq 1 \quad \text{a.e.}$$

Proof: By Lemma 13 we have

$$\sum_{m\geq 1} \mathbb{P}\left(\overline{A}_m\right) < +\infty,$$

and therefore the Borel-Cantelli lemma implies

$$\liminf_{m\to\infty} \frac{\left|\sum_{k=1}^{\theta^m} p(n_{\sigma(k)}x)\right|}{\sqrt{(2+\varepsilon)\theta^m \log\log\theta^m}} \leq 1 \quad \text{a.e.} \tag{30}$$

It remains to fill the gaps between θ^m and θ^{m+1}, $m \geq 1$. Using Lemma 10 we can show, e.g. by using the method from [7, Section 4], that

$$\limsup_{m\to\infty} \max_{\theta^m \leq M \leq \theta^{m+1}} \frac{\left|\sum_{k=\theta^m}^M p(n_{\sigma(k)}x)\right|}{\sqrt{2(\theta^{m+1}-\theta^m)\log\log(\theta^{m+1}-\theta^m)}} \leq C_4 \quad \text{a.e.,}$$

where C_4 may only depend on p and the growth factor q. Combining this with (30) we have

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{\left|\sum_{k=1}^N p(n_{\sigma(k)}x)\right|}{\sqrt{2N\log\log N}} \\
\leq\ &\limsup_{m\to\infty} \max_{\theta^m\leq M\leq\theta^{m+1}} \frac{\left|\sum_{k=1}^M p(n_{\sigma(k)}x)\right|}{\sqrt{2\theta^m\log\log\theta^m}} \\
\leq\ &\limsup_{m\to\infty} \frac{\left|\sum_{k=1}^{\theta^m} p(n_{\sigma(k)}x)\right|}{\sqrt{2\theta^m\log\log\theta^m}} + \limsup_{m\to\infty} \max_{\theta^m\leq M\leq\theta^{m+1}} \frac{\left|\sum_{k=\theta^m}^M p(n_{\sigma(k)}x)\right|}{\sqrt{2\theta^m\log\log\theta^m}}
\end{aligned}$$

$$\leq \quad (2+\varepsilon) + C_4(\theta - 1) \quad \text{a.e.}$$

Since $\varepsilon > 0$ and $\theta > 1$ can be chosen arbitrarily, this concludes the proof of Lemma 14. $\square$

Next we prove the lower bound in the LIL.

Lemma 15

$$\limsup_{N\to\infty} \frac{\left|\sum_{k=1}^{N} p(n_{\sigma(k)}x)\right|}{\sqrt{2N\log\log N}} \geq 1 \quad \text{a.e.}$$

Proof: By Lemma 11 we have

$$\begin{aligned}
& \sum_{n=1}^{N}\sum_{m=1}^{N} \hat{P}(A_m A_n) \\
\geq \quad & -C_5 + 2\sum_{n=1}^{N} \sum_{m=n^{2/3}}^{n-\log_\theta n} \Big(R\left((1+\tau_m)^{-1} z_m\right)\Big) R\left((1+\tau_n)^{-1} z_n\right) \\
& + (\log m)^2 (\log n)^2 \frac{1}{m^4 n^4}\Big) \\
\geq \quad & -C_6 + 2\sum_{n=1}^{N} \sum_{m=1}^{n} R\left((1+\tau_m)^{-1} z_m\right)\Big) R\left((1+\tau_n)^{-1} z_n\right) \\
& -2\sum_{n=1}^{N} \sum_{m=1}^{n^{2/3}} R\left((1+\tau_m)^{-1} z_m\right)\Big) R\left((1+\tau_n)^{-1} z_n\right) \\
& -2\sum_{n=1}^{N} \sum_{m=n-\log_\theta n}^{n} R\left((1+\tau_m)^{-1} z_m\right)\Big) R\left((1+\tau_n)^{-1} z_n\right)
\end{aligned}$$

for some positive constants C_5 and C_6.

In the sequel we will assume that ε and θ are chosen in such a way that there exists some $\rho > 0$ such that

$$(2-\varepsilon)(\theta/(\theta-1))/2 < 1 - \rho.$$

For given ε this is possible by choosing θ large. Some calculations show that

$$\begin{aligned}
\exp\left(-\left((1+\tau_m)^{-1} z_m + 1\right)^2/2\right) \quad & \leq \quad \sqrt{2\pi}\, R\left((1+\tau_m)^{-1} z_m\right) \\
& \leq \quad \exp\left(-\left((1+\tau_m)^{-1} z_m\right)^2/2\right),
\end{aligned}$$

and therefore

$$\frac{(m\log\theta)^{-(1+\tau_m)^{-1}(2-\varepsilon)(\theta/(\theta-1))/2}}{e^{-(1+\tau_m)^{-1}\sqrt{(2-\varepsilon)(\theta/(\theta-1))\log\log\theta^m}/2-1/2}} \tag{31}$$

$$
\begin{aligned}
&\le \sqrt{2\pi}\; R\left((1+\tau_m)^{-1}z_m\right) \\
&\le (m\log\theta)^{-(1+\tau_m)^{-1}(2-\varepsilon)(\theta/(\theta-1))/2},
\end{aligned}
$$

which implies

$$
\begin{aligned}
& \sum_{n=1}^{N}\sum_{m=1}^{n^{2/3}} R\left((1+\tau_m)^{-1}z_m\right) R\left((1+\tau_n)^{-1}z_n\right) \\
& + \sum_{n=1}^{N}\sum_{m=n-\log_\theta n}^{n} R\left((1+\tau_m)^{-1}z_m\right) R\left((1+\tau_n)^{-1}z_n\right) \\
= {} & o\left(\sum_{n=1}^{N}\sum_{m=1}^{n} R\left((1+\tau_m)^{-1}z_m\right) R\left((1+\tau_n)^{-1}z_n\right)\right) \quad \text{as} \quad m\to\infty.
\end{aligned}
$$

Thus

$$
\liminf_{N\to\infty}\frac{\sum_{n=1}^{N}\sum_{m=1}^{N}\hat{P}(A_mA_n)}{\left(\sum_{m=1}^{N}\hat{P}(A_m)\right)^2}=1.
$$

Then, by (31), for sufficiently large m

$$
\hat{P}(A_m)\ge m^{-1+\rho/2}.
$$

Therefore

$$
\sum_{m=1}^{\infty}\hat{P}(A_m)=+\infty,
$$

and by Lemma 1 there occur infinitely events A_m with probability 1, which implies

$$
\limsup_{m\to\infty}\frac{\left|\hat{\eta}_m+\hat{h}_m\right|}{\sqrt{(2-\varepsilon)(\theta/(\theta-1))\log\log\theta^m}}\ge 1 \quad \text{a.s.}
$$

Using the classical LIL for i.i.d. random variables we easily get

$$
\limsup_{m\to\infty}\frac{\left|\hat{h}_m\right|}{\sqrt{(2-\varepsilon)(\theta/(\theta-1))\log\log\theta^m}}=0 \quad \text{a.s.},
$$

(recall that $\tau_m\to 0$) and therefore

$$
\limsup_{m\to\infty}\frac{|\hat{\eta}_m|}{\sqrt{(2-\varepsilon)(\theta/(\theta-1))\log\log\theta^m}}\ge 1 \quad \text{a.s.}
$$

This implies the similar result for the original random variables $\eta_1,\eta_2,\dots$, i.e.

$$
\limsup_{m\to\infty}\frac{|\eta_m|}{\sqrt{(2-\varepsilon)(\theta/(\theta-1))\log\log\theta^m}}\ge 1 \quad \text{a.e.}
$$

or

$$
\limsup_{m\to\infty}\frac{\left|\sum_{k\in\Delta_m}p(\nu_k x)\right|}{\sqrt{(2-\varepsilon)(\theta/(\theta-1))|\Delta_m|\log\log\theta^m}}\ge 1 \quad \text{a.e.}
$$

Using Lemma 10, it is not difficult to show

$$\limsup_{m\to\infty} \frac{\left|\sum_{k\in\overline{\overline{\Delta}}_m\setminus\Delta_m} p(\nu_k x)\right|}{\sqrt{(2-\varepsilon)(\theta/(\theta-1))|\Delta_m|\log\log\theta^m}} = 0 \quad \text{a.e.}$$

and since by (13) and (14)

$$\frac{|\Delta_m|}{\theta^m(\theta-1)} \to 1$$

this implies

$$\limsup_{m\to\infty} \frac{\sum_{k=\theta^m}^{\theta^{m+1}} p(\nu_k x)}{\sqrt{(2-\varepsilon)\theta^{m+1}\log\log\theta^{m+1}}} \geq 1 \quad \text{a.e.}$$

By the results from the previous section,

$$\limsup_{m\to\infty} \frac{\left|\sum_{k=1}^{\theta^m} p(\nu_k x)\right|}{\sqrt{2\theta^m\log\log\theta^m}} \leq 1 \quad \text{a.e.,}$$

and therefore

$$\limsup_{m\to\infty} \frac{\sum_{k=1}^{\theta^{m+1}} p(\nu_k x)}{\sqrt{2\theta^{m+1}\log\log\theta^{m+1}}} \geq \frac{\sqrt{2-\varepsilon}}{\sqrt{2}} - \frac{1}{\sqrt{\theta}} \quad \text{a.e.}$$

Choosing $\varepsilon > 0$ small and $\theta > 1$ large this proves Lemma 6, and therefore the proof of Theorem 1 is complete. $\square$

To conclude this section, we justify the remark made after the statement of Theorem 1. Assume there exist integers $a \neq 0, b \neq 0, c$, such that the Diophantine equation

$$an_k - bn_l = c \tag{32}$$

has infinitely many solutions $(k,l), k \neq l$ (by an easy observation we can assume $a > 0, b > 0$). We will construct a trigonometric polynomial $p(x)$ and a permutation $\sigma : \mathbb{Z}^+ \to \mathbb{Z}^+$ such that

$$\limsup_{N\to\infty} \frac{\sum_{k=1}^{N} p(n_{\sigma(k)}x)}{\sqrt{2N\log\log N}} \neq \|p\| \quad \text{a.e.} \tag{33}$$

We define

$$p(x) = \cos(2\pi a x) + \cos(2\pi b x).$$

Let

$$(k_1, l_1), (k_2, l_2), \ldots$$

denote a sequence of solutions of (32), chosen in such a way that

- $k_j > k_i, \quad l_j > l_i \qquad \text{for } j > i$
- $k_{i+1}/k_i > 2, \quad l_{i+1}/l_i > 2, \qquad i \geq 1$

- $k_{i+1}/k_i \to \infty, \quad l_{i+1}/l_i \to \infty.$

Clearly there exists a permutation $\sigma: \mathbb{Z}^+ \to \mathbb{Z}^+$ such that

$$\limsup_{N\to\infty} \frac{\left|\sum_{k=1}^N p(n_{\sigma(k)}x)\right|}{\sqrt{2N\log\log N}} = \limsup_{N\to\infty} \frac{\left|\sum_{i=1}^{N/2} p(n_{k_i}x) + p(n_{l_i}x)\right|}{\sqrt{2N\log\log N}} \tag{34}$$

For example, we can construct σ such that for every even N

$$\begin{aligned}\{\sigma(k),\ 1 \le k \le N\} &= \{k_i, 1 \le i \le N/2 - \lfloor \log_{10} N \rfloor\} \\ &\quad \cup \{l_i, 1 \le i \le N/2 - \lfloor \log_{10} N \rfloor\} \\ &\quad \cup \{1 \le k \le M\},\end{aligned}$$

where M is chosen such that the set on the right-hand side really consists of N elements. Since always $M \le 2\log N$, relation (34) will hold for σ. Thus it suffices to calculate

$$\limsup_{N\to\infty} \frac{\left|\sum_{i=1}^{N/2} p(n_{k_i}x) + p(n_{l_i}x)\right|}{\sqrt{2N\log\log N}}.$$

Using standard trigonometric identities we have

$$\begin{aligned}&p(n_{k_i}x) + p(n_{l_i}x)\\ =\ & \cos(2\pi a n_{k_i}x) + \cos(2\pi b n_{k_i}x) + \cos(2\pi a n_{l_i}x) + \cos(2\pi b n_{l_i}x)\\ =\ & 2\cos(\pi(a n_{k_i} + b n_{l_i})x)\cos(\pi(a n_{k_i} - b n_{l_i})x) + \cos(2\pi b n_{k_i}x) + \cos(2\pi a n_{l_i}x)\\ =\ & 2\cos(\pi c x)\cos\left(\pi(a n_{k_i} + b n_{l_i})x\right) + \cos(2\pi b n_{k_i}x) + \cos(2\pi a n_{l_i}x).\end{aligned}$$

Clearly, a sequence consisting of the elements

$$(a n_{k_i} + b n_{l_i})/2, \quad a n_{l_i}, \quad b n_{k_i}, \qquad i \ge 1,$$

arranged in increasing order, is a lacunary sequence for i sufficiently large. Using the methods of [1] we can show

$$\begin{aligned}&\limsup_{N\to\infty} \frac{\sum_{k=1}^N 2\cos(\pi c x)\cos(\pi(a n_{k_i} + b n_{l_i})x) + \cos(2\pi b n_{k_i}x) + \cos(2\pi a n_{l_i}x)}{\sqrt{2N\log\log N}}\\ =\ & \sqrt{2\cos^2(\pi c x) + 1} \quad \text{a.e.}\end{aligned}$$

and therefore

$$\begin{aligned}\limsup_{N\to\infty} \frac{\left|\sum_{i=1}^{N/2} p(n_{k_i}x) + p(n_{l_i}x)\right|}{\sqrt{2N\log\log N}} &= \sqrt{\cos^2(\pi c x) + 1/2} \quad \text{a.e.}\\ &= \sqrt{\frac{\cos(2\pi c x) + 2}{2}} \quad \text{a.e.},\end{aligned}$$

which verifies (33).

References

[1] C. Aistleitner. Irregular discrepancy behavior of lacunary series. *Monatsh. Math.*, 160(1):1–29, 2010.

[2] C. Aistleitner. Irregular discrepancy behavior of lacunary series II. *Monatsh. Math.*, 161(3):255–270, 2010.

[3] C. Aistleitner. On the law of the iterated logarithm for the discrepancy of Lacunary sequences. *Trans. Amer. Math. Soc.*, 362(11):5967–5982, 2010.

[4] C. Aistleitner and I. Berkes. On the central limit theorem for $f(n_k x)$. *Probab. Theory Related Fields*, 146(1-2):267–289. 2010.

[5] C. Aistleitner, I. Berkes and R. Tichy. On permutations of lacunary series. Submitted.

[6] J.-P. Conze and S. Le Borgne. Limit law for some modified ergodic sums. *Stoch. Dyn.* 11(1):107–133, 2011.

[7] P. Erdös and I. S. Gál. On the law of the iterated logarithm. I, II. *Nederl. Akad. Wetensch. Proc. Ser. A.* **58** = *Indag. Math.*, 17:65–76, 77–84, 1955.

[8] K. Fukuyama. The law of the iterated logarithm for discrepancies of $\{\theta^n x\}$. *Acta Math. Hungar.*, 118(1-2):155–170, 2008.

[9] K. Fukuyama. The law of the iterated logarithm for the discrepancies of a permutation of $\{n_k x\}$. *Acta Math. Hungar.*, 123(1-2):121–125, 2009.

[10] K. Fukuyama. A central limit theorem and a metric discrepancy result for sequences with bounded gaps. Berkes, István (ed.) et al., Dependence in probability, analysis and number theory. A volume in memory of Walter Philipp. Heber City, UT: Kendrick Press. 233-246, 2010.

[11] K. Fukuyama and S. Miyamoto. Metric discrepancy results for Erdős-Fortet sequence. *Studia Sci. Math. Hung.*, to appear.

[12] V.F. Gaposhkin. Lacunary series and independent functions. *Russian Math. Surveys* **21**, 3-82, 1966.

[13] V.F. Gaposhkin. The central limit theorem for some weakly dependent sequences. *Theory Probab. Appl.* **15**, 649-666, 1970.

[14] S.-i. Izumi. Notes on Fourier analysis. XLIV. On the law of the iterated logarithm of some sequences of functions. *J. Math. Tokyo*, 1:1–22, 1951.

[15] M. Kac. On the distribution of values of sums of the type $\sum f(2^k t)$. *Ann. of Math.* **47**, 33–49, 1946.

[16] M. Kac. Probability methods in some problems of analysis and number theory. Bull. Amer. Math. Soc. **55** (1949), 641–665.

[17] G. Maruyama. On an asymptotic property of a gap sequence. *Kōdai Math. Sem. Rep.*, 2:31–32, 1950.

[18] Y. Peres and W. Schlag. Two Erdős problems on lacunary sequences: chromatic number and Diophantine approximation. *Bull. Lond. Math. Soc.*, 42(2):295–300, 2010.

[19] W. Philipp. Limit theorems for lacunary series and uniform distribution mod 1. *Acta Arith.*, 26(3):241–251, 1974/75.

[20] A. Rényi. *Probability theory.* North-Holland Publishing Co., Amsterdam, 1970. Translated by László Vekerdi, North-Holland Series in Applied Mathematics and Mechanics, Vol. 10.

[21] P. Révész. The law of the iterated logarithm for multiplicative systems. *Indiana Univ. Math. J.*, 21:557–564, 1971/72.

[22] G. R. Shorack and J. A. Wellner. *Empirical processes with applications to statistics.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York, 1986.

[23] S. Takahashi. An asymptotic property of a gap sequence. *Proc. Japan Acad.*, 38:101–104, 1962.

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

[25] A. Zygmund. *Trigonometric series. Vol. I–II,* 3^{rd} *Edition.* Cambridge University Press, 2002.