On permutations of lacunary series

By

Christoph AISTLEITNER *, István BERKES**, and Robert TICHY***

Abstract

It is a well known fact that for periodic measurable f and rapidly increasing $(n_k)_{k \geq 1}$ the sequence $(f(n_k x))_{k \geq 1}$ behaves like a sequence of independent, identically distributed random variables. For example, if f is a periodic Lipschitz function, then $(f(2^k x))_{k \geq 1}$ satisfies the central limit theorem, the law of the iterated logarithm and several further limit theorems for i.i.d. random variables. Since an i.i.d. sequence remains i.i.d. after any permutation of its terms, it is natural to expect that the asymptotic properties of lacunary series are also permutation-invariant. Recently, however, Fukuyama (2009) showed that a rearrangement of the sequence $(f(2^k x))_{k \geq 1}$ can change substantially its asymptotic behavior, a very surprising result. The purpose of the present paper is to investigate this interesting phenomenon in detail and to give necessary and sufficient criteria for the permutation-invariance of the CLT and LIL for $f(n_k x)$.

*Graz University of Technology, Department for Analysis and Computational Number Theory, Steyrergasse 30, 8010 Graz, Austria. Research supported by FWF grant S9603-N23.
email: `aistleitner@math.tugraz.at`
**Graz University of Technology, Institute of Statistics, Münzgrabenstraße 11, 8010 Graz, Austria. Research supported by FWF grant S9603-N23 and OTKA grants K 67961 and K 81928.
e-mail: `berkes@tugraz.at`
***Graz University of Technology, Department for Analysis and Computational Number Theory, Steyrergasse 30, 8010 Graz, Austria. Research supported by FWF grant S9603-N23.
e-mail: `tichy@tugraz.at`

§1.   Introduction

By classical results of Salem and Zygmund [19] and Erdős and Gál [8], if $(n_k)_{k \geq 1}$ satisfies the Hadamard gap condition

$$(1.1) \qquad n_{k+1}/n_k \geq q > 1 \qquad (k = 1, 2, \ldots)$$

then $(\cos 2\pi n_k x)_{k \geq 1}$ obeys the central limit theorem and the law of the iterated logarithm, i.e.

$$(1.2) \qquad \lim_{N \to \infty} \lambda \left\{ x \in (0,1) : \sum_{k=1}^{N} \cos 2\pi n_k x \leq t\sqrt{N/2} \right\} = (2\pi)^{-1/2} \int_{-\infty}^{t} e^{-u^2/2} du,$$

and

$$(1.3) \qquad \limsup_{N \to \infty} \frac{\sum_{k=1}^{N} \cos 2\pi n_k x}{\sqrt{N \log \log N}} = 1 \qquad \text{a.e.}$$

where λ denotes the Lebesgue measure. Philipp and Stout [17] showed that under (1.1) on the probability space $([0,1], \mathcal{B}, \lambda)$ there exists a Brownian motion process $\{W(t), \ t \geq 0\}$ such that

$$(1.4) \qquad \sum_{k=1}^{N} \cos 2\pi n_k x = W(N/2) + O\left(N^{1/2-\rho}\right) \qquad \text{a.e.}$$

for some $\rho > 0$. This relation implies not only the CLT and LIL for $(\cos 2\pi n_k x)_{k \geq 1}$, but a whole class of further limit theorems for independent, identically distributed random variables; for examples and discussion we refer to [17]. Similar results hold for lacunary sequences $f(n_k x)$, where f is a measurable function on $\mathbb{R}$ satisfying the conditions

$$(1.5) \qquad f(x+1) = f(x), \qquad \int_0^1 f(x)\,dx = 0, \qquad \int_0^1 f^2(x)\,dx < \infty.$$

For example, Kac [14] proved that if (1.5) holds and f is either a Lipschitz function or is of bounded variation in $(0,1)$, then $f(2^k x)$ satisfies the central limit theorem, more precisely

$$\lim_{N \to \infty} \lambda \left\{ x \in (0,1) : \sum_{k=1}^{N} f(2^k x) \leq \sigma t \sqrt{N} \right\} = (2\pi)^{-1/2} \int_{-\infty}^{t} e^{-u^2/2} du,$$

where

$$\sigma^2 = \int_0^1 f^2(x)\,dx + 2\sum_{k=1}^{\infty} \int_0^1 f(x) f(2^k x)\,dx.$$

The corresponding LIL

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

was proved by Maruyama [16] and Izumi [13] and the analogue of (1.4) for $f(2^k x)$ was proved by Berkes and Philipp [6]. These results show that lacunary subsequences of the trigonometric system and of the system $(f(nx))_{n\geq 1}$ behave like sequences of i.i.d. random variables and since an i.i.d. sequence remains i.i.d. after any permutation of its terms, it is natural to expect that limit theorems for $f(n_k x)$ are also permutation-invariant. Recently, however, Fukuyama [10] showed the surprising fact that rearrangement of the sequence $(\{2^k x\})_{k\geq 1}$, where $\{\cdot\}$ denotes fractional part, can change the LIL behavior of its discrepancy. The purpose of this paper is to show that the permutation-invariance of $f(n_k x)$ is intimately connected with the number theoretic properties of $(n_k)_{k\geq 1}$ and we will give necessary and sufficient criteria for the rearrangement-invariant CLT and LIL for $f(n_k x)$ in terms of the number of solutions of the Diophantine equation

$$a_1 n_{k_1} + \ldots + a_p n_{k_p} = b, \qquad 1 \leq k_1,,\ldots,k_p \leq N.$$

In Section 2 we will formulate our results for the trigonometric system and in Section 3 for the system $f(nx)$. The proofs will be given in Section 4.

§2. The trigonometric system

In this section we deal with permutation-invariance of lacunary trigonometric series.

Theorem 2.1. *Let $(n_k)_{k\geq 1}$ be a sequence of positive integers satisfying the Hadamard gap condition (1.1) and let $\sigma : \mathbb{N} \to \mathbb{N}$ be a permutation of the positive integers. Then we have*

$$(2.1) \qquad N^{-1/2} \sum_{k=1}^{N} \cos 2\pi n_{\sigma(k)} x \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1/2)$$

and

$$(2.2) \qquad \limsup_{N\to\infty} \frac{\sum_{k=1}^{N} \cos 2\pi n_{\sigma(k)} x}{\sqrt{N \log\log N}} = 1 \qquad \text{a.e.}$$

Here we used probabilistic terminology and $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution in the probability space $([0,1], \mathcal{B}, \lambda)$.

Theorem 2.1 shows that for Hadamard lacunary trigonometric series the CLT and LIL are permutation-invariant without any additional number theoretic condition on n_k.

We note also that the unpermuted CLT (1.2) and LIL (1.3) hold for $\cos 2\pi n_k x$ under the weaker gap condition

$$n_{k+1}/n_k \geq 1 + ck^{-\alpha}, \qquad 0 \leq \alpha < 1/2$$

(see Erdős [7] and Takahashi [21, 22]). However, as the next theorem shows (which is contained in [4]), for having the permuted CLT and LIL the Hadamard gap condition (1.1) in Theorem 2.1 cannot be weakened.

Theorem 2.2. *For any positive sequence $(\varepsilon_k)_{k \geq 1}$ tending to 0, there exists a sequence $(n_k)_{k \geq 1}$ of positive integers satisfying*

$$n_{k+1}/n_k \geq 1 + \varepsilon_k, \qquad k \geq k_0$$

and a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the positive integers such that

$$a_N^{-1} \sum_{k=1}^{N} \cos 2\pi n_{\sigma(k)} x - b_N \xrightarrow{\mathcal{D}} G$$

where G is the Cauchy distribution with density $\frac{1}{\pi(1+x^2)}$, $a_N = \sqrt{N/\log N}$ and $b_N = O(1)$. Moreover, there exists a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the positive integers such that

$$(2.3) \qquad \limsup_{N \to \infty} \frac{\sum_{k=1}^{N} \cos 2\pi n_{\sigma(k)} x}{\sqrt{N \log N}} > 0 \quad \text{a.e.}$$

For subexponentially growing $(n_k)_{k \geq 1}$ the permuted CLT and LIL are much harder problems and we do not have a precise charaterization. Our next theorem gives sufficient Diophantine conditions in this case. Let us say that a sequence $(n_k)_{k \geq 1}$ of positive integers satisfies

Condition R, if for any $p \geq 2$ and any fixed nonzero integers $a_1, \ldots a_p$ satisfying $|a_1| \leq p, \ldots, |a_p| \leq p$, the Diophantine equation

$$(2.4) \qquad a_1 n_{k_1} + \ldots + a_p n_{k_p} = 0, \qquad k_1 < \ldots < k_p$$

has only a finite number of solutions. If the number of solutions of (2.4) is at most $C \exp(p^\alpha)$ for some constants $C, \alpha > 0$, we say that $(n_k)_{k \geq 1}$ satisfies condition $\mathbf{R}^*$.

Theorem 2.3. *If $(n_k)_{k \geq 1}$ is a sequence of positive integers satisfying condition $\mathbf{R}$, then for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the positive integers we have the central limit theorem (2.1). If $(n_k)_{k \geq 1}$ satisfies condition $\mathbf{R}^*$, we also have the permuted LIL (2.2).*

It is not easy to decide if an explicitly given subexponential sequence satisfies condition $\mathbf{R}$ or $\mathbf{R}^*$. A simple example is the sequence $(n_k)_{k \geq 1}$ consisting of the numbers

of the form $p_1^{k_1} \cdots p_r^{k_r}$ $(k_1, \ldots, k_r \geq 0)$, where $\{p_1, \ldots, p_r\}$ is a finite set of coprime integers. In this case the validity of $\mathbf{R}^*$ follows from a recent deep bound, due to Amoroso and Viada [5], in the subspace theorem of W. Schmidt [20]. On the other hand, it is easy to obtain sequences satisfying condition $\mathbf{R}^*$ via random constructions. Let $\omega_k = (\log k)^\alpha$, $\alpha > 0$ and let n_k be independent random variables, defined on some probability space, such that n_k is uniformly distributed over the integers of the interval $[a(k-1)^{\omega_k}, ak^{\omega_k}]$, where a is a large constant. Then with probability one, $(n_k)_{k \geq 1}$ satisfies condition $\mathbf{R}^*$. The so constructed sequence $(n_k)_{k \geq 1}$ grows much slower than exponentially, in fact its growth speed is only slightly faster than polynomial. In contrast to the CLT (1.2) which can hold even for sequences (n_k) with $n_{k+1} - n_k = O(1)$ (see Fukuyama [9]) we do not know if there are polynomially growing sequences $(n_k)_{k \geq 1}$ satisfying the permutation-invariant CLT and LIL.

§3.   The system $f(kx)$

Let f be a measurable function satisfying

$$f(x+1) = f(x), \qquad \int_0^1 f(x)\, dx = 0, \qquad \int_0^1 f^2(x)\, dx < \infty$$

and let $(n_k)_{k \geq 1}$ be a sequence of integers satisfying the Hadamard gap condition (1.1). The central limit theorem for $f(n_k x)$ has a long history. Kac [14] proved that if f is Lipschitz continuous then $f(n_k x)$ satisfies the CLT for $n_k = 2^k$ and not much later Erdős and Fortet (see [15, p. 655]) showed that this is not any more valid if $n_k = 2^k - 1$. Gaposhkin [11] proved that $f(n_k x)$ obeys the CLT if $n_{k+1}/n_k \to \alpha$ where α^r is irrational for $r = 1, 2 \ldots$ and the same holds if all the fractions n_{k+1}/n_k are integers. To formulate more general criteria, let us say that a sequence $(n_k)_{k \geq 1}$ of positive integers satisfies

Condition $\mathbf{D}_2$, if for any fixed nonzero integers a, b, c the number of solutions of the Diophantine equation

$$(3.1) \qquad\qquad an_k + bn_l = c, \qquad k, l \geq 1$$

is bounded by a constant $K(a,b)$, independent of c.

Condition $\mathbf{D}_2^*$ (strong $\mathbf{D}_2$), if for any fixed integers $a \neq 0$, $b \neq 0$, c the number of solutions of the Diophantine equation (3.1) is bounded by a constant $K(a,b)$, independent of c, where for $a = b, c = 0$ we require also $k \neq l$.

Condition $\mathbf{D}_2$ is a variant of Sidon's $\mathbf{B_2}$ condition, where the bound for the number of solutions of (3.1) is assumed only for $a = b = 1$. Gaposhkin [12] proved that

under minor smoothness assumptions on f, condition $\mathbf{D}_2$ implies the CLT for $f(n_k x)$. Recently, Aistleitner and Berkes [1] proved that the CLT holds for $f(n_k x)$ provided for any fixed nonzero integers a, b, c the number of solutions (k, l) of

$$an_k + bn_l = c, \qquad 1 \leq k, l \leq N$$

is $o(N)$, uniformly in c, and this condition is also necessary. This criterion settles the CLT problem for $f(n_k x)$, but, as we noted, the validity of the CLT does not imply permutation-invariant behavior of $f(n_k x)$. The purpose of this section is to give a precise characterization for the CLT and LIL for permuted sums $\sum_{k=1}^{N} f(n_{\sigma(k)} x)$.

Our first result gives a necessary and sufficient condition for permuted partial sums $\sum_{k=1}^{N} f(n_{\sigma(k)} x)$ to have only Gaussian limit distributions and gives precise criteria this to happen for a specific permutation σ.

Theorem 3.1. *Let $(n_k)_{k \geq 1}$ be a sequence of positive integers satisfying the Hadamard gap condition (1.1) and condition $\mathbf{D}_2$. Let f satisfy (1.5) and let σ be a permutation of $\mathbb{N}$. Then $N^{-1/2} \sum_{k=1}^{N} f(n_{\sigma(k)} x)$ has a limit distribution iff*

$$(3.2) \qquad \gamma = \lim_{N \to \infty} N^{-1} \int_0^1 \left(\sum_{k=1}^{N} f(n_{\sigma(k)} x) \right)^2 dx$$

exists, and then

$$(3.3) \qquad N^{-1/2} \sum_{k=1}^{N} f(n_{\sigma(k)} x) \xrightarrow{\mathcal{D}} N(0, \gamma).$$

If $\gamma = 0$, the limit distribution is degenerate.

Theorem 3.1 is best possible in the following sense:

Theorem 3.2. *If condition $\mathbf{D}_2$ fails, there exists a permutation σ such that the normed partial sums in (3.3) have a nongaussian limit distribution.*

In other words, under the Hadamard gap condition and condition $\mathbf{D}_2$, the limit distribution of $N^{-1/2} \sum_{k=1}^{N} f(n_{\sigma(k)} x)$ can only be Gaussian, but the variance of the limit distribution depends on the constant γ in (3.2) which, as simple examples show, is not permutation-invariant. For example, if $n_k = 2^k$ and σ is the identity permutation, then (3.2) holds with

$$\gamma = \gamma_f = \int_0^1 f^2(x) dx + 2 \sum_{k=1}^{\infty} \int_0^1 f(x) f(2^k x) dx$$

(see Kac [14]). Using an idea of Fukuyama [10], one can construct permutations σ of $\mathbb{N}$ such that

$$\lim_{N\to\infty} \frac{1}{N} \int_0^1 \left(\sum_{k=1}^N f(n_{\sigma(k)}x) \right)^2 dx = \gamma_{\sigma,f}$$

with $\gamma_{\sigma,f} \neq \gamma_f$. Actually, the set of possible values $\gamma_{\sigma,f}$ belonging to all permutations σ contains the interval $I_f = [\|f\|^2, \gamma_f]$ and it is equal to this interval provided the Fourier coefficients of f are nonnegative. For general f this is false (for details cf. [2]).

Under the slightly stronger condition $\mathbf{D}_2^*$ we get

Theorem 3.3.   *Let $(n_k)_{k\geq 1}$ be a sequence of positive integers satisfying the Hadamard gap condition (1.1) and condition $\mathbf{D}_2^*$. Let f satisfy (1.5) and let σ be a permutation of $\mathbb{N}$. Then the central limit theorem (3.3) holds with $\gamma = \|f\|^2$.*

We now pass to the problem of the LIL.

Theorem 3.4.   *Let $(n_k)_{k\geq 1}$ be a sequence of positive integers satisfying the Hadamard gap condition (1.1) and condition $\mathbf{D}_2$. Let f be a measurable function satisfying (1.5), let σ be a permutation of $\mathbb{N}$ and assume that*

$$\gamma = \lim_{N\to\infty} N^{-1} \int_0^1 \left(\sum_{k=1}^N f(n_{\sigma(k)}x) \right)^2 dx$$

for some $\gamma \geq 0$. Then we have

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

If instead of condition $\mathbf{D}_2$ we assume $\mathbf{D}_2^*$, in (3.4) we have $\gamma = \|f\|$.

Theorem 3.2 shows that condition $\mathbf{D_2}$ in Theorem 3.1 is best possible: if there exist nonzero integers a, b and a sequence $(c_m)_{m\geq 1}$ of integers such that the number of solutions (k, l), $k \neq l$ of

$$an_k + bn_l = c_m$$

tends to infinity as $m \to \infty$, the CLT for $f(n_{\sigma(k)}x)$ fails for a suitable permutation σ and a trigonometric polynomial f. We cannot prove the analogous statement in Theorem 3.4. However, the conclusion of Theorem 3.4 fails to hold for appropriate f and σ provided there exist nonzero integers a, b, and a fixed integer c, such that the Diophantine equation (3.1) has infinitely many solutions $(k, l), k \neq l$.

The previous results describe quite precisely the permutation-invariant CLT and LIL under the Hadamard gap condition (1.1). If $(n_k)_{k\geq 1}$ grows faster than exponentially, i.e.

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

then condition $\mathbf{D}_2^*$ is satisfied, and thus so are the permutational-invariant CLT and LIL. In [3] we have shown that given a 1-periodic Lipschitz α function ($0 < \alpha \le 1$) satisfying $\int_0^1 f(x)\,dx = 0$, under the slightly stronger gap condition

$$\sum_{k=1}^{\infty} (n_k/n_{k+1})^{\alpha} < \infty, \qquad 0 < \alpha < 1$$

there exists a sequence $(g_k(x))_{k\ge 1}$ of measurable functions on $(0,1)$, i.i.d. in the probabilistic sense, such that

$$\sum_{k=1}^{\infty} |f(n_k x) - g_k(x)| < \infty \qquad \text{a.e.}$$

This implies not only the CLT and LIL for any rearrangement of $(f(n_k x))_{k\ge 1}$, but also the permutation-invariance of practically all asymptotic properties of $(f(n_k x))_{k\ge 1}$.

§4.   Proofs

To keep our paper at reasonable length, we will give here the proofs only for the CLT case. The arguments for the LIL use similar ideas, but they are considerably more complicated and will be given in a subsequent paper. As the proof of Theorem 3.1 will show, in the case when f is a trigonometric polynomial of degree d, it suffices to assume condition $\mathbf{D}_2$ with coefficients a, b satisfying $|a| \le d, |b| \le d$. In particular, in the trigonometric case $f(x) = \cos 2\pi x$ it suffices to allow only coefficients ± 1 in condition $\mathbf{D}_2$, i.e. condition $\mathbf{D_2}$ reduces to Sidon's $\mathbf{B_2}$ condition mentioned in Section 3. This condition is satisfied for any Hadamard lacunary sequence $(n_k)_{k\ge 1}$ (see e.g. Zygmund [23, pp. 203-204]) and thus Theorem 2.1 is contained in the proof of Theorem 3.1.

Theorem 3.3 follows from Theorem 3.1 and the following

Lemma 4.1.   *Let $(n_k)_{k\ge 1}$ be a lacunary sequence of positive integers satisfying condition $\mathbf{D}_2^*$, and let f be a function satisfying (1.5). Then for any permutation σ of $\mathbb{N}$*

$$\lim_{N\to\infty} N^{-1} \int_0^1 \left(\sum_{k=1}^{N} f(n_{\sigma(k)}x) \right)^2 dx = \|f\|^2.$$

Proof of Lemma 4.1: For the simplicity of writing we assume that f is an even function, i.e. the Fourier series of f is of the form

$$f \sim \sum_{j=1}^{\infty} a_j \cos 2\pi j x.$$

Let $\varepsilon > 0$ be arbitrary. We choose $d > \varepsilon^{-1}$ and define

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

Since by assumption $\mathrm{Var}_{[0,1]} f < \infty$ we have

$$|a_j| = \mathcal{O}\left(j^{-1}\right), \qquad j \to \infty$$

(cf. Zygmund [23, p. 48]), and therefore

$$\text{(4.1)} \qquad \|r\| = \sum_{j=d+1}^{\infty} a_j^2/2 \ll \sum_{j=d+1}^{\infty} j^{-2} \ll d^{-1} \ll \varepsilon$$

By a classical norm inequality for lacunary series we have

$$\text{(4.2)} \qquad \left\| \sum_{k=1}^{N} r(n_{\sigma(k)}x) \right\| \ll d^{-1/2}\sqrt{N},$$

where the implied constant depends only on f and the growth factor q (see e.g. [1, Lemma 2.1]). Thus we have

$$\left\| \sum_{k=1}^{N} r(n_{\sigma(k)}x) \right\| \ll \sqrt{\varepsilon N}.$$

By the orthogonality of the trigonometric system and Minkowski's inequality we have

$$\left\| \sum_{k=1}^{N} f(n_{\sigma(k)}x) \right\| \leq \left\| \sum_{k=1}^{N} p(n_{\sigma(k)}x) \right\| + \left\| \sum_{k=1}^{N} r(n_{\sigma(k)}x) \right\|$$

$$\leq \left(\sum_{j=1}^{d}\sum_{k=1}^{N} \frac{a_j^2}{2} + \underbrace{\sum_{k_1,k_2=1}^{N}\sum_{j_1,j_2=1}^{d}}_{(j_1,k_1)\neq(j_2,k_2)} \frac{a_{j_1}a_{j_2}}{2} \mathbf{1}(j_1 n_{\sigma(k_1)} = j_2 n_{\sigma(k_2)}) \right)^{1/2} + \mathcal{O}\left(\sqrt{\varepsilon N}\right).$$

Since by condition $\mathbf{D}_2^*$ the number of nontrivial solutions of $j_1 n_{k_1} - j_2 n_{k_2} = 0$ is bounded by a constant (where we can choose the same constant for all finitely many possible values of j_1, j_2), we get

$$\text{(4.3)} \qquad \left\| \sum_{k=1}^{N} f(n_{\sigma(k)}x) \right\| \leq \sqrt{N}\|p\| + \mathcal{O}(\sqrt{\varepsilon N}).$$

A lower bound can be found in exactly the same way, and since the implied constant in (4.3) does not depend on ε (and d, resp.), we obtain, utilizing (4.1),

$$\left| \left\| \sum_{k=1}^{N} f(n_{\sigma(k)}x) \right\| - \sqrt{N}\|f\| \right| \ll \sqrt{\varepsilon N},$$

where the constant implied by $\ll$ is independent of ε. Since ε can be chosen arbitrarily small, this proves the lemma. $\square$

The proof of Theorem 2.3 is implicit in our paper [2]. Let us say that an increasing sequence $(n_k)_{k\geq 1}$ of positive integers satisfies Condition $\mathbf{R}^{**}$ if for any $p \geq 2$ the number of nondegenerate solutions of the Diophantine equation

$$(4.4) \qquad\qquad \pm n_{k_1} \pm \ldots \pm n_{k_p} = 0, \qquad k_1, \ldots, k_p \geq 1$$

is at most $C\exp(p^\alpha)$ for some constants $C, \alpha > 0$ (here a solution of (4.4) is called nondegenerate if no subsum of the left hand side equals 0). As an analysis of the proof of the main theorem in [2] shows, the proof uses only condition $\mathbf{R}^{**}$, and by collecting the terms in (4.4) with equal indices shows that condition $\mathbf{R}^*$ implies condition $\mathbf{R}^{**}$.

To prove the remark made at the end of Section 2, let $\omega_k = (\log k)^\alpha$ and let n_k, $k = 1, 2, \ldots$ be independent random variables such that n_k is uniformly distributed over the integers of the interval $I_k = [a(k-1)^{\omega_{k-1}}, ak^{\omega_k}]$. Note that the length of I_k is at least $a\omega_k(k-1)^{\omega_{k-1}} \geq a\omega_1$ for $k = 2, 3, \ldots$ and equals a for $k = 1$ and thus choosing a large enough, each I_k contains at least one integer. We claim that, with probability 1, the sequence $(n_k)_{k\geq 1}$ satisfies condition $\mathbf{R}^*$. To see this, set $\eta_k = \frac{1}{2}\omega_k^{1/2}$, then

$$(4.5) \qquad\qquad (2k)^{\eta_k^2 + 2\eta_k} \leq (2k)^{\omega_k/2} \leq k^{-2}|I_k| \qquad \text{for } k \geq k_0$$

since, as we noted, $|I_k| \geq a\omega_k(k-1)^{\omega_k-1} \geq (k/2)^{\omega_k-1}$ for large k. Let $k \geq 1$ and consider the numbers of the form

$$(4.6) \qquad\qquad (a_1 n_{i_1} + \ldots + a_s n_{i_s})/d$$

where $1 \leq s \leq \eta_k$, $1 \leq i_1, \ldots, i_s \leq k-1$, $a_1, \ldots a_s, d$ are nonzero integers with $|a_1|, \ldots, |a_s|, |d| \leq \eta_k$. Since the number of values in (4.6) is at most $(2k)^{\eta_k^2 + 2\eta_k}$, (4.5) shows that the probability that n_k equals any of these numbers is at most k^{-2}. Thus by the Borel-Cantelli lemma, with probability 1 for $k \geq k_1$, n_k will be different from all the numbers in (4.6) and thus the equation

$$a_1 n_{i_1} + \ldots + a_s n_{i_s} + a_{s+1} n_k = 0$$

has no solution with $1 \leq s \leq \eta_k$, $1 \leq i_1 < \ldots < i_s \leq k-1$, $0 < |a_1|, \ldots, |a_{s+1}| \leq \eta_k$. By monotonicity, the equation

$$(4.7) \qquad\qquad a_1 n_{i_1} + \ldots + a_s n_{i_s} = 0, \qquad i_1 < \ldots < i_s$$

has no solutions provided the number of terms is at most η_k, $0 < |a_1|, \ldots, |a_s| \leq \eta_k$ and $i_s \geq k$. Thus the number of solutions of (4.7) where $s \leq \eta_k$ and $0 < |a_1|, \ldots, |a_s| \leq \eta_k$ is at most $k^{\eta_k} \ll \exp(\eta_k^\beta)$ for some $\beta > 0$. Since the sequence $[\eta_k]$ takes all sufficiently large integers p, the sequence $(n_k)_{k \geq 1}$ satisfies condition $\mathbf{R}^*$.

It remains now to prove Theorems 3.1 and 3.2, which will be done in the next two sections of the paper.

§ 4.1.   Proof of Theorem 3.1.

In this section we give the proof of Theorem 3.1. The proof of the main lemma (Lemma 4.3) uses ideas of Révész [18]. We will need the following simple smoothing inequality.

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 $P_1^* = P_1 \star H$, $P_2^* = P_2 \star H$, where H is a normal distribution with mean zero and variance τ^2. Then for all $y > 0, T > 0$,*

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

Proof: Letting $h(s) = \exp(-\tau^2 s^2/2)$, the characteristic functions of P_1^* and P_2^* are $p_1^* = p_1 h$ and $p_2^* = p_2 h$ and thus for the densities ρ_1 and ρ_2 of P_1^* and P_2^* we have, letting $w(s) = p_1(s) - p_2(s)$,

$$|\rho_1(u) - \rho_2(u)| \leq (2\pi)^{-1}\left|\int_{\mathbb{R}} e^{-isu}\,(p_1^*(s) - p_2^*(s))\, ds\right|$$
$$\leq (2\pi)^{-1}\int_{s \in [-T,T]} |w(s)| ds + (2\pi)^{-1}\, 2\int_{s \notin [-T,T]} |h(s)| ds.$$

Thus

$$|P_1^*([-y,y]) - P_2^*([-y,y])| \leq \int_{[-y,y]} |\rho_1(u) - \rho_2(u)| du$$
$$\leq y\int_{s \in [-T,T]} |w(s)|\, ds + y\int_{s \notin [-T,T]} |h(s)|\, ds,$$

and the statement of the lemma follows from

$$\int_{s \notin [-T,T]} |h(s)|\, ds = 2\int_T^\infty e^{-\tau^2 s^2/2}\, ds \leq 4\tau^{-1} e^{-\tau^2 T^2/2}. \qquad \square$$

Let f be a function satisfying (1.5) and $(n_k)_{k \geq 1}$ a sequence satisfying the Hadamard gap condition (1.1) and condition $\mathbf{D}_2$, and let

$$p(x) = \sum_{j=1}^d (a_j \cos 2\pi j x + b_j \sin 2\pi j x)$$

be the d-th partial sum of the Fourier series of f for some $d \geq 1$. By condition $\mathbf{D}_2$ we can find a number C_1 (depending on d) such that for any a, b satisfying $0 < |a| \leq d$, $0 < |b| \leq d$

$$(4.8) \qquad \# \{k_1, k_2 \geq 1 : \ an_{k_1} - bn_{k_2} = c\} \leq C_1,$$

for all $c \in \mathbb{Z}\backslash\{0\}$. We set

$$\gamma_N = N^{-1/2} \left\| \sum_{k=1}^{N} f(n_{\sigma(k)}x) \right\|.$$

Lemma 4.3. *There exists an N_0, depending only on p, d and the constants implied by condition $\mathbf{D}_2$ (but not on σ) such that for any $N \geq N_0$ there exists a set $A \subset \{1, \ldots, N\}$ with $\#A \geq N - N(\log_q(2d))/(\log N)^{1/2}$ such that*

$$\left| \mathbb{E}\left(\exp\left(isN^{-1/2} \sum_{k \in A} p(n_{\sigma(k)}x) \right) \right) - e^{-s^2\delta_N^2/2} \right| \leq 2N^{-1/8}$$

for all $s \in [-(\log N)^{1/8}, (\log N)^{1/8}]$, where

$$\delta_N = N^{-1} \left\| \sum_{k \in A} p(n_{\sigma(k)}x) \right\|.$$

Proof. For the simplicity of writing we assume that

$$p(x) = \sum_{j=1}^{d} a_j \cos 2\pi jx$$

is an even function; the proof in the general case is similar. Let $(\nu_k)_{1 \leq k \leq N}$ denote the sequence $(n_{\sigma(k)})_{1 \leq k \leq N}$ arranged in increasing order. Put

$$\Delta = \left\{ k \in \{1, \ldots, N\} : \ k \mod \lceil (\log N)^{1/2} + \log_q(2d) \rceil \notin \{0, \ldots, \lceil \log_q(2d) \rceil\} \right\}$$

$$\Delta^{(h)} = \left\{ k \in \Delta : \ \frac{k}{\lceil (\log N)^{1/2} + \log_q 2d \rceil} \in [h, h+1) \right\}, \qquad h \geq 0,$$

$$A = \{1 \leq l \leq N : \ n_{\sigma(l)} \in (\nu_k)_{k \in \Delta}\},$$

$$\alpha(s) = \prod_{h \geq 0} \left(1 + isN^{-1/2} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right).$$

and

$$\eta_N = N^{-1/2} \sum_{k \in \Delta} p(\nu_k x).$$

The sets $\Delta^{(h)}$ are constructed in such a way that for $k_1 \in \Delta^{(h_1)}$, $k_2 \in \Delta^{(h_2)}$, where $h_1 < h_2$, we have

$$(4.9) \qquad \nu_{k_2}/\nu_{k_1} > q^{k_2-k_1} \geq q^{\lceil \log_q(2d) \rceil} \geq 2d.$$

Using

$$(4.10) \qquad e^{ix} = (1+ix)e^{-x^2/2+w(x)}, \qquad |w(x)| \leq |x|^3$$

we get

$$(4.11) \qquad e^{is\eta_N} = \prod_{h \geq 0} \exp\left(isN^{-1/2} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right)$$

$$(4.12) \qquad = \alpha(s) \, \exp\left(-(2N)^{-1} \sum_{h \geq 0} s^2 \left(\sum_{k \in \Delta^{(h)}} p(\nu_k x) \right)^2 \right)$$

$$(4.13) \qquad \times \, \exp\left(\sum_{h \geq 0} w\left(isN^{-1/2} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right) \right).$$

We have

$$(4.14) \quad \sum_{h \geq 0} \left(\sum_{k \in \Delta^{(h)}} p(\nu_k x) \right)^2$$

$$= \sum_{h \geq 0} \sum_{k_1,k_2 \in \Delta^{(h)}} \sum_{j_1,j_2=1}^{d} \frac{a_{j_1} a_{j_2}}{2} \left(\cos(2\pi(j_1\nu_{k_1} + j_2\nu_{k_2})x) + \cos(2\pi(j_1\nu_{k_1} - j_2\nu_{k_2})x) \right)$$

$$= \sum_{h \geq 0} \sum_{k_1,k_2 \in \Delta^{(h)}} \sum_{j_1,j_2=1}^{d} \frac{a_{j_1} a_{j_2}}{2} \left(\cos(2\pi(j_1\nu_{k_1} + j_2\nu_{k_2})x) \right)$$

$$+ \sum_{h \geq 0} \underbrace{\sum_{k_1,k_2 \in \Delta^{(h)}} \sum_{j_1,j_2=1}^{d}}_{j_1\nu_{k_1} - j_2\nu_{k_2}=0} \frac{a_{j_1} a_{j_2}}{2}$$

$$+ \sum_{h \geq 0} \underbrace{\sum_{k_1,k_2 \in \Delta^{(h)}} \sum_{j_1,j_2=1}^{d}}_{j_1\nu_{k_1} - j_2\nu_{k_2} \neq 0} \frac{a_{j_1} a_{j_2}}{2} \left(\cos(2\pi(j_1\nu_{k_1} - j_2\nu_{k_2})x) \right)$$

$$(4.15) \quad = \delta_N^2 N + \beta(x),$$

where

$$\delta_N = N^{-1} \left(\sum_{h \geq 0} \underbrace{\sum_{k_1, k_2 \in \Delta^{(h)}} \sum_{j_1, j_2 = 1}^{d}}_{j_1 \nu_{k_1} - j_2 \nu_{k_2} = 0} \frac{a_{j_1} a_{j_2}}{2} \right)^{1/2}$$

and

$$\beta(x) = \sum_{h \geq 0} \sum_{k_1, k_2 \in \Delta^{(h)}} \sum_{j_1, j_2 = 1}^{d} \frac{a_{j_1} a_{j_2}}{2} \left(\cos(2\pi (j_1 \nu_{k_1} + j_2 \nu_{k_2}) x) \right)$$

$$+ \sum_{h \geq 0} \underbrace{\sum_{k_1, k_2 \in \Delta^{(h)}} \sum_{j_1, j_2 = 1}^{d}}_{j_1 \nu_{k_1} - j_2 \nu_{k_2} \neq 0} \frac{a_{j_1} a_{j_2}}{2} \left(\cos(2\pi (j_1 \nu_{k_1} - j_2 \nu_{k_2}) x) \right).$$

Note that

$$\delta_N^2 = N^{-1} \int_0^1 \left(\sum_{h \geq 0} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right)^2 dx = N^{-1} \int_0^1 \left(\sum_{k \in \Delta} p(\nu_k x) \right)^2 dx,$$

since by (4.9) the functions

$$\sum_{k \in \Delta^{(h_1)}} p(\nu_k x) \qquad \text{and} \qquad \sum_{k \in \Delta^{(h_2)}} p(\nu_k x)$$

are orthogonal for $h_1 \neq h_2$. From (4.11), (4.12), (4.13), (4.14) and (4.15) we conclude

$$e^{is\eta_N} = \alpha(s) \, \exp \left(-\frac{s^2}{2} \left(\delta_N^2 + \frac{\beta}{N} \right) + \sum_{h \geq 0} w \left(is N^{-1/2} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right) \right)$$

and, writing

$$w^*(s, x) = \sum_{h \geq 0} w \left(is N^{-1/2} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right),$$

(4.10) implies

$$|w^*(s, x)| \leq N \left| s N^{-1/2} \sum_{k \in \Delta^{(h)}} p(\nu_k x) \right|^3$$

$$\text{(4.16)} \qquad\qquad \leq |s|^3 \, \|p\|_\infty^3 \, (\log N)^{3/2} \, N^{-1/2}.$$

We note that

$$(4.17) \qquad |\alpha(s)| \le \prod_{h \ge 0} \left(1 + N^{-1} s^2 \left(\sum_{k \in \Delta^{(h)}} p(\nu_k x) \right)^2 \right)^{1/2}$$

$$\le \exp \left((2N)^{-1} \sum_{h \ge 0} s^2 \left(\sum_{k \in \Delta^{(h)}} p(\nu_k x) \right)^2 \right)$$

$$(4.18) \qquad \le \exp \left(\frac{s^2}{2} \left(\delta_N^2 + \frac{\beta(x)}{N} \right) \right)$$

and obtain, for $\varphi(s) = \mathbb{E} e^{i s \eta_N}$, using $\mathbb{E}\alpha(s) = 1$ and (4.17), (4.18),

$$\left| \varphi(s) - e^{-s^2 \delta_N^2 / 2} \right|$$

$$= \left| \mathbb{E} \left(\alpha(s) \exp \left(-\frac{s^2}{2} \left(\delta_N^2 + \frac{\beta(x)}{N} \right) + w^*(s,x) \right) \right) - e^{-s^2 \delta_N^2 / 2} \right|$$

$$= \left| \mathbb{E} \left(\alpha(s) \left(\exp \left(-\frac{s^2}{2} \left(\delta_N^2 + \frac{\beta(x)}{N} \right) + w^*(s,x) \right) - e^{-s^2 \delta_N^2 / 2} \right) \right) \right|$$

$$\le \left| \mathbb{E} \left(|\alpha(s)| \left| \exp \left(-\frac{s^2}{2} \left(\delta_N^2 + \frac{\beta(x)}{N} \right) + w^*(s,x) \right) - e^{-s^2 \delta_N^2 / 2} \right| \right) \right|$$

$$\le \mathbb{E} \left| e^{w^*(s,x)} - 1 \right| + \mathbb{E} \left| \exp \left(\frac{s^2}{2} \frac{\beta(x)}{N} \right) - 1 \right|.$$

If $|s| \le (\log N)^{1/8}$, then by (4.16)

$$\|p\|_\infty^3 \, (\log N)^{15/8} \, N^{-1/2} \le N^{-1/4} \qquad \text{for} \qquad N \ge N_1$$

with N_1 depending only on p, and hence

$$\left| e^{w^*(s,x)} - 1 \right| \le 2N^{-1/4}, \qquad \text{for } N \ge N_1.$$

On the other hand, the function $\beta(x)$ is a sum of at most

$$2 \sum_{h \ge 0} d^2 \left| \Delta^{(h)} \right|^2 \le 2 d^2 N (\log N)^{1/2}$$

trigonometric functions, and the coefficient of each of this summands is bounded by $\max_{1 \le j \le d} a_j^2 / 2 \le C_2$, where C_2 depends only on p and d. Since by assumption $(n_k)_{k \ge 1}$ satisfies condition $\mathbf{D}_2$, there can be at most $4d^2 C_1$ summands giving the same frequency (the constant C_1 is defined in (4.8)). This means, writing $\beta(x)$ in the form

$$\beta(x) = \sum_{j=1}^{\infty} c_j \cos 2\pi j x$$

we have

$$(4.19) \qquad |c_j| \le 4d^2 C_1 C_2, \quad j \ge 1, \qquad \text{and} \qquad \|\beta\|_\infty \le \sum_{j=1}^{\infty} |c_j| \le 2C_2 d^2 N \sqrt{\log N}$$

Therefore

$$\|\beta\|^2 = \sum_{j=1}^{\infty} c_j^2/2 \le \max_{j \ge 1} |c_j| \sum_{j=1}^{\infty} |c_j| \le (4d^2 C_1 C_2)^2 \, 2C_2 d^2 N \sqrt{\log N},$$

and

$$\mathbb{P}(|\beta| > N^{2/3}) \le N^{-1/4} \qquad \text{for} \qquad N \ge N_2,$$

where N_2 depends only on p, d, C_1. Hence by (4.19)

$$\mathbb{E} \left| \exp\left(\frac{s^2}{2} \frac{\beta(x)}{N} \right) - 1 \right|$$

$$\le \left| \exp\left(\frac{s^2}{2} \frac{\|\beta\|_\infty}{N} \right) - 1 \right| N^{-1/4} + \left| \exp\left(\frac{s^2}{2} \frac{N^{2/3}}{N} \right) - 1 \right|,$$

and, assuming

$$|s| \le (\log N)^{1/8},$$

we get

$$\mathbb{E} \left| \exp\left(\frac{s^2}{2} \frac{\beta(x)}{N} \right) - 1 \right|$$

$$\le \left| \exp\left(C_2 d^2 (\log N)^{3/4} \right) - 1 \right| N^{-1/4} + \left| \exp\left(\frac{(\log N)^{1/4} N^{2/3}}{N} \right) - 1 \right|$$

$$\le N^{-1/8} \qquad \text{for} \qquad N \ge N_3,$$

where N_3 depends on d, p, C_1. Combining these estimates, we have

$$(4.20) \qquad\qquad |\varphi(s) - e^{s^2 \delta_N^2/2}| \le 2N^{-1/8} \quad \text{for} \quad N \ge N_4$$

where N_4 also depends on p, d, C_1. This proves Lemma 4.3. $\quad\square$

Proof of Theorem 3.1. Let y and $\varepsilon > 0$ be given. We choose $d > \varepsilon^{-2}$ and write p and r, respectively, for the d-th partial sum and d-th remainder term of the Fourier series of f. Assume that $N \ge N_0$ with the N_0 in Lemma 4.3, and let A be the set in Lemma 4.3. By Lemma 4.3 we have

$$|\varphi_N(s) - e^{s^2 \delta_N^2/2}| \le 2N^{-1/8},$$

where

$$\varphi_N(s) = \mathbb{E}e^{is\eta_N}, \qquad \eta_N = \sum_{k\in\Delta} p(n_{\sigma(k)}x),$$

$$\delta_N^2 = N^{-1} \int_0^1 \left(\sum_{k\in A} p(n_{\sigma(k)}x) \right)^2 dx.$$

We recall that

$$\gamma_N^2 = N^{-1} \int_0^1 \left(\sum_{k=1}^N f(n_{\sigma(k)}x) \right)^2 dx.$$

By (4.2) we have

$$(4.21) \qquad \left\| \sum_{k=1}^N r(n_{\sigma(k)}x) \right\| \le C_4 d^{-1/2} \sqrt{N},$$

where C_4 depends on f, q. By Minkowski's inequality, and using (4.21) for our choice of $d > \varepsilon^{-2}$,

$$\gamma_N \sqrt{N} \le \left\| \sum_{k=1}^N p(n_{\sigma(k)}x) \right\| + \left\| \sum_{k=1}^N r(n_{\sigma(k)}x) \right\| \le \left\| \sum_{k=1}^N p(n_{\sigma(k)}x) \right\| + C_4 \varepsilon \sqrt{N},$$

and, since $\left\| \sum_{k=1}^N p(n_{\sigma(k)}x) \right\| \le C_5 \sqrt{N}$ (for C_5 depending on f, q, again by [1, Lemma 2.1]), we obtain

$$\gamma_N^2 \le N^{-1} \left\| \sum_{k=1}^N p(n_{\sigma(k)}x) \right\|^2 + C_6 \varepsilon$$

(C_6 depends on f, q). Since $N - \#A \le N(\log_q(2d))/\sqrt{\log N}$ we have

$$\gamma_N^2 - \delta_N^2 \le C_6 \varepsilon + 2N^{-1} \int_0^1 \sum_{k_1, k_2 \in \{1,\ldots,N\}, \; k_2 \notin \Delta} p(\nu_{k_1} x) p(\nu_{k_2} x) \, dx$$

$$\le C_6 \varepsilon + 2N^{-1} \|p\|^2 \sum_{k_2 \in \{1,\ldots,N\}, k_2 \notin \Delta} \#\{k_1 \in \{1,\ldots,N\} : \; k_1/k_2 \in [1/d, d]\}$$

$$\le C_6 \varepsilon + 2N^{-1} \|p\|^2 N (\log N)^{-1/2} \lceil \log_q 2d \rceil^2$$

$$(4.22) \qquad \le C_7 \varepsilon \qquad \text{for} \qquad N \ge N_5,$$

where C_7 depends on f, q and N_5 depends on f, q, d. We assume $s \le \varepsilon^{-1/2}$, and, in view of (4.20), we arrive at

$$|\varphi_N(s) - e^{s^2\gamma_N^2/2}| \le 2N^{-1/8} + |e^{s^2\gamma_N^2/2} - e^{s^2\delta_N^2/2}| \le C_8 \varepsilon$$

for $N \geq N_6$, where C_8 depends on f, q and N_6 depends on f, q, d. The proof of (4.22) also shows

$$\left\| \sum_{1 \leq k \leq N, \ k \notin A} p(n_{\sigma(k)}x) \right\| \leq C_9 \varepsilon, \qquad N \geq N_7,$$

and thus

$$(4.23) \qquad \mathbb{P} \left(\left| \sum_{1 \leq k \leq N, \ k \notin A} p(n_{\sigma(k)}x) \right| > \varepsilon^{1/6} \right) \leq C_9 \varepsilon^{2/3}, \qquad N \geq N_7,$$

for C_9 depending on f, q and N_7 depending on f, q, d.

We use Lemma 4.2 with

$$p_1(s) = \varphi(s), \qquad p_2(s) = e^{s^2 \gamma_N^2 / 2}, \qquad T = \varepsilon^{-1/2}, \qquad \tau = \varepsilon^{1/3}$$

to get (using the notation from this lemma) for all $y > 0$

$$|P_1^*([-y, y]) - P_2^*([-y, y])| \leq y \int_{s \in [-T, T]} C_8 \varepsilon \, ds + 4y \varepsilon^{-1/3} e^{-\varepsilon^{-1/3}/2}$$

for sufficiently large N (depending on f, q, d), provided ε is sufficiently small, which we can assume. Thus, if P_1, P_2 are the measures corresponding to p_1, p_2, we get for sufficiently large N (depending on f, q, d), using (4.21) and (4.23),

$$\mathbb{P} \left(N^{-1/2} \sum_{k=1}^{N} f(n_{\sigma(k)}x) \in [-y, y] \right)$$

$$\leq P_1([-y - 2\varepsilon^{1/6}, y + 2\varepsilon^{1/6}]) + \mathbb{P} \left(\left| N^{-1/2} \sum_{k=1}^{N} r(n_{\sigma(k)}x) \right| > \varepsilon^{1/6} \right)$$

$$+ \mathbb{P} \left(\left| \sum_{1 \leq k \leq N, \ k \notin A} p(n_{\sigma(k)}x) \right| > \varepsilon^{1/6} \right)$$

$$\leq P_1^*([-y - 3\varepsilon^{1/6}, y + 3\varepsilon^{1/6}]) + H(\mathbb{R} \backslash [-\varepsilon^{1/6}, \varepsilon^{1/6}]) + C_4 \varepsilon^{2/3} + C_9 \varepsilon^{2/3}$$

$$\leq P_2^*([-y - 3\varepsilon^{1/6}, y + 3\varepsilon^{1/6}]) + 2C_8(y + 3\varepsilon^{1/6})\varepsilon^{1/2}$$

$$\quad + 4(y + 3\varepsilon^{1/6})\varepsilon^{-1/3} e^{-\varepsilon^{-1/3}/2} + \frac{2}{\sqrt{2\pi\tau^2}} \int_{\varepsilon^{1/6}}^{\infty} e^{-u^2/(2\tau^2)} \, du + (C_4 + C_9)\varepsilon^{2/3}$$

$$\leq P_2([-y - 4\varepsilon^{1/6}, y + 4\varepsilon^{1/6}]) + 2C_8(y + 3\varepsilon^{1/6})\varepsilon^{1/2}$$

$$\quad + 4(y + 3\varepsilon^{1/6})\varepsilon^{-1/3} e^{-\varepsilon^{-1/3}/2} + \frac{4}{\sqrt{2\pi\tau^2}} \int_{\varepsilon^{1/6}}^{\infty} e^{-u^2/(2\tau^2)} \, du + (C_4 + C_9)\varepsilon^{2/3}$$

$$\leq P_2([-y, y])$$

$$(4.24) \quad + \frac{2}{\sqrt{2\pi\gamma_N^2}} \int_y^{y+4\varepsilon^{1/6}} e^{-u^2/(2\gamma_N^2)} \, du$$

$$(4.25) \quad + 2C_8(y + 3\varepsilon^{1/6})\varepsilon^{1/2} + 4(y + 3\varepsilon^{1/6})\varepsilon^{-1/3}e^{-\varepsilon^{-1/3}/2}$$

$$(4.26) \quad + \frac{4}{\sqrt{2\pi\varepsilon^{2/3}}} \int_{\varepsilon^{1/6}}^\infty e^{-u^2/(2\varepsilon^{2/3})} \, du + (C_4 + C_9)\varepsilon^{2/3}.$$

It is clear that for $\gamma_N \to 0$ the limit distribution of $N^{-1/2} \sum_{k=1}^N f(n_{\sigma(k)}x)$ is the distribution concentrated at 0. Now assume

$$\gamma_N \to \gamma \qquad \text{as} \quad N \to \infty \qquad \text{for some} \quad \gamma > 0.$$

Then $\liminf_{N\to\infty} \gamma_N > 0$, and since ε can be chosen arbitrarily, the value of (4.24), (4.26) and (4.25) can also be made arbitrarily small. This means that for any given $\hat{\varepsilon} > 0$ and sufficiently large N

$$(4.27) \qquad \mathbb{P}\left(x : \ N^{-1/2} \sum_{k=1}^N f(n_{\sigma(k)}x) \in [-y, y] \right)$$

$$\leq \frac{1}{\sqrt{2\pi\gamma_N^2}} \int_{-x}^x e^{-u^2/(2\gamma_N^2)} \, du + \hat{\varepsilon}$$

$$\leq \frac{1}{\sqrt{2\pi\gamma^2}} \int_{-x}^x e^{-u^2/(2\gamma^2)} \, du + 2\hat{\varepsilon}.$$

In the same way we can get a lower bound for (4.27), proving the first part of Theorem 3.1. On the other hand, the sequence γ_N is bounded since $(n_k)_{k\geq 1}$ is lacunary and thus if γ_N is not convergent, it has at least two accumulation points $\gamma^{(1)} \neq \gamma^{(2)}$. Thus there exist two subsequences of $\mathbb{N}$ along which $N^{-1/2} \sum_{k=1}^N f(n_{\sigma(k)}x)$ converges to different distributions $\mathcal{N}(0, \gamma^{(1)})$, $\mathcal{N}(0, \gamma^{(2)})$ (one of them may be the measure concentrated at 0, in case $\gamma^{(1)} = 0$ or $\gamma^{(2)} = 0$). But then $N^{-1/2} \sum_{k=1}^N f(n_{\sigma(k)}x)$ does not have a single limit distribution as $N \to \infty$.

§ 4.2.   Proof of Theorem 3.2.

In this section we prove Theorem 3.2 stating that the Diophantine condition $\mathbf{D}_2$ in Theorem 3.1 is best possible. In other words we show that if a lacunary sequence $(n_k)_{k\geq 1}$ does not satisfy condition $\mathbf{D}_2$, then there exist a trigonometric polynomial f satisfying (1.5) and a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ such that

$$N^{-1/2} \sum_{k=1}^N f(n_{\sigma(k)}x)$$

has a non-Gaussian limiting distribution.

Let $(n_k)_{k \geq 1}$ be given, and assume that condition $\mathbf{D}_2$ does not hold for this sequence. Then there exist integers $0 < a \leq b$ and a sequence $(c_\ell)_{\ell \geq 1}$ of different positive integers such that

$$\# \{k_1, k_2 \in \mathbb{N} : \ an_{k_1} - bn_{k_2} = c_\ell\} \to \infty \qquad \text{as} \qquad \ell \to \infty.$$

Throughout this section we will assume that $a < b$; the case $a = b$ can be handled in a similar way, with some minor changes.

We divide the set of positive integers into consecutive blocks $\Delta_1, \Delta_2, \ldots, \Delta_m, \ldots$ of lenghts $2^{2^1}, 2^{2^2}, \ldots, 2^{2^m}, \ldots$; we will write $|\Delta_m|$ for the number of elements of a block Δ_m. We can find integers c_m, $m \geq 1$ and a permutation $\sigma : \ \mathbb{N} \to \mathbb{N}$ such that the following holds:

$$\sigma(k+1) > \sigma(k), \qquad\qquad k \geq 1$$
$$\sigma(2k) > 2b\, \sigma(2k-1), \qquad k \geq 1$$
(4.28) $$an_{\sigma(2k)} - bn_{\sigma(2k-1)} = c_m, \qquad 2k, 2k-1 \in \Delta_m, \ m \geq 1.$$

Let

$$f(x) = \cos 2\pi ax + \cos 2\pi bx.$$

To show that

$$N^{-1/2} \sum_{k=1}^{N} f\big(n_{\sigma(k)}x\big)$$

does not converge to a Gaussian distribution, we will show that

$$\frac{\sum_{k \in \Delta_m} f(n_{\sigma(k)}x)}{\sqrt{|\Delta_m|}}$$

has a non-Gaussian limit distribution as $m \to \infty$. Since the lengths of the block Δ_m dominates $|\Delta_1| + \cdots + |\Delta_{m-1}|$, this means that

$$\frac{\sum_{m=1}^{M} \sum_{k \in \Delta_m} f(n_{\sigma(k)}x)}{\sqrt{\sum_{m=1}^{M} |\Delta_m|}}$$

also has a non-Gaussian limiting distribution.

Lemma 4.4.

$$\left| \mathbb{E}\left(\exp\left(\frac{is \sum_{k \in \Delta_m} f(n_{\sigma(k)}x)}{\sqrt{|\Delta_m|}} \right) \right) - \mathbb{E}\left(e^{-s^2(1 + \cos 2\pi x)/2} \right) \right| \ll |\Delta_m|^{-1/8},$$

for all $s \in [-(\log |\Delta_m|)^{1/8}, -(\log |\Delta_m|)^{1/8}]$.

Proof. We write $(\nu_k)_{k\geq 1}$ for $(n_{\sigma(k)})_{k\geq 1}$ and define

$$\eta_m = \frac{\sum_{k\in\Delta_m} f(\nu_k x)}{\sqrt{|\Delta_m|}}.$$

For different k_1, k_2, for which $2k_1 - 1, 2k_2 - 1 \in \Delta_m$, the functions

$$1 + \frac{is\left(f(\nu_{2k_1-1}x) + f(\nu_{2k_1}x)\right)}{\sqrt{2}} \qquad \text{and} \qquad 1 + \frac{is\left(f(\nu_{2k_2-1}x) + f(\nu_{2k_2}x)\right)}{\sqrt{2}}$$

are orthogonal. This means, writing

$$\alpha_m(s) = \prod_{k\geq 1:\ 2k-1\in\Delta_m} \left(1 + \frac{is\left(f(\nu_{2k-1}x) + f(\nu_{2k}x)\right)}{\sqrt{2|\Delta_m|}}\right),$$

we have

$$\mathbb{E}\Big(\alpha_m(s)\Big) = 1.$$

Using again (4.10), we get

$$e^{is\eta_m} = \prod_{k\geq 1:\ 2k-1\in\Delta_m} \exp\left(\frac{is\left(f(\nu_{2k-1}x) + f(\nu_{2k}x)\right)}{\sqrt{2|\Delta_m|}}\right)$$

$$= \alpha_m(s)\ \exp\left(\sum_{k\geq 1:\ 2k-1\in\Delta_m} \frac{-s^2\left(f(\nu_{2k-1}x) + f(\nu_{2k}x)\right)^2}{2|\Delta_m|}\right)\ W_m(s),$$

where

$$W_m(s) = \exp\left(\sum_{k\geq 1:\ 2k-1\in\Delta_m} w\left(\frac{is\left(f(\nu_{2k-1}x) + f(\nu_{2k}x)\right)}{\sqrt{2|\Delta_m|}}\right)\right).$$

By (4.28) we have

$$\sum_{k\geq 1:\ 2k-1\in\Delta_m} \left(f(\nu_{2k-1}x) + f(\nu_{2k}x)\right)^2$$

$$= \sum_{k\geq 1:\ 2k-1\in\Delta_m} \left(\cos 2\pi a\nu_{2k-1}x + \cos 2\pi b\nu_{2k-1}x + \cos 2\pi a\nu_{2k}x + \cos 2\pi b\nu_{2k}x\right)^2$$

$$= |\Delta_m| + \left(\sum_{k\geq 1:\ 2k-1\in\Delta_m} \frac{\cos 2\pi(a\nu_{2k} - b\nu_{2k_1})x}{2}\right) + R_m(x)$$

$$= |\Delta_m| + \left(\sum_{k\geq 1:\ 2k-1\in\Delta_m} \frac{\cos 2\pi c_m x}{2}\right) + R_m(x)$$

$$= |\Delta_m| \left(1 + \frac{\cos 2\pi c_m x}{4}\right) + R_m(x),$$

where $R_m(x)$ is a sum of cosine functions with coefficients $1/2$ and frequencies

$$a\nu_{2k-1}, \ b\nu_{2k-1}, \ a\nu_{2k}, \ b\nu_{2k}, \ a\nu_{2k-1} \pm b\nu_{2k-1}, \ a\nu_{2k-1} \pm a\nu_{2k}, \ a\nu_{2k-1} \pm b\nu_{2k},$$

$$b\nu_{2k-1} + a\nu_{2k}, \ b\nu_{2k-1} \pm b\nu_{2k}, \ a\nu_{2k} \pm b\nu_{2k},$$

where k runs through the set $\{k \geq 1 : \ 2k - 1 \in \Delta_m\}$. If we write

$$R_m^{(1)}(x) = \frac{1}{2} \sum_{k \geq 1: \ 2k-1 \in \Delta_m} \cos 2\pi a\nu_{2k-1}x,$$

$$R_m^{(2)}(x) = \frac{1}{2} \sum_{k \geq 1: \ 2k-1 \in \Delta_m} \cos 2\pi b\nu_{2k-1}x,$$

$$\vdots$$

$$R_m^{(14)}(x) = \frac{1}{2} \sum_{k \geq 1: \ 2k-1 \in \Delta_m} \cos 2\pi (a\nu_{2k} + b\nu_{2k})x,$$

$$R_m^{(15)}(x) = \frac{1}{2} \sum_{k \geq 1: \ 2k-1 \in \Delta_m} \cos 2\pi (a\nu_{2k} - b\nu_{2k})x,$$

the function $R_m(x)$ can be split into 15 lacunary cosine series, each consisting of $\Delta_m/2$ elements. Observing

$$(4.29) \qquad |\alpha_m(s)| \leq \prod_{k \geq 1: \ 2k-1 \in \Delta_m} \left(1 + \frac{4s^2}{2|\Delta_m|} \right)^{1/2} \leq e^{s^2/2}$$

and, in view of (4.10),

$$|W_m(s) - 1| \leq \left| \exp\left(\frac{|\Delta_m|}{2} \left| \frac{2s}{\sqrt{2|\Delta_m|}} \right|^3 \right) - 1 \right|$$

$$\leq e^{4|\Delta_m|^{-1/2}} - 1$$

$$\leq \frac{8}{|\Delta_m|^{1/2}}.$$

We have

$$(4.30) \quad \left| \mathbb{E}e^{is\eta_m} - \mathbb{E}\exp\left(-\frac{s^2}{2}\left(1 + \frac{\cos 2\pi x}{4} \right) \right) \right|$$

$$= \left| \mathbb{E}\left(\alpha_m(s) \ \exp\left(-\frac{s^2}{2}\left(1 + \frac{\cos 2\pi c_m x}{4} \right) \right) \ \exp\left(\frac{-s^2 R_m(x)}{2|\Delta_m|} \right) \ W_m(s) \right) \right.$$

$$\left. - \mathbb{E}\exp\left(-\frac{s^2}{2}\left(1 + \frac{\cos 2\pi c_m x}{4} \right) \right) \right|$$

$$(4.31) \quad \leq \left| \mathbb{E}\left(\alpha_m(s) \ \exp\left(-\frac{s^2}{2}\left(1 + \frac{\cos 2\pi c_m x}{4} \right) \right) \right. \times \right.$$

$$(4.32) \qquad \times \left(\exp \left(\frac{-s^2 R_m(x)}{2|\Delta_m|} \right) \; W_m(s) - 1 \right) \right) \Bigg|$$

$$(4.33) \qquad + \left| \mathbb{E} \left((\alpha_m(s) - 1) \; \exp \left(-\frac{s^2}{2} \left(1 + \frac{\cos 2\pi c_m x}{4} \right) \right) \right) \right|$$

In view of (4.29), the term in lines (4.31), (4.32) is at most

$$\max_{x \in (0,1)} \left(\exp \left(-\frac{s^2 \cos 2\pi c_m x}{8} \right) \right) \mathbb{E} \left| \exp \left(\frac{-s^2 R_m(x)}{2|\Delta_m|} \right) \; W_m(s) - 1 \right|$$

$$(4.34) \qquad \leq e^{s^2/8} \; \mathbb{E} \left| \exp \left(\frac{-s^2 R_m(x)}{2|\Delta_m|} \right) \; W_m(s) - 1 \right|.$$

It is easy to see that

$$\|R_m\| \leq 15|\Delta_m|$$

and therefore

$$\mathbb{P} \left\{ |R_m| \geq 15|\Delta_m|^{2/3} \right\} \leq \Delta_m^{-1/3}.$$

Thus

$$(4.35) \qquad \mathbb{E} \left| \exp \left(\frac{-s^2 R_m(x)}{2|\Delta_m|} \right) \; W_m(s) - 1 \right|$$

$$\leq \left(1 + \frac{15s^2}{|\Delta_m|^{1/3}} \right) \left(1 + \frac{8}{|\Delta_m|^{1/2}} \right) - 1 + \frac{15s^2|\Delta_m|}{4|\Delta_m|^{4/3}}$$

$$(4.36) \qquad \ll s^2 |\Delta_m|^{-1/3},$$

where the implied constant does not depend on m, s.

The function $\alpha_m(s)$ is a sum of the constant term 1 plus at most $2^{2|\Delta_m|}$ cosine functions with coefficients at most 1 (provided $|s| \leq |\Delta_m|^{1/2}$) and frequencies at least $2^{4|\Delta_m|}$. Thus (4.33) is at most

$$2^{2|\Delta_m|} 2^{-4|\Delta_m|} \max_{x \in (0,1)} \left| \frac{d}{dx} \; \exp \left(-\frac{s^2}{2} \left(1 + \frac{\cos 2\pi c_m x}{4} \right) \right) \right|$$

$$(4.37) \qquad \leq 2^{-2|\Delta_m|} \; e^{5s^2/8} \; \frac{\pi s^2 c_m}{4}.$$

Combining (4.34), (4.35), (4.36) and (4.37) we see that (4.30) is at most

$$\ll e^{s^2/8} s^2 |\Delta_m|^{-1/3} + 2^{-2|\Delta_m|} \; e^{5s^2/8} \; \frac{\pi s^2 c_m}{4}.$$

In particular, since we assumed $|s| \leq (\log |\Delta_m|)^{1/8}$, we get that the expression (4.30) is at most

$$\ll |\Delta_m|^{-1/8},$$

proving Lemma 4.4.

Proof of Theorem 3.2. The proof of Theorem 3.2 can be obtained from Lemma 4.4 like the proof of Theorem 3.1 was obtained from Lemma 4.3 in Section 4.1. The normed partial sums

$$|\Delta_m|^{-1/2} \sum_{k \in \Delta_m} f(n_{\sigma(k)}x)$$

have a limiting distribution which is non-Gaussian, and since the set Δ_m is much larger than the union of the sets $\Delta_1, \ldots, \Delta_{m-1}$, the same distribution is the limit distribution of

$$N^{-1/2} \sum_{k=1}^{N} f(n_{\sigma(k)}x).$$

This proves Theorem 3.2.

References

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

[2] Aistleitner, C., Berkes, I., and Tichy, R., On permutations of Hardy-Littlewood-Pólya sequences, *Trans. Amer. Math. Soc.* **363** (2011), 6219–6244

[3] Aistleitner, C., Berkes, I., and Tichy, R., On the system $f(nx)$ and probabilistic number theory, To appear in *Analytic and probabilistic methods in number theory*, Proceedings of the 5th International Conference in honour of J. Kubilius held in Palanga.

[4] Aistleitner, C., Berkes, I., and Tichy, R., On the asymptotic behavior of weakly lacunary series, *Proc. Amer. Math. Soc.*, **139** (2011), 2505–2517.

[5] Amoroso, F. and Viada, E., Small points on subvarieties of a torus, *Duke Math. J.*, **150** (2009), 407–442.

[6] Berkes, I. and Philipp, W., An a.s. invariance principle for lacunary series $f(n_k x)$. *Acta Math. Acad. Sci. Hungar.*, **34**(1-2) (1979), 141–155.

[7] Erdős, P., On trigonometric sums with gaps, *Magyar Tud. Akad. Mat. Kutató Int. Közl*, **7** (1962) 37–42.

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

[9] Fukuyama, F., A central limit theorem for trigonometric series with bounded gaps. *Probab. Theory Related Fields.*, **149**(1-2) (2011) 139–148.

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

[11] Gapoškin, V. F., Lacunary series and independent functions. *Uspehi Mat. Nauk*, **21**(6 (132)), (1966), 3–82.

[12] Gapoškin, V. F., The central limit theorem for certain weakly dependent sequences. *Teor. Verojatnost. i Primenen.*, **15** (1970), 666–684.

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

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

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

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

[17] Philipp, W. and Stout, W., Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc. 2*, (issue 2, 161):iv+140, 1975.

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

[19] Salem, R. and Zygmund, A., On lacunary trigonometric series. *Proc. Nat. Acad. Sci. U. S. A.*, **33** (1947), 333–338.

[20] Schmidt, W. M., *Diophantine approximation*, volume 785 of *Lecture Notes in Mathematics*. Springer, Berlin, 1980.

[21] Takahashi, S., On the law of the iterated logarithm for lacunary trigonometric series. *Tôhoku Math. J. (2)*, **24** (1972), 319–329.

[22] Takahashi, S., On the law of the iterated logarithm for lacunary trigonometric series. II. *Tôhoku Math. J. (2)*, **27** (1975) 391–403.

[23] Zygmund, A. *Trigonometric series. Vol. I, II*. Cambridge Mathematical Library. Cambridge University Press, Cambridge, third edition, 2002.