Limit distributions in metric discrepancy theory

Christoph Aistleitner,* István Berkes†

Abstract

Let $(n_k)_{k\geq 1}$ be a lacunary sequence of integers, satisfying certain number-theoretic conditions. We determine the limit distribution of $\sqrt{N}D_N(n_kx)$ as $N\to\infty$, where $D_N(n_kx)$ denotes the discrepancy of the sequence $(n_kx)_{k\geq 1}$ mod 1.

1 Introduction and statement of results

An infinite sequence $(x_k)_{k\geq 1}$ of real numbers is called uniformly distributed mod 1 if

$$\lim_{N\to\infty}\frac{1}{N}\sum_{k=1}^{N}\mathbb{1}_{[a,b)}(x_k)=b-a \tag{1}$$

for any $0\leq a\leq b\leq 1$; here $\mathbb{1}_{[a,b)}$ denotes the indicator function of the interval $[a,b)$, extended with period 1. It is known that (1) is equivalent to the relations $D_N(x_k)\to 0$ or $D_N^*(x_k)\to 0$, where

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

and

$$D_N^*(x_k):=\sup_{0\leq a\leq 1}\left|\frac{1}{N}\sum_{k=1}^{N}\mathbb{1}_{[0,a)}(x_k)-a\right|$$

denote the discrepancy, resp. star discrepancy of the first N terms of $(x_k)_{k\geq 1}$. By a classical result of Weyl [15], for any increasing sequence $(n_k)_{k\geq 1}$ of positive integers the sequence

*Graz University of Technology, Institute of Mathematics A, 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 grant K 67961.

AMS2010 Subject Classification: 42C15, 42A55, 42A61, 30B50, 11K38, 60G50

Keywords: lacunary series, discrepancy, covariance kernel, Brownian bridge

$(n_k x)_{k\geq 1}$ is uniformly distributed mod 1 for almost all x in the sense of Lebesgue measure. Computing the order of magnitude of the discrepancy of $(n_k x)_{1\leq k\leq N}$ is a difficult problem and precise results exist only in a few cases. Philipp [12] proved that if $(n_k)_{k\geq 1}$ satisfies the Hadamard gap condition

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

then the law of the iterated logarithm (LIL)

$$\frac{1}{4\sqrt{2}} \leq \limsup_{N\to\infty} \left(\frac{N}{2\log\log N}\right)^{1/2} D_N(n_k x) \leq C \qquad \text{a.e.} \tag{3}$$

holds with some constant $C = C(q)$. Note that if $(X_k)_{k\geq 1}$ is a sequence of independent random variables in (0, 1) with $\mathbb{P}(X_k \leq x) = x \quad (0 \leq x \leq 1)$, then by the Chung-Smirnov LIL we have

$$\limsup_{N\to\infty} \left(\frac{N}{2\log\log N}\right)^{1/2} D_N(X_k) = \frac{1}{2} \tag{4}$$

with probability 1, see e.g. Shorack and Wellner [14, p. 504]. A comparison of (3) and (4) shows that the sequence $(n_k x)_{k\geq 1}$ mod 1 behaves like a sequence of i.i.d. random variables. The analogy, however, is not complete. Fukuyama [10] determined the limsup Σ_a in (3) in the case $n_k = a^k$ for $a > 1$; in particular he proved that

$$\begin{array}{lll} \Sigma_a = \sqrt{42}/9 \quad \text{a.e.} & & \text{if } a = 2, \\ \Sigma_a = \frac{\sqrt{(a+1)a(a-2)}}{2\sqrt{(a-1)^3}} & \text{a.e.} & \text{if } a \geq 4 \text{ is an even integer,} \\ \Sigma_a = \frac{\sqrt{a+1}}{2\sqrt{a-1}} \quad \text{a.e.} & & \text{if } a \geq 3 \text{ is an odd integer.} \end{array} \tag{5}$$

Thus the limsup in (3) is generally different from the value 1/2 obtained in the i.i.d. case. For further pathologies of the LIL behavior of $D_N(n_k x)$, see [2, 3, 7].

Given a sequence $(n_k)_{k\geq 1}$ of positive integers, define

$$L(N,d,\nu) = \#\{1 \leq a,b \leq d,\ 1 \leq k,\ell \leq N:\ an_k - bn_\ell = \nu\},$$

where we exclude the trivial solutions $k = \ell$ in the case $a = b,\ \nu = 0$. Aistleitner [4] proved that if $(n_k)_{k\geq 1}$ satisfies (2) and

$$L(N,d,\nu) = \mathcal{O}\left(N/(\log N)^{1+\varepsilon}\right) \qquad \text{as} \qquad N \to \infty \tag{6}$$

for all $d \geq 2, \nu \in \mathbb{Z}$ and some $\varepsilon > 0$, then we have

$$\limsup_{N\to\infty} \left(\frac{N}{2\log\log N}\right)^{1/2} D_N(n_k x) = \frac{1}{2} \qquad \text{a.e.} \tag{7}$$

Thus, under the Diophantine condition (6), the discrepancy behavior of $(n_k x)_{k\ge1}$ follows exactly the i.i.d. case. Condition (6) holds e.g. if $n_{k+1}/n_k \to \infty$ or if $n_{k+1}/n_k \to \alpha$ for some $\alpha > 1$ such that α^r is irrational for $r = 1, 2, \ldots$.

In this paper we will prove the following results.

Theorem 1. *Let $(n_k)_{k\ge1}$ be a sequence of positive integers satisfying (2) and*

$$L(N, d, \nu) = o(N) \qquad as \qquad N \to \infty \tag{8}$$

for any $d \ge 2$ and $\nu \in \mathbb{Z}$. Then

$$\sqrt{N} D_N(n_k y) \xrightarrow{\mathcal{D}} K_1, \qquad \sqrt{N} D_N^*(n_k y) \xrightarrow{\mathcal{D}} K_2$$

where K_1, K_2 are the distributions on $(0, \infty)$ with densities

$$(8/\pi)^{1/2} \sum_{k=1}^{\infty} (-1)^{k-1} k^2 e^{-k^2 y^2/2}, \qquad 8y \sum_{k=1}^{\infty} (-1)^{k-1} k e^{-2k^2 y^2},$$

respectively.

The distribution K_2 in Theorem 1 is called *Kolmogorov distribution.*

Note that Theorem 1 does not cover the case $n_k = a^k$, $a \in \mathbb{N}$, $a \ge 2$. In this case (8) holds for all $\nu \neq 0$, but not for $\nu = 0$: we have namely $n_{k+1} - a n_k = 0$ for all $k \ge 1$. Our next theorem determines the limit distribution of $\sqrt{N} D_N(n_k x)$ and $\sqrt{N} D_N^*(n_k x)$ in this case. For $0 \le t \le 1$ and $x \in \mathbb{R}$, put

$$\mathbf{I}_t(x) = \mathbb{1}_{[0,t]}(x) - t.$$

Theorem 2. *Let $a \ge 2$ be an integer. Then the series*

$$\Gamma(s, t) = \int_0^1 \mathbf{I}_s(x) \mathbf{I}_t(x)\, dx + \sum_{k=1}^{\infty} \int_0^1 \big(\mathbf{I}_s(x) \mathbf{I}_t(a^k x) + \mathbf{I}_s(a^k x) \mathbf{I}_t(x)\big)\, dx \tag{9}$$

converges absolutely on $[0, 1]^2$ and

$$\sqrt{N} D_N(a^k x) \xrightarrow{\mathcal{D}} K_\Gamma^{(1)} \qquad \sqrt{N} D_N^*(a^k x) \xrightarrow{\mathcal{D}} K_\Gamma^{(2)}$$

where $K_\Gamma^{(1)}$, $K_\Gamma^{(2)}$ denote the distribution of $\sup_{0\le x,y\le1} |G_\Gamma(x) - G_\Gamma(y)|$ and $\sup_{0\le x\le1} |G_\Gamma(x)|$, respectively, and where G_Γ is a Gaussian process over $[0, 1]$ with mean 0 and covariance function Γ.

In contrast to Theorem 1, we cannot give an explicit formula for the distribution or density function of the limit distributions.

As mentioned before, Fukuyama recently calculated the value of the limsup in Philipp's discrepancy LIL (3) for sequences of the form $n_k = a^k,\ k \geq 1$, see (5). With the notations of Theorem 2 the value Σ_a of the limsup equals

$$\sup_{0 \leq s \leq 1} \sqrt{\Gamma(s,s)} \tag{10}$$

for a.e. x, and thus Theorem 2 is the distributional analogue of Fukuyama's LIL. A graph of $\Gamma(s,s)$ resp. $\Gamma(s,t)$ is given below.

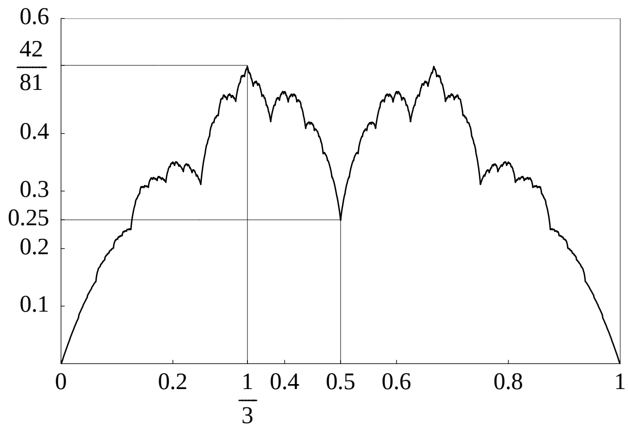

Figure 1: $\Gamma(s,s)$ for $n_k = 2^k,\ k \geq 1$. The maximum of the function is $\Gamma(1/3,1/3) = 42/81$, which leads to the value $\sqrt{42}/9$ in Fukuyama's result (5). The functions $\mathbf{I}_{[0,1/2)}(2^k x)$ are independent for $k \geq 1$ (similar to the Rademacher functions), and thus $\Gamma(1/2,1/2) = \|\mathbf{I}_{[0,1/2)}\|^2 = 1/4$.

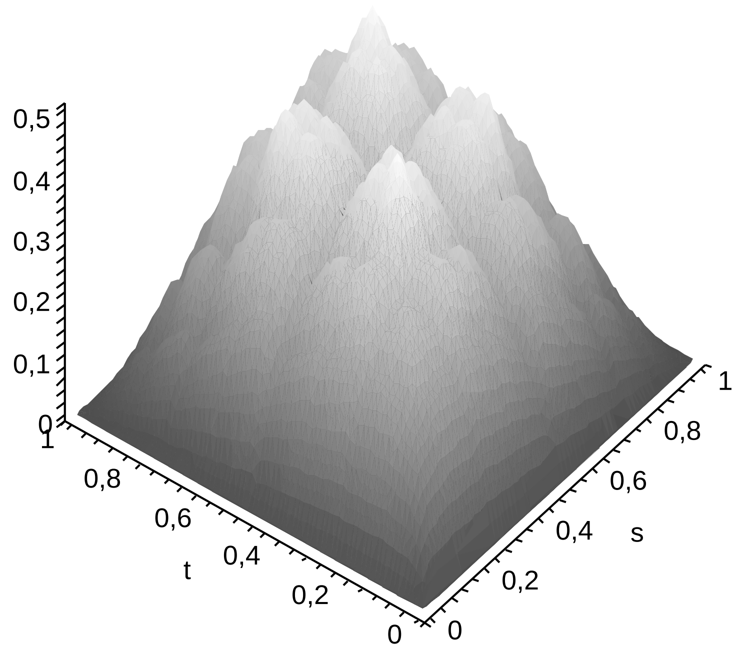

Figure 2: Covariance function $\Gamma(s,t)$ for $n_k = 2^k,\ k \geq 1$.

Let

$$D^*_{N,2}(x_k) := \left(\int_0^1 \left(\frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_{[0,t)}(x_k) - t \right)^2 dt \right)^{1/2}$$

denote the L^2 star discrepancy of the sequence $(x_k)_{k\geq 1}$. The following theorem is the analogue of Theorems 1 and 2 for $D^*_{N,2}(n_k x)$.

Theorem 3. *Under the conditions of Theorem 1 we have*

$$N(D^*_{N,2})^2(n_k x) \xrightarrow{\mathcal{D}} L$$

where L is the distribution with characteristic function

$$\prod_{k=1}^{\infty} (1 - 2it\lambda_k)^{-1/2} \tag{11}$$

and $\lambda_k = (\pi^2 k^2)^{-1}$ are the eigenvalues of the covariance kernel $s \wedge t - st$ of the Brownian bridge. The result remains valid for $n_k = a^k$, $a \geq 2$, just in this case we have to replace the numbers λ_k in (11) by the eigenvalues of the kernel $\Gamma(s,t)$ in (9).

In the case of the kernel $s \wedge t - st$, L is the limit distribution appearing in the classical Cramér-von Mises test, see e.g. Anderson and Darling [6]. Its distribution function $L(y)$ can be calculated explicitly (see [6, p. 202]):

$$L(y) = \frac{1}{\pi\sqrt{y}} \sum_{k=1}^{\infty} (-1)^k \binom{-\frac{1}{2}}{k} (4k+1)^{1/2} e^{-(4k+1)^{1/2}/(16y)} K_{1/4}((4k+1)^{1/2}/(16y)),$$

where $K_{1/4}$ is a Bessel function. We do not know a similar formula for L in the case of the kernel Γ in (9). However, the variance σ_L^2 of L equals $\int_0^1 \Gamma(s,t)^2 \, ds \, dt$ and can be computed explicitly for $n_k = a^k, \ k \geq 1$. We obtained

$$\sigma_L^2 = \frac{a(10 + a(7 + a(13 + a(7 + 4a))))}{180(a-1)^2(a+1)(a^2+1)}. \tag{12}$$

The proof uses Fourier analysis and is very laborious. It will be omitted. For $a \to \infty$ the variance σ_L^2 converges to $\mathbb{E}(\int_0^1 B^2(t)dt)^2 - (\mathbb{E}\int_0^1 B^2(t)dt)^2 = 1/45$, which is clear from the sum representation for $\Gamma(s,t)$.

Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function satisfying

$$f(x+1) = f(x), \qquad \int_0^1 f(x) \, dx = 0, \qquad \mathrm{Var}_{[0,1]} f < \infty. \tag{13}$$

In Aistleitner and Berkes [5] it is proved that under the Diophantine conditions of Theorem 1 the central limit theorem for $(f(n_k x))_{k \geq 1}$ holds. More precisely, we have the following

Theorem A. *Let f be a function satisfying (13), and let $(n_k)_{k\geq 1}$ be a sequence of positive integers satisfying (2) and (8) for any $d \geq 2$ and $\nu \in \mathbb{Z}$. Then for all $t \in \mathbb{R}$*

$$\lim_{N\to\infty} \mathbb{P} \left\{ x \in (0,1) : \ \sum_{k=1}^{N} f(n_k x) \leq t \|f\| \sqrt{N} \right\} = \Phi(t),$$

where Φ is the standard normal distribution function.

Moreover, it is shown in [5] that condition (8) is optimal for the CLT: replacing (8) by

$$L(N,d,\nu) \leq \delta N \qquad N \geq 1$$

the CLT becomes generally false. Thus condition (8) is the precise condition for the CLT for $f(n_k x)$. One can show that (8) is also optimal in Theorem 1. However, the proof is complicated and will not be given here.

A functional LIL for the empirical process of $(n_k x)_{k\geq 1}$ was proved by Philipp [13]; this enables one to get laws of the iterated logarithm for various functionals of the empirical process. Theorems 1–3 will be deduced from a functional CLT for the empirical process, which has a number of further applications. However, in the present paper we will deal only with the asymptotics of the discrepancy of $(n_k x)_{k\geq 1}$.

2 Proofs

Set

$$F_N(t) = F_N(x;t) = \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \mathbf{I}_t(n_k x), \qquad 0 \le t \le 1.$$

We show that under the conditions of Theorem 1 and Theorem 2 we have $F_N \Rightarrow B$ and $F_N \Rightarrow G_\Gamma$, respectively, where B is the Brownian bridge and $\Rightarrow$ denotes weak convergence in the Skorokhod space $D[0,1]$ (for basic facts on weak convergence on metric spaces see [8].) Since the functionals $f \to \sup_{0\le t\le 1} |f(t)|$ and $f \to \sup_{0\le s,t\le 1} |f(s) - f(t)|$ are continuous in $D[0,1]$ and the limit distributions in Theorem 1 are the same as the distributions of $\sup_{0\le s,t\le 1} |B(s) - B(t)|$ and $\sup_{0\le t\le 1} |B(t)|$, this will prove Theorems 1 and 2. Theorem 3 follows similarly, using the continuity of the functional $f \to \int_0^1 f(t)^2 dt$ on $D[0,1]$ and the fact that the limit distributions in Theorem 3 are the same as the distribution of $\int_0^1 B(t)^2 dt$ and $\int_0^1 G_\Gamma(t)^2 dt$, see [6, p. 198 and p. 202.].

Assume first the conditions of Theorem 2. We show that for any $r \ge 1$, $(c_1, \dots, c_r) \in \mathbb{R}^r$ and $0 \le t_1 < \dots < t_r \le 1$ we have

$$c_1 F_N(t_1) + \dots + c_r F_N(t_r) \xrightarrow{\mathcal{D}} c_1 K_\Gamma(t_1) + \dots + c_r K_\Gamma(t_r) \qquad \text{as} \quad N \to \infty. \tag{14}$$

By the Cramér-Wold theorem (see [8, Theorem 7.7]) this will imply the convergence of the finite dimensional distributions of F_N to those of G_Γ. Setting

$$f(x) = \sum_{m=1}^{r} c_m \mathbf{I}_{t_m}(x), \tag{15}$$

we have by a classical central limit theorem of Kac [11]

$$\frac{1}{\sqrt{N}} \sum_{k=1}^{N} f(a^k x) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \sigma_f^2),$$

where

$$\begin{aligned}
\sigma_f^2 &= \|f\|^2 + 2 \sum_{k=1}^{\infty} \int_0^1 f(x) f(a^k x)\, dx \\
&= \int_0^1 \left(\sum_{m=1}^{r} c_m \mathbf{I}_{t_m}(x) \right)^2 dx \\
&\quad + 2 \sum_{k=1}^{\infty} \int_0^1 \left(\sum_{m=1}^{r} c_m \mathbf{I}_{t_m}(x) \right) \left(\sum_{n=1}^{r} c_n \mathbf{I}_{t_n}(a^k x) \right) dx \\
&= \sum_{m=1}^{r} \int_0^1 c_m^2 \mathbf{I}_{t_m}(x)^2 dx + 2 \sum_{1\le m<n\le r} \int_0^1 c_m c_n \mathbf{I}_{t_m}(x) \mathbf{I}_{t_n}(x)\, dx
\end{aligned}$$

$$
\begin{aligned}
&+2\sum_{m=1}^{r}\sum_{k=1}^{\infty}\int_0^1 c_m^2\mathbf{I}_{t_m}(x)\mathbf{I}_{t_m}(a^kx)\\
&+2\sum_{1\le m<n\le r}\sum_{k=1}^{\infty}\int_0^1 c_mc_n\left(\mathbf{I}_{t_m}(x)\mathbf{I}_{t_n}(a^kx)+\mathbf{I}_{t_m}(a^kx)\mathbf{I}_{t_n}(x)\right)\,dx\\
=\ &\sum_{m=1}^{r}c_m^2\Gamma(t_m,t_m)+2\sum_{1\le m<n\le r}c_mc_n\Gamma(t_m,t_n).
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\mathbb{E}\left(c_1K_\Gamma(t_1)+\cdots+c_rK_\Gamma(t_r)\right)^2\\
=\ &\sum_{m=1}^{r}\mathbb{E}(c_m^2K_\Gamma(t_m)^2)+2\sum_{1\le m<n\le r}\mathbb{E}(c_mc_nK_\Gamma(t_m)K_\Gamma(t_n))\\
=\ &\sum_{m=1}^{r}c_m^2\Gamma(t_m,t_m)+2\sum_{1\le m<n\le r}c_mc_n\Gamma(t_m,t_n),
\end{aligned}
$$

proving (14). Hence by a well known criterion (see [8, p. 128]), for the weak convergence of F_N to G_Γ it suffices to prove the following

Lemma 1. *For any $(n_k)_{k\ge1}$ satisfying (2) there exists a constant c (depending only the growth factor q in (2)) such that for $N\ge1$ and $t_1,t_2,t_3\in[0,1]$, $t_1\le t_2\le t_3$,*

$$\mathbb{E}\left(|F_N(t_1)-F_N(t_2)|^3\,|F_N(t_2)-F_N(t_3)|^3\right)\le c(t_3-t_1)^2.$$

Proof. Let $Q\ge1$ be a number for which

$$q^Q>4 \tag{16}$$

(here q is the growth factor from (2)). To shorten formulas we assume that $\mathbf{I}_{t_1}-\mathbf{I}_{t_2}$ is an even function, i.e. that it can be expanded into a pure cosine-series (the proof in the general case is exactly the same). Write

$$\mathbf{I}_{t_1}(x)-\mathbf{I}_{t_2}(x)\sim\sum_{j=1}^{\infty}a_j\cos 2\pi jx$$

for the Fourier series of $\mathbf{I}_{t_1}-\mathbf{I}_{t_2}$. Then

$$\sum_{j=1}^{\infty}\frac{a_j^2}{2}=\|\mathbf{I}_{t_1}-\mathbf{I}_{t_2}\|^2\le|t_1-t_2|,$$

and, since the total variation of the function $\mathbf{I}_{t_1}-\mathbf{I}_{t_2}$ on the unit interval is at most 2, by a classical estimate from Fourier analysis (see [16, Vol. I, p.48])

$$|a_j|\le\frac{\mathrm{Var}_{[0,1]}(\mathbf{I}_{t_1}-\mathbf{I}_{t_2})}{j}\le\frac{2}{j},\qquad j\ge1. \tag{17}$$

Let $k_1, \dots, k_6$ be mutually different from each other, and assume that $k_1 \equiv k_2 \equiv k_3 \equiv k_4 \equiv k_5 \equiv k_6 \mod Q$. Let $j_1, j_2, j_3, j_4, j_5, j_6 \in [2^n, 2^{n+1})$ for some $n \geq 0$. Then by (16)

$$j_1 n_{k_1} \pm j_2 n_{k_2} \pm j_3 n_{k_3} \pm j_4 n_{k_4} \pm j_5 n_{k_5} \pm j_6 n_{k_6} \neq 0, \tag{18}$$

no matter how the signs $\pm$ are chosen. Thus by Markov's inequality and the orthogonality of the trigonometric system

$$\begin{aligned}
& \left(\mathbb{E}\left(F_N(t_1) - F_N(t_2)\right)^6\right)^{1/6} \\
= & \left(\int_0^1 \left(\frac{1}{\sqrt{N}} \sum_{k=1}^N \mathbf{I}_{t_1}(n_k x) - \mathbf{I}_{t_2}(n_k x)\right)^6 dx\right)^{1/6} \\
\leq & \frac{1}{\sqrt{N}} \sum_{m=0}^{Q-1} \left(\int_0^1 \left(\sum_{\substack{1 \leq k \leq N, \\ k \equiv m \mod Q}} \mathbf{I}_{t_1}(n_k x) - \mathbf{I}_{t_2}(n_k x)\right)^6 dx\right)^{1/6} \\
\leq & \frac{1}{\sqrt{N}} \sum_{m=0}^{Q-1} \sum_{n=0}^{\infty} \left(\int_0^1 \left(\sum_{\substack{1 \leq k \leq N, \\ k \equiv m \mod Q}} \sum_{j=2^n}^{2^{n+1}-1} a_j \cos 2\pi j n_k x\right)^6 dx\right)^{1/6} \\
= & \frac{1}{\sqrt{N}} \sum_{m=0}^{Q-1} \sum_{n=0}^{\infty} \left(\sum_{\substack{1 \leq k_1,k_2,k_3,k_4,k_5,k_6 \leq N \\ k_1,k_2,k_3,k_4,k_5,k_6 \equiv m \mod Q}} \sum_{2^n \leq j_1,j_2,j_3,j_4,j_5,j_6 < 2^{n+1}} \frac{a_{j_1} a_{j_2} a_{j_3} a_{j_4} a_{j_5} a_{j_6}}{32} \right. \\
& \left. \sum_{\pm} \mathbb{1}(j_1 n_{k_1} \pm j_2 n_{k_2} \pm j_3 n_{k_3} \pm j_4 n_{k_4} \pm j_5 n_{k_5} \pm j_6 n_{k_6} = 0) \right)^{1/6},
\end{aligned} \tag{19}$$

where the sum $\sum_{\pm}$ is meant as a sum over all possible choices of signs "+" and "−" in the indicator $\mathbb{1}(j_1 n_{k_1} \pm j_2 n_{k_2} \pm j_3 n_{k_3} \pm j_4 n_{k_4} \pm j_5 n_{k_5} \pm j_6 n_{k_6} = 0)$. Now by (18) the only solutions of $j_1 n_{k_1} \pm j_2 n_{k_2} \pm j_3 n_{k_3} \pm j_4 n_{k_4} \pm j_5 n_{k_5} \pm j_6 n_{k_6} = 0$, subject to the given restrictions of the coefficients, are of the form

$$\underbrace{j_1 n_{k_1} - j_1 n_{k_1}}_{=0} \pm \underbrace{j_2 n_{k_2} - j_2 n_{k_2}}_{=0} \pm \underbrace{j_3 n_{k_3} - j_3 n_{k_3}}_{=0}$$

(where we have $\binom{6}{2}$ possible combinations of the pairs). Thus by (17), the expression in (19) is bounded by

$$\frac{1}{\sqrt{N}} \sum_{m=0}^{Q-1} \sum_{n=0}^{\infty} \left(\underbrace{4 \binom{6}{2}}_{=60} \sum_{\substack{1 \leq k_1,k_2,k_3 \leq N \\ k_1,k_2,k_3 \equiv m \mod Q}} \frac{1}{32} \left(\sum_{2^n \leq j < 2^{n+1}} a_j^2 \right)^3 \right)^{1/6}$$

$$\leq \quad Q\sum_{n=0}^{\infty}\left(2\underbrace{\left(\sum_{2^n\le j<2^{n+1}} a_j^2\right)^2}_{\le 8(t_2-t_1)^2}\underbrace{\left(\sum_{2^n\le j<2^{n+1}}\frac{4}{j^2}\right)}_{\le 2^{-n+2}}\right)^{1/6}$$

$$\leq \quad 17Q(t_2-t_1)^{1/3}.$$

Hence

$$\mathbb{E}\left(F_N(t_1)-F_N(t_2)\right)^6 \le 17^6 Q^6 (t_2-t_1)^2. \tag{20}$$

In the same way we obtain

$$\mathbb{E}\left(F_N(t_2)-F_N(t_3)\right)^6 \le 17^6 Q^6 (t_3-t_2)^2. \tag{21}$$

By (20), (21) and Hölders inequality

$$\begin{aligned}
&\mathbb{E}\left(\left|F_N(t_1)-F_N(t_2)\right|^3\left|F_N(t_2)-F_N(t_3)\right|^3\right)\\
\le\ & \left(\mathbb{E}\left(F_N(t_1)-F_N(t_2)\right)^6\right)^{1/2}\left(\mathbb{E}\left(F_N(t_2)-F_N(t_3)\right)^6\right)^{1/2}\\
\le\ & 17^6Q^6(t_2-t_1)(t_3-t_2)\\
\le\ & 17^6Q^6(t_3-t_1)^2,
\end{aligned}$$

which proves the lemma and the relation $F_N \Rightarrow G_\Gamma$. Thus the proof of Theorem 2 is complete.

Assume now the conditions of Theorem 1. For a function f of the form (15) we have

$$\begin{aligned}
\|f\|^2 &= \int_0^1\left(\sum_{m=1}^r c_m\mathbf{I}_{t_m}(x)\right)^2 dx\\
&= \sum_{m=1}^r\int_0^1 c_m^2\mathbf{I}_{t_m}(x)\ dx + 2\sum_{1\le m<n\le r}\int_0^1 c_mc_n\mathbf{I}_{t_m}(x)\mathbf{I}_{t_n}(x)\ dx\\
&= \sum_{m=1}^r c_m^2 t_m(1-t_m) + 2\sum_{1\le m<n\le r} c_mc_nt_m(1-t_n)\\
&= : V(t_1,\dots,t_r)
\end{aligned}$$

and thus Theorem A implies

$$\frac{1}{\sqrt{N}}\sum_{k=1}^N f(n_kx) \xrightarrow{\mathcal{D}} \mathcal{N}(0,V). \tag{22}$$

On the other hand,

$$\mathbb{E}\left(c_1B(t_1)+\cdots+c_rB(t_r)\right)^2 = \sum_{m=1}^r \mathbb{E}\left(c_jB(t_j)\right) + 2\sum_{1\le m<n\le r}\mathbb{E}\left(c_mB(t_m)c_nB(t_n)\right)$$

$$= \sum_{m=1}^{r} c_j^2 t_j(1-t_j) + 2 \sum_{1\le m<n\le r} c_m c_n t_m (1-t_n) = V(t_1,\dots,t_r),$$

and hence $c_1 B(t_1) + \cdots + c_r B(t_r)$ has $\mathcal{N}(0,V)$ distribution. Thus (22) implies (14) which, together with the already proved tightness (Lemma 1), implies $F_N \Rightarrow B$. This proves Theorem 1.

Theorem 3 also follows from $F_N \Rightarrow B$ and $F_N \Rightarrow G_\Gamma$, respectively. By the Karhunen-Loève theorem the Brownian bridge can be represented in the form

$$B(t) = \sum_{k=1}^{\infty} Z_k \frac{\sqrt{2}\sin(k\pi t)}{k\pi},$$

where the Z_k's are independent random variables having $\mathcal{N}(0,1)$ distribution, and a general Gaussian process G_Γ can be represented in the form

$$G_\Gamma(t) = \sum_{k=1}^{\infty} Z_k \sqrt{\lambda_k} e_k(t),$$

where again the Z_k's are i.i.d. $\mathcal{N}(0,1)$-random variables, $(e_k(t))_{k\ge1}$ is an orthonormal system of eigenfunctions of the covariance kernel $\Gamma(s,t)$, and $\lambda_k,\ k\ge 1$ are the corresponding eigenvalues (see e.g. [1, Chapter 3.2]). Thus

$$N\left(D^*_{N,2}\right)^2 \xrightarrow{\mathcal{D}} \int_0^1 (B(t))^2 \ dt = \sum_{k=1}^{\infty} \frac{Z_k^2}{k^2\pi^2}$$

and

$$N\left(D^*_{N,2}\right)^2 \xrightarrow{\mathcal{D}} \int_0^1 (G_\Gamma(t))^2 \ dt = \sum_{k=1}^{\infty} \lambda_k Z_k^2,$$

respectively, and Theorem 3 follows from the well-known formula for the characteristic function of the chi-square distribution. By Mercer's theorem (see again [1, Chapter 3.2]) we also have the representation

$$\Gamma(s,t) = \sum_{k=1}^{\infty} \lambda_k e_k(s) e_k(t),$$

and thus the variance of

$$\sum_{k=1}^{\infty} \lambda_k Z_k^2$$

is

$$2\sum_{k=1}^{\infty} \lambda_k^2 = 2\int_0^1 \Gamma(s,t)^2 \ ds\ dt,$$

which leads to formula (12).

References

[1] R. J. Adler and J. E. Taylor. *Random fields and geometry.* Springer Monographs in Mathematics. Springer, New York, 2007.

[2] C. Aistleitner. Irregular discrepancy behavior of lacunary series. *Monatshefte Math.* 160 (2010), 1-29.

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

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

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

[6] T.W. Anderson and D.A. Darling. Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Ann. Math. Statist.* 23 (1952), 193–212.

[7] I. Berkes and W. Philipp. The size of trigonometric and Walsh series and uniform distribution mod 1, *J. London Math. Soc.* 50 (1994), 454-464.

[8] P. Billingsley. *Convergence of probability measures.* Wiley, 1968.

[9] W. Feller. The asymptotic distribution of the range of sums of independent random variables. *Ann. Math. Statist.* 22 (1951), 427–432.

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

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

[12] W. Philipp. Limit theorems for lacunary series and uniform distribution mod 1. *Acta Arith.*, 26 (1975), 241–251.

[13] W. Philipp. A functional law of the iterated logarithm for empirical distribution functions of weakly dependent random variables. *Annals of Probability*, 5 (1977), 319-350.

[14] R. Shorack and J. Wellner. *Empirical Processes with Applications to Statistics*, Wiley, 1986.

[15] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins, *Math. Ann.* 77 (1916), 313-352.

[16] A. Zygmund. *Trigonometric series. Vol. I, II.* Cambridge Mathematical Library. Cambridge University Press, Cambridge, 1988. Reprint of the 1979 edition.