On the system $f(nx)$ and probabilistic number theory

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

Dedicated to the memory of Professor Jonas Kubilius.

Abstract

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

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

The asymptotic properties of series $\sum c_k f(kx)$ have been studied extensively in the literature and turned out to be, in general, quite different from those of the trigonometric system. As the theory shows, the behavior of such series is determined by a combination of analytic, probabilistic and number theoretic effects, resulting in highly interesting phenomena not encountered in classical harmonic analysis. In this paper we survey some recent results in the field and prove asymptotic results for the system $\{f(nx), n \geq 1\}$ in the case when the function f is not square integrable.

[*]Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: aistleitner@math.tugraz.at. Research supported by FWF grant S9610-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 81928.

[‡]Graz University of Technology, Institute of Mathematics A, Steyrergasse 30, 8010 Graz, Austria. e-mail: tichy@tugraz.at. Research supported by FWF grant S9603-N23.

1 Introduction

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

$$f(x+1) = f(x), \qquad \int_0^1 f(x)\,dx = 0, \qquad \int_0^1 f^2(x)\,dx < \infty. \tag{1.1}$$

The asymptotic properties of the system $\{f(nx), n \geq 1\}$ have been studied extensively in the literature and turned out to be, in general, very different from those of the trigonometric system. Khinchin [51] conjectured that (1.1) (even without the last condition) implies

$$\lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(kx) = 0 \qquad \text{a.e.} \tag{1.2}$$

This remained open for almost 50 years until it was disproved by Marstrand [54]; he showed that there exist even bounded counterexamples f. At about the same time, Nikishin [57] constructed a continuous function f satisfying (1.1) such that $\sum c_k f(kx)$ diverges a.e. for some (c_k) with $\sum c_k^2 < \infty$. Gaposhkin [43] showed that if f satisfies (1.1) and belongs to the Lipschitz α class for some $\alpha > 1/2$, then $\sum c_k f(kx)$ converges a.e. provided $\sum c_k^2 < \infty$, i.e. the analogue of Carleson's theorem holds for the system $f(nx)$. Berkes [18] showed that this result becomes false, in general, for Lipschitz 1/2 functions. There exists no characterization of functions f for which the analogue of the Carleson convergence theorem holds for $f(nx)$ and, despite the profound work of Bourgain [24] connecting Khinchin's conjecture with metric entropy behavior, we have no characterization of functions f for which (1.2) holds.

The asymptotic properties of the system $\{f(nx), n \geq 1\}$ play also an important role in the metric theory of uniform distribution. A sequence $(x_n)_{n\geq 1}$ of real numbers is called uniformly distributed mod 1 if for any interval $[a, b)$ on the real line we have

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

where $\mathbb{1}_{[a,b)}$ denotes the indicator function of the interval $[a, b)$, extended with period 1. Given a sequence $(x_1, \ldots, x_N)$ of real numbers, the value

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

is called the discrepancy of the sequence. It is easily seen that an infinite sequence $(x_n)_{n\geq 1}$ is uniformly distributed mod 1 iff $D_N(x_1, \ldots x_N) \to 0$ as $N \to \infty$. By a classical result of Weyl [68], for any increasing sequence (n_k) of integers, $\{n_k x\}_{k\geq 1}$ is uniformly distributed mod 1 for all $x \in \mathbb{R}$, with the exception of a set having Lebesgue measure 0. Improving the results of Erdős and Koksma [32] and Cassels [25], Baker [15] proved that for any increasing sequence (n_k) of positive integers the discrepancy $D_N(\{n_k x\})$ of the first N terms of the sequence $\{n_k x\}$ satisfies

$$D_N(\{n_k x\}) \ll \frac{(\log N)^{3/2+\varepsilon}}{\sqrt{N}} \qquad \text{a.e.} \tag{1.3}$$

for any $\varepsilon > 0$. On the other hand, Berkes and Philipp [20] constructed an increasing sequence (n_k) of integers such that for almost all x the relation

$$D_N(\{n_k x\}) \geq \frac{(\log N)^{1/2}}{\sqrt{N}} \tag{1.4}$$

holds for infinitely many N. These results describe the extremal behavior of $D_N(\{n_k x\})$ rather precisely; on the other hand, the exact asymptotics of $D_N(\{n_k x\})$ is known only in a few special cases, e.g. for $n_k = k$ (Khinchin [52], Kesten [50]) and for exponentially growing n_k (Philipp [58]). Let us note that for every $x \in \mathbb{R}$

$$\frac{1}{4} \sup_{V_f \leq 2} \left| \sum_{k=1}^{N} f(n_k x) \right| \leq D_N(\{n_k x\}) \leq \sup_{V_f \leq 2} \left| \sum_{k=1}^{N} f(n_k x) \right|, \tag{1.5}$$

where V_f denotes the total variation of f on $[0,1]$. The second inequality in (1.5) is obvious from the definition of $D_N(\{n_k x\})$, while the first one follows from Koksma's inequality (see e.g. [27]), stating

$$\left| \frac{1}{N} \sum_{k=1}^{N} f(x_k) - \int_0^1 f(x)\, dx \right| \leq 2V_f \cdot D_N(x_1, \ldots x_N) \tag{1.6}$$

for any function f with $V_f < \infty$ and for every set $x_1, \ldots, x_N$ of points from the unit interval. The inequality (1.6) plays a crucial role in the theory of Monte Carlo and quasi-Monte Carlo integration. By (1.5), determining the precise order of magnitude of $D_N(\{n_k x\})$ requires sharp bounds for sums $\sum_{k=1}^{N} f(n_k x)$, which, in turn, is closely connected with estimating the integral

$$\int_0^1 \left(\sum_{k=1}^{N} f(n_k x) \right)^2 dx \tag{1.7}$$

for functions f of bounded variation. Koksma [53] showed that the integral in (1.7) is bounded by $V_f \cdot G(n_1, n_2, \ldots, n_N)$, where

$$G(n_1, n_2, \ldots, n_N) = \sum_{1 \leq i \leq j \leq N} \frac{(n_i, n_j)}{[n_i, n_j]}. \tag{1.8}$$

Here $(\cdot, \cdot)$ denotes the greatest common divisor, while $[\cdot, \cdot]$ stands for the least common multiple. Equation (1.8) shows that the order of magnitude of the discrepancy $D_N(\{n_k x\})$ depends not only on the growth speed of the sequence, but also its number theoretic properties.

Relations (1.3) and (1.5) imply that for any f with $V_f < \infty$ and any increasing sequence (n_k) of integers we have

$$\left| \sum_{k=1}^{N} f(n_k x) \right| \ll \sqrt{N} (\log N)^{3/2+\varepsilon} \qquad \text{a.e.} \tag{1.9}$$

for any $\varepsilon > 0$. On the other hand, the proof of (1.4) in Berkes and Philipp [20] provides an increasing sequence (n_k) of integers such that for $f(x) = \{x\} - 1/2$ we have for almost all x

$$\left| \sum_{k=1}^{N} f(n_k x) \right| \geq \sqrt{N} (\log N)^{1/2} \tag{1.10}$$

for infinitely many N. The gap between (1.3) and (1.4), as well as the gap between (1.9) and (1.10) remain open until today. Very recently, Aistleitner, Mayer and Ziegler [13] improved (1.9) to

$$\left| \sum_{k=1}^{N} f(n_k x) \right| \ll \sqrt{N} (\log N)^{3/2} (\log \log N)^{-1/2+\varepsilon} \qquad \text{a.e.} \tag{1.11}$$

for any $\varepsilon > 0$. Their proof uses the estimate

$$\sum_{1 \leq k \leq l \leq N} \frac{(n_k, n_l)}{\sqrt{n_k n_l}} \ll N \exp\left(\frac{c \log N}{\log \log N} \right) \tag{1.12}$$

valid for any set $\{n_1, \ldots, n_N\}$ of distinct positive integers, which is due to Dyer and Harman [30]. Dyer and Harman also conjectured that $\log N$ on the right hand side of (1.12) can be replaced by $\sqrt{\log N}$. Assuming the validity of this conjecture, the bound in (1.11) can be improved to $\sqrt{N} (\log N)^{1+\varepsilon}$.

The previous results show that there is a crucial difference between the asymptotic behavior of the system $f(nx)$ for functions f with bounded variation and for $f(x) = e^{2\pi i x}$. In the latter case Berkes and Philipp [20] proved that if $\psi(n)$ is a nondecreasing sequence satisfying $\psi(n^2) \ll \psi(n)$, then

$$\left| \sum_{k=1}^{N} e^{2\pi i n_k x} \right| \ll \sqrt{N} \psi(N) \qquad \text{a.e.} \tag{1.13}$$

holds for all increasing (n_k) if and only if

$$\sum_{k=1}^{\infty} \frac{1}{k \psi(k)^2} < \infty.$$

In particular, (1.13) holds for $\psi(N) = (\log N)^{1/2+\varepsilon}$ for $\varepsilon > 0$ and fails for $\psi(N) = (\log N)^{1/2}$

By a profound result of Gál [38], for any sequence $(n_1, \ldots, n_N)$ we have

$$G(n_1, n_2, \ldots, n_N) \ll N (\log \log N)^2$$

and this result is best possible. Since under $V_f < \infty$ the integral (1.7) is $\ll G(n_1, n_2, \ldots, n_N)$, for functions f with bounded variation the L^2 norm of $\sum_{k=1}^{N} f(n_k x)$ is $O(\sqrt{N} \log \log N)$, a bound only slightly weaker than the bound $O(\sqrt{N})$ valid for orthogonal series. Thus one can expect that under $V_f < \infty$ the convergence properties of $\sum_{k=1}^{\infty} c_k f(n_k x)$ are also not much worse than those of orthogonal series, described by the Rademacher-Mensov convergence theorem. This is indeed the case: Berkes and Weber [22] showed that for any increasing sequence (n_k) of positive integers and any function f satisfying (1.1) and $V_f < \infty$, the series $\sum_{k=1}^{\infty} c_k f(n_k x)$ converges a.e. provided

$$\sum_{k=1}^{\infty} c_k^2 (\log k)^{3+\varepsilon} < \infty$$

for some $\varepsilon > 0$. On the other hand, Nikishin [57] showed that for $f(x) = \operatorname{sgn} \sin 2\pi x$ (a function with bounded variation on $[0, 1]$) the series $\sum_{k=1}^{\infty} c_k f(kx)$ diverges on a set with positive measure for some (c_k) with $\sum_{k=1}^{\infty} c_k^2 < \infty$. These two results characterize, up to a logarithmic factor, the a.e. convergence of $\sum_{k=1}^{\infty} c_k f(n_k x)$ for functions f with bounded variation. For other classes of functions the convergence properties of the series are completely different. Recall that if f is a Lip α function with $\alpha > 1/2$ satisfying (1.1), then the analogue of Carleson's theorem holds for $\sum_{k=1}^{\infty} c_k f(kx)$, and this theorem generally fails for $\alpha = 1/2$. In the case $0 < \alpha < 1/2$ Weber [66] proved, improving results of Gaposhkin [41], that a sufficient convergence criterion is

$$\sum_{k=1}^{\infty} c_k^2 d(k) (\log k)^2 < \infty, \tag{1.14}$$

where $d(k) = \#\{1 \le i \le k : i | k\}$ is the divisor function. It is known that

$$d(k) \ll \exp(C \log k / \log \log k),$$

for some $C > 0$ and thus $\sum_{k=1}^{\infty} c_k f(kx)$ converges a.e. provided

$$\sum c_k^2 \exp(C_1 \log k / \log \log k) < \infty,$$

a fact proved independently also by Aistleitner [7] for $1/4 < \alpha < 1/2$. We note that, as Weber [66] showed, (1.14) is sufficient for the a.e. convergence of $\sum_{k=1}^{\infty} c_k f(kx)$ even if instead of the Lipschitz character of f we assume only that the Fourier coefficients of f are $O(k^{-1/2} (\log k)^{-(1+\varepsilon)})$ for some $\varepsilon > 0$, a criterion allowing a much larger class of functions f. In the case when the Fourier coefficients of f are $O(k^{-\gamma})$, $1/2 < \gamma < 1$, a sufficient convergence criterion is

$$\sum_{k=1}^{\infty} c_k^2 \rho_\gamma(k) (\log k)^2 < \infty$$

where

$$\rho_\gamma(n) = \sum_{d|n} d^{-(2\gamma-1)}.$$

See Berkes and Weber [23]. These results show that even in the case $n_k = k$ the convergence behavior of $\sum_{k=1}^{\infty} c_k f(n_k x)$ is intimately connected with number theory. For a detailed study of the convergence properties of sums $\sum_{k=1}^{\infty} c_k f(n_k x)$, see Berkes and Weber [22].

The previous results show that even for "nice" functions f, the a.e. convergence of $\sum_{k=1}^{\infty} c_k f(kx)$ is a highly delicate question, far from being solved. In contrast, convergence in L^2 norm is essentially solved by a theorem of Wintner [69], who proved that if f has the Fourier series

$$f \sim \sum_{k=1}^{\infty} (a_k \cos 2\pi kx + b_k \sin 2\pi kx),$$

then $\sum_{k=1}^{\infty} c_k f(kx)$ converges in norm for all (c_k) with $\sum_{k=1}^{\infty} c_k^2 < \infty$ iff the Dirichlet series

$$\sum_{k=1}^{\infty} a_k k^{-s}, \quad \text{and} \quad \sum_{k=1}^{\infty} b_k k^{-s} \tag{1.4}$$

are regular and bounded in the half-plane $\Re(s) > 0$. This remarkable criterion shows again the complexity of the convergence
problem studied here.

2 Lacunary series

In the previous chapter we have seen that that the convergence and growth properties of series $\sum_{k=1}^{\infty} c_k f(n_k x)$ are closely connected with the number theoretic properties of (n_k), and even for $n_k = k$ the convergence problem has an arithmetic character. In this chapter we investigate lacunary series, i.e. the behavior of $f(n_k x)$, where (n_k) satisfies the Hadamard gap condition

$$n_{k+1}/n_k \geq q > 1 \qquad (k = 1, 2, \ldots). \tag{2.1}$$

In the case $f(x) = \cos 2\pi x$ and $f(x) = \sin 2\pi x$, Salem and Zygmund [60] and Erdős and Gál [31] proved the central limit theorem and the law of the iterated logarithm, i.e.

$$\frac{1}{\sqrt{N/2}} \sum_{k=1}^{N} f(n_k x) \xrightarrow{d} N(0,1) \tag{2.2}$$

and

$$\limsup_{N\to\infty} \frac{1}{\sqrt{N \log\log N}} \sum_{k=1}^{N} f(n_k x) = 1 \qquad \text{a.e.} \tag{2.3}$$

with respect to the probability space $([0,1], \mathcal{B}, \mu)$, where $\mathcal{B}$ is the Borel σ-field in $[0,1]$ and μ is the Lebesgue measure. This shows that Hadamard lacunary subsequences of the trigonometric system behave like independent random variables. Extensions for weighted sums $\sum_{k=1}^{N} c_k f(n_k x)$ were proved by Salem and Zygmund [60, 61] and Weiss [67] under the same coefficient conditions as assumed for independent random variables. For general f satisfying (1.1), the situation is considerably more complex. Kac [48] showed that if f is a Lipschitz function or a function with bounded variation satisfying (1.1), then the CLT (2.2) holds in the case $n_k = 2^k$ with a limit distribution $N(0, \sigma^2)$, 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 law of the iterated logarithm was proved by Izumi [47] and Maruyama [55]. On the other hand, Erdős and Fortet showed (see [49], p. 646) that both the CLT and LIL fail for the system $f(n_k x)$, where

$$f(x) = \cos 2\pi x + \cos 4\pi x \quad \text{and} \quad n_k = 2^k - 1, \quad k \geq 1.$$

Specifically, in the case of the CLT a limit distribution in (2.2) still exists, but its distribution function equals

$$\pi^{-1/2} \int_0^1 \int_{-\infty}^{x/2|\cos \pi t|} e^{-u^2} du dt$$

(i.e. it is a mixture of Gaussian distributions) and the limsup in (2.3) is $\sqrt{2} \cos \pi x$, i.e. the limsup depends on x (cf. also [26]). These results show that under (2.1) the behavior of $f(n_k x)$ still resembles to that of independent random variables, but it is influenced substantially by the number theoretic properties of the sequence (n_k) as well. Gaposhkin [40] showed that the CLT for $f(n_k x)$ remains valid if all the fractions n_{k+1}/n_k are integers or if $n_{k+1}/n_k \to \alpha$, where α^r is irrational for $r = 1, 2, \ldots$. He also showed (see [44]) that the validity of the CLT is intimately connected with the number of solutions (k, l) of Diophantine equations of the form

$$a n_k \pm b n_l = c, \qquad \text{where } a, b, c \in \mathbb{Z}. \tag{2.4}$$

Improving Gaposhkin's results, Aistleitner and Berkes [8] gave a complete characterization for the CLT under (2.1). They proved, namely, that under (2.1) $f(n_k x)$ satisfies the CLT for all functions f satisfying (1.1) if and only if

$$L(N, d, \nu) = o(N) \qquad \text{as} \qquad N \to \infty \tag{2.5}$$

uniformly in $\nu \neq 0$, where

$$L(N, d, \nu) = \#\{1 \leq a, b \leq d,\ 1 \leq k, l \leq N :\ a n_k - b n_l = \nu\}, \tag{2.6}$$

where we exclude the trivial solutions $k = l$ in the case $a = b,\ \nu = 0$. Allowing also $\nu = 0$ in (2.5), the CLT will hold with norming factor $\|f\|\sqrt{N}$. A similar criterion holds for the LIL, see [5, 6].

In his classical paper, Philipp [58] investigated the law of the iterated logarithm for the discrepancy $D_N(\{n_k x\})$ under the lacunarity condition (2.1). For an i.i.d. sequence (ξ_n) of random variables, uniformly distributed on $(0, 1)$, the Chung-Smirnov law of the iterated logarithm (see e.g. [62], p. 504) states

$$\limsup_{n \to \infty} \frac{N D_N(\xi_1, \ldots \xi_N)}{\sqrt{2N \log \log N}} = \frac{1}{2} \tag{2.7}$$

with probability 1. Philipp proved that under (2.1) we have

$$\frac{1}{4\sqrt{2}} \leq \limsup_{N \to \infty} \frac{N D_N(\{n_k x\})}{\sqrt{2N \log \log N}} \leq C_q \qquad \text{a.e.} \tag{2.8}$$

where the constant C_q depends only on the growth factor q in (2.1). Again, this shows that under (2.1) the sequence $\{n_k x\}$ behaves like a sequence of independent random variables, but the value of the limsup in (2.8), as well as the question whether the limsup is a constant almost everywhere, remained open. Berkes and Philipp [20] showed that for any $q > 1$ there exists a sequence (n_k) satisfying (2.1) such that the limsup in (2.8) exceeds $c \log\log \frac{1}{q}$, showing that the limsup in (2.8) can be different from the classical value 1/2 in (2.7). Aistleitner proved that the lim sup does not have to be a constant a.e. for lacunary (n_k) (cf. [2, 3, 4]). The (nonconstant) limsup in the case $n_k = 2^k - 1$ was determined by Fukuyama [36].

Very recently Fukuyama [33] developed a powerful technique to calculate the exact value of the limsup in (2.8). In particular, he showed

$$\limsup_{N\to\infty} \frac{ND_N(\{\theta^k x\})}{\sqrt{2N\log\log N}} = C_\theta \quad \text{a.e.},$$

where the constants C_θ (which are explicitly known) depend on the precise value and the number-theoretic properties of θ in a highly interesting way. For example, we have

$$\begin{aligned} C_\theta &= \sqrt{42}/9, \quad \text{if } \theta = 2 \\ C_\theta &= \frac{\sqrt{(\theta+1)\theta(\theta-2)}}{2\sqrt{(\theta-1)^3}} \quad \text{if } \theta \geq 4 \text{ is an even integer,} \\ C_\theta &= \frac{\sqrt{\theta+1}}{2\sqrt{\theta-1}} \quad \text{if } \theta \geq 3 \text{ is an odd integer.} \end{aligned}$$

Of particular interest is the case when θ has no rational powers (which is the case e.g. if θ is transcendental), where we have $C_\theta = 1/2$, i.e. the same constant as in (2.7). Aistleitner [5] showed that the limsup is also equal 1/2 if the counting function defined in (2.6) satisfies

$$L(N, d, \nu) = O(N/(\log N)^{1+\varepsilon}) \qquad \text{as} \qquad N \to \infty \tag{2.9}$$

for some $\varepsilon > 0$, uniformly in $\nu \in \mathbb{Z}$. That is, under a condition only slightly stronger than the necessary and sufficient condition for the CLT for $f(n_k x)$, the discrepancy behavior of $\{n_k x\}$ also follows i.i.d. behavior precisely. Note, however, that although the asymptotic order of the discrepancy of $\{n_k x\}$ is very well understood for exponentially growing (n_k) from a probabilistic point of view, there exist hardly any results for the corresponding problem for concrete values of x. For example, the classical problems asking for uniform distribution of the sequences $\{2^k\sqrt{2}\}$ and $\{(3/2)^k\}$ are still completely open, and there is little hope that they can be solved within the next decades (for a discussion of these problems and recent contributions, see Bailey and Crandall [14] and Dubickas [28, 29]).

Let $\xi_1, \xi_2, \ldots$ be i.i.d. random variables, uniformly distributed on $[0, 1]$, let

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}(\xi_k \leq x)$$

denote the empirical distribution function of the sample $(\xi_1, \ldots, \xi_n)$ and let

$$T_n = \sup_{0\leq x\leq 1} \sqrt{n}|F_n(x) - x|$$

be the Kolmogorov-Smirnov statistic. T_n plays an important role in nonparametric statistics, see e.g. [62]. Using probabilistic terminology, the Chung-Smirnov LIL (2.7) can be formulated as

$$\limsup_{n\to\infty} \frac{T_n}{\sqrt{2\log\log n}} = \frac{1}{2} \qquad \text{a.s.}$$

The limit distributional behavior of T_n is described by Kolmogorov's theorem

$$\lim_{n\to\infty} P(T_n \le t) = K(t), \tag{2.10}$$

where

$$K(t) = 1 - 2\sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2t^2}. \tag{2.11}$$

It is natural to ask if an analogue of the Kolmogorov limit theorem (2.10)-(2.11) holds for discrepancies. Aistleitner and Berkes [9] showed that if (2.5) holds uniformly in $\nu \in \mathbb{Z}$ (including $\nu = 0$), then the limit distribution of $\sqrt{N} D_N(\{n_k x\})$ exists, namely we have

$$\sqrt{N} D_N(\{n_k y\}) \xrightarrow{d} K.$$

Note that the just mentioned Diophantine condition is not satisfied for $n_k = a^k$ (in this case relation (2.5) fails for $\nu = 0$), but the limit distribution of $\sqrt{N} D_N(\{a^k x\})$ still exists; the limit distribution is the same as the distribution of $\sup_{0\le t\le 1} |G(x)|$, where G is a Gaussian process with covariance function

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

where

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

As we pointed out above, assuming (2.5) uniformly in $\nu \in \mathbb{Z}$, $f(n_k x)$ satisfies the CLT and replacing $o(N)$ by $O(N/(\log N)^{1+\varepsilon})$, the LIL also holds for $f(n_k x)$. Hence under these conditions the behavior of $f(n_k x)$ follows precisely that of i.i.d. random variables. However, as Fukuyama [34] observed, the validity of the CLT and LIL can break down after a permutation of the terms of the sequence (n_k), even though an i.i.d. sequence remains i.i.d. after any permutation. We investigated this surprising phenomenon in a series of papers [10, 11, 12], and found necessary and sufficient Diophantine conditions for the permutation-invariant behavior of $f(n_k x)$.

In conclusion we note that most results discussed in this chapter break down for sublacunary sequences (i.e. sequences (n_k) satisfying $n_{k+1}/n_k \to 1$), except that the upper half of the LIL for $f(n_k x)$, i.e.

$$\limsup_{N\to\infty} \left| \frac{\sum_{k=1}^{N} f(n_k x)}{\sqrt{N \log\log N}} \right| < \infty \qquad \text{a.e.}$$

still holds for some classes of sub-lacunary sequences satisfying strong number-theoretic conditions (Philipp [59], Berkes, Philipp and Tichy [21], Fukuyama and Nakata [35], Aistleitner [1]; cf. also Furstenberg [37], who studied denseness properties of such sequences from an ergodic point of view).

The case of superlacunary sequences will be investigated in the next chapter.

3 The case $f \notin L^2$

In the previous chapter we saw that under (1.1) and the Hadamard gap condition (2.1) the asymptotic properties of partial sums $\sum_{k=1}^N f(n_k x)$ are determined by a combination of probabilistic and number theoretic effects. In particular, the behavior of the system $f(n_k x)$ is strongly influenced by the number of solutions of the Diophantine equation (2.4). Assuming

$$n_{k+1}/n_k \to \infty, \tag{3.1}$$

both Diophantine conditions (2.5) and (2.9) are satisfied and consequently $f(n_k x)$ satisfies the central limit theorem and the law of the iterated logarithm in their classical form, a result established by Takahashi [64, 65]. In other words, under the gap condition (3.1) the sequence $f(n_k x)$ behaves precisely as an i.i.d. sequence, without any number theoretic assumptions on (n_k). It is natural to ask about the asymptotic properties of lacunary sequences $f(n_k x)$ when the square integrability condition $\int_0^1 f^2(x)dx < \infty$ does not hold. Gaposhkin [42] was the first one to investigate this question; he proved the following result.

Theorem 3.1 *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function with period 1. Then there exists an increasing sequence (n_k) of positive integers and measurable functions $g_k(x)$, $\psi_k(x)$, $\eta_k(x)$, $k = 1, 2, \dots$ on $(0,1)$ such that the g_k are stochastically independent and*

$$f(n_k x) = g_k(x) + \psi_k(x) + \eta_k(x) \tag{3.2}$$

where

$$\sum_{k=1}^{\infty} \|\psi_k\|_M < \infty \quad \text{and} \quad \sum_{k=1}^{\infty} \mu\{x : \eta_k(x) \neq 0\} < \infty. \tag{3.3}$$

Here $\|\cdot\|_M$ denotes the norm in the space $M(0,1)$ of measurable functions on $(0,1)$ defined by $\|\psi\|_M = \inf\{\epsilon > 0 : \mu(x : |\psi(x)| \geq \epsilon) \leq \epsilon\}$. If $f \in L^p(0,1)$ $(p \geq 1)$ or $f \in C(0,1)$, then the conclusion remains valid with the g_k belonging to the corresponding spaces and $\|\cdot\|_M$ replaced by $\|\cdot\|_p$ or $\|\cdot\|_C$, respectively.

As an immediate consequence, we get

$$\sum_{k=1}^{\infty} |f(n_k x) - g_k(x)| < \infty \qquad \text{a.e.} \tag{3.4}$$

in all cases covered by the theorem. Relation (3.4) has powerful consequences. Most limit theorems of probability theory are invariant for small perturbations, i.e. if they are valid for some sequence (ξ_k) of random variables, then they remain valid for all sequences (ξ_k') satisfying

$$\sum_{k=1}^{\infty} |\xi_k - \xi_k'| < \infty \ \text{ a.s.}$$

Thus going beyond the limit theorems studied in the previous section, Gaposhkin's theorem extends a very large class of limit theorems of independent r.v.'s for lacunary sequences $f(n_k x)$. On the other hand, his theorem provides (except an unproved remark in the case $f \in L^p(0,1)$, $p \geq 1$) no estimate of the growth rate of (n_k) in (3.4) and in particular, it provides no explicit lacunarity condition for many interesting limit theorems such as versions of the central limit theorem with stable limits, laws of large numbers and their generalizations, extremal limit theorems, etc. The purpose of the present chapter is to obtain explicit growth rates in Gaposhkin's theorem, leading to concrete lacunarity conditions for a large class of limit theorems for $f(n_k x)$. As we will see, for many important limit theorems (including the ones mentioned before) these lacunarity conditions are actually quite close to (3.1). Our results will be deduced from the following approximation theorem.

Theorem 3.2 *Let (n_k) be an increasing sequence of positive integers. Then there exists a probability space $(\Omega, \mathcal{F}, P)$ and two sequences (X_k) and (Y_k) of random variables with the following properties.*

(a) The sequence $(X_k)_{k\geq 1}$ is a probabilistic replica of the sequence $\{n_k x)\}_{k\geq 1}$ in the sense that the distribution of the two sequences in the spaces $(\Omega, \mathcal{F}, P)$ and $((0,1), \mathcal{B}, \mu)$ are the same.

(b) $(Y_k)_{k\geq 1}$ is an i.i.d. sequence with uniform distribution over $[0,1]$.

(c) We have

$$P(|X_k - Y_k| \geq \delta_k) \leq \delta_k \qquad k = 1, 2, \dots \tag{3.5}$$

where $\delta_1 = 1$ and

$$\delta_k = 5\,(n_{k-1}/n_k + n_k/n_{k+1}) \qquad k = 2, 3, \dots . \tag{3.6}$$

By the identical distribution of the sequences $(X_k)_{k\geq 1}$ and $(f(n_k x))_{k\geq 1}$ the asymptotic properties of the two sequences are the same. Using standard probabilistic language, we can say that we "redefined" the sequence $\{n_k x\}_{k\geq 1}$ (without changing its distribution) on the probability space $(\Omega, \mathcal{F}, P)$ together with an i.i.d. uniform sequence $(Y_k)_{k\geq 1}$ such that relation (3.5) holds with the δ_k in (3.6). In other words, the sequence $\{n_k x\}_{k\geq 1}$ is, after a suitable redefinition, a small perturbation of an i.i.d. uniform sequence. This fact implies that if $\sum_{k=1}^\infty \delta_k < \infty$ (or equivalently if $\sum_{k=1}^\infty n_k/n_{k+1} < \infty$), then most limit theorems valid for the i.i.d. sequence $(Y_k)_{k\geq 1}$ will be valid for $\{n_k x\}_{k\geq 1}$ as well. Related, weaker approximation theorems were obtained in Hawkes [45] and in Berkes [17] dealing with the trigonometric case.

Note that Theorem 3.2 concerns the specific sequence $\{n_k x\}$, but depending on the properties of the function f, it leads automatically to a corresponding approximation theorem for general sequences $f(n_k x)$. For example, if f is continuous with continuity modulus $\omega(f, \delta)$, then (3.5) and (3.6) imply

$$\sum_{k=1}^\infty |f(X_k) - f(Y_k)| < \infty \qquad \text{a.e.} \tag{3.7}$$

provided

$$\sum_{k=1}^\infty n_k/n_{k+1} < \infty, \qquad \sum_{k=1}^\infty \omega(f, n_k/n_{k+1}) < \infty.$$

A much more general consequence of Theorem 3.2 is the following

Theorem 3.3 *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function with period 1 and let (n_k) be an increasing sequence of positive integers. Let (T_k) be positive numbers such that $\mu\{x \in (0,1) : |f(x)| \geq T_k\} \leq k^{-2}$ and assume that*

$$\sum_{k=1}^{\infty} \left(T_k \delta_k^{1/4} + \omega_2^{1/2}(f_{T_k}, 8\delta_k^{1/2}) \right) < \infty. \tag{3.8}$$

Then on a suitable probability space there exists a probabilistic replica $(X_k)_{k\geq 1}$ of $(f(n_k x))_{k\geq 1}$ together with an i.i.d. sequence $(Y_k)_{k\geq 1}$ such that Y_k are distributed as $f(x)$ on $((0,1), \mathcal{B}, \mu)$ and

$$\sum_{k=1}^{\infty} |X_k - Y_k| < \infty \qquad a.s.$$

Here δ_k is defined by (3.6),

$$\omega_2(f, \delta) = \left(\sup_{0\leq h\leq \delta} \int_0^1 |f(x+h) - f(x-h)|^2 \, dx \right)^{1/2}$$

is the L^2 modulus of continuity of f and f_T is the truncated function $f \cdot \mathbb{1}\{|f| \leq T\}$. There exist several classical limit theorems for $f(n_k x)$ involving this modulus of continuity $\omega_2(f, h)$. For example, Ibragimov [46] proved the CLT for $f(2^k x)$ under (1.1) and the assumption

$$\sum_{k=1}^{\infty} \omega_2(f, 2^{-k}) < \infty.$$

Takahashi [64] proved the CLT for $f(n_k x)$ under (1.1), $n_{k+1}/n_k \to \infty$ and

$$\omega_2(f, h) = \mathcal{O}\left(\log \frac{1}{h} \right)^{-\alpha}$$

for some $\alpha > 1$ and Matsuyama and Takahashi [56] proved the corresponding LIL under similar, slightly stronger assumptions. Gaposhkin [40, 41] proved that under (1.1) and

$$\sum_{k=1}^{\infty} \omega_2^2(f, n_k/n_{k+1}) < \infty \tag{3.9}$$

the sum $\sum_{k=1}^{\infty} c_k f(n_k x)$ is a.e. convergent provided $\sum_{k=1}^{\infty} c_k^2 < \infty$ and also that $f(n_k x)$ satisfies the LIL, provided (3.9) holds with ω_2 replaced by the ordinary modulus of continuity ω.

Given any periodic measurable function f, we can choose T_k so that $\mu\{|f| \geq T_k\} \leq k^{-2}$ for all $k \geq 1$ and then condition (3.8) is satisfied if δ_k tends to 0 sufficiently rapidly or, equivalently, if (n_k) grows sufficiently rapidly. More importantly, however, Theorem 3.3 enables one to give a concrete gap condition implying the validity of i.i.d. limit theorems for lacunary sequences $f(n_k x)$. We illustrate the procedure on two classical limit theorems for i.i.d. random variables.

Corollary 3.1 *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function with period 1 such that the distribution function*

$$F(x) = \mu\{t \in (0,1) : f(t) \le x\} \tag{3.10}$$

of f satisfies

$$1 - F(x) \sim px^{-\alpha}L(x), \qquad F(-x) \sim qx^{-\alpha}L(x) \qquad \text{as } x \to \infty \tag{3.11}$$

for some constants $p, q \ge 0$, $p + q = 1$, $0 < \alpha < 2$ and a slowly varying function L. Let (n_k) be an increasing sequence of positive integers satisfying (3.8). Then letting $S_n = \sum_{k=1}^n f(n_k x)$ we have

$$(S_n - a_n)/b_n \xrightarrow{d} G \tag{3.12}$$

for some numerical sequences $(a_n), (b_n)$ and an α-stable distribution G.

Corollary 3.2 *Let $f : \mathbb{R} \to \mathbb{R}$ be a measurable function with period 1 such that the distribution function F in (3.10) satisfies $F(x) < 1$ for all x and $1 - F$ is regularly varying at $+\infty$ with a negative exponent. Let (n_k) be an increasing sequence of positive integers satisfying (3.8). Then letting $M_n = \max_{1 \le k \le n} f(n_k x)$, we have*

$$(M_n - a_n)/b_n \xrightarrow{d} G \tag{3.13}$$

where $G(x) = \exp(-x^{-\alpha})\mathbb{1}_{(0,\infty)}(x)$.

Note that (3.11) is the classical necessary and sufficient condition for the partial sums S_n of an i.i.d. sequence with distribution function F to satisfy the limit theorem (3.12) with suitable norming and centering sequences a_n, b_n. Corollary 3.1 shows that if the distribution function F of the periodic function f satisfies (3.11), then the partial sums of $f(n_k x)$ for any (n_k) satisfying (3.8) obey the limit theorem (3.12). Similarly, the assumption on F in Corollary 3.2 is the well-known necessary and sufficient condition for the centered and normed maxima of an i.i.d. sequence with distribution F to converge weakly to the distribution $G(x) = \exp(-x^{-\alpha})\mathbb{1}_{(0,\infty)}(x)$, the so called Fréchet distribution. As we know (see e.g. [39]), the limit distribution in (3.13) for any i.i.d. sequence can be only one of the Fréchet, Weibull and Gumbel distributions with respective distribution functions $\exp(-x^{-\alpha})\mathbb{1}_{(0,\infty)}(x)$, its analogue on the negative axis and $\exp(-e^{-x})$; the analogue of Corollary 3.2 holds for the other two limiting classes, too.

The growth speed of (n_k) in (3.8) depends on f; clearly, "nice" functions f require less rapidly growing (n_k). For example, in Corollary 3.1 condition (3.11) with $L(x) = 1$ and $p = q = 1/2$ can be realized with the function

$$f(x) = \begin{cases} -|x - 1/2|^{-1/\alpha} & \text{if } \ 0 < x < 1/2 \\ |x - 1/2|^{-1/\alpha} & \text{if } \ 1/2 < x < 1. \end{cases} \tag{3.14}$$

Then we can choose $T_k = k^{1/\alpha}$ and a bound for $|f'|$ on the set $\{|f| \le T_k\}$ is $Ck^{2(1+\alpha)/\alpha}$ and thus

$$\omega_2(f_{T_k}, \delta) \le \omega(f_{T_k}, \delta) \le Ck^{2(1+\alpha)/\alpha}\delta.$$

Thus a simple calculation shows that (3.8) is satisfied if

$$n_{k+1}/n_k \geq k^\gamma \tag{3.15}$$

for some $\gamma = \gamma(\alpha) > 0$. Note that (3.15) is only slightly stronger than (3.1): relation (3.1) requires that n_k grows faster than exponential, while (3.15) is satisfied for $n_k \sim e^{Ck \log k}$ for a sufficiently large C. There are, of course, many other choices of the function f leading to the same distribution function F in (3.10), which lead in general to faster growing (n_k).

Besides covering a large class of limit theorems, Theorem 3.2 leads also to permutation-invariant results. As we have seen, under (1.1) and suitable smoothness conditions, $f(2^k x)$ satisfies the central limit theorem and the law of the iterated logarithm, but, as Fukuyama [34] showed, these results are not permutation-invariant: both the CLT and LIL break down after a suitable permutation of the terms of the sequence $f(2^k x)$. In contrast, relation (3.4) is clearly permutation-invariant and so are its consequences discussed above.

4 Proofs

Theorem 3.2 will be proved by using the strong approximation technique developed in Berkes and Philipp [19]. More precisely, the result will be deduced from the following extension of Theorem 2 in [19].

Lemma 4.1 *Let $(\Omega, \mathcal{F}, P)$ be a probability space, $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots \subset \mathcal{F}$ a sequence of σ-fields and $X_1, X_2, \ldots$ a sequence of discrete random variables such that X_k is $\mathcal{F}_k$-measurable and*

$$P(\pi(\text{dist}(X_k|\mathcal{F}_{k-1}), \text{dist}(X_k)) \geq \gamma_k) \leq \gamma_k \tag{4.1}$$

for some γ_k, $k = 1, 2, \ldots$. Assume that on $(\Omega, \mathcal{F}, P)$ there exists a random variable Z, independent of $\sigma\{\mathcal{F}_1, \mathcal{F}_2, \ldots\}$ and uniformly distributed over $[0, 1]$. Then on $(\Omega, \mathcal{F}, P)$ there exist independent random variables $Y_1, Y_2, \ldots$ such that $X_k \stackrel{d}{=} Y_k$ $(k = 1, 2, \ldots)$ and

$$P(|X_k - Y_k| \geq \gamma_k) \leq 2\gamma_k \qquad (k = 1, 2, \ldots).$$

Here $\stackrel{d}{=}$ denotes equality in distribution, $\text{dist}(X_k)$ and $\text{dist}(X_k|\mathcal{F}_{k-1})$ denote, respectively, the distribution of X_k and its conditional distribution relative to $\mathcal{F}_{k-1}$, and $\pi(P_1, P_2)$ denotes the Prohorov distance of the probability measures P_1 and P_2 defined by

$$\pi(P_1, P_2) = \inf\big\{\varepsilon > 0 : P_1(A) \leq P_2(A^\varepsilon) + \varepsilon \text{ and } P_2(A) \leq P_1(A^\varepsilon) + \varepsilon \text{ for all Borel sets } A \subset \mathbb{R}\big\}$$

where A^ε is the open ε-neighborhood of A, i.e.,

$$A^\varepsilon = \big\{x \in \mathbb{R} : |x - y| < \varepsilon \text{ for some } y \in A\big\}.$$

Proof of Lemma 4.1. We start with recalling some well known facts from probability theory. Given an atomless probability space $(\Omega, \mathcal{F}, P)$ and a distribution function F, there always exists on $(\Omega, \mathcal{F}, P)$ a r.v. X with distribution F. As an immediate consequence, if X is a discrete r.v. on an atomless space $(\Omega, \mathcal{F}, P)$ with distribution function F and G is a two-dimensional distribution function with first marginal F (i.e. $G(x, +\infty) = F(x)$), then on $(\Omega, \mathcal{F}, P)$ there exists a r.v. Y such that the distribution of the vector (X, Y) is G.

Trivially, if X and Y are r.v.'s defined on the same probability space satisfying $P(|X - Y| \geq \varepsilon) \leq \varepsilon$, then the Prohorov distance of the distribution of X and Y is $\leq \varepsilon$. By a theorem of Strassen [63], the converse is also true: if P_1 and P_2 are probability measures on the real line with $\pi(P_1, P_2) \leq \varepsilon$, then on any atomless probability space $(\Omega, \mathcal{F}, P)$ there exist r.v.'s X and Y with distributions P_1 and P_2 such that $P(|X - Y| \geq \varepsilon) \leq \varepsilon$. Combining this with the previous remarks it follows that if the distribution P_1 is discrete and $\pi(P_1, P_2) \leq \varepsilon$, then one can even prescribe a r.v. X on $(\Omega, \mathcal{F}, P)$ with distribution P_1 and there still exists a r.v. Y on $(\Omega, \mathcal{F}, P)$ with distribution P_2 such that $P(|X - Y| \geq \varepsilon) \leq \varepsilon$.

Turning to the proof of Lemma 4.1, we will construct the r.v.'s $Y_1, Y_2, \ldots$ by induction. We first enlarge the σ-fields $\mathcal{F}_k$ by setting $\mathcal{F}_k^* = \sigma\{\mathcal{F}_k, Z\}$, where Z is the random variable in the formulation of the lemma. Clearly $\operatorname{dist}(X_k | \mathcal{F}_{k-1}^*) = \operatorname{dist}(X_k | \mathcal{F}_{k-1})$ and thus (4.1) implies

$$P(\pi\left(\operatorname{dist}(X_k | \mathcal{F}_{k-1}^*), \operatorname{dist}(X_k)\right) \geq \gamma_k) \leq \gamma_k \qquad (k = 1, 2, \ldots). \tag{4.2}$$

Let $Y_1 = X_1$ and assume that $Y_1, Y_2, \ldots, Y_{k-1}$ are already constructed and satisfy the statements of the lemma, moreover, Y_j is $\mathcal{F}_j^*$ measurable for $1 \leq j \leq k - 1$. Since $X_j \stackrel{d}{=} Y_j$ for $1 \leq j \leq k - 1$, the r.v.'s $Y_1, Y_2, \ldots, Y_{k-1}$ are discrete. Letting $\mathcal{F}_{k-1}^{**} = \sigma\{Y_1, \ldots, Y_{k-1}\}$, clearly $\mathcal{F}_{k-1}^{**} \subset \mathcal{F}_{k-1}^*$ and thus (4.2) implies

$$P(\pi\left(\operatorname{dist}(X_k | \mathcal{F}_{k-1}^{**}), \operatorname{dist}(X_k)\right) \geq \gamma_k) \leq \gamma_k \qquad (k = 1, 2, \ldots). \tag{4.3}$$

Consider the sets D of the form

$$D = \{Y_1 = b_1, \ldots, Y_{k-1} = b_{k-1}\} \tag{4.4}$$

where $b_1, \ldots, b_{k-1}$ are in the range of $Y_1, \ldots, Y_{k-1}$, respectively. The union of these sets is clearly Ω; we will construct Y_k on each such set separately. We clearly have $\operatorname{dist}(X_k | \mathcal{F}_{k-1}^{**}) = \operatorname{dist}(X_k | D)$ on the set D in (4.4), and thus (4.3) implies that the sets D can be distributed into two classes Γ_1 and Γ_2 such that $\sum_{D \in \Gamma_2} P(D) \leq \gamma_k$ and for any $D \in \Gamma_1$ we have

$$\pi(\operatorname{dist}(X_k | D), \operatorname{dist}(X_k)) \leq \gamma_k. \tag{4.5}$$

Let first $D \in \Gamma_1$, let $P^{(D)}$ denote conditional probability with respect to D and define the probability measures P_1 and P_2 on the Borel sets of $\mathbb{R}$ by

$$P_1(A) = P(X_k \in A | D), \quad P_2(A) = P(X_k \in A), \qquad A \in \mathcal{B}.$$

By (4.5) we have $\pi(P_1, P_2) \leq \gamma_k$ and since the probability space $(D, \mathcal{F}_k^*, P^{(D)})$ is atomless, the remarks at the beginning of the proof imply that on the probability space $(D, \mathcal{F}_k^*, P^{(D)})$ there exists a random variable Y_k with distribution $\text{dist}(X_k)$ such that $P^{(D)}(|X_k - Y_k| \geq \gamma_k) \leq \gamma_k$. This defines the random variable on each set $D \in \Gamma_1$; for a set $D \in \Gamma_2$ let Y_k be any r.v. on the probability space $(D, \mathcal{F}_k^*, P^{(D)})$ with distribution $\text{dist}(X_k)$. Thus we defined Y_k on the whole probability space Ω; clearly, the so defined Y_k is $\mathcal{F}_k^*$ measurable and has the property that its conditional distribution on any set of the form (4.4) equals its unconditional distribution and thus Y_k is independent of the vector $(Y_1, \ldots, Y_{k-1})$. Also,

$$P(|X_k - Y_k| \geq \gamma_k) \leq 2\gamma_k.$$

This completes the induction step and thus the proof of Lemma 4.1.

Proof of Theorem 3.2. Let $\varepsilon_k = n_k/n_{k+1}$ and $\psi(x) = \{x\}$. Set $A_k = \{i/n_{k+1} : 1 \leq i \leq n_{k+1}\}$, $B_k = \bigcup_{j=0}^k A_j$ and let $J_1, J_2, \ldots, J_{m_k}$ be the left closed intervals to which the points of B_k divide the interval $[0,1)$. Let $\mathcal{F}_k$ be the σ-field generated by the intervals $J_1, \ldots, J_{m_k}$ and set

$$T_k = T_k(x) = \psi(n_k x),$$
$$X_k = E(T_k | \mathcal{F}_k).$$

Clearly $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots$ and X_k is an $\mathcal{F}_k$-measurable discrete r.v. taking the constant value $\mu(A)^{-1} \int_A \psi(n_k x)\, dx$ on any atom A of $\mathcal{F}_k$. Since the atoms of $\mathcal{F}_k$ are intervals of length $\leq 1/n_{k+1}$, we have

$$|T_k - X_k| \leq n_k/n_{k+1} = \varepsilon_k\,. \tag{4.6}$$

Let $\mathcal{G}_k$ denote the set of intervals $I = \big[i/n_{k+1}, (i+1)/n_{k+1}\big)$, $0 \leq i \leq n_{k+1} - 1$ which contain in their interior no point of B_{k-1}. Clearly each interval $I \in \mathcal{G}_k$ is an atom of $\mathcal{F}_k$ and

$$\sum_{I \in \mathcal{G}_k} \mu(I) \geq 1 - 2\varepsilon_k \qquad k \geq k_0. \tag{4.7}$$

To see the last inequality observe that the number of those intervals

$$I = \big[i/n_{k+1}, (i+1)/n_{k+1}\big),\ 0 \leq i \leq n_{k+1} - 1,$$

which contain in their interior a point of B_{k-1} is at most $\text{card}\, B_{k-1} \leq n_1 + \cdots + n_k$ and thus the total measure of these intervals is at most $(n_1 + \cdots + n_k)/n_{k+1}$. Now (3.1) implies $n_1 + \cdots + n_k \leq 2n_k$ for $k \geq k_0$, proving (4.7).

Let H denote the uniform distribution over $(0,1)$. Clearly the conditional distribution of $T_k = \psi(n_k x)$ relative to each interval $I \in \mathcal{G}_{k-1}$ is H and thus by (4.6) we have

$$\pi\left(\text{dist}(X_k|I), H\right) = \pi\left(\text{dist}(X_k|I), \text{dist}(T_k|I)\right) \leq \varepsilon_k \qquad \text{for } I \in \mathcal{G}_{k-1}.$$

In view of (4.7), the last relation means

$$\pi\left(\text{dist}(X_k|\mathcal{F}_{k-1}), H\right) \leq \varepsilon_k \quad \text{with probability } \geq 1 - 2\varepsilon_{k-1}. \tag{4.8}$$

Since $\operatorname{dist}(X_k)$ is obtained from $\operatorname{dist}(X_k|\mathcal{F}_{k-1})$ by integration, (4.8) easily implies

$$\pi\left(\operatorname{dist}(X_k), H\right) \le \varepsilon_k + 2\varepsilon_{k-1},$$

which, together with (4.8), yields

$$\pi\left(\operatorname{dist}(X_k|\mathcal{F}_{k-1}), \operatorname{dist}(X_k)\right) \le 2(\varepsilon_k + \varepsilon_{k-1}) \quad \text{with probability } \ge 1 - 2(\varepsilon_k + \varepsilon_{k-1}).$$

We thus showed that condition (4.1) of Lemma 4.1 holds with $\delta_k = 2(\varepsilon_k + \varepsilon_{k-1})$. Hence if on the probability space $((0,1), \mathcal{B}, \mu)$ there exists a uniformly distributed random variable Z independent of $\sigma\{\mathcal{F}_1, \mathcal{F}_2, \dots,\}$, then Lemma 4.1 applies and yields a sequence (Y_k) of independent random variables on this space such that $Y_k \stackrel{d}{=} X_k$ and

$$\mu\big(|X_k - Y_k| \ge \delta_k\big) \le 2\delta_k \qquad k \ge k_0.$$

By (4.6) this implies

$$\mu\big(|\{n_k x\} - Y_k| \ge \delta_k'\big) \le \delta_k' \qquad k \ge k_0$$

with $\delta_k' = 5(\varepsilon_k + \varepsilon_{k-1})$. If a random variable Z with the desired properties does not exist, replace $((0,1), \mathcal{B}, \mu)$ by the product space $((0,1), \mathcal{B}, \mu) \times ((0,1), \mathcal{B}, \mu)$ and redefine all random variables and σ-fields on the new space in an obvious fashion. In the new space the required Z obviously exists and all arguments of our proof remain valid in the new space.

To complete the proof of Theorem 3.2 it suffices now to show that there exists an i.i.d. sequence (Z_k) with $Z_k \stackrel{d}{=} \psi$ such that $|Y_k - Z_k| \le \varepsilon_k$ for $k = 1, 2, \dots$. By passing to a suitable product space as before, we can assume without loss of generality that there exist independent random variables $\eta_1, \eta_2, \dots$, having uniform distribution over $(0,1)$ and independent also of $X_1, Y_1, X_2, Y_2, \dots$; let $\mathcal{H}_k = \sigma\{Y_k, \eta_k\}$. By the remarks made at the beginning of the proof, on the atomless probability space $(\Omega, \mathcal{H}_k, P)$ there exists a random variable Z_k such that the joint distribution of Y_k and Z_k is the same as the joint distribution of Y_k and ψ in $((0,1), \mathcal{B}, \mu)$. Clearly the Z_k are independent, $Z_k \stackrel{d}{=} \psi$ and by (4.6) we have $|Z_k - Y_k| \le \varepsilon_k$. This completes the proof of Theorem 3.2.

Proof of Theorem 3.3. We first consider the case when f satisfies the additional condition $\int_0^1 f^2(x)\,dx < \infty$; let

$$f \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos 2\pi kx + b_k \sin 2\pi kx)$$

be its Fourier series. Then for any $n \ge 1$ we have

$$\|f - s_n(f)\| \le \omega_2(f, \pi/n)$$

where $s_n(f)$ denotes the nth partial sum of the Fourier series of f (see e.g. [16], Vol. II, p. 156). Write $f = f_1 + f_2$, where

$$f_1(x) = s_m(f, x) = \frac{a_0}{2} + \sum_{k=1}^{m} (a_k \cos 2\pi kx + b_k \sin 2\pi kx),$$

and $f_2 = f - s_m$ with m to be determined later. By Theorem 3.2 there exists a probability space $(\Omega, \mathcal{F}, P)$ and sequences (X_k^*) and (Y_k^*) of random variables such that (X_k^*) is a probabilistic replica of the sequence $\{n_k x\}$, Y_k^* are i.i.d. random variables with uniform distribution over $(0,1)$ and

$$P(|X_k^* - Y_k^*| \geq \delta_k) \leq \delta_k \qquad k = 1, 2, \ldots, \tag{4.9}$$

with δ_k defined by (3.6). Then

$$|f(X_k^*) - f(Y_k^*)| \leq |f_1(X_k^*) - f_1(Y_k^*)| + |f_2(X_k^*) - f_2(Y_k^*)| =: V_1 + V_2.$$

Clearly

$$\begin{aligned} |f_1'(x)| &\leq 2\pi \sum_{k=1}^m k(|a_k| + |b_k|) \\ &\leq 2\pi \left[\left(\sum_{k=1}^m a_k^2 \right)^{1/2} + \left(\sum_{k=1}^m b_k^2 \right)^{1/2} \right] \left(\sum_{k=1}^m k^2 \right)^{1/2} \leq 4\pi \|f\| m^{3/2} \end{aligned}$$

and thus (4.9) and the mean value theorem imply that $|V_1| \leq 4\pi \|f\| m^{3/2} \delta_k$ with probability $\geq 1 - \delta_k$. On the other hand, X_k^* and Y_k^* are uniformly distributed r.v.'s over $(0,1)$ and thus $\|f_2(X_k^*)\| = \|f_2(Y_k^*)\| = \|f_2\|$. Hence we have

$$\|V_2\| \leq 2\|f_2\| \leq 2\omega_2(f, \pi/m)$$

and thus the Markov inequality yields that $|V_2| \leq \omega_2^{1/2}(f, \pi/m)$ with the exception of a set with measure not exceeding $4\omega_2(f, \pi/m)$. We thus proved that

$$|f(X_k^*) - f(Y_k^*)| \leq 4\pi \|f\| m^{3/2} \delta_k + \omega_2^{1/2}(f, \pi/m) \tag{4.10}$$

with probability exceeding $1 - (\delta_k + 4\omega_2(f, \pi/m))$. In the case of a general periodic measurable f, apply the previous argument for the truncated function $f_{T_k} = f \cdot \mathbb{1}\{|f| \leq T_k\}$, with the choice $m = [\delta_k^{-1/2}]$. Clearly, $\|f_{T_k}\| \leq T_k$, and thus using (4.10) for f_{T_k} and applying the Borel-Cantelli lemma, Theorem 3.3 follows with $X_k = f(X_k^*)$, $Y_k = f(Y_k^*)$.

References

[1] C. Aistleitner. Diophantine equations and the LIL for the discrepancy of sublacunary sequences. Illinois J. Math. 53 (2010), 785–815.

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

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

[4] C. Aistleitner. On the class of limits of lacunary trigonometric series. Acta Math. Hungar. 129 (2010), 1–23.

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

[6] C. Aistleitner. On the law of the iterated logarithm for the discrepancy of lacunary sequences II. Trans. Amer. Math. Soc., to appear.

[7] C. Aistleitner. Convergence of $\sum c_k f(kx)$ and the Lip α class. Proc. Amer. Math. Soc., to appear.

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

[9] C. Aistleitner and I. Berkes. Limit distributions in metric discrepancy theory. Monatshefte Math., to appear.

[10] C. Aistleitner, I. Berkes and R. F. Tichy. On the law of the iterated logarithm for permuted lacunary sequences. Proc. Steklov Inst. Math. 276 (2012), 3–20.

[11] C. Aistleitner, I. Berkes and R. F. Tichy. On permutations of lacunary series. RIMS Kôkyûroku Bessatsu, Kyoto University, to appear.

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

[13] C. Aistleitner, P. Mayer and V. Ziegler. Metric discrepancy theory, functions of bounded variation and GCD sums. Unif. Distrib. Theory 5 (2010), 95–109.

[14] D.H. Bailey and R.E. Crandall. On the random character of fundamental constant expansions. Experiment. Math. 10 (2001), no. 2, 175–190.

[15] R. C. Baker. Metric number theory and the large sieve. J. London Math. Soc. 24 (1981), 34-40.

[16] N. K. Bari. A treatise on trigonometric series, I-II. Pergamon Press, New York 1964.

[17] I. Berkes. On almost i.i.d. subsequences of the trigonometric system. In: Functional analysis, Univ. of Texas, Austin, 1986–87. Lecture Notes in Math., Vol. 1332, pp. 54–63. Springer, 1988.

[18] I. Berkes. On the convergence of $\sum c_n f(nx)$ and the Lip 1/2 class. Trans. Amer. Math. Soc. 349 (1997), 4143–4158.

[19] I. Berkes and W. Philipp. Approximation theorems for independent and weakly dependent random vectors. Ann. Probab. 7 (1979), 29–54.

[20] 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.

[21] I. Berkes, W. Philipp, and R. F. Tichy. Empirical processes in probabilistic number theory: the LIL for the discrepancy of $(n_k \omega)$ mod 1. Illinois J. Math. 50 (2006), 107–145.

[22] I. Berkes and M. Weber. On the convergence of $\sum c_k f(n_k x)$. Memoirs of the AMS 201 (2009), No. 943.

[23] I. Berkes and M. Weber. On series of dilated functions. Preprint.

[24] J. Bourgain. Almost sure convergence and bounded entropy. Israel J. Math. 63 (1988), 79–97.

[25] J. W. S. Cassels. Some metrical theorems in Diophantine approximation III. Proc. Cambridge Philos. Soc. 46 (1950), 219-225.

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

[27] M. Drmota and R. F. Tichy. Sequences, discrepancies and applications. Lecture Notes in Math., Vol. 1651. Springer, Berlin, 1997.

[28] A. Dubickas. On the powers of 3/2 and other rational numbers. Math. Nachr. 281 (2008), no. 7, 951–958.

[29] A. Dubickas. Powers of rational numbers modulo 1 lying in short intervals. Results Math. 57 (2010), no. 1-2, 23–31.

[30] T. Dyer and G. Harman. Sums involving common divisors. J. London Math. Soc. 34 (1986), 1–11.

[31] P. Erdős and I. S. Gál. On the law of the iterated logarithm. Proc. Kon. Nederl. Akad. Wetensch. 58 (1955), 65-84.

[32] P. Erdős and J. Koksma. On the uniform distribution modulo 1 of sequences $(f(n,\theta))$. Proc. Kon. Nederl. Akad. Wetensch. 52 (1949), 851-854.

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

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

[35] K. Fukuyama and K. Nakata. A metric discrepancy result for the Hardy-Littlewood-Pólya sequences. Monatsh. Math. 160 (2010), 41–49.

[36] K. Fukuyama and S. Miyamoto. Metric discrepancy results for Erdős-Fortet sequence. Studia Sci. Math. Hung. 49 (2012), 52–78.

[37] H. Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation. Math. Systems Theory 1 (1967), 1–49.

[38] I. S. Gál. A theorem concerning Diophantine approximations. Nieuw Arch. Wiskunde 23 (1949), 13–38.

[39] J. Galambos. The asymptotic theory of extreme order statistics. Robert E. Krieger Publishing Co., Melbourne, FL, second edition, 1987.

[40] V. F. Gaposhkin. Lacunary series and independent functions. Uspehi Mat. Nauk, 21/6 (1966), 3–82.

[41] V. F. Gaposhkin. On the convergence of series of the form $\sum a_n \phi(nx)$. Mat. Sbornik 74 (1967), 93–99.

[42] V. F. Gaposkin. Certain systems of almost independent functions. Sibirsk. Mat. J. 9 (1968), 264–279.

[43] V. F. Gaposhkin. Convergence and divergence systems. Mat. Zametki 4 (1968), 253–260.

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

[45] J. Hawkes. Probabilistic behaviour of some lacunary series. Z. Wahrsch. Verw. Gebiete 53 (1980), 21–33.

[46] I. A. Ibragimov. On asymptotic distribution of values of certain sums. Vestnik Leningrad. Univ. 15 (1960), 55–69.

[47] S. Izumi. Notes on Fourier analysis XIV. On the law of the iterated logarithm of some series of functions. J. Math. Tokyo 1 (1951), 1–22.

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

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

[50] H. Kesten. The discrepancy of random sequences $\{kx\}$. Acta Arith. 10 (1964/1965), 183–213.

[51] A. Khinchin. Ein Satz über Kettenbrüche mit arithmetischen Anwendungen. Math. Zeitschrift 18 (1923), 289–306.

[52] A. Khinchin. Einige Sätze über Kettenbrüche mit Anwendungen auf die Theorie der Diophantischen Approximationen. Math. Ann. 92 (1924), 115–125.

[53] J. Koksma. On a certain integral in the theory of uniform distribution. Indagationes Math. 13 (1951), 285–287.

[54] J. M. Marstrand. On Khinchin's conjecture about strong uniform distribution. Proc. London Math. Soc. 21 (1970), 540–556.

[55] G. Maruyama. On an asymptotic propoerty of a gap sequence. Kodai Math. Sem. Rep. 2 (1950), 31–32.

[56] N. Matsuyama and S. Takahashi. The law of the iterated logarithms. Sci. Rep. Kanazawa Univ. 7 (1961), 35–39.

[57] E. M. Nikishin. Resonance theorems and superlinear operators. Russian Math. Surveys 25/6 (1970), 125-187.

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

[59] W. Philipp. Empirical distribution functions and strong approximation theorems for dependent random variables. A problem of Baker in probabilistic number theory. Trans. Amer. Math. Soc. 345 (1994), 705–727.

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

[61] R. Salem and A. Zygmund. La loi du logarithme itéré pour les séries trigonométriques lacunaires. Bull. Sci. Math. 74 (1950), 209–224.

[62] G. R. Shorack and J. A. Wellner. Empirical processes with applications to statistics. Wiley, New York, 1986.

[63] V. Strassen. The existence of probability measures with given marginals. Ann. Math. Statist. 36 (1965), 423–439.

[64] S. Takahashi. A gap sequence with gaps bigger than the Hadamards. Tôhoku Math. J. 13 (1961), 105–111.

[65] S. Takahashi. The law of the iterated logarithm for a gap sequence with infinite gaps. Tôhoku Math. J. 15 (1963), 281–288.

[66] M. Weber. On systems of dilated functions. Comptes Rendus Mathématiques, to appear.

[67] M. Weiss. The law of the iterated logarithm for lacunary trigonometric series. Trans. Amer. Math. Soc. 91 (1959), 444–469.

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

[69] A. Wintner. Diophantine approximations and Hilbert's space. Amer. J. Math. 66 (1944), 564–578.