LACUNARY SEQUENCES IN ANALYSIS, PROBABILITY AND NUMBER THEORY

CHRISTOPH AISTLEITNER, ISTVÁN BERKES AND ROBERT TICHY

ABSTRACT. In this paper we present the theory of lacunary trigonometric sums and lacunary sums of dilated functions, from the origins of the subject up to recent developments. We describe the connections with mathematical topics such as equidistribution and discrepancy, metric number theory, normality, pseudorandomness, Diophantine equations, and the subsequence principle. In the final section of the paper we prove new results which provide necessary and sufficient conditions for the central limit theorem for subsequences, in the spirit of Nikishin's resonance theorem for convergence systems. More precisely, we characterize those sequences of random variables which allow to extract a subsequence satisfying a strong form of the central limit theorem.

CONTENTS

1. INTRODUCTION

The word "lacunary" has its origin in the Latin *lacuna* (ditch, gap), which is a diminutive form of *lacus* (lake). Accordingly, a lacunary sequence is a sequence with gaps, and a lacunary trigonometric sum is a sum of trigonometric functions with gaps between the frequencies of consecutive summands. The origin of the theory of lacunary sums might lie in Weierstrass' famous example of a continuous, nowhere differentiable function (1872). Since then the subject has evolved into many very different directions, reflecting for example the emergence of modern measure theory and axiomatic probability theory in the early twentieth century, profound

developments in harmonic analysis and Diophantine approximation, the establishment of ergodic theory as one of the key instruments of number theory, or the interest in notions of pseudo-randomness which are associated with the evolution of theoretical computer science. Throughout this paper we will be concerned with convergence/divergence properties of infinite trigonometric series

$$\sum_{k=1}^{\infty} c_k \cos(2\pi n_k x) \qquad \text{or} \qquad \sum_{k=1}^{\infty} c_k \sin(2\pi n_k x),$$

as well as with the asymptotic order and the distributional behavior of finite trigonometric sums

$$\sum_{k=1}^{N} c_k \cos(2\pi n_k x) \qquad \text{or} \qquad \sum_{k=1}^{N} c_k \sin(2\pi n_k x)$$

(the latter often in the simple case where $c_k \equiv 1$), and with their generalizations

$$\sum_{k=1}^{\infty} c_k f(n_k x) \qquad \text{and} \qquad \sum_{k=1}^{N} c_k f(n_k x).$$

Here $(c_k)_{k\geq 1}$ is a sequence of coefficients, and $(n_k)_{k\geq 1}$ is a sequence of positive integers (typically increasing), which satisfies some gap property such as the classical Hadamard gap condition

$$\frac{n_{k+1}}{n_k} > q > 1, \qquad k \geq 1,$$

or the "large gap condition" (also called "super-lacunarity property")

$$\frac{n_{k+1}}{n_k} \to \infty, \qquad k \to \infty.$$

Furthermore, f is a 1-periodic function which is usually assumed to satisfy some regularity properties (such as being of bounded variation, being Lipschitz-continuous, etc.), and which for simplicity is usually assumed to be centered such that $\int_0^1 f(x)\, dx = 0$.

Early appearances of such lacunary sums include the following.

- Sums of the form $\sum_{k=1}^{N} f(2^k x)$, where f is an indicator function of a dyadic sub-interval of $[0, 1]$, extended periodically with period 1. Borel used such sums in 1909 to show that almost all reals are "normal"; more on this topic is contained in Section 6 below.

- Uniform distribution of sequences $(\{n_k x\})_{k\geq 1}$ in Weyl's seminar paper of 1916; more on this in Section 2. Here and throughout the paper, we write $\{\cdot\}$ for the fractional part function.

- Kolmogorov's theorem on the almost everywhere convergence of lacunary trigonometric series if the sequence of coefficients is square-summable (1924),

a result related to his Three Series Theorem for the almost sure convergence of series of independent random variables. Later it turned out that Kolmogorov's convergence theorem for trigonometric series actually remains true without any gap condition whatsoever, a result which was widely believed to be "too good to be true" before being established by Carleson [77] in 1966. More on this in Section 2.

- Foundational work on the distribution of normalized lacunary trigonometric sums, in particular the central limit theorems of Kac (1946) and Salem and Zygmund (1947), and the laws of the iterated logarithms of Salem and Zygmund (1950) and of Erdős and Gál (1955). More on this in Sections 4 and 5.

A fundamental observation is that the unit interval, equipped with Borel sets and Lebesgue measure, forms a probability space, and that consequently a sequence of functions such as $(\cos(2\pi n_k x))_{k \geq 1}$ or $(f(n_k x))_{k \geq 1}$ can be viewed as a sequence of random variables over this space; if f is 1-periodic and if $(n_k)_{k \geq 1}$ is a sequence of positive integers then these random variables are identically distributed, but in general they are not independent. However, under appropriate circumstances the gap condition which is imposed upon $(n_k)_{k \geq 1}$ can ensure that these random variables have a low degree of stochastic dependence. Consequently lacunary sums often mimic the behavior of sums of independent and identically distributed random variables. This viewpoint was in particular taken by Steinhaus, Kac, and Salem and Zygmund in their fundamental work on the subject. In a particularly striking situation, the dyadic functions considered by Borel actually turn out to be a version of a sequence of Bernoulli random variables which are truly stochastically independent; accordingly, Borel's result on the normality of almost all reals is nowadays usually read as the historically very first version of the strong law of large numbers in probability theory.

When taking this probabilistic viewpoint, the theory of lacunary sums could be seen as a particular segment of the much wider field of the theory of weakly dependent random systems in probability theory, which is associated with notions such as mixing, martingales, and short-range dependence. However, it should be noted that the precise dependence structure in a lacunary function system $(f(n_k x))_{k \geq 1}$ is controlled by the analytic properties of the function f, in conjunction with arithmetic properties of the sequence $(n_k)_{k \geq 1}$. It is precisely this interplay between probabilistic, analytic and arithmetic aspects which makes the theory of lacunary sums so interesting, so challenging and so rewarding. In the following sections we want to illustrate some instances of these phenomena in more detail.

2. Uniform distribution and discrepancy

Let $(x_n)_{n\geq 1}$ be a sequence of real numbers in the unit interval. Such a sequence is called uniformly distributed modulo one (in short: u.d. mod 1) if

$$(1) \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_A(x_n) = \lambda(A)$$

for all sub-intervals $A \subset [0,1]$ of the unit interval. The word "equidistributed" is also used for this property, synonymously with "uniformly distributed modulo one". In this definition, and in the sequel, $\mathbb{1}$ denotes an indicator function, and λ denotes Lebesgue measure. In informal language, this definition means that a sequence is u.d. mod 1 if every interval A asymptotically receives its fair share of elements of the sequence, which is proportional to the length of the interval. Note that (for example as a consequence of the Glivenko–Cantelli theorem) for a sequence of independent, uniformly $(0,1)$-distributed random variables $(U_n)_{n\geq 1}$ one has

$$\frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_A(U_n) = \lambda(A) \qquad \text{almost surely}$$

for all intervals $A \subset [0,1]$, so that in a vague sense uniform distribution of a deterministic sequence can be interpreted in the sense that the sequence shows "random" behavior; more on this aspect in Section 6 below. Uniform distribution theory can be said to originate with Kronecker's approximation theorem and with work of Bohl, Sierpiński and Weyl on the sequence $(\{n\alpha\})_{n\geq 1}$ for irrational α. However, the theory only came into its own with Hermann Weyl's [222] seminal paper of 1916. Among many other fundamental insights, Weyl realized that Definition (1), which in earlier work had only be read in terms of counting points in certain intervals, can be interpreted in a "functional" way and can equivalently be written as

$$(2) \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(x_n) = \int_0^1 f(x) \; dx$$

for all continuous functions f. This viewpoint suggests that uniformly distributed sequences can be used as quadrature points for numerical integration; in the multi-dimensional setting and together with quantitative error estimates this observation forms the foundation of the so-called Quasi-Monte Carlo integration method, a concept which today forms a cornerstone of numerical methods in quantitative finance and other fields of applied mathematics (more on this below). Furthermore, Weyl realized that the indicator functions in (1) or the continuous functions in (2) could also be replaced by complex exponentials, as a consequence of the Weierstrass approximation theorem; thus by the famous Weyl Criterion a sequence is u.d. mod 1 if and only if

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} e^{2\pi i h x_n} = 0$$

for all fixed non-zero integers h, thereby tightly connecting uniform distribution theory with the theory of exponential sums.

For the particular sequence $(\{n\alpha\})_{n\geq 1}$ it can be easily seen from the Weyl criterion that this sequence is u.d. mod 1 if and only if $\alpha \notin \mathbb{Q}$. However, for other parametric sequences of the form $(\{n_k\alpha\})_{k\geq 1}$ the situation is much more difficult, and in general it is completely impossible to determine whether for some particular value of α the sequence is u.d. or not. It turns out that in a *metric* sense the situation is quite different. Metric number theory arose after the clarification of the concept of real numbers, the realization that the reals drastically outnumber the integers and the rationals, and the development of modern measure theory. Loosely speaking, the purpose of metric number theory is to determine properties which hold for a set of reals which is "typical" with respect to a certain measure; here "typical" means that the measure of the complement is small. In the present paper the measure under consideration will always be the Lebesgue measure, and a set of reals will be considered typical if its complement has vanishing Lebesgue measure; however, metric number theory has for example also been intensively studied with respect to the Hausdorff dimension or other fractal measures.

Returning to Weyl's results, what he proved in the metric setting is the following. For every sequence of distinct integers $(n_k)_{k\geq 1}$, the sequence $(\{n_k\alpha\})_{k\geq 1}$ is u.d. mod 1 for (Lebesgue-) almost all reals α. In other words, even if we cannot specify the set of α's for which uniform distribution holds, at least we know that the set of such α's has full Lebesgue measure. It is amusing that after formulating the result, Weyl continues to write:

> *Wenn ich nun freilich glaube, daß man den Wert solcher Sätze, in denen eine unbestimmte Ausnahmemenge vom Maße 0 auftritt, nicht eben hoch einschätzen darf, möchte ich diese Behauptung hier doch kurz begründen.*[1]

One should recall that Weyl's paper was written in a time of intense conflict of formalists vs. constructivists (with Weyl favoring the latter ones), and only very briefly after the notion of a set of zero (Lebesgue) measure had been introduced at all. Today, Weyl's theorem is seen as one of the foundational results of metric number theory, together with the work of Borel, Koksma, Khinchin and others.

While uniform distribution modulo one is a qualitative asymptotic property, it is natural that one is also interested in having a corresponding quantitative concept which applies to finite sequences (or finite truncations of infinite sequences). Such

---

[1]Even if I think that the value of theorems, which contain an unspecified exceptional set of measure zero, is not particularly high, I still want to give a short justification.

a concept is the discrepancy of a sequence, which is defined by

$$D_N(x_1, \ldots, x_N) = \sup_{A \subset [0,1]} \left| \frac{1}{N} \sum_{n=1}^{N} \mathbb{1}_A(x_n) - \lambda(A) \right|.$$

Here the supremum is taken over all sub-intervals $A \subset [0,1]$, and it is easy to see that an infinite sequence $(x_n)_{n \geq 1}$ is u.d. mod 1 if and only if the discrepancy $D_N(x_1, \ldots, x_N)$ tends to 0 as $N \to \infty$. With a slight abuse of notation, we will write throughout the paper $D_N(x_n) = D_N(x_1, \ldots, x_N)$ for the discrepancy of the first N elements of an infinite sequence $(x_n)_{n \geq 1}$. From a probabilistic perspective, the discrepancy is a variant of the (two-sided) Kolmogorov–Smirnov statistic, where one tests the empirical distribution of the point set $x_1, \ldots, x_N$ against the uniform distribution on $[0,1]$. Without going into details, we note that $D_N(x_1, \ldots, x_N)$ can be bounded above in terms of exponential sums by the Erdős–Turán inequality, and that the error when using $x_1, \ldots, x_N$ as a set of quadrature points to approximate $\int_0^1 f(x)\, dx$ by $\frac{1}{N} \sum_{n=1}^{N} f(x_n)$ can be bounded above by Koksma's inequality in terms of the variation of f and the discrepancy D_N; for details see the monographs [96, 161], which contain all the basic information on uniform distribution theory and discrepancy. See also [181] for a discussion of equidistribution and discrepancy from the viewpoint of analytic number theory, and [164, 165, 192] for expositions which put particular emphasis on the numerical analysis aspects.

Weyl's metric result from above can be written as

$$\lim_{N \to \infty} D_N(\{n_k \alpha\}) = 0 \qquad \text{for almost all } \alpha,$$

for any sequence $(n_k)_{k \geq 1}$ of distinct itegers. Strikingly, the precise answer to the corresponding quantitative problem is still open more than a hundred years later. It is known that for every strictly increasing sequence of integers $(n_k)_{k \geq 1}$ one has

$$(3) \qquad D_N(\{n_k \alpha\}) = O\left(\frac{(\log N)^{3/2+\varepsilon}}{\sqrt{N}} \right) \qquad \text{for almost all } \alpha.$$

This is a result of R.C. Baker [38], who improved earlier results of Cassels [78] and of Erdős and Koksma [104] by using Carleson's celebrated convergence theorem in the form of the Carleson–Hunt inequality [140]. In his paper Baker wrote that

[...] probably the exponent $3/2 + \varepsilon$ could be replaced by ε [...]

but it turned out that this is not actually the case. Instead, Berkes and Philipp [64] constructed an example of an increasing integer sequence $(n_k)_{k \geq 1}$ for which

$$(4) \qquad \limsup_{N \to \infty} \frac{\left| \sum_{k=1}^{N} \cos(2\pi n_k x) \right|}{\sqrt{N \log N}} = +\infty \qquad \text{for almost every } x.$$

By the Erdős–Turán inequality this gives a corresponding lower bound for the discrepancy, which implies that the optimal exponent of the logarithmic term in an upper bound of the form (3) has to be at least $1/2$. But the actual size of this optimal exponent, one of the most fundamental problems in metric discrepancy theory,

still remains open. Note that for pure cosine-sums $\sum_{k=1}^{N} \cos(2\pi n_k x)$ it is easily seen that one has a metric upper bound with exponent $1/2 + \varepsilon$ in the logarithmic term; this follows from the orthogonality of the trigonometric system, together with Carleson's inequality and the Chebyshev inequality. Thus, in connection with (4), the optimal upper bound in a metric estimate for pure cosine sums is known. For sums $\sum_{k=1}^{N} f(n_k x)$ with f being a 1-periodic function of bounded variation, the optimal exponent also is $1/2 + \varepsilon$, but this is a much deeper result than the one for the pure cosine case, and was established only recently in [15, 169]. By Koksma's inequality, an upper bound for the discrepancy implies an upper bound for sums of function values for a (fixed) function of bounded variation, but the opposite is not true. So while the case of a fixed function f is solved and is an important test case for the discrepancy, the problem of the discrepancy itself (which requires a supremum over a whole class of test functions) is more involved and remains open.

3. ARITHMETIC EFFECTS: DIOPHANTINE EQUATIONS AND SUMS OF COMMON DIVISORS

One of the most classical tools of probability theory is the calculation of expectations, variances, and higher moments of sums of random variables. Due to trigonometric identities such as

$$(5) \qquad \cos a \cos b = \frac{\cos(a + b) + \cos(a - b)}{2},$$

the calculation of moments of sums of trigonometric functions (with integer frequencies) reduces to a counting of solutions of certain Diophantine equations. Indeed, while the first and second moments

$$\int_0^1 \sum_{k=1}^{N} \cos(2\pi n_k x)\, dx = 0 \qquad \text{and} \qquad \int_0^1 \left(\sum_{k=1}^{N} \cos(2\pi n_k x) \right)^2 dx = \frac{N}{2}$$

are trivial and do not depend on the particular sequence $(n_k)_{k \geq 1}$ (as long as the elements of the sequence are assumed to be distinct), interesting arithmetic effects come into play when one has to compute higher moments, and it can be clearly seen how the presence of a gap condition leads to a behavior of the moments which is similar to that of sums of independent random variables. More precisely, assume that we try to calculate

$$\int_0^1 \left(\sum_{k=1}^{N} \cos(2\pi n_k x) \right)^m dx$$

for some integer $m \geq 3$. By (5) this can be written as a sum

$$2^{-m} \sum_{\pm} \sum_{1 \leq k_1, \ldots, k_m \leq N} \mathbb{1}\left(\pm n_{k_1} \pm \cdots \pm n_{k_m} = 0 \right).$$

Here the first sum is meant as a sum over all positive combinations of "+" and "-" signs inside the indicator function at the end. Now assume that, for simplicity, we consider the particular sequence $n_k = 2^k$, $k \geq 1$, which is a prototypical example

of a sequence satisfying the Hadamard gap condition. Then the majority of the solutions of

$$\pm n_{k_1} \pm \cdots \pm n_{k_m} = 0$$

arise from cancellation in a pairwise way; that is, after a suitable re-ordering of the indices, one has $k_1 = k_2, \ k_3 = k_4, \ldots, \ k_{m-1} = k_m$, and suitable opposing signs. There are further solutions, coming for example from the fact that $n_{k+1} - n_k - n_k = 0$ for all k, but one can show that they play an overall insignificant role. Thus one is led to calculating the combinatorial quantity which stems from the solutions with pairwise cancellation, which turns out to be exactly the same combinatorial quantity that arises when calculating an m-th moment of a sum of independent random variables. Thus the moments of the trigonometric lacunary sum converge to those of a suitable Gaussian distribution, which gives rise to the classical limit theorems for lacunary trigonometric sums. The situation is more delicate if one only has the Hadamard gap condition $n_{k+1}/n_k > q > 1$ rather than exact exponential growth, and again more delicate if one considers a sum of dilated functions $\sum f(n_k x)$ instead of a pure trigonometric sum, but the principle described here is very powerful also in these more general situations. For a long time this was the key ingredient in most of the proofs of limit theorems for lacunary sums; see for example [103, 145, 193, 201, 214, 221]. A different method is based on the approximation of a lacunary sum by a martingale difference; here the "almost independent" behavior is not captured by controlling the moments of the sum, but in the fact that later terms of the sum (functions with high frequency) oscillate quickly in small regions where earlier summands (functions with much lower frequency) are essentially constant. As far as we can say, this method was first used in the context of lacunary sums by Berkes [53] and, independently, by Philipp and Stout [196]. We will come back to this topic in Section 4.

Broadly speaking, the "almost independent" behavior of sums of dilated functions breaks down when the lacunarity condition is relaxed. Many papers have been devoted to this effect; see in particular [57, 59, 102, 184]. In order to maintain the "almost independent" behavior of the sum, there are two natural routes to take. On the one hand, one could randomize the construction of the sequence $(n_k)_{k \geq 1}$, and assume that the undesired effects disappear almost surely with respect to the underlying probability measure – it turns out that this is a very powerful method, and we will come back to it in Section 7 below. On the other hand, when adapting the viewpoint that the "almost independence" property is expressed in the small number of solutions of certain Diophantine equations, one could try to compensate the weaker growth assumption by stronger arithmetic assumptions. A prominent example of a class of sequences for which the latter approach has been very successfully used are the so-called Hardy–Littlewood–Pólya sequences, which consist of all the elements of the multiplicative semigroup generated by a finite set of primes, sorted in increasing order. These sequences are in several ways a natural analogue of lacunary sequences; note that the sequence $(2^k)_{k \geq 1}$ actually also falls into this framework by consisting of all elements of the semigroup generated by a single prime. Such sequences generated by a finite set of primes have attracted the attention of number theorists again

and again, a particularly interesting instance being Fürstenberg's [126] paper on disjointness in ergodic theory. It is known that Hardy–Littlewood–Pólya sequences (if generated by two or more primes) grow sub-exponentially, and the precise (only slightly sub-exponential) growth rates are known (Tijdeman [216]). What is more striking (and a much deeper fact) is that also the number of solutions of the relevant linear Diophantine equations can be bounded efficiently – this is Schmidt's celebrated Subspace Theorem [207] in a quantitative form such as that of Evertse, Schlickewei and Schmidt [107] or Amoroso and Viada [30]. By a combination of the (slightly weaker) growth condition with the (strong) arithmetic information available for Hardy–Littlewood–Pólya sequences, much of the machinery that is used for Hadamard lacunary sequences can be rescued for this generalized setup; see [194] as well as [19, 65, 123].

We briefly come back to the case of sums of dilated functions $\sum f(n_k x)$ without the presence of a growth condition on $(n_k)_{k \geq 1}$. We assume for simplicity that $\int_0^1 f(x) \, dx = 0$, so trivially

$$\int_0^1 \sum_{k=1}^N f(n_k x) \, dx = 0,$$

but already the calculation of the variance

(6)
$$\int_0^1 \left(\sum_{k=1}^N f(n_k x) \right)^2 dx$$

is in general quite non-trivial. If $f(x) = \cos(2\pi x)$, then one can simply use the orthogonality of the trigonometric system. If f is a more general function, then one can still express f by its Fourier series, expand the square and integrate, and thus translate the problem of calculating (6) into a problem of counting the solutions of certain linear Diophantine equations. When carrying out this approach, one is naturally led to the problem of estimating a certain sum involving greatest common divisors. For example, assume that $f(x) = \{x\} - 1/2$. In this case a classical formula (first stated by Franel and first proved by Landau) asserts that

$$\int_0^1 f(mx)f(nx) \, dx = \frac{1}{12} \frac{(\gcd(m,n))^2}{mn},$$

and consequently

$$\int_0^1 \left(\sum_{k=1}^N (\{n_k x\} - 1/2) \right)^2 dx = \frac{1}{12} \sum_{1 \leq k, \ell \leq N} \frac{(\gcd(n_k, n_\ell))^2}{n_k n_\ell}.$$

The sum on the right-hand side of this equation is called a GCD sum. A similar identity holds for example for the Hurwitz zeta function $\zeta(1 - \alpha, \cdot)$, where

$$\int_0^1 \zeta(1 - \alpha, \{mx\})\zeta(1 - \alpha, \{nx\}) \, dx = 2\Gamma(\alpha)^2 \frac{\zeta(2\alpha)}{(2\pi)^{2\alpha}} \frac{(\gcd(m,n))^{2\alpha}}{(mn)^\alpha}$$

for $\alpha > 1/2$, thus leading to a GCD sum

$$\sum_{1 \leq k, \ell \leq N} \frac{(\gcd(n_k, n_\ell))^{2\alpha}}{(n_k n_\ell)^\alpha}$$

with parameter α. If $f(x)$ is a general 1-periodic function, then one usually does not obtain such a nice exact representation of the variance of a sum of dilated function values, but typically the variance (6) can be bounded above by a GCD sum, which together with Chebyshev's inequality and the Borel–Cantelli lemma allows to make a statement on the almost everywhere asymptotic behavior of a sum of dilated function values.

This connection between sums of dilated functions and GCD sums is explained in great detail in Chapter 3 in Harman's monograph on *Metric Number Theory* [136], where mainly the context of metric Diophantine approximation is treated (see also [127, 155]). Recently this connection has also led to a solution of the problem of the almost everywhere convergence of series of dilated functions. Recall that Carleson's theorem [77] asserts that the series

$$\sum_{k=1}^{\infty} c_k \cos(2\pi n_k x)$$

is almost everywhere convergent provided that $\sum_k c_k^2 < \infty$. It is natural to ask which assumption on the sequence of coefficients $(c_k)_{k \geq 1}$ is necessary to ensure the almost everywhere convergence of the more general series

$$(7) \qquad \sum_{k=1}^{\infty} c_k f(n_k x),$$

under some regularity assumptions on f. Gaposhkin [129, 130] obtained some partial results, but a satisfactory understanding of the problem was only achieved very recently, when the connection with GCD sums was fully understood and optimal upper bounds for such sums were obtained. Exploiting this connection with GCD sums, it was shown in [15, 169] that for 1-periodic f which is of bounded variation on $[0, 1]$ the series (7) is almost everywhere convergent provided that

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

for some $\gamma > 2$, and this result is optimal in the sense that the same assumption with $\gamma = 2$ would not be sufficient. In [16] it was shown that for the class C_α of 1-periodic square integrable functions f with Fourier coefficients a_j, b_j satisfying

$$a_j = O(j^{-\alpha}), \quad b_j = O(j^{-\alpha})$$

for $1/2 < \alpha < 1$, a sharp criterion for the almost everywhere convergence of (7) is that

$$(8) \qquad \sum_{k=1}^{\infty} c_k^2 \exp\left(\frac{K (\log k)^{1-\alpha}}{(\log \log k)^\alpha} \right) < \infty$$

with a suitable $K = K(\alpha)$. In the case of 1-periodic Lipschitz α functions f, Gaposhkin [130] proved that for $\alpha > 1/2$, the series (7) converges a.e. under $\sum_k c_k^2 < \infty$ (just like in the case of Carleson's theorem) and Berkes [60] showed that this result is sharp, i.e. for $\alpha = 1/2$ the exact analogue of Carleson's theorem is not valid. No sharp convergence criteria exists in the case $0 < \alpha \le 1/2$; for sufficient criteria for see Gaposhkin [129]. See also [4, 68, 128] for a general discussion and several further results for the convergence of series $\sum_k c_k f(n_k x)$.

For general periodic $f \in L^2$ the direct connection between the integral (6) and GCD sums breaks down, but upper bounds for (6) as well as for

$$(9) \qquad \int_0^1 \left(\sum_{k=1}^N c_k f(n_k x) \right)^2 dx$$

can be given in terms of the coefficients c_k, of the Fourier coefficients of f, and arithmetic functions such as $d(n) = \sum_{d|n} 1$, $\sigma_s(n) = \sum_{d|n} d^s$, or the Erdős-Hooley function $\Delta(n) = \sup_{u \in \mathbb{R}} \sum_{d|n, u \le d \le eu} 1$. See Koksma [156, 157], Weber [220], and Berkes and Weber [69, 70]. A typical example (see [220]) is the bound

$$\int_0^1 \left(\sum_{k \in H} c_k f(kx) \right)^2 dx \le \left(\sum_{\nu=1}^\infty a_\nu^2 \Delta(\nu) \right) \sum_{k \in H} c_k^2 d(k)$$

valid for any set H of disjoint positive integers lying in some interval $[e^r, e^{r+1}]$, $r \ge 1$. Here a_k are the complex Fourier coefficients of f. Using standard methods, such bounds lead easily to a.e. convergence criteria for sums $\sum_k c_k f(kx)$, see the papers cited above.

In Wintner [223] it was proved that if f is a periodic L^2 function with Fourier coefficients a_k, b_k, then the series $\sum_k c_k f(kx)$ converges in L^2 norm for all coefficient sequences $(c_k)_{k \ge 1}$ satisfying $\sum_k c_k^2 < \infty$ if and only if the functions defined by the Dirichlet series

$$\sum_{k=1}^\infty a_k k^{-s}, \qquad \sum_{k=1}^\infty b_k k^{-s},$$

are bounded and regular in the half plane $\Re(s) > 0$. There is also a remarkable connection between the maximal order of magnitude of GCD sums with the order of extreme values of the Riemann zeta function in the critical strip; see [72, 94, 138, 209].

Naturally, estimating the integral (6) provides important information also on the asymptotic behavior of averages

$$(10) \qquad \frac{1}{N} \sum_{k=1}^N f(n_k x).$$

By the Weyl equidistribution theorem, for any 1-periodic f with bounded variation in $(0,1)$ we have

$$(11) \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{k=1}^{N} f(kx) = \int_0^1 f(x)\,dx \quad \text{a.e.}$$

(actually for every irrational x). Khinchin [150] conjectured that (11) holds for every 1-periodic Lebesgue integrable f as well. This conjecture remained open for nearly 50 years and was finally disproved by Marstrand [178]. An example for a periodic integrable f and a sequence $(n_k)_{k\geq 1}$ of positive integers such that the averages (10) do not converge almost everywhere had already been given earlier by Erdős [100]. On the other hand, Koksma [157] proved that (11) holds if $f \in L^2$ and the Fourier coefficients a_k, b_k of f satisfy

$$\sum_{k=1}^{\infty} \left((a_k^2 + b_k^2) \sum_{d|k} \frac{1}{d} \right) < \infty,$$

and Berkes and Weber [70] proved that the last condition is optimal. No similarly sharp criteria are known in the case $f \in L^1$. For further results related to the Khinchin conjecture, see [37, 50, 70, 74, 185].

4. THE CENTRAL LIMIT THEOREM FOR LACUNARY SEQUENCES

Salem and Zygmund [201] proved the first central limit theorem (CLT) for lacunary trigonometric sums. More specifically, they showed that for any integer sequence $(n_k)_{k\geq 1}$ satisfying the Hadamard gap condition one has

$$\lim_{N\to\infty} \lambda \left(x \in (0,1) : \sum_{k=1}^{N} \cos(2\pi n_k x) \leq t\sqrt{N/2} \right) = \Phi(t),$$

where Φ denotes the standard normal distribution. Note that

$$\int_0^1 \left(\sum_{k=1}^{N} \cos(2\pi n_k x) \right)^2 dx = \frac{N}{2},$$

so the result above contains the "correct" variance for the limit distribution, exactly as it should also be expected in the truly independent case. This result has been significantly strengthened since then; for example, Philipp and Stout [196] showed that under the Hadamard gap condition the function

$$S(t,x) = \sum_{k\leq t} \cos(2\pi n_k x),$$

considered as a stochastic process over the space $([0,1], \mathcal{B}[0,1], \lambda)$, is a small perturbation of a Wiener process, a characterization which allows to deduce many fine asymptotic results for this sum. It is also known that the central limit theorem for pure trigonometric lacunary sums remains valid under a slightly weaker gap condition than Hadamard's: as Erdős [102] proved, it is sufficient to assume that

$n_{k+1}/n_k \geq 1 + ck^{-\alpha}$, $\alpha < 1/2$, while such an assumption with $\alpha = 1/2$ is not sufficient.

The whole situation becomes very different when the cosine-function is replaced by a more general 1-periodic function, even if it is such a well-behaved one as a trigonometric polynomial. For example, consider

$$(12) \qquad f(x) = \cos(2\pi x) - \cos(4\pi x), \qquad n_k = 2^k, \ k \geq 1.$$

In this case the lacunary sum is telescoping, and it can be immediately seen that there cannot be a non-trivial limit distribution. A more delicate example is attributed to Erdős and Fortet[2], and goes as follows. Let

$$f(x) = \cos(2\pi x) + \cos(4\pi x), \qquad n_k = 2^k - 1, \ k \geq 1.$$

Then it can be shown that $N^{-1/2} \sum_{k=1}^N f(n_k x)$ does indeed have a limit distribution, but one which is actually non-Gaussian. More precisely, for this example one has

$$\lim_{N \to \infty} \lambda \left(x \in (0,1) : \sum_{k=1}^N f(n_k x) \leq t\sqrt{N/2} \right) = \frac{1}{\sqrt{\pi}} \int_0^1 \int_{-\infty}^{t/2|\cos(\pi s)|} e^{-u^2} du\, ds.$$

Thus the limit distribution in this case is a so-called "variance mixture Gaussian", which can be seen as a normal distribution whose variance is a function rather than a constant. This limiting behavior can be explained from the observation that

$$(13) \qquad f(n_k x) = \cos((2^{k+1} - 2)\pi x) + \cos((2^{k+2} - 4)\pi x)$$

and

$$(14) \qquad f(n_{k+1} x) = \cos((2^{k+2} - 2)\pi x) + \cos((2^{k+3} - 4)\pi x).$$

Combining the second term on the right-hand side of (13) with the first term on the right-hand side of (14) we obtain

$$\cos((2^{k+2} - 4)\pi x) + \cos((2^{k+2} - 2)\pi x) = 2\cos(\pi x)\cos((2^{k+2} - 3)\pi x),$$

so the whole lacunary sum $\sum_{k=1}^N f(n_k x)$ can essentially be written as $2\cos(\pi x)$ multiplied with a pure cosine lacunary sum. This is exactly what the "variance mixture Gaussian" indicates: the limit distribution is actually that of $2\cos(\pi x)$ independently multiplied with a Gaussian. The failure of a Gaussian central limit theorem in the example above can be seen as a consequence of the fact that the Diophantine equation

$$n_{k+1} - 2n_k = 1$$

---

[2]The Erdős–Fortet example is first mentioned in print in a paper of Salem and Zygmund [202]. They mention the example without proof, and write: "This remark is essentially due to Erdős.". Later the example was mentioned in a paper of Kac [146], who wrote: "It thus came as a surprise when simultaneously and independently of each other, Erdős and Fortet constructed an example showing that the limit [...] need not be Gaussian", with a footnote: "In Salem and Zygmund this example is erroneously credited to Erdős alone." No proof is given in Kac's paper either, but he writes: "Details will be given in [a forthcoming] paper by Erdős, Ferrand, Fortet and Kac". Such a joint paper never appeared.

possesses many solutions k for this particular choice of sequence. Equipped with this observation, one could readily construct similar examples with other trigonometric polynomials f, and other variance mixture Gaussians as limit distributions, by creating situations where there are many solutions k, ℓ to

$$(15) \qquad\qquad an_k - bn_\ell = c$$

for some fixed a, b, c. However, interestingly, a special role is played by such equations when c has the particular value $c = 0$; very roughly speaking, solutions of the equation for $c = 0$ effect only the limiting variance (in a Gaussian distribution), but not the structure of the limiting distribution itself. This is visible in a paper of Kac [145], who studied the sequence $n_k = 2^k$, $k \geq 1$, where indeed the only equations that have many solution are of the form $2^m n_k - n_\ell = 0$ for some m (the solutions being $\ell = k + m$). Kac proved that for this sequence and any 1-periodic f of bounded variation and zero mean one has

$$(16) \qquad \lim_{N \to \infty} \lambda \left(x \in (0, 1) : \sum_{k=1}^{N} f(n_k x) \leq t \sigma_f \sqrt{N} \right) = \Phi(t)$$

with a limiting variance σ_f^2, provided that

$$(17) \qquad \sigma_f^2 := \int_0^1 f^2(x) \, dx + 2 \sum_{m=1}^{\infty} \int_0^1 f(x) f(2^m x) \, dx \neq 0.$$

Thus in this case the limit distribution is always a Gaussian, and the failure of the trivial example in (12) to produce such a Gaussian limit comes from the fact that the limiting variance is degenerate.

These observations show that there is a delicate interplay between arithmetic, analytic and probabilistic effects; in particular, it is obviously not only the order of growth of $(n_k)_{k \geq 1}$ which is responsible for the fine probabilistic behavior of a lacunary sum. Takahashi [211] proved a CLT (with pure Gaussian limit) under the assumption that $n_{k+1}/n_k \to \infty$, and Gaposhkin [128] proved that a CLT (with pure Gaussian limit) holds when n_{k+1}/n_k is an integer for all k, or if $n_{k+1}/n_k \to \alpha$ for some α such that $\alpha^r \notin \mathbb{Q}$, $r = 1, 2, \dots$ (and if additionally the variance does not degenerate). A general framework connecting Diophantine equations and the distribution of lacunary sums was established in Gaposhkin's profound paper [131], where he proved that a CLT (with pure Gaussian limit) holds if for all fixed positive integers a, b the number of solutions k, ℓ of the Diophantine equation (15) is bounded by a constant which is independent of c (where only $c \neq 0$ needs to be considered, provided that the variance does not degenerate). One can check the validity of this general condition for sequences satisfying the assumptions mentioned earlier in this paragraph, such as $n_{k+1}/n_k \to \infty$ or $n_{k+1}/n_k \to \alpha$ for $\alpha^r \notin \mathbb{Q}$. Finally, an optimal result was established in [13]: For $(n_k)_{k \geq 1}$ satisfying the Hadamard gap condition, the limit distribution of $N^{-1/2} \sum_{k=1}^{N} f(n_k x)$ is Gaussian provided that the number of solutions (k, ℓ) of (15), subject to $k, \ell \leq N$, is of order $o(N)$ (for all fixed a, b, uniformly in $c \neq 0$). If, on the other hand, for some a, b, c the number of solutions

is $\Omega(N)$, then the CLT generally fails to hold. If the number of solutions with $c = 0$ also is of order $o(N)$, then the CLT has the "correct" variance $\int_0^1 f(x)^2 dx$, in perfect accordance with the independent case. Even if the number of solutions is of order $\Omega(N)$ for some a, b, c, then the deviation of the distribution of $N^{-1/2} \sum_{k=1}^N f(n_k x)$ from the Gaussian distribution can be quantified in terms of the ratio "(number of solutions)$/N$". This shows for example that while the CLT generally fails in the case $n_{k+1}/n_k \to p/q$, one obtains an "almost CLT" if both p and q are assumed to be large. Another example for such an "almost CLT" is when the growth constant in the Hadamard gap condition is assumed to be very large. See the statement of [13, Theorem 1.3] and the subsequent discussion for more details.

Note that Gaposhkin's condition implies that the CLT also holds for all subsequences that are picked out of $(n_k)_{k \geq 1}$. This is not the case under the assumptions from [13], where one might be able to extract a subsequence along which the CLT fails (by choosing a subsequence which allows a large number of solutions of the relevant Diophantine equations). It is interesting that the probabilistic behavior of lacunary sums might change when one passes to a subsequence of the original sequence – this is in clear contrast to the bevahior of sums of independent random variables, where any subsequence of course is independent as well. A similar remark holds for permutations of lacunary sums resp. permutations of sums of independent random variables. These phenomena have received strong attention during the last years; see for example [17, 18, 19, 21, 114]. To give only one sample result, in [17] the following is shown. As noted above, the CLT is true for pure trigonometric sums under the Erdős gap condition $n_{k+1}/n_k \geq 1 + ck^{-\alpha}$ for some $\alpha < 1/2$. However, this is only true for the unpermuted sequence (i.e. sorted in increasing order). If permutations of the sequence are allowed, then this gap condition is not sufficient anymore for the validity of the CLT, as is no other gap conditon weaker than Hadamard's. More precisely, for any sequence $(\varepsilon_k)_{k \geq 1}$ with $\varepsilon_k \to 0$ there exists a sequence of positive integers satisfying $n_{k+1}/n_k \geq 1 + \varepsilon_k$, together with a permutation $\sigma : \mathbb{N} \mapsto \mathbb{N}$, such that the permuted (pure trigonometric) sum $N^{-1/2} \sum_{k=1}^N \cos(2\pi n_{\sigma(k)} x)$ converges in distribution to a non-Gaussian limit. One can also construct such examples where the norming sequence $N^{-1/2}$ has to be replaced by $(\log N)^{1/2} N^{-1/2}$ and the limit is a Cauchy distribution, and examples where no limit distribution exists at all. See [17] for details on this particular result, and Chapter 3 of [61] for a detailed discussion of permutation-invariance of limit theorems for lacunary (trigonometric) systems.

We close this section with some further references. For Hadamard lacunary $(n_k)_{k \geq 1}$, the limit distribution of $N^{1/2} D_N(n_k x)$ was calculated in [14]; under suitable Diophantine assumptions it coincides with the Kolmogorov distribution, which is the distribution of the range of a Brownian bridge. A central limit theorem for Hardy–Littlewood–Pólya sequences was established in [124]. In [87] the Erdős–Fortet example was revisited from the perspective of ergodic theory, and was interpreted in terms of the limiting behavior of certain modified ergodic sums, and generalized to cases such as expanding maps, group actions, and chaotic dynamical systems under

the assumption of multiple decorrelation. See also [86, 88]. The limit distribution of $N^{-1/2} \sum_{k=1}^{N} \cos(2\pi n_k x)$ for the special sequence $n_k = k^2$, $k \geq 1$, was determined by Jurkat and Van Horne in [141, 142, 143], and turned out to have finite moments of order < 4, but not of order 4. The theory of such sums is closely related to theta sums, and goes back to Hardy and Littlewood [135]. For further related results, see [80, 108, 219]. For non-Gaussian limit distributions of $N^{-1/2} \sum_{k=1}^{N} \cos(2\pi n_k x)$ near the Erdős gap condition $n_{k+1}/n_k \geq 1 + ck^{-1/2}$ see [58]. For a multidimensional generalization of Kac's results see [112, 120], and for a multidimensional generalization of the CLT for Hardy–Littlewood–Pólya sequences (considering a semi-group generated by powers of matrices instead) see [167, 168]. See also [85, 90] for generalizations of the CLT for Hardy–Littlewood–Pólya sequences to a very general setup of sums over powers of transformations/automorphisms.

5. THE LAW OF THE ITERATED LOGARITHM FOR LACUNARY SEQUENCES

Together with the law of large numbers (LLN) and the central limit theorem (CLT), the law of the iterated logarithm (LIL) is one of the fundamental results of probability theory. Very roughly speaking, the (strong) law of large numbers says that when scaling by N^{-1} one has almost sure convergence of a sum of random variables, and the central limit theorem says that when scaling by $N^{-1/2}$ one has a (Gaussian) limit distribution. The law of the iterated logarithm operates between these two other asymptotic limit theorems; in its simplest form, it says that for a sequence $(X_n)_{n \geq 1}$ of centered i.i.d. random variables (under suitable extra assumptions, such as boundedness) one has

$$\limsup_{N \to \infty} \frac{\sum_{n=1}^{N} X_n}{\sqrt{2N \log \log N}} = \sigma \qquad \text{almost surely,}$$

where σ is the standard deviation. Heuristically, the law of the iterated logarithm identifies the threshold between convergence in distribution and almost sure convergence for sums of i.i.d. random variables; indeed, while $\frac{\sum_{n=1}^{N} X_n}{\sqrt{2N \log \log N}}$ converges to 0 in distribution by the CLT, it does not converge to 0 almost surely by the LIL. The first version of the LIL was given by Khinchin in 1924, and a more general variant was established by Kolmogorov in 1929. Note that the law of large numbers for trigonometric sums or sums of dilated functions is rather unproblematic: for *any* sequence of distinct integers $(n_k)_{k \geq 1}$ one has

$$(18) \qquad \lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} f(n_k x) = \int_0^1 f(x) \, dx,$$

as long as one can assume a bit of regularity for f (such as f being a trigonometric polynomial, being Lipschitz-continuous, being of bounded variation on $[0,1]$, etc.). Only if one is not willing to impose any regularity assumptions upon f the situation becomes quite different; see the remarks at the end of Section 3.

The most basic law of the iterated logarithm for lacunary systems is

$$(19) \qquad \limsup_{N \to \infty} \frac{\sum_{n=1}^{N} \cos(2\pi n_k x)}{\sqrt{2N \log \log N}} = \frac{1}{\sqrt{2}} \qquad \text{a.e.}$$

under the Hadamard gap condition on $(n_k)_{k \geq 1}$; this was obtained by Salem and Zygmund (upper bound) [203] and Erdős and Gál (lower bound) [103]. Generally speaking, as often in probability theory the lower bound is more difficult to establish than the upper bound, since the latter can be proved by an application of the first Borel–Cantelli lemma (convergence part), while the former is proved by the second Borel–Cantelli lemma (divergence part, which needs some sort of stochastic independence as an extra assumption). Note that (19) is a perfect analogue of (18) with the "correct" constant on the right-hand side.

As in the case of the CLT, replacing pure trigonometric sums by sums of more general 1-periodic functions makes the situation much more delicate. As in the previous section, a key role is played by Diophantine equations. However, while for the CLT it is crucial that the number of solutions of Diophantine equations "stabilizes" in some way to allow for a limit distribution (albeit a potentially non-Gaussian one), no such property is necessary for the validity of a form of the LIL (since, as noted above, this is defined as a $\limsup$, not as a $\lim$). Instructive examples are the following. In all examples, we assume that f is 1-periodic with mean zero and bounded variation on $[0, 1]$.

- If $n_{k+1}/n_k \to \infty$ as $k \to \infty$, then

$$\limsup_{N \to \infty} \frac{\sum_{n=1}^{N} f(n_k x)}{\sqrt{2N \log \log N}} = \left(\int_0^1 f^2(x) \, dx \right)^{1/2} \qquad \text{a.e.}$$

- If $n_k = 2^k$, $k \geq 1$, then

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

with

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

- Assume that $n_{k+1}/n_k \geq q > 1$, $k \geq 1$. Then there exists a constant C (depending on f and on q) such that

$$(20) \qquad \limsup_{N \to \infty} \frac{\sum_{n=1}^{N} f(n_k x)}{\sqrt{2N \log \log N}} \leq C \qquad \text{a.e.}$$

- If $n_k = 2^k - 1$, $k \geq 1$, and if $f(x) = \cos(2\pi x) + \cos(4\pi x)$, then

$$(21) \qquad \limsup_{N \to \infty} \frac{\sum_{n=1}^{N} f(n_k x)}{\sqrt{2N \log \log N}} = \sqrt{2}|\cos(\pi x)| \qquad \text{a.e.}$$

The first result in this list (due to Takahashi [213]) is in perfect accordance with the LIL for truly independent random sums, in accordance with the fact that

also the CLT holds in the "truly independent" form under the large gap condition $n_{k+1}/n_k \to \infty$. The second result is an analogue of Kac's CLT in Equations (16) and (17): as with the CLT, also the LIL holds for the sequence $(2^k)_{k \geq 1}$, but the limiting variance deviates from the one in the "truly independent" case. Note that in contrast to the CLT case we now do not need to require that $\sigma_f \neq 0$ for the validity of the statement. The third result (Takahashi [212]) asserts that there is an upper-bound version of the LIL for lacunary sums (even for sequences where there is no convergence of distributions, and any form of the CLT fails). Finally, the fourth result (the Erdős–Fortet example for the LIL instead of the CLT) shows the remarkable fact that the lim sup in the LIL for Hadamard lacunary sums might actually be *non-constant* – this is very remarkable, and a drastic deviation from what one can typically observe for sequences of independent random variables. In particular this example shows that under the Hadamard gap condition an upper-bound version of the LIL is in general the best that one can hope for. Not very surprisingly, the source of all these phenomena are (as in the previous section) Diophantine equations such as (15), and their number of solutions within the sequence $(n_k)_{k \geq 1}$. So in the LIL there is again a complex interplay between probabilistic, analytic and arithmetic aspects which controls the fine asymptotic behavior of lacunary sums.

In probability theory there is a version of the LIL for the Kolmogorov–Smirnov statistic of an empirical distribution. This is called the Chung–Smirnov LIL, and in the special case of a sequence $(X_n)_{n \geq 1}$ of i.i.d. random variables having uniform distribution on $[0, 1]$ (where the Kolmogorov–Smirnov statistic coincides with the discrepancy) it asserts that

$$\limsup_{N \to \infty} \frac{N D_N(X_n)}{\sqrt{2N \log \log N}} = \frac{1}{2} \qquad \text{almost surely.}$$

Here the number $1/2$ on the right-hand side arises essentially as the maximal L_2 norm ("standard deviation") of a centered indicator function of an interval $A \subset [0, 1]$ (namely the indicator function of an interval of length $1/2$). Based on the principle that lacunary sequences tend to "imitate" the behavior of truly independent sequences, it was conjectured that an analogue of the Chung–Smirnov LIL should also hold for the discrepancy of $(\{n_k x\})_{k \geq 1}$, where $(n_k)_{k \geq 1}$ is a Hadamard lacunary sequence. This was known as the Erdős–Gál conjecture, and was finally solved by Philipp [193], who proved that for any $q > 1$ there exists a constant C_q such that for $(n_k)_{k \geq 1}$ satisfying $n_{k+1}/n_k \geq q$ we have

$$(22) \qquad \frac{1}{\sqrt{32}} \leq \limsup_{N \to \infty} \frac{N D_N(\{n_k x\})}{\sqrt{2N \log \log N}} \leq C_q \qquad \text{a.e.}$$

An admissible value of C_q was specified in [193] as $C_q = 166/\sqrt{2} + 664/(\sqrt{2q} - \sqrt{2})$. The first inequality in (22) follows from (a complex version of) Koksma's inequality together with (19), so the novelty is the second inequality (upper bound). Note also that the upper bound in (22) implies Takahashi's "upper bound" LIL in (20), again as a consequence of Koksma's inequality.

Philipp's result has been extended and refined into many different directions. The most precise results, many of which were obtain by Fukuyama, show again a fascinating interplay between arithmetic, analytic and probabilistic effects. As a sample we state the following results (all from Fukuyama's paper [113]):

Let $n_k = \theta^k$, $k \geq 1$. Then

$$(23) \qquad \limsup_{N \to \infty} \frac{N D_N(\{n_k x\})}{\sqrt{2N \log \log N}}$$

exists and is constant (almost everywhere). Denoting the value of this $\limsup$ by Σ_θ, we have:

- If $\theta^r \notin \mathbb{Q}$ for all $r = 1, 2, \ldots$, then $\Sigma_\theta = 1/2$ a.e.
- If r denotes the smallest positive integer such that $\theta^r = p/q$ for some coprime p, q, then $1/2 \leq \Sigma_\theta \leq \sqrt{(pq+1)/(pq-1)}/2$ a.e.
- If $\theta^r = p/q$ as above and both p and q are odd, then $\Sigma_\theta = \sqrt{(pq+1)/(pq-1)}/2$ a.e.
- If $\theta = 2$, then $\Sigma_\theta = \sqrt{42}/9$ a.e.
- If $\theta > 2$ is an even integer, then $\Sigma_\theta = \sqrt{(p+1)p(p-2)/(p-1)^3}/2$ a.e.
- If $\theta = 5/2$, then $\Sigma_\theta = \sqrt{22}/9$ a.e.

All these results were obtained by very delicate calculations involving Fourier analysis and Diophantine equations. The calculations from [113] were continued by the same author and his group in [119, 121, 122, 125], so that now we have a relatively comprehensive picture on the behavior of these $\limsup$'s in the case when $(n_k)_{k \geq 1}$ is (exactly) a geometric progression.

In [7, 8] for general Hadamard lacunary sequences $(n_k)_{k \geq 1}$ a direct connection was established which links the number of solutions of (15) with the value of the $\limsup$ in the LIL, in the same spirit as this was done before in [13] for the CLT (as described in the previous section). In particular, if the number of solutions of (15) is sufficiently small, then the LIL holds with the constant $1/2$ on the right-hand side, exactly as in the truly independent case. Another interesting observation is that if $(n_k)_{k \geq 1}$ is Hadamard lacunary with growth factor $q > 1$, and if Σ denotes the value of the $\limsup$ in (23), then the difference $|\Sigma - 1/2|$ can be quantified in terms of q and tends to zero a.e. as $q \to \infty$. Thus there is a smooth transition towards the "truly independent" LIL as the growth factor q increases, and under the large gap condition $n_{k+1}/n_k \to \infty$ the value of Σ actually equals $1/2$. Another remarkable fact is that there exist Hadamard lacunary sequences for which the $\limsup$ in the LIL for the discrepancy is not a constant almost everywhere, but rather a function of x, similar to what happened in (21) for the LIL for $\sum f(n_k x)$. In some cases the limit functions in the LIL for the discrepancy can be explicitly calculated, and are "surprisingly exotic" (in the words of Ben Green's MathSciNet review of [6]).

As noted above, Philipp's LIL for the discrepancy has been extended into many different directions. For example, while it is known that the result can fail as soon

as the Hadamard gap condition is relaxed to any sub-exponential growth condition, it turns out to be possible to obtain an LIL for the discrepancy when a weaker growth condition is compensated by stronger arithmetic assumptions. In particular, an analogue of Philipp's result has been proved for Hardy–Littlewood–Pólya sequences [195]; see also [5, 65, 123, 215]. As a closing remark concerning the LIL, it is interesting that the optimal value of the lower bound in (22) is still unknown; cf. [25] for more context.

While much effort has been spent towards understanding the probabilistic behavior of lacunary sums at the scales of the CLT and LIL, it seems that investigations at other scales (such as in particular at the large deviations scale) were only started recently. The few results which are currently available point once again towards an intricate connection between probabilistic, analytic and arithmetic effects; see the very recent papers [22, 111].

6. NORMALITY AND PSEUDORANDOMNESS

Normal numbers were introduced by Borel [73] in 1909. From the very beginning the concept of normality of real numbers was associated with "randomness". While normality of real numbers was originally defined in terms of counting the number of blocks of digits, it is not difficult to see[3] that a number x is normal in base b if and only if the sequence $(\{b^n x\})_{n \geq 1}$ is equidistributed. As Borel proved, Lebesgue-almost all real numbers are normal in a (fixed) integer base $b \geq 2$, and thus almost all reals are normal in *all* bases $b \geq 2$ (such numbers are called *absolutely normal*). While normal numbers are ubiquitous from a measure-theoretic perspective,[4] it is difficult to construct normal numbers. The most fundamental construction is due to Champernowne, who proved (using combinatorial arguments) that the number

$$0.1\ 2\ 3\ 4\ 5\ 6\ 7\ 8\ 9\ 10\ 11\ 12\ 13\ 14\ \ldots,$$

which is obtained by a concatenation of the decimal expansions of the integers, is normal in base 10. The idea of creating normal numbers by a concatenation of the (b-ary) expansions of the values of (simple) functions at integers (or primes) is still the most popular, and probably most powerful, method in this field. We only note that Copeland and Erdős [89] proved that

$$0.2\ 3\ 5\ 7\ 11\ 13\ 17\ 19\ 23\ \ldots,$$

which is obtained by concatenating the decimal expansions of the primes, is normal in base 10, and refer to [92, 93, 171, 174, 186, 187] for more results of this flavor. It should be noted, however, that there have been earlier constructions of a conceptually very different nature, such as that of Sierpinski [208] in 1917. See [43]

---

[3]Probably first explicitly mentioned by D.D. Wall in his PhD thesis, 1949.

[4]Interestingly, while the set of normal numbers is large from a measure-theoretic point of view, it turns out to be small from a topological point of view. More precisely, the set of normal numbers is meager (of first Baire category), see e.g. [133]. In the words of Edmund Hlawka [139, p. 78]: "Thus whereas the normal numbers almost force themselves on to the measure theorist, the topologist is apt to overlook them entirely."

for an exposition of Sierpinski's construction and more context, and see also [44] on an early (unpublished) algorithm of Turing for the construction of normal numbers. We finally mention a very recent idea for the construction of a normal number by Drmota, Mauduit and Rivat [95], which is not based on the concatenation of decimal blocks as above, but rather on the evaluation of an automatic sequence along a subsequence of the index set (in this particular case, the Thue–Morse sequences evaluated along the squares); see also [183, 210].

While Lebesgue-almost all numbers are normal and there are some constructions of normal numbers, it is generally considered to be completely hopeless to prove that natural constants such as $\pi, e, \sqrt{2}, \ldots$ are normal in a given base (although the experimental evidence clearly points in that direction: [189, 218, 224]). Many such open problems "for the next millennium" are contained in Harman's survey article [137]; see also [33]. However, it is quite clear that the mathematical machinery which would be necessary to prove the normality of $\sqrt{2}$ or other such constants is completely lacking; compare the rather deplorable current state of knowledge on the binary digits of $\sqrt{2}$ as given in [34, 98, 217]. A small spark of hope is provided by the very remarkable formulas of Bailey, Borwein and Plouffe (now widely known as BBP formulas), which allow to calculate deep digits of π (and other constants) without the need of computing all previous digits. See [51] for a very comprehensive "source book" covering computational aspects of π, and see [35] for a very rare example of a possible strategy of what a proof of the normality of π could possibly like (cf. also [162]).

Since normality of x in a base b can be expressed in terms of the equidistribution of the sequence $(\{b^n x\})_{n \geq 1}$, it is very natural to consider the discrepancy of $D_N(\{b^n x\})$ and call this (with a slight abuse of language) the discrepancy of x (as a normal number, with respect to a base b). Remarkably, it is still unknown how small the discrepancy of a normal number can be (this is known as *Korobov's problem*). Levin [166] constructed (for given base b) a number x such that

$$D_N(\{b^n x\}) = O\left(\frac{(\log N)^2}{N}\right);$$

by Schmidt's general lower bound the exponent of the logarithm cannot be reduced below 1, but the optimal size of this exponent remains open.

One of the most interesting, and most difficult, aspects of normal numbers is normality with respect to two or more different bases. Extending work of Cassels [79], Schmidt [206] characterized when normality with respect to a certain base implies normality with respect to another base, and when this is not the case. See also [48, 76]. However, generally speaking it is very difficult to construct numbers which are normal with respect to several different bases, and the "constructions" are much less explicit than the ones of Champernowne and Copeland–Erdős mentioned above. The problem of the minimal order of the discrepancy of normal numbers seems to be very difficult when different bases are considered simultaneously.

Aistleitner, Becher, Scheerer and Slaman [12] constructed a number x such that

$$D_N(\{b^n x\}) = O_b\left(N^{-1/2}\right)$$

for all integer bases $b \geq 2$; this is considered to be an "unexpectedly small" order of the discrepancy by Bugeaud in his MathSciNet review of [12]. It is not known if the exponent $-1/2$ of N in this estimate is optimal or not; indeed, no non-trivial lower bounds whatsoever (beyond the general lower bound $(\log N)/N$ of Schmidt) are known for this problem, but it is quite possible that whenever simultaneous normality with respect to different (multiplicatively independent) bases is considered, there must be at least one base for which the discrepancy is "large".

In a recent years, there has been a special focus on algorithmic aspects of the construction of normal numbers. A particularly striking contribution was a polynomial-time algorithm for the construction of absolutely normal numbers due to Becher, Heiber and Slaman [45]. See also [29, 47, 205]. Related to such algorithmic and computational problems are questions on the complexity of the set of normal numbers from the viewpoint of descriptive set theory in mathematical logic; in this framework, the rank of the set of normal numbers [152] and absolutely normal numbers [46] within the Borel hierarchy has been determined.

The notion of normality can be extended in a natural way to many other situations, where it is always understood that normality should be the typical behaviour. For example, one can consider normal continued fractions, where the "expected" number of occurences of each partial quotient is prescribed by the Gauss–Kuzmin measure; see for example [1, 49]. Other generalizations consider for example normality with respect to β-expansions [39, 173], a numeration system which generalizes the b-ary expansion to non-integral bases β, or normality with respect to Cantor expansions [3, 109, 175], a numeration system which allows a different set of "digits" at each position. For a particularly general framework, see [172]. Interestingly, in such generalized numeration systems there can be more than one natural definition of normality, using as starting point for exampe either the idea of counting blocks of digits, or the idea of equidistribution of an associated system. The relation between such different (sometimes conflicting) notions of normality has been studied in particular detail for Cantor expansions [2, 170, 176].

Normal numbers feature prominently in the chapter on random numbers in Volume 2 of Knuth's celebrated series on *The Art of Computer Programming* [153]. There he tries to come to terms with the notion of "random" sequences of numbers, and introduces an increasingly restrictive scheme of "randomness" of deterministic sequences. The concept of normality is also the starting point for one of the (quantitative) measures of pseudorandomness, which were introduced by Mauduit and Sárközy [179] and then studied in a series of papers. Note in this context that the transformation $T : x \mapsto bx \mod 1$, which is at the foundation of the concept of normal numbers, can in some sense be seen as the continuous analogue of the recursive formula which defines a linear congruential generator (LCG), one of the most classical devices for

pseudorandom number generation. For this connection between normal numbers and pseudo-random number generators, see for example [36]. Another very fruitful aspect of normal numbers is the connection with ergodic theory, which comes from the observation that the sequence $(\{b^n x\})_{n \geq 1}$ is the orbit of x under the transformation T from above, and that this transformation is measure-preserving (with respect to the Lebesgue measure) and ergodic. We will not touch upon this connection in any detail, and instead refer to [91, 149].

Another sequence which is often associated with "randomness" is the sequence $(\{x^n\})_{n \geq 1}$ for real $x > 1$, or more generally $(\{\xi x^n\})_{n \geq 1}$ for $\xi \neq 0$ and $x > 1$. This looks quite similar to a (Hadamard) lacunary sequence such as $(b^n x)_{n \geq 1}$, but its metric theory is of a very different nature in several respects. Both sequences are variants of a geometric progression, but while in the lacunary sequence the base b is fixed and x is assumed to be a "parameter", now ξ is assumed to be fixed and the base x of the geometric progression is the parameter. While $(b^n x)_{n \geq 1}$ can in many ways be easily interpreted in terms of harmonic analysis, digital expansions, ergodic theory, etc., such simple interpretations fail for $(\{x^n\})_{n \geq 1}$. Note in particular that in contrast to lacunary sequences there now is no periodicity when replacing $x \mapsto x+1$, there is no "orthogonality", and the calculation of moments of sums $\sum f(x^n)$ does not simply reduce to the counting of solutions of Diophantine equations. Still, what is preserved from the setup of lacunary sequences is that x^n (as a function of x) oscillated quickly on intervals where x^m is essentially constant, provided that n is significantly larger than m, and there are good reasons to consider systems such as $(\cos(2\pi x^n))_{n \geq 1}$ to be "quasi-orthogonal" and "almost independent" in some appropriate sense.

One of the most fundamental results on this type of sequence is due to Koksma [154]: assuming that $\xi \neq 0$ is fixed, the sequence $(\{\xi x^n\})_{n \geq 1}$ is uniformly distributed mod 1 for almost all $x > 1$. In particular, when $\xi = 1$, the sequence $(\{x^n\})_{n \geq 1}$ is uniformly distributed mod 1 for almost all $x > 1$. In very sharp contrast with Koksma's metric result is the fact that until today *not a single example* of a number x is known for which $(\{x^n\})_{n \geq 1}$ is indeed uniformly distributed. This problem is related with Mahler's problem on the range of $(\{(3/2)^n\})_{n \geq 1}$, which also seems to be completely out of reach for current methods (cf. [97, 110]). The sequence $(\{x^n\})_{n \geq 1}$ is discussed at length in Knuth's book, where it is conjectured that this sequence is a good candidate to pass several very strict pseudorandomness criteria for almost all x. For example, Knuth conjectured that for all sequences of distinct integers $(s_n)_{n \geq 1}$ the sequence $(\{x^{s_n}\})_{n \geq 1}$ (a subsequence of the original sequence) has a strong equidistribution property called complete uniform distribution, for almost all $x > 1$; this was indeed established by Niederreiter and Tichy [188]. It is also known that $(\{x^n\})_{n \geq 1}$ satisfies an law of the iterated logarithm in the "truly independent" form

$$\limsup_{N \to \infty} \frac{N D_N(\{x^n\})}{\sqrt{2N \log \log N}} = \frac{1}{2} \qquad \text{a.e.,}$$

and similarly satisfies a central limit theorem which is perfectly analogous to the one for truly independent systems [9]. Note that Knuth's assertion that the sequence $(\{x^n\})_{n\geq 1}$ shows good pseudo-random behavior for almost all $x > 1$ is of limited practical use, as long as no such value of x is found. The discrete analogue would be to study the pseudo-randomness properties of $a^n \bmod q$ for $n = 1, 2, \ldots$, where a and q are fixed integers. Investigations on the pseudo-randomness properties of such sequence were carried out for example by Arnol'd [32], who experimentally observed good pseudorandom behavior; cf. also [10].

To close this section, we note that equidistribution is of course just one property which can be used to characterize "pseudorandom" behavior (essentially by analogy with the Glivenko–Cantelli theorem). There are many other statistics which could be applied to a sequence in $[0, 1]$ to determine whether it behaves in a "random" way or not. One class of such statistics are gap statistics at the level of the average gap (which is of order $1/N$ when considering the first N elements of a sequence in $[0, 1]$), such as the distribution of nearest-neighbor gaps, or the pair correlation statistics. We do not give formal definitions of these concepts here, but note that they are inspired by investigations of the statistics of quantum energy eigenvalues in the context of the Berry–Tabor conjecture in theoretical physics; see [177] for more context. Pseudorandom behavior with respect to such statistics is called "Poissonian", since it agress with the corresponding statistics for the Poisson process. The general principle that lacunary systems show pseudorandom behavior is also valid in this context. For example, Rudnick and Zaharescu [199] showed that for $(n_k)_{k\geq 1}$ satisfying the Hadamard gap condition the sequence $(\{n_k x\})_{k\geq 1}$ is Poissonian for almost all x, and Aistleitner, Baker, Technau and Yesha [11] showed that the same holds for $(\{x^n\})_{n\geq 1}$ for almost all $x > 1$.

This section on normal numbers and sequences of the form $(\{\xi x^n\})_{n\geq 1}$ gives of course only a very brief overview of the subject, and has to leave out many interesting aspects. For a much more detailed exposition we refer the reader to the book of Bugeaud [75].

7. Random sequences

In the previous sections we have illustrated the philosophy that gap sequences exhibit many probabilistic properties which are typical for sequences of i.i.d. random variables. In many cases the large gap condition $n_{k+1}/n_k \to \infty$ gives "true" random limit theorems, the Hadamard gap condition $n_{k+1}/n_k \geq q > 1$ is a critical transition point where a mixture of probabilistic, analytic and arithmetic effects comes into play, and the "almost independent" behavior is lost when the gap condition is relaxed below Hadamard's. There are results which hold under weaker gap conditions such as the Erdős gap condition $n_{k+1}/n_k \geq 1 + ck^{-\alpha}$, $0 < \alpha < 1/2$, or under additional arithmetic assumptions, but as a whole the Hadamard gap condition is the critical point where the "almost independent" behavior of systems of dilated functions starts to break down.

However, while almost independent behavior is generally lost under a weaker gap condition (without strong arithmetic information), there is another possible perspective on the problem. As noted, for a fixed sequence $(n_k)_{k\geq 1}$ one cannot expect "almost independent" behavior of $(\{n_k x\})_{k\geq 1}$, say, without assuming a strong growth condition on $(n_k)_{k\geq 1}$. However, even without such a growth condition one might expect that $(\{n_k x\})_{k\geq 1}$ shows independent behavior for "typical" sequences $(n_k)_{k\geq 1}$. Here the word "typical" of course implies that the sequence has to be taken from a generic set in some appropriate space which possesses a measure, so quite naturally this idea leads to considering "random" sequences $(n_k)_{k\geq 1} = (n_k(\omega))_{k\geq 1}$ which are constructed in a randomized way over some probability space.

Of course there are many possible ways how a random sequence can be constructed. From results of Salem and Zygmund [204] for trigonometric sums with random signs it follows easily that if we define a sequence $(n_k)_{k\geq 1}$ by flipping a coin (independently) for every positive integer to decide whether it should be contained in the sequence or not, and let $\mathbb{P}$ denote the probability measure on the space over which the "coins" are defined, then for $\mathbb{P}$-almost all sequences as defined above one has

$$(24) \qquad \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \cos(2\pi n_k x) \xrightarrow{\mathcal{D}} N(0, 1/4)$$

and

$$(25) \qquad \limsup_{N \to \infty} \frac{1}{\sqrt{2N \log \log N}} \sum_{k=1}^{N} \cos(2\pi n_k x) = \frac{1}{2} \qquad \text{for almost all } x,$$

where $N(0, \sigma^2)$ denotes the normal distribution with mean 0 and variance σ^2 and $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. Note that (24) and (25) are not exactly matching with the truly independent case, where the limit distribution would be $N(0, 1/2)$ and the limsup in the LIL would be $1/\sqrt{2}$. The "loss" on the right-hand sides of (24) and (25) comes from the fact that a Dirichlet kernel is "hiding" in this linearly growing sequence, and this kernel is highly localized near 0 and 1 so that its contribution is lost in the CLT and LIL. By the strong law of large numbers (SLLN) clearly $n_k \sim 2k$ as $k \to \infty$, $\mathbb{P}$-almost surely, so the sequences constructed here are very far from satisfying any substantial gap condition; in contrast, their (typical) order of growth is only linear. It should be noted that the gaps $n_{k+1} - n_k$ in this sequence are not bounded: with full $\mathbb{P}$-probability, $n_{k+1} - n_k = 1$ for infinitely many k (roughly, in half of the cases), but for infinitely many k, the gap $n_{k+1} - n_k$ has order of magnitude $c \log k$; this follows from the "pure heads" theorem of Erdős and Rényi, see [198].

We call an increasing sequence $(n_k)_{k\geq 1}$ of positive integers a B_2 sequence if there exists a constant $C > 0$ such that for any integer $\nu > 0$ the number of representations of ν in the form $\nu = n_k \pm n_\ell$, $k > \ell \geq 1$, is at most C. By a result of Gaposhkin [131] already mentioned in Section 4, the sequence $(f(n_k x))_{k\geq 1}$ satisfies the CLT for all Hadamard lacunary $(n_k)_{k\geq 1}$ and all 1-periodic Lipschitz continuous f if and only if

for any $m \geq 1$, the set-theoretic union of the sequences $(n_k)_{\geq 1}, (2n_k)_{\geq 1}, \ldots, (mn_k)_{\geq 1}$ satisfies the B_2 condition.[5] No similarly complete result is known for sequences $(n_k)_{k \geq 1}$ growing slower than exponentially, but Berkes [54] proved that if $(n_k)_{k \geq 1}$ is a B_2 sequence satisfying the gap condition

$$(26) \qquad n_{k+1}/n_k \geq 1 + ck^{-\alpha}, \qquad k \geq 1,$$

for some $c > 0$, $\alpha > 0$, then $(\cos(2\pi n_k x))_{k \geq 1}$ satisfies the CLT and LIL. To verify the B_2 property for a concrete sequence $(n_k)_{k \geq 1}$ is generally a difficult problem, but the situation is quite different for random constructions. Let $I_1, I_2, \ldots$ be disjoint blocks of consecutive integers and let $n_1, n_2, \ldots$ be independent random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that n_k is uniformly distributed over I_k. Clearly, the number of different sums $\pm n_{k_1} \pm n_{k_2} \pm n_{k_3}$, $1 \leq k_1, k_2, k_3 \leq k-1$, is at most $8(k-1)^3$, and thus if the size of I_k is $\geq k^5$, then the probability that n_k is equal to any of these sums is $\leq 8k^3 k^{-5} = O(k^{-2})$. Thus by the Borel-Cantelli lemma, with $\mathbb{P}$-probability 1, such a coincidence can occur only for finitely many k. Thus the equation

$$\pm n_{k_1} \pm n_{k_2} \pm n_{k_3} \pm n_{k_4} = 0, \quad k_1 \leq k_2 \leq k_3 < k_4$$

has only finitely many solutions, which implies that $(n_k)_{k \geq 1}$ is a B_2 sequence.

Let us recall now that by a result of Erdős [102], $(\cos(2\pi n_k x))_{k \geq 1}$ satisfies the CLT with limit distribution $N(0, 1/2)$, provided that (26) holds with $\alpha < 1/2$, and this result is sharp, i.e. there exists a sequence $(n_k)_{k \geq 1}$ satisfying (26) with $\alpha = 1/2$ such that the CLT fails for $(\cos(2\pi n_k x))_{k \geq 1}$. Note that the counterexample is irregular: while $n_{k+1}/n_k - 1$ is of the order $O(k^{-1/2})$ for most k, there is also a subsequence along which $n_{k+1}/n_k \to \infty$. One may therefore wonder if regular behavior of n_{k+1}/n_k implies the CLT; in particular, Erdős [102] conjectured that the CLT holds for $(\cos(2\pi n_k x))_{k \geq 1}$ if $n_k = \lfloor e^{(k^\beta)} \rfloor$ for some β in the range $0 < \beta \leq 1/2$. (Note that for $\beta > 1/2$ condition (26) is satisfied with $\alpha < 1/2$, so the CLT follows from Erdős' result.) This conjecture was proved by Murai [184] for $\beta > 4/9$, but for smaller β the problem is still open. Random constructions provide here important information. Kaufman [148] proved that if c is chosen at random, with uniform distribution on a finite interval $(a, b) \subset (0, \infty)$, then $(\cos(2\pi n_k x))_{k \geq 1}$ with $n_k = e^{(ck^\beta)}$ satisfies the CLT with probability 1 for any fixed $\beta > 0$. An even wider class of random sequences with the CLT property is obtained by choosing the blocks I_k in the random construction above as the integers in the interval

$$(27) \qquad J_k = \left(e^{(ck^\beta)}(1 - r_k), \ e^{(ck^\beta)}(1 + r_k) \right), \qquad r_k = o(k^{-(1-\beta)}).$$

A simple calculation shows that these intervals are disjoint for $k \geq k_0$ and for $n_k \in J_k$ we have (26) with $\alpha = 1 - \beta$, in fact we even have

$$n_{k+1}/n_k = 1 + c_1(1 + o(1))/k^{1-\beta}$$

---

[5]Note that the definition of the B_2 property used in [131] is slightly different from the standard usage in number theory (see e.g. [134]) requiring that the number of solutions of $\nu = n_k + n_\ell$, $k > \ell \geq 1$, is bounded by C, but this does not affect the discussion below.

with some constant $c_1 > 0$. Now if r_k decreases like a negative power of k, then the length of J_k will be $\geq k^5$ and thus the constructed random sequence $(n_k)_{k \geq 1}$ will be a B_2 sequence with probability 1, so $(\cos(2\pi n_k x))_{k \geq 1}$ satisfies the CLT. In other words, the CLT for $(\cos(2\pi n_k x))_{k \geq 1}$ holds for a huge class of sequences $n_k \sim e^{(ck^\beta)}$ for any $c > 0$, $\beta > 0$.

Concerning B_2 sequences, it is worth pointing out that Erdős [101] proved, decades before Carleson's convergence theorem, that $\sum_{k=1}^\infty (a_k \cos(2\pi n_k x) + b_k \sin(2\pi n_k x))$ is almost everywhere convergent if $(n_k)_{k \geq 1}$ is a B_2 sequence. The question of how slowly B_2 sequences can grow is a much investigated problem of number theory, see e.g. Halberstam and Roth [134], Chapters II and III. It is easily seen that a B_2 sequence $(n_k)_{k \geq 1}$ cannot be $o(k^2)$ and Erdős and Rényi [105] proved by a random construction that for any $\varepsilon > 0$ there exists a B_2 sequence $(n_k)_{k \geq 1}$ with $n_k = O(k^{2+\varepsilon})$. Changing the B_2 property slightly and requiring that all numbers $n_k \pm n_\ell$, $k > \ell$, are actually different, makes the problem considerably harder. The "greedy algorithm" yields a B_2 sequence $(n_k)_{k \geq 1}$ with $n_k = O(k^3)$, see [180], and it took nearly 40 years to improve this to $n_k = o(k^3)$, see [26]. The best currently known (random) construction is due to Ruzsa [200], and satisfies $n_k = k^{1/(\sqrt{2}-1)+o(1)}$.

Let $(\omega_n)_{n \geq 1}$ be a nondecreasing sequence of positive integers tending to $+\infty$ and let us divide the set of positive integers into disjoint blocks $I_1, I_2, \ldots$ such that the cardinality of I_k is ω_k. Using these blocks in the random construction above, the resulting random sequence $(n_k)_{k \geq 1}$ cannot be a B_2 sequence if $(\omega_n)_{n \geq 1}$ grows slower than any power of n, but it is proved in Berkes [55] that with $\mathbb{P}$-probability 1, $(\cos(2\pi n_k x))_{k \geq 1}$ still satisfies the CLT and LIL. The limit distribution here is $N(0, 1/2)$ and the limsup in the LIL is $1/\sqrt{2}$, so that the "loss of mass" phenomenon observed in the case of the random sequence $(n_k)_{k \geq 1}$ in the Salem-Zygmund paper [204] does not occur here. The gaps in this sequence satisfy $n_{k+1} - n_k \leq 2\omega_{k+1}$, i.e. they can grow arbitrarily slowly. An LIL for the discrepancy of $(\{n_k x\})_{k \geq 1}$ under the same gap condition was given in Fukuyama [118]. In [55] the question was raised if there exists a sequence $(n_k)_{k \geq 1}$ with bounded gaps $n_{k+1} - n_k = O(1)$ such that the CLT holds. Bobkov and Götze [71] showed that if we want no loss of mass in the CLT, the answer is negative: if $(n_k)_{k \geq 1}$ is any increasing sequence of positive integers with $n_{k+1} - n_k \leq L$, $k \geq 1$, such that $N^{-1/2} \sum_{k=1}^N \cos(2\pi n_k x)$ has a Gaussian limit distribution $N(0, \sigma^2)$, then necessarily $\sigma^2 < 1/2$ and $L \geq 1/(1-2\sigma^2)$. On the other hand, Fukuyama [115, 116, 117] showed that for any $\sigma^2 < 1/2$ there exists indeed a random subsequence $(\cos(2\pi n_k x))_{k \geq 1}$ of the trigonometric system satisfying the CLT with limit distribution $N(0, \sigma^2)$ and with bounded gaps $n_{k+1} - n_k \leq L$ with $L \sim 4/(1-2\sigma^2)$ as $\sigma^2 \to 1/2$. This shows that the result of Bobkov and Götze is optimal up to a factor 4. This remarkable result is the "small gaps" counterpart of Erdős' central limit theorem [102]: the latter determines the smallest gap sizes in $(n_k)_{k \geq 1}$ implying the CLT for $(\cos(2\pi n_k x))_{k \geq 1}$, while Fukuyama's result determines the smallest gap size which still allows a CLT with limit distribution $N(0, \sigma^2)$ to hold.

It is worth pointing out that the bounded gap sequences in [115, 116, 117] are
obtained by rather complicated random constructions, while using the previously
discussed simple construction and choosing the n_k as independent random variables
uniformly distributed over adjoining blocks I_k with equal length results in a random
sequence $(n_k)_{k\geq 1}$ satisfying almost surely

$$(28) \qquad \frac{1}{\sqrt{N}} \sum_{k=1}^{N} \cos(2\pi n_k x) \xrightarrow{\mathcal{D}} N(0, Y),$$

where $Y \geq 0$ is a random variable and $N(0, Y)$ is a "variance mixture" normal
distribution with characteristic function $\mathbb{E} \exp(-Y t^2/2)$, see [71]. We also note that
there is generally no "loss of mass" phenomenon for the LIL for trigonometric series
with bounded gaps, see [23, 24]. For further results for trigonometric series with
bounded/random gaps, see [42, 41, 62].

8. THE SUBSEQUENCE PRINCIPLE

The purpose of the previous sections was to illustrate the principle that thin subse-
quences of the trigonometric system, or thin subsequences of a more general system
of dilated functions, exhibit properties which are typical for sequences of indepen-
dent random variables. However, an analogous principle holds in a much wider
framework: it is known that, under suitable technical assumptions, sufficiently thin
subsequences of general systems of random variables behave like genuine indepen-
dent sequences, in the sense that a general sequence of random variables allows
to extract a subsequence showing independent behavior. For example, Gaposhkin
[128, 132] and Chatterji [83, 84] proved that if $(X_n)_{n\geq 1}$ is any sequence of random
variables satisfying $\sup_n \mathbb{E} X_n^2 < \infty$, then there exist a subsequence $(X_{n_k})_{k\geq 1}$ and
random variables $X \in L_2$, $Y \in L_1$, $Y \geq 0$, such that

$$(29) \qquad \frac{1}{\sqrt{N}} \sum_{k\leq N} (X_{n_k} - X) \xrightarrow{\mathcal{D}} N(0, Y)$$

and

$$(30) \qquad \limsup_{N\to\infty} \frac{1}{\sqrt{2N \log\log N}} \sum_{k\leq N} (X_{n_k} - X) = Y^{1/2} \qquad \text{a.s.,}$$

where as at the end of the previous section $N(0, Y)$ denotes the "variance mix-
ture"normal distribution with characteristic function $\mathbb{E} \exp(-Y t^2/2)$, and where
again $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution. A functional (Strassen type) ver-
sion of (30) was proved by Berkes [52]. By a result of Komlós [159], from any
sequence $(X_n)_{n\geq 1}$ of random variables satisfying $\sup_n \mathbb{E}|X_n| < \infty$ one can select a
subsequence $(X_{n_k})_{k\geq 1}$ such that

$$(31) \qquad \lim_{N\to\infty} \frac{1}{N} \sum_{k\leq N} X_{n_k} = X \qquad \text{a.s.}$$

for some $X \in L_1$. Chatterji [81] proved that if $(X_n)_{n \geq 1}$ is a sequence of random variables satisfying $\sup_n \mathbb{E}|X_n|^p < \infty$ for some $0 < p < 2$, then there exist a subsequence $(X_{n_k})_{k \geq 1}$ and a random variable X with $\mathbb{E}|X|^p < \infty$ such that

$$(32) \qquad \lim_{N \to \infty} \frac{1}{N^{1/p}} \sum_{k \leq N} (X_{n_k} - X) = 0 \qquad \text{a.s.}$$

These results establish the analogues of the central limit theorem (CLT), the law of the iterated logarithm (LIL), the strong law of large numbers (SLLN) and Marczinkiewicz' strong law for subsequences $(X_{n_k})_{k \geq 1}$. Note the mixed (or randomized) character of (29)–(32): the limit X in the strong law of large numbers, the centering factor X in Marczinkiewicz' strong law, and the limiting variance Y in the CLT (which also determines the limsup in the LIL) all become random. For further limit theorems for subsequences of arbitrary random variable sequences, see Gaposhkin [128]. On the basis of these and several other examples, Chatterji [82] formulated the following heuristic principle:

Subsequence Principle. *Let T be a probability limit theorem valid for all sequences of i.i.d. random variables belonging to an integrability class L defined by the finiteness of a norm $\| \cdot \|_L$. Then if $(X_n)_{n \geq 1}$ is an arbitrary (dependent) sequence of random variables satisfying $\sup_n \|X_n\|_L < +\infty$ then there exists a subsequence $(X_{n_k})_{k \geq 1}$ satisfying T in a mixed form.*

In a profound study, Aldous [27] proved the validity of the subsequence principle for all distributional and almost sure limit theorems subject to minor technical conditions. To formulate his results, let $\mathcal{M}$ denote the class of probability measures on the Borel sets of $\mathbb{R}$, equipped with the Lévy metric. By [27], a subset $A \subset \mathcal{M} \times \mathbb{R}^\infty$ is called a *limit statute* if:

(a) $P((\lambda, X_1(\omega), X_2(\omega), \ldots) \in A) = 1$ provided $X_1, X_2, \ldots$ are i.i.d. random variables with distribution λ.

(b) $(\lambda, x_1, x_2, \ldots) \in A$ and $\sum |x_i - x_i'| < \infty$ implies that $(\lambda, x_1', x_2', \ldots) \in A$.

An a.s. limit theorem can thus be identified with a limit statute, where the analytic statement of the theorem is expressed by (a), while relation (b) means that a small perturbation of the sequence $X_1, X_2, \ldots$ does not change the validity of the limit theorem. Let us give two examples of limit statutes representing the strong law of large numbers and the law of the iterated logarithm:

$A_1 = \left\{ (\lambda, \mathbf{x}) \in A : \lim_{N \to \infty} N^{-1} \sum_{k=1}^N x_k = |\lambda|_1 \right\} \cup \{(\lambda, \mathbf{x}) : |\lambda|_1 = \infty\},$

$A_2 = \left\{ (\lambda, \mathbf{x}) \in A : \limsup_{N \to \infty} (2N \log \log N)^{-1/2} \left(\sum_{k=1}^N x_k - N|\lambda|_1 \right) = |\lambda|_2 \right\}$
$\qquad \cup \{(\lambda, \mathbf{x}) : |\lambda|_2 = \infty\}.$

Here $|\lambda|_1$ and $|\lambda|_2$ denote the mean and variance of λ provided they are finite, and we write $|\lambda|_1 = \infty$, resp. $|\lambda|_2 = \infty$ if $\int_{\mathbb{R}} |x| d\lambda(x) = \infty$, resp. $\int_{\mathbb{R}} |x|^2 d\lambda(x) = \infty$.

On the other hand, by the definitions in [27], a weak limit theorem for i.i.d. random variables is a system

$$T = (f_1, f_2, \ldots, \{G_\lambda, \lambda \in \mathcal{M}_0\})$$

where

(a) $\mathcal{M}_0$ is a measurable subset of $\mathcal{M}$.

(b) For each $k \geq 1$, $f_k = f_k(\lambda, x_1, x_2, \ldots)$ is a real function on $\mathcal{M} \times \mathbb{R}^\infty$, measurable in the product topology, satisfying the smoothness condition

$$|f_k(\lambda, \mathbf{x}) - f_k(\lambda, \mathbf{x}')| \leq \sum_{k=1}^{\infty} c_{k,i} |x_i - x_i'|$$

where $0 \leq c_{k,i} \leq 1$ and $\lim_{k\to\infty} c_{k,i} = 0$ for each i.

(c) For each $\lambda \in \mathcal{M}_0$, G_λ is a probability distribution on the real line such that the map $\lambda \to G_\lambda$ is measurable (with respect to the Borel σ-field in $\mathcal{M}_0$).

(d) If $\lambda \in \mathcal{M}_0$ and $X_1, X_2, \ldots$ are independent random variables with common distribution λ then

$$f_k(\lambda, X_1, X_2, \ldots,) \xrightarrow{\mathcal{D}} G_\lambda \quad \text{as } k \to \infty.$$

For example, the central limit theorem corresponds to the case when $\mathcal{M}_0$ is the class of distributions with mean 0 and finite variance,

$$(33) \qquad f_k(\lambda, x_1, x_2, \ldots) = \frac{x_1 + \ldots + x_k - k\mathbb{E}(\lambda)}{\sqrt{k}}$$

and $G_\lambda = N(0, \text{Var}(\lambda))$.

Let now $(X_n)_{n\geq 1}$ be a sequence of random variables with $\sup_n \|X_n\|_L < \infty$ with any norm $\|\cdot\|_L$ on $\mathbb{R}$. Then $(X_n)_{n\geq 1}$ is bounded in probability, i.e.

$$\lim_{K\to\infty} P(|X_n| > K) = 0 \quad \text{uniformly in } n.$$

By an extension of the Helly–Bray theorem (see e.g. [66]), $(X_n)_{n\geq 1}$ has a subsequence $(X_{n_k})_{k\geq 1}$ having a limit distribution conditionally on any event in the probability space with positive probability, i.e. for any $A \subset \Omega$ with $P(A) > 0$ there exists a distribution function F_A such that

$$\lim_{k\to\infty} P(X_{n_k} \leq t \mid A) = F_A(t)$$

for all continuity points t of F_A. According to the terminology of [66], such a subsequence is called *determining*. Thus when investigating asymptotic properties of sufficiently thin subsequences of sequences $(X_n)_{n\geq 1}$ with bounded norms, we can assume, without loss of generality, that $(X_n)_{n\geq 1}$ itself is determining. As is shown in [27, 66], for any determining sequence $(X_n)_{n\geq 1}$ there exists a random measure μ

(i.e. a measurable map from the underlying probability space $(\Omega, \mathcal{F}, \mathcal{P})$ to $\mathcal{M}$) such that for any A with $P(A) > 0$ and any continuity point t of F_A we have

$$(34) \qquad\qquad F_A(t) = \mathbb{E}_A(\mu(-\infty, t])$$

where $\mathbb{E}_A$ denotes conditional expectation given A. This measure μ is called the *limit random measure* of $(X_n)_{n \geq 1}$; see Section 9 below for more details.

With these preparations, we are now in a position to formulate the subsequence theorems of Aldous.

Theorem A (Aldous [27]). *Let $(X_n)_{n \geq 1}$ be a determining sequence with limit random measure μ and let A be a limit statute. Then there exists a subsequence $(X_{n_k})_{k \geq 1}$ such that for any further subsequence $(X_{m_k})_{k \geq 1} \subset (X_{n_k})_{k \geq 1}$ we have*

$$P((\lambda(\omega), X_{m_1}(\omega), X_{m_2}(\omega), \ldots) \in A) = 1.$$

Theorem B (Aldous [27]). *Let $(X_n)_{n \geq 1}$ be a determining sequence with limit random measure μ and let*

$$T = (f_1, f_2, \ldots, \{G_\lambda, \lambda \in \mathcal{M}_0\})$$

be a weak limit theorem. Assume that $P(\mu \in \mathcal{M}_0) = 1$. Then there exists a subsequence $(X_{n_k})_{k \geq 1}$ such that for any further subsequence $(X_{m_k})_{k \geq 1} \subset (X_{n_k})_{k \geq 1}$ we have

$$\lim_{k \to \infty} P(f_k(X_{m_1}(\omega), X_{m_2}(\omega), \ldots \mu(\omega)) \leq t) = \mathbb{E} G_{\mu(\omega)}(t)$$

at all continuity points t of the distribution function on the right hand side.

Writing out Theorem A and B in the case of the limit statutes A_1, A_2 above and the weak limit theorem defined by (33), we get the CLT, LIL and SLLN for thin subsequences of determining sequences, as stated in (29), (30), (31) above.

The proof of Komlós' result (31) exemplifies the technique used in the field of subsequence behavior before Aldous' paper [27], and in particular in proving the results (29)–(32) mentioned above. As Komlós showed, if $(X_n)_{n \geq 1}$ is a sequence of random variables with bounded L_1 norms, then its sufficiently thin subsequences $(X_{n_k})_{k \geq 1}$ are, after a random centering and small perturbation, an identically distributed martingale difference sequence with finite means and thus, by classical martingale theory, they satisfy the SLLN. Martingale versions of the CLT and LIL yield also relations (29), (30) and their functional versions. While this method yields several further limit theorems for lacunary sequences, martingale difference sequences certainly do not satisfy all i.i.d. limit theorems in a randomized form and thus the general subsequence principle cannot be proved in such a way. The proof of Theorems A and B in [27] uses a different way and utilizes near exchangeability properties of subsequences of general sequences of random variables. Let $(X_n)_{n \geq 1}$ be a determining sequence with limit random measure μ and let $(Y_n)_{n \geq 1}$ be a sequence of random variables, defined on the same probability space as the X_n's, conditionally i.i.d. with respect to μ, with conditional distribution μ. (For the construction of such an $(Y_n)_{n \geq 1}$ one may need to enlarge the probability space.) Clearly, $(Y_n)_{n \geq 1}$

is exchangeable, i.e. for any permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the positive integers, the sequence $(Y_{\sigma(n)})_{n \geq 1}$ has the same distribution as $(Y_n)_{n \geq 1}$, and it satisfies limit theorems of i.i.d. random variables in a mixed form. For example, if $\mathbb{E} Y_1^2 < \infty$ and $Y = \mathbb{E}(Y_1 \mid \mu)$, $Z = \mathrm{Var}\,(Y_1 \mid \mu)$, then

$$N^{-1/2} \sum_{k=1}^{N} (Y_k - Y) \xrightarrow{\mathcal{D}} N(0, Z)$$

and

$$\limsup_{N \to \infty} (2N \log \log N)^{-1/2} \sum_{k=1}^{N} (Y_k - Y) = Z^{1/2} \quad \text{a.s.}$$

This principle holds in full generality, i.e. for all a.s. and distributional limit theorems in the above formalization. Indeed, if the Y_n are conditionally i.i.d. with respect to μ and with conditional distribution μ (a random probability measure on $\mathbb{R}$) and if A is a limit statute, then

$$(35) \qquad P((\mu, Y_1, Y_2, \ldots) \in A | \mu)(\omega) = P(\mu(\omega), Y_1^*, Y_2^*, \ldots) \in A) \quad \text{a.e.}$$

where $(Y_n^*)_{n \geq 1}$ is an i.i.d. sequence with marginal distribution $\mu(\omega)$. By the definition of limit statute, the last probability in (35) equals 1 and taking expectations we get

$$P((\mu, Y_1, Y_2, \ldots) \in A) = 1,$$

which is exactly our claim. Specializing to the case of the limit statutes A_1 and A_2 above, we get relations (29) and (30). A similar argument works for distributional limit theorems. Now, as is shown in [27], for every $k \geq 1$ we have

$$(36) \quad (X_{n_1}, X_{n_2}, \ldots X_{n_k}) \xrightarrow{\mathcal{D}} (Y_1, Y_2, \ldots, Y_k) \quad \text{as} \quad n_1 < n_2 < \ldots < n_k, \ n_1 \to \infty.$$

In other words, for large indices the finite dimensional distributions of the sequence $(X_{n_k})_{k \geq 1}$ are close to those of the limiting exchangeable sequence $(Y_k)_{k \geq 1}$ and thus one may expect that limit theorems of $(Y_k)_{k \geq 1}$ (which, as we have just seen, are mixed versions of i.i.d. limit theorems) continue to hold for sufficiently thin subsequences $(X_{n_k})_{k \geq 1}$ as well. Of course, a limit theorem for $(X_{n_k})_{k \geq 1}$ can describe a complicated analytic property of the infinite vector $(X_{n_1}, X_{n_2}, \ldots, X_{n_k}, \ldots)$ which does not follow from the weak convergence of the finite dimensional distributions of the sequence, but with a suitable thinning procedure and delicate analytic arguments, Aldous showed an infinite dimensional extension of (36), leading to the validity of Theorems A and B.

Although the theorems of Aldous are of exceptional generality, there are important results for lacunary sequences which are not covered by them. As was shown by Gaposhkin [128], for every uniformly bounded sequence $(X_n)_{n \geq 1}$ of random variables there exists a subsequence $(X_{n_k})_{k \geq 1}$ and bounded random variables X and $Y \geq 0$ such that for any numerical sequence $(a_n)_{n \geq 1}$ satisfying

$$(37) \qquad\qquad A_N := \sum_{k=1}^{N} a_k^2 \to \infty, \quad a_N = o(A_N^{1/2})$$

we have

$$(38) \qquad \frac{1}{A_N} \sum_{k=1}^{N} a_k (X_{n_k} - X) \xrightarrow{\mathcal{D}} N(0, Y),$$

and if the second relation of (37) is replaced by

$$(39) \qquad a_N = o(A_N / (\log \log A_N)^{1/2})$$

then we have

$$(40) \qquad \limsup_{N \to \infty} \frac{1}{\sqrt{2A_N^2 \log \log A_N}} \sum_{k=1}^{N} a_k (X_{n_k} - X) = Y^{1/2} \quad \text{a.s.}$$

The difference of these results from (29) and (30) is that in the CLT and LIL we have weighted sums $\sum_{k=1}^{N} a_k (X_{n_k} - X)$ instead of ordinary sums $\sum_{k=1}^{N} (X_{n_k} - X)$. For every fixed coefficient sequence $(a_n)_{n \geq 1}$ the CLT and LIL in (38) and (39) follow from Theorems A and B, but the subsequence $(X_{n_k})_{k \geq 1}$ provided by the proofs depends on $(a_k)_{k \geq 1}$ and it is not clear that we can select a subsequence $(X_{n_k})_{k \geq 1}$ satisfying (38) and (39) simultaneously for all considered coefficient sequences $(a_k)_{k \geq 1}$.

Another important situation not covered by Aldous' general theorems is when we investigate permutation-invariance of limit theorems for subsequences. Since the asymptotic properties of an exchangeable sequence $(Y_n)_{n \geq 1}$ do not change after any permutation of its terms, it is natural to expect that the conclusions in Theorem A and B remain valid after an arbitrary permutation of the subsequence $(X_{n_k})_{k \geq 1}$ in the theorems. However, the proofs of Theorem A and B are not permutation-invariant and it does not follow that, e.g., any sequence $(X_n)_{n \geq 1}$ of random variables with bounded L_1 norms contains a subsequence $(X_{n_k})_{k \geq 1}$ satisfying the strong law of large numbers after any permutation of its terms. Using ad hoc methods, the latter result has been proved by Berkes [56] and another classical case, namely the unconditional a.e. convergence of series $\sum c_k (X_{n_k} - X)$ under $\sum c_k^2 < \infty$ for subsequences $(X_{n_k})_{k \geq 1}$ of L_2 bounded sequences $(X_n)_{n \geq 1}$, has been settled by Komlós [160] (see [27] for another proof via exchangeability). It clearly would be desirable to provide further general results in this direction.

We now formulate some structure theorems for lacunary sequences enabling one to handle problems of the kind discussed above. Recall that if $(X_n)_{n \geq 1}$ is a determining sequence with limit random measure μ and $(Y_n)_{n \geq 1}$ is a sequence conditionally i.i.d. with respect to the σ-algebra generated by μ and with conditional marginal distributions μ, then there exists a subsequence $(X_{n_k})_{k \geq 1}$ such that (36) holds. This shows that, in some sense, for large indices the sequence $(X_{n_k})_{k \geq 1}$ resembles the sequence $(Y_k)_{k \geq 1}$, but this property is far too weak to deduce limit theorems for $(X_{n_k})_{k \geq 1}$ from those valid for the exchangeable sequence $(Y_k)_{k \geq 1}$. The following theorem, proved by Berkes and Péter [63], shows that with a suitable choice of the subsequence $(n_k)_{k \geq 1}$, the variables $(X_{n_k})_{k \geq 1}$ can be chosen to be close to the Y_k in a pointwise sense. We call a sequence $(X_n)_{n \geq 1}$ of random variables *ε-exchangeable*

if on the same probability space there exists an exchangeable sequence $(Y_n)_{n\geq 1}$ such that $P(|X_n - Y_n| \geq \varepsilon) \leq \varepsilon$ for all n. Then we have

Theorem C (Berkes and Péter [63])**.** *Let $(X_n)_{n\geq 1}$ be a sequence of random variables bounded in probability, and let $(\varepsilon_n)_{n\geq 1}$ be a sequence of positive reals tending to zero. Then, if the underlying probability space is large enough, thee exists a subsequence $(X_{n_k})_{k\geq 1}$ such that, for all $l \geq 1$, the sequence $X_{n_l}, X_{n_{l+1}}, \ldots$ is ε_l-exchangeable.*

Note that Theorem C provides a different approximating exchangeable sequence $(Y_j^{(l)})_{j\geq 1}$ for each tail sequence $(X_{n_l}, X_{n_{l+1}}, \ldots)$, with termwise approximating error ε_l. The following theorem describes precisely the structure of the the sequences $(Y_j^{(l)})_{j\geq 1}$.

Theorem D (Berkes and Péter [63])**.** *Let $(X_n)_{n\geq 1}$ be a determining sequence of random variables, and let $(\varepsilon_n)_{n\geq 1}$ be a sequence of positive reals. Then there exists a subsequence $(X_{m_k})_{k\geq 1}$ and a sequence $(Y_k)_{k\geq 1}$ of discrete random variables such that*

$$(41) \qquad P\big(|X_{m_k} - Y_k| \geq \varepsilon_k\big) \leq \varepsilon_k \quad k = 1, 2\ldots,$$

and for each $k > 1$ the atoms of the finite σ-field $\sigma\{Y_1, \ldots, Y_{k-1}\}$ can be divided into two classes Γ_1 and Γ_2 such that the following holds. Firstly,

$$(42) \qquad \sum_{B\in\Gamma_1} P(B) \leq \varepsilon_k.$$

Secondly, for any $B \in \Gamma_2$ there exist P_B-independent random variables $\{Z_j^{(B)}, j = k, k+1, \ldots\}$ defined on B with common distribution function F_B such that

$$(43) \qquad P_B\big(|Y_j - Z_j^{(B)}| \geq \varepsilon_k\big) \leq \varepsilon_k, \qquad j = k, k+1, \ldots$$

Here F_B denotes the limit distribution of $(X_n)_{n\geq 1}$ relative to B and P_B denotes conditional probability given B.

We now give applications of Theorem D to the problems discussed above. First we note that using Theorem D it is a simple exercise to prove, for suitable subsequences of a uniformly bounded sequences $(X_n)_{n\geq 1}$, the weighted CLT and LIL in (38), (40) simultaneously for all permitted coefficient sequences $(a_n)_{n\geq 1}$. Next we give a permutation-invariant form of Theorem B for distributional limit theorems.

Definition 1. *We call the weak limit theorem $T = (f_1, f_2, \ldots, S, \{G_\mu, \mu \in \mathcal{M}_0\})$ regular if there exist sequences $p_k \leq q_k$ of positive integers tending to $+\infty$ and a sequence $\omega_k \to +\infty$ such that*

(i) $f_k(\lambda, x_1, x_2, \ldots)$ depends only on $\lambda, x_{p_k}, \ldots, x_{q_k}$.

(ii) f_k satisfies the Lipschitz condition

$$|f_k(\lambda, x_{p_k}, \ldots, x_{q_k}) - f_k(\lambda', x'_{p_k}, \ldots, x'_{q_k})| \leq$$

$$\leq \frac{1}{\omega_k} \sum_{i=p_k}^{q_k} |x_i - x'_i|^\alpha + \varrho^*(\lambda, \lambda')$$

for some $0 < \alpha \leq 1$, where ϱ^* is a metric on $\mathcal{M}_0$ generating the same topology as the Prohorov metric ϱ.

For example, the central limit theorem can be formalized by the functions

$$f_k(\lambda, x_{[k^{1/4}]}, \ldots, x_k) = \frac{x_{[k^{1/4}]} + \ldots + x_k - k\mathbb{E}(\lambda)}{\sqrt{k}},$$

leading to a regular limit theorem. Note that originally we formalized the CLT with the functions f_k in (33) containing all variables $x_1, x_2, \ldots$, but under bounded second moments the first $k^{1/4}$ terms here are irrelevant and hence we can always switch to the present version. The same procedure applies in the general case.

Theorem E (Aistleitner, Berkes and Tichy [20]). *Let $(X_n)_{n \geq 1}$ be a determining sequence with limit random measure $\tilde{\mu}$. Let $T = (f_1, f_2, \ldots, S, \{G_\mu, \mu \in \mathcal{M}_0\})$ be a regular weak limit theorem and assume that $P(\tilde{\mu} \in \mathcal{M}_0) = 1$. Then there exists a subsequence $(X_{n_k})_{k \geq 1}$ such that for any permutation $(X_k^*)_{k \geq 1}$ of $(X_{n_k})_{k \geq 1}$ we have*

$$(44) \qquad f_k(X_1^*, X_2^*, \ldots, \tilde{\mu}) \to_d \int G_{\tilde{\mu}} dP.$$

In case of the CLT formalized above, assuming $\sup_n \mathbb{E}X_n^2 < +\infty$ implies easily that $\tilde{\mu}$ has finite variance almost surely, and thus denoting its mean and variance by X and Y, respectively, we see that the integral in (44) is the distribution $N(0, Y)$. Hence (44) states in the present case that

$$\frac{1}{\sqrt{N}} \sum_{k=1}^N (X_k^* - X) \xrightarrow{\mathcal{D}} N(0, Y),$$

which is the permutation-invariant form of the CLT.

Concerning a.s. limit theorems, a permutation-invariant form of the strong law of large numbers for subsequences of an L_1-bounded sequence was proved, as already mentioned, in Berkes [56], and a similar argument yields the corresponding result for the LIL. No permutation-invariant version of the general result in Theorem A has been proved in the literature, but there is no need for that, since a.s. limit theorems can be reformulated in a distributional form and thus the proof of Theorem B applies with obvious changes. For illustration, we give here the reformulation of the LIL:

Theorem F. *Let $(X_n)_{n \geq 1}$ be a sequence of random variables with $\mathbb{E}|X_n| \leq 1$, $n = 1, 2, \ldots$ Put $S_n = \sum_{i=1}^n X_i$, $S_{k,l} = \sum_{i=k+1}^l X_i$, and $L(N) = (2N \log \log N)^{1/2}$. Then $\limsup_{N \to \infty} S_N/L(N) = 1$ a.s. iff for any $\varepsilon > 0$ there exists a sequence $m_1 < m_2 < \cdots$ of positive integers such that $m_k \geq 5^k$ and*

$$P\left(\max_{m_k \leq j \leq m_{k+1}} \frac{S_{k,j}}{L(j)} > 1 + \varepsilon\right) \leq 2^{-k}, \qquad k \geq k_0,$$

and

$$P\left(\max_{m_k \leq j \leq m_{k+1}} \frac{S_{k,j}}{L(j)} < 1 - \varepsilon\right) \leq 2^{-k}, \qquad k \geq k_0.$$

It is worth pointing out that given a sequence $(X_n)_{n \geq 1}$ of random variables, finding a subsequence $(X_{n_k})_{k \geq 1}$ satisfying the permutation-invariant form of some limit theorem generally requires a much faster growing sequence $(n_k)_{k \geq 1}$ than to find a subsequence to satisfy the original limit theorem. This is a phenomenon which also occurs for lacunary trigonometric sums or lacunary sums of dilated functions; compare the last paragraph of Section 4 above.

In conclusion we note that if $(X_n)_{n \geq 1}$ is a sequence of random variables with finite means over the probability space $(0, 1)$ equipped with the Borel σ-algebra and Lebesgue measure such that for all $n \geq 1$ and $(a_1, \ldots, a_n) \in \mathbb{R}^n$ we have

$$(45) \qquad C_1 \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p} \leq \mathbb{E} \left| \sum_{k=1}^{n} a_k X_k \right| \leq C_1 \left(\sum_{k=1}^{n} |a_k|^p \right)^{1/p}$$

for some $p \geq 1$ and positive constants C_1, C_2, then the closed subspace of $L_1(0, 1)$ spanned by the X_n is isomorphic with the ℓ_p space (Hilbert space if $p = 2$). Relation (45) holds, in particular, if the X_n are i.i.d. symmetric p-stable random variables with $p > 1$, i.e. their characteristic function (Fourier transform) is given by $\exp(-c|t|^p)$ with some $c > 0$. Thus applying the subsequence principle to the "limit theorem" (45) provides important information on the subspace structure of $L_1(0, 1)$. Using this method, Aldous [28] proved the famous conjecture that every infinite dimensional closed subspace of $L_1(0, 1)$ contains an isomorphic copy of ℓ_p for some $1 \leq p \leq 2$. For a further application of this method, see an improvement of the classical theorem of Kadec and Pełczyński [147] on the subspace structure on L_p, $p > 2$, in Berkes and Tichy [67].

9. New results: Exact criteria for the central limit theorem for subsequences

By the classical resonance theorem of Landau [163], for a real sequence $(x_n)_{n \geq 1}$ the series $\sum_{n=1}^{\infty} a_n x_n$ converges for all $(a_n)_{n \geq 1} \in \ell_p$ $(1 \leq p \leq \infty)$ if and only if $(x_n)_{n \geq 1} \in \ell_q$, where $1/p + 1/q = 1$. A deep extension of this result to the case of function series was given by Nikishin [190]. We call a sequence $(f_n)_{n \geq 1}$ of measurable functions on $(0, 1)$ a *convergence system in measure for ℓ_p* if for any real sequence $(a_n)_{n \geq 1} \in \ell_p$ the series $\sum_{n=1}^{\infty} a_n f_n$ converges in measure. In the case $p = 2$ Nikishin proved the following result.

Theorem G (Nikishin [190, 191]). *A function system $(f_n)_{n \geq 1}$ over $(0, 1)$ is a convergence system in measure for ℓ_2 if and only if for any $\varepsilon > 0$ there exists a measurable set $A_\varepsilon \subset (0, 1)$ with measure exceeding $1 - \varepsilon$ and a constant $K_\varepsilon > 0$ such that for all $N \geq 1$ and all $(a_1, \ldots, a_N) \in \mathbb{R}^N$ we have*

$$(46) \qquad \int_{A_\varepsilon} \left(\sum_{n=1}^{N} a_n f_n \right)^2 dx \leq K_\varepsilon \sum_{n=1}^{N} a_n^2.$$

The sufficiency of (46) is obvious, so the essential (and highly remarkable) statement is the converse: if a sequence $(f_n)_{n \geq 1}$ is a convergence system in measure for

ℓ_2, then, except for a subset of $(0,1)$ with arbitrary small measure, $(f_n)_{n\geq 1}$ behaves like an orthonormal sequence.

In the previous section we discussed the subsequence principle stating that sufficiently thin subsequences of arbitrary sequences of random variables, subject to mild boundedness conditions, satisfy "all" limit theorems for i.i.d. random variables in a mixed (randomized) form. A typical special case of this principle is the following result:

Theorem H (Gaposhkin [132]). *Let $(X_n)_{n\geq 1}$ be a sequence of random variables satisfying*

$$(47) \qquad \sup_n \mathbb{E}X_n^2 < +\infty.$$

Then there exists a subsequence $(X_{n_k})_{k\geq 1}$ together with random variables X and $Y \geq 0$ such that for any further subsequence $(X_{m_k})_{k\geq 1}$ of $(X_{n_k})_{k\geq 1}$ we have

$$(48) \qquad \frac{1}{\sqrt{N}}\sum_{k=1}^{N}(X_{m_k} - X) \xrightarrow{\mathcal{D}} N(0, Y),$$

where $N(0,Y)$ denotes the "variance mixture" normal distribution with characteristic function $\mathbb{E}\exp(-Yt^2/2)$.

If $(X_n^2)_{n\geq 1}$ is uniformly integrable then by well-known compactness results (see e.g. [99]) there exist a subsequence $(X_{m_k})_{k\geq 1}$ and random variables $X \in L_2$ and $Y \in L_{1/2}$, $Y \geq 0$, such that

$$(49) \qquad X_{m_k} \to X \text{ weakly in } L_2, \qquad (X_{m_k} - X)^2 \to Y^2 \text{ weakly in } L_1.^6$$

As Gaposhkin [128] showed, in this case the random variables X, Y in (48) can be chosen as in (49).

In Theorem H, condition (47) is not necessary: simple examples show (see below) that there exist sequences $(X_n)_{n\geq 1}$ of random variables without any finite moments, but having subsequences satisfying (48). The purpose of this section is to give necessary and sufficient conditions for the existence of subsequences $(X_{n_k})_{k\geq 1}$ satisfying the randomized CLT (48), and it will turn out that our conditions have the same character as Nikishin's conditions for the existence of a subsequence being a convergence system, i.e. "nice" behavior of the sequence on subsets of the probability space with measure as close to 1 as we wish.

To formulate our results, call a sequence $(X_n)_{n\geq 1}$ of random variables *nontrivial* if it has no subsequence converging with positive probability. It is easily seen that for non-degenerate sequences the random variable Y in Theorem H is almost surely

---

[6]A sequence $(\xi_n)_{n\geq 1}$ of random variables in L_p, $p \geq 1$, is said to converge weakly to $\xi \in L_p$ if $\mathbb{E}(\xi_n\eta) \to \mathbb{E}(\xi\eta)$ for any $\eta \in L_q$, where $1/p + 1/q = 1$. This type of convergence should not be confused with weak convergence of probability measures and distributions, called generally convergence in distribution, and denoted by $\xrightarrow{\mathcal{D}}$.

positive and Gaposhkin's theorem can be rewritten in a form involving a pure (i.e. not mixed) Gaussian limit distribution.

Theorem J. *Let $(X_n)_{n\geq 1}$ be a nontrivial sequence of random variables satisfying (47). Then there exists a subsequence $(X_{n_k})_{k\geq 1}$ and random variables X, Y with $Y > 0$ such that for all subsequences $(X_{m_k})_{k\geq 1}$ of $(X_{n_k})_{k\geq 1}$ and for any set A in the probability space with $P(A) > 0$ we have*

$$(50) \qquad P_A\left(\frac{\sum_{k=1}^{N}(X_{m_k} - X)}{Y\sqrt{N}} < t\right) \to \Phi(t) \quad \text{for all } t.$$

Here P_A denotes the conditional probability with respect to A, and Φ is the cumulative distribution function of the standard normal distribution.

The nontriviality of $(X_n)_{n\geq 1}$ is assumed here to avoid degenerate cases. If $X_{n_k} \to X$ on some set A with positive probability then for any sufficiently thin subsequence $(X_{m_k})_{k\geq 1}$ of $(X_{n_k})_{k\geq 1}$ we have $\sum |X_{m_k} - X| < +\infty$ a.s. on A, and consequently

$$a_N^{-1} \sum_{k=1}^{N}(X_{m_k} - X) \to 0 \quad \text{a.s. on } A$$

for any norming sequence $a_N \to \infty$ (random or not). Since for any sequence $(X_n)_{n\geq 1}$ satisfying (47) (and in fact any tight sequence $(X_n)_{n\geq 1}$) there is a subsequence $(X_{n_k})_{k\geq 1}$ and a measurable partition $A \cup B$ of the probability space such that X_{n_k} converges on A and is nontrivial on B, there is no loss of generality in assuming that $(X_n)_{n\geq 1}$ is nontrivial.

Clearly, if $(X_n)_{n\geq 1}$ satisfies the conclusion of Theorem J, then so does the sequence $(X_n + 2^{-n}Z)_{n\geq 1}$ for any a.s. finite random variable Z, and thus the assumption (47) is, as stated above, not necessary in Theorem J. Below we will give necessary and sufficient condition for the CLT for lacunary subsequences of a given sequence $(X_n)_{n\geq 1}$ of random variables without any moment assumption on $(X_n)_{n\geq 1}$. To formulate our results, let us note that if all subsequences $(X_{m_k})_{k\geq 1}$ of a sequence $(X_n)_{n\geq 1}$ satisfy (50) for some random variables X, Y, then $(X_n)_{n\geq 1}$ is bounded in probability (see Lemma 2 below). As mentioned in the previous section, every sequence $(X_n)_{n\geq 1}$ of random variables bounded in probability has a subsequence $(X_{n_k})_{k\geq 1}$ which has a limit distribution relative to every set A of the probability space with $P(A) > 0$. Such a sequence was called *determining*. This concept is the same as that of stable convergence, introduced by Rényi [197]; our terminology follows that of functional analysis. Hence in our investigations we can assume without loss of generality that the original sequence $(X_n)_{n\geq 1}$ is determining. Now if $(X_n)_{n\geq 1}$ is determining and F_A denotes its limit distribution relative to the set A, then, as we noted in the previous section, there exists a random measure μ (called the limit random measure of (X_n)) such that

$$(51) \qquad F_A(t) = \mathbb{E}_A(\mu(-\infty, t])$$

for any continuity point t of F_Ω, where $\mathbb{E}_A$ denotes conditional expectation relative to A. Let $F_\bullet$ denote the distribution function of μ; we shall call it the *limit random distribution* of $(X_n)_{n\geq1}$. We can state now our first new theorem.

Theorem 1. *Let $(X_n)_{n\geq1}$ be a nontrivial sequence of random variables. Then the following statements are equivalent:*

A) *There exist a subsequence $(X_{n_k})_{k\geq1}$ and random variables X, Y with $Y > 0$ such that (50) holds for all subsequences $(X_{m_k})_{k\geq1}$ of $(X_{n_k})_{k\geq1}$ and for any set $A \subset \Omega$ with $P(A) > 0$.*

B) *For every $\varepsilon > 0$ there is a subsequence $(X_{n_k})_{k\geq1}$ and a set $A_\varepsilon \subset \Omega$ with $P(A_\varepsilon) \geq 1 - \varepsilon$ such that*

(52)
$$\sup_k \int_{A_\varepsilon} X_{n_k}^2 dP < +\infty.$$

If $(X_n)_{n\geq1}$ is determining, then two further equivalent statements are:

C) *We have*

(53)
$$\int_{-\infty}^{+\infty} x^2 dF_\bullet(x) < +\infty \quad \text{almost surely.}$$

D) *For every $\varepsilon > 0$ there exists a set $A_\varepsilon \subset \Omega$ with $P(A_\varepsilon) \geq 1 - \varepsilon$ such that*

(54)
$$\int_{-\infty}^{+\infty} x^2 dF_{A_\varepsilon}(x) < +\infty.$$

Our second new theorem characterizes sequences $(X_n)_{n\geq1}$ for which (50) holds with $X \in L_2$, $Y \in L_{1/2}$.

Theorem 2. *Let $(X_n)_{n\geq1}$ be a nontrivial sequence of random variables defined on an atomless probability space $(\Omega, \mathcal{F}, P)$. Then the following statements are equivalent:*

A) *There exists a subsequence $(X_{n_k})_{k\geq1}$ and random variables X, Y with $Y > 0$, $X \in L_2$, $Y \in L_{1/2}$ such that (50) holds for all subsequences $(X_{m_k})_{k\geq1}$ of $(X_{n_k})_{k\geq1}$ and all sets $A \subset \Omega$ with $P(A) > 0$.*

B) *There exists a subsequence $(X_{n_k})_{k\geq1}$ and sequences $(Y_k)_{k\geq1}$, $(\tau_k)_{k\geq1}$ of random variables satisfying*

(55)
$$X_{n_k} = Y_k + \tau_k,$$

where

(56)
$$\sup_k \mathbb{E}Y_k^2 < +\infty, \quad \sum_k |\tau_k| < +\infty \qquad a.s.$$

If $(X_n)_{n \geq 1}$ has a limit distribution F, then a third equivalent statement is:

C) We have

$$(57) \qquad \int_{-\infty}^{+\infty} x^2 dF(x) < +\infty.$$

In other words, for the validity of (50) with $X \in L_2$, $Y \in L_{1/2}$, assumption (47) is necessary and sufficient after a small perturbation of $(X_n)_{n \geq 1}$, and for identically distributed $(X_n)_{n \geq 1}$ even this perturbation is not needed. A particularly simple case when $X \in L_2$, $Y \in L_{1/2}$ is satisfied is when X, Y are nonrandom.

A trivial example showing the difference between condition (D) of Theorem 1 and condition (C) of Theorem 2 is the following. Let $\{H_k, k \geq 1\}$ be a partition of the probability space with $P(H_k) = 2^{-k}$ for $k = 1, 2, \ldots$, and let $(X_n)_{n \geq 1}$ be a sequence of random variables on this space which is conditionally i.i.d. given each H_k with mean 0 and variance 2^k. Then $(X_n)_{n \geq 1}$ is nontrivial, determining and clearly satisfies condition (D) of Theorem 1, but since it is identically distributed (in fact exchangeable) and since $\mathbb{E}X_1^2 = +\infty$, condition (C) of Theorem 2 is not satisfied.

9.1. **Some lemmas.** The key for the proof of our theorems is a general structure theorem for lacunary sequences which was proved in [63], and which was stated as Theorem D in the previous section. Furthermore, we need the following lemmas.

Lemma 1. *Let $(X_n)_{n \geq 1}$ be a sequence of random variables such that for some random variables X, Y with $Y > 0$ and for all subsequences $(X_{n_k})_{k \geq 1}$ we have*

$$(58) \qquad \frac{\sum_{k=1}^{N}(X_{n_k} - X)}{Y\sqrt{N}} \xrightarrow{\mathcal{D}} N(0, 1).$$

Then $(X_n)_{n \geq 1}$ is bounded in probability.

Proof. Clearly (58) implies that the sequence $(X_{n_N} - X)/(Y\sqrt{N})$ is bounded in probability as $N \to \infty$, and thus $X_{n_N}/\sqrt{N}$ is bounded in probability for any subsequence $(n_k)_{k \geq 1}$. If $(X_n)_{n \geq 1}$ were not bounded in probability then one could find a subsequence $(m_k)_{k \geq 1}$ and a constant $c > 0$ such that $P(|X_{m_k}| \geq k) \geq c$ for $k = 1, 2, \ldots$, i.e. $X_{m_k}/\sqrt{k}$ would not be bounded in probability, a contradiction. $\square$

Lemma 2. *Let $(X_n)_{n \geq 1}$ be a sequence of random variables and assume that for some random variables X and $Y > 0$ and all sets $A \subset \Omega$ with $P(A) > 0$ we have*

$$(59) \qquad P_A\left(\frac{\sum_{k=1}^{N}(X_k - X)}{Y\sqrt{N}} < t\right) \to \Phi(t) \quad \text{for all } t.$$

Assume further that (59) *remains valid if we replace* X, Y *by some random variables* X^* *and* $Y^* > 0$. *Then* $X = X^*$ *a.s. and* $Y = Y^*$ *a.s.*

Proof. From the assumption it follows that the sequences

$$N^{-1/2} \sum_{k=1}^{N} (X_k - X) \qquad \text{and} \qquad N^{-1/2} \sum_{k=1}^{N} (X_k - X^*)$$

are bounded in probability, and thus the same holds for their difference $\sqrt{N}(X - X^*)$, whence $X = X^*$ a.s. To prove $Y = Y^*$, fix $c > 1$ and set $A = \{Y^* \geq cY\}$. If $P(A) > 0$ then clearly we cannot have both (59) and the analogous relation with Y replaced by Y^*. Thus $P(A) > 0$ for all $c > 1$ whence $Y^* \leq Y$ a.s. The same argument yields $Y \leq Y^*$ a.s., completing the proof. $\qquad\square$

Lemma 3. *Let* $X_1, X_2, \ldots, X_n$ *be i.i.d. random variables with distribution function* F *and set* $S_n = X_1 + \cdots + X_n$. *Then for any* $t > 0$ *we have*

$$(60) \qquad P(|S_n| \leq 2t) \leq A \frac{t}{\sqrt{n}} \left(\int_{|x| \leq t} x^2 dF(x) - 2 \left(\int_{|x| \leq t} x dF(x) \right)^2 \right)^{-1/2},$$

provided the difference on the right-hand side is positive and $\int_{|x| \leq t} dF(x) \geq 1/2$. *Here* A *is an absolute constant.*

Proof. Let F^* denote the distribution function obtained from F by symmetrization. From a well-known concentration function inequality of Esseen [106, Theorem 2] it follows that the left-hand side of (60) cannot exceed

$$A_1 \frac{t}{\sqrt{n}} \left(\int_{|x| \leq 2t} x^2 dF^*(x) \right)^{-1/2},$$

where A_1 is an absolute constant. Hence to prove (60) it suffices to show that $\int_{|x| \leq t} dF(x) \geq 1/2$ implies

$$(61) \qquad \int_{|x| \leq 2t} x^2 dF^*(x) \geq \int_{|x| \leq t} x^2 dF(x) - 2 \left(\int_{|x| \leq t} x dF(x) \right)^2.$$

Let ξ and η be independent random variables with distribution function F, and set

$$C = \{|\xi - \eta| \leq 2t\}, \qquad D = \{|\xi| \leq t, |\eta| \leq t\}.$$

Then

$$\int_{|x| \leq 2t} x^2 dF^*(x) = \int_C (\xi - \eta)^2 dP$$

$$\geq \int_D (\xi - \eta)^2 dP$$

$$= 2 \int_{|\xi| \leq t} \xi^2 dP \cdot P(|\eta| \leq t) - 2 \left(\int_{|\xi| \leq t} \xi dP \right)^2$$

$$\geq \int_{|\xi| \leq t} \xi^2 dP - 2 \left(\int_{|\xi| \leq t} \xi dP \right)^2,$$

provided $P(|\eta| \leq t) \geq 1/2$. Thus (61) is valid. $\qquad \square$

Lemma 4. *Let $(\Omega, \mathcal{F}, P)$ be an atomless probability space and $X_1, X_2, \ldots$ a sequence of random variables on $(\Omega, \mathcal{F}, P)$ with limit distribution F. Then there exist a subsequence $(X_{n_k})_{k \geq 1}$ and sequences $(Y_k)_{k \geq 1}$ and $(\tau_k)_{k \geq 1}$ of random variables on $(\Omega, \mathcal{F}, P)$ such that $X_{n_k} = Y_k + \tau_k$, $k = 1, 2, \ldots$, such that the random variables Y_k have distribution function F, and such that $\sum_k |\tau_k| < +\infty$ a.s.*

Proof. Let $(\hat{X}_n)_{n \geq 1}$ be discrete random variables such that $P(|X_n - X'_n| \geq 2^{-n}) \leq 2^{-n}$, $n = 1, 2, \ldots$, and denote by F_n the distribution function of X_n. Clearly $F_n \to F$ and thus $\varepsilon_n := \varrho(F_n, F) \to 0$, where ϱ denotes the Prohorov distance. By a theorem of Strassen [8] there exists a probability measure μ_n on $\mathbb{R}^2$ with marginals F_n and F such that

$$\mu_n((x, y) : |x - y| \geq \varepsilon_n) \leq \varepsilon_n.$$

Let c be a possible value of $\hat{X}_n$. Since the probability space restricted to $A = \{\hat{X}_n = c\}$ is atomless, there exists a random variable V_n on this space such that

$$P_A(V_n < t) = \frac{\mu_n((x, y) : x = c, y < t)}{\mu_n((x, y) : x = c)}$$

for all t. Carrying out this construction for all possible values of c in the range of $\hat{X}_n$, we get a random variable V_n defined on the whole probability space such that the joint distribution of $\hat{X}_n$ and V_n is μ_n. Clearly the distribution of V_n is F and $P(|\hat{X}_n - V_n| \geq \varepsilon_n) \leq \varepsilon_n$. Choosing $(n_k)_{k \geq 1}$ so that $\varepsilon_{n_k} \leq 2^{-k}$ we get

$$P\big(|\hat{X}_{n_k} - V_{n_k}| \geq 2^{-k}\big) \leq 2^{-k},$$

i.e.

$$P\big(|X_{n_k} - V_{n_k}| \geq 2 \cdot 2^{-k}\big) \leq 2 \cdot 2^{-k}$$

and thus $\sum_k |X_{n_k} - V_{n_k}| < +\infty$ a.s. by the Borel–Cantelli lemma. Thus the decomposition $X_{n_k} = Y_k + \tau_k$, where $Y_k = V_{n_k}$ and $\tau_k = X_{n_k} - V_{n_k}$, satisfies the requirements. $\qquad \square$

Our final two lemmas concern the properties of the limit random distribution of determining sequences.

Lemma 5. *Let $(X_n)_{n\geq 1}$ be a determining sequence of random variables with limit random distribution F. Then for any set $A \subset \Omega$ with $P(A) > 0$ we have*

$$(62) \qquad \mathbb{E}_A\left(\int_{-\infty}^{+\infty} x^2 dF_\bullet(x)\right) = \int_{-\infty}^{+\infty} x^2 dF_A(x),$$

in the sense that if one side is finite then the other side is also finite and the two sides are equal. The statement remains valid if in (62) we replace the interval of integrations by $(-t, t)$, provided t and $-t$ are continuity points of F_A.

Proof. This lemma follows easily from (51) by integration by parts. $\qquad\square$

Lemma 6. *Let $X, X_1, X_2, \dots$ be random variables such that both sequences $(X_n)_{n\geq 1}$ and $(X_n - X)_{n\geq 1}$ are determining; let $F_\bullet$ and $G_\bullet$ denote, respectively, their limit random distributions. Then $\int_{-\infty}^{+\infty} x^2 dG_\bullet(x) < +\infty$ a.s. implies $\int_{-\infty}^{+\infty} x^2 dF_\bullet(x) < +\infty$ a.s. and conversely.*

Proof. Let $\varepsilon > 0$ and choose a set $A \subset \Omega$ such that $P(A) \geq 1 - \varepsilon$ and on A both X and $\int_{-\infty}^{+\infty} x^2 dF_\bullet(x)$ are bounded. Let F_A and G_A denote the limit random distribution of $(X_n)_{n\geq 1}$ resp. $(X_n - X)_{n\geq 1}$ relative to A. Replacing X_n by $X_n + \tau_n$ where $\tau_n \to 0$ a.s. clearly does not change the limit distributions $F_A, G_A, F_\bullet, G_\bullet$, and thus by passing to a subsequence and using Lemma 4 we can assume, without loss of generality, that the X_n are identically distributed on A. Then

$$\mathbb{E}_A X_1^2 = \mathbb{E}_A X_2^2 = \cdots = \int_{-\infty}^{+\infty} x^2 dF_A(x),$$

where the last integral is finite by the boundedness of $\int_{-\infty}^{+\infty} x^2 dF_\bullet(x)$ on A and Lemma 5. By Minkowski's inequality and the boundedness of X on A it follows that $\mathbb{E}_A((X_n - X)^2)$ is also bounded, and thus Fatou's lemma implies that

$$\int_{-\infty}^{+\infty} x^2 dG_A(x) \leq \liminf_{n\to\infty} \mathbb{E}_A\left((X_n - X)^2\right) < +\infty.$$

Using Lemma 5 again it follows that $\int_{-\infty}^{+\infty} x^2 dG_\bullet(x) < +\infty$ a.s. on A. As the measure of A can be chosen arbitrarily close to 1, we get $\int x^2 dG_\bullet(x) < +\infty$ a.s., as required. $\qquad\square$

9.2. **Proof of the theorems.** We begin with the proof of Theorem 1. Using diagonalization and Chebyshev's inequality it follows that if a sequence $(X_n)_{n\geq 1}$ satisfies (B), then it has a subsequence bounded in probability and thus also a determining subsequence. By Lemma 1 the same conclusion holds if $(X_n)_{n\geq 1}$ satisfies (A). Thus

to prove our theorem it suffices to prove the equivalence of (A), (B), (C), (D) for determining sequences $(X_n)_{n \geq 1}$. In what follows we shall prove the implications (A) $\Longrightarrow$ (C) $\Longrightarrow$ (D) $\Longrightarrow$ (B); since (B) $\Longrightarrow$ (A) follows easily from Theorem D in the previous section by diagonalization, this will prove Theorem 1.

Assume that $(X_n)_{n \geq 1}$ is a determining sequence satisfying (A), i.e. there exists a subsequence $(X_{n_k})_{k \geq 1}$ and random variables X, Y with $Y > 0$ such that for any further subsequence $(X_{m_k})_{k \geq 1}$ of $(X_{n_k})_{k \geq 1}$ and any set $A \subset \Omega$ with $P(A) > 0$ we have

$$(63) \qquad P_A \left(\frac{\sum\limits_{k=1}^{N} (X_{m_k} - X)}{Y \sqrt{N}} < t \right) \to \Phi(t) \quad \text{for all } t.$$

We claim that $(X_n)_{n \geq 1}$ satisfies (C). Clearly we can assume without loss of generality that $(X_{n_k})_{k \geq 1} = (X_k)_{k \geq 1}$ and since $(X_n - X)_{k \geq 1}$ contains a determining subsequence, we can assume also that $(X_n - X)_{n \geq 1}$ itself is determining. Moreover, since $(X_n - X)_{n \geq 1}$ satisfies (C) if and only if $(X_n)_{n \geq 1}$ does (see Lemma 6), we can assume that $X = 0$. Assume indirectly that $(X_n)_{n \geq 1}$ does not satisfy (C), i.e. there exists a set $B \subset \Omega$ with $P(B) > 0$ such that

$$(64) \qquad \lim_{t \to \infty} \int_{-t}^{t} x^2 dF_\bullet(x) = +\infty \quad \text{on } B.$$

Then there exists a set $B^* \subset B$ with $P(B^*) \geq P(B)/2$ such that on B^* the random variable Y is bounded and (63) holds uniformly, i.e. there exists a constant $K > 0$ and a numerical sequence $K_t \to +\infty$ such that

$$\int_{-t}^{t} x^2 dF_\bullet(x) \geq K_t \ \text{ and } \ Y \leq K \ \text{ on } \ B^*.$$

Also, $1 - F_\bullet(t) + F_\bullet(-t) \to 0$ a.s. as $t \to \infty$, and thus we can choose a set $B^{**} \subset B^*$ with $P(B^{**}) \geq P(B^*)/2$ such that on B^{**} the last convergence relation holds uniformly, i.e. there exists a positive numerical sequence $\varepsilon_t \searrow 0$ such that

$$(65) \qquad\qquad 1 - F_\bullet(t) + F_\bullet(-t) \leq \varepsilon_t \ \text{ on } \ B^{**}.$$

We show that there exists a subsequence $(X_{m_k})_{k \geq 1}$ of $(X_{n_k})_{k \geq 1}$ such that (63) fails for $A = B^{**}$. Since our argument will involve the sequence $(X_n)_{n \geq 1}$ only on the set B^{**}, in the sequel we can assume, without loss of generality, that $B^{**} = \Omega$. That is, we may assume that (65) holds on the whole probability space.

Let C be an arbitrary set in the probability space with $P(C) > 0$. Integrating (65) and using (51) and Lemma 5 we get

$$(66) \qquad \int_{-t}^{t} x^2 dF_C(x) \geq K_t, \quad 1 - F_C(t) + F_C(-t) \leq \varepsilon_t, \quad t \in H,$$

where H denotes the set of continuity points of F_Ω. Choose $t_0 \in H$ so large that $\varepsilon_{t_0} \leq 1/16$ and then choose t_1 so large that

$$K_t^{1/2}/4 \geq 2t_0^2 \ \text{ for } \ t \geq t_1, \ t \in H.$$

Then for $t \geq t_1$, $t \in H$ we have, using the second relation of (66),

$$\left| \int_{-t}^{t} x dF_C(x) \right| \leq 2t_0^2 + \int_{t_0 \leq |x| \leq t} |x| dF_C(x)$$

$$\leq 2t_0^2 + \left(\int_{|x| \geq t_0} dF_C(x) \right)^{1/2} \left(\int_{|x| \leq t} x^2 dF_C(x) \right)^{1/2}$$

$$\leq 2t_0^2 + \frac{1}{4} \left(\int_{|x| \leq t} x^2 dF_C(x) \right)^{1/2}$$

$$\leq \frac{1}{2} \left(\int_{|x| \leq t} x^2 dF_C(x) \right)^{1/2},$$

and thus we proved that for any $C \subset \Omega$ with $P(c) > 0$ we have

$$(67) \qquad \int_{-t}^{t} x^2 dF_C(x) - 2 \left(\int_{-t}^{t} x dF_C(x) \right)^2 \geq \frac{1}{2} K_t, \qquad t \geq t_1, \ t \in H.$$

Since $(X_n)_{n \geq 1}$ is bounded in probability, there exists a function $\psi(x) \nearrow \infty$ such that

$$(68) \qquad \sup_n E\psi(X_n) \leq 1$$

(see [9]). Let $(a_k)_{k \geq 1}$ be a sequence of integers tending to $+\infty$ so slowly that $a_k \leq \log k$ and

$$(69) \qquad \delta_k := a_k/\psi(k^{1/4}) \to 0.$$

Let further $(\varepsilon_n)_{n \geq 1}$ tend to 0 so rapidly that $\varepsilon_{a_k} \leq 2^{-k}$. By Theorem D there exists a subsequence $(X_{m_k})_{k \geq 1}$ and a sequence $(Y_k)_{k \geq 1}$ of discrete random variables such

that (41) holds and for each $k > 1$ the atoms of the finite σ-field $\sigma\{Y_1, \ldots, Y_{a_k}\}$ can be divided into two classes Γ_1 and Γ_2 such that

$$(70) \qquad \sum_{B \in \Gamma_1} P(B) \leq \varepsilon_{a_k} \leq 2^{-k}$$

and for each $B \in \Gamma_2$ there exist i.i.d. random variables $Z_{a_k+1}^{(B)}, \ldots, Z_k^{(B)}$ on B with distribution F_B such that

$$(71) \qquad P_B\left(|Y_j - Z_j^{(B)}| \geq 2^{-k}\right) \leq 2^{-k}, \qquad j = a_k + 1, \ldots, k.$$

We show that

$$(72) \qquad P\left(\frac{\sum_{k=1}^{N} X_{m_k}}{Y\sqrt{N}} < t\right) \to \Phi(t) \quad \text{for all } t$$

cannot hold; this will complete our indirect proof of (A) $\Longrightarrow$ (C). Set

$$S_{a_k,k}^{(B)} = \sum_{j=a_k+1}^{k} Z_j^{(B)}, \qquad B \in \Gamma_2,$$

$$\overline{S}_{a_k,k} = \sum_{B \in \Gamma_2} S_{a_k,k}^{(B)} \mathbb{1}_B,$$

where $\mathbb{1}_B$ denotes the indicator function of B. By (71),

$$P_B\left(\left|\sum_{j=a_k+1}^{k} Y_j - \sum_{j=a_k+1}^{k} Z_j^{(B)}\right| \geq 1\right) \leq 2^{-a_k}, \qquad B \in \Gamma_2,$$

and thus using (70) we get

$$(73) \qquad P\left(\left|\sum_{j=a_k+1}^{k} Y_j - \overline{S}_{a_k,k}\right| \geq 1\right) \leq 2 \cdot 2^{-k}.$$

By (68), (69) and the Markov inequality we have

$$P\left(\left|\sum_{j=1}^{a_k} X_{m_j}\right| \geq a_k k^{1/4}\right) \leq a_k \sup_{1 \leq j \leq a_k} P\left(|x_{m_j}| \geq k^{1/4}\right)$$
$$\leq a_k \psi(k^{1/4})^{-1}$$
$$= \delta_k,$$

which, together with (73) and (41), yields

$$(74) \qquad P\left(\left|\sum_{j=1}^{k} X_{m_j} - \overline{S}_{a_k,k}\right| \geq 2_{a_k} k^{1/4}\right) \leq 3 \cdot 2^{-a_k} + \delta_k.$$

Applying Lemma 3 to the i.i.d. sequence $\{Z_j^{(B)}, a_k + 1 \leq j \leq k\}$ and using (67) we get

$$P_B\left(\left|\frac{S_{a_k,k}^{(B)}}{\sqrt{k}}\right| \leq 1\right) \leq P_B\left(\frac{S_{a_k,k}^{(B)}}{\sqrt{k-a_k}} \leq 2\right)$$

$$\leq \text{const.} K_{\sqrt{k}}^{-1/2},$$

where the constant is absolute. Thus using (70) and $Y \leq K$ it follows that

(75) $$P\left(\left|\frac{\overline{S}_{a_k,k}}{Y\sqrt{k}}\right| \leq \frac{1}{K}\right) \leq \text{const.} K_{\sqrt{k}}^{-1/2} + 2^{-k}.$$

If (72) were true then by (74) and $a_k \leq \log k$ we would also have

$$\overline{S}_{k,a_k}/Y\sqrt{k} \xrightarrow{\mathcal{D}} N(0,1),$$

which clearly contradicts (75) for large k, since the right-hand side tends to zero. This completes the proof of (A) $\Longrightarrow$ (C).

The remaining implications (C) $\Longrightarrow$ (D) and (D) $\Longrightarrow$ (B) of Theorem 1 are easy. Assume first that (C) holds, then for any $\varepsilon > 0$ there exists a set $A \subset \Omega$ with $P(A) \geq 1 - \varepsilon$ and a constant $K = K_\varepsilon$ such that

$$\int_{-\infty}^{+\infty} x^2 dF_\bullet(x) \leq K \quad \text{on} \quad A.$$

Integrating the last relation on A and using Lemma 5 we get (52), i.e. (D) holds. Assume now that (D) holds, i.e. for any $\varepsilon > 0$ there exists a set $A \subset \Omega$ with $P(A) \geq 1 - \varepsilon$ such that (52) is valid. Applying Lemma 4 for $(X_n)_{n\geq 1}$ on the set A it follows that there exists a subsequence $(X_{n_k})_{k\geq 1}$ and random variables Y_k and τ_k, $k = 1, 2, \ldots$, defined on A such that $X_{n_k} = Y_k + \tau_k$, $k = 1, 2, \ldots$, and such that the random variables Y_k have distribution F_A on A and $\tau_k \to 0$ a.s. on A. Choose a set $B \subset A$ with $P(B) \geq 1 - 2\varepsilon$ such that $\tau_k \to 0$ uniformly on B. Then clearly $(\tau_n)_{n\geq 1}$ is uniformly bounded on B, and further

$$\int_B Y_k^2 dP \leq \int_A Y_k^2 dP$$

$$= P(A)\int_{-\infty}^{+\infty} x^2 dF_A(x)$$

$$\leq \int_{-\infty}^{+\infty} x^2 dF_A(x) < +\infty$$

for each $k \geq 1$ by the identical distribution of the Y_k's and (52). Thus on B the sequences $(Y_k)_{k\geq 1}$ and $(\tau_k)_{k\geq 1}$ have bounded L_2 norms and thus the same holds for

$X_{m_k} = Y_k + \tau_k$, i.e.

$$\sup_k \int_B X^2_{m_k} dP < +\infty.$$

In view of $P(B) \geq 1 - 2\varepsilon$ this shows that $(X_n)_{n\geq1}$ satisfies statement (B). This completes the proof of Theorem 1.

Proof of Theorem 2. Theorem 2 follows from Theorem 1 and a slightly sharper form of Theorems H and J which was proved in [147]. We already mentioned the fact that in Theorem H the random variables X, Y appearing in (50) actually satisfy $X \in L_2$, $Y \in L_{1/2}$. Moreover, if instead of (47) we make the slightly stronger assumption that the sequence $(X^2_n)_{n\geq1}$ is uniformly integrable then by the weak compactness criteria in L_1 and L_2 it follows that there exists a subsequence (X_{n_k}) and random variables $X \in L_2$, $Y \in L_{1/2}$ such that

(76) $X_{n_k} \to X$ weakly in L_2, $(X_{n_k} - X)^2 \to Y^2$ weakly in L_1.

As is shown in [163], in this case (50) holds with the random variables X, Y determined by (76). We turn now to the proof of Theorem 2. As in the case of Theorem 1, it suffices to prove the equivalence of statements (A), (B), (C) in the case when $(X_n)_{n\geq1}$ is determining. Also, since replacing X_n by $X_n + \tau_n$ where $\sum |\tau_n| < +\infty$ a.s. does not affect the validity of (50), the conclusion (B) $\implies$ (A) of Theorem (2) is contained in the stronger form of Theorem H mentioned above. Thus it suffices to verify the implications (A) $\implies$ (C) and (C) $\implies$ (B). To prove (A) $\implies$ (C) let us assume that $(X_n)_{n\geq1}$ is determining with limit distribution F, and that there exist a subsequence $(X_{n_k})_{k\geq1}$ and random variables $X \in L_2$, $Y \in L_{1/2}$, $Y > 0$, such that for all subsequences $(X_{m_k})_{k\geq1}$ of $(X_{n_k})_{k\geq1}$ and any set $A \subset \Omega$ with $P(A) > 0$ equation (50) holds. We show $\int_{-\infty}^{+\infty} x^2 dF(x) < +\infty$. Clearly we can assume without loss of generality that $(X_{n_k})_{k\geq1} < (X_k)_{k\geq1}$. Fix $\varepsilon > 0$. By the implication (A) $\implies$ (B) $\implies$ (D) of Theorem 1 there is a set $A \subset \Omega$ with $P(A) \geq 1 - \varepsilon$ and a subsequence $(X_{n_k})_{k\geq1}$ such that

(77) $\sup_k \int_A X^2_{n_k} dP < +\infty$ and $\int_A x^2 dF_A(x) < +\infty,$

where F_A is the limit distribution of $(X_n)_{n\geq1}$ on A. Applying Lemma 4 for $(X_{n_k})_{k\geq1}$ on A it follows that there exists a subsequence $(X_{m_k})_{k\geq1}$ of $(X_{n_k})_{k\geq1}$ admitting the decomposition

(78) $X_{m_k} = Y_k + \tau_k$ on A,

where the Y_k are identically distributed on A with distribution function F_A and $\sum |\tau_k| < +\infty$ a.s. on A. Being an identically distributed sequence with finite expectation, the sequence $(Y^2_k)_{k\geq1}$ is uniformly integrable on A, and thus by the sharper form of Theorem H mentioned above it follows that there exists a subsequence $(Y_{p_k})_{k\geq1}$ of $(Y_k)_{k\geq1}$ such that

$Y_{p_k} \to U$ weakly in $L_2(A)$, $(Y_{p_k} - U)^2 \to V^2$ weakly in $L_1(A),$

and for any $B \subset A$ with $P(B) > 0$ we have

$$P_B \left(\frac{\sum_{k=1}^{N}(Y_{p_k} - U)}{V\sqrt{N}} < t \right) \to \Phi(t) \ \text{ for all } t,$$

where U, V are random variables such that $U \in L_2(A)$, $V \in L_{1/2}(A)$, $V > 0$. Thus by (78) and $\sum |\tau_k| < +\infty$ a.s. on A we get

$$P_B \left(\frac{\sum_{k=1}^{N}(X_{m_{p_k}} - U)}{V\sqrt{N}} < t \right) \to \Phi(t) \ \text{ for all } t$$

for any $B \subset A$ with $P(B) > 0$. Comparing with (50) and using Lemma 2 we get $U = X$, $V = Y$ a.s. on A, and thus we proved that

$$Y_{p_k} \to X \text{ weakly in } L_2(A), \qquad (Y_{p_k} - X)^2 \to Y^2 \text{ weakly in } L_1(A).$$

Hence

$$\mathbb{E}_A Y^2 = \lim_{k \to \infty} \mathbb{E}_A (Y_{p_k} - X)^2 = \lim_{k \to \infty} (\mathbb{E}_A Y_{p_k}^2 - 2\mathbb{E}_A Y_{p_k} X + \mathbb{E}_A X^2)$$

$$(79) \qquad = \lim_{k \to \infty} \mathbb{E}_A Y_{p_k}^2 - \mathbb{E}_A X^2 = \int_{-\infty}^{+\infty} x^2 dF_A(x) - \mathbb{E}_A X^2,$$

where in the last step we used the fact that the Y_k's have distribution F_A on A. Hence

$$(80) \qquad P(A)^{-1} \int_A Y^2 dP = \int_{-\infty}^{+\infty} x^2 dF_A(x) - P(A)^{-1} \int_A X^2 dP.$$

Since $X \in L_2(\Omega)$, $Y^2 \in L_1(\Omega)$, the left-hand side of (80) and the second term on the right-hand side approach finite limits as $P(A) \to 1$ and thus $\int_{-\infty}^{+\infty} x^2 dF_A(x)$ also converges to a finite limit. On the other hand, $F_A \to F$ as $P(A) \to 1$ and thus Fatou's lemma implies

$$\int_{-\infty}^{+\infty} x^2 dF(x) \le \liminf_{P(A) \to 1} \int_{-\infty}^{+\infty} x^2 dF_A(x) < +\infty,$$

proving the implication (A) $\implies$ (C). Now if (C) holds then by Lemma 4 there exists a subsequence $(X_{n_k})_{k \ge 1}$ permitting the decomposition (55) where $\sum |\tau_k| < +\infty$ a.s. and Y_k are identically distributed with distribution F; since F has finite variance by (C), the first relation of (56) holds. Thus $(X_n)_{n \ge 1}$ satisfies (B) and the proof of Theorem 2 is completed.

ACKNOWLEDGMENTS

References

[1] R. Adler, M. Keane, and M. Smorodinsky. A construction of a normal number for the continued fraction transformation. *J. Number Theory*, 13(1):95–105, 1981.

[2] D. Airey and B. Mance. Normality of different orders for Cantor series expansions. *Nonlinearity*, 30(10):3719–3742, 2017.

[3] D. Airey, B. Mance, and J. Vandehey. Normal number constructions for Cantor series with slowly growing bases. *Czechoslovak Math. J.*, 66(141)(2):465–480, 2016.

[4] C. Aistleitner. Convergence of $\sum c_k f(kx)$ and the Lip α class. *Proc. Amer. Math. Soc.*, 140(11):3893–3903, 2010.

[5] C. Aistleitner. Diophantine equations and the LIL for the discrepancy of sublacunary sequences. *Illinois J. Math.*, 53(3):785–815, 2009.

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

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

[8] C. Aistleitner. On the law of the iterated logarithm for the discrepancy of lacunary sequences II. *Trans. Amer. Math. Soc.*, 365(7):3713–3728, 2013.

[9] C. Aistleitner. Quantitative uniform distribution results for geometric progressions. *Israel J. Math.*, 204(1):155–197, 2014.

[10] C. Aistleitner. On some questions of V. I. Arnold on the stochasticity of geometric and arithmetic progressions. *Nonlinearity*, 28(10):3663–3675, 2015.

[11] C. Aistleitner, S. Baker, N. Technau, and N. Yesha. Gap statistics and higher correlations for geometric progressions modulo one. *Math Ann.*, 385(1-2):845-861, 2023.

[12] C. Aistleitner, V. Becher, A.-M. Scheerer, and T. A. Slaman. On the construction of absolutely normal numbers. *Acta Arith.*, 180(4):333–346, 2017.

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

[14] C. Aistleitner and I. Berkes. Limit distributions in metric discrepancy theory. *Monatsh. Math.*, 169(3-4):253–265, 2013.

[15] C. Aistleitner, I. Berkes, and K. Seip. GCD sums from Poisson integrals and systems of dilated functions. *J. Eur. Math. Soc. (JEMS)*, 17(6):1517–1546, 2015.

[16] C. Aistleitner, I. Berkes, K. Seip and M. Weber. Convergence of series of dilated functions and spectral norms of GCD matrices. *Acta Arithmetica* 168:221–246, 2015.

[17] C. Aistleitner, I. Berkes, and R. Tichy. On the asymptotic behavior of weakly lacunary series. *Proc. Amer. Math. Soc.* 139(7):2505–2517, 2011.

[18] C. Aistleitner, I. Berkes, and R. F. Tichy. Lacunary sequences and permutations. In *Dependence in probability, analysis and number theory*, pages 35–49. Kendrick Press, Heber City, UT, 2010.

[19] C. Aistleitner, I. Berkes, and R. F. Tichy. On permutations of Hardy-Littlewood-Pólya sequences. *Trans. Amer. Math. Soc.* 363(12):6219–6244, 2011.

[20] C. Aistleitner, I. Berkes, and R. F. Tichy. On permutations of lacunary series. *RIMS Kôkyûroku Bessatsu*, B34 (2012), 1–25.

[21] C. Aistleitner and C. Elsholtz. The central limit theorem for subsequences in probabilistic number theory. *Canad. J. Math.* 64(6):1201–1221, 2012.

[22] C. Aistleitner, N. Gantert, Z. Kabluchko, J. Prochno, and K. Ramanan. Large deviation principles for lacunary sums. *Trans. Amer. Math. Soc.* 376(1):507–553, 2023.

[23] C. Aistleitner, K. Fukuyama. On the law of the iterated logarithm for trigonometric series with bounded gaps. *Probab. Theory Related Fields* 154(3-4):607–620, 2012.

[24] C. Aistleitner, K. Fukuyama. On the law of the iterated logarithm for trigonometric series with bounded gaps II. *J. Théor. Nombres Bordeaux* 28(2):391–416, 2016.

[25] C. Aistleitner, K. Fukuyama, and Y. Furuya. Optimal bound for the discrepancies of lacunary sequences. *Acta Arith.* 158(3):229–243, 2013.

[26] M. Ajtai, J. Komlós, E. Szemerédi. A dense infinite Sidon sequence. *European J. Combin.* 2(1):1-11, 1981.

[27] D. J. Aldous. Limit theorems for subsequences of arbitrarily-dependent sequences of random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 40(1):59–82, 1977.

[28] D. J. Aldous. Subspaces of L^1, via random measures. *Trans. Amer. Math. Soc.* 267(2):445–463, 1981.

[29] J. I. Almarza and S. Figueira. Normality in non-integer bases and polynomial time randomness. *J. Comput. System Sci.* 81(7):1059–1087, 2015.

[30] F. Amoroso and E. Viada. Small points on subvarieties of a torus. *Duke Math. J.* 150(3):407–442, 2009.

[31] J. Angst and G. Poly. Variations on Salem–Zygmund results for random trigonometric polynomials. Application to almost sure nodal asymptotics. *Electron. J. Probab.* 26:1–36, 2021.

[32] V. I. Arnol'd. Stochastic and deterministic statistics of orbits in chaotically looking dynamical systems. *Tr. Mosk. Mat. Obs.* 70:46–101, 2009.

[33] D. H. Bailey and J. M. Borwein. Pi Day is upon us again and we still do not know if pi is normal. *Amer. Math. Monthly* 121(3):191–206, 2014.

[34] D. H. Bailey, J. M. Borwein, R. E. Crandall, and C. Pomerance. On the binary expansions of algebraic numbers. *J. Théor. Nombres Bordeaux* 16(3):487–518, 2004.

[35] D. H. Bailey and R. E. Crandall. On the random character of fundamental constant expansions. *Experiment. Math.* 10(2):175–190, 2001.

[36] D. H. Bailey and R. E. Crandall. Random generators and normal numbers. *Experiment. Math.* 11(4):527–546, 2002.

[37] R. C. Baker. Khinchin's conjecture and Marstrand's theorem. *Mathematika* 21:248–260, 1974.

[38] R. C. Baker. Metric number theory and the large sieve. *J. London Math. Soc. (2)* 24(1):34–40, 1981.

[39] S. Baker and D. Kong. Numbers with simply normal β-expansions. *Math. Proc. Cambridge Philos. Soc.* 167(1):171–192, 2019.

[40] P. Balister, B. Bollobás, R. Morris, J. Sahasrabudhe, and M. Tiba. Flat Littlewood polynomials exist. *Ann. of Math. (2)* 192(3):977–1004, 2020.

[41] A. Bazarova, I. Berkes, and M. Raseta. On trigonometric sums with random frequencies. *Studia Sci. Math. Hungar.* 55(1):141–152, 2018.

[42] A. Bazarova, I. Berkes, and M. Raseta. Strong approximation of lacunary series with random gaps. *Monatsh. Math.* 186(3):393–403, 2018.

[43] V. Becher and S. Figueira. An example of a computable absolutely normal number. *Theoret. Comput. Sci.* 270(1-2):947–958, 2002.

[44] V. Becher, S. Figueira, and R. Picchi. Turing's unpublished algorithm for normal numbers. *Theoret. Comput. Sci.* 377(1-3):126–138, 2007.

[45] V. Becher, P. A. Heiber, and T. A. Slaman. A polynomial-time algorithm for computing absolutely normal numbers. *Inform. and Comput.* 232:1–9, 2013.

[46] V. Becher, P. A. Heiber, and T. A. Slaman. Normal numbers and the Borel hierarchy. *Fund. Math.* 226(1):63–78, 2014.

[47] V. Becher, P. A. Heiber, and T. A. Slaman. A computable absolutely normal Liouville number. *Math. Comp.* 84(296):2939–2952, 2015.

[48] V. Becher and T. A. Slaman. On the normality of numbers to different bases. *J. Lond. Math. Soc. (2)* 90(2):472–494, 2014.

[49] V. Becher and S. A. Yuhjtman. On absolutely normal and continued fraction normal numbers. *Int. Math. Res. Not. IMRN* (19):6136–6161, 2019.

[50] J. Beck. From Khinchin's conjecture on strong uniformity to superuniform motions. *Mathematika* 61(3):591–707, 2015.

[51] L. Berggren, J. Borwein, and P. Borwein. *Pi: a source book*. Springer-Verlag, New York, third edition, 2004.

[52] I. Berkes. The law of the iterated logarithm for subsequences of random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 30:209–215, 1974.

[53] I. Berkes. On the asymptotic behaviour of $\sum f(n_k x)$. I. Main theorems. II. Applications. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 34(4):319–345 & 347–365, 1976.

[54] I. Berkes. On the central limit theorem for lacunary trigonometric series. *Analysis Math.* 4:159-180, 1978.

[55] I. Berkes. A central limit theorem for trigonometric series with small gaps. *Z. Wahrsch. Verw. Gebiete* 47(2):157–161, 1979.

[56] I. Berkes. An extension of the Komlós subsequence theorem. *Acta Math. Hungar.* 55(1-2):103–110, 1990.

[57] I. Berkes. A note on lacunary trigonometric series. *Acta Math. Hungar.* 57(1-2):181–186, 1991.

[58] I. Berkes. Non-Gaussian limit distributions of lacunary trigonometric series. *Canad. J. Math.* 43(5):948–959, 1991.

[59] I. Berkes. Critical LIL behavior of the trigonometric system. *Trans. Amer. Math. Soc.* 338(2):553–585, 1993.

[60] I. Berkes. On the convergence of $\sum c_n f(nx)$ and the Lip 1/2 class. *Trans. Amer. Math. Soc.* 349(10):4143–4158, 1997.

[61] I. Berkes. On the uniform theory of lacunary series. In *Number theory—Diophantine problems, uniform distribution and applications*, pages 137–167. Springer, Cham, 2017.

[62] I. Berkes and B. Borda. On the law of the iterated logarithm for random exponential sums. *Trans. Amer. Math. Soc.* 371(5):3259–3280, 2019.

[63] I. Berkes and E. Péter. Exchangeable random variables and the subsequence principle. *Probab. Theory Related Fields* 73(3):395–413, 1986.

[64] I. Berkes and W. Philipp. The size of trigonometric and Walsh series and uniform distribution mod 1. *J. London Math. Soc. (2)* 50(3):454–464, 1994.

[65] 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(1-4):107–145, 2006.

[66] I. Berkes and H. P. Rosenthal. Almost exchangeable sequences of random variables. *Z. Wahrsch. Verw. Gebiete* 70(4):473–507, 1985.

[67] I. Berkes and R. F. Tichy. The Kadec-Pełczyński theorem in L^p, $1 \leq p < 2$. *Proc. Amer. Math. Soc.* 144(5):2053–2066, 2016.

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

[69] I. Berkes and M. Weber. On series of dilated functions. *Quarterly J. Math.* 65:25–52, 2014.

[70] I. Berkes and M. Weber. On series $\sum c_k f(kx)$ and Khinchin's conjecture. *Israel J. Math.* 201:591–609, 2014.

[71] S. G. Bobkov and F. Götze. Concentration inequalities and limit theorems for randomized sums. *Probab. Theory Related Fields* 137(1-2):49–81, 2007.

[72] A. Bondarenko and K. Seip. Large greatest common divisor sums and extreme values of the Riemann zeta function. *Duke Math. J.* 166(9):1685–1701, 2017.

[73] E. Borel. Les probabilités denombrables et leurs applications arithmétiques. *Rend. Circ. Mat. Palermo* 27:247–271, 1909.

[74] Z. Buczolich, B. Hanson, B. Maga, G. Vértesy. Type 1 and 2 sets for series of translates of functions. *Acta Math. Hungar.* 158(2):271–293, 2019.

[75] Y. Bugeaud. *Distribution modulo one and Diophantine approximation*, volume 193 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge, 2012.

[76] Y. Bugeaud. On the expansions of a real number to several integer bases. *Rev. Mat. Iberoam.* 28(4):931–946, 2012.

[77] L. Carleson. On convergence and growth of partial sums of Fourier series. *Acta Math.* 116:135–157, 1966.

[78] J. W. S. Cassels. Some metrical theorems of Diophantine approximation. III. *Proc. Cambridge Philos. Soc.* 46:219–225, 1950.

[79] J. W. S. Cassels. On a problem of Steinhaus about normal numbers. *Colloq. Math.* 7:95–101, 1959.

[80] F. Cellarosi, J. Marklof. Quadratic Weyl sums, automorphic functions and invariance principles. *Proc. Lond. Math. Soc.* 113(6):775–828, 2016.

[81] S. D. Chatterji. A general strong law. *Invent. Math.* 9:235–245, 1969/70.

[82] S. D. Chatterji. Un principe de sous-suites dans la théorie des probabilités. In *Séminaire de Probabilités, VI (Univ. Strasbourg, année universitaire 1970-1971; Journées Probabilistes de Strasbourg, 1971)*, pages 72–89. Lecture Notes in Math., Vol. 258. 1972.

[83] S. D. Chatterji. A principle of subsequences in probability theory: the central limit theorem. *Advances in Math.* 13:31–54; correction, ibid. 14 (1974), 266–269, 1974.

[84] S. D. Chatterji. A subsequence principle in probability theory. II. The law of the iterated logarithm. *Invent. Math.* 25:241–251, 1974.

[85] G. Cohen and J.-P. Conze. CLT for random walks of commuting endomorphisms on compact abelian groups. *J. Theoret. Probab.* 30(1):143–195, 2017.

[86] J.-P. Conze, S. Isola, and S. Le Borgne. Diffusive behavior of ergodic sums over rotations. *Stoch. Dyn.* 19(2):1950016, 26, 2019.

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

[88] J.-P. Conze, S. Le Borgne, and M. Roger. Central limit theorem for stationary products of toral automorphisms. *Discrete Contin. Dyn. Syst.* 32(5):1597–1626, 2012.

[89] A. H. Copeland and P. Erdös. Note on normal numbers. *Bull. Amer. Math. Soc.* 52:857–860, 1946.

[90] C. Cuny, J. Dedecker, and D. Volný. A functional CLT for fields of commuting transformations via martingale approximation. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* 441(Veroyatnost i Statistika. 22):239–262, 2015.

[91] K. Dajani and C. Kraaikamp. *Ergodic theory of numbers* volume 29 of *Carus Mathematical Monographs.* Mathematical Association of America, Washington, DC, 2002.

[92] H. Davenport and P. Erdös. Note on normal decimals. *Canad. J. Math.* 4:58–63, 1952.

[93] J.-M. De Koninck and I. Kátai. On a problem on normal numbers raised by Igor Shparlinski. *Bull. Aust. Math. Soc.* 84(2):337–349, 2011.

[94] R. de la Bretèche and G. Tenenbaum. Sommes de Gál et applications. *Proc. Lond. Math. Soc. (3)*, 119(1):104–134, 2019.

[95] M. Drmota, C. Mauduit, and J. Rivat. Normality along squares. *J. Eur. Math. Soc. (JEMS)* 21(2):507–548, 2019.

[96] M. Drmota and R. F. Tichy. *Sequences, discrepancies and applications*, volume 1651 of *Lecture Notes in Mathematics.* Springer-Verlag, Berlin, 1997.

[97] A. Dubickas. Powers of a rational number modulo 1 cannot lie in a small interval. *Acta Arith.* 137(3):233–239, 2009.

[98] A. Dubickas. On the size of a restricted sumset with application to the binary expansion of $\sqrt{d}$. *Appl. Anal. Discrete Math.* 13(2):346–360, 2019.

[99] N. Dunford and J. T. Schwartz. *Linear Operators. I. General Theory.* Pure and Applied Mathematics, Vol. 7. Interscience Publishers, 1958.

[100] P. Erdős. On the strong law of large numbers. *Trans. Amer. Math. Soc.* 67:51–56, 1949.

[101] P. Erdős. On the convergence of trigonometric series. *J. Math. Phys. Mass. Inst. Tech.* 22:37–39, 1943.

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

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

[104] P. Erdös and J. F. Koksma. On the uniform distribution modulo 1 of sequences $(f(n, \theta))$. *Nederl. Akad. Wetensch., Proc.* 52:851–854 = Indagationes Math. 11, 299–302 (1949), 1949.

[105] P. Erdős and A. Rényi. Additive properties of random sequences of positive integers. *Acta Arith.* 6: 83–110, 1960.

[106] C. G. Esseen. On the concentration function of a sum of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* 9:290–308, 1968.

[107] J.-H. Evertse, H. P. Schlickewei, and W. M. Schmidt. Linear equations in variables which lie in a multiplicative group. *Ann. of Math. (2)*, 155(3):807–836, 2002.

[108] H. Fiedler, W. Jurkat, O. Körner. Asymptotic expansions of finite theta series. *Acta Arith.* 32(2):129–146, 1977.

[109] F. Filip and J. Šustek. Normal numbers and Cantor expansions. *Unif. Distrib. Theory* 9(2):93–101, 2014.

[110] L. Flatto, J. C. Lagarias, and A. D. Pollington. On the range of fractional parts $\{\xi(p/q)^n\}$. *Acta Arith.* 70(2):125–147, 1995.

[111] L. Frühwirth, M. Juhos, and J. Prochno. The large deviation behavior of lacunary sums. *Monatsh. Math.* 199(1):113-133, 2022.

[112] K. Fukuyama. The central limit theorem for Riesz-Raikov sums. *Probab. Theory Related Fields* 100(1):57–75, 1994.

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

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

[115] K. Fukuyama. A central limit theorem and metric discrepancy result for sequences with bounded gaps. *Dependence in Probability, Analysis and Number Theory,* pp. 233–246. Kendrick Press 2010.

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

[117] K. Fukuyama. Pure Gaussian limit distributions of trigonometric series with bounded gaps. *Acta Math. Hungar.* 129(4):303–313, 2010.

[118] K. Fukuyama. A metric discrepancy result for a lacunary sequence with small gaps. *Monatsh. Math.* 162(3):277–288, 2011.

[119] K. Fukuyama. Metric discrepancy results for alternating geometric progressions. *Monatsh. Math.* 171(1):33–63, 2013.

[120] K. Fukuyama. The central limit theorem for Riesz-Raikov sums II. *Trans. Amer. Math. Soc.* 372(2):1193–1211, 2019.

[121] K. Fukuyama and N. Hiroshima. Metric discrepancy results for subsequences of $\{\theta^k x\}$. *Monatsh. Math.* 165(2):199–215, 2012.

[122] K. Fukuyama and S. Miyamoto. Metric discrepancy results for Erdös-Fortet sequence. *Studia Sci. Math. Hungar.* 49(1):52–78, 2012.

[123] K. Fukuyama and K. Nakata. A metric discrepancy result for the Hardy-Littlewood-Pólya sequences. *Monatsh. Math.* 160(1):41–49, 2010.

[124] K. Fukuyama and B. Petit. Le théorème limite central pour les suites de R. C. Baker. *Ergodic Theory Dynam. Systems* 21(2):479–492, 2001.

[125] K. Fukuyama and M. Yamashita. Metric discrepancy results for geometric progressions with large ratios. *Monatsh. Math.* 180(4):731–742, 2016.

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

[127] I. S. Gál. A theorem concerning Diophantine approximations. *Nieuw Arch. Wiskunde (2)*, 23:13–38, 1949.

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

[129] V. F. Gaposhkin. On series with respect to the system $\{\varphi(nx)\}$. *Mat. Sb. (N.S.)* 69 (111):328–353, 1966.

[130] V. F. Gaposhkin. Convergence and divergence systems. *Mat. Zametki* 4:253–260, 1968.

[131] V. F. Gaposhkin. The central limit theorem for certain weakly dependent sequences. *Teor. Verojatnost. i Primenen* 15:666–684, 1970.

[132] V. F. Gaposhkin. Convergence and limit theorems for subsequences of random variables. *Teor. Verojatnost. i Primenen* 17:401–423, 1972.

[133] M. Goldstern, J. Schmeling, and R. Winkler. Further Baire results on the distribution of subsequences. *Unif. Distrib. Theory* 2(1):127–149, 2007.

[134] H. Halberstam and K. F. Roth. Sequences, Vol. I. *Clarendon Press, Oxford* 1966.

[135] G. H. Hardy ad J. Littlewood. Some problems in diophantine approximation. *Acta Math.* 37:193–214, 1914.

[136] G. Harman. *Metric number theory*, volume 18 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1998.

[137] G. Harman. One hundred years of normal numbers. In *Number theory for the millennium, II (Urbana, IL, 2000)*, pages 149–166. A K Peters, Natick, MA, 2002.

[138] T. Hilberdink. An arithmetical mapping and applications to Ω-results for the Riemann zeta function. *Acta Arith.* 139(4):341–367, 2009.

[139] E. Hlawka. *The theory of uniform distribution*. Academic Publishers, 1984.

[140] R. A. Hunt. On the convergence of Fourier series. In *Orthogonal Expansions and their Continuous Analogues (Proc. Conf., Edwardsville, Ill., 1967)*, pages 235–255. Southern Illinois Univ. Press, Carbondale, Ill., 1968.

[141] W. B. Jurkat, J. W. Van Horne. The proof of the central limit theorem for theta sums. *Duke Math. J.* 48(4):873–885, 1981.

[142] W. B. Jurkat, J. W. Van Horne. On the central limit theorem for theta series. *Michigan Math. J.* 29(1):65–77, 1982.

[143] W. B. Jurkat, J. W. Van Horne. The uniform central limit theorem for theta sums. *Duke Math. J.* 50(3):649–666, 1983.

[144] M. Kac. Convergence of certain gap series. *Ann. of Math.* 44:411–415, 1943.

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

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

[147] M. I. Kadec and A. Pełczyński. Bases, lacunary sequences and complemented subspaces in the spaces L_p. *Studia Math.* 21:161–176, 1961/62.

[148] R. Kaufman. On the approximation of lacunary series by Brownian motion. *Acta. Math. Acad. Sci. Hung.* 35(1-2): 61–66 , 1980.

[149] M. Kesseböhmer, S. Munday, and B. O. Stratmann. *Infinite ergodic theory of numbers*. De Gruyter, Berlin, 2016.

[150] A. J. Khinchin. Ein Satz über Kettenbrüche, mit arithmetischen Anwendungen. *Mathematische Zeitschrift* 18:289-306, 1923.

[151] A. J. Khinchin. Über einen Satz der Wahrscheinlichkeitsrechnung. *Fund. Math.* 6:9-20, 1924.

[152] H. Ki and T. Linton. Normal numbers and subsets of **N** with given densities. *Fund. Math.* 144(2):163–179, 1994.

[153] D. E. Knuth. *The art of computer programming. Vol. 2: Seminumerical algorithms*. Addison-Wesley Publishing Co., 1969.

[154] J. F. Koksma. Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins. *Compositio Math.* 2:250–258, 1935.

[155] J. F. Koksma. On a certain integral in the theory of uniform distribution. *Nederl. Akad. Wetensch., Proc. Ser. A.* **54** = *Indagationes Math.*, 13:285–287, 1951.

[156] J. F. Koksma. A Diophantine property of summable functions. *J. Indian. Math. Soc.* 15:87-96, 1951.

[157] J. F. Koksma. Estimations de fonctions à l'aide d'intégrales de Lebesgue, *Bull. Soc. Math. Belg.* 6:4–13, 1953.

[158] A.N. Kolmogorov. Über das Gesetz des iterierten Logarithmus. *Math. Ann.* 101:126–135, 1929.

[159] J. Komlós. A generalization of a problem of Steinhaus. *Acta Math. Acad. Sci. Hungar.* 18:217–229, 1967.

[160] J. Komlós. Every sequence converging to 0 weakly in L_2 contains an unconditional convergence sequence. *Ark. Mat.* 12:41–49, 1974.

[161] L. Kuipers and H. Niederreiter. *Uniform distribution of sequences.* Wiley-Interscience, 1974.

[162] J. C. Lagarias. On the normality of arithmetical constants. *Experiment. Math.* 10(3):355–368, 2001.

[163] E. Landau. Über einen Konvergenzsatz. *Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl.* 1907:25–27, 1907.

[164] C. Lemieux. *Monte Carlo and quasi-Monte Carlo sampling.* Springer Series in Statistics. Springer, New York, 2009.

[165] G. Leobacher and F. Pillichshammer. *Introduction to quasi-Monte Carlo integration and applications.* Compact Textbooks in Mathematics. Birkhäuser/Springer, Cham, 2014.

[166] M. B. Levin. On the discrepancy estimate of normal numbers. *Acta Arith.* 88(2):99–111, 1999.

[167] M. B. Levin. Central limit theorem for $\mathbb{Z}_+^d$-actions by toral endomorphisms. *Electron. J. Probab.* 18:no. 35, 42, 2013.

[168] M. B. Levin. A multiparameter variant of the Salem-Zygmund central limit theorem on lacunary trigonometric series. *Colloq. Math.* 131(1):13–27, 2013.

[169] M. Lewko and M. Radziwiłł. Refinements of Gál's theorem and applications. *Adv. Math.* 305:280–297, 2017.

[170] B. Li and B. Mance. Number theoretic applications of a class of Cantor series fractal functions, II. *Int. J. Number Theory* 11(2):407–435, 2015.

[171] M. G. Madritsch. Construction of normal numbers via pseudo-polynomial prime sequences. *Acta Arith.* 166(1):81–100, 2014.

[172] M. G. Madritsch and B. Mance. Construction of μ-normal sequences. *Monatsh. Math.* 179(2):259–280, 2016.

[173] M. G. Madritsch, A.-M. Scheerer, and R. F. Tichy. Computable absolutely Pisot normal numbers. *Acta Arith.* 184(1):7–29, 2018.

[174] M. G. Madritsch, J. M. Thuswaldner, and R. F. Tichy. Normality of numbers generated by the values of entire functions. *J. Number Theory* 128(5):1127–1145, 2008.

[175] B. Mance. Typicality of normal numbers with respect to the Cantor series expansion. *New York J. Math.* 17:601–617, 2011.

[176] B. Mance. Number theoretic applications of a class of Cantor series fractal functions. I. *Acta Math. Hungar.* 144(2):449–493, 2014.

[177] J. Marklof. The Berry-Tabor conjecture. In *European Congress of Mathematics, Vol. II (Barcelona, 2000)*, volume 202 of *Progr. Math.*, pages 421–427. Birkhäuser, Basel, 2001.

[178] J. M. Marstrand. On Khinchin's conjecture about strong uniform distribution. *Proc. London Math. Soc. (3)*, 21:540–556, 1970.

[179] C. Mauduit and A. Sárközy. On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol. *Acta Arith.* 82(4):365–377, 1997.

[180] A. M. Mian, S. Chowla. On the B_2 sequences of Sidon. *Proc. Nat. Acad. Sci. India Sect. A* 14:3–4, 1944.

[181] H. L. Montgomery. *Ten lectures on the interface between analytic number theory and harmonic analysis*, volume 84 of *CBMS Regional Conference Series in Mathematics.* American Mathematical Society, Providence, RI, 1994.

[182] G. W. Morgenthaler. A central limit theorem for uniformly bounded orthonormal systems. *Trans. Amer. Math. Soc.* 79:281–311, 1955.

[183] C. Müllner and L. Spiegelhofer. Normality of the Thue-Morse sequence along Piatetski-Shapiro sequences, II. *Israel J. Math.* 220(2):691–738, 2017.

[184] T. Murai. The central limit theorem for trigonometric series. *Nagoya Math. J.* 87:79–94, 1982.

[185] R. Nair. On strong uniform distribution. *Acta Arith.* 56(3):183–193, 1990.

[186] Y. Nakai and I. Shiokawa. A class of normal numbers. *Japan. J. Math. (N.S.)* 16(1):17–29, 1990.

[187] Y. Nakai and I. Shiokawa. Normality of numbers generated by the values of polynomials at primes. *Acta Arith.* 81(4):345–356, 1997.

[188] H. Niederreiter and R. F. Tichy. Solution of a problem of Knuth on complete uniform distribution of sequences. *Mathematika* 32(1):26–32, 1985.

[189] J. S. B. Nielsen and J. G. Simonsen. An experimental investigation of the normality of irrational algebraic numbers. *Math. Comp.* 82(283):1837–1858, 2013.

[190] E. M. Nikishin. Resonance theorems and superlinear operators. *Uspehi Mat. Nauk* 25(6(156)):129–191, 1970.

[191] E. M. Nikishin. On systems of convergence in measure for ℓ_2. *Mat. Zametki* 13:337–340, 1973.

[192] E. Novak and H. Woźniakowski. *Tractability of multivariate problems. Volume II: Standard information for functionals*, volume 12 of *EMS Tracts in Mathematics*. European Mathematical Society (EMS), Zürich, 2010.

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

[194] 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(2):705–727, 1994.

[195] 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(2):705–727, 1994.

[196] W. Philipp and W. Stout. Almost sure invariance principles for partial sums of weakly dependent random variables. *Mem. Amer. Math. Soc.* 2(no. 161):iv+140, 1975.

[197] A. Rényi. On stable sequences of events. *Sankhyā Ser. A* 25:293 302, 1963.

[198] A. Rényi. Foundations of probability. *Holden-Day*, 1970.

[199] Z. Rudnick and A. Zaharescu. The distribution of spacings between fractional parts of lacunary sequences. *Forum Math.* 14(5):691–712, 2002.

[200] I. Z. Ruzsa. An infinite Sidon sequence. *J. Number Theory* 68(1):63–71, 1998.

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

[202] R. Salem and A. Zygmund. On lacunary trigonometric series. II. *Proc. Nat. Acad. Sci. U.S.A.* 34:54–62, 1948.

[203] R. Salem and A. Zygmund. La loi du logarithme itéré pour les séries trigonométriques lacunaires. *Bull. Sci. Math. (2)*, 74:209–224, 1950.

[204] R. Salem and A. Zygmund. Some properties of trigonometric series whose terms have random signs. *Acta Math.* 91:245–301, 1954.

[205] A.-M. Scheerer. Computable absolutely normal numbers and discrepancies. *Math. Comp.* 86(308):2911–2926, 2017.

[206] W. M. Schmidt. Über die Normalität von Zahlen zu verschiedenen Basen. *Acta Arith.* 7:299–309, 1961/62.

[207] W. M. Schmidt. Norm form equations. *Ann. of Math. (2)*, 96:526–551, 1972.

[208] W. Sierpinski. Démonstration élémentaire du théorème de M. Borel sur les nombres absolument normaux et détermination effective d'une tel nombre. *Bull. Soc. Math. France* 45:125–132, 1917.

[209] K. Soundararajan. Extreme values of zeta and L-functions. *Math. Ann.* 342(2):467–486, 2008.

[210] L. Spiegelhofer. Normality of the Thue-Morse sequence along Piatetski-Shapiro sequences. *Q. J. Math.* 66(4):1127–1138, 2015.

[211] S. Takahashi. A gap sequence with gaps bigger than the Hadamards. *Tohoku Math. J. (2)*, 13:105–111, 1961.

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

[213] S. Takahashi. The law of the iterated logarithm for a gap sequence with infinite gaps. *Tohoku Math. J. (2)*, 15:281–288, 1963.

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

[215] N. Technau and A. Zafeiropoulos. The discrepancy of $(n_k x)_{k=1}^{\infty}$ with respect to certain probability measures. *Q. J. Math.* 71(2):573–597, 2020.

[216] R. Tijdeman. On integers with many small prime factors. *Compositio Math.* 26:319–330, 1973.

[217] J. Vandehey. On the binary digits of $\sqrt{2}$. *Integers* 18:Paper No. A30, 7, 2018.

[218] S. Wagon. Is π normal? *Math. Intelligencer* 7(3):65–67, 1985.

[219] A. Walfisz. Ein metrischer Satz über diophantische Approximationen. *Fund. Math.* 16:361–365, 1930.

[220] M. Weber. On systems of dilated functions. *C. R. Math. Acad. Sci. Paris* 349:1261–1263, 2011.

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

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

[223] A. Wintner. Diophantine approximations and Hilbert's space. *Amer. J. Math.* 66: 564–578, 1944.

[224] K. F. Xylogiannopoulos, P. Karampelas, and R. Alhajj. Experimental analysis on the normality of π, e, ϕ, $\sqrt{2}$ using advanced data-mining techniques. *Exp. Math.* 23(2):105–128, 2014.