Equidistribution, uniform distribution: a probabilist's perspective

Vlada Limic and Nedžad Limić[*]

IRMA, UMR 7501 de l'Université
de Strasbourg et du CNRS,
7 rue René Descartes,
67084 Strasbourg Cedex, France
e-mail: vlada@math.unistra.fr

University of Zagreb,
Bijenička cesta 30,
10000 Zagreb, Croatia
e-mail: nlimic@math.hr

Abstract: The theory of equidistribution is about hundred years old, and has been developed primarily by number theorists and theoretical computer scientists. A motivated uninitiated peer could encounter difficulties perusing the literature, due to various synonyms and polysemes used by different schools. One purpose of this note is to provide a short introduction for probabilists. We proceed by recalling a perspective originating in a work of the second author from 2002. Using it, various new examples of completely uniformly distributed mod 1 sequences, in the "metric" (meaning almost sure stochastic) sense, can be easily exhibited. In particular, we point out natural generalizations of the original p-multiply equidistributed sequence $k^p t$ mod 1, $k \geq 1$ (where $p \in \mathbb{N}$ and $t \in [0,1]$), due to Hermann Weyl in 1916. In passing, we also derive a Weyl-like criterion for weakly completely equidistributed (also known as WCUD) sequences, of substantial recent interest in MCMC simulations.

The translation from number theory to probability language brings into focus a version of the strong law of large numbers for weakly correlated complex-valued random variables, the study of which was initiated by Weyl in the aforementioned manuscript, followed up by Davenport, Erdős and LeVeque in 1963, and greatly extended by Russell Lyons in 1988. In this context, an application to ∞-distributed Koksma's numbers t^k mod 1, $k \geq 1$ (where $t \in [1, a]$ for some $a > 1$), and an important generalization by Niederreiter and Tichy from 1985 are discussed.

The paper contains negligible amount of new mathematics in the strict sense, but its perspective and open questions included in the end could be of considerable interest to probabilists and statisticians, as well as certain computer scientists and number theorists.

1. Introduction

This is certainly neither the first nor the last time that equidistribution is viewed using a "probabilistic lense". A probability and number theory enthusiast will likely recall in this context the famous Erdös-Kac central limit theorem [EK] (see also [Du], Ch. 2 (4.9)), or the celebrated monograph by Kac [Ka1].

Unlike in [Ka1, Ka2, EK, Ke1, Ke2] our main concern here are the completely equidistributed sequences and their "generation". In the abstract we deliberately alternate between equidistribution (or equidistributed) and two of its synonyms. In particular, the *completely uniformly distributed* (sometimes followed by mod 1) and ∞-*distributed* in the abstract, the keywords, and the references (see for example any of [Ho1, Ho2, DT, Kn1, Kn2, Kr5, KN, NT1, Lac, Lo, Le2, TO, OT, SP]) mean the same as *completely equidistributed*. Uniform distribution is a fundamental probability theory concept. To minimize the confusion, in the rest of this note we shall:

(i) always write the *uniform law* when referring to the distribution of a uniform random variable, and

(ii) almost exclusively write equidistributed (to mean equidistributed, uniformly distributed, uniformly distributed mod 1 or $\cdot$-distributed), usually preceded by one of the following attributes: *simply, d-multiply or completely.*

Let $D = [0,1]^d$ be the d-dimensional unit cube. Let $r \in \mathbb{N}$ and assume that a bounded domain G is given in $\mathbb{R}^r$. Denote by λ the restriction of the d-dimensional Lebesgue measure on D, as well as the Lebesgue measure on $\mathbb{R}^r$. In particular $\lambda(G)$ is the r-dimensional volume of G. In most of our examples r will equal 1, and G will be an interval.

Throughout this note, a sequence of measurable (typically continuous) functions $(\mathbf{x}_k)_{k \geq 1}$, where $\mathbf{x}_k : G \mapsto \mathbb{R}^d$ will define a sequence of points $\boldsymbol{\beta}_k \in D$ as follows:

$$\boldsymbol{\beta}_k \equiv \boldsymbol{\beta}(\mathbf{t}) := \mathbf{x}_k(\mathbf{t}) \text{ mod } \mathbf{1}, \quad \mathbf{t} \in G, \tag{1.1}$$

understanding that $\mathbf{1} = (1, \ldots, 1) \in \mathbb{R}^d$, and that the modulo operation is naturally extended to vectors in the component-wise sense. While each $\boldsymbol{\beta}_k$ is a (measurable) function of $\mathbf{t}$, this dependence is typically omitted from the notation. The $(\mathbf{x}_k)_k$ is called the *generating sequence* or simply the *generator*, and the elements $\mathbf{t}$ of G are called *seeds*.

Any $A = \prod_j (a_j, b_j] \subset D$, $0 \leq a_j < b_j \leq 1$ will be called a *d-dimensional box*, or just a *box* in D. As usual, we denote by $\mathbb{1}_S$ the indicator of a set S. The sequence $\boldsymbol{\beta} = \{\boldsymbol{\beta}_k : k \in \mathbb{N}\}$ is *simply equidistributed in D* if

$$\lim_{N \to \infty} \frac{1}{N} \sum_{k=1}^{N} \mathbb{1}_A(\boldsymbol{\beta}_k(\mathbf{t})) = \lambda(A), \tag{1.2}$$

for each A box in D, and almost every $\mathbf{t} \in G$.

1.1. Notes on the literature

Our general setting is mostly inherited from [Li]. The few changes in the notation and the jargon aim to simplify reading of the present work by an interested probabilist or statistician. In [DT, Kn2, KN, Kr5, SP] and other standard references, equidistribution is defined as a property of a single (deterministic) sequence of real numbers. Nevertheless, in any concrete example discussed here (e.g. the sequences mentioned in the abstract, as well as all the examples given in the forthcoming sections), this property is verified only up to a null set over a certain parameter space. So there seems to be no loss of generality in integrating the almost everywhere/surely aspect in Definitions 1.1 and 1.2 below. There is one potential advantage: the Weyl criterion (viewed in a.s. sense) reduces to (countably many applications of) the strong law of large numbers (SLLN) for specially chosen sequences of dependent complex-valued random variables. The just made observation is the central theme of this note. Note in addition that the study of R4 (and R6) types of randomness (according to Knuth [Kn2], Section 3.5, see also Sections 2.4.1 and 3 below) makes sense only in the stochastic setting.

Davenport, Erdös and LeVeque [DELV] are strongly motivated by the Weyl equidistribution analysis [We], yet they do not mention any connections to probability theory. Lyons [Ly] clearly refers to the main result of [DELV] as a SLLN criterion, but is not otherwise interested in equidistribution.

The Weyl variant of the SLLN (see Section 2.2) is central to the analysis of Koksma [Kok], who seems to ignore its probabilistic aspect. The breakthrough on the complete equidistribution of Koksma's numbers (and variations) by Franklin [Fr] relies on the main result in [Kok], but without (any need of) recalling the Weyl variant of the SLLN in the background (see the proof of [Fr], Theorems 14 and 15). On the surface, this line of research looks more and more distant from the probability theory. The classical mainstream number theory textbooks (e.g. [KN]), as well as modern references (e.g. [DT, SP]), corroborate this point of view of equidistribution in the "metric" (soon to be called "stochastic a.s.") sense. And yet, the Niederreiter and Tichy [NT1] metric theorem, considered by many as one of the highlights on this topic, consists of lengthy and clever (calculus based) covariance calculations, followed by an application of the SLLN from [DELV] (see Section 2.4.1 for more details).

To the best of our knowledge, [Li] is the first (and at present the only) study of complete (or other) equidistribution in the "metric" sense, which identifies the verification of Weyl's criterion as a stochastic problem (equivalent to countably many SLLN, recalled in Section 2), without any additional restriction on the nature of the generator $(\mathbf{x}_k)_k$. In comparison, Holewijn [Ho1, Ho2] made analogous connection (and in [Ho2] even applied the SLLN criterion of [DELV] to the sequence of rescaled Weyl's sums (2.4)), but only under particularly nice probabilistic assumptions, which are not satisfied in any of the examples discussed in the present survey. In addition, both Lacaze [Lac] and Loynes [Lo] use the Weyl variant (or a slight modification) of the SLLN, again under rather restrictive probabilistic hypothesis (to be recalled in Section 2.1.2).

Kesten in [Ke1, Ke2], as well as in his subsequent articles on similar probabilistic number theory topics, works in the setting analogous to that of [Ho1, Ho2], however his analysis is not directly connected to the Weyl criterion. On a related topic, in a pioneer study of (non-)normal numbers and their relation to equidistribution, Mendes France [MF] also applies the SLLN from [DELV] as a purely analytic result, though it is not clear that a probabilistic interpretation brings a new insight in that setting. Interested readers are referred to [Kem, Kh] for probabilistic perspectives on normal numbers.

The study of complete equidistribution in the standard (deterministic) sense, and its close link to normal numbers, was initiated by Korobov [Kr1] in 1948. Korobov used Weyl's criterion [We] for (multiple and complete) equidistribution, and exhibited an explicit function f (in a form of a power series) such that $(f(k) \bmod 1)_k$ is a completely equidistributed sequence (see also [Kr5], Theorem 28). Knuth was aware of [Fr], but apparently unaware of either [We, Kr1], when he exhibited in [Kn1] a (deterministic) completely equidistributed (there called "random") sequence of numbers in $[0, 1]$, by extending the method of Champernowne's that previously served to find an explicit normal number. Knuth [Kn1, Kn2] uses his own (computer science inspired) criterion for complete equidistribution. All the examples given in [Kr1, Kr2] and [Kn1, Kn2], as well as those obtained later on by the "Korobov school" (see Levin [Le2], Korobov [Kr5], and also the historical notes in [KN], and [SP]), are practical to a varying degree. To find out more about the deterministic setting, readers are encouraged to use the pointers (to synonyms and references) given above as a guide to the literature.

We finally wish to make note of a recent spur of interest in completely equidistributed sequences, and their generalizations (definable only in the stochastic setting), in relation to Markov chain Monte-Carlo simulations [CMNO, CDO, OT, TO]. Section 2.3 makes a brief digression in this direction. Importance of complete equidistribution for MCMC is not surprising, in view of a long list of empirical tests that these sequences satisfy (see [Fr, Kn2] and Remark 2(d,e)). The first author is convinced that anyone who regularly runs or even looks at pseudo-random simulations should benefit from reading a note of this kind.

Disclaimer and motivation This review and tutorial is highly inclusive, but by no means exhaustive. The theory of equidistribution (or uniform distribution) is rich, complex and fast evolving, and it would be very difficult to point to a single book volume, let alone a survey paper, which covers all of its interesting aspects (for example, the list of references in the specialized survey [AB2] overlaps with ours in only four items). Even when focusing on complete equidistribution in the metric (stochastic a.s.) sense, it seems hard to find a single expository article aimed at specialists, let alone at the probability and statistics community at large. We hope to have accomplished here an "order of magnitude" effect, citing several dozens of original research papers and surveys, textbooks or monographs, in a brief attempt to shed a probability-friendly light on the concepts and ideas presented, as well as to point out a number of natural and interesting open questions. We wish we had come across such a paper at our very encounter with this important topic. The latter thought gave the impetus to our writing.

1.2. One-dimensional examples

Suppose that $d = r = 1$, and that G is an interval. We recall several well-known examples of functions $x_k : G \mapsto \mathbb{R}$, that generate equidistributed sequences in $[0, 1]$. The first class of examples is as follows: for each $p \in \mathbb{N}$ define

$$x_k(t) = k^p\, t, \quad t \in G := (0, 1),\ k \geq 1. \tag{1.3}$$

Then $(\beta_k)_{k \geq 1} \in [0, 1]^{\mathbb{N}}$ from (1.1) are *Weyl's numbers* of parameter p. The second class of examples are the so-called *multiplicatively generated numbers*: for each $M \in \{2, 3, \dots, \}$ let

$$x_k(t) = M^k\, t, \quad t \in G := (0, 1),\ \ k \geq 1, \tag{1.4}$$

and $(\beta_k)_{k \geq 1} \in [0, 1]^{\mathbb{N}}$ as in (1.1). Finally, if one defines for each k, and any $a > 1$

$$x_k(t) = t^k, \quad t \in G := (1, a),\ \ k \geq 1, \tag{1.5}$$

then $(\beta_k)_{k \geq 1}$ from (1.1) are known as *Koksma's numbers*. It is well-known that for each of the above three examples, the sequences $(\beta_k)_k$ are (at least simply) equidistributed for almost all seeds [We, Fr, Kok] (see also [KN, DT, SP]). In fact if $p = 1$, then the Weyl (Sierpinski, Bohl) equidistribution theorem is a stronger claim: $k \cdot t$ mod 1 is simply equidistributed for all irrational t.

One could **refine** the notion of a **seed**, and call t a seed only if $\beta_k(t)$ is (sufficiently) equidistributed. The new set of seeds T would then be well-defined up to a null-set. Note that (1.3) and (1.4) are both linear in t, which is not true for (1.5). The methodology developed by the second author in [Li] was motivated by the particularly simple analysis of the linear case, which can be extended (under certain hypotheses on $(x_k)_k$) to the non-linear setting.

1.3. Multiple equidistribution with examples

Let $d \geq 2$ be fixed. A sequence of real measurable functions $(x_k)_{k \geq 1}$ can be used to form sequences $(\mathbf{x}_k)_{k \geq 1}$ of d-dimensional vector-valued (measurable) functions, and therefore the corresponding sequences $(\boldsymbol{\beta}_k)_{k \geq 1}$ (see (1.1)), in at least three different natural ways:

(a) The set of seeds could be equally d-dimensional (here $r = d$). More precisely, for $\mathbf{t} = (t_1, t_2, \dots, t_d) \in G$, let $\mathbf{x}_k(\mathbf{t}) := (x_{k1}(\mathbf{t}), x_{k2}(\mathbf{t}), \dots, x_{kd}(\mathbf{t}))$, where $x_{kj}(\mathbf{t}) := x_{(k-1)d+j}(t_j)$, for all $k \in \mathbb{N}$, $j = 1, \dots, d$. In this case

$$\beta_{kj}^d = x_{(k-1)d+j}(t_j) \text{ mod } 1, \quad k \in \mathbb{N},\ j = 1, \dots, d. \tag{1.6}$$

(b) The set of seeds could be r-dimensional, and the successive $\mathbf{x}$ (and $\boldsymbol{\beta}$) could be formed by shifting the window of observation by 1. More precisely, for $\mathbf{t} \in G \subset \mathbb{R}^r$, define $\mathbf{x}_k(\mathbf{t}) := (x_{k1}(\mathbf{t}), x_{k2}(\mathbf{t}), \dots, x_{kd}(\mathbf{t}))$, where $x_{kj}(\mathbf{t}) := x_{k+j-1}(\mathbf{t})$, for all $k \in \mathbb{N}$, $j = 1, \dots, d$. In this case

$$\beta_{kj}^d = x_{k+j-1}(\mathbf{t}) \text{ mod } 1, \quad k \in \mathbb{N},\ j = 1, \dots, d. \tag{1.7}$$

(c) Let $d \geq 1$. The set of seeds could be r-dimensional, and the successive $\mathbf{x}$ (and $\boldsymbol{\beta}$) could be formed by shifting the window of observation by $h \in \mathbb{N}$. More precisely, for $\mathbf{t} \in G \subset \mathbb{R}^r$, define $\mathbf{x}_k(\mathbf{t}) := (x_{k1}(\mathbf{t}), x_{k2}(\mathbf{t}), \ldots, x_{kd}(\mathbf{t}))$, where $x_{kj}(\mathbf{t}) := x_{(k-1)h+j}(\mathbf{t})$, for all $k \in \mathbb{N}$, $j = 1, \ldots, d$. In this case

$$\beta_{kj}^{d,h} = x_{(k-1)h+j}(\mathbf{t}) \bmod 1, \quad k \in \mathbb{N}, j = 1, \ldots, d. \qquad (1.8)$$

Note that class (c) comprises class (b) (if h is set to 1), and that if $h > d$, then $(\boldsymbol{\beta}_k^{d,h})_k$ is formed from a strict subsequence of $(x_k)_k$. We shall consider the above definitions with d varying over $\mathbb{N}$. The sequences in (1.6) will be included in the analysis of Section 2.1.

The equidistribution analysis is typically done on the sequences of vectors $\boldsymbol{\beta}^d$ from (1.7). Yet (see Remark 1(c) below) the construction in (1.8) is particularly interesting from the perspective of comparison with (pseudo-)random simulations.

Simple equidistribution in D was defined in the paragraph preceding Section 1.1. Let again G be a bounded domain in $\mathbb{R}^r$ for some $r \geq 1$.

DEFINITION **1.1.** *Assume that $(x_k)_{k \geq 1}$ is a sequence of real measurable functions on G. (i) The sequence $\{\beta_k : k \in \mathbb{N}\}$ defined in (1.1) is said to be d-multiply equidistributed in $[0, 1]$ if there exists a measurable subset of seeds T_d, of full measure (or equivalently, $\lambda(T_d) = \lambda(G)$), such that the sequence of vectors $\boldsymbol{\beta}_k^d$ in (1.7) is simply equidistributed in $[0, 1]^d$ for each $t \in T_d$.*

(ii) If $\{\beta_k : k \in \mathbb{N}\}$ is d-multiply equidistributed in $[0, 1]$ for all $d \geq 1$, then $\{\beta_k : k \in \mathbb{N}\}$ is called completely equidistributed.

We shall often abbreviate completely equidistributed sequence in $[0, 1]$ by *c.e.s* in $[0, 1]$. Occasionally we may drop the descriptive "in $[0, 1]$". It is natural to also consider shifts of arbitrary fixed length.

DEFINITION **1.2.** *If for some measurable subset of seeds $T_{d,h}$, again of full measure, we have that the sequence of vectors $\boldsymbol{\beta}_k^{d,h}$ in (1.8) is simply equidistributed in $[0, 1]^d$ for all $t \in T_{d,h}$ then we say that $\{\beta_k : k \in \mathbb{N}\}$ defined in (1.1) is d-multiply equidistributed in $[0, 1]$ with respect to shift by h.*

Naturally, we omit "with respect to shift..." if $h = 1$, and use attribute *simply* instead of 1-multiply if $d = 1$ and $h > 1$.

Remark 1. (a) Fix some $h \in \mathbb{N}$ and $o \in \mathbb{N}_0$. It is easy to check that if $\{\beta_k : k \in \mathbb{N}\}$ is d-multiply equidistributed in $[0, 1]$ with respect to shift by h, then the same property holds (on the same subset of seeds) for $\{\beta_{k+o} : k \in \mathbb{N}\}$. For this reason the definitions above are stated only in the case $o = 0$.

(b) Due to the just mentioned easy fact, one can quickly check (by averaging over h different offsets) that if $\{\beta_k : k \in \mathbb{N}\}$ is d-multiply equidistributed in $[0, 1]$ with respect to shift by $h \geq 2$, then it is also d-multiply equidistributed in $[0, 1]$ (in the sense of Definition 1.1). Note however the following example: suppose $(\alpha_k)_k$ and $(\gamma_k)_k$ are (simply) equidistributed in $[0, 1]$ so that $(\alpha_k/2)_k$ and $((1 + \gamma_k)/2)_k$ are equidistributed in $[0, 1/2]$ and $[1/2, 1]$, respectively. Then

$(\beta_k)_k$ defined by

$$\beta_{2k-1} := \alpha_k, \; \beta_{2k} := \gamma_k, \quad k \geq 1,$$

is (simply) equidistributed in $[0, 1]$, but not even simply equidistributed in $[0, 1]$ with respect to shift by 2. One could construct similar examples in the multiply equidistributed setting.

(c) As already noted, the interest of including shifts becomes apparent if one compares the equidistribution with the law of large numbers (LLN), or with Monte-Carlo simulations using pseudo-random numbers. If $X_1, X_2, \ldots$ is a sequence of i.i.d. uniform random variables, and $f : [0, 1]^d \to \mathbb{R}$ a bounded measurable map, then $Ef(X_1, \ldots, X_d)$ is the theoretical (LLN) limit (in the L^p and in the almost sure sense) of

$$\frac{1}{n} \sum_{k=1}^n f(X_{(k-1)d+1}, \ldots, X_{(k-1)d+d}). \tag{1.9}$$

Given a pseudo-random sequence $a = (a_k)_{k \geq 1}$, a direct analogue of (1.9) is

$$\frac{1}{n} \sum_{k=1}^n f(a_{(k-1)d+1}, \ldots, a_{(k-1)d+d}), \tag{1.10}$$

and the shift by $h = d$ is typical in doing simulations. While it is also true that for each choice of $h, o \in \mathbb{N}$

$$\lim_n \frac{1}{n} \sum_{k=1}^n f(X_{(k-1)h+o}, \ldots, X_{(k-1)h+d+o-1}) = Ef(X_1, \ldots, X_d),$$

the variance (and therefore the theoretical error) of the approximation is the smallest if $h \geq d$. This variance (error) bound is constant over shifts $h \geq d$ and offsets $o \geq 1$, hence (1.9) is an "economical" approximation (each element of $(X_k)_k$ is used once and only once), and (1.10) is its direct analog.

(d) We can point the reader to at least two different derivations ([Kn2] Section 3.5, Theorem C and the final Note in [Kr5] Ch. III, §20) of the following important fact: if $\{\beta_k : k \in \mathbb{N}\}$ is completely equidistributed, then it is also ("omega-by-omega", on the same set of seeds) d-multiply equidistributed in $[0, 1]$ with respect to shift by h for each $d \geq 1$ and each $h \geq 2$. Due to this fact, and the observations made in (b), defining completely equidistributed with respect to shifts is superfluous. (e) As a consequence of (d) and various other properties derived in [Fr], Knuth [Kn1, Kn2] concludes that any c.e.s. passes numerous empirical tests (see [Kn2], Section 3.5, the comment following Definition R1), and is therefore an exemplary pseudo-random sequence. □

The Weyl sequence $(\beta_k)_{k \geq 1}$ from (1.3) is p-multiply equidistributed, but it is not $p+1$-multiply equidistributed; the multiplicatively generated sequence from (1.4) is only simply equidistributed (see also Remark 2(c)). Section 2.1.3 serves to point out that already in the linear setting, numerous examples of c.e.s. that generalize the Weyl sequence in a natural way, can be easily constructed.

Koksma's numbers from (1.5) generally serve (see e.g. [Kn2, KN, TO, SP]) as the prototype of a metric c.e.s. Most of the sequel is organized in connection to this example. In particular, Section 2.4 gives a page-long proof of their complete equidistribution, as an extension of the technique from Section 2.1 to the non-linear setting. The discussion in Section 2.4.1 serves to put this into perspective with respect to [Kok, Fr] and [NT1]. Section 3 recalls the main observations made in this and the next section, and discusses several natural open problems.

2. Equidistribution via probabilistic reasoning

The purpose of this section is to describe the approach of [Li], as well as to put it into perspective with respect to [Kok, Fr, NT1].

2.1. A lesson from the linear case

Consider the set of multi-indices $\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}^d$. Let $\{x_k : k \in \mathbb{N}\}$ be a sequence of functions (soon taken to be linear) as in Definitions 1.1-1.2, and let points $\boldsymbol{\beta}_k \in D$ be defined by (1.6).

Define the sequence $(\nu_N)_N$ of purely atomic finite (probability) measures on $[0,1]^d$ via

$$\nu_N := \sum_{k=1}^{N} \frac{1}{N} \delta_{\boldsymbol{\beta}_k}, \quad N \geq 1, \tag{2.1}$$

where the dependence in $\mathbf{t} \in G$ is implicit. Then clearly, the (simple) equidistribution in $[0,1]^d$ (1.2) is equivalent to the weak convergence of $(\nu_N)_N$ to the uniform law on $[0,1]^d$ (denoted in (1.2) by λ), as N goes to ∞. This in turn is equivalent to saying that for any polynomial p in d-variables we have $\lim_N \int p(\mathbf{s})\nu_N(d\mathbf{s}) = \int_{[0,1]^d} p(\mathbf{s})\lambda(d\mathbf{s})$.

From an analyst's perspective, choosing the class of polynomials in the above characterization of weak convergence is suboptimal. Indeed, if $d = 1$, the class of complex exponentials $t \mapsto \{\exp(2\pi i\, m \cdot t)\}$, $m \in \mathbb{Z}$, is orthogonal (even orthonormal) in $L^2[0,1]$, and (due to the Stone-Weierstrass theorem) the algebra they generate is dense in the (periodic) continuous functions on $[0,1]$, with respect to the usual sup-norm. The same is true in the d-multiple setting, this time with respect to the class of multi-dimensional complex exponentials $D \ni \mathbf{s} \mapsto \{\exp(2\pi i\, \mathbf{m} \cdot \mathbf{s})\}$, $\mathbf{m} \in \mathbb{Z}^d$, where $\cdot$ is the dot (or scalar) product.

In order to check that $\lim_N \nu_N = \lambda$ in law, it is necessary and sufficient that for each $\mathbf{m} \in \mathbb{Z}^d$

$$\frac{1}{N} \sum_{k=1}^{N} \exp(2\pi i\, \mathbf{m} \cdot \boldsymbol{\beta}_k) = \int_D \exp(2\pi i\, \mathbf{m} \cdot \mathbf{s})\nu_N(d\mathbf{s})$$

converges as $N \to \infty$ to

$$\int_{[0,1]^d} \exp(2\pi i\, \mathbf{m} \cdot \mathbf{s})\lambda(d\mathbf{s}) = \left\{ \begin{array}{ll} 1, & \mathbf{m} = \mathbf{0}, \\ 0, & \mathbf{m} \neq \mathbf{0}. \end{array} \right.$$

The just obtained characterization for equidistribution of $(\beta_k)_k$ is the well-known *Weyl criterion* [We]: consider the quantities

$$W_N(\boldsymbol{\beta}, \mathbf{m}) \ := \ \frac{1}{N} \sum_{k=1}^{N} \exp\left(2\pi i\, \mathbf{m} \cdot \boldsymbol{\beta}_k\right), \tag{2.2}$$

then $(\boldsymbol{\beta}_k)_k$ is (simply) equidistributed in D if and only if $W_N(\boldsymbol{\beta}, \mathbf{m}) \to 0$ for $N \to \infty$ and each $\mathbf{m} \neq \mathbf{0}$. Recalling that in our setting each $\boldsymbol{\beta}_k$ and therefore ν_N is in fact a (measurable) function of $\mathbf{t}$, the criterion reads

$$\lim_N W_N(\boldsymbol{\beta}, \mathbf{m}) = 0, \text{ almost everywhere in } G, \text{ for each } \mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}. \tag{2.3}$$

Denote by $\mathbb{R}^d \ni \mathbf{z} \mapsto e(\mathbf{m}, \mathbf{z}) = \exp(2\pi i\, \mathbf{m} \cdot \mathbf{z})$. Due to (1.1), and the periodicity of the complex exponential, we have the identity

$$W_N(\boldsymbol{\beta}, \mathbf{m}) \equiv \frac{1}{N} \sum_{k=1}^{N} e(\mathbf{m}, \mathbf{x}_k). \tag{2.4}$$

A crucial point is that if $(\mathbf{x}_k)_k$ is defined by (1.6), where the sequence of real (measurable) functions $(x_k)_k$ is given either in (1.3) or in (1.4), then the functions $G^d \ni \mathbf{t} \mapsto e(\mathbf{m}, \mathbf{x}_k(\mathbf{t}))$ form a bounded orthogonal sequence in $L_2(G, \mathbb{C})$. Note that this is the first time that the linearity in t (resp. $\mathbf{t}$) of the functions in (1.3,1.4) (resp. (1.6)) is being called for.

Unlike [Li], we continue the discussion using probabilistic wording and notation. Given $z \in \mathbb{C}$, denote by $\overline{z}$ its complex conjugate. For a fixed $\mathbf{m} \in \mathbb{Z}^d \setminus \{0\}$ and each $k \in \mathbb{N}$, define $Y_k(\mathbf{m}\,;\mathbf{t}) := e(\mathbf{m}, \mathbf{x}_k(\mathbf{t}))$, $\mathbf{t} \in G$. Then $(Y_k(\mathbf{m}))_k$ is a sequence of complex-valued random variables on the probability space $(G, \mathcal{B}, P)$, where $\mathcal{B}$ is the Borel σ-field on G, and $P(d\mathbf{t}) = \lambda(d\mathbf{t})/\lambda(G)$, such that $EY_k(\mathbf{m}) = 0$, for all $k \geq 1$, and

$$E(Y_k(\mathbf{m})\overline{Y_l(\mathbf{m})}) = \left\{ \begin{array}{ll} 1, & k = l, \\ 0, & k \neq l, \end{array} \right. \quad l, k \in \mathbb{N}. \tag{2.5}$$

The standard proof of the strong LLN for i.i.d. variables with finite second moment (see for example the exercise concluding [Du] Ch. I, Section 7) can clearly be extended to sequences of pairwise uncorrelated centered random variables with constant (or uniformly bounded) variance. Anticipating some readers with non-probabilist background, we include a sketch of the argument in points LLN.(a)-LLN.(d) below. The above sequence $(Y_k(\mathbf{m}))_k$ has the just stated properties, and therefore if $S_N \equiv S_N(\mathbf{m}) := \sum_{k=1}^{N} Y_k(\mathbf{m})$, then

$$\lim_N \frac{S_N}{N} = \lim_N \frac{1}{N} \sum_{k=1}^{N} Y_k(\mathbf{m}) = 0, \quad \text{almost surely.} \tag{2.6}$$

Recalling (2.4) and the fact that $\mathbf{m} \in \mathbb{Z}^d \setminus \{0\}$ was fixed but arbitrary, gives (2.3) (λ-a.e. is the same as P-a.s,) and therefore the above stated (simple) equidistribution in $[0, 1]^d$ of the sequence of vectors $(\boldsymbol{\beta}_k)_k$.

LLN.(a) Note that $ES_N = 0$ and $\text{var}(S_N) = O(N)$.

LLN.(b) Use the Chebyshev (or the Markov) inequality, and the Borel-Cantelli lemma on the subsequence $N_n = n^2$ to conclude that

$$\sum_{n=1}^{\infty} P(|S_{n^2}| > n^2 \varepsilon) = \sum_{n=1}^{\infty} O(1/n^2) < \infty, \quad \forall \varepsilon > 0,$$

and therefore that $S_{N_n}/N_n \to 0$, almost surely.

LLN.(c) Since $E(\sum_{N \in [n^2+1,(n+1)^2]} |S_N - S_{n^2}|^2) = \sum_{N \in [n^2+1,(n+1)^2]} O(N - n^2) = O(n^2)$, and therefore $E(\max_{N \in [n^2+1,(n+1)^2]} |S_N - S_{n^2}|^2) \leq Dn^2$, another application of the Markov inequality gives that with probability greater than $1 - \frac{2D}{\varepsilon^2 n^2}$

$$|S_N - S_{n^2}| \in [-\varepsilon n^2, \varepsilon n^2], \quad \forall N \in [n^2 + 1, (n+1)^2 - 1], \tag{2.7}$$

and again by the Borel-Cantelli lemma, that (2.7) happens with probability 1 for all but finitely many n.

LLN.(d) Divide (2.7) by n^2, noting that $N = n^2 + O(n)$. Use the fact $\varepsilon > 0$ was arbitrary, and the conclusion of (b).

Remark 2. (a) In our special setting, each member of the sequence $(Y_k)_k$ is uniformly bounded below (and above), so (2.7) could have been replaced by a simpler estimate: with probability 1 for all n

$$|S_N - S_{n^2}| \in [-(N - n^2)c, ((n+1)^2 - N)c], \quad N \in [n^2 + 1, (n+1)^2 - 1].$$

(b) If $r = d = 1$, then (1.6) and (1.7) coincide, leading to the conclusion that if $(x_k)_k$ is again either from (1.3) or in (1.4), then the corresponding $(\beta_k)_k$ is simply equidistributed. The study of multiple equidistribution can be similarly set up (see also Section 2.1.3 below): here (2.1) has the same form, but we take $d \geq 2$ and $\boldsymbol{\beta} := (\boldsymbol{\beta}_k^d)_k$ defined as in (1.7) with $(x_k)_k$ linear (as in (1.3,1.4)), and study the averaged Weyl sums of terms $Y_k(\mathbf{m}; t) = e(\mathbf{x}_k(t), \mathbf{m})$, for $t \in G$, for each $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$.

(c) Recall the multiplicatively generated sequence from (1.4). To see that it is not 2-multiply equidistributed, take $\mathbf{m} = (M, -1) \neq (0, 0)$ and note that with this choice of $\mathbf{m}$ the corresponding criterion (2.3–2.4) converges nowhere to 0. Indeed, $m_1 x_k + m_2 x_{k+1} \equiv 0$, for any $k \in \mathbb{N}$.

Similarly, it is possible to find a non-trivial $p + 1$-dimensional vector $\mathbf{m}$ such that

$$m_1 k^p + m_2(k+1)^p + \ldots + m_p(k+p-1)^p + m_{p+1}(k+p)^p \text{ is constant over } k.$$

This amounts to solving a linear system $A\mathbf{x} = \mathbf{0}$, where A has p rows and $p + 1$ columns. The corresponding criterion (2.3–2.4) applied to Weyl numbers (1.3) will again converge nowhere to 0 (non-convergence to 0 on an event of positive probability would already be sufficient).

2.1.1. Equivalent formulations of complete equidistribution

Let $d \geq 1$ be fixed. The d-dimensional (or multiple) *discrepancy* at level N is defined by

$$D_N^d(\mathbf{t}) := \sup_{A \in J} \left| \frac{\sum_{k=1}^N \mathbb{1}_A(\boldsymbol{\beta}_k^d(\mathbf{t}))}{N} - \lambda(A) \right|, \quad \mathbf{t} \in G,$$

where J is the family of boxes as in definition (1.2). The d-dimensional *"star" discrepancy* at level N is defined by

$$D_N^{*,d}(\mathbf{t}) := \sup_{A \in J^*} \left| \frac{\sum_{k=1}^N \mathbb{1}_A(\boldsymbol{\beta}_k^d(\mathbf{t}))}{N} - \lambda(A) \right|, \quad \mathbf{t} \in G,$$

where J^* is the family of boxes in $[0,1]^d$ as above, with lower left corner equal to $\mathbf{0}$. Equivalently, the supremum above is taken over all boxes A of the form $\prod_{i=1}^d (0, b_i]$, where $0 \leq b_i \leq 1$, for each i. Note that we again deviate slightly from the standard definitions (see Kuipers and Niederreiter [KN]), where discrepancy sequences are defined for deterministic sets or sequences. Easy properties of measurability of functions make each D_N^d and $D_N^{d,*}$ a random variable in the current (stochastic a.s./metric) setting. Note that $D_N^{*,d}$ is a direct extension of the *Kolmogorov-Smirnov* statistic to the equidistribution setting.

Recall (2.1) with $\boldsymbol{\beta}_k = \boldsymbol{\beta}_k^d$, $k \geq 1$. As any probabilist knows (think about *cumulative distribution functions*), the weak convergence of the sequence of (random) measures ν_N to the uniform law on $[0,1]^d$ is, omega-by-omega (where "omega" is typically denoted by $\mathbf{t}$), equivalent to the convergence of $(D_N^d)_N$ (or equivalently of $(D_N^{*,d})_N$) to 0 (a clear generalization of the Glivenko-Cantelli theorem). Let us paraphrase: $(\beta_k)_k$ is d-multiply equidistributed in $[0,1]$ if and only if $\lim_N D_N^d = \lim_N D_N^{*,d} = 0$, almost surely. In particular, a sequence $(\beta_k)_k$ is a c.e.s. in $[0,1]$ if and only if

$$\lim_N D_N^d = \lim_N D_N^{*,d} = 0, \ \forall d \in \mathbb{N}, \ \text{almost surely.} \tag{2.8}$$

Note that $I_\infty := [0,1]^{\mathbb{N}}$ is a product of compact spaces, and therefore compact itself. Let $\mathcal{B}_m$ be the Borel σ-field on $[0,1]^m$. Consider the *cylinder sets* $C \subset I_\infty$, such that $C = C_b \times [0,1]^{\mathbb{N}}$, for some $C_b \in \mathcal{B}_m$ and $m \in \mathbb{N}$. Let $\mathcal{B}_\infty$ be the σ-field on I_∞ generated by the cylinders. Instead of d-dimensional vectors, one could consider straight away the "∞-dimensional" (random) vector sequence

$$\mathbf{x}_k(\mathbf{t}) = (x_k(\mathbf{t}), x_{k+1}(\mathbf{t}), x_{k+2}(\mathbf{t}), \ldots), \ k \geq 1, \ \mathbf{t} \in G,$$

and its corresponding $\boldsymbol{\beta}_k^\infty := \mathbf{x}_k \ \mathsf{mod} \ \mathbf{1}$, where again "modulo" operation is applied component-wise (a.s.). Let ν_N^∞ be as ν_N in (2.1), but with $\boldsymbol{\beta}$ redefined as $\boldsymbol{\beta}^\infty$. Another formulation of complete equidistribution can be read off from an "abstract fundamental theorem" (see e.g. [KN], Ch. 3, Theorem 1.2 and the remark following it) or easily proved by approximating all closed sets in

$\mathcal{B}_\infty$ by closed cylinders: $(\beta_k)_k$ is a c.e.s. in $[0,1]$ if and only if $(\nu_N^\infty)_N$ converges weakly to the uniform law on I_∞. A realization Γ from this limiting uniform law (that is, the random object having that law) is also called the *infinite statistical sample* or the *i.i.d. family of uniform* $[0,1]$ *random variables* $(U_1, U_2, \ldots)$. The just made statement could also be reformulated as follows: $(\beta_k)_k$ is a c.e.s. in $[0,1]$ if and only if for any f bounded and continuous function on I_∞ (equipped with the product topology), we have

$$\lim_N \frac{1}{N} \sum_{k=1}^{N} f(\beta_k^\infty) = \mathbf{E}\left(f(\Gamma)\right), \text{ almost surely.} \qquad (2.9)$$

As indicated in Section 1.1, another formulation/criterion of complete equidistribution for a (deterministic) sequence of real numbers in $[0,1]$ was derived by Knuth [Kn1, Kn2]. There is no doubt that this could also be turned into a stochastic (a.s.) formulation for a sequence (1.1), the details are left to an interested reader.

2.1.2. Why not simply take an i.i.d. family of uniforms?

The first author can guess that, especially on the first reading, a non-negligible fraction of fellow probabilists could be asking the above or a similar question. It is clear that an i.i.d. sequence Γ of uniform random variables is a c.e.s. in $[0,1]$ in the sense of Definition 1.1, or any of its equivalent formulations described in the previous section. Studying complete equidistribution, starting from i.i.d. (or similar) random families is precisely what probability oriented works [Ho1, Ho2, Lac, Lo] did. However, to most non-probabilist mathematicians this will not mean much, especially due to the fact that the rigorous probability theory is axiomatic, and the rigorous construction of Γ rather abstract.

More precisely, Γ is not presented in a neat (classical) functional format, like any of the sequences $(x_k)_k$ and their corresponding $(\beta_k) = (x_k \bmod 1)_k$ in (1.3,1.4,1.5). Instead we usually start with an abstract infinite product space (or more concretely, with $[0,1]^\mathbb{N}$), and the kth uniform random variable equals the identity map from the kth space to itself. This suffices for the most purposes of modern probability theory, but may not seem very convincing to a non-probabilist who wishes to "see a concrete example of" Γ. A related wish to "see an explicit outcome from a concrete example of Γ" quickly leads to philosophical discussions around the question "What is a random sequence?" (the reader is referred to the section in [Kn2] carrying that very title).

The Kac [Ka1] approach to the i.i.d. discrete-valued sequence is also revolutionary from the point of view of the just mentioned drawback. Indeed, an infinite sequence of independent Bernoulli(1/2) random variables (also called the *Bernoulli scheme*) can be explicitly constructed on $\Omega = [0,1]$, with $\mathcal{F}$ equal to the Borel σ-field $\mathcal{B}$, and P equal to λ. This was done in [Ka1], by applying a simple transformation to the classical Rademacher system. We leave on purpose

the link with (1.4), and other details, to interested readers, as well as the discovery of the related b-adic Rademacher system and its relation to the discrete uniform law on $\{0, 1, \ldots, b-1\}$.

Given an infinite sequence $X := (X_i)_{i=1}^{\infty}$, which has the Bernoulli scheme distribution (take for example the one from the above recalled construction by Kac), one can define Γ on the same $(\Omega, \mathcal{F}, P)$ by "redistributing the digits" via a triangular scheme, for example $U_1 := X_1/2 + X_3/2^2 + X_6/2^3 + X_{10}/2^4 + \cdots$, $U_2 := X_2/2 + X_5/2^2 + X_9/2^3 + \cdots$, $U_3 := X_4/2 + X_8/2^2 + \cdots$, $U_4 := X_7/2 + \cdots$, and so on, and finally $\Gamma = (U_i)_i$. Is this asymptotic definition of Γ on $([0,1], \mathcal{B}, P)$ sufficiently explicit? Or are the examples like (1.5,2.10,2.11) - all leading (after mod 1 application) to c.e.s. but none to i.i.d. uniforms - more reassuring to think about (more amenable to analysis)? The answer will likely vary from one peer to another, depending not only on their mathematical background and research interests, but also on their personal perception of randomness. Note that such and related questions challenged the very founders of probability theory only 80 years ago (see e.g. the historical notes of Knuth [Kn2], Section 3.3.5, or [Lam, Bu1, Bu2]).

For reader's benefit, we mention in this paragraph two other prominent approaches to randomness, nowadays practically forgotten by the probability and statistics communities: the von Mises collectives, and Martin-Löf sequences. The original definition [Ms] of *Kollektiv* (or *collective*) is not transparent. Following [Bu2] (e.g. page 49), a collective is an infinite sequence of observations such that the relative frequency of an event converges to the same number along every subsequence chosen without prophetic powers; this common limit is called the probability of the event in the collective. For further interpretations, discussions and more see [Lam, Bu1, Bu2]. Continued search for (more concrete) examples of mathematically generated randomness led to *Kolmogorov's complexity* [Kol2], and the *Martin-Löf sequences* (see [ML] and [Le1]). Informally speaking, Martin-Löf [ML] proves existence of binary sequences that pass the "super-test for randomness". Any such sequence can be rightfully called random. Concreteness of this definition is again relative, and therefore left to an individual appraisal (Burdzy gives an interesting commentary in [Bu2], pp. 77–78).

2.1.3. Novel classes of examples of c.e.s.

Define

$$x_k(t) = k^k t, \quad t \in G := (0,1), \ k \geq 1. \tag{2.10}$$

and

$$x_k(t) = k! t, \quad t \in G := (0,1), \ k \geq 1. \tag{2.11}$$

We claim that (1.1) defined using either (2.10) or (2.11) is completely equidistributed in the sense of Definition 1.1.

Let us take (2.11) for example. Simple equidistribution can be quickly deduced as in the previous section (see Remark 2(b) in particular). Now fix some $d \geq 2$ and $\mathbf{m} \in \mathbb{Z} \setminus \{\mathbf{0}\}$. WLOG we can assume that the Weyl criterion (2.3)

has been shown for the $(d-1)$-multiply case. Moreover, due to symmetry, we can and will assume $m_d > 0$. We have

$$Y_k(t) \equiv Y_k(\mathbf{m}\,;t) := e(\mathbf{m}, \mathbf{x}_k(t)), \tag{2.12}$$

which in this special case reads $Y_k(t) = \exp(2\pi i\, t \sum_{i=1}^{d} m_i(k+i-1)!)$, $t \in [0,1]$. It is clearly true that $E(Y_k) = 0$, and $E|Y_k|^2 = 1$, for each $k \in \mathbb{N}$. It is furthermore easy to see that whenever $k - l > c(\mathbf{m})$ for some $c(\mathbf{m}) \in \mathbb{N}$, then $\sum_i m_i[(k+i-1)! - (l+i-1)!]$ is a strictly positive integer, and due to the periodicity of sine and cosine, we have again $E(Y_k\overline{Y_l}) = \overline{E(\overline{Y_k}Y_l)} = 0$ (or equivalently, Y_k and Y_l are uncorrelated). Moreover, $|Y|$s are uniformly bounded by 1. Therefore

$$\mathrm{var}\left(\sum_{k=1}^{N} Y_k\right) = N + \sum_{k=1}^{N} \sum_{l=k-c(\mathbf{m})}^{k-1} 1 = N(1 + c(\mathbf{m})),$$

and again the argument LLN.(b)-LLN.(d) (using the simplification from Remark 2(a)) applies. By induction on d, we obtain the above stated complete equidistribution. The proof for (2.10) follows the very same steps. The reader is undoubtedly able to construct further new examples of c.e.s. for which the elements of $(Y_k)_k$ are pairwise uncorrelated (except for nearby indices).

2.2. The Weyl variant of the SLLN and generalizations

It is easy to see that the steps in LLN.(a)-LLN.(d) could be applied to show simple equidistribution in $[0,1]$ of (1.1), whenever $x_k(t) := a_k t$, $k \geq 1$ and $(a_k)_k$ is a sequence of distinct integers, yielding an alternative derivation of [KN] I, Theorem 4.1. This result is already due to Weyl [We], and its original proof made a profound impact on the development of (analytic) number theory. Indeed, in the above and analogous situations (see e.g. [KN, MF, DT], as well as [Ho1], Theorem 2), it is standard to apply the following variant of the SLLN, due to Weyl [We]: since $E|S_N|^2 = O(N)$, then

$$\sum_n E\left(\frac{|S_{n^2}|}{n^2}\right)^2 < \infty,$$

and therefore by Tonelli's theorem $\sum_n \frac{|S_{n^2}|}{n^2} < \infty$ almost surely, and in particular $S_{n^2}/n^2 \to 0$, as $n \to \infty$. Finally use Remark 2(a) to get (2.6).

The Weyl variant of the SLLN has the following important extension, due to Davenport, Erdös and LeVeque [DELV], that has been since used extensively in number theoretic studies (see e.g. [KN] I, Theorem 4.2):

$$\text{if} \quad \sum_n \frac{1}{n} E\left(\frac{|S_n|}{n}\right)^2 < \infty \quad \text{then (2.6).} \tag{2.13}$$

Its proof strongly relies on the uniform boundedness of the individual random variables Y, but not at all on their particular (complex exponential) form (see Lyons [Ly], Theorem 1). Lyons [Ly] takes moreover general complex-valued sequence $(Y_k)_k$, and derives different generalizations of the above SLLN criterion of [DELV], where the uniform boundedness condition is replaced by various bounded moment conditions.

Remark 3. In view of the SLLN stated here, the fact that the sequences of functions from Section 2.1.3 generate c.e.s. is again trivial. Just like with the usual SLLN (of LLN.(a)-(d) and Remark 2(a)), it is important here that all the covariances can be computed or adequately (uniformly) estimated.

2.3. Weyl-like criterion for weakly c.e.s.

Recall the (random) discrepancy sequences introduced in Section 2.1.1. Owen and Tribble (e.g. [OT] Definition 2 or [TO] Definition 5) define *weak complete equidistribution* (aka WCUD) in terms of the weak(er) convergence of the discrepancy sequence as follows: $(\beta_k)_k$ is WCUD if and only if $D_N^{*,d} \Rightarrow 0$, or equivalently (since the limit is deterministic), if $D_N^{*,d} \to 0$ in probability, for each $d \geq 1$.

The following "a.s.-convergence along subsequences" characterization is well-known: a sequence of random variables $(X_n)_n$ converges in probability to X, if and only if for any subsequence $(n_k)_k$, one can find a further subsequence $(n_{k(j)})_j$ such that $X_{n_{k(j)}}$ converges to X almost surely. Let us apply it to the discrepancy sequences, and conclude (due to the countability of $\mathbb{N}$, and the Arzelà-Ascoli diagonalization scheme) that for any subsequence $(n_k)_k$ one can find a further subsequence $(n_{k(j)})_j$ and a single event (Borel measurable set) G_f of full probability such that

$$D_{n_{k(j)}}^{*,d}(\mathbf{t}) \to 0, \quad \text{for all } d \geq 1, \text{ and all } \mathbf{t} \in G_f \subset G.$$

But now we recall (as in Section 2.1.1) that the above convergence is equivalent ($\mathbf{t}$-by-$\mathbf{t}$) to the simultaneous (for each $d \in \mathbb{N}$ and on the full probability event G_f) weak convergence of the random sequence $(\nu_{n_{k(j)}}^d)_j$ to the corresponding uniform law λ on $[0,1]^d$, where

$$\nu_N^d := \sum_{k=1}^N \frac{1}{N} \delta_{\boldsymbol{\beta}_k^d}, \ d \in \mathbb{N}.$$

The last made claim is in turn equivalent (recall the reasoning of Section 2.1 and definition (2.12)) to the statement

$$\lim_j \frac{1}{n_{k(j)}} \sum_{k=1}^{n_{k(j)}} Y_k(\mathbf{m}) = 0, \text{ on } G_f, \text{ for each } \mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}, \text{ and each } d \in \mathbb{N}.$$

(2.14)

This leads to the following conclusion: for each subsequence $(n_k)_k$ one can find a further subsequence $(n_{k(j)})_j$ such that for each $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ the rescaled Weyl

sums indexed by $\mathbf{m}$ converge to 0 along that sub(sub)sequence, almost surely. In other words, for each d and $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$, instead of (2.6) we arrived to

$$\lim_N \frac{S_N}{N} = \lim_N \frac{1}{N} \sum_{k=1}^N Y_k(\mathbf{m}) = 0, \quad \text{in probability.} \tag{2.15}$$

Finally recalling that, in our special setting, $|S_N|/N$ is uniformly bounded by 1, the (adaptation of) Lebesgue dominated convergence theorem says that (2.15) is equivalent to

$$\lim_N E \frac{S_N}{N} = \lim_N E \frac{|S_N|}{N} = 0. \tag{2.16}$$

We record the above sequentially made equivalences as

LEMMA **2.1** (Sufficiency and necessity for weak c.e.s. (aka WCUD)). *Let* $(x_k)_k$ *be a sequence of random variables (or measurable functions on G), $Y_k(\mathbf{m})$ be defined by (2.12), and $(\beta_k)_k$ by (1.1). Then $(\beta_k)_k$ is weakly completely equidistributed in $[0,1]$ if and only if (2.16) is valid for each d and $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$.*

In view of Lemma 2.1, the claim from [TO] that WCUD sequences are hard to construct seems rather surprising.

2.4. Extensions to the non-linear setting

As already noted, the reasoning of Section 2.1 is crucially connected to linearity only through (2.5). In particular, if we were to take (1.5) or some other sequence of functions, then apply (1.7) to obtain d-dimensional vector sequences, and plug them into (2.2,2.4,2.12), the criterion (2.3), to be verified through (2.6), would stay the same. Note however that the correlation structure may (and typically does) become much more complicated than that given in (2.5) or in Section 2.1.3. For a detailed discussion of problems with non-linearity see [Ai].

In the special case of Koksma's numbers, the probability P is set to λ renormalized by $(a-1)$ on $[1,a]$, or equivalently, P is the uniform law on $[1,a]$. Before considering covariance, one could already note that Y_k may not (necessarily) be even centered. Since $Y_k(t) \equiv Y_k(\mathbf{m};t) := e(\mathbf{m}, \mathbf{x}_k(t))$, $t \in [1,a]$, we have that

$$E(Y_k) = \frac{1}{a-1} \left[\int_1^a \cos(2\pi \mathbf{m} \cdot \mathbf{x}_k(t)) \, dt + i \int_1^a \sin(2\pi \mathbf{m} \cdot \mathbf{x}_k(t)) \, dt \right]. \tag{2.17}$$

The non-zero expectation does not matter much, if one could show that the long term average of $(Y.)$ is approximately centered, or equivalently, that

$$\lim_N \frac{1}{N} \sum_{k=1}^N EY_k = 0.$$

The covariance analysis to be done below will imply an even stronger estimate $E(|S_N|) = O(N^{1-1/d})$, but already from Jensen's inequality we get that if $E(|S_N|^2) = o(N^2)$, then $E|S_N| = o(N)$, which would be sufficient.

The argument LLN.(a)-LLN.(d) will continue to work for the sequence $(Y_k - EY_k)_k$, and will imply (2.6) for $(Y_k)_k$, even in the presence of correlation, provided that the total covariance

$$\sum_{k=1}^{n} \sum_{l=k+1}^{n} |E(Y_k \overline{Y_l}) + E(Y_l \overline{Y_k})|$$

grows sufficiently slowly in N. Indeed, a little thought is needed to see that if

$$|E(Y_k \overline{Y_l} + Y_l \overline{Y_k})| = \frac{1}{a-1} \left| \int_{[1,a]} \cos(2\pi \, \mathbf{m} \cdot (\mathbf{x}_k(t) - \mathbf{x}_l(t)) \, dt \right| \le c|k-l|^{-\delta}, \ k \ne l,$$

(2.18)

for some $c = c(a, \mathbf{m}, d) < \infty$, $\delta = \delta(a, \mathbf{m}, d) > 0$, then the conclusion (2.6) remains (this time choose $N_n = \lfloor n^{2/\delta} \rfloor$, and then apply the sandwiching argument). Alternatively, this can be immediately deduced form the SLLN criterion (2.13).

We record the just made observations in form of a lemma.

LEMMA **2.2** (Sufficiency for complete equidistribution I)**.** *Let $(x_k)_k$ be a sequence of random variables (or measurable functions on G), $Y_k(\mathbf{m})$ be defined by (2.12), and $(\beta_k)_k$ by (1.1). If for each $\mathbf{m}$, $|E(Y_k(\mathbf{m})\overline{Y_l}(\mathbf{m}) + Y_l(\mathbf{m})\overline{Y_k}(\mathbf{m}))| = O(|k-l|^{-\delta})$, for some $\delta(\mathbf{m}) > 0$, then $(\beta_k)_k$ is completely equidistributed in $[0,1]$.*

Deriving (2.18) is a longer calculus exercise, sketched here for reader's benefit. Fix $d \ge 1$, $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$, and consider $\boldsymbol{\beta}_k^d = (\beta_{k1}, \ldots, \beta_{kd})$ defined in (1.7). Recalling that $x_k(t) = t^k$ we have $\mathbf{m} \cdot (\mathbf{x}_k(t) - \mathbf{x}_l(t)) = \sum_i m_i(t^{k+i-1} - t^{l+i-1})$. If $p(t) \equiv p(t; \mathbf{m}) = \sum_i m_i t^{i-1}$ and $g(t) \equiv g(t; k, l) = t^k - t^l$, then

$$\mathbf{m} \cdot (\mathbf{x}_k(t) - \mathbf{x}_l(t)) = p(t)g(t), \quad t \in G = [1, a]. \tag{2.19}$$

One can assume WLOG that $k > l$, so that $g(t)$ is non-negative on G, with a single zero $t = 1$. Let d_* equal to the maximal index i such that $m_i \ne 0$. The polynomial p does not depend on k, l. It has degree $d_* - 1$ and therefore at most $d_* - 1 \le d - 1$ real zeros, each of which may fall into G. Therefore $p(\cdot)g(\cdot)$ takes value 0 at 1, and at most $d - 1$ other points in G. Let us denote these zeros by $(z_j)_{j=1}^r$, where $1 =: z_0 \le z_1 < \ldots < z_r \le a =: z_{r+1}$. For $j = 1, \ldots, r$, let l_j be the multiplicity of z_j for p. Extend this definition to $l_0 = 0$ if $z_1 > 1$, and $l_{r+1} = 0$ if $z_r < a$. Define

$$b_1 := \max_{t \in [1,a]} |p'(t)|, \quad b_2 := \max_{t \in [1,a]} |p''(t)|.$$

Keeping in mind that $r \le d - 1$, we now let $I_i := [z_i, z_{i+1}]$, and estimate

$$\left| \int_{I_i} \cos(2\pi \, g(t)p(t)) \, dt \right| \tag{2.20}$$

separately for each $i = 0, \ldots, r$. The non-zero polynomial p does not change sign on I_i, so WLOG we can assume that it takes positive values in the interior of I_i. Moreover

$$\min_{z \in I_i} p(z)/((z - z_i)^{l_i}(z_{i+1} - z)^{l_{i+1}}) =: c_i > 0. \tag{2.21}$$

If $z_{i+1} - z_i \leq \frac{3}{(k-l)^{1/d}}$ then (2.20) is clearly bounded by $3/(k-l)^{1/d}$. Otherwise, define $I'_i := [z_i + \frac{1}{(k-l)^{1/d}}, z_{i+1} - \frac{1}{(k-l)^{1/d}}]$. Split the domain of integration in (2.20) into three disjoint pieces: $[z_i, z_i + \frac{1}{(k-l)^{1/d}})$, I'_i and $(z_{i+1} - \frac{1}{(k-l)^{1/d}}, z_{i+1}]$. The integral of $\cos(\cdot)$ over the first and the final piece is again trivially bounded above by $1/(k-l)^{1/d}$. For the middle piece, it suffices to show that:

(a) $(pg)'(z_i + \frac{1}{(k-l)^{1/d}}) \geq \bar{c}_i |k-l|^{1/d}$, where $\bar{c}_i > 0$ and that
(b) $(pg)''(t) = g''(t)p(t) + 2g'(t)p'(t) + g(t)p''(t) > 0$ on I'_i.

Indeed, (a)-(b) would imply that $2\pi g(\cdot)p(\cdot)$ increases more and more rapidly on I'_i, and in turn that $\cos(2\pi\, g(t)p(t))$ makes shorter and shorter "excursions" of alternating sign, away from 0. This would yield an upper bound for the integral over I'_i in terms of $t^*_i - (z_i + \frac{1}{(k-l)^{1/d}})$ where t^*_i is the second zero of $\cos(2\pi\, g(\cdot)p(\cdot))$ in $[z_i + \frac{1}{(k-l)^{1/d}}, \infty]$. From (a) and (b) one could easily conclude that $t^*_i - (z_i + \frac{1}{(k-l)^{1/d}}) \leq (2\bar{c}_i)^{-1}(k-l)^{-1/d}$. The just made reasoning justifies the upper bound $(2 + (2\bar{c}_i)^{-1})(k-l)^{-1/d}$, which in turn implies the upper bound (2.18) with $\delta = 1/d$ and $c(a, \mathbf{m}, d) := \sum_{i=0}^{r}(2 + (2\bar{c}_i)^{-1})$. Alternatively, one could use [KN] I, Lemma 2.1.

To show (b), note that $g''(t) = k(k-1)t^{k-2} - l(l-1)t^{l-2}$ is greater than $(k^2 - l^2 + O(k+l))t^{k-2}$. We also have $g'(t) = kt^{k-1} - lt^l \leq ka \cdot t^{k-2}$ and $g(t) \leq a^2 \cdot t^{k-2}$. Due to (2.21) and the fact $l_i + l_{i+1} \leq d_* - 1 \leq d - 1$, the leading term of $(pg)''(t)$ is positive and bounded below by $c_i |k-l|^{1/d}(k+l)t^{k-2}$ on I'_i, while the two other terms are bounded above by $b_1 kat^{k-2}$ and $b_2 a^2 t^{k-2}$, respectively, yielding (b). We leave to the reader a similar (and easier) argument for (a).

Finally note that the derivation of the bound in (2.18) applies also to estimating both the real and the imaginary part of (2.17), and leads to an analogous bound $EY_k = O(k^{-1/d})$.

2.4.1. *Koksma's numbers have even stronger properties*

Koksma's derivation [Kok] of the simple equidistribution of Koksma's numbers was similar to that included above (see also [KN] I, Theorem 4.3), and easier due to the fact that a simpler polynomial $m \cdot g$ replaces $p \cdot g$ in (2.19). As already mentioned in the introduction, Franklin [Fr] found a way (again similar to the calculus exercise above) of building on Koksma's analysis, without directly applying Weyl's criterion for multiple equidistribution.

Koksma [Kok] showed in addition a stronger type of equidistribution: suppose that $(a_k)_k$ is a sequence of positive distinct integers, and consider the reordering (with possible deletion) of Koksma's numbers obtained via (1.1) from $(x_{a_k})_{k \geq 1}$ defined as in (1.5), the obtained sequence is again simply equidistributed. Niederreiter and Tichy [NT1] proved that the above arbitrarily permuted (sub)sequence of Koksma numbers is in fact completely equidistributed, thus confirming a guess of Knuth's [Kn2], who called the above property pseudorandom of R4 type. As it turns out, the arguments in [NT1] are equally based

on probabilistic reasoning (using the [DELV] variant (2.13) of the SLLN and covariance calculations). For the benefit of the reader we sketch it next.

Suppose initially that, for each $\mathbf{m} \neq \mathbf{0}$, $|E(Y_k(\mathbf{m})\overline{Y_l}(\mathbf{m}) + Y_l(\mathbf{m})\overline{Y_k}(\mathbf{m}))|$ is bounded by $c(\mathbf{m})/\log^2 N$, unless $|k - l| \leq \sqrt{N}$ or $\min\{k, l\} \leq \sqrt{N}$. Then it is easy to see that $E|S_N|^2 \leq N + 2\sum_{k=1}^{N}\sqrt{N} + \sum_{k=\sqrt{N}}^{N}\sum_{l=k+\sqrt{N}}^{N} c(\mathbf{m})/\log^2 N = N + N^{3/2} + N^2/\log^2 N$ which is $O(N^2/\log^2 N)$. The criterion (2.13) implies the required convergence in the Weyl criterion. We again record the just made observations in form of a lemma.

LEMMA **2.3** (Sufficiency for complete equidistribution II). *let $(x_k)_k$ be a sequence of random variables (or measurable functions on G), $Y_k(\mathbf{m})$ be defined by (2.12), and $(\beta_k)_k$ by (1.1). If for each $\mathbf{m}$, $|E(Y_k(\mathbf{m})\overline{Y_l}(\mathbf{m}) + Y_l(\mathbf{m})\overline{Y_k}(\mathbf{m}))| = O(1/\log N^2)$, whenever $\min\{l, k - l\} \geq \sqrt{N}$, then $(\beta_k)_k$ is completely equidistributed in $[0, 1]$.*

The proof in [NT1] uses Lemma 2.3 with a small (and natural) twist: instead of the original indexing/ordering, for each N, $\mathbf{m}$, to each $k \in \{1, \ldots, N\}$ one assigns $b(k) := \max_{j=1}^{d}\{a_{k+j} : m_j \neq 0\}$, and then estimates $|E(Y_k(\mathbf{m})\overline{Y_l}(\mathbf{m}) + Y_l(\mathbf{m})\overline{Y_k}(\mathbf{m}))|$ whenever $\min\{|b(k) - b(l)|, b(k), b(l)\} \geq \sqrt{N}$. The rest of the argument is exactly as described above (without the probabilistic rescaling by $a - 1$).

Further qualitative and quantitative improvements and extensions, to be recalled soon, were done in the late 1980s by Niederreiter and Tichy [NT2], Tichy [Ti], Drmota, Tichy and Winkler [DTW], and Goldstern [Go] (see also [DT], Section 1.6). And yet, interesting non-trivial open problems remain, as indicated in the next section.

3. Concluding remarks with open problems

The main theorem of [Li] also yields novel generators of (stochastic a.s.) completely equidistributed sequences in the non-linear setting, of which the prototype is $x_k(t) = (t(\log t)^{r_k})^{w_k}$, $k \geq 1$, where $(r_k)_k$ (resp. $(w_k)_k$) are sequences of positive (resp. natural) numbers satisfying certain hypotheses. This may not be so interesting (even though there is no obvious reduction of such sequences to exponential sequences of Section 2.4.1) in view of the main theorem of Niederreiter and Tichy [NT1, NT2]. However, the approach to (complete) equidistribution, implicit in [Li] and made explicit in Section 2, is interesting. In particular, it leads to the realization that [NT1] is based on probability arguments.

A careful reader must have noticed that (2.13) is only a sufficiency condition for the SLLN (2.6). Davenport, Erdös and LeVeque also state (in [DELV], main/only Theorem) "On the contrary...", however these counter-examples do not either imply nor disprove the necessity of assumption

$$\sum_n \frac{1}{n}E\left(\frac{|S_n|}{n}\right)^2 < \infty,$$

for the SLLN of the corresponding Weyl sums. Indeed, it could be that x_k are purely constant (random variables), and such that for some $\mathbf{m} \in \mathbb{Z}^d \setminus \mathbf{0}$ we have

that $\frac{E|S_n(\mathbf{m})|^2}{n^2} = \frac{|S_n(\mathbf{m})|^2}{n^2}$ is both $\Omega(1/\log n)$ and $o(1)$ (Korobov [Kr4] guarantees that such examples exist), so that the SLLN happens even though the series in the criterion diverges. However, here we would trivially have, for the same d and $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$, that

$$\sum_n \frac{1}{n}\mathrm{var}\left(\frac{|S_n(\mathbf{m})|}{n}\right) < \infty, \text{ and } \lim_n \frac{E|S_n(\mathbf{m})|^2}{n^2} = 0. \qquad (3.1)$$

While the above example is arguably contrived, it already illustrates the following straight-forward consequence of (2.13) (or more precisely, of its generalization [Ly], Theorem 1): (3.1) is a (strictly) stronger sufficiency criterion for the SLLN of the corresponding (indexed by $\mathbf{m}$) Weyl sums, than that given in [DELV] (see (2.13)). Is it necessary? If not, is the whole d-dimensional collection of them (meaning that (3.1) is true for all $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$) necessary for d-multiple equidistribution of the corresponding $\boldsymbol{\beta}$? If not, is the complete collection of them (meaning that (3.1) is true for all $\mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$ and all $d \in \mathbb{N}$) necessary for complete equidistribution of the corresponding $\boldsymbol{\beta}$? If not, are there natural additional (easy to check) hypotheses on $(x_k)_k$ under which these families of criteria become necessary (they are always sufficient)?

As pointed out in Section 2.4.1, we know from [NT1] that Koksma's numbers are of R4 type and that the same is true, as shown in [NT2], for a large class of other exponentially generated sequences. A trivial example of a c.e.s. which is also R4 (and even R6 in Knuth's terminology) is the infinite statistical sample Γ, but as explained in Section 2.1.2, this is likely not considered sufficiently explicit from a non-probabilist's perspective. The examples from Section 2.1.3 must have been known to the modern analytic number theory community, in particular as examples of so-called *lacunary sequences*. We could not find them on the list of c.e.s. in any the standard references, however (2.11) was studied by Korobov [Kr3] in connection to an explicit example of a simply equidistributed sequence. They seem amenable to analysis, and natural for continuing the investigation in Knuth's framework, who anticipated the result of [NT1], but at the time knew only the work of Franklin and Koksma about Koksma's numbers (and variations thereof) [Kok, Fr]. Are linear c.e.s. from Section 2.1.3 also of R4 type? Is it possible that any sufficiently random c.e.s. (e.g. $\mathrm{var}(\beta_k \mathbb{1}_{[c,d]}) > 0$ for all k and and all $0 \le c < d \le 1$) is of R4 type? If not, is there a natural and easy to check characterization of when the complete equidistribution (type R1) implies R4? By the way, note that if an infinite sequence of random variables $(\beta_k)_k$ on $([0,1], \mathcal{B}, \lambda)$ is both completely equidistributed in $[0,1]$ (that is, of type R1) and *exchangeable* (see e.g. [Du], Example 5.6.4), then it must be equal in law to Γ. Is there a natural exchangeability-like property (clearly stronger than R4 type), though weaker than exchangeability, that would jointly with c.e.s. still imply that the sequence $(\beta_k)_k$ has the law of Γ? Is there a natural condition (stronger than c.e.s. but weaker than i.i.d.) that would imply, jointly with R4 type, that the sequence $(\beta_k)_k$ has the law of Γ?

Once equidistribution is established, it is natural to ask about the rate of convergence in (1.2), or in analogous multi-dimensional expressions that corre-

spond to Definitions 1.1-1.2. The discrepancy sequences (as recalled in Section 2.1.1) seem to be the principal object for this analysis, already in the focus of classical studies by the founders of analytic number theory (see [KN], Chapter 2 and [DT] Sections 1.1–1.2). From the stochastic point of view, it seems more interesting to study discrepancy as a random variable. In particular, the well-known *Erdős-Turán-Koksma inequality* (see e.g. [DT], Theorem 1.21) can be restated in the present setting as follows: if H is a positive integer and, for $\mathbf{m} \in \mathbb{Z}^d$, we let $r(\mathbf{m}) := \prod_{i=1}^{d} \max\{1, |m_i|\}$, then it is true (**t**-by-**t**) that

$$D_N^{*,d} \quad \leq \quad \left(\frac{3}{2}\right)^d \left(\frac{2}{H+1} + \sum_{0 < \|\mathbf{m}\|_\infty \leq H} \frac{1}{r(\mathbf{m})} \left| \frac{1}{N} \sum_{k=1}^{N} Y_k(\mathbf{m}) \right| \right).$$

Substantial work by the "Tichy school" has been done on multi-dimensional discrepancy estimation of Koksma's numbers and variations. In particular, Tichy [Ti] obtains, in the context of [NT2] (more precisely, $x_k(t) = t^{a_k}$, $t > 1$, for some $a_k \in \mathbb{R}$, and the minimal distance between different powers is bounded below by $\delta > 0$), bounds on $D_N^{*,d}$ of the form $O(N^{-1/2+\eta})$ for any $\eta > 0$, even if the multiplicity d is not fixed but diverges as $\log N$ raised to a small power; and Goldstern [Go] extends these to the setting where the replication between the exponents a_n is possible but infrequent, and the minimal distance between the exponents may slowly converge to 0 (see also notes on the literature in [DT], Section 1.6).

Recall that already in the setting of linear generators, the random variables $Y_k(\mathbf{m})$ are not mutually independent, however they do have other nice properties. For concreteness, one could take the two examples of completely equidistributed sequences from Section 2.1.3. Is there a central limit theorem, or another type of concentration result that would apply, and give good estimates on d-dimensional (star) discrepancy of these sequences, with or without modifications related to R4 type of randomness? When is $\sqrt{N/\log\log N} D_N^{*,d}$ a tight sequence of random variables? The above LIL-type result for the simple (1-multiple) discrepancy sequences $(D_N^{*,1})_N$ is well-known (see e.g. [DT] Section 1.6.2 or [AB2]) in the context of lacunary sequences, and in particular for (2.10,2.11), even if permuted (without possible deletion), as shown by Aistleitner et al. [ABT]. But as soon as one starts increasing the multiplicity d, no specific study of the corresponding LIL seem to exist.

In view of (2.9) and Remark 1(d), most probabilists and statisticians would likely ask if any of the following is true for any (or all) of the c.e.s. discussed here: given any $d \geq 1$ and any $f : \mathbb{R}^d \to \mathbb{R}$ a nice enough function, define $f(\boldsymbol{\beta}_k^\infty) := f(\boldsymbol{\beta}_k^d)$ and $f(\Gamma) := f(U_1, U_2, \ldots, U_d)$. Are the sequences of random variables (recall (1.8))

$$\sqrt{N} \left(\frac{1}{N} \sum_{k=0}^{N-1} f(\boldsymbol{\beta}_{kh+1}^\infty) - \mathbf{E}\left(f(\Gamma)\right) \right), \ N \geq 1, \ \text{where } h \geq d, \qquad (3.2)$$

and/or

$$\sqrt{N} \left(\frac{1}{N} \sum_{k=1}^{N} f(\boldsymbol{\beta}_k^\infty(t)) - \mathbf{E}\left(f(\Gamma)\right) \right), N \geq 1, \qquad (3.3)$$

tight? Do they converge in law to a centered Gaussian random variable? These questions are similar in spirit to those of the preceding paragraph, but not quite the same (see [AB3] for a functional CLT for 1-dimensional discrepancies). In particular, the well-known *Koksma* and *Koksma-Hlawka* inequalities (see e.g. [KN], Ch. 2, Section 5 or [DT], Theorem 1.14) serve to give universal bounds on the error in (MC) numerical integration in terms of the discrepancy of the sequence and (a multi-dimensional extension of) the bounded variation of the integrand. Though essential in various applications, they are too crude for studying weak convergence properties (3.2,3.3). In the one-dimensional setting, again a substantial progress has been made for lacunary sequences, starting from the classical work of Salem and Zygmund [SZ], and ending in the recent studies by Aistleitner et al. [AB1, ABT] (see also [Ai, AB2]). Satisfying the CLT analogues (3.2,3.3) (with permutations and possible deletions permitted) would undoubtedly be a strong "evidence of randomness". How is it (or is it) related to the strongest Knuth's [Kn2] R6 type of pseudo-randomness?

Acknowledgments

References

[Ai] C. AISTLEITNER, Quantitative uniform distribution results for geometric progressions. *Israel J. Math.*, **204**, 1: 155–197, 2014. MR3273455

[AB1] 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. MR2550364

[AB2] C. AISTLEITNER AND I. BERKES, Probability and metric discrepancy theory. *Stoch. Dyn.*, **11**, 183–207, 2011. MR2771348

[AB3] C. AISTLEITNER AND I. BERKES, Limit distributions in metric discrepancy theory. *Monatsh. Math.*, **169**, 3–4: 253—265, 2013. MR3019283

[ABT] C. AISTLEITNER, I. BERKES, AND R.F. TICHY, Lacunary sequences and permutations. *Dependence in probability, analysis and number theory. A volume in memory of Walter Philipp*, Kendrick Press. 35–49, 2010. MR2731043

[Bu1] K. BURDZY, *The Search for Certainty*. World Scientific, 2009. MR2510150

[Bu2] K. BURDZY, *Resonance: From Probability to Epistemology and Back.* Imperial College Press, 2016. MR3468703

[Du] R. DURRETT, *Probability: theory and examples*, 3rd edition. Duxbury advanced series, 2005. MR1609153

[CDO] S. CHEN, J. DICK, AND A.B. OWEN, Consistency of Markov chain quasi-Monte Carlo on continuous state spaces. *Ann. Statist.*, **39**, 2:673–701, 2011. MR2816335

[CMNO] S. CHEN, M. MATSUMOTO, T. NISHIMURA, AND A.B. OWEN, New inputs and methods for Markov chain quasi-Monte Carlo. In *Monte Carlo and quasi-Monte Carlo methods 2010.* Springer Proc. Math. Stat. **23**, 313–327, Springer, 2012. MR3173841

[DELV] H. DAVENPORT, P. ERDŐS, AND W.J. LEVEQUE, On Weyl's criterion for uniform distribution. *Michigan Math. J.*, **10**, 3:311–314, 1963. MR0153656

[DT] M. DRMOTA AND R.F. TICHY, *Sequences, Discrepancies, and Applications.* Lecture Notes in Mathematics **1651**, Springer, Berlin, 1997. MR1470456

[DTW] M. DRMOTA, R.F TICHY, AND R. WINKLER, Completely uniformly distributed sequences of matrices. In: *Number-Theoretic Analysis.* Lecture Notes in Math. **1452**, pp. 43–57, Springer, 1990. MR1084637

[EK] P. ERDŐS AND M. KAC, The Gaussian law of errors in the theory of additive number theoretic functions. *Amer. J. Math.*, **62**, 1:738–742, 1940. MR0002374

[Fr] J.N. FRANKLIN, Deterministic simulation of random processes, *Math. Comp.*, **17**, 28–59, 1963. MR0149640

[Go] M. GOLDSTERN, Eine Klasse volständig gleichverteiler Folgen. In *Zahlentheoretishe Analysis, II* Lecture Notes in Math. **1262**, pp. 37–45, Springer, 1987. MR1012958

[Ho1] P.J. HOLEWIJN, Note on Weyl's criterion and the uniform distribution of independent random variables. *Ann. Math. Stat.* **40**, 1124–1125, 1969. MR0240855

[Ho2] P.J. HOLEWIJN, On the Uniform Distribution of Sequences of Random Variables. *Z. Wahrscheinlichkeitstheorie verw. Geb.*, **14**, 89–92, 1969. MR0264735

[Ka1] M. KAC, *Statistical Independence in Probability, Analysis and Number Theory.* Mathematical Association of America, 1959. MR0110114

[Ka2] M. KAC, Probability methods in some problems of analysis and number theory. *Bull. Amer. Math. Soc.*, **55**, 641–665, 1949. MR0031504

[Kem] J.H.B. KEMPERMAN, Probability methods in the theory of distributions modulo one. *Compositio Mathematica*, **16** 106–137, 1964. MR0173281

[Ke1] H. KESTEN, Uniform Distribution mod 1. *Ann. Math.*,**71**, 2:445–471, 1960. MR0113864

[Ke2] H. KESTEN, Uniform Distribution mod 1 (II). *Acta Arith.*, **7**, 4:355–380, 1962. MR0142532

[Kh] D. Khoshnevisan, Normal numbers are normal. *Clay Mathematics Institute Annual Report*, 15, continued pp. 27—31, 2006.

[Kn1] D.E. Knuth, Construction of a random sequence. *BIT Numerical Mathematics*, **5**, 4:246–250, 1965. MR0197434

[Kn2] D.E. Knuth, *The Art of Computer Programming*, volume 2/Seminumerical Algorithms, 3rd edition. Addison-Wesley, 1997. MR3077152

[Kol1] A.N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung, Ergebnisse der Mathematik und ihrer Grenzgebeite.* J. Springer, 1933. MR0362415

[Kol2] A.N. Kolmogorov, Tri podhoda k opredeleniju ponjatija "količevstvo informacii" (Three approaches to the definition of the concept "quantity of information".). *Problemy peredači informacii* **1**, 3–11, 1965.

[Kok] J.F. Koksma, Ein mengentheoretischer Satz über die Gleichverteilung modulo Eins. *Compositio Math.*, **2**, 250–258, 1935. MR1556918

[Kr1] N.M. Korobov, On functions with uniformly distributed fractional parts (Russian). *Doklady Akad. Nauk SSSR (N.S.)*, **62**, 21–22, 1948. MR0027012

[Kr2] N.M. Korobov, Some problems on the distribution of fractional parts (Russian). *Uspehi Matem. Nauk SSSR (N.S.)*, **4**, 1(29):189–190, 1949. MR0031948

[Kr3] N.M. Korobov, Concerning some questions of uniform distribution (Russian). *Izvestiya Akad. Nauk SSSR. Ser. Mat.* **14**, 215–238, 1950. MR0037876

[Kr4] N.M. Korobov, Bounds of trigonometric sums involving completely uniformly distributed functions (Russian). *Soviet Math. Dokl.*, **1**, 923–926, 1960. MR0130527

[Kr5] N.M. Korobov, *Exponential Sums and their Applications.* Mathematics and Its Applications, **80**, Kluwer, 1992. MR1162539

[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences.* Wiley and Sons, 1974. MR0419394

[Lac] M.B. Lacaze, Estimation de moyennes et fonctions de répartition de suites d'échantillonnage. *Ann. Inst. Henri Poincaré Probab. Stat.*, **9**, 2:145–165, 1973. MR0356199

[Lam] M. van Lambalgen, Randomness and Foundations of Probability: von Mises' axiomatisation of random sequences. *Statistics, probability and game theory: Papers in honor of David Blackwell*, 347–367, IMS, 1996. MR1481789

[Le1] L. Levin, On the notion of a random sequence. *Soviet Math. Dok.* **14** 1413–1416, 1973.

[Le2] M.B. Levin, Discrepancy Estimates of Completely Uniformly Distributed and Pseudorandom Number Sequences. *Internat. Math. Res. Notices* **22**, 1231–1251, 1999. MR1731474

[Li] N. Limić, On completely equidistributed numbers. *Math. Commun.* **7**, 103–111, 2002. MR1952752

[Lo] R.M. LOYNES, Some results in the probabilistic theory of asymptotic uniform distribution modulo 1. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **26**, 33–41, 1973. MR0329051

[Ly] R. LYONS, Strong laws of large numbers for weakly correlated random variables. *Michigan Math. J.*, **35**, 3:353–359, 1988. MR0978305

[ML] P. MARTIN-LÖF, The definition of random sequences. *Inform. Control* **9**, 602–619, 1966. MR0223179

[MF] M. MENDÈS FRANCE, Nombres normaux. Applications aux fonctions pseudo-aléatoires. *J. Anal. Math.*, **20**, 1:1–56, 1967. MR0220683

[Ms] R. VON MISES, *Probability, Statistics and Truth*, 2nd revised English edn., Dover, New York, 1957. MR0090158

[NT1] H. NIEDERREITER AND R.F. TICHY, Solution of a problem of Knuth on complete uniform distribution of sequences. *Mathematika*, **32**, 26–32, 1985. MR0817103

[NT2] H. NIEDERREITER AND R.F. TICHY, Metric theorems on uniform distribution and approximation theory. Journées arithmétiques de Besançon (Besançon, 1985). *Astérisque* **147–148**, 319–323, 1987. MR0891439

[OT] A.B. OWEN AND S.D. TRIBBLE, A quasi-Monte Carlo Metropolis algorithm. *Proc. Natl. Acad. Sci. USA*, **102**, 25:8844–8849, 2005. MR2168266

[SZ] R. SALEM AND A. ZYGMUND, On Lacunary Trigonometric Series. *Proc. Natl. Acad. Sci. USA*, **33**, 11:333–338, 1947. MR0022263

[SP] O. STRAUCH AND S. PORUBSKÝ, *Distribution of Sequences: A Sampler*. Peter Lang, 2005. MR2290224

[Ti] R.F. TICHY, Ein metrischer Satz über vollständig gleichverteile Folgen. *Acta Arith.* **48**, 2:197–207, 1987. MR0895440

[TO] S.D. TRIBBLE AND A.B. OWEN, Construction of weakly CUD sequences for MCMC sampling. *Electron. J. Stat.*, **2**, 634–660, 2008. MR2426105

[We] H. WEYL, *Über die Gleichverteilung von Zahlen mod. Eins.* Math. Ann., **77**, 313–352, 1916. MR1511862