

Pseudorandom processes

Tamás F. Móri^{a,*}, Gábor J. Székely^{b,c}

^a*Department of Probability Theory and Statistics, Eötvös Loránd University,
Pázmány Péter s. 1/C, H-1117 Budapest, Hungary*

^b*National Science Foundation, 2415 Eisenhower Avenue, Alexandria, VA 22314*

^c*Rényi Institute of Mathematics, Hungarian Academy of Sciences,
Reáltanoda u. 13–15, H-1053 Budapest, Hungary*

Abstract

Pseudorandom numbers turned out to be very widely applicable substitutes of real random number sequences. In this paper we introduce pseudorandom processes that appear to be random but they just imitate certain properties of random/stochastic processes. This paper is only the first step in replacing random coins of random walks by trigonometric functions. More general functions (“coins”) lead to unexpected new challenges, e.g. in case of Haar coins.

Keywords: stochastic process, limit theorem, Fourier series, weak convergence, pseudorandom numbers

1. Introduction

Our starting point is the random walk approximation of Brownian motion $B(t) = \lim s_{[nt]}/\sqrt{n}$ as $n \rightarrow \infty$ where s_n is a random walk, that is, s_n is the sum of n iid random variables that take the values $+1$ and -1 with the same probability $1/2$. We can call them random coins or real coins. In this paper we replace them by the most classical orthogonal functions like sine and cosine. It turns out that their partial sums, S_n and T_n , have limit distributions without dividing them by the usual \sqrt{n} factor. This led us to a conjecture on more general 1-periodic functions. In the final sections we discuss orthogonal polynomials and also the Haar series as substitutes of

*Corresponding author

Email addresses: mori@math.elte.hu (Tamás F. Móri), gszekely@nsf.gov (Gábor J. Székely)

random coins. Our “pseudo-coins” are always identically distributed but not independent, not even exchangeable because then they would be conditionally independent, that is conditionally “real coins”. The price we need to pay for replacing “real coins” with “pseudo-coins” is that limit distributions of pseudo-random series might depend on the order of terms. In many cases our “pseudo-coins” are uncorrelated/orthogonal, for example the trigonometric coins below.

2. Trigonometric coins and processes

Let the probability space be the interval $[0, 1)$ with the Lebesgue measure, $\omega \in [0, 1)$ and let U, V be independent random variables, uniformly distributed on $[0, 1)$. Then we have the results below.

Theorem 1 (sine coins). *The sine coins, $\sin(2k\pi\omega)$, $k = 1, 2, \dots$, are identically distributed, $S_n := \sum_{k=1}^n \sin(2k\pi\omega)$ as $n \rightarrow \infty$ has a limit distribution without any normalization (!), and this limit distribution is the same as the distribution of*

$$X := \frac{\cos(\pi U) - \cos(2\pi V)}{2 \sin(\pi U)}.$$

Corollary 1. *The limit distribution of S_n is a Cauchy distribution with scale parameter $1/2$ and its pdf*

$$f_X(x) = \frac{2}{\pi(1 + 4x^2)}.$$

A pseudorandom number sequence is typically determined by the first or the first two numbers of this sequence. Here we have the same: the first two sine coin values determine all others.

Theorem 2 (cosine coins). *The cosine coins, $\cos(2k\pi\omega)$, $k = 1, 2, \dots$, are identically distributed and $T_n := \sum_{k=1}^n \cos(2k\pi\omega)$ as $n \rightarrow \infty$ has a limit distribution which is the same as the distribution of*

$$Y := \frac{\sin(2\pi V) - \sin(\pi U)}{2 \sin(\pi U)}.$$

Corollary 2. *The pdf of the limit distribution of T_n is*

$$f_Y(x) = \frac{2}{\pi^2 |2x + 1|} \left| \log \left| 1 + \frac{1}{x} \right| \right|, \quad x \notin \left\{ -1, -\frac{1}{2}, 0 \right\}.$$

This function, unlike the Cauchy pdf, is neither bounded, nor unimodal. It is symmetric around $-\frac{1}{2}$.

Remark 1. Trigonometric coins are not exchangeable and thus it can easily happen that a permutation of the coins $\sin(2k\pi\omega)$ is such that the sum of the first n of them, S_n^* has a weak limit different from Cauchy. For example, it is known that a sufficiently rare subsequence of $\sin(2k\pi\omega)$ behaves as if they were independent (see Gaposhkin, 1966; Berkes, 2017), and if we insert the remaining terms into the essentially independent sequence such that their number until the n -th term is $o(\sqrt{n})$ then S_n^*/\sqrt{n} tends to a Gaussian distribution because the CLT applies. We do not know the set of all weak accumulation points of $S_n^*/a(n)$ where $a(n)$ is a suitable normalizing sequence.

For the proofs of our theorems we will need the following lemma where $\{.\}$ denotes the fractional part.

Lemma 1. *Let t_1, \dots, t_r be positive real numbers, linearly independent over the rational numbers \mathbb{Q} . Let $\varphi_1, \dots, \varphi_r$ be real valued functions such that $\varphi_j(\nu) \sim \nu t_j$ as $\nu \rightarrow \infty$, that is, the ratio of the two sides tends to 1. Finally, let U, V be random variables; U being uniformly distributed on $[0, 1]$, and V having an absolutely continuous distribution. Then the limit distribution of the vector $(U, \{\varphi_1(\nu)V\}, \dots, \{\varphi_r(\nu)V\})$ as $\nu \rightarrow \infty$ is uniform on the $r+1$ dimensional unit cube.*

Proof. It is well known, cf. Ch.1, §7 of Billingsley (2013), that for the limit distribution we only need to show

$$\lim_{\nu \rightarrow \infty} \mathbb{E} \left(e^{2\pi i(m_0 U + m_1 \varphi_1(\nu)V + \dots + m_r \varphi_r(\nu)V)} \right) = 0 \quad (1)$$

for arbitrary integers m_0, m_1, \dots, m_r provided not all of them are zero.

Now, (1) is obviously true if $m_1 = \dots = m_r = 0$ (and $m_0 \neq 0$). If m_1, \dots, m_r are not all zero, then decompose $e^{im_0 2\pi U}$ into the sum of real and imaginary parts, then take the positive and negative parts of both terms. In this way the expectation above is expressed as a sum of four expectations; the first of them is

$$\mathbb{E} \left(e^{2\pi i(m_1 \varphi_1(\nu)V + \dots + m_r \varphi_r(\nu)V)} (\cos(2m_0 \pi U))^+ \right).$$

This can be expressed as a constant multiple of the integral

$$\int e^{2\pi i(m_1\varphi_1(\nu) + \cdots + m_r\varphi_r(\nu))} V d\mu, \quad (2)$$

where μ is the probability measure with Radon–Nikodym derivative $d\mu/d\mathbb{P}$ proportional to $(\cos(2\pi m_0 U))^+$. In the exponent

$$m_1\varphi_1(\nu) + \cdots + m_r\varphi_r(\nu) = \nu(m_1t_1 + \cdots + m_rt_r) + o(\nu)$$

diverges to $-\infty$ or $+\infty$ because the coefficient of ν is not 0 by the condition imposed on the numbers t_i . The distribution of V with respect to μ is still absolutely continuous. Since the characteristic function of an absolutely continuous probability measure disappears at infinity by the Riemann–Lebesgue lemma, we obtain that (2) tends to 0 as $\nu \rightarrow \infty$. The remaining three integrals can be handled similarly. \square

Proofs of Theorems 1, 2, and Corollaries 1, 2. It is well known that

$$S_n(\omega) = \frac{\sin((n+1)\pi\omega) \sin(n\pi\omega)}{\sin(\pi\omega)} = \frac{\cos(\pi\omega) - \cos((n+\frac{1}{2})2\pi\omega)}{2\sin(\pi\omega)}.$$

Here $U(\omega) = \omega$ is uniformly distributed on $[0, 1)$, and the limit distribution of $(U, \{(n+\frac{1}{2})U\})$ as $n \rightarrow \infty$ is uniform on the unit square $[0, 1) \times [0, 1)$. This is implied by Lemma 1 with $r = 1$, $\varphi_1(\nu) = \nu + \frac{1}{2}$, and $V = U$. Hence Theorem 1 follows. For Corollary 1 observe that $-\cos(2\pi V)$ and $\cos(\pi V)$ have the same distribution, thus the limit distribution of S_n can also be written in the form

$$\frac{\cos(\pi U) + \cos(\pi V)}{2\sin(\pi U)}.$$

Moreover, any 1-periodic function $f(\omega)$ restricted to the Lebesgue probability space over the unit interval $[0, 1)$ has the same distribution as $f(2\omega)$, thus we can equally deal with the limit distribution of $S_n(2\omega)$, which is

$$\frac{\cos(2\pi U) + \cos(2\pi V)}{2\sin(2\pi U)}$$

where U and V are independent and uniformly distributed on $[0, 1)$. Let

$\xi := U + V$ and $\eta := U - V$, then we have

$$\begin{aligned}
\frac{\cos(2\pi U) + \cos(2\pi V)}{2 \sin(2\pi U)} &= \frac{\cos(\pi(\xi + \eta)) + \cos(\pi(\xi - \eta))}{2 \sin(\pi(\xi + \eta))} \\
&= \frac{\cos(\pi\xi) \cos(\pi\eta)}{\sin(\pi\xi) \cos(\pi\eta) + \cos(\pi\xi) \sin(\pi\eta)} \\
&= \frac{\frac{1}{2}}{\frac{1}{2} [\tan(\pi\xi) + \tan(\pi\eta)]} \\
&= \frac{\frac{1}{2}}{\frac{1}{2} [\tan(\pi\{\xi\}) + \tan(\pi\{\eta\})]}
\end{aligned}$$

Here $\{\xi\}$ and $\{\eta\}$ are also independent and uniformly distributed on $[0, 1)$. Therefore, the denominator, being the arithmetic mean of two independent standard Cauchy distributed random variables, is also standard Cauchy and the reciprocal of a Cauchy random variable is Cauchy again. This proves Corollary 1.

The sum of cosines can be treated similarly. From the well-known formula

$$T_n(\omega) = \frac{\sin\left((n + \frac{1}{2})2\pi\omega\right)}{2 \sin(\pi\omega)} - \frac{1}{2}$$

it immediately follows that the limit distribution of T_n is equal to the distribution of

$$Y = \frac{\sin(2\pi V) - \sin(\pi U)}{2 \sin(\pi U)},$$

where U and V are independent and uniformly distributed on $[0, 1)$.

For computing the probability density function of Y let us first consider the pdf of $|2Y + 1|$. It is the same as that of $\sin(\pi V)/\sin(\pi U)$, because $|\sin(2\pi V)|$ and $\sin(\pi V)$ are identically distributed. The numerator and the denominator are independent and their common pdf is

$$g(x) = \frac{2}{\pi} \cdot \frac{1}{\sqrt{1-x^2}}, \quad 0 < x < 1,$$

therefore the pdf of $|2Y + 1|$ at $x \in (0, 1]$ is

$$h(x) = \int_{-\infty}^{+\infty} |t|g(t)g(tx) dt = \frac{4}{\pi^2} \int_0^1 \frac{t}{\sqrt{(1-t^2)(1-x^2t^2)}} dt.$$

Let us substitute $s = t^2$. After some easy calculations we get

$$\begin{aligned} h(x) &= \frac{2}{\pi^2} \int_0^1 \frac{1}{\sqrt{(1-s)(1-x^2s)}} ds \\ &= \left[-\frac{4}{\pi^2 x} \log \left(x\sqrt{1-s} + \sqrt{1-x^2s} \right) \right]_{s=0}^{s=1} \\ &= \frac{2}{\pi^2 x} \log \frac{1+x}{1-x}. \end{aligned}$$

Since $2Y + 1 = \sin(2\pi V)/\sin(\pi U)$ and its reciprocal are identically distributed, $h(1/x) = x^2 h(x)$. Therefore

$$h(x) = \frac{2}{\pi^2 x} \log \frac{x+1}{x-1} = \frac{2}{\pi^2 x} \log \left| \frac{1+x}{1-x} \right| \quad (3)$$

for $x > 1$. Moreover, the distribution of $2Y + 1$ is symmetric around the origin, its pdf is equal to $\frac{1}{2} h(|x|)$, $x \in \mathbb{R}$. Finally, the pdf of Y is obviously $h(|2x + 1|)$, from which straightforward calculation leads to the formula of Corollary 2. \square

Remark 2. If the random variables $\sin(2k\pi\omega)$ were iid, then by Lévy's equivalence theorem (Chung, 2000, Theorem 9.5.5) the weak convergence of S_n would imply its almost sure convergence. But our S_n , $n = 1, 2, \dots$ sequence is divergent almost surely. The same holds for T_n .

Remark 3. More than ten years ago one of our PhD students, Vidyadhar S. Phadke, gave less transparent proofs of these results in his dissertation (Phadke, 2008) based on a theorem of Rényi (1958). In 2014 we proposed Corollary 1 for a college level contest problem in Hungary (János Bolyai Math. Soc., 2015).

We are going to characterize the weak limits of the finite dimensional distributions of the processes $X_n(t) = S_{[nt]}$ and $Y_n(t) = T_{[nt]}$, $t \geq 0$, as $n \rightarrow \infty$. Since $S_n(0) = T_n(0) \equiv 0$, we can focus on positive values of t .

First, introduce the process $\gamma_n(t, \omega) = \left\{ \left([nt] + \frac{1}{2} \right) 2\pi\omega \right\}$, $t > 0$, $\omega \in [0, 1)$.

Theorem 3. *Let \mathcal{H} be a Hamel basis (maximal subset of the real numbers that is linearly independent over \mathbb{Q} , the field of rational numbers). For this*

the axiom of choice is needed. Let t_1, \dots, t_k be different positive numbers, and consider the unique representation

$$t_j = \sum_{h \in \mathcal{H}} \alpha_{jh} h, \quad 1 \leq j \leq k, \quad \alpha_{jh} \in \mathbb{Q}.$$

Note that the sum can have finitely many nonzero terms only. Put

$$\mathcal{H}' = \{h \in \mathcal{H} : \alpha_{jh} \neq 0 \text{ for some } j = 1, \dots, k\}.$$

Suppose $\alpha_{jh} = p_{jh}/N$, where p_{jh} and $N > 0$ are integers (thus N is a common denominator for $1 \leq j \leq k$ and $h \in \mathcal{H}'$). Then $(\gamma_n(t_1), \dots, \gamma_n(t_k))$ converges in distribution to the vector $(\gamma(t_1), \dots, \gamma(t_k))$, where

$$\gamma(t_j) = \left\{ \sum_{h \in \mathcal{H}'} p_{jh} V_h \right\}, \quad 1 \leq j \leq k,$$

and V_h , $h \in \mathcal{H}'$ are independent and uniformly distributed on the interval $[0, 1)$. For every fixed t the distribution of $\gamma(t)$ is uniform on $[0, 1)$, and $\gamma^*(t) := \gamma(e^{-t})$, $t \in \mathbb{R}$, is (strongly) stationary.

Proof. Since

$$(\lfloor nt_j \rfloor + \tfrac{1}{2}) \omega = \left(\left\lfloor \frac{n}{N} \sum_{h \in \mathcal{H}'} p_{jh} h \right\rfloor + \frac{1}{2} \right) \omega = \left(\sum_{h \in \mathcal{H}'} p_{jh} \varphi_h \left(\frac{n}{N} \right) \right) \omega,$$

where $\varphi_h(\nu) = \nu h + O(1)$, Lemma 1 can be applied.

By Kolmogorov's extension theorem (Billingsley, 2012, Section 36) there exists a probability distribution on the product space $[0, 1)^{\mathbb{R}^+}$ with finite dimensional marginals given by the above theorem. This process $\gamma(t)$ can be considered the “limiting process” of $\gamma_n(t)$, $t > 0$, as $n \rightarrow \infty$, where the word “limit” is meant in the sense of weak convergence of finite dimensional distributions. For more details on this kind of definition see Doob (1953) and Gettoor (2009).

The transformed process $\gamma^*(t) = \gamma(e^{-t})$ is stationary, that is, the joint distribution of $(\gamma^*(t_1), \dots, \gamma^*(t_k))$ coincides with the distribution of the vector $(\gamma^*(t_1 + h), \dots, \gamma^*(t_k + h))$, $h > 0$. This follows from the fact that using the notation $s_j = e^{-t_j}$, $c = e^{-h}$, the distributions of $(\gamma(s_1), \dots, \gamma(s_k))$ and $(\gamma(cs_1), \dots, \gamma(cs_k))$ are also identical because the coordinates of s_1, \dots, s_k in the Hamel basis \mathcal{H} coincide with those of cs_1, \dots, cs_k in the Hamel basis $c\mathcal{H}$. \square

The following lemma will imply important properties of our pseudorandom processes.

Lemma 2.

(i) *The process $\gamma(t)$ is not continuous stochastically at any time t (thus $\gamma(t)$ is “more pathological” than typical continuous or càdlàg realization processes that are known to be stochastically continuous).*

(ii) *$\gamma(t)$ is separable.*

Separability ensures that there exists a dense countable subset $\mathcal{S} \subset \mathbb{R}^+$ such that sample path functionals are essentially determined by the random variables $\gamma(t)$, $t \in \mathcal{S}$. By a theorem of Doob every real-valued continuous-time stochastic process has a separable modification. Part (ii) of Lemma 2 says that our construction itself is separable.

Proof. (i) $\gamma(t)$ cannot be continuous stochastically, because any positive real number t can be approximated with arbitrary precision by numbers s such that t/s is irrational. In this case $\gamma(t)$ and $\gamma(s)$ are independent and identically distributed random variables thus $\gamma(s)$ cannot converge to $\gamma(t)$ in probability.

It seems the sample paths of $\gamma(t) = \gamma(t, \omega)$ are not measurable for a.e. ω because the Hamel basis is not measurable but for this extreme “lamentable scourge” in the sense of Hermite we do not have a rigorous proof.

(ii) It is enough to find an event N with zero probability, and an everywhere dense countable subset $\mathcal{S} \subset \mathbb{R}^+$ (the separant) such that for every open interval $I \subset \mathbb{R}^+$ on the complement of N we have

$$\begin{aligned} \inf\{\gamma(t) : t \in I\} &= \inf\{\gamma(t) : t \in I \cap \mathcal{S}\}, \\ \sup\{\gamma(t) : t \in I\} &= \sup\{\gamma(t) : t \in I \cap \mathcal{S}\}. \end{aligned}$$

Now, let \mathcal{S} be an everywhere dense countable set the elements of which are linearly independent over \mathbb{Q} , the field of rational numbers. Then $\gamma(t)$, $t \in \mathcal{S}$ are (stochastically) independent. Indeed, let us fix a Hamel basis \mathcal{H} and let $t_1, \dots, t_k \in \mathcal{S}$ be different positive numbers.

By definition, $\gamma(t_j) = \{\sum_{h \in \mathcal{H}} p_{jh} V_h\}$, where V_h , $h \in \mathcal{H}$ are iid and uniform on $[0, 1)$; furthermore, p_{jh} are integers, and $p_{jh} \neq 0$ if and only the coefficient of h in the representation

$$t_j = \sum_{h \in \mathcal{H}} \alpha_{jh} h, \quad 1 \leq j \leq k, \quad \alpha_{jh} \in \mathbb{Q},$$

differs from 0. For every $j = 1, \dots, k$ there exists an $h_j \in \mathcal{H}$ such that $\alpha_{j,h_i} = 0$ for $i \neq j$ but $\alpha_{j,h_j} \neq 0$. Therefore the joint conditional distribution of $\gamma(t_j)$, $1 \leq j \leq k$, given V_h , $h \in \mathcal{H} \setminus \{h_1, \dots, h_k\}$, is uniform on the k dimensional hypercube, hence the same holds for the unconditional distribution.

Consequently, for every diadic interval $I \subset \mathbb{R}^+$ we have

$$\inf\{\gamma(t) : t \in I\} = 0, \quad \sup\{\gamma(t) : t \in I\} = 1 \quad (4)$$

with probability 1. The number of diadic intervals is countable, so there exists an event N with zero probability such that (4) holds simultaneously for all diadic intervals on the complement of N . Every open interval contains diadic intervals, thus on the complement of N we have

$$\begin{aligned} \inf\{\gamma(t) : t \in I \cap \mathcal{S}\} &= 0 = \inf\{\gamma(t) : t \in I\}, \\ \sup\{\gamma(t) : t \in I \cap \mathcal{S}\} &= 1 = \sup\{\gamma(t) : t \in I\} \end{aligned}$$

for every open interval I , as needed. \square

As we have seen before,

$$S_{[nt_j]}(\omega) = \frac{\cos(\pi\omega) - \cos\left(\left([nt_j] + \frac{1}{2}\right)2\pi\omega\right)}{2\sin(\pi\omega)},$$

and

$$T_{[nt_j]}(\omega) = \frac{\sin\left(\left([nt_j] + \frac{1}{2}\right)2\pi\omega\right) - \sin(\pi\omega)}{2\sin(\pi\omega)},$$

hence we immediately obtain the following limit result for the processes $X_n(t)$ and $Y_n(t)$.

Corollary 3. *The “limiting processes” of $X_n(t) = S_{[nt]}$ and $Y_n(t) = T_{[nt]}$, is*

$$X(t) = \frac{\cos(\pi U) - \cos(2\pi\gamma(t))}{2\sin(\pi U)}, \quad Y(t) = \frac{\sin(2\pi\gamma(t)) - \sin(\pi U)}{2\sin(\pi U)},$$

where U is independent of the process $\gamma(t)$ and uniformly distributed on the interval $[0, 1)$. By the previous lemma $X(t), Y(t)$ are separable but not continuous stochastically. \square

3. Fourier coins and processes

Theorem 4 (Fourier coins). *Let f be a 1-periodic, twice continuously differentiable function, and let U be uniformly distributed on the interval $[0, 1]$. Suppose $\mathbb{E}f(U) = 0$. Then $F_n = \sum_{k=1}^n f(kU)$ converges in distribution.*

Proof. Expand f into Fourier series:

$$f(\omega) = \sum_{j=1}^{\infty} [a_j \sin(2j\pi\omega) + b_j \cos(2j\pi\omega)], \quad \omega \in [0, 1];$$

the constant term is missing because $f(U)$ has zero mean. It is well known that this Fourier series converges to f uniformly. Moreover, since f belongs to the Sobolev space $H^2 = W^{2,2}$, it follows that

$$\sum_{j=1}^{\infty} j^4 (a_j^2 + b_j^2) < \infty,$$

see Adams and Fournier (2003). Then by Hölder's inequality we have

$$\begin{aligned} \sum_{j=1}^{\infty} (|a_j| + |b_j|)^{1/2} &= \sum_{j=1}^{\infty} j (|a_j| + |b_j|)^{1/2} \cdot j^{-1} \\ &\leq \left[\sum_{j=1}^{\infty} j^4 (|a_j| + |b_j|)^2 \right]^{1/4} \left[\sum_{j=1}^{\infty} j^{-4/3} \right]^{3/4} < \infty. \end{aligned} \quad (5)$$

Consider the partial sums

$$\begin{aligned} f_N(\omega) &= \sum_{j=1}^N (a_j \sin(2j\pi\omega) + b_j \cos(2j\pi\omega)), \\ F_{n,N} &= \sum_{k=1}^n f_N(kU) = \sum_{k=1}^n \sum_{j=1}^N (a_j \sin(2jk\pi U) + b_j \cos(2jk\pi U)), \\ Z_N &= \sum_{j=1}^N (a_j X_j + b_j Y_j), \end{aligned}$$

where

$$X_j = \frac{\cos(j\pi\xi) - \cos(2j\pi\eta)}{2 \sin(j\pi\xi)}, \quad Y_j = \frac{\sin(2j\pi\eta) - \sin(j\pi\xi)}{2 \sin(j\pi\xi)}, \quad (6)$$

and ξ, η are independent and uniformly distributed on $[0, 1)$. The X 's and the Y 's are identically distributed but dependent.

First we show that the infinite series $Z_\infty = \sum_{j=1}^\infty [a_j X_j + b_j Y_j]$ converges almost surely. For arbitrary random variables φ, ψ define

$$\varrho(\varphi, \psi) = \min\{\varepsilon \geq 0 : \mathbb{P}(|\varphi - \psi| > \varepsilon) \leq \varepsilon\}.$$

This is a complete metric that induces convergence in probability. Now,

$$\begin{aligned} \varrho(Z_N, Z_{N-1}) &= \min\{\varepsilon \geq 0 : \mathbb{P}(|a_N X_N + b_N Y_N| > \varepsilon) \leq \varepsilon\} \\ &\leq \min\{\varepsilon \geq 0 : \mathbb{P}((|a_N| + |b_N|)/|\sin(N\pi\xi)| > \varepsilon) \leq \varepsilon\}. \end{aligned}$$

Since $|\sin(N\pi\xi)|$ and $\sin(\pi\xi)$ are identically distributed, and $\mathbb{P}(c > \sin(\pi\xi)) \leq c$, we have

$$\varrho(Z_N, Z_{N-1}) \leq (|a_N| + |b_N|)^{1/2}.$$

Similar estimation holds for $\varrho(F_{n,N}, F_{n,N-1})$. Now apply the Borel–Cantelli lemma to the events $A_N = \{|Z_N - Z_{N-1}| > (|a_N| + |b_N|)^{1/2}\}$. By (5), the infinite sum of $\mathbb{P}(A_N)$ is convergent, hence $|Z_N - Z_{N-1}| \leq (|a_N| + |b_N|)^{1/2}$ if N large enough, and (5) implies almost sure convergence.

Let $\mathcal{L}(\varphi, \psi)$ denote the Lévy distance between the distributions of the random variables φ, ψ , i.e.,

$$\mathcal{L}(\varphi, \psi) = \min\{\varepsilon \geq 0 : \mathbb{P}(\varphi \leq t - \varepsilon) - \varepsilon \leq \mathbb{P}(\psi \leq t) \leq \mathbb{P}(\varphi \leq t + \varepsilon) + \varepsilon, \forall t \in \mathbb{R}\}.$$

The Lévy distance is a complete metric that induces the convergence in distribution. Then $\mathcal{L}(\varphi, \psi) \leq \varrho(\varphi, \psi)$, thus

$$\begin{aligned} \mathcal{L}(F_n, F_{n,N}) &\leq \sum_{j=N+1}^\infty (|a_j| + |b_j|)^{1/2}, \\ \mathcal{L}(Z_N, Z_\infty) &\leq \sum_{j=N+1}^\infty (|a_j| + |b_j|)^{1/2}. \end{aligned}$$

As we have seen in Theorems 1 and 2,

$$\begin{aligned}
F_{n,N} &= \sum_{k=1}^n \sum_{j=1}^N [a_j \sin(2jk\pi U) + b_j \cos(2jk\pi U)] \\
&= \sum_{j=1}^N \sum_{k=1}^n [a_j \sin(2jk\pi U) + b_j \cos(2jk\pi U)] \\
&= \sum_{j=1}^N \left[a_j \frac{\cos(j\pi U) - \cos((n + \frac{1}{2})2j\pi U)}{2 \sin(j\pi U)} \right. \\
&\quad \left. + b_j \frac{\sin((n + \frac{1}{2})2j\pi U) - \sin(j\pi U)}{2 \sin(j\pi U)} \right] \\
&\rightarrow Z_N
\end{aligned}$$

in distribution, because if $\{(n + \frac{1}{2})U\} \rightarrow V$, then also

$$\{(n + \frac{1}{2})jU\} = \{j\{(n + \frac{1}{2})U\}\} \rightarrow \{jV\}$$

in distribution (and the convergence holds for joint distributions, too). Therefore

$$\begin{aligned}
\limsup_{n \rightarrow \infty} \mathcal{L}(F_n, Z_\infty) &\leq \limsup_{n \rightarrow \infty} [\mathcal{L}(F_n, F_{n,N}) + \mathcal{L}(F_{n,N}, Z_N) + \mathcal{L}(Z_N, Z_\infty)] \\
&\leq 2 \sum_{j=N+1}^{\infty} (|a_j| + |b_j|)^{1/2},
\end{aligned}$$

which can be arbitrarily small if N is large enough. \square

Remark 4. One might conjecture that for all 1-periodic L^2 functions $f(\omega) = \sum_{j=1}^{\infty} [a_j \sin(2j\pi\omega) + b_j \cos(2j\pi\omega)]$ we have that $\sum_{k=1}^n f(k\omega)$ has a limit distribution as $n \rightarrow \infty$, and this limit distribution is the same as the distribution of $Z_\infty = \sum_{j=1}^{\infty} (a_j X_j + b_j Y_j)$, where X_j, Y_j are defined in (6).

Unfortunately this is not true, shown by the following Theorem. Heuristically, if the non-zero coefficients of a purely sine trigonometric series are sufficiently rare then the terms are essentially independent (see Gaposhkin, 1966; Berkes, 2017) and thus Z_∞ becomes a weighted sum of (nearly) independent Cauchy variables. This is convergent if and only if the sum of the absolute values of the coefficients is convergent. More precisely, we will prove the following result.

Theorem 5. *Let (n_j) be a strictly increasing sequence of positive integers such that*

$$\sum_{j=1}^{\infty} \left(\frac{n_j}{n_{j+1}} \right)^{1/3} < \infty$$

(e.g., $n_j = (j!)^4$ will do). Then the infinite series $\sum_{j=1}^{\infty} (a_j X_{n_j} + b_j Y_{n_j})$ converges in distribution if and only if $\sum_{j=1}^{\infty} (|a_j| + |b_j|) < \infty$.

Thus the function

$$f(\omega) = \sum_{j=1}^{\infty} (a_j \sin(2n_j \pi \omega) + b_j \cos(2n_j \pi \omega))$$

with coefficients in ℓ^2 but not in ℓ^1 disproves the conjecture.

Proof. Similarly to the proof of Corollary 1, it is sufficient to deal with the function $f(2\omega)$, or equivalently, with the infinite series $\sum_{j=1}^{\infty} (a_j X_{2n_j} + b_j Y_{2n_j})$.

By (Berkes, 2017, Theorem 4.3) there exist iid random variables U_j and V_j , $j = 1, 2, \dots$, uniformly distributed on $[0, 1)$, such that

$$\mathbb{P}(|\{n_j \xi\} - U_j| \geq 2c_j) \leq 2c_j, \quad \mathbb{P}(|\{n_j \eta\} - V_j| \geq 2c_j) \leq 2c_j,$$

$j = 1, 2, \dots$, where $c_j = n_j/n_{j+1}$. For $0 \leq u < 1$, $0 \leq v < 1$ put

$$h_1(u, v) = \frac{\cos(2\pi u) - \cos(4\pi v)}{2 \sin(2\pi u)}, \quad h_2(u, v) = \frac{\sin(4\pi v) - \sin(2\pi u)}{2 \sin(2\pi u)}.$$

Then $X_{2n_j} = h_1(\{n_j \xi\}, \{n_j \eta\})$ and $Y_{2n_j} = h_2(\{n_j \xi\}, \{n_j \eta\})$. We will show that

$$\sum_{j=1}^{\infty} |a_j| |X_{2n_j} - h_1(U_j, V_j)| < \infty, \quad \sum_{j=1}^{\infty} |b_j| |Y_{2n_j} - h_2(U_j, V_j)| < \infty \quad (7)$$

almost surely. From this it follows that the series

$$\sum_{j=1}^{\infty} (a_j X_{2n_j} + b_j Y_{2n_j}) \quad \text{and} \quad \sum_{j=1}^{\infty} (a_j h_1(U_j, V_j) + b_j h_2(U_j, V_j))$$

are equiconvergent in distribution. As the summands of the latter one are independent, the sum converges in distribution if and only if it does almost surely. Let us split the second series into two. The first part is

$\sum_{j=1}^{\infty} a_j h_1(U_j, V_j)$; this is a sum of independent Cauchy distributed random variables with scale parameters $2a_j$. Therefore the sum from 1 to N is also Cauchy with scale parameter $2(|a_1| + \dots + |a_N|)$, thus it converges in distribution if and only if $\sum_{j=1}^{\infty} |a_j| < \infty$.

The second part can be written in the form $\sum_{j=1}^{\infty} b_j (\zeta_j - \frac{1}{2})$, where ζ_1, ζ_2, \dots are symmetrically distributed iid random variables, and the pdf of $|\zeta_j|$ is given in (3). Hence the distribution of $\sum_{j=1}^N b_j (\zeta_j - \frac{1}{2})$ is symmetric around $(b_1 + \dots + b_N)/2$, thus for the convergence in distribution it is necessary that the numerical series $\sum_{j=1}^{\infty} b_j$ be convergent. To the symmetric part $\sum_{j=1}^{\infty} b_j \zeta_j$ one can apply the three series theorem. The necessary and sufficient condition for the convergence of this series is the convergence of the numerical series

$$\sum_{j=1}^{\infty} \mathbb{P}(|b_j \zeta_j| > 1), \quad \sum_{j=1}^{\infty} \mathbb{E}(b_j \zeta_j I(|b_j \zeta_j| \leq 1)), \quad \sum_{j=1}^{\infty} \text{Var}(b_j \zeta_j I(|b_j \zeta_j| \leq 1)),$$

where $I(\cdot)$ stands for the indicator of the event in brackets. By (3),

$$G_j(x) := \mathbb{P}(|b_j \zeta_j| > x) = \mathbb{P}(|\zeta_j| > x/|b_j|) \sim \frac{4|b_j|}{\pi^2 x}$$

as $j \rightarrow \infty$ and x remains fixed. Moreover, $\mathbb{E}(b_j \zeta_j I(|b_j \zeta_j| \leq 1)) = 0$ by the symmetry of ζ_j . Finally,

$$\text{Var}(b_j \zeta_j I(|b_j \zeta_j| \leq 1)) = \mathbb{E}(b_j^2 \zeta_j^2 I(|b_j \zeta_j| \leq 1)) = \int_0^1 2x G_j(x) dx = O(|b_j|).$$

Therefore $\sum_{j=1}^{\infty} |b_j| < \infty$ is sufficient and necessary for the convergence of $\sum_{j=1}^{\infty} b_j h_2(U_j, V_j)$.

For the proof of (7) apply the mean value theorem.

$$|X_{2n_j} - h_1(U_j, V_j)| \leq \sup \|h'_1\| \cdot \|(\{n_j \xi\} - U_j, \{n_j \eta\} - V_j)\|,$$

where the supremum of the derivative h'_1 is taken over the segment connecting $(\{n_j \xi\}, \{n_j \eta\})$ and (U_j, V_j) . Here

$$h'_1(u, v) = \left(-\pi \frac{1 - \cos(2\pi u) \cos(4\pi v)}{\sin^2(2\pi u)}, 2\pi \frac{\sin(4\pi v)}{\sin(2\pi u)} \right),$$

thus

$$\|h'_1(u, v)\| \leq \frac{2\pi}{\sin^2(2\pi u)} \leq 2\pi c_j^{-2/3},$$

provided $|\sin(2\pi u)| \geq c_j^{1/3}$. Now, let A_j denote the event that at least one of the following conditions is violated:

$$\begin{aligned} |\sin(2\pi\{n_j\xi\})| &\geq c_j^{1/3}, & |\sin(2\pi U)| &\geq c_j^{1/3}, \\ |\{n_j\xi\} - U_j| &< 2c_j, & |\{n_j\eta\} - V_j| &< 2c_j. \end{aligned}$$

Then $\mathbb{P}(A_j) \leq 2c_j^{1/3} + 4c_j$, hence $\sum_{j=1}^{\infty} \mathbb{P}(A_j) < \infty$. By the Borel–Cantelli lemma it follows almost surely that only finitely many of the events A_j can occur. On the other hand, on the complement of A_j we have

$$|X_{2n_j} - h_1(U_j, V_j)| \leq 2\pi c_j^{-2/3} 8^{1/2} c_j < 6\pi c_j^{1/3},$$

thus $\sum_{j=1}^{\infty} |a_j| |X_{2n_j} - h_1(U_j, V_j)| < \infty$ holds almost surely by assumption.

In the same way one can show that $|Y_{2n_j} - h_2(U_j, V_j)| = O(c_j^{1/3})$, thus completing the proof. \square

Motivated by this counterexample, we conjecture the following.

Conjecture 1. *If $f(\omega) = \sum_{j=1}^{\infty} (a_j \sin(2j\pi\omega) + b_j \cos(2j\pi\omega))$ where $\sum_{j=1}^{\infty} (|a_j| + |b_j|) < \infty$, then $\sum_{k=1}^n f(k\omega)$ has a limit distribution as $n \rightarrow \infty$, and this limit distribution is the same as the distribution of $Z_{\infty} = \sum_{j=1}^{\infty} (a_j X_j + b_j Y_j)$, where*

$$X_j = \frac{\cos(j\pi\xi) - \cos(2j\pi\eta)}{2\sin(j\pi\xi)}, \quad Y_j = \frac{\sin(2j\pi\eta) - \sin(j\pi\xi)}{2\sin(j\pi\xi)}.$$

and ξ, η are independent and uniformly distributed on $[0, 1)$.

The condition $\sum_{j=1}^{\infty} (|a_j| + |b_j|) < \infty$ in the conjecture is equivalent to the absolute convergence of the Fourier series. This condition holds, e.g., if f is Hölder continuous with exponent $a > 1/2$ or if f is of bounded variation and Hölder continuous with exponent $a > 0$ (Katznelson, 2004). (a -Hölder continuity means that $|f(x) - f(y)| \leq C|x - y|^a$.)

The conjecture does not apply e.g. to the square wave function $f(\omega) = \text{sgn} \sin(2\pi\omega)$ because

$$f(\omega) = \frac{4}{\pi} \sum_{j=1}^{\infty} \frac{\sin((2j-1)2\pi\omega)}{2j-1},$$

and thus the sum of the Fourier coefficients diverges. We do know that an exponentially rare subsequence, the Rademacher system: $r_n(\omega) = \text{sgn} \sin(2^n \pi \omega)$

is iid thus for a suitable permutation of the terms $\text{sgn} \sin(2n\pi\omega)$ the CLT applies.

In a recent paper Pillai and Meng (2016) the authors prove that if (A_1, \dots, A_n) and (B_1, \dots, B_n) are iid multivariate normal $N(0, \Sigma)$ with positive main diagonal in Σ , and $a_j \geq 0$, $j = 1, 2, \dots, n$, then the sum

$$\sum_{j=1}^n a_j \frac{A_j}{B_j}$$

is Cauchy distributed with scale parameter $\sum_{j=1}^n a_j$.

This led us to the following conjecture.

Conjecture 2. *If $f(\omega) = \sum_{j=1}^{\infty} a_j \sin(2j\pi\omega)$ is a pure sine Fourier series with $a_j \geq 0$, $j = 1, 2, \dots$, and $a := \sum_{j=1}^{\infty} a_j < \infty$, then $\sum_{k=1}^n f(k\omega)$ has a limit distribution as $n \rightarrow \infty$, and this limit distribution is $a/2$ times standard Cauchy.*

Simulations, e.g. on square wave functions, suggest the following

Conjecture 3. *Let f be a 1-periodic L^2 function with $\int_0^1 f(\omega) d\omega = 0$, and let $c(n)$ denote the sum of the absolute values of the first n Fourier coefficients of f . Then $(1/c(n)) \sum_{k=1}^n f(k\omega)$ has a non-degenerate limit distribution.*

Remark 5. Kesten (1960; 1962) showed that if $f(\omega) = I(a \leq \omega \leq b) - (b-a)$ is a centered and periodically extended indicator function of a subinterval $[a, b] \subset [0, 1)$ than the limit distribution of

$$\frac{1}{\log n} \sum_{k=1}^n f(k\pi U + V)$$

is Cauchy whenever U, V are independent and uniformly distributed on $[0, 1)$. According to Kesten: “It seems still impossible to say anything about the asymptotic behavior of $\sum_{k=1}^n f(k\pi U + V)$ [...] for fixed V and only one random variable U .”

Theorem 6 (Fourier process). *Under the conditions of Theorem 4 the finite dimensional distributions of $F_n(t) = F_{[nt]}$ weakly converge to those of the process*

$$F(t) = \sum_{j=1}^{\infty} \left[a_j \frac{\cos(j\pi U) - \cos(2j\pi\gamma(t))}{2 \sin(j\pi U)} + b_j \frac{\sin(2j\pi\gamma(t)) - \sin(j\pi U)}{2 \sin(j\pi U)} \right]$$

where $\gamma(t)$ is defined in Theorem 3, U is uniformly distributed on $[0, 1)$ and it is independent of the process $\gamma(t)$.

Proof. This theorem can be proved by the same method we applied for the proof of Theorem 4. This time F_n , $F_{n,N}$ and Z_N are replaced by vectors, and we have to use the multivariate version of the distances ϱ and \mathcal{L} . For random vectors $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ let

$$\begin{aligned}\varrho(\boldsymbol{\varphi}, \boldsymbol{\psi}) &= \min\{\varepsilon \geq 0 : \mathbb{P}(\max_i |\varphi_i - \psi_i| > \varepsilon) \leq \varepsilon\}, \\ \mathcal{L}(\boldsymbol{\varphi}, \boldsymbol{\psi}) &= \min\{\varepsilon \geq 0 : \mathbb{P}(\varphi_i \leq t_i - \varepsilon, \forall i) - \varepsilon \\ &\leq \mathbb{P}(\psi_i \leq t_i, \forall i) \leq \mathbb{P}(\varphi_i \leq t_i + \varepsilon, \forall i) + \varepsilon, \forall \mathbf{t} \in \mathbb{R}^k\},\end{aligned}$$

then $\mathcal{L}(\boldsymbol{\varphi}, \boldsymbol{\psi}) \leq \varrho(\boldsymbol{\varphi}, \boldsymbol{\psi})$ remains valid. \square

4. Chebyshev coins and processes

Here we give two examples for orthogonal polynomial coins, Chebyshev polynomials of the first kind and Chebyshev polynomials of the second kind. This time let the probability space be the interval $[-1, 1]$ equipped with the σ -field of Lebesgue measurable sets and half of the Lebesgue measure as probability.

Theorem 7 (Chebyshev coins). *Let $p_n(\omega) = \cos(n \arccos \omega)$, $n = 0, 1, 2, \dots$ be Chebyshev polynomials of the first kind that are orthogonal with respect to the weight function $(1 - \omega^2)^{-1/2}$ on $[-1, 1]$ then $p_1 + \dots + p_n$ has a limit distribution as $n \rightarrow \infty$, with pdf $2h_1(2x + 1)$, where*

$$h_1(x) = \begin{cases} \frac{1}{2|x|^3}, & \text{if } |x| \geq 1, \\ \frac{1}{\pi|x|^3} \left(\arcsin x - x\sqrt{1-x^2} \right), & \text{if } 0 < |x| < 1. \end{cases} \quad (8)$$

A similar result is true if we consider Chebyshev polynomials of the second kind, q_n , that are orthogonal with respect to the weight function $(1 - \omega^2)^{1/2}$ on $[-1, 1]$, the limit distribution, however, will be different.

Proof. Let us start with the summation formula

$$p_1(\omega) + \dots + p_n(\omega) = \frac{\sin\left(\left(n + \frac{1}{2}\right) \arccos \omega\right)}{2 \sin\left(\frac{1}{2} \arccos \omega\right)} - \frac{1}{2} = \frac{\sin\left(\left(n + \frac{1}{2}\right) \arccos \omega\right)}{\sqrt{2(1 - \omega)}} - \frac{1}{2}.$$

Consider the fraction in the right-hand side. First we prove that the numerator and the denominator are asymptotically independent. More precisely, the joint limit distribution of the numerator and the denominator is equal to the distribution of $(\sin(2\pi V), 2\sqrt{U})$, where U and V are independent and uniformly distributed on $[0, 1]$. This follows from Lemma 1 with $r = 1$, $\varphi(n) = n + \frac{1}{2}$, $U(\omega) = \frac{1}{2}(1 - \omega)$ and $V(\omega) = \frac{1}{2\pi} \arccos(\omega)$.

The pdf of $\sin(2\pi V)$ is

$$\frac{1}{\pi} \cdot \frac{1}{\sqrt{1-x^2}}, \quad -1 < x < 1,$$

while that of \sqrt{U} is $2x$, $0 < x < 1$. Let h_1 denote the pdf of $\sin(2\pi V)/\sqrt{U}$. This random variable is symmetrically distributed, thus $h_1(x) = h_1(-x)$. Let $x > 0$ and $z = \min\{1, x\}$. Then we have

$$\begin{aligned} h_1(x) &= \int_0^{z/x} t \cdot \frac{1}{\pi\sqrt{1-t^2x^2}} \cdot 2t \, dt = \frac{2}{\pi x^3} \int_0^z \frac{y^2}{\sqrt{1-y^2}} \, dy \\ &= \frac{1}{4\pi x^3} \left(\arcsin z - z\sqrt{1-z^2} \right) \end{aligned}$$

after some computation. This immediately gives (8).

The sum of Chebyshev polynomials of the second kind,

$$q_n(\omega) = \frac{\sin((n+1) \arccos \omega)}{\sin(\arccos \omega)},$$

can be obtained from the sine summation formula.

$$\begin{aligned} q_1(\omega) + \dots + q_n(\omega) &= \frac{\cos\left(\frac{1}{2} \arccos \omega\right) - \cos\left((n + \frac{3}{2}) \arccos \omega\right)}{2 \sin\left(\frac{1}{2} \arccos \omega\right) \sin(\arccos \omega)} - 1 \\ &= \frac{1}{2(1-\omega)} - \frac{\cos\left((n + \frac{3}{2}) \arccos \omega\right)}{(1-\omega)\sqrt{2(1+\omega)}} - 1. \end{aligned}$$

Again, the cosine term in the right-hand side is asymptotically independent of the rest, thus the limit distribution is of the form

$$\frac{1}{4(1-U)} \left(1 - \frac{\cos(2\pi V)}{\sqrt{U}} \right) - 1,$$

where U and V are independent and uniformly distributed on $[0, 1]$. The explicit form of the pdf is left to the reader. \square

Remark 6. A formal substitution $t = 1$ into the well-known generating functions

$$\sum_{n=1}^{\infty} p_n(\omega) t^n = \frac{1 - t\omega}{1 - 2t\omega + t^2} - 1, \quad \sum_{n=1}^{\infty} q_n(\omega) t^n = \frac{1}{1 - 2t\omega + t^2} - 1$$

(see Abramowitz and Stegun, 1972, p.783) would lead to the (erroneous) deduction that the limit distribution of the sums of Chebyshev polynomials of the first kind is degenerate, concentrated on $-1/2$, or the limit distribution corresponding to Chebyshev polynomials of the second kind is equal to the distribution of $\frac{1}{2U} - 1$, where U is uniform on $[0, 1)$.

Theorem 8 (Chebyshev process). *Define the summation processes of Chebyshev polynomials of first and second kind as*

$$P_n(t) = \sum_{j=1}^{[nt]} p_j, \quad t > 0, \quad Q_n(t) = \sum_{j=1}^{[nt]} q_j, \quad t > 0,$$

respectively. Then $P_n(t) \rightarrow P(t)$ and $Q_n(t) \rightarrow Q(t)$ as $n \rightarrow \infty$ in the sense of weak convergence of finite dimensional distributions, where

$$P(t) = \frac{\sin(2\pi\gamma(t))}{2\sqrt{U}} - \frac{1}{2}, \quad Q(t) = \frac{1}{4(1-U)} \left(1 - \frac{\cos(2\pi\gamma(t))}{\sqrt{U}} \right) - 1,$$

$\gamma(t)$ is defined in Theorem 3, U is uniformly distributed on $[0, 1)$ and it is independent of the process $\gamma(t)$.

Proof. For Chebyshev polynomials of the first kind we can follow the first part of the proof of Theorem 7, applying Theorem 3. In the case of Chebyshev polynomials of the second kind, we only have to repeat the proof of Theorem 3 to show that $\{([nt] + \frac{3}{2}) \arccos \omega\} \rightarrow \gamma(t)$, as $n \rightarrow \infty$, and the limit process is asymptotically independent of $U(\omega) = \frac{1}{2}(1 + \omega)$. Details of the computations are straightforward, therefore we omit them. \square

5. Haar coins

In general, a pseudorandom process is the weak limit of suitably normalized partial sums of “coins” that is of identically distributed orthogonal functions. We cannot characterize all possible limit distributions. The difficulties are illustrated by the Haar coins, the orthogonal functions in the Haar

system where the limiting behavior is very much different from the behaviors in Sections 2–4.

Change the probability space back to $[0, 1)$.

Put $\chi_0(\omega) \equiv 1$. For $n > 0$ let $n = 2^k + m$, where $0 \leq m < 2^k$, and define

$$\chi_n(\omega) = \begin{cases} 2^{k/2}, & \text{if } m2^{-k} \leq \omega < (m + \frac{1}{2})2^{-k} \\ -2^{-k/2}, & \text{if } (m + \frac{1}{2})2^{-k} \leq \omega < (m + 1)2^{-k}, \\ 0 & \text{otherwise.} \end{cases}$$

Clearly,

$$\sum_{n=2^{k-1}}^{2^k-1} \chi_n = 2^{(k-1)/2} r_k,$$

where r_k stands for the k -th Rademacher function defined as $r_k(\omega) = \text{sgn} \sin(2^k \pi \omega)$.

We are going to discuss the asymptotic distribution of the normalized partial sums

$$K_n = \frac{1}{\sqrt{n}} \sum_{i=0}^n \chi_i,$$

and will see that K_n/\sqrt{n} does not have a limit distribution, it is only “merging”, in other words it has many accumulation distributions. None of them are scale mixtures of normals.

For a more precise statement introduce

$$W_k = 2^{-k/2} \sum_{n=0}^{2^k-1} \chi_n = 2^{-k/2} + \sum_{i=1}^k 2^{-i/2} r_{k-i+1}.$$

Obviously, W_k converges in distribution to the infinite sum $W = \sum_{k=1}^{\infty} 2^{-k/2} r_k$.

One can easily see that

$$K_n = \begin{cases} \sqrt{\frac{2^{k+1}}{n}} W_{k+1} & \text{on the interval } A = [0, (m+1)2^{-k}), \\ \sqrt{\frac{2^k}{n}} W_k & \text{on the interval } \bar{A} = [(m+1)2^{-k}, 1). \end{cases}$$

In a suitably enlarged probability space define the random variables $Z(a)$, $0 \leq a < 1$ by

$$Z(a) = \begin{cases} 2^{(1-a)/2} & \text{with probability } 2^a - 1, \\ 2^{-a/2} & \text{with probability } 2 - 2^a, \end{cases}$$

and let $Z(a)$ be independent of W . Finally, let $w(\xi, \eta)$ denote the Wasserstein distance between the distributions of the integrable random variables ξ and η , that is, $w(\xi, \eta)$ is the infimum of $\mathbb{E}|\xi' - \eta'|$ over all pairs (ξ', η') , where ξ and ξ' are identically distributed, just like η and η' .

Theorem 9.

$$w(K_n, Z(\{\log_2 n\})W) = O(n^{-1/3}),$$

where $\{\cdot\}$ denotes the fractional part. Hence, the accumulation distributions of the sequence (K_n) coincide with the distributions of $Z(a)W$, $0 \leq a < 1$.

Remark 7. The accumulation distributions cannot be scale mixtures of normal laws, because they are bounded: $|W| < \sqrt{2} - 1$, $Z(a) \leq \sqrt{2}$.

Proof of Theorem 9. Approximate the quantity $\sqrt{2^{-k}n} K_n = W_k + r_{k+1}I(A)$ in two steps. Put $\ell = \lfloor k/3 \rfloor$ and let j be defined by $j2^{-\ell} \leq (m+1)2^{-k} < (j+1)2^{-\ell}$, finally let B denote the interval $[0, j2^{-\ell})$. With this notation, our first approximation is

$$\sum_{i=1}^{k-\ell} 2^{-i/2} r_{k-i+1} + \sum_{i=k-\ell+1}^k 2^{-i/2} r_{\ell+i+1} + r_{k+1}I(B).$$

It has the same distribution as $Z = W + X I(B)$, where W , X , and B are independent, and $\mathbb{P}(X = 1) = \mathbb{P}(X = -1) = 1/2$. The error of approximation is bounded as follows.

$$\left| \sqrt{2^{-k}n} K_n - Z \right| \leq 2^{-k/2} + \sum_{i=k-\ell+1}^k 2^{-i/2} + \sum_{i=k-\ell+1}^{\infty} 2^{-i/2} + I(A \setminus B).$$

Its expectation is less than $2^{-\ell} + 5 \cdot 2^{-(k-\ell)/2} = O(2^{-k/3}) = O(n^{-1/3})$.

The second approximation will be $Z' = W + X I(B')$, where $B' = [0, m2^{-k})$ (and the independence of W , X and B' is supposed). Then $\mathbb{E}|Z - Z'| \leq 2^{-\ell}$, thus we have

$$w(\sqrt{2^{-k}n} S_n, Z') = O(n^{-1/3}),$$

and therefore

$$w(S_n, \sqrt{2^k n^{-1}} Z') = O(n^{-1/3}).$$

Finally, note that $\mathbb{P}(B') = 2^a - 1$ and $\sqrt{2^k n^{-1}} = 2^{-a/2}$, where $a = \{\log_2 n\}$; furthermore, the distribution of $W + X$ coincides with that of $\sqrt{2}W$. The proof is completed. \square

6. Conclusion

Pseudorandom numbers turned out to be very effective substitutes of real random numbers e.g. in Monte Carlo method applications, see von Neumann (1951) and Knuth (1997, Ch.3). More recently other pseudorandom versions of important random objects, like random graphs, have also been studied, see e.g. Krivelevich and Sudakov (2006). This paper is the first attempt to provide pseudorandom alternatives of random/stochastic processes in the hope of important applications.

Acknowledgment

Part of this research was based on G. J. Székely's work supported by the National Science Foundation, while working at the Foundation.

T. F. Móri was supported by the Hungarian National Research, Development and Innovation Office NKFIH – Grant No. K125569.

Finally, the authors are indebted to Craig L. Zirbel and István Berkes for their constructive suggestions and to the editor, Philip Ernst, for his excellent work in handling our manuscript.

References

- Abramowitz, M; Stegun, I. A. (eds.) (1972) *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York. ISBN 978-0-486-61272-0
- Adams, R. A.; Fournier, J. (2003) *Sobolev Spaces*. Pure and Applied Mathematics 140 (2nd ed.) Academic Press, Boston, MA. ISBN 978-0-120-44143-3
- Berkes, I. (2017) On the uniform theory of lacunary series In: Elsholtz C., Grabner P. (eds) *Number Theory – Diophantine Problems, Uniform Distribution and Applications*. Springer, Cham, pp 137–167. doi: 10.1007/978-3-319-55357-3_6
- Billingsley, P. (2012) *Probability and Measure* (4th ed.) Wiley, New York ISBN 978-1-118-34191-9
- Billingsley, P. (2013) *Convergence of Probability Measures* (2nd ed.) Wiley, New York. ISBN 978-1-118-62596-5

- Chung, K. L. (2000) *A Course in Probability Theory* (3rd ed.) Academic Press, San Diego, CA. ISBN 978-0-121-74151-8
- Doob, J. L. (1953) *Stochastic Processes* Wiley, New York. ISBN 978-0-471-52369-0
- Gaposhkin, V. F. (1966) Lacunary series and independent events. *Russ. Math. Surv.*, **21/6**, 1–82. doi: 10.1070/RM1966v021n06ABEH001196
- Getoor, R. (2009) J. L. Doob: Foundations of stochastic processes and probabilistic potential theory. *Ann. Probab.* **37/5**, 1647–1663. doi: 10.1214/09-AOP465
- János Bolyai Math. Soc. (2015) Schweitzer Contest in Higher Mathematics 2014. (Hungarian) *Mat. Lapok (N.S.)* **21/1**, 77–88.
- Katznelson, Y. (2004) *An Introduction to Harmonic Analysis* (3rd ed.) Cambridge Mathematical Library, New York. ISBN 978-0-521-54359-0
- Kesten, H. (1960) Uniform distribution mod 1. *Ann. of Math. (2)* **71/3**, 445–471. doi: 10.2307/1969938
- Kesten, H. (1962) Uniform distribution mod 1 (II). *Acta Arith.* **7**, 355–380. doi: 10.4064/aa-7-4-355-380
- Knuth, D. (1997) *The Art of Computer Programming, Vol 2: Seminumerical Algorithms* (3rd ed.) Addison-Wesley, Boston, MA. ISBN 978-0-201-89684-8
- Krivelevich, M.; Sudakov, B. (2006) Pseudo-random graphs. In: Györi, E.; Katona, G. O. H.; Lovász, L.; Fleiner, T. (eds) *More Sets, Graphs and Numbers*. Bolyai Society Mathematical Studies, vol 15. Springer, Berlin, Heidelberg, pp.199–262. doi: 10.1007/978-3-540-32439-3_10
- Phadke, V. (2008) *Non-classical convergence results for sums of dependent random variables*, Electronic Thesis or Dissertation, Bowling Green State University. https://etd.ohiolink.edu/rws_etd/document/get/bgsu1224514478/inline Accessed 16 Aug 2018
- Pillai, N. S.; Meng, X. L. (2016) An unexpected encounter with Cauchy and Lévy. *Ann. Statist.* **44/5**, 2089–2097. doi: 10.1214/15-AOS1407

Rényi, A. (1958) On mixing sequences of sets. *Acta Math. Acad. Sci. Hungar.*, **9**, 215–228. doi: 10.1007/BF02023873

von Neumann, J. (1951) Various techniques used in connection with random digits. *J. Res. Nat. Bur. Stand. Appl. Math. Ser.* **12**, 36–38.