HAL open science

Random subsequences of αn with asymptotically independent successive terms

Bruno Massé

▶ To cite this version:

Bruno Massé. Random subsequences of αn with asymptotically independent successive terms. [Research Report] Université du Littoral - Côte d'Opale. 2021. hal-03502439

HAL Id: hal-03502439
https://hal.science/hal-03502439v1

Submitted on 25 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Random subsequences of $\{\alpha n\}$ with asymptotically independent successive terms

Bruno Massé [1]

November 18, 2021

Abstract

Fix an irrational number α and let Y_n be the number of attempts needed to get the nth success in a non-stationary sequence of independent Bernoulli trials. It is known that the law of the fractional part of αY_n converges weakly to the uniform distribution on $[0,1)$, as $n \to +\infty$, when the probabilities of success decrease to 0 and sum to $+\infty$. We provide sufficient conditions on the probabilities of success ensuring that the fractional parts of αY_n and αY_{n+1} are asymptotically independent. We extend our results to any number of successive terms, compute upper bounds of the convergence rates depending on a measure of irrationality of α and on the probabilities of success and apply our results to discuss the mantissae of 2^{Y_n} and $2^{Y_{n+1}}$.

1 Introduction

We denote by U^k the uniform distribution on $[0,1)^k$ and by $\{y\}$ the fractional part of a real y. Since we deal with fractional parts, we identify $[0,1)^k$ to the the k-dimensional torus equipped with its natural topology. That allows us to use Lemmas 1 and 5 below on which our proofs rely.

The uniform probability measure (u.p.m. in the sequel) on the set $\big\{\{\alpha n\} : n = 1,\dots,N\big\}$ converges weakly to U^1 as $N \to +\infty$ when α is irrational [9, p. 8]. In other words, the sequences $(\alpha n)_n$ with irrational α are uniformly distributed modulo 1 in the sense of the natural density.

If we are given α and the value x of $\{\alpha n\}$ (but not the value of n), then we know that $\{\alpha(n+1)\} = \big\{\{\alpha n\} + \{\alpha\}\big\} = \big\{x + \{\alpha\}\big\}$. By contrast, it is proved in [5] that the u.p.m. on the set $\big\{\big(\{\alpha n^\gamma\}, \{\alpha(n+1)^\gamma\}\big) : n = 1,2,\dots,N\big\}$ converges weakly to U^2 when $\gamma > 1$. This can be interpreted in the following manner: if $\gamma > 1$ and if we are given the values of α and $\{\alpha n^\gamma\}$ (but not the value of n), that does not give us any information on the value of $\{\alpha(n+1)^\gamma\}$; the possible values of $\{\alpha(n+1)^\gamma\}$ are still uniformly distributed in $[0,1)$. Another example of a sequence uniformly distributed in $[0,1)^2$ is $\big(\{\alpha_1 n\}, \{\alpha_2 n\}\big)_n$ where α_1 and α_2 are linearly independent over the rationals [9, p. 48–49].

Univ. Littoral Côte d'Opale, UR 2597, LMPA, Laboratoire de Mathématiques Pures et Appliquées Joseph Liouville, F-62100 Calais, France

Key words and phrases: Modulo 1 distribution, Benford law, Fourier coefficients, Weak convergence, Asymptotic independence, Measure of irrationality.

2010 Mathematics subject classification: 60B10, 62E20, 11K06, 11J82

Consider an irrational number α and a sequence of independent Bernoulli trials whose probabilities of success sum to $+\infty$ and are either all equal or decrease to zero as the process moves forward. For all n, denote by Y_n the number of attempts needed to get the n-th success. We are interested in the present paper in the sequence $(\{\alpha Y_n\})_n$ which is the random subsequence of $(\{\alpha n\})_n$ obtained by removing from $(\{\alpha n\})_n$ all the points $\{\alpha j\}$ for which the j-th trial fails. See [2] and [15] for the investigation of random subsequences of $(\{\alpha n\})_n$ of another kind.

It is proved in [11] that, as $n \to +\infty$, the law of $\{\alpha Y_n\}$ converges weakly to U^1 whenever α is irrational. As shown at the beginning of Section 2 below, $\{\alpha Y_n\}$ and $\{\alpha Y_{n+1}\}$ are not asymptotically independent (the law of $\big(\{\alpha Y_n\}, \{\alpha Y_{n+1}\}\big)_n$ does not converge weakly to U^2) when $q_j = q \in (0,1)$ for all j. Moreover, some computer simulations we have made on the mantissa in base 10 of 2^{Y_n} when $q_j = 1/\log j$ ($\log j$ denotes the natural logarithm of j) suggest that this is still true for $\{\alpha Y_n\}$ and $\{\alpha Y_{n+1}\}$ if $q_j = 1/\log j$ and α is the logarithm of 2 in base 10 (these simulations were motivated by the fact that Y_n is the n-th Cramér's random pseudo-primes [7] when $q_j = 1/\log j$). See Section 4 for the definition of the mantissa of a positive real number and for the link between distribution modulo 1 and mantissa distribution.

We prove below that, when $q_j = 1/j^\beta$ with $1/3 < \beta \leq 1$, the law of the random vector $\big(\{\alpha Y_n\}, \{\alpha Y_{n+1}\}\big)_n$ converges weakly, as $n \to +\infty$, to U^2, thus that the random variables $\{\alpha Y_n\}$ and $\{\alpha Y_{n+1}\}$ are asymptotically independent. This can be interpreted in the following manner: if $q_j = 1/j^\beta$ with $1/3 < \beta \leq 1$ and if we are given the values of α, n_0 and $\{\alpha Y_{n_0}(\omega)\}$, that gives us almost no information on the value of $\{\alpha Y_{n_0+1}(\omega)\}$ when n_0 is large; the possible values of $\{\alpha Y_{n_0+1}(\omega)\}$ are still almost uniformly distributed in $[0,1)$.

We then extend this result to the case of any number of successive terms and compute some bounds of the convergence rates depending on the *type* (see Section 3) of α and the q_j. We finally use our results to investigate the asymptotic independence of the *mantissae* (see Section 4) of 2^{Y_n} and $2^{Y_{n+1}}$. Note that the bounds obtained in [2] and [15] depend on the type of α too.

1.1 Definitions and notation

In all the present paper, *independent* will mean *mutually independent*. We will use the standard notation: $\log_b a$ for the logarithm to the base b of a and $e(x)$ for $\exp(2i\pi x)$ with $i^2 = -1$. The type of an irrational number and the mantissa of a positive real are defined respectively in Sections 3 and 4.

Let $(X_j)_{j\geq 1}$ be a sequence of independent Bernoulli random variables. We suppose that the probabilities of success $q_j = P(X_j = 1)$ sum to $+\infty$. According to Borel-Cantelli Lemma, this is necessary and sufficient to ensure that $S_N = \sum_{j=1}^N X_j \to +\infty$ almost surely as $N \to +\infty$ and so that the number of trials needed to get the n-th success $Y_n = \min\{N : S_N = n\}$ is almost surely well defined for all n. The q_j are all equal in Proposition 1 and Corollary 2 and decrease to zero and sum to $+\infty$ in our other results.

1.2 Content

After demonstrating that $\{\alpha Y_n\}$ and $\{\alpha Y_{n+1}\}$ are not asymptotically independent when the q_j are all equal, we prove in Section 2 that Condition (1) below ensures that $\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\}$ are asymptotically independent for all $k \geq 1$ when the q_j

decrease to zero and sum to $+\infty$. Section 3 is devoted to the convergence rates of the law of the random vector $\big(\{\alpha Y_n\},\dots,\{\alpha Y_{n+k}\}\big)$ to U^{k+1}. In Section 4, we apply the results of Sections 2 and 3 to the case $\alpha=\log_b 2$, where b is not a power of 2, in order to investigate the random subsequence of $(2^n)_n$ generated by our Bernoulli trials.

2 Main result

We begin with a situation where $\{\alpha Y_n\}$ and $\{\alpha Y_{n+1}\}$ are not asymptotically independent.

Proposition 1. *Fix $q\in(0,1)$ and suppose that $q_j=q$ for all j. Then $\{\alpha Y_n\}$ and $\{\alpha Y_{n+1}\}$ are not asymptotically independent.*

Proof. In situations where the law of $\big(\{\alpha Y_n\},\{\alpha Y_{n+1}\}\big)$ converges weakly to U^2, the law of $\big\{\{\alpha Y_{n+1}\}-\{\alpha Y_n\}\big\}$ converges weakly to U (see [3, p. 379]) since U is the probability distribution of $\{Z_2-Z_1\}$ where the random vector (Z_1,Z_2) is distributed following U^2. We suppose now that $q_j=q\in(0,1)$ for all j and prove that then the law of $\big\{\{\alpha Y_{n+1}\}-\{\alpha Y_n\}\big\}$ does not converge weakly to an absolutely continuous probability distribution.

The X_n being independent, for all (n,m) satisfying $n\le m$,

$$\begin{aligned}P\big((Y_{n+1}-Y_n=1)|(Y_n=m)\big)&=\frac{P\big((Y_n=m)\cap(Y_{n+1}=m+1)\big)}{P(Y_n=m)}\\&=\frac{qP(S_{m-1}=n-1)q}{qP(S_{m-1}=n-1)}=q\end{aligned}$$

does not depend on m and so $P(Y_{n+1}-Y_n=1)=q$ for all n. Hence

$$P\Big(\big\{\{\alpha Y_{n+1}\}-\{\alpha Y_n\}\big\}=\{\alpha\}\Big)=q>0$$

for all n because $\{x-y\}=\{\{x\}-\{y\}\}$ and because the values of $\{\alpha k\}$ for $k=1,2,\dots$ are distinct from one another when α is irrational. So the law of $\big\{\{\alpha Y_{n+1}\}-\{\alpha Y_n\}\big\}$ does not converge weakly to an absolutely continuous probability distribution and Proposition 1 is proved.

□

When the probabilities of success q_j decrease to zero and sum to $+\infty$, the laws of $\{\alpha Y_n\}$, ... and $\{\alpha Y_{n+k}\}$ converge weakly to U^1 whenever α is irrational [11]. Therefore $\{\alpha Y_n\},\dots$ and $\{\alpha Y_{n+k}\}$ are asymptotically independent if and only if the law of $\big(\{\alpha Y_n\},\dots,\{\alpha Y_{n+k}\}\big)$ converges weakly to U^{k+1}. The condition

$$\sum_{m=2}^{+\infty}(q_1+\cdots+q_{m-1})^{-1/2}q_mq_{m+1}<+\infty \tag{1}$$

will be called *Condition* (1) in the sequel.

Note that Condition (1) is not satisfied in Proposition 1. Here are our main result and a direct consequence of it.

Theorem 1. *Suppose that the q_j decrease to zero and sum to $+\infty$. For any $k \geq 1$, the $k+1$ consecutive terms $\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\}$ are asymptotically independent when Condition* (1) *is satisfied.*

If $q_j \leq 1/j$, then $(q_1 + \cdots + q_{m-1})^{-1/2} q_m q_{m+1} \leq 1/(q_1^{1/2} m^2)$ and Condition 1 is satisfied. Set now $q_j = 1/j^\beta$ with $0 < \beta < 1$. Then $(q_1 + \cdots + q_{m-1}) \sim (1-\beta)^{-1} m^{1-\beta}$ and $q_m q_{m+1} \sim 1/m^{2\beta}$ and thus $(q_1 + \cdots + q_{m-1})^{-1/2} q_m q_{m+1} = \mathcal{O}\left(1/m^{\frac{3}{2}\beta + \frac{1}{2}}\right)$. This shows that Condition 1 is satisfied if $\beta > 1/3$. This demonstrates the following corollary.

Corollary 1. *The random variables $\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\}$ are asymptotically independent in particular when $q_j = 1/j^\beta$ with $1/3 < \beta < 1$ and when $q_j = 1/j$ or $q_j = 1/j \log j$ and so on.*

Theorem 1 is demonstrated below. Before that, we present the main lemmas used in our proofs.

2.1 Lemmas

We collect here the main tools used in our proofs. The weak convergence of a sequence of probability measures on the k-dimensional torus $[0,1)^k$ is characterized by the convergence of the corresponding Fourier coefficients. We present here the case where the limit distribution is U^k (Weyl criterion). A direct proof can be obtained by extending the arguments in [9, p. 7] to the k-dimensional case and general sequences of probability measures. It is also a simple consequence of Lemma 5 below. Let $A_1, A_2, \ldots$ be some random vectors taking their values in the k-dimensional torus and denote by $\mathbb{Z}$ the set of integers.

Lemma 1. *In order that the law of A_n converges weakly to U^k as $n \to +\infty$ it is necessary and sufficient that, for every $h \in \mathbb{Z}^k$ satisfying $h \neq (0, \ldots, 0)$,*

$$\lim_{n \to +\infty} \mathbb{E}(e(h \cdot A_n)) = 0$$

where $h \cdot A_n$ designates the dot product of h and A_n.

Lemma 2 is a particular case of the Kolmogorov-Rogozin inequality on the Lévy concentration function (Theorem 4 in [13, p. 44]). It will avoid us to upper-bound roughly $P(S_{m-1} = n-1)$ by 1 in the proofs of Theorems 1 and 2.

Lemma 2. *When the q_j decrease to* 0,

$$\max_{0 \leq l \leq N} P(S_N = l) = \mathcal{O}\left((q_1 + \cdots + q_N)^{-1/2}\right)$$

where $\mathcal{O}$ denotes the Landau big O.

The next lemma is elementary but we do not have a reference.

Lemma 3. *If $q_j \in (0,1)$ for all j and $\sum q_j = +\infty$, then, for all fixed m,*

$$\sum_{n=m+1}^{+\infty} q_n \prod_{j=m+1}^{n-1} (1 - q_j) = 1$$

with the value of an empty product is one.

Proof. By induction on N,

$$\sum_{n=m+1}^{m+N} q_n \prod_{j=m+1}^{n-1} (1-q_j) = 1 - \prod_{n=m+1}^{m+N} (1-q_n).$$

Now

$$\lim_{N\to+\infty} \prod_{n=m+1}^{m+N} (1-q_n) = 0$$

when $0 < q_n < 1$ and $\sum q_n = +\infty$.

Another argument is that

$$q_n \prod_{j=m+1}^{n-1} (1-q_j) = P(Z_m = n)$$

where Z_m is the rank of the first success beyond the m-th attempt (Z_m is almost surely well defined when $\sum q_n = +\infty$). □

The next lemma is demonstrated in the proof of Theorem 1 in [11].

Lemma 4. *Let α be any irrational. For any non-nul integer h,*

$$|\mathbb{E}(e(h\alpha Y_n))| \le \frac{3q_n}{|\sin(\pi h\alpha)|} \max_{0\le j\le n-1} P(S_{n-1} = j). \tag{2}$$

2.2 Proof of Teorem 1 for $k = 1$

In this section, we show that Condition (1) ensures the asymptotic independence of two consecutive terms.

Let α be any irrational. By Lemma 4, if $h_1 = 0$ and $h \neq 0$, then

$$\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| = \big|\mathbb{E}\big(e(h\alpha Y_n)\big)\big|$$

tends to 0 as $n \to +\infty$ because $\lim_n q_n = 0$ (see also Lemma 2).

Having Lemma 1 in mind, we verify now that, for all (h, h_1) such that $h_1 \neq 0$,

$$\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big) = \sum_{m=n}^{+\infty} \sum_{m_1=m+1}^{+\infty} e(h\alpha m + h_1\alpha m_1) P((Y_n = m) \cap (Y_{n+1} = m_1))$$

tends to 0 as $n \to +\infty$ when condition (1) is satisfied.

The random variables X_n being independent,

$$P\big((Y_n = m) \cap (Y_{n+1} = m+1)\big) = P(S_{m-1} = n-1) q_m q_{m+1}$$

for all $m \ge n$ and

$$P\big((Y_n = m) \cap (Y_{n+1} = m_1)\big) = P(S_{m-1} = n-1) q_m q_{m_1} \prod_{j=m+1}^{m_1-1} (1-q_j)$$

for all $m_1 > m+1$. Hence, since $|e(h\alpha m)| = 1$,

$$\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| \le \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) \left| \sum_{m_1=m+1}^{+\infty} e(h_1\alpha m_1) Q_{m_1}(m) \right| \tag{3}$$

where $Q_{m+1}(m) = q_{m+1}$ and, for $m_1 > m+1$,

$$Q_{m_1}(m) = q_{m_1} \prod_{j=m+1}^{m_1-1} (1-q_j).$$

For all $m_1 \geq m$,

$$\left| \sum_{l=m+1}^{m_1} e(h_1 \alpha l) \right| \leq \frac{2}{|e(h_1\alpha)-1|} = \frac{1}{|\sin(\pi h_1 \alpha)|}.$$

So a summation by parts leads to

$$\left| \sum_{m_1=m+1}^{+\infty} e(h_1 \alpha m_1) Q_{m_1}(m) \right| \leq \frac{1}{|\sin(\pi h_1 \alpha)|} \sum_{m_1=m+1}^{+\infty} |Q_{m_1+1}(m) - Q_{m_1}(m)|. \tag{4}$$

For all m, the sequence $(Q_{m_1}(m))_{m_1}$ is decreasing (recall that $(q_n)_n$ is decreasing). Therefore

$$\sum_{m_1=m+1}^{+\infty} |Q_{m_1+1}(m) - Q_{m_1}(m)| = Q_{m+1}(m) = q_{m+1}.$$

Combining this with (3) and (4) gives

$$\left|\mathbb{E}\big(e(h\alpha Y_n + h_1 \alpha Y_{n+1})\big)\right| \leq \frac{1}{|\sin(\pi h_1 \alpha)|} \sum_{m=n}^{+\infty} q_m q_{m+1} P(S_{m-1} = n-1).$$

By Lemma 2, $P(S_{m-1} = n-1) = \mathcal{O}\left((q_1 + \cdots + q_{m-1})^{-1/2}\right)$ where the constant implied by $\mathcal{O}$ is independent of m. The proof is completed.

2.3 Proof of Teorem 1 for $k = 2$

In this section, we show that Condition (1) ensures the asymptotic independence of three consecutive terms.

If $h_2 = 0$ and $(h, h_1) \neq (0,0)$ and if condition (1) is satisfied, the $k = 1$ case of Theorem 1 indicates that

$$|\mathbb{E}(e(h\alpha Y_n + h_1 \alpha Y_{n+1} + h_2 \alpha Y_{n+2}))| = |\mathbb{E}(e(h\alpha Y_n + h_1 \alpha Y_{n+1}))|$$

tends to 0 as $n \to +\infty$.

Fix $h_2 \neq 0$ and (h, h_1). We are going to verify that

$$A_n := \mathbb{E}\big(e(h\alpha Y_n + h_1 \alpha Y_{n+1} + h_2 \alpha Y_{n+2})\big)$$

tends to 0 as $n \to +\infty$ when condition (1) is satisfied.

The random variables X_n being independent, for all $m_2 > m_1 > m \geq n$,

$$P\big((Y_n = m) \cap (Y_{n+1} = m_1) \cap (Y_{n+2} = m_2)\big) = q_m P(S_{m-1} = n-1) Q_{m_1}(m) Q_{m_2}(m_1)$$

where $Q_t(s) = q_t \prod_{j=s+1}^{t-1}(1-q_j)$. This leads to

$$|A_n| = \left| \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) e(h\alpha m) \sum_{\substack{m_1 \geq m+1 \\ m_2 \geq m_1+1}} Q_{m_1}(m) Q_{m_2}(m_1) e(h_1 \alpha m_1 + h_2 \alpha m_2) \right|$$
$$\leq \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) \left| \sum_{m_1=m+1}^{+\infty} Q_{m_1}(m) e(h_1 \alpha m_1) \sum_{m_2=m_1+1}^{+\infty} Q_{m_2}(m_1) e(h_2 \alpha m_2) \right|$$
$$\leq \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) \sum_{m_1=m+1}^{+\infty} Q_{m_1}(m) \left| \sum_{m_2=m_1+1}^{+\infty} Q_{m_2}(m_1) e(h_2 \alpha m_2) \right|.$$

Using similar arguments to those above, we get

$$\left| \sum_{m_2=m_1+1}^{+\infty} Q_{m_2}(m_1) e(h_2 \alpha m_2) \right| \leq \frac{q_{m_1+1}}{|\sin(\pi h_2 \alpha)|}.$$

Therefore

$$|A_n| \leq \frac{1}{|\sin(\pi h_2 \alpha)|} \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) \sum_{m_1=m+1}^{+\infty} Q_{m_1}(m) q_{m_1+1}$$
$$\leq \frac{1}{|\sin(\pi h_2 \alpha)|} \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) q_{m+1} \sum_{m_1=m+1}^{+\infty} Q_{m_1}(m)$$

because (q_n) is non-increasing.

This and Lemmas 2 and 3 conclude our proof.

2.4 Proof of Teorem 1 for $k \geq 3$

We proceed by induction. Fix $(h, h_1, \ldots, h_k) \neq (0, 0, \ldots, 0)$ and suppose that

$$\lim_{n \to +\infty} \mathbb{E}(e(h\alpha Y_n + h_1 \alpha Y_{n+1} + \cdots + h_{k-1} \alpha Y_{n+k-1})) = 0$$

for all $(h, h_1, \ldots, h_{k-1}) \neq (0, 0, \ldots, 0)$. It remains to verify that

$$A_n^k := \mathbb{E}\big(e(h\alpha Y_n + h_1 \alpha Y_{n+1} + \cdots + h_k \alpha Y_{n+k})\big)$$

tends to 0 as $n \to +\infty$ when condition (1) is satisfied and $h_k \neq 0$. Set again $Q_t(s) = q_t \prod_{j=s+1}^{t-1}(1-q_j)$. Then, using similar arguments to those above, we get

$$|A_n^k| \leq \sum_{m=n}^{+\infty} q_m P(S_{m-1} = n-1) \sum_{m_1=m+1}^{+\infty} Q_{m_1}(m) \cdots \sum_{m_{k-1}=m_{k-2}+1}^{+\infty} Q_{m_{k-1}}(m_{k-2}) B_{m_{k-1}}$$

where

$$B_{m_{k-1}} := \left| \sum_{m_k=m_{k-1}+1}^{+\infty} Q_{m_k}(m_{k-1}) e(h_k \alpha m_k) \right| \leq \frac{q_{m_{k-1}+1}}{|\sin(\pi h_k \alpha)|}.$$

We get

$$|A_n^k| \leq \frac{1}{|\sin(\pi h_k \alpha)|} \sum_{m=n}^{+\infty} q_m q_{m+1} P(S_{m-1} = n-1) \sum_{m_1=m+1}^{+\infty} Q_{m_1}(m) \cdots \sum_{m_{k-1}=m_{k-2}+1}^{+\infty} Q_{m_{k-1}}(m_{k-2})$$

since $q_{m_{k-1}+1} \leq q_{m+1}$.

Indeed $P(S_{m-1} = n-1) = \mathcal{O}\left((q_1 + \cdots + q_{m-1})^{-1/2}\right)$ by Lemma 2 and

$$\sum_{m_1=m+1}^{+\infty} Q_{m_1}(m) = \cdots = \sum_{m_{k-1}=m_{k-2}+1}^{+\infty} Q_{m_{k-1}}(m_{k-2}) = 1$$

by Lemma 3. Thus

$$|A_n^k| = \frac{1}{|\sin(\pi h_k \alpha)|} \mathcal{O}\left(\sum_{m=n}^{+\infty} q_m q_{m+1} (q_1 + \cdots + q_{m-1})^{-1/2}\right)$$

and this shows that $\lim_{n\to+\infty} A_n^k = 0$ when condition (1) is satisfied. Our proof is completed.

3 Rates of convergence

Aiming at clarity, we treat only the case of two consecutive terms and of $q_n = 1/n^\beta$ with $1/3 < \beta < 1$, $q_n = 1/n$ and $q_n = 1/n \log n$ and so on. However our methods can easily be extended (at the cost of less and less readable computations as the number of terms grows).

Let Q_n designates the law of $(\{\alpha Y_n\}, \{\alpha Y_{n+1}\})$. We are now concerned with the Kolmogorov-Smirnov distance between Q_n and U^2

$$\Delta(Q_n, U^2) := \sup_{(s,t)\in[0,1)^2} |Q_n\big([0,s) \times [0,t)\big) - st|.$$

We have no reasons to think that the bounds presented below are sharp, but their proofs have the merit to show why the convergence rates of the probability distribution of $(\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\})$ to U^{k+1} depend most likely on the values of the q_n and on a measure of irrationality of α which is defined in [9, p. 161] by

$$\eta(\alpha) = \sup\{\gamma : \liminf_{h\to+\infty} h^\gamma \langle h\alpha\rangle = 0\}.$$

It measures the closeness of α to the rationals with *reasonable* denominators. The Liouville numbers [14, p. 310] are of infinite type; they are very well approximated by rationals. The algebraic numbers are of type 1 [4, p. 169]; they are badly approximated by rationals. Here is what is known on the type (rounded to one decimal place) of some common transcendental numbers:

$$\eta(e) = 1\,,\ \eta(\pi) < 6.2\,,\ \eta(\pi^2) < 4.5\,,\ \eta(\log 2) < 2.6 \ \text{ and } \ \eta(\log 3) < 4.2.$$

See Section 4 for some bounds of $\eta(\log_b a)$ where a and b are positive integers and see [11] for references and more information about $\eta(\alpha)$.

The main result of the present section is the following theorem. It can be slightly improved for quadratic irrationals α because they are of *of constant type* (they are in a way more badly approximated by rationals than e and that the other irrational algebraic numbers). See [9, p. 161] and [11] for the definition and more details. Aiming at lighten the present text, we will only use the fact that these numbers are of type 1 like the other algebraic numbers and like e.

Theorem 2. *Set $q_n = 1/n^\beta$ with $1/3 < \beta \leq 1$. Then, for all $\eta > \eta(\alpha)$,*

$$\Delta(Q_n, U^2) = \mathcal{O}\left(n^{-(3\beta-1)/2\eta}\right)$$

Here are four direct consequences of Theorem 2 and of the information about $\eta(\alpha)$ given above. If $\alpha = e$ or α is an irrational algebraic number, then

$$\Delta(Q_n, U^2) = \mathcal{O}\left(n^{-\gamma}\right) \quad \text{for all } \gamma < 1/4,$$

when $q_n = 1/\sqrt{n}$ and

$$\Delta(Q_n, U^2) = \mathcal{O}\left(n^{-\gamma}\right) \quad \text{for all } \gamma < 1$$

when $q_n = 1/n$. If $\alpha = \pi$, then

$$\Delta(Q_n, U^2) = \mathcal{O}\left(n^{-1/24.8}\right),$$

when $q_n = 1/\sqrt{n}$ and

$$\Delta(Q_n, U^2) = \mathcal{O}\left(n^{-1/6.2}\right).$$

when $q_n = 1/n$.

A proof of Theorem 2 is given in Section 3.2. It will use the three following lemmas.

3.1 Lemmas

Lemma 5 is the Erdős-Turán-Koksma inequality. It is a kind of Berry-Esseen theorem on the 2-dimensional torus. Theorem 2 in [12] is the most general version in the univariate case. Here is a simplified formulation which is sufficient in our network.

Lemma 5. *Let Q stands for the law of a 2-dimensional random vector $Z = (Z_1, Z_2)$ and set $H = \max(|h|, |h_1|)$ and $R = \max(|h|, 1)\max(1, |h_1|)$. Then, for every positive integer M,*

$$\Delta(Q, U^2) \leq C\left(\frac{1}{M} + \sum_{0<H\leq M} \frac{\left|\mathbb{E}(e(hZ_1 + h_1 Z_2))\right|}{R}\right)$$

where the constant C is independent of M and Q.

We give now some bounds for $|\mathbb{E}(e(h\alpha Y_n))|$ and $|\mathbb{E}(e(h\alpha Y_n + h_1\alpha Y_{n+1}))|$ in order to make use of Lemma 5 in the case $q_n = 1/n^\beta$ with $0 < \beta \leq 1$.

Lemma 6. *Set $q_n = 1/n^\beta$ with $0 < \beta \leq 1$. Then, if $h \neq 0$,*

$$|\mathbb{E}(e(h\alpha Y_n))| = \frac{1}{|\sin(\pi h\alpha)|}\mathcal{O}\left(1/n^{(\beta+1)/2}\right)$$

and, if $h_1 \neq 0$ and $1/3 < \beta \leq 1$,

$$\left|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\right| = \frac{1}{|\sin(\pi h_1\alpha)|}\mathcal{O}\left(1/n^{(3\beta-1)/2}\right)$$

where the constants implied by $\mathcal{O}$ are independent of h and h_1.

Proof. We know that, if $h \neq 0$,

$$|\mathbb{E}(e(h\alpha Y_n))| \leq \frac{3q_n}{|\sin(\pi h\alpha)|} \max_{0\leq j\leq n-1} P(S_{n-1}=j)$$

(see Lemma 4) and that, if $h_1 \neq 0$

$$\left|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\right| \leq \frac{1}{|\sin(\pi h_1\alpha)|} \sum_{m=n}^{+\infty} q_m q_{m+1} P(S_{m-1}=n-1)$$

(see Section 2.2 above). By Lemma 2,

$$\max_{0\leq j\leq n-1} P(S_{n-1}=j) = \mathcal{O}\left((q_1+\cdots+q_{n-1})^{-1/2}\right)$$

and

$$P(S_{m-1}=n-1) = \mathcal{O}\left((q_1+\cdots+q_{m-1})^{-1/2}\right).$$

Moreover $(q_1+\cdots+q_n) \sim (1-\beta)n^{1-\beta}$ when $q_n = 1/n^\beta$ with $0<\beta<1$ and $(q_1+\cdots+q_n) \sim \log n$ when $q_n = 1/n$.

Set $q_n = 1/n^\beta$ with $0<\beta<1$. We get

$$\sup_{h\neq 0} |\sin(\pi h\alpha)||\mathbb{E}(e(h\alpha Y_n))| = \mathcal{O}\left(1/n^{(\beta+1)/2}\right)$$

and

$$\begin{aligned}\sup_{h_1\neq 0} |\sin(\pi h_1\alpha)|\left|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\right| &= \mathcal{O}\left(\sum_{m=n}^{+\infty} m^{-2\beta} m^{(\beta-1)/2}\right)\\ &= \mathcal{O}\left(1/n^{(3\beta-1)/2}\right).\end{aligned}$$

Set now $q_n = 1/n$. We get

$$\sup_{h\neq 0} |\sin(\pi h\alpha)||\mathbb{E}(e(h\alpha Y_n))| = \mathcal{O}\left(1/n(\log n)^{-1/2}\right) = \mathcal{O}\left(1/n\right)$$

and

$$\begin{aligned}\sup_{h_1\neq 0} |\sin(\pi h_1\alpha)|\left|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\right| &= \mathcal{O}\left(\sum_{m=n}^{+\infty} m^{-2}(\log m)^{-1/2}\right)\\ &= \mathcal{O}\left(\sum_{m=n}^{+\infty} m^{-2}\right) = \mathcal{O}\left(1/n\right).\end{aligned}$$

□

In connection with Lemma 6, the following lemma explains the influence of $\eta(\alpha)$ on the convergence rates. It is proved in [9, p. 123].

Lemma 7. *We have*

$$\sum_{h=1}^{M} \frac{1}{h|\sin(\pi h\alpha)|} = \mathcal{O}(M^{\eta-1})$$

for all $\eta > \eta(\alpha)$.

3.2 Proof of Theorem 2

Set $q_n = 1/n^\beta$ with $0 < \beta \le 1$. Recall the notation $H = \max(|h|, |h_1|)$ and $R = \max(|h|, 1)\max(1, |h_1|)$ and that $(h, h_1) \neq (0, 0)$ and note that $R = |h|$ when $h_1 = 0$, $R = |h_1|$ when $h = 0$ and $R = |h||h_1|$ when $h \neq 0$ and $h_1 \neq 0$. For all pairs (M, n) of positive integers,

$$\sum_{\substack{0<H\le M\\ h_1=0}} R^{-1}\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| = 2\sum_{h=1}^{M} |h|^{-1}\big|\mathbb{E}(e(h\alpha Y_n))\big|$$

because $h \neq 0$ when $h_1 = 0$ and $\big|\mathbb{E}(e(h\alpha Y_n))\big| = \big|\mathbb{E}(e(-h\alpha Y_n))\big|$.

By Lemma 6 we get, for all positive integers M,

$$\sum_{\substack{0<H\le M\\ h_1=0}} R^{-1}\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| = 2\left(\sum_{h=1}^{M} \frac{1}{h|\sin(\pi h\alpha)|}\right)\mathcal{O}\left(\frac{1}{n^{(\beta+1)/2}}\right) \tag{5}$$

and

$$\sum_{\substack{0<H\le M\\ h_1\neq 0}} R^{-1}\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| = \left(\sum_{\substack{0<H\le M\\ h_1\neq 0}} \frac{1}{R|\sin(\pi h_1\alpha)|}\right)\mathcal{O}\left(\frac{1}{n^{(3\beta-1)/2}}\right) \tag{6}$$

where the constants implied by $\mathcal{O}$ are independent of M. Now

$$\begin{aligned}\sum_{\substack{0<H\le M\\ h_1\neq 0}} \frac{1}{R|\sin(\pi h_1\alpha)|} &= \sum_{0<|h_1|\le M} \frac{1}{|\sin(\pi h_1\alpha)|}\sum_{0\le|h|\le M}\frac{1}{R}\\ &= \left(1 + 2\left(1 + \frac{1}{2} + \cdots + \frac{1}{M}\right)\right)\sum_{0<|h_1|\le M}\frac{1}{|h_1||\sin(\pi h_1\alpha)|}\\ &\le \mathcal{O}(\log M)\sum_{h_1=1}^{M}\frac{1}{h_1|\sin(\pi h_1\alpha)|}. \end{aligned}\tag{7}$$

Fix $\eta > \eta(\alpha)$ and set $\eta' = (\eta(\alpha)+\eta)/2$. There exists a constant C_1, independent of M and n, such that

$$\sum_{\substack{0<H\le M\\ h_1=0}} R^{-1}\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| \le C_1\frac{M^{\eta'-1}}{n^{(3\beta-1)/2}}$$

(see (5) and Lemma 7 and note that $(\beta+1)/2 \ge (3\beta-1)/2$ when $\beta \le 1$) and

$$\sum_{\substack{0<H\le M\\ h_1\neq 0}} R^{-1}\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| \le C_1\frac{M^{\eta'-1}\log M}{n^{(3\beta-1)/2}}$$

(see (6), (7) and Lemma 7). Therefore

$$\sum_{0<H\le M} R^{-1}\big|\mathbb{E}\big(e(h\alpha Y_n + h_1\alpha Y_{n+1})\big)\big| \le 2C_1\frac{M^{\eta'-1}\log M}{n^{(3\beta-1)/2}}. \tag{8}$$

As a consequence of (8) and Lemma 5, we get

$$\Delta(Q_n, U^2) \leq C \left(\frac{1}{M} + \frac{M^{\eta'-1} \log M}{n^{(3\beta-1)/2}} \right)$$

where the constant C is independent of M and n. Now we replace M by the integer part of $n^{(3\beta-1)/2\eta'}$. In this case,

$$\frac{M^{\eta'-1}}{n^{(3\beta-1)/2}} \sim \frac{1}{n^{(3\beta-1)/2\eta'}} \quad \text{and} \quad \log M \sim \frac{3\beta-1}{2\eta'} \log n$$

and we get

$$\Delta(Q_n, U^2) = \mathcal{O}\left(\frac{\log n}{n^{(3\beta-1)/2\eta'}} \right).$$

Theorem 2 is proved since $\eta > \eta'$.

4 Application to Benford law

All along the present section, a and b are two integers greater than 1 such that $\log_b a$ is irrational, that is to say such that a is not a rational power of b. We consider the sequence $(a^{Y_n})_n$ which is the random subsequence of $(a^n)_n$ generated by our Bernoulli trials. More precisely, we are interested in $\mathcal{M}_b(a^{Y_n})$ where $\mathcal{M}_b(x)$ designates the mantissa of the positive real x in the numeration base b, that is to say the unique number in $[1, b)$ such that $x = \mathcal{M}_b(x) b^m$ for some integer m.

The investigations on fractional parts and on mantissae are closely related. Indeed, for any positive real number x, $\{\log_b x\} = \log_b(\mathcal{M}_b(x))$ and so

$$(\mathcal{M}_b(x) < t) \iff (\{\log_b x\} < \log_b t) \qquad (t \in [1, b)). \tag{9}$$

In particular, if Z is a positive random variable, $\{\log_b Z\}$ is distributed following U^1 if and only if $\mathcal{M}_b(Z)$ is distributed following the Benford law in base b, denoted B_b here and defined by $B_b([1, t)) = \log_b t$ for $t \in [1, b)$. Moreover, for positive random variables Z_n, the law of $\mathcal{M}_b(Z_n)$ converges weakly to B_b as $n \to +\infty$ if and only if the law of $\{\log_b Z_n\}$ converges weakly to U^1 and, for any positive integer k, the law of $(\mathcal{M}_b(Z_n), \ldots, \mathcal{M}_b(Z_{n+k}))$ converges weakly to the product probability measure B_b^{k+1} as $n \to +\infty$ if and only if the law of $(\{\log_b Z_n\}, \ldots, \{\log_b Z_{n+k}\})$ converges weakly to U^{k+1}.

As already said above, the u.p.m. on the set $\{\{\alpha n\} : n = 1, \ldots, N\}$ converges weakly to U^1 as $N \to +\infty$ when α is irrational. Thus what precedes shows that the u.p.m. on the set $\{\mathcal{M}_b(a^n) : n = 1, \ldots, N\}$ converges weakly to B_b as $N \to +\infty$ since $\log_b a$ is supposed irrational. That is why the sequence $(a^n)_n$ is said to be a Benford sequence in base b when $\log_b a$ is irrational. Among many others [10], the sequences $(n!)_n$ and $(n^n)_n$ are Benford in any base b. The main property of the Benford sequences is the so-called *firts digit phenomenon*: 1 appears as the leading significant digit about 30.1 percent of the time (instead of 11.1 percent as it might seem at first glance), 2 about 17.6 percent of the time, and so on. See for example [1] and [6, Appendix] for relevant background on Benford law.

If we are given the value of $\mathcal{M}_b(a^n)$ (but not the value of n), then we know precisely the value of $\mathcal{M}_b(a^{n+1})$ because $\mathcal{M}_b(xy) = \mathcal{M}_b\big(\mathcal{M}_b(x)\mathcal{M}_b(y)\big)$. We will see

in particular that, by contrast, the random variables $\mathcal{M}_b(a^{Y_n})$ and $\mathcal{M}_b(a^{Y_{n+1}})$ are asymptotically independent when Condition (1) is satisfied.

The two following corollaries are direct consequences of what precedes and of Proposition 1, Theorem 1 and Corollary 1.

Corollary 2. *Fix $q \in (0,1)$ and suppose that $q_n = q$ for all n and that $\log_b a$ is irrational. Then the law of $\big(\mathcal{M}_b(a^{Y_n}), \mathcal{M}_b(a^{Y_{n+1}})\big)$ does not converge weakly to B_b^2 as $n \to +\infty$ and thus $\mathcal{M}_b(a^{Y_n})$ and $\mathcal{M}_b(a^{Y_{n+1}})$ are not asymptotically independent.*

Corollary 3. *Suppose that the q_n decrease to zero and sum to $+\infty$ and that $\log_b a$ is irrational. If condition (1) is satisfied, then, for any positive integer k, the law of $\big(\mathcal{M}_b(a^{Y_n}), \ldots, \mathcal{M}_b(a^{Y_{n+k}})\big)$ converges weakly to B_b^{k+1} as $n \to +\infty$ and thus $\mathcal{M}_b(a^{Y_n})$, $\ldots$, $\mathcal{M}_b(a^{Y_{n+k}})$ are asymptotically independent.*

This is the case in particular if $q_n = 1/n^\beta$ with $1/3 < \beta < 1$ and if $q_n = 1/n$ or $q_n = 1/n \log n$ and so on.

4.1 Rates of convergence

In the present section, b is such that $\log_b 2$ is irrational, that is to say: b is not a power of 2. Let F_b denotes the set of prime factors of b. Then $\eta(\log_b 2) \leq 7.62$ when $F_b \subset \{2,3\}$, $\eta(\log_b 2) \leq 15.28$ when $F_b \subset \{2,3,5\}$ and $\eta(\log_b 2) \leq 256.87$ when $F_b \subset \{2,3,5,7\}$ (see [11]).

Recall that Q_n designates the law of $(\{\alpha Y_n\}, \{\alpha Y_{n+1}\})$ and that

$$\Delta(Q_n, U^2) := \sup_{(s,t)\in[0,1)^2} |Q_n\big([0,s) \times [0,t)\big) - st|.$$

Let Q_n^* denotes the law of $\big(\mathcal{M}_b(2^{Y_n}), \mathcal{M}_b(2^{Y_{n+1}})\big)$ and set

$$\Delta(Q_n^*, B_b^2) = \sup_{(u,v)\in[1,b)^2} |Q_n^*\big([1,u) \times [1,v)\big) - \log_b u \log_b v|.$$

Note that, by (9),

$$\Delta(Q_n, U^2) = \Delta(Q_n^*, B_b^2) \tag{10}$$

if $\alpha = \log_b 2$.

The following corollary is a direct consequence of what precedes and of Theorem 2.

Corollary 4. *Set $q_n = 1/n^\beta$ with $1/3 < \beta \leq 1$. Then*

$$\Delta(Q_n^*, B_b^2) = \mathcal{O}\left(\frac{\log n}{n^{(3\beta-1)/16}}\right)$$

if $F_b \subset \{2,3\}$,

$$\Delta(Q_n^*, B_b^2) = \mathcal{O}\left(\frac{\log n}{n^{(3\beta-1)/31}}\right)$$

$F_b \subset \{2,3,5\}$ and

$$\Delta(Q_n^*, B_b^2) = \mathcal{O}\left(\frac{\log n}{n^{(3\beta-1)/514}}\right)$$

$F_b \subset \{2,3,5,7\}$.

For example, if $b = 10$ and $q_n = 1/\sqrt{n}$, then $\Delta(Q_n^*, B_{10}^2) = \mathcal{O}\left(\frac{\log n}{n^{1/62}}\right)$ and, if $b = 10$ and $q_n = 1/n$, then $\Delta(Q_n^*, B_{10}^2) = \mathcal{O}\left(\frac{\log n}{n^{1/31}}\right)$.

5 Concluding remark

Here $q_n = 1/n^\beta$ with $0 < \beta < 1$. It is possible to evaluate the almost sure order of magnitude of Y_n by using Proposition 6.2 in [8] which states that S_n is almost surely equivalent to $q_1 + \cdots + q_n$ whenever $\sum q_n = +\infty$. The sequence (Y_n) being a (random) subsequence of (n), $n = S_{Y_n}$ is almost surely equivalent to $q_1 + \cdots + q_{Y_n}$. In particular $(1-\beta)^{-1} Y_n^{1-\beta}$ is almost surely equivalent to n. So we can say that Y_n is almost surely of the same order of magnitude than $n^{(1-\beta)^{-1}}$.

On the other hand, it is proved in [5] that, if $u_n = \{\alpha n^\gamma\}$, the u.p.m. on the set $\{(u_n, \ldots, u_{n+k}) : n = 1, 2, \ldots, N\}$ converges weakly to U^{k+1} when $\gamma > k$ *but not when* $\gamma \leq k$. That why we have been surprised to find that Condition (1) ensures the asymptotic independence of $\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\}$ whatever is the value of k. In view of the results in [5] and the above evaluation, we expected something like: $\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\}$ are asymptoticly independent if $(1-\beta)^{-1} > k$ but not if $(1-\beta)^{-1} \leq k-1$.

However the comparison between the two situations would be more pertinent if we have investigated the weak convergence as $N \to +\infty$ of the u.p.m. on the set

$$\{(\{\alpha Y_n(\omega)\}, \ldots, \{\alpha Y_{n+k}(\omega)\}) : n = 1, 2, \ldots, N\}$$

instead of the distribution of the possible values of $(\{\alpha Y_n\}, \ldots, \{\alpha Y_{n+k}\})$ for fixed n.

References

[1] Berger A. and Hill T.P. (2011), *A basic theory of Benford's Law*, Probab. Surv., 8, 1–126.

[2] Berkes I. and Borda B. (2018), *Berry-Esseen bounds and Diophantine approximation*, Anal. Math., 44, 149–161.

[3] Billingsley P., Probability and Measure, Wiley, New-York, 1979.

[4] Bugeaud Y. (2018), *Linear Forms in Logarithmes and applications*, IRMA Lect. Math. Theor. Phys. 28, Eur. Math. Soc., Zürich.

[5] Cai Z. et al (2019), *A local Benford law for a class or arithmetic sequences*, Int. J. Number Theory, 15-3, 613-638.

[6] Chenavier N., Massé B. and Schneider D. (2018), *Products of random variables and the first digit phenomenon*, Stoch. Process. Appl., 128, 1615–1634.

[7] Cramér H. (1935), *Prime numbers and probabilities*, Skand. Math. Kongr., 8, 107–115.

[8] Fan A. and Schneider D. (2010), *Recurrence properties of sequences of integers*, Sci. China Math., 53-3, 1–16.

[9] Kuipers L. and Niederreiter H. (2006), *Uniform Distribution of Sequences*, Dover Publications New-York.

[10] Massé B. and Schneider D. (2015), *Fast growing sequences of numbers and the first digit phenomenon*, Int. J. Number Theory, 11-3, 705?719.

[11] Massé B. (2021), *Random walks on the circle and measure of irrationality*, https://hal.archives-ouvertes.fr/hal-03274061.

[12] Niederreiter H. and Philipp W. (1973), *Berry-Esseen bounds and a theorem of Erdős and Turán on uniform distribution mod 1*, Duke Math. J., 40, 633–649.

[13] Petrov V. (1975), *Sums of Independent Random Variables*, Springer-Verlag.

[14] Ribenboim P. (2000), *My Numbers, My Friends : Popular Lectures on Number Theory*, Springer-Verlag.

[15] Su F. E. (1998), *Convergence of random walks on the circle generated by an irrational rotation*, Trans. Amer. Math. Soc., 350-9, 3717–3741.