

# Uniform distribution mod 1 for sequences of ergodic sums and continued fractions\*

Albert M. Fisher<sup>†</sup>      Xuan Zhang<sup>‡</sup>

## Abstract

We establish a coboundary condition for a sequence of ergodic sums (i.e. Birkhoff partial sums) to be almost surely uniformly distributed mod 1. Applications are given when the sequence is generated by a Gibbs-Markov map. In particular, we show that for almost every real number, the sequence of denominators of the convergents of its continued fraction expansion satisfies Benford's law.

**Keywords:** Benford law; uniform distribution mod 1; sequence of ergodic sums; continued fractions; Gibbs-Markov map.

**MSC2020 subject classifications:** 11K50; 11K0; 37A50.

Submitted to ECP on July 3, 2024, final version accepted on February 6, 2025.

## 1 Introduction

For  $x \in (0, 1)$  irrational, denote by  $[a_1, a_2, \dots]$  its continued fraction expansion and write  $p_n/q_n$  for the  $n$ -th convergent  $[a_1, \dots, a_n] = 1/(a_1 + \dots + 1/a_n)$ . The sequence  $(\log q_n(x))_{n \in \mathbb{N}}$  behaves much like a sum of i.i.d. random variables in that it obeys a range of classical limit theorems ([14]), in this note we are interested in uniform distribution (u.d.) mod 1. Recalling the definition, a sequence  $(x_n)$  of real numbers is said to be u.d. mod 1 if and only if for every  $t \in [0, 1)$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N : \{x_n\} \leq t\} = t,$$

where  $\{x_n\}$  represents the fractional part of  $x_n$ . Equivalently, on the circle  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  with Lebesgue measure  $dt$ , then for every continuous function  $f \in C(\mathbb{T})$  we have:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f(x_n) = \int_{\mathbb{T}} f(t) dt.$$

Uniform distribution mod 1 is closely related to *Benford's law* or *the first-digit law*, that is, the proportion of entries with the first digit  $d$ ,  $d \in \{1, \dots, 9\}$ , tends to  $\lg(1 + 1/d)$ . Here  $\lg = \log_{10}$  is the decimal logarithm. First noted by Benford as well as Newcomb, this specific distribution has been observed to occur in a variety of social science statistics and data sets as well as in many number-theoretical experiments. We refer to [4] for

---

\*Partially supported by FAPESP grants #2016/25053-8 and #2018/15088-4.

<sup>†</sup>Universidade de São Paulo, Brazil. E-mail: afisher@ime.usp.br

<sup>‡</sup>Universidade Federal Fluminense, Brazil. E-mail: xuanz@id.uff.br

a comprehensive survey of Benford’s law. For a nonzero real number  $x$ , let  $s(x)$  be its decimal significand, which is the unique number  $t \in [1, 10)$  such that  $|x| = 10^k t$  holds for some  $k \in \mathbb{Z}$ . A sequence of real numbers  $(x_n)_{n \in \mathbb{N}}$  is said to be a *Benford sequence* if the corresponding sequence of decimal significands  $(s(x_n))_{n \in \mathbb{N}}$  is distributed in  $[1, 10)$  with density  $\lg$ , that is,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{1 \leq n \leq N : s(x_n) \leq t\} = \lg(t), \quad \forall t \in [1, 10).$$

Many sequences are known to be Benford; examples are  $(2^n)_{n \in \mathbb{N}}$  and the Fibonacci sequence. The sequences of all natural numbers and of the prime numbers are not Benford, but are *weakly Benford* (with the usual Cesàro density being replaced by logarithmic density). Benford’s law is related to uniform distribution mod 1 via the following equivalent condition (see e.g. [4, Theorem 4.2]):

$$(x_n)_{n \in \mathbb{N}} \text{ is a Benford sequence if and only if } (\lg x_n)_{n \in \mathbb{N}} \text{ is u.d. mod 1.}$$

Hence the question under consideration can be rephrased by whether  $(q_n(x))_{n \in \mathbb{N}}$  is a Benford sequence.

When  $x$  is a quadratic irrational, several authors at about the same time ([10, 11, 17]) showed that  $(q_n(x))_{n \in \mathbb{N}}$  is indeed a Benford sequence. In particular this reconfirms that the Fibonacci sequence, being equal to  $(q_n(\frac{1+\sqrt{5}}{2}))_{n \in \mathbb{N}}$ , is Benford. Jager (as mentioned in [11]) asked whether  $(q_n(x))_{n \in \mathbb{N}}$  is Benford for almost every  $x$ . In this note we answer this question affirmatively.

**Theorem 1.1.**  *$(q_n(x))_{n \in \mathbb{N}}$  is a Benford sequence for almost every  $x$ . Equivalently,  $(\lg q_n(x))_{n \in \mathbb{N}}$  is u.d. mod 1 for almost every  $x$ .*

We note that Schatte in [18] presented a proof by way of a direct analysis of the continued fractions. Some of his estimates were illustrated via examples, rather than being given full proofs valid for the general cases. Here we present a new, detailed proof, based on ergodic theory methods. We believe this more general and hopefully more transparent approach will lead to new insights on this and other related questions. For one example, it follows immediately from our proof that for any constant  $\rho \in \mathbb{R}$  the sequence  $(\lg q_n(x) + n\rho)_{n \in \mathbb{N}}$  is also u.d. mod 1 for almost every  $x$ ; the same holds for the sequences  $(a_1 + \dots + a_n)_{n \in \mathbb{N}}$  and  $(\lg(a_1 \dots a_n))_{n \in \mathbb{N}}$ .

In what follows will use the natural logarithm  $\log$  in place of  $\lg$ ; it will be clear that the proofs work for both.

The almost sure behavior of the sequence  $(\log q_n(x))_{n \in \mathbb{N}}$  is somewhat different from the situation where  $x$  is a quadratic irrational, which equivalently has an eventually periodic continued fraction expansion. In this case, as shown by Jager-Liardet ([10]),  $\log q_n(x) = \frac{n}{l(x)}\alpha(x) + c_n(x) + o(1)$  for large  $n$ , where  $\alpha(x)$  is an irrational number,  $l(x)$  is the period in the continued fraction expansion of  $x$ , and  $c_n = c_{n+l}$  takes a finite number of values periodically. One can then deduce uniform distribution mod 1 of the sequence  $(\log q_n(x))_{n \in \mathbb{N}}$  from that of the irrational rotation  $(n\alpha(x))_{n \in \mathbb{N}}$ , the classical Weyl’s Theorem. In the almost sure case,  $\lim_{n \rightarrow \infty} \log q_n(x)/n$  converges to the Lévy constant  $\frac{\pi^2}{12 \log 2}$  and the rate of convergence is of the order  $O(\sqrt{n^{-1} \log \log n})$  by the law of iterated logarithm of Philipp-Stackelberg ([15]). However we do not know if the Lévy constant is irrational; moreover this rate of convergence is too weak to show almost sure uniform distribution of  $(\log q_n(x))_{n \in \mathbb{N}}$  as a direct consequence of that for a rotation; an explanation is that this gives  $\limsup$  rather than the sought-for time average information. Nevertheless, one might still guess that  $(\log q_n(x))_{n \in \mathbb{N}}$  is u.d. mod 1 for almost every  $x$ , given that the statistics of  $\log q_n(x)$  resembles well that of a partial sum of i.i.d. random variables and that, as shown by Robbins ([16]), the sequence of partial

sums of i.i.d. random variables (if not supported on a rational lattice) is almost surely uniformly distributed (a.s.u.d.) mod 1.

The main idea of our proof is to approximate the sequence  $(\log q_n(x))_{n \in \mathbb{N}}$  fast enough by a series of sequences of ergodic sums (in other words, sequences of partial sums of stationary processes), each of which is a.s.u.d. mod 1 (Lemma 3.4). Sequences of partial sums of stationary processes were studied by Holewijn ([9]) and more recently by Chenavier, Massé and Schneider ([6]) where they found a necessary and sufficient condition for such a sequence to be a.s.u.d. mod 1. We use a classical argument of Furstenberg to find a condition in terms of coboundaries for sequences of the partial sums generated by ergodic maps (Theorem 2.1). When the sequence is generated by a Gibbs-Markov map we give examples where the coboundary condition is verified (Theorem 3.2), and these provide the approximating sequences to  $(\log q_n)$ .

## 2 Sequence of ergodic sums

Let  $(X, \mu)$  be a standard probability space,  $T : X \rightarrow X$  be a  $\mu$ -preserving transformation, and  $f : X \rightarrow \mathbb{R}$  be a measurable function (also called here an *observable*). For any  $n \in \mathbb{N}$ ,  $x \in X$ , form the ergodic sum  $\mathcal{S}_n f(x) = f(x) + \dots + f(T^{n-1}x)$ .

Assuming ergodicity of  $T$ , we apply an argument of Furstenberg ([7, Lemma 2.1]; see also [3, Theorem 4.4]) to show that  $\mathcal{S}_n f(x)$  is almost surely uniformly distributed (a.s.u.d.) mod 1 under a coboundary condition for  $f$ . Note that we do not assume the compactness of  $X$  as done in [7].

**Theorem 2.1.** *Let  $T$  be an ergodic  $\mu$ -preserving transformation and  $f : X \rightarrow \mathbb{R}$  measurable. Suppose for all integers  $k \neq 0$ , the equation  $e^{2\pi i k f(x)} = g(x)/g(Tx)$  (a.s.) has no solution  $g \in L^2(X)$ . Then  $\mathcal{S}_n f(x)$  is a.s.u.d. mod 1.*

*Proof.* We write  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$  with the Lebesgue measure  $dt$ . Let  $Y = X \times \mathbb{T}$  and define the skew product  $T_f : Y \rightarrow Y$  by  $T_f(x, t) = (Tx, t + f(x))$ . Then  $T_f^n(x, t) = (T^n x, t + \mathcal{S}_n f(x))$ . Now firstly  $\lambda = \mu \times dt$  is an invariant measure for  $T_f$ ; this is because for any  $F \in L^\infty(Y)$

$$\begin{aligned} \int_Y F \circ T_f d\lambda &= \iint_Y F(Tx, t + f(x)) dt d\mu = \iint_Y F(Tx, t) dt d\mu \\ &= \iint_Y F(x, t) dt d\mu = \int_Y F d\lambda \end{aligned}$$

by the invariance of  $dt$  under translation and the invariance of  $\mu$  under  $T$ . Moreover the skew product transformation is ergodic; we follow Furstenberg's proof of this part. If not, then there is a non-constant invariant function  $F \in L^2_\lambda(Y)$  for  $T_f$ , that is,  $F \circ T_f = F$  almost everywhere. Now for almost every  $x$ ,  $F(x, \cdot) \in L^2(\mathbb{T})$ . We write its Fourier expansion,  $F(x, t) = \sum_{n \in \mathbb{Z}} a_n(x) e^{2\pi i n t}$ , with  $a_n \in L^2_\mu(X)$ . Then

$$F \circ T_f(x, t) = \sum_{n \in \mathbb{Z}} a_n(Tx) e^{2\pi i n f(x)} e^{2\pi i n t}.$$

$F \circ T_f = F$  implies  $a_n(x) = a_n(Tx) e^{2\pi i n f(x)}$ , hence  $|a_n(x)| = |a_n(Tx)|$ , for every  $n \in \mathbb{Z}$  and almost every  $x$ . The ergodicity of  $T$  implies that  $|a_n(x)|$  is a.s. a constant for every  $n \in \mathbb{Z}$ . Since  $F$  is not constant, one of  $|a_n(x)|$  with  $n \neq 0$  is not 0. Say for  $k \neq 0$ ,  $|a_k(x)|$  is a nonzero constant. Then  $e^{2\pi i k f(x)} = a_k(x)/a_k(Tx)$  a.s., a contradiction to the hypothesis, proving ergodicity of  $T_f$ .

Next we show how u.d. in  $\mathbb{T}$  follows from this. Let  $h \in C(\mathbb{T})$ . We extend this to  $H : Y \rightarrow \mathbb{R}$  by  $H(x, t) = h(t)$ . This is a measurable function which is continuous on the fiber over  $x$ , for each  $x$ . Note that  $H \in L^1_\lambda(Y)$ .

The Birkhoff ergodic theorem implies that for every such  $H$ , hence for each continuous function  $h \in C(\mathbb{T})$ , there is a full measure set  $E_h$  such that for every  $(x, t) \in E_h \subseteq Y$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} H \circ T_f^n(x, t) = \int_Y H d\lambda = \int_{\mathbb{T}} h dt. \tag{2.1}$$

Since  $\mathbb{T}$  is compact,  $C(\mathbb{T})$  is separable. Suppose  $\{h_i\}_{i \in \mathbb{N}}$  is a dense subset of  $C(\mathbb{T})$ , then (2.1) holds for every  $h_i$  and all  $(x, t)$  in  $\cap_{i \in \mathbb{N}} E_{h_i}$ . This passes over to every  $h \in C(\mathbb{T})$ : we approximate  $h$   $\varepsilon$ -uniformly by functions  $h_i$ , and both sides of (2.1) are within  $\varepsilon$ .

In conclusion,  $\lambda$ -almost every  $(x, t)$  is a *fiber generic point*, that is, (2.1) holds for every  $h \in C(\mathbb{T})$ .

Suppose  $(x, t)$  is a fiber generic point, then  $(x, t + s)$  is also a fiber generic point for any  $s \in \mathbb{T}$ . This is because, for every  $h \in C(\mathbb{T})$  defining  $h_s(\cdot) = h(\cdot + s) \in C(\mathbb{T})$  and  $H_s(x, t) = h_s(t)$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} H \circ T_f^n(x, t + s) = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} H_s \circ T_f^n(x, t) = \int_{\mathbb{T}} h_s dt = \int_{\mathbb{T}} h dt.$$

So the set of all fiber generic points is  $E \times \mathbb{T}$  for some  $E \subset X$ . Since this set has full  $\lambda$ -measure,  $\mu(E) = 1$ . Now for every  $x \in E$   $(x, 0)$  is a fiber generic point. This means that for every  $h \in C(\mathbb{T})$ ,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=0}^{N-1} h(\mathcal{S}_n f(x)) = \int_{\mathbb{T}} h dt,$$

equivalently  $\mathcal{S}_n f(x)$  is u.d. mod 1 for every  $x \in E$ . □

The coboundary condition is close to being a necessary and sufficient condition. Recall Weyl’s criterion (e.g. [13]):  $(x_n)$  is u.d. mod 1 if and only if for every integer  $k \neq 0$

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i k x_n} = 0.$$

**Proposition 2.2.** *Let  $T$  be an ergodic  $\mu$ -preserving transformation and  $f : X \rightarrow \mathbb{R}$  measurable. Then if  $\mathcal{S}_n f(x)$  is a.s.u.d. mod 1, the equation  $e^{2\pi i k f(x)} = g(x)/g(Tx)$  (a.s.) has no solution  $g \in L^1$  with  $\int g d\mu \neq 0$ , for each integer  $k \neq 0$ .*

*Proof.* Suppose for some integer  $k \neq 0$  and  $g \in L^1$ ,  $e^{2\pi i k f(x)} = g(x)/g(Tx)$  a.s. Then  $|g| = |g \circ T|$ . We can assume that  $|g|=1$  by ergodicity of  $T$ . So

$$\frac{1}{N} \sum_{n=1}^N e^{2\pi i k \mathcal{S}_n f(x)} = \frac{1}{N} \sum_{n=1}^N g(x) \bar{g}(T^n x) \rightarrow g(x) \int \bar{g} d\mu \text{ a.s.}$$

by ergodic theorem. It follows from Weyl’s criterion that if  $\mathcal{S}_n f(x)$  is a.s.u.d. mod 1 then  $\int g d\mu = 0$ . □

For circle-valued skew products the coboundary condition is called *aperiodicity* in [1]. This terminology makes sense in view of [3, Proposition 4.3]. However that fact for the skew product does not directly imply the uniform distribution, which is why we give this argument.

**2.1 Gibbs-Markov maps**

When  $T$  is a Gibbs-Markov map, we give some sufficient conditions for the coboundary condition to hold that are easier to verify. First recall the definition of a Gibbs-Markov map from [2, 1]. Let  $(\Omega, \mathcal{B}, \mu)$  be a probability space and let  $T$  be a non-singular transformation as defined in ergodic theory; this weakens the usual condition of measure-preserving, to being measure-class-preserving, i.e. preserving sets of measure zero: that is for  $A \in \mathcal{B}$ ,  $\mu(T^{-1}A) = 0$  if and only if  $\mu(A) = 0$ . In the setting of Aaronson, Denker and Urbanski, one generally studies maps which are non-invertible with a countable number of branches. The map restricted to each branch is non-singular while it may or may not be measure-preserving when taken as a whole.

Consider a countable partition  $\alpha$  of  $\Omega \pmod{\mu}$ . Denote by  $\alpha_0^{n-1}$  the refined partition  $\bigvee_{i=0}^{n-1} T^{-i}\alpha$  and by  $\sigma(\cdot)$  the generated  $\sigma$ -algebra.

**Definition 2.3.**  $T$  is called a Gibbs-Markov map if the following conditions are satisfied.

- (a)  $\alpha$  is a strong generator of  $\mathcal{B}$  under  $T$ , i.e.  $\sigma(\{T^{-n}\alpha : n \in \mathbb{N} \cup \{0\}\}) = \mathcal{B} \pmod{\mu}$ .
- (b)  $\inf_{a \in \alpha} \mu(Ta) > 0$ .
- (c) For every  $a \in \alpha$ ,  $Ta \in \sigma(\alpha) \pmod{\mu}$ , moreover the restriction  $T|_a$  is invertible and non-singular.
- (d) For every  $n \in \mathbb{N}$  and  $a \in \alpha_0^{n-1}$ , denote the non-singular inverse branch of  $T^{-n}$  on  $T^n a$  by  $v_a : T^n a \rightarrow a$ . We denote the Radon-Nikodym derivative  $\frac{d\mu \circ v_a}{d\mu}$  by  $v'_a$ . There exist  $r \in (0, 1)$  and  $M > 0$  such that for any  $n \in \mathbb{N}$ ,  $a \in \alpha_0^{n-1}$  and  $x, y \in T^n a$  a.e.

$$\left| \frac{v'_a(x)}{v'_a(y)} - 1 \right| \leq M \cdot r(x, y),$$

where  $r(x, y)$  is the metric  $r^{m(x,y)}$  with

$$m(x, y) = \min\{n \in \mathbb{N} : T^{n-1}(x) \text{ and } T^{n-1}(y) \text{ belong to different elements of } \alpha\}.$$

$T$  is said to be topologically mixing if for any  $a, b \in \alpha$  there is  $n_{a,b} \in \mathbb{N}$  such that for any  $n \geq n_{a,b}$ ,  $T^n a \supset b$ . Note that a topologically mixing and measure preserving Gibbs-Markov map is always ergodic ([1, Corollary]). A function  $f$  is said to be locally Lipschitz if

$$D_\alpha(f) = \sup_{a \in \alpha} \sup_{x, y \in a} \frac{|f(x) - f(y)|}{r(x, y)} < \infty,$$

and Lipschitz if

$$D(f) = \sup_{x, y \in \Omega} \frac{|f(x) - f(y)|}{r(x, y)} < \infty.$$

Let  $\beta$  be the finest partition such that  $\sigma(T\alpha) = \sigma(\beta)$ .

**Proposition 2.4.** Let  $T$  be a topologically mixing and  $\mu$ -preserving Gibbs-Markov map. Suppose  $e^{2\pi i k f(x)} = g(x)/g(Tx)$  a.s. for some integer  $k \neq 0$  and measurable  $g$ .

1. If  $f$  is  $\alpha$ -measurable, then  $g$  is  $\beta$ -measurable.
2. Suppose  $p$  is a fixed point of  $T$  and, with respect to the metric  $r$ ,  $p$  lies in the interior of a partition member and every open ball around  $p$  has positive measure. If  $p$  is a continuity point of both  $T$  and  $f$  and if  $f$  is locally Lipschitz, then  $e^{2\pi i k f(p)} = 1$ .

*Proof.* 1. This is [1, Theorem 3.1].

2. We can assume that  $|g|=1$  by ergodicity of  $T$ . According to [1, Corollary 2.2], the condition that  $f$  is locally Lipschitz implies that  $g$  is Lipschitz. For a fixed point  $p$  in the assumption, there is a sequence  $x_n$  tending to  $p$  such that  $|g(x_n) - g(Tx_n)| \leq D(g)r(x_n, Tx_n)$  and  $e^{2\pi i k f(x_n)} = g(x_n)/g(Tx_n)$ . Since both  $Tx_n$  and  $x_n$  tend to  $p$  and since  $f$  is continuous at  $p$ , it follows that  $e^{2\pi i k f(p)} = 1$ .  $\square$

### 3 Applications to continued fractions

We return to continued fractions and prove Theorem 1.1 in this section. Let  $\Omega = (0, 1)$ ,  $\mu$  be the Gauss measure  $d\mu = \frac{1}{\log 2} \frac{1}{1+x} dx$ , and  $T$  be the continued fraction map  $x \mapsto \{1/x\}$ . Denoting the continued fraction expansion of an irrational number  $x$  by  $[a_1(x), a_2(x), \dots]$ , then  $T$  is equal to the shift map  $[a_1, a_2, \dots] \mapsto [a_2, a_3, \dots]$ . Let  $\alpha$  be the partition into cylinder sets determined by  $a_1$ , i.e.  $\alpha = \{\{x : a_1(x) = n\} : n \in \mathbb{N}\}$ . Let  $r(x, y) = (1/2)^{\max\{n: a_j(x)=a_j(y), 1 \leq j \leq n\}}$ . The constant  $r = 1/2$  is chosen so that  $r^2 \cdot |(T^2)'| \geq 1$ . One can check that  $T$  is a topologically mixing,  $\mu$ -preserving Gibbs-Markov map ([1, Example 2]), so we can apply results from the previous section to the continued fraction map.

**Proposition 3.1.** *The equation  $e^{2\pi i k f(x)} = g(x)/g(Tx)$  has no solution for any integer  $k \neq 0$  and measurable  $g$  in each of the following cases.*

1.  $f(x) = a_1(x)$
2.  $f(x) = \log(a_1(x))$
3.  $f(x) = -\log(T^l x)$ , where  $l \in \mathbb{N} \cup \{0\}$ .

*Proof.* This follows immediately from Proposition 2.4, since

1.  $a_1(x)$  is  $\alpha$ -measurable and  $\beta = \{\Omega\}$ . For  $g$  to be  $\beta$ -measurable, it has to be constant, but  $e^{2\pi i k a_1}$  is not.
2. The same reason as item 1.
3. First let  $l = 0$ . Note that  $\log x$  is locally Lipschitz, because  $D_{[n]}(\log)$  has the order of  $\frac{1}{n}$ . Consider the fixed point  $p = \frac{\sqrt{5}+1}{2} = [1, 1, \dots]$ , which is in the interior of the cylinder set  $\{x : a_1(x) = 1\}$ . Both  $T$  and  $\log$  are continuous at  $p$ , but  $e^{-2\pi i k \log(p)} \neq 1$  for any integer  $k \neq 0$ . For  $l \in \mathbb{N}$ ,  $\log \circ T^l$  is locally Lipschitz, as  $D_{[n]}(\log \circ T^l)$  has the order of  $\frac{1}{r^l n}$ . But again  $e^{-2\pi i k \log(T^l p)} \neq 1$  for any integer  $k \neq 0$ .  $\square$

Consequently, applying Theorem 2.1,

**Theorem 3.2.** *With respect to  $\mu$ , hence also with respect to Lebesgue measure, each of the following sequences is a.s.u.d. mod 1.*

- (i)  $(a_1(x) + \dots + a_n(x))_{n \in \mathbb{N}}$
- (ii)  $(\log(a_1(x) \cdots a_n(x)))_{n \in \mathbb{N}}$
- (iii)  $(-\mathcal{S}_n \log(T^l x))_{n \in \mathbb{N}}$ , where  $l \in \mathbb{N} \cup \{0\}$ .

**Remark 3.3.** Item (ii) implies that  $(a_1(x) \cdots a_n(x))_{n \in \mathbb{N}}$  is a Benford sequence a.s.

Now we prepare for the proof of Theorem 1.1. Recalling that  $q_n$  denotes the denominator of the convergent  $[a_1, \dots, a_n]$ , to extend the previous theorem to sequences such as  $(\log q_n)$  that are not necessarily ergodic sums, we make use of a simple approximation lemma. Given a sequence  $(u_n(x))_{n \in \mathbb{N}}$  and a series of sequences  $(u_n^{(k)}(x))_{n \in \mathbb{N}}$ , where  $k \in \mathbb{N}$  and  $x \in \Omega$ ,

**Lemma 3.4.** *Suppose that for every  $k \in \mathbb{N}$  the sequence  $(u_n^{(k)}(x))_{n \in \mathbb{N}}$  is a.s.u.d. mod 1 and that*

$$\lim_{k \rightarrow \infty} \sup_{n > k} \|u_n^{(k)} - u_n\|_\infty = 0,$$

*then  $(u_n(x))_{n \in \mathbb{N}}$  is also a.s.u.d. mod 1.*

*Proof.* For every  $m \in \mathbb{N}$ , choose  $k = k(m)$  such that

$$2\pi \sup_{n > k} \|u_n^{(k)} - u_n\|_\infty < 1/m.$$

So there is a full-measure set  $\Omega'_m$  such that for every  $x \in \Omega'_m$ ,

$$2\pi \sup_{n > k} |u_n^{(k)}(x) - u_n(x)| < 1/m.$$

Again by assumption, there is a full-measure set  $\Omega_m$  such that for every  $x \in \Omega_m$ ,  $(u_n^{(k(m))}(x))_{n \in \mathbb{N}}$  is u.d. mod 1. Put  $\Omega_0 = \bigcap_{m \in \mathbb{N}} \Omega'_m \cap \Omega_m$ , which is a full-measure set. Let  $x \in \Omega_0$ . Since  $(u_n^{(k(m))}(x))_{n \in \mathbb{N}}$  is u.d. mod 1, in view of Weyl's criterion, choose  $N_0 > 2km$  such that for all  $N \geq N_0$

$$\frac{1}{N} \left| \sum_{n=1}^N e^{2\pi i u_n^{(k)}(x)} \right| \leq 1/m.$$

So for all  $N \geq N_0$

$$\begin{aligned} \frac{1}{N} \left| \sum_{n=1}^N e^{2\pi i u_n(x)} \right| &\leq \frac{1}{N} \left| \sum_{n=1}^N e^{2\pi i u_n^{(k)}(x)} \right| + \frac{1}{N} \sum_{n=1}^N |e^{2\pi i u_n^{(k)}(x)} - e^{2\pi i u_n(x)}| \\ &\leq \frac{1}{N} \left| \sum_{n=1}^N e^{2\pi i u_n^{(k)}(x)} \right| + \frac{2k}{N} + 2\pi \sup_{n > k} |u_n^{(k)}(x) - u_n(x)| \\ &\leq 3/m. \end{aligned}$$

Hence  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N e^{2\pi i u_n(x)} = 0$ , and  $(u_n(x))_{n \in \mathbb{N}}$  is u.d. mod 1. □

Now consider the difference  $\delta_n(x) = \log q_n(x) - (-\mathcal{S}_n \log(x))$ . It is well-known that  $\delta_n(x)$  is uniformly bounded; in fact (see e.g. [12]),

**Lemma 3.5.**

$$\delta_n(x) = -\log(1 + T^n x \cdot [a_n, \dots, a_1])$$

where  $[a_n, \dots, a_1]$  represents the finite continued fraction  $1/(a_n + \dots + 1/a_1)$ .

*Proof.* Because

$$x = \frac{p_{n-1}r_n + p_{n-2}}{q_{n-1}r_n + q_{n-2}},$$

where  $r_n = 1/(T^{n-1}x)$ , one finds

$$T^n x = \frac{xq_n - p_n}{p_{n-1} - xq_{n-1}}.$$

So

$$q_n \cdot x \cdots T^{n-1} x = (-1)^{n-1} q_n (xq_{n-1} - p_{n-1}).$$

Because  $p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}$  and  $\frac{q_{n-1}}{q_n} = [a_n, \dots, a_1]$ , it follows that

$$\begin{aligned} \delta_n &= \log(q_n \cdot x \cdots T^{n-1} x) \\ &= -\log \left( 1 + \frac{x - \frac{p_n}{q_n}}{\frac{p_{n-1}}{q_{n-1}} - x} \right) \\ &= -\log(1 + T^n x \cdot [a_n, \dots, a_1]). \end{aligned} \quad \square$$

Moreover, we can approximate  $(\delta_n(x))_{n \in \mathbb{N}}$  by a dynamical sequence up to a uniformly small error.

**Lemma 3.6.** For all  $k \in \mathbb{N}$ ,  $n > k$  and  $x \in (0, 1)$ ,

$$|\delta_k(T^{n-k}x) - \delta_n(x)| < -\log(1 - 2^{-k/2}).$$

*Proof.* Fix  $n \in \mathbb{N}$ . For every  $1 \leq k \leq n$ , let  $\tilde{p}_k/\tilde{q}_k = [a_n, \dots, a_{n-k+1}]$  be the  $k$ -th convergent of the finite continued fraction  $[a_n, \dots, a_1]$ ,  $\tilde{r}_k = 1/[a_{n-k+1}, \dots, a_1]$  be the  $k$ -th remainder and let  $\tilde{p}_0 = 0, \tilde{q}_0 = 1$ , then

$$[a_n, \dots, a_1] = \frac{\tilde{p}_n}{\tilde{q}_n} = \frac{\tilde{p}_k \tilde{r}_{k+1} + \tilde{p}_{k-1}}{\tilde{q}_k \tilde{r}_{k+1} + \tilde{q}_{k-1}}, \quad 1 \leq k < n.$$

One calculates

$$\begin{aligned} & \delta_k(T^{n-k}x) - \delta_n(x) \\ &= -\log(1 + T^n x \cdot [a_n, \dots, a_{n-k+1}]) + \log(1 + T^n x \cdot [a_n, \dots, a_1]) \\ &= -\log\left(1 + \frac{T^n x \cdot ([a_n, \dots, a_{n-k+1}] - [a_n, \dots, a_1])}{1 + T^n x \cdot [a_n, \dots, a_1]}\right) \\ &= -\log\left(1 + \frac{T^n x \cdot (\tilde{p}_k \tilde{q}_{k-1} - \tilde{p}_{k-1} \tilde{q}_k)}{\tilde{q}_k (\tilde{q}_k \tilde{r}_{k+1} + \tilde{q}_{k-1}) + T^n x \cdot \tilde{q}_k (\tilde{p}_k \tilde{r}_{k+1} + \tilde{p}_{k-1})}\right) \\ &= -\log\left(1 + \frac{T^n x \cdot (-1)^{k-1}}{\tilde{q}_k (\tilde{q}_k \tilde{r}_{k+1} + \tilde{q}_{k-1}) + T^n x \cdot \tilde{q}_k (\tilde{p}_k \tilde{r}_{k+1} + \tilde{p}_{k-1})}\right), \end{aligned}$$

hence the lemma follows from the estimate

$$\begin{aligned} & \left| \frac{T^n x \cdot (-1)^{k-1}}{\tilde{q}_k (\tilde{q}_k \tilde{r}_{k+1} + \tilde{q}_{k-1}) + T^n x \cdot \tilde{q}_k (\tilde{p}_k \tilde{r}_{k+1} + \tilde{p}_{k-1})} \right| \\ & < \frac{1}{\tilde{q}_k (\tilde{q}_k \tilde{r}_{k+1} + \tilde{q}_{k-1})} \leq \frac{1}{\tilde{q}_k (\tilde{q}_k a_{n-k} + \tilde{q}_{k-1})} = \frac{1}{\tilde{q}_k \tilde{q}_{k+1}} \leq 2^{-k/2}. \quad \square \end{aligned}$$

*Proof of Theorem 1.1.* Let  $h(x) = -\log(x)$ , and for every  $k \in \mathbb{N}$  let

$$h^{(k)} = h \circ T^k - \delta_k + \delta_k \circ T.$$

Note that for all  $n > k$ ,

$$\log q_n(x) = \mathcal{S}_n h + \delta_n = \mathcal{S}_{n-k} h^{(k)} + \mathcal{S}_k h + \delta_k - \delta_k \circ T^{n-k} + \delta_n.$$

In order to show that  $(\log q_n(x))_{n \in \mathbb{N}}$  is a.s.u.d. mod 1, applying Lemmas 3.4 and 3.6, it suffices to show that for every  $k$  the sequence

$$(\mathcal{S}_{n-k} h^{(k)}(x) + \mathcal{S}_k h(x) + \delta_k(x))_{n > k}$$

is a.s.u.d. mod 1. For every  $k$ , subtract  $\mathcal{S}_k h(x) + \delta_k(x)$  from every entry of the latter sequence. Since a translation does not change equidistribution, the equidistribution of the sequence  $(\mathcal{S}_{n-k} h^{(k)}(x))_{n > k} = (\mathcal{S}_n h^{(k)}(x))_{n \in \mathbb{N}}$  will complete the proof.

According to Proposition 3.1, the equation  $e^{2\pi i m h \circ T^k} = g/g \circ T$  has no solution for any integer  $m \neq 0$  and measurable  $g$ , so  $e^{2\pi i m h^{(k)}} = g/g \circ T$  also has no solution. Thus, by Theorem 2.1,  $(\mathcal{S}_n h^{(k)}(x))_{n \in \mathbb{N}}$  is a.s.u.d. mod 1.  $\square$

**Remark 3.7.** Lemma 3.6 may be of independent interest. With  $T$  being  $\mu$ -preserving, Lemma 3.6 implies that  $\lim_{n \rightarrow \infty} \int \delta_n(x) d\mu$  exists and together with the ergodic theorem this implies that for almost every  $x$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \delta_j(x) = \lim_{n \rightarrow \infty} \int \delta_n d\mu.$$

The limit on the left hand side can be explicitly calculated. Let  $t_n(x) = T^n x \cdot \frac{q_{n-1}(x)}{q_n(x)}$ , then Bosma, Jager and Wiedijk ([5, Theorem 3]) found that for a.e.  $x$ ,  $\frac{1}{n} \sum_{j=1}^n \mathbf{1}_{(0,z]}(t_j(x))$  converges to  $F(z)$  at every  $z \in [0, 1]$ , where

$$F(z) = \frac{1}{\log 2} \left( \log(1+z) - \frac{z}{1+z} \log z \right).$$

Since  $\delta_n(x) = -\log(1+t_n(x))$  and since  $\log(1+z)$  is a bounded and continuous function on  $[0, 1]$ , one has that for almost every  $x$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \delta_j(x) = \int_0^1 -\log(1+z) dF(z) = -1 - \frac{1}{2} \log 2 + \frac{\pi^2}{12 \log 2}.$$

So  $\lim_{n \rightarrow \infty} \int \delta_n d\mu$  is equal to this number as well.

Let

$$\theta_n(x) = q_n |xq_n - p_n|.$$

Since  $\delta_n = \log \theta_{n-1} + \log \frac{q_n}{q_{n-1}}$ , it follows that for almost every  $x$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n \log \theta_j(x) = -1 - \frac{1}{2} \log 2.$$

This gives an alternative proof of a result of Haas ([8]).

## References

- [1] Jon Aaronson and Manfred Denker, *Local limit theorems for partial sums of stationary sequences generated by Gibbs-Markov maps*, Stoch. Dyn. **1** (2001), no. 2, 193–237. MR1840194
- [2] Jon Aaronson, Manfred Denker, and Mariusz Urbański, *Ergodic theory for Markov fibred systems and parabolic rational maps*, Trans. Amer. Math. Soc. **337** (1993), no. 2, 495–548. MR1107025
- [3] Tim Bedford, Albert M. Fisher, and Mariusz Urbański, *The scenery flow for hyperbolic Julia sets*, Proc. London Math. Soc. (3) **85** (2002), no. 2, 467–492. MR1912058
- [4] Arno Berger and Theodore P. Hill, *A basic theory of Benford’s law*, Probab. Surv. **8** (2011), 1–126. MR2846899
- [5] W. Bosma, H. Jager, and F. Wiedijk, *Some metrical observations on the approximation by continued fractions*, Nederl. Akad. Wetensch. Indag. Math. **45** (1983), no. 3, 281–299. MR718069
- [6] Nicolas Chenavier, Bruno Massé, and Dominique Schneider, *Products of random variables and the first digit phenomenon*, Stochastic Process. Appl. **128** (2018), no. 5, 1615–1634. MR3780691
- [7] H. Furstenberg, *Strict ergodicity and transformation of the torus*, Amer. J. Math. **83** (1961), 573–601. MR133429
- [8] Andrew Haas, *An ergodic sum related to the approximation by continued fractions*, New York J. Math. **11** (2005), 345–349. MR2154360
- [9] P. J. Holewijn, *On the uniform distribution of sequences of random variables*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **14** (1969/70), 89–92. MR264735
- [10] Hendrik Jager and Pierre Liardet, *Distributions arithmétiques des dénominateurs de convergents de fractions continues*, Nederl. Akad. Wetensch. Indag. Math. **50** (1988), no. 2, 181–197. MR952514
- [11] Shigeru Kanemitsu, Kenji Nagasaka, Gérard Rauzy, and Jau-Shyong Shiue, *On Benford’s law: the first digit problem*, Probability theory and mathematical statistics (Kyoto, 1986), Lecture Notes in Math., vol. 1299, Springer, Berlin, 1988, pp. 158–169. MR935987

## Uniform distribution mod 1 for continued fractions

- [12] A. Ya. Khinchin, *Continued fractions*, translated from the third (1961) russian edition ed., Dover Publications, Inc., Mineola, NY, 1997, With a preface by B. V. Gnedenko, Reprint of the 1964 translation. MR1451873
- [13] L. Kuipers and H. Niederreiter, *Uniform distribution of sequences*, Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974, Pure and Applied Mathematics. MR0419394
- [14] Walter Philipp, *Mixing sequences of random variables and probabilistic number theory*, American Mathematical Society, Providence, R. I., 1971, Memoirs of the American Mathematical Society, No. 114. MR0437481
- [15] Walter Philipp and Olaf P. Stackelberg, *Zwei Grenzwertsätze für Kettenbrüche*, Math. Ann. **181** (1969), 152–156. MR244186
- [16] Herbert Robbins, *On the equidistribution of sums of independent random variables*, Proc. Amer. Math. Soc. **4** (1953), 786–799. MR56869
- [17] P. Schatte and K. Nagasaka, *A note on Benford's law for second order linear recurrences with periodical coefficients*, Z. Anal. Anwendungen **10** (1991), no. 2, 251–254. MR1155374
- [18] Peter Schatte, *On Benford's law for continued fractions*, Math. Nachr. **148** (1990), 137–144. MR1127337