Check for
updates

Prime numbers in typical continued fraction expansions

Tanja I. Schindler[1] · **Roland Zweimüller[1]**

Abstract

We study, from the viewpoint of metrical number theory and (infinite) ergodic theory, the probabilistic laws governing the occurrence of prime numbers as digits in continued fraction expansions of real numbers.

1 Introduction

Ever since Gauss [10] declared his interest in the intriguing statistical properties of sequences of *digits* $a_n(x)$, $n \geq 1$, in the *continued fraction (CF) expansion* of real numbers $x \in I := (0, 1]$,

$$x = [a_1(x), a_2(x), \ldots] = \cfrac{1}{a_1(x) + \cfrac{1}{a_2(x) + \cfrac{1}{a_3(x) + \cdots}}}$$

(and, in particular, mentioned that this led to questions he could not answer), the *metrical theory of continued fractions* has attracted the attention of many mathematicians. In the present paper we will be interested in the *prime digits* of x, i.e. those $a_n(x)$ which happen to belong to the set $\mathbb{P}$ of prime numbers. To single them out, we define, for $x \in I$ and $n \geq 1$,

$$a'_n(x) := \mathbb{1}_{\mathbb{P}}(a_n(x)) \cdot a_n(x) = \begin{cases} a_n(x) & \text{if } a_n(x) \in \mathbb{P}, \\ 0 & \text{otherwise.} \end{cases}$$

(There is hardly any danger of misinterpreting this phonetically perfect symbol as a derivative.) The purpose of this note is to point out that it is in fact possible—with the aid of the

✉ Tanja I. Schindler
tanja.schindler@univie.ac.at

Roland Zweimüller
roland.zweimueller@univie.ac.at

[1] Fakultät für Mathematik, Universität Wien, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria

prime number theorem and recent work in (infinite) ergodic theory and in the probability theory of dynamical systems—to derive a lot of information about the occurrences and values of prime digits in CF-expansions of (Lebesgue) typical numbers. Besides stating the theorems themselves it is also our aim to show some newer more general results in ergodic theory in action. While many analogous versions of the following statements have directly been proven for the continued fraction digits, today, it is possible to deduce them or the version for the prime digits from more general theorems.

The paper is structured as follows. In Sect. 2 we give results about the pointwise behaviour of the prime digits in the continued fraction expansion and in Sect. 3 about their distributional behaviour. Then we start with the proof sections, i.e. in Sect. 4 we state general results about the Gauss map and in Sects. 5 and 6 we give the proofs of the results from Sects. 2 and 3 respectively.

2 Main results—pointwise matters

We first consider questions about the pointwise behaviour of the sequence $(a'_n)_{n \geq 1}$ on I. Throughout, λ denotes Lebesgue measure on (the Borel σ-field $\mathcal{B}_I$ of) I, and *almost everywhere* (*a.e.*) is meant w.r.t. λ. For the sake of completeness, we also include a few easy basic facts, e.g. that for a.e. $x \in I$, the proportion of those $k \in \{1, \ldots, n\}$ for which $a_k(x)$ is prime converges:

Proposition 2.1 (Asymptotic frequency of prime digits) *We have*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\mathbb{P}} \circ a_k = \frac{1}{\log 2} \, \log \prod_{p \in \mathbb{P}} \left(1 + \frac{1}{p(p+2)} \right) \quad a.e.$$

The results to follow can best be understood (and proved) by regarding (a'_n) as a stationary sequence with respect to the Gauss measure (cf. §4 below). The first statement of the next theorem is parallel to the classical Borel-Bernstein theorem (cf. [5, 7]) the third and fourth statements are in accordance with [16]. As usual, *i.o.* is short for "infinitely often", i.e. for "infinitely many indices". We denote the *iterated logarithms* by $\log_1 := \log$ and $\log_{m+1} := \log \circ \log_m$, $m \geq 1$. Furthermore, we define the maximal entry $M'_n := \max_{1 \leq k \leq n} a'_k$, $n \geq 1$.

Theorem 2.1 (Pointwise growth of prime digits and maxima)
a) Assume that $(b_n)_{n \geq 1}$ is a sequence in $(1, \infty)$. Then

$$\lambda(\{a'_n > b_n \text{ i.o.}\}) = \begin{cases} 1 \text{ if } \sum_{n \geq 1} \frac{1}{b_n \log b_n} = \infty, \\ 0 \text{ otherwise.} \end{cases} \tag{2.1}$$

b) Moreover, if $(b_n)_{n \geq 1}$ is non-decreasing, then

$$\lambda(\{a'_n > b_n \text{ i.o.}\}) = \lambda(\{M'_n > b_n \text{ i.o.}\}). \tag{2.2}$$

c) Let $(c_n)_{n \geq 1}$ and $(d_n)_{n \geq 1}$ be sequences in $(1, \infty)$ with $d_n \to \infty$ and $c_n \leq d_n^{0.475}$ for large n. Then

$$\lambda(\{d_n \leq a'_n \leq d_n(1 + 1/c_n) \text{ i.o.}\}) = \begin{cases} 1 \text{ if } \sum_{n \geq 1} \frac{1}{c_n d_n \log(d_n)} = \infty. \\ 0 \text{ otherwise.} \end{cases}$$

d) Let $(d_n)_{n \geq 1}$ be a sequence of primes, then

$$\lambda(\{a'_n = d_n \text{ i.o.}\}) = \begin{cases} 1 \text{ if } \sum_{n \geq 1} \frac{1}{d_n^2} = \infty, \\ 0 \text{ otherwise.} \end{cases}$$

Remark 2.1 The exponent 0.475 in c) comes from estimates for the error term in the prime number theorem and might be improved by future research.

Example 2.1 A straightforward calculation shows that

$$\lambda(\{a'_n > n \log_2^\gamma n \text{ i.o.}\}) = \begin{cases} 1 \text{ if } \gamma \leq 1, \\ 0 \text{ otherwise}, \end{cases}$$

and this remains true if a'_n is replaced by M'_n. We thus find that

$$\overline{\lim_{n\to\infty}} \frac{\log a'_n - \log n}{\log_3 n} = \overline{\lim_{n\to\infty}} \frac{\log M'_n - \log n}{\log_3 n} = 1 \quad a.e.$$

As a consequence of Theorem 2.1 b), observing that the series $\sum_{n\geq 1} 1/(b_n \log b_n)$ converges iff $\sum_{n\geq 1} 1/(\rho b_n \log(\rho b_n))$ converges for every $\rho \in (0, \infty)$, we get

Corollary 2.1 *If* $(b_n)_{n\geq 1}$ *is non-decreasing, then*

$$\overline{\lim_{n\to\infty}} \frac{a'_n}{b_n} = \overline{\lim_{n\to\infty}} \frac{M'_n}{b_n} = \begin{cases} \infty \ a.e. \ \text{if } \sum_{n\geq 1} \frac{1}{b_n \log b_n} = \infty, \\ 0 \ a.e. \quad otherwise. \end{cases} \tag{2.3}$$

In particular,

$$\overline{\lim_{n\to\infty}} \frac{a'_n}{n \log_2 n} = \overline{\lim_{n\to\infty}} \frac{M'_n}{n \log_2 n} = \infty \ a.e. \tag{2.4}$$

A convenient condition for the criterion above is provided by

Lemma 2.1 *Let* $(b_n)_{n\geq 1}$ *be a sequence in* $(1, \infty)$ *for which* b_n/n *increases. Then*

$$\overline{\lim_{n\to\infty}} \frac{n \log_2 n}{b_n} > 0 \ implies \ \sum_{n\geq 1} \frac{1}{b_n \log b_n} = \infty.$$

As in the case of the full digit sequence $(a_n)_{n\geq 1}$, the peculiar properties of $(a'_n)_{n\geq 1}$ are due to the fact that these functions are not integrable. A general fact for non-integrable non-negative stationary sequences is the non-existence of a non-trivial strong law of large numbers, made precise in a), c) and d) of the next result, where c) is in the spirit of [22]. However, it is sometimes possible to recover a meaningful limit by *trimming*, i.e. by removing maximal terms. In the case of $(a_n)_{n\geq 1}$, this was first pointed out in [8]. Assertion b) below gives the proper version for the $(a'_n)_{n\geq 1}$.

Theorem 2.2 (Strong laws of large numbers)
a) The prime digits satisfy

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} a'_k = \infty \ a.e.$$

b) Subtracting M'_n, *we obtain a trimmed strong law,*

$$\lim_{n\to\infty} \frac{\log 2}{n \log_2 n} \left(\sum_{k=1}^{n} a'_k - M'_n \right) = 1 \ a.e. \tag{2.5}$$

c) For sequences $(b_n)_{n\geq 1}$ *in* $(1, \infty)$ *satisfying* $b_n/n \nearrow \infty$ *as* $n \to \infty$,

$$\overline{\lim_{n\to\infty}} \frac{1}{b_n} \sum_{k=1}^{n} a'_k = \infty \ a.e. \ iff \ \sum_{n\geq 1} \frac{1}{b_n \log b_n} = \infty, \tag{2.6}$$

while otherwise

$$\lim_{n\to\infty} \frac{1}{b_n} \sum_{k=1}^{n} a_k' = 0 \ a.e. \tag{2.7}$$

d) But, defining $\bar{n}(j) := e^{j \log^2 j}$, $j \geq 1$, *and* $d_n' := \bar{n}(j) \log_2 \bar{n}(j)/\log 2$ *for* $n \in (\bar{n}(j - 1), \bar{n}(j)]$ *gives a normalizing sequence for which*

$$\overline{\lim_{n\to\infty}} \frac{1}{d_n'} \sum_{k=1}^{n} a_k' = 1 \ a.e. \tag{2.8}$$

The trimmed law from b) shows that the bad pointwise behaviour described in c) is due to a few exceptionally large individual terms a_n' which, necessarily, have to be of the order of the preceding partial sum $\sum_{k=1}^{n-1} a_k'$. In fact, almost surely, the partial sum will infinitely often be of strictly smaller order than the following term, see statement a) below. We can also ask whether, or to what extent, the terms from the thinner sequence $(a_n')_{n\geq 1}$ come close to the partial sums $(\sum_{k=0}^{n-1} a_k)_{n\geq 1}$ of the unrestricted one. The answer is given by the dichotomy rule in statement b) of the next result.

We shall tacitly interpret real sequences $(g_n)_{n\geq 0}$ as functions on $\mathbb{R}_+$ via $t \longmapsto g_{[t]}$, and write $g(t) \sim h(t)$ as $t \to \infty$ if $g(t)/h(t) \to 1$. Moreover, $g(t) \asymp h(t)$ means $0 < \underline{\lim}_{t\to\infty} g(t)/h(t) \leq \overline{\lim}_{t\to\infty} g(t)/h(t) < \infty$.

Theorem 2.3 (Relative size of digits and partial sums)
a) We have

$$\overline{\lim_{n\to\infty}} \frac{a_n'}{\sum_{k=1}^{n-1} a_k'} = \infty \ a.e. \tag{2.9}$$

Generally, for functions $g : [0, \infty) \to (3, \infty)$ *fulfilling* $g(\eta(t)) \asymp g(t)$ *if* $\eta(t) \sim t$ *as* $t \to \infty$, *we have*

$$\overline{\lim_{n\to\infty}} \frac{g(a_n')}{\sum_{k=1}^{n-1} a_k'} = \infty \ a.e. \ iff \ \int_c^\infty \frac{g(y)}{\log_2 g(y)} \frac{dy}{y^2 \log y} = \infty, \tag{2.10}$$

while otherwise

$$\lim_{n\to\infty} \frac{g(a_n')}{\sum_{k=1}^{n-1} a_k'} = 0 \ a.e.$$

b) In contrast, comparing to the unrestricted digit sum $\sum_{k=1}^{n-1} a_k$, *one has*

$$\lim_{n\to\infty} \frac{a_n'}{\sum_{k=1}^{n-1} a_k} = 0 \ a.e.$$

Generally, for functions $g : [0, \infty) \to (3, \infty)$ *fulfilling* $g(\eta(t)) \asymp g(t)$ *if* $\eta(t) \sim t$ *as* $t \to \infty$, *we have*

$$\overline{\lim_{n\to\infty}} \frac{g(a_n')}{\sum_{k=1}^{n-1} a_k} = \infty \ a.e. \ iff \ \int_c^\infty \frac{g(y)}{\log g(y)} \frac{dy}{y^2 \log y} = \infty,$$

while otherwise

$$\lim_{n\to\infty} \frac{g(a_n')}{\sum_{k=0}^{n-1} a_k} = 0 \ a.e.$$

c) *Turning to a comparison of partial sums, we find that*

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} a_k'}{\sum_{k=1}^{n} a_k} = 0 \ a.e.$$

Remark 2.2 A broad class of functions which satisfy $g(\eta(t)) \asymp g(t)$ if $\eta(t) \sim t$ as $t \to \infty$, are the regularly varying functions. Recall that a measurable function $g : (L, \infty) \to (0, \infty)$ is *regularly varying of index* $\rho \in \mathbb{R}$ at infinity, written $g \in \mathcal{R}_\rho$, if $g(ct)/g(t) \to c^\rho$ as $t \to \infty$ for all $c > 0$ (see Chapter 1 of [6] for more information).

Whether or not the integrals diverge can easily be checked for many specific g's:

Example 2.2 **a)** Taking $g(t) := t \, \log^\rho t$, $\rho \in \mathbb{R}$, part b) gives

$$\overline{\lim_{n \to \infty}} \frac{a_n' \log a_n'}{\sum_{k=1}^{n-1} a_k} = \infty \ \text{a.e. for } \rho > 1 \quad \text{while} \quad \lim_{n \to \infty} \frac{a_n' \log^\rho a_n'}{\sum_{k=1}^{n-1} a_k} = 0 \ \text{a.e. for } \rho \le 1.$$

b) In case $g(t) := t \, \log t / \log_2^\gamma t$, $\gamma \in \mathbb{R}$, we find for $\gamma \le 1$

$$\overline{\lim_{n \to \infty}} \frac{a_n' \log a_n' / \log_2 a_n'}{\sum_{k=1}^{n-1} a_k} = \infty \ \text{a.e.}$$

while, for $\gamma > 1$,

$$\lim_{n \to \infty} \frac{a_n' \log a_n' / \log_2^\gamma a_n'}{\sum_{k=1}^{n-1} a_k} = 0 \ \text{a.e.}$$

On the other hand, if we look at primes to some power γ we obtain—as a counterpart to Theorem 2.2 b)—the following result:

Theorem 2.4 **a)** *For $\gamma < 1$ there exists $K_\gamma > 0$ such that*

$$\lim_{n \to \infty} \frac{\sum_{k=1}^{n} (a_k')^\gamma}{n} = K_\gamma < \infty \ a.e.$$

b) *Let $\sigma := \sigma_{(n,x)} \in \mathcal{S}_n$ be a pointwise permutation, i.e. $\sigma : I \times \{1, \ldots, n\} \to \{1, \ldots, n\}$, such that $a_{\sigma(1)}' \ge \ldots \ge a_{\sigma(n)}'$ and $S_n^k := \sum_{j=k+1}^{n} a_{\sigma(j)}'$. If $\gamma > 1$, then for all $(b_n) \in \mathbb{N}^{\mathbb{N}}$ fulfilling $b_n = o(n^{1-\epsilon})$ for some $\epsilon > 0$ and*

$$\lim_{n \to \infty} \frac{b_n}{\log \log n} = \infty \tag{2.11}$$

we have

$$\lim_{n \to \infty} \frac{S_n^{b_n}}{d_n} = 1 \ a.e. \tag{2.12}$$

where

$$d_n \sim \frac{1}{(\gamma - 1)(\log 2)^\gamma} \cdot \frac{n^\gamma b_n^{1-\gamma}}{(\log n)^\gamma}. \tag{2.13}$$

Remark 2.3 It is not proven that a trimming rate slower than the one given in (2.11) is possible. However, by [11] one can deduce that for i.i.d. random variables with the same distribution function and $b_n \asymp \log \log n$ a strong law of large numbers as in (2.12) is no longer possible.

However, if we only ask for convergence in probability, the picture looks much simpler and we refer the reader to Theorem 3.1 in the next section.

3 Main results—distributional matters

The second set of results we present focuses on the distributions of (various functions of) the digits a'_n. If (M, d) is a separable metric space with Borel σ-field $\mathcal{B}_M$, a sequence $(\nu_n)_{n \geq 1}$ of probability measures on $(M, \mathcal{B}_M)$ *converges weakly* to the probability measure ν on $(M, \mathcal{B}_M)$, written $\nu_n \Longrightarrow \nu$, if the integrals of bounded continuous function $\psi : M \to \mathbb{R}$ converge, i.e. $\int \psi \, d\nu_n \longrightarrow \int \psi \, d\nu$ as $n \to \infty$. If $R_n : I \to M$, $n \geq 1$, Borel measurable functions and ν a Borel probability on M (or R another *random element* of M, not necessarily defined on I, with distribution ν) then $(R_n)_{n \geq 1}$ *converges in distribution* to ν (or to R) *under the probability measure P* on $\mathcal{B}_I$, if the distributions $P \circ R_n^{-1}$ of the R_n w.r.t. P converge weakly to ν. Explicitly specifying the underlying measure, we denote this by

$$R_n \overset{P}{\Longrightarrow} \nu \text{ or } R_n \overset{P}{\Longrightarrow} R.$$

For sequences (R_n) defined on an ergodic dynamical system, it is often the case that a distributional limit theorem $R_n \overset{P}{\Longrightarrow} R$ automatically carries over to a large collection of other probability measures: *strong distributional convergence*, written

$$R_n \overset{\mathcal{L}(\lambda)}{\Longrightarrow} \nu \text{ or } R_n \overset{\mathcal{L}(\lambda)}{\Longrightarrow} R,$$

means that $R_n \overset{P}{\Longrightarrow} R$ for all probability measures $P \ll \lambda$, see [24].

We start by giving a counterpart to Theorem 2.4 for weak convergence, where b) is in the spirit of [19].

Theorem 3.1 *a) For $\gamma < 1$ there exists $K_\gamma > 0$ such that*

$$\frac{\sum_{k=1}^{n} (a'_k)^\gamma}{n} \overset{\mathcal{L}(\lambda)}{\Longrightarrow} K_\gamma.$$

b) For the case $\gamma = 1$ we have

$$\frac{\sum_{k=1}^{n} a'_k}{n \log_2 n} \overset{\mathcal{L}(\lambda)}{\Longrightarrow} \log 2.$$

c) If $\gamma > 1$ we have

$$\frac{S_n^{b_n}}{d_n} \overset{\mathcal{L}(\lambda)}{\Longrightarrow} 1,$$

where $S_n^{b_n}$ is defined as in Theorem 2.4, (d_n) is given as in (2.13) and $\lim_{n \to \infty} b_n = \infty$ and $b_n = o(n^{1-\epsilon})$.

Remark 3.1 Indeed by [17] the stronger result of convergence in mean follows for c). It is not proven that for the situation in c) convergence in probability can not hold for a lightly trimmed sum, i.e. a sum from which only a finite number of large entries, being independent of n is removed. However, it follows from [1] that $\sum_{k=1}^{n} (a'_k)^\gamma$ normed by the right norming sequence converges to a non-degenerate Mittag-Leffler distribution if $\gamma > 1$. On the other hand, by [18] it follows that light trimming does not have any influence on distributional convergence if the random variables considered are i.i.d.

As we have seen in the previous section, the maximum M'_n has a large influence then the whole system, in the following we will give its distributional convergence. We let Θ denote a positive random variable with $\Pr[\Theta \leq y] = e^{-1/y}$, $y > 0$ and get the following counterpart to [21].

Theorem 3.2 (Distributional convergence of M_n')*The maximum M_n' of the prime digits converges in distribution,*

$$\frac{\log 2 \, \log n}{n} \cdot M_n' \overset{\mathcal{L}(\lambda)}{\Longrightarrow} \Theta \quad as \; n \to \infty. \tag{3.1}$$

A related classical topic, introduced by Doeblin [9], is the Poissonian nature of occurrences of very large CF-digits. For $l \geq 1$ let $\varphi_l = \varphi_{l,1} := \inf\{k \geq 1 : a_k \geq l\}$, the first position in the CF-expansion at which a digit $\geq l$ shows up, and $\varphi_{l,i+1} := \inf\{k \geq 1 : a_{\varphi_{l,i}+k} \geq l\}$ the distance between the ith and $(i+1)$st occurrence. Defining $\Phi_l : I \to [0, \infty]^{\mathbb{N}}$ as $\Phi_l := (\varphi_{l,1}, \varphi_{l,2}, \ldots)$ and letting Φ_{Exp} denote an i.i.d. sequence of normalized exponentially distributed random variables, we can express this classical result by stating that

$$\frac{1}{\log 2} \frac{1}{l} \cdot \Phi_l \overset{\lambda}{\Longrightarrow} \Phi_{\text{Exp}} \quad as \; l \to \infty.$$

Turning to prime digits, we shall consider the corresponding quantities $\varphi_{l,i}'$ with $\varphi_{l,0}' := 0$ and $\varphi_{l,i+1}' := \inf\{k \geq 1 : a_{\varphi_{l,i}'+k}' \geq l\}$, $i \geq 0$, and the processes $\Phi_l' := (\varphi_{l,1}', \varphi_{l,2}', \ldots)$ of distances between consecutive occurrences of prime digits of size at least l. In fact, we also provide refined versions of the limit theorem which show that, asymptotically, both the relative size compared to l of such a large prime digit $a_{\varphi_{l,i}'}'$ and its residue class for a given modulus m, are stochastically independent of the positions $\varphi_{l,i}'$ at which they occur. (These statements are parallel to Propositions 10.1 and 10.2 of [25]. A $(q_1, \ldots, q_d)$-Bernoulli sequence is an iid sequence of random variables which can assume d different values with respective probabilities $q_1, \ldots, q_d$.)

Theorem 3.3 (Poisson limits for large prime CF-digits) *The sequences Φ_l' of positions at which large prime digits occur satisfy the following.*
a) Their distances converge to an i.i.d. sequence of exponential variables,

$$\frac{1}{\log 2} \frac{1}{l \log l} \cdot \Phi_l' \overset{\mathcal{L}(\lambda)}{\Longrightarrow} \Phi_{\text{Exp}} \quad as \; l \to \infty. \tag{3.2}$$

b) Take any $\vartheta \in (0, 1)$, let $\psi_{l,i}'$ be the indicator function of $\{a_{\varphi_{l,i}'}' \geq l/\vartheta\}$ and set $\Psi_l' := (\psi_{l,1}', \psi_{l,2}', \ldots)$, which identifies those prime digits $\geq l$ which are in fact $\geq l/\vartheta$. Then

$$\left(\frac{1}{\log 2} \frac{1}{l \log l} \cdot \Phi_l', \Psi_l' \right) \overset{\mathcal{L}(\lambda)}{\Longrightarrow} (\Phi_{\text{Exp}}, \Psi') \quad as \; l \to \infty, \tag{3.3}$$

where $(\Phi_{\text{Exp}}, \Psi')$ is an independent pair with Ψ' a $(1 - \vartheta, \vartheta)$-Bernoulli sequence.
c) Fix an integer $m \geq 2$. For $l > m$ define $\upsilon_{l,i}' : I \to \{j \in \{1, \ldots, m\} : j \text{ relatively prime to } m\}$ by $\upsilon_{l,i}'(x) := j$ if $a_{\varphi_{l,i}'}'(x) \equiv j \mod m$, so that $\Upsilon_l' := (\upsilon_{l,1}', \upsilon_{l,2}', \ldots)$ identifies the residue classes mod m of the prime digits $a_{\varphi_{l,i}'}'$. Then

$$\left(\frac{1}{\log 2} \frac{1}{l \log l} \cdot \Phi_l', \Upsilon_l' \right) \overset{\mathcal{L}(\lambda)}{\Longrightarrow} (\Phi_{\text{Exp}}, \Upsilon') \quad as \; l \to \infty, \tag{3.4}$$

where $(\Phi_{\text{Exp}}, \Upsilon')$ is an independent pair with Υ' a $\left(\frac{1}{\phi(m)}, \ldots, \frac{1}{\phi(m)} \right)$-Bernoulli sequence. (Here $\phi(m)$ denotes the Euler totient.)

We finally look at the distribution of a function which counts how many a_n' fall into particular sets A_n giving a limit theorem in the spirit of [16, 20]. We let $\mathcal{N}$ denote a positive random variable with $\Pr[\mathcal{N} \leq y] = \int_0^y e^{-t^2/2} \, dt/\sqrt{2\pi}$, $y > 0$.

Theorem 3.4 (A CLT for counting primes in CF) *Suppose that either*

(A) $A_n := \{a'_n \geq b_n\}$ *with* $(b_n) \in \mathbb{R}^{\mathbb{N}}$ *and* $\sum_{n:b_n > 1} 1/b_n \log b_n = \infty$,

(B) $A_n := \{a'_n = d_n\}$ *with* (d_n) *a sequence of primes and* $\sum_{n \in \mathbb{N}} 1/d_n^2 = \infty$,

(C) $A_n := \left\{ d_n \leq a'_n \leq d_n \cdot \left(1 + \frac{1}{c_n}\right) \right\}$ *with* (d_n) *a sequence of natural numbers tending to infinity,* (c_n) *a sequence of positive numbers with* $c_n \leq d_n^{0.475}$ *and* $\sum_{n=1}^{\infty} 1/(c_n d_n \log(d_n)) = \infty$.

Then, for $S_n := \sum_{k=1}^{n} \mathbb{1}_{A_k}$ *the following central limit theorem holds:*

$$\frac{S_n - \int S_n \, d\mu_{\mathfrak{G}}}{\sqrt{\int \left(S_n - \int S_n \, d\mu_{\mathfrak{G}}\right)^2}} \overset{\mathcal{L}(\lambda)}{\Longrightarrow} \mathcal{N} \quad as \; n \to \infty.$$

4 The Gauss map and the prime digit function

The results announced above express properties of certain stochastic processes derived from the exceptionally well understood dynamical system generated by the ergodic *continued fraction map* (or *Gauss map*)

$$S : (0, 1] \to [0, 1], \; Sx := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor = \frac{1}{x} - k \; \text{for} \; x \in \left(\frac{1}{k+1}, \frac{1}{k} \right] =: I_k, k \geq 1$$

which, since [10], is known to preserve the probability density

$$h_{\mathfrak{G}}(x) := \frac{1}{\log 2} \frac{1}{1+x}, \; x \in I.$$

The invariant *Gauss measure* $\mu_{\mathfrak{G}}$ on $\mathcal{B}_I$ defined by the latter, $\mu_{\mathfrak{G}}(B) := \int_B h_{\mathfrak{G}}(x) \, dx$, is exact (and hence ergodic). As hardly any textbook on ergodic theory fails to point out, iteration of S reveals the continued fraction digits of any $x \in I$, in that

$$x = [a_1(x), a_2(x), \ldots] \; \text{with} \; a_n(x) = a \circ S^{n-1}(x), n \geq 1,$$

where $a : I \to \mathbb{N}$ is the *digit function* corresponding to the partition $\xi := \{I_k : k \geq 1\}$, i.e. $a(x) := \lfloor 1/x \rfloor = k$ for $x \in I_k$. The stationary sequence $(a \circ S^n)_{n \geq 0}$ on the probability space $(I, \mathcal{B}_I, \mu_{\mathfrak{G}})$ thus obtained exhibits interesting properties since a has infinite expectation, $\int_I a \, d\mu_{\mathfrak{G}} = \sum_{k \geq 1} k \, \mu_{\mathfrak{G}}(I_k) = \infty$, as $\mu_{\mathfrak{G}}(I_k) = \log(\frac{(k+1)^2}{k(k+2)})/\log 2 \sim 1/(\log 2 \cdot k^2)$ for $k \to \infty$. As in classical probability theory, the *tail behaviour of the distribution*, given by

$$\mu_{\mathfrak{G}}(\{a \geq K\}) = \frac{1}{\log 2} \cdot \log \left(\frac{K+1}{K} \right) \sim \frac{1}{\log 2} \cdot \frac{1}{K} \; as \; K \to \infty$$

(which entails $L(N) := \int_I (a \wedge N) \, d\mu_{\mathfrak{G}} = \sum_{K=1}^{N} \mu_{\mathfrak{G}}(\{a \geq K\}) \sim \log N / \log 2$ as $N \to \infty$), is the key to fine asymptotic results. However, the study of the CF digit sequence goes beyond standard results, since the random variables $a \circ S^n$ are not independent. Yet, it is well known that they still satisfy a strong form of *asymptotic independence* or *mixing* in the following sense:

Given any measure preserving transformation T on a probability space $(\Omega, \mathcal{B}, P)$, and a countable measurable partition γ (mod P), the ψ-*mixing coefficients of* γ are defined as

$$\psi_\gamma(n) := \sup_{k \geq 1} \left\{ \left| \frac{P(V \cap W)}{P(V)P(W)} - 1 \right| : \begin{array}{l} V \in \sigma(\bigvee_{j=0}^{k-1} T^{-j} \gamma), P(V) > 0, \\ W \in T^{-(n+k-1)} \mathcal{B}, P(W) > 0 \end{array} \right\}, n \geq 1.$$

The partition γ is said to be *continued-fraction (CF-) mixing* for the probability preserving system $(\Omega, \mathcal{B}, P, T)$ if it is generating, and if $\psi_\gamma(1) < \infty$ as well as $\psi_\gamma(n) \to 0$ for $n \to \infty$. (Note that $(\psi_\gamma(n))_{n \geq 1}$ is non-increasing.) Of course, the nomenclature is due to the fact that

$$\xi \text{ is CF-mixing for } (I, \mathcal{B}_I, \mu_\mathfrak{G}, S). \tag{4.1}$$

Actually, this system is *exponentially CF-mixing*, in that there are constants $C > 0$ and $\rho \in (0, 1)$ such that

$$\psi_\xi(n) \leq C \rho^n \text{ for } n \geq 1$$

(which is related to Gauss' famous question mentioned in the introduction, see e.g. [13] or [23]).

We are going to study occurrences of *prime digits* by considering the *restricted digit function* $\mathrm{a}' := (\mathbb{1}_\mathbb{P} \circ \mathrm{a}) \cdot \mathrm{a} : I \to \{0\} \cup \mathbb{P}$. As in the case of a, this function, as a random variable on $(I, \mathcal{B}_I, \mu_\mathfrak{G})$, still has infinite expectation. Indeed, the prime number theorem (PNT) enables us to quickly determine the all-important tail asymptotics for the distribution of a'. The following lemma is the key to our analysis of the prime digit sequence.

Lemma 4.1 (Tail behaviour and truncated expectation of a')*The distribution of* a' *(with respect to the Gauss measure) satisfies*

$$\mu_\mathfrak{G}\left(\{\mathrm{a}' \geq K\}\right) \sim \frac{1}{\log 2} \cdot \frac{1}{K \log K} \text{ as } K \to \infty. \tag{4.2}$$

In particular, a' *is not integrable,* $\int_I \mathrm{a}' \, d\mu_\mathfrak{G} = \infty$. *Moreover,*

$$L'(N) := \int_I (\mathrm{a}' \wedge N) \, d\mu_\mathfrak{G} \sim \frac{\log_2 N}{\log 2} \text{ as } N \to \infty. \tag{4.3}$$

so that $a'(N) := N/L'(N) \sim \log 2 \cdot N/\log_2 N$ *is asymptotically inverse to* $b'(N) := (N \log_2 N)/\log 2$.

Proof First, the PNT is easily seen (cf. [12], Theorem 1.8.8) to imply that

$$\mathrm{p}_n \sim n \log n \text{ as } n \to \infty, \tag{4.4}$$

where p_n denotes the *n*th prime number. Therefore,

$$\sum_{n \geq N} \frac{1}{\mathrm{p}_n^2} \sim \sum_{n \geq N} \frac{1}{n^2 (\log n)^2} \sim \frac{1}{N (\log N)^2} \text{ as } N \to \infty.$$

Letting $N(K)$ denote the least n with $\mathrm{p}_n \geq K$, we have, as $K \to \infty$,

$$\mu_\mathfrak{G}\left(\{\mathrm{a}' \geq K\}\right) = \sum_{\mathrm{p} \geq K, \mathrm{p} \in \mathbb{P}} \mu_\mathfrak{G}(I_\mathrm{p}) \sim \frac{1}{\log 2} \sum_{\mathrm{p} \geq K, \mathrm{p} \in \mathbb{P}} \frac{1}{\mathrm{p}^2} = \frac{1}{\log 2} \sum_{n \geq N(K)} \frac{1}{\mathrm{p}_n^2}$$

and, by PNT, $N(K) \sim K/\log K$. Combining these observations yields (4.2). The second statement is an easy consequence thereof, since

$$L'(N) = \sum_{K=1}^N \mu_\mathfrak{G}\left(\{\mathrm{a}' \geq K\}\right) \sim \frac{1}{\log 2} \sum_{K=2}^N \frac{1}{K \log K} \sim \frac{1}{\log 2} \int_2^N \frac{dx}{x \log x} \text{ as } N \to \infty.$$

Straightforward calculation verifies the assertions about a' and b'. □

Remark 4.1 Several of the results allow for analogues in which prime digits are replaced by digits belonging to other subsets $\mathbb{M}$ of the integers for which $\pi_{\mathbb{M}}(n) := \#\mathbb{M} \cap \{1, \ldots, n\}$ is regularly varying with $\sum_{m \in \mathbb{M}} \frac{1}{m} = \infty$, like, for example, the set of integers which are the product of exactly k prime numbers, see Theorem 3.5.11 of [14]. (M. Thaler, personal communication.)

5 Proofs of the results on a.e. convergence

We are now ready for the proofs of our pointwise convergence results. We can always work, without further mention, with the invariant measure $\mu_{\mathfrak{G}}$, since it has the same null-sets as λ.

Proof of Proposition 2.1 This, of course, is just the ergodic theorem,

$$\frac{1}{n} \sum_{k=1}^{n} \mathbb{1}_{\mathbb{P}} \circ a_k = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{\mathbb{P}} \circ S^k \longrightarrow \mu_{\mathfrak{G}}(\mathbb{P}) = \sum_{p \in \mathbb{P}} \mu_{\mathfrak{G}}(I_p) \text{ a.e. as } n \to \infty.$$

$\square$

In the following we will repeatedly appeal to the following version of Rényi's Borel-Cantelli Lemma (BCL) (as in Lemma 1 of [3]):

Lemma 5.1 (Rényi's Borel-Cantelli Lemma) *Assume that $(E_n)_{n \geq 1}$ is a sequence of events in the probability space $(\Omega, \mathcal{B}, P)$ for which there is some $r \in (0, \infty)$ such that*

$$\frac{P(E_j \cap E_k)}{P(E_j) P(E_k)} \leq r \quad \text{whenever } j, k \geq 1, \, j \neq k.$$

Then $P(\{E_n \text{ infinitely often}\}) > 0$ iff $\sum_{n \geq 1} P(E_n) = \infty$.

This lemma enables us to prove Theorem 2.1.

Proof of Theorem 2.1 a) Note that $\{a'_j > c\} = S^{-(j-1)}\{a' > c\}$ with $\{a' > c\}$ measurable w.r.t. ξ. As a consequence of the CF-mixing property (4.1), we see that Rényi's BCL applies to show that

$$\mu_{\mathfrak{G}}(\{a'_n > b_n \text{ i.o.}\}) > 0 \text{ iff } \sum_{n \geq 1} \mu_{\mathfrak{G}}(\{a'_n > b_n\}) = \infty. \tag{5.1}$$

By S-invariance of $\mu_{\mathfrak{G}}$ and Lemma 4.1, we have $\mu_{\mathfrak{G}}(\{a'_n \geq b_n\}) = \mu_{\mathfrak{G}}(\{a' \geq b_n\}) \sim 1/(b_n \log b_n)$, so that divergence of the right-hand series in (5.1) is equivalent to that of $\sum_{n \geq 1} (b_n \log b_n)^{-1}$. Finally, again because of $\{a'_j > c\} = S^{-(j-1)}\{a' > c\}$, the set $\{a'_n > b_n \text{ i.o.}\}$ is easily seen to belong to the tail-σ-field $\bigcap_{n \geq 0} S^{-n} \mathcal{B}_I$ of S. The system $(I, \mathcal{B}_I, \mu_{\mathfrak{G}}, S)$ being exact, the latter is trivial mod $\mu_{\mathfrak{G}}$. Hence $\mu_{\mathfrak{G}}(\{a'_n > b_n \text{ i.o.}\}) > 0$ implies $\mu_{\mathfrak{G}}(\{a'_n > b_n \text{ i.o.}\}) = 1$.

b) Statement (2.2) is seen by an easy routine argument, as in the proof of Proposition 3.1.8 of [13].

c) Without loss of generality we first assume that $c_n \leq 0.5$, for all n. If this doesn't hold, we can easily switch to a subsequence in which this holds and consider the subsequences

separately. By the prime number theorem we have

$$\lambda(a_1' \in [d_n, d_n(1 + 1/c_n)]) = \lambda(a_1' \geq d_n) - \lambda(a_1' \geq d_n(1 + 1/c_n))$$

$$\asymp \frac{1}{d_n \log d_n} - \frac{1}{d_n(1 + 1/c_n) \log(d_n(1 + 1/c_n))}$$

$$\sim \frac{1}{c_n d_n \log d_n}.$$

Next, we assume that $c_n > 0.5$. We note that

$$\#\mathbb{P} \cap (d_n, d_n(1 + 1/c_n)] \cdot \lambda(a_1 = d_n) \leq \lambda(a_1' \in [d_n, d_n(1 + 1/c_n)])$$
$$\leq \#\mathbb{P} \cap (d_n, d_n(1 + 1/c_n)] \cdot \lambda(a_1 = d_n(1 + 1/c_n)).$$
$$(5.2)$$

Furthermore,

$$\lambda(a_1 = d_n) \asymp \frac{1}{d_n^2} \asymp \lambda(a_1 = d_n(1 + 1/c_n)). \tag{5.3}$$

On the other hand, we have by [4], p. 562 that there exists $K > 0$ such that

$$\#\mathbb{P} \cap (d_n, d_n(1 + 1/c_n)] \sim \pi(d_n(1 + 1/c_n)) - \pi(d_n) \leq K \cdot \frac{d_n}{c_n \log d_n}.$$

Combining this with (5.2) and (5.3) yields the statement of c).
d) This follows immediately from [16, Theorem 6a]. □

Proof of Lemma 2.1 By assumption there is some $\varepsilon \in (0, 1)$ such that the set $M := \{n \geq 1 : (n \log_2 n)/b_n \geq \varepsilon\}$ is infinite. Define $c(x) := \exp(\sqrt{\log x})$ and $f(x) := x \log x \log_2 x$ for $x > 1$, and note that $c(x) < x$ for $x > e$. Suppose that $n \in M$, $n \geq 4$, and $c(n) < k \leq n$. Since $k \leq n$, we have $b_k = (b_k/k)k \leq (b_n/n)k \leq (1/\varepsilon)k \log_2 n$, and thus

$$b_k \log b_k \leq (1/\varepsilon)k \log_2 n(\log(1/\varepsilon) + \log k + \log_3 n)$$

$$= \frac{f(k)}{\varepsilon} \frac{\log_2 n}{\log_2 k} \left(\frac{-\log(\varepsilon)}{\log k} + 1 + \frac{\log_3 n}{\log k} \right).$$

On the other hand, $c(n) < k$ implies $\log k > \sqrt{\log n}$ and hence

$$\log_2 k > (1/2) \log_2 n, \quad \log k > 1, \quad \text{and} \quad \log k > \log_3 n.$$

Using these estimates we see that

$$b_k \log b_k \leq \frac{f(k)}{C(\varepsilon)} \quad \text{with} \quad C(\varepsilon) := \frac{\varepsilon}{2(-\log(\varepsilon) + 2)} > 0.$$

Taking into account that $\log_3 x$ is a primitive of $1/f(x)$ we get

$$\sum_{k>c(n)} \frac{1}{b_k \log b_k} \geq C(\varepsilon) \sum_{c(n)<k\leq n} \frac{1}{f(k)}$$

$$\geq C(\varepsilon) \left(\int_{c(n)}^n \frac{dx}{f(x)} - \frac{1}{f([c(n)])} \right) = C(\varepsilon) \left(\log 2 - \frac{1}{f([c(n)])} \right).$$

Since this estimate holds for infinitely many n, we see that

$$\varliminf_{n \to \infty} \sum_{k>n} \frac{1}{b_k \log b_k} \geq C(\varepsilon) \log 2,$$

proving that $\sum_{k \geq 1} \frac{1}{b_k \log b_k}$ diverges.

Proof of Theorem 2.2 a) Since $\int_I a' \, d\mu_{\mathfrak{G}} = \infty$ by Lemma 4.1, this is immediate from the ergodic theorem. **b)** We apply Theorem 1.1 of [2] to $(I, \mathcal{B}_I, \mu_{\mathfrak{G}}, S)$ and a', observing that (in the notation of that paper), $\mathfrak{N}_{a'} = 1$ since $J_1 = \sum_{n \geq 1} \left(n^2 \log n \log_2 n \right)^{-1} < \infty$. Furthermore, by using the estimate of Lemma 4.1 and setting $a'(N) := N/L'(N) \sim \log 2 \cdot N / \log_2 N$ we get that its asymptotic inverse can be written as $b'(N) := (N \log_2 N)/\log 2$ which by the statement of the paper coincides with the norming sequence. **c)** Using Theorem 2.1 a), we first note that $\sum_{n \geq 1} 1/(b_n \log b_n) = \infty$ implies $\overline{\lim}_{n \to \infty} b_n^{-1} \sum_{k=1}^n a_k' = \infty$ a.e. since $a_n' \leq \sum_{k=1}^n a_k'$. For the converse, assume that $\sum_{n \geq 1} 1/(b_n \log b_n) < \infty$, which by Lemma 2.1 entails $(n \log_2 n)/b_n \to 0$. In view of Theorem 2.1 a), our assumption implies that $M_n'/b_n \to 0$ a.e. Together with statement b) above, these observations prove (2.7), because

$$\frac{1}{b_n} \sum_{k=1}^n a_k' = \frac{M_n'}{b_n} + \frac{n \log_2 n}{\log 2 \cdot b_n} \cdot \frac{\log 2}{n \log_2 n} \left(\sum_{k=1}^n a_k' - M_n' \right).$$

d) Note first that letting $c_n' := d_n'/\log_2(10j)$ for $n \in (\bar{n}(j-1), \bar{n}(j)]$, provides us with a non-decreasing sequence satisfying $\sum_{n \geq 1} 1/(c_n' \log c_n') < \infty$ (use generous estimates). By Theorem 2.1 therefore $\lambda(\{M_n' > c_n' \text{ i.o.}\}) = 0$. Since $c_n' = o(d_n')$, we see that for every $\varepsilon > 0$, $\{M_n' > \varepsilon \, d_n' \text{ i.o.}\} \subseteq \{M_n' > c_n' \text{ i.o.}\}$. Combining these observations shows that

$$\frac{M_n'}{d_n'} \longrightarrow 0 \text{ a.e.} \tag{5.4}$$

Together with (2.5) and $n \log_2 n/(\log 2 \cdot d_n') \leq 1$, this proves, via

$$\frac{1}{d_n'} \sum_{k=1}^n a_k' = \frac{n \log_2 n}{\log 2 \cdot d_n'} \cdot \frac{\log 2}{n \log_2 n} \left(\sum_{k=1}^n a_k' - M_n' \right) + \frac{M_n'}{d_n'}, \tag{5.5}$$

that

$$\overline{\lim}_{n \to \infty} \frac{1}{d_n'} \sum_{k=1}^n a_k' \leq 1 \text{ a.e.}$$

Specializing (5.5), and using (2.5) and (5.4) again, we find that

$$\frac{1}{d_{\bar{n}(j)}'} \sum_{k=1}^{\bar{n}(j)} a_k' = \frac{\log 2}{\bar{n}(j) \log_2 \bar{n}(j)} \left(\sum_{k=1}^{\bar{n}(j)} a_k' - M_{\bar{n}(j)}' \right) + \frac{M_{\bar{n}(j)}'}{d_{\bar{n}(j)}'} \longrightarrow 1 \text{ a.e.}$$

as $j \to \infty$, and our claim (2.8) follows. $\qquad \square$

Proof of Theorem 2.3 a) Apply Theorem 4 of [3] to the system $(I, \mathcal{B}_I, \mu_{\mathfrak{G}}, S)$ with CF-mixing partition $\gamma := \xi$. Statement (2.9) is immediate if we take (a', a') as our pair (φ, ψ) of γ-measurable functions, cf. Remark 3 in [3]. Turning to the general version (2.10), we consider $\varphi := g \circ a'$ and $\psi := a'$. According to the result cited,

$$\overline{\lim}_{n \to \infty} \frac{g(a_n')}{\sum_{k=1}^{n-1} a_k'} = \infty \text{ a.e. iff } \int_I a' \circ g \circ a' \, d\mu_{\mathfrak{G}} = \infty,$$

(with a' from Lemma 4.1), while otherwise $\lim_{n \to \infty} g(a_n')/(\sum_{k=1}^{n-1} a_k') = 0$ a.e. The present assertion merely reformulates the divergence condition above: We see (using (4.4) and the

regularity properties on g) that (for some constant $c > 0$)

$$\int_I a \circ g \circ a'\, d\mu_\mathfrak{G} \asymp \sum_{n \geq 1} \frac{g(\mathrm{p}_n)}{\log_2 g(\mathrm{p}_n)} \mu_\mathfrak{G}(I_{\mathrm{p}_n}) \asymp \sum_{n \geq 1} \frac{g(n \log n)}{\log_2 g(n \log n)} \frac{1}{(n \log n)^2}$$

$$\asymp \int_c^\infty \frac{g(x \log x)}{\log_2 g(x \log x)} \frac{dx}{(x \log x)^2} \asymp \int_c^\infty \frac{g(y)}{\log_2 g(y)} \frac{dy}{y^2 \log y}.$$

b) Same argument as in a), this time with $\varphi := g \circ a'$ and $\psi := a$, and replacing a' above by $a(t) := t/L(t) \sim \log 2 \cdot t/\log t$ as $t \to \infty$.

c) We have

$$\lim_{n \to \infty} \frac{\sum_{k=1}^n a_k'}{\sum_{k=1}^n a_k} \leq \lim_{n \to \infty} \frac{\sum_{k=1}^n a_k' - \mathrm{M}_n'}{\sum_{k=1}^n a_k - \mathrm{M}_n} + \lim_{n \to \infty} \frac{\mathrm{M}_n'}{\sum_{k=1}^n a_k - \mathrm{M}_n}$$

$$\leq \lim_{n \to \infty} \frac{n \log_2 n \log 2}{n \log n \log 2} + \lim_{n \to \infty} \frac{n \log_2 n \log 2}{n \log n \log 2} = 0$$

which follows by b) of Theorem 2.2 together with the Diamond-Vaaler trimmed law, $\log 2 \left(\sum_{k=1}^n a_k - \mathrm{M}_n \right) / (n \log n) \to 1$ a.e. and finally by a) of Corollary 2.1. $\qquad \square$

Proof of Theorem 2.4 **a)** We have that $\int (a')^\gamma\, d\lambda < \infty$ and the statement follows by the ergodic theorem.

b) We may apply [15, Theorem 1.7 & erratum]. That Property $\mathfrak{C}$ is fulfilled with the bounded variation norm $\|\cdot\|_{BV}$ is a standard result. For Property $\mathfrak{D}$, we notice that $\|a \cdot \mathbb{1}_{\{a \leq \ell\}}\|_{BV} \leq 2\ell$ and $\|\mathbb{1}_{\{a \leq \ell\}}\|_{BV} \leq 2$ implying that this property is fulfilled.

In order to calculate the norming sequence (d_n) we notice that

$$\mu_\mathfrak{G} \left((a')^\gamma > n \right) = \mu_\mathfrak{G} \left(a' > n^{1/\gamma} \right)$$

$$\sim \frac{1}{\log 2\, n^{1/\gamma} \log n^{1/\gamma}} = \frac{\gamma}{\log 2\, n^{1/\gamma} \log n} = \frac{L(n)}{n^{1/\gamma}},$$

where $L(n) = \gamma/(\log 2 \log n)$ is a slowly varying function.

Using then [15, Theorem 1.7 & erratum] we obtain that (2.12) holds for (b_n) fulfilling $b_n = o(n)$ and $\lim_{n \to \infty} b_n \log_2 n = \infty$ and for (d_n) fulfilling

$$d_n \sim \frac{1/\gamma}{1 - 1/\gamma} n^\gamma b_n^{1-\gamma} \left(L^{-\gamma} \right)^\# \left(\left(\frac{n}{b_n} \right)^\gamma \right),$$

where $\ell^\#$ denotes the de Bruijn conjugate of a slowly varying function ℓ, see e.g. [6] for a precise definition. In our case $\left(L^{-\gamma} \right)^\# (n) = ((\log n)^\gamma (\log 2)^\gamma / \gamma^\gamma)^\# \sim \gamma^\gamma / ((\log n)^\gamma (\log 2)^\gamma)$. Hence,

$$d_n \sim \frac{\gamma^\gamma}{(\gamma - 1)(\log 2)^\gamma} n^\gamma b_n^{1-\gamma} \frac{1}{(\log (n/b_n)^\gamma)^\gamma} \sim \frac{1}{(\gamma - 1)(\log 2)^\gamma} \cdot \frac{n^\gamma b_n^{1-\gamma}}{(\log n)^\gamma},$$

where the last assymptotic follows from the assumption $b_n = o(n^{1-\epsilon})$. $\qquad \square$

6 Proofs of the results on distributional convergence

We are now ready for the proofs of our distributional convergence results.

Proof of Theorem 3.1 In all cases we only need to check convergence in law w.r.t. $\mu_\mathfrak{G}$.

a) This follows directly from Theorem 2.4.

b) This statement follows directly by [1]. We use the expression for $L'(N)$ from (4.3) which is a slowly varying function. Since for $b_n = n \log_2 n / \log 2$ we have $n L(b_n) \sim b_n$ the statement follows.

c) This follows from [17, Theorem 1.8]. The conditions on the system and the asymptotic of the norming sequence (d_n) we have already considered in the proof of Theorem 2.4. □

Proof of Theorem 3.3 In each of the three statements it suffices to prove distributional convergence under the invariant measure $\mu_\mathfrak{G}$ (see Propositions 3.1 and 5.1 in [25]).

a) For $A \in \mathcal{B}_I$ with $\lambda(A) > 0$, the *(first) hitting time* function of A under the Gauss map S, $\varphi_A : I \to \overline{\mathbb{N}} := \{1, 2, \ldots, \infty\}$ is given by $\varphi_A(x) := \inf\{n \geq 1 : S^n x \in A\}$, which is finite a.e. on I. Define $S_A x := S^{\varphi_A(x)} x$ for a.e. $x \in I$, which gives the *first entrance map* $S_A : I \to A$. Letting $A'_l := \{a' \geq l\}$, $l \geq 1$, we see that $\varphi'_l = \varphi_{A'_l}$ and, more generally, $\varphi'_{l,i} = \varphi_{A'_l} \circ S^{i-1}_{A'_l}$ for $i \geq 1$. It is clear that A'_l is ξ-measurable, and according to Lemma 4.1, $\mu_\mathfrak{G}(A'_l) \sim (\log 2 \cdot l \log l)^{-1}$ as $l \to \infty$. Therefore, Theorem 10.2.a) of [25] immediately implies statement a).

b) This is a straightforward consequence of Theorem 10.2.b) in [25], because $\{a' \geq l/\vartheta\} = A'_{\lfloor l/\vartheta \rfloor}$ is ξ-measurable and (4.2) entails $\mu_\mathfrak{G}(a' \geq l/\vartheta) \sim \vartheta \, \mu_\mathfrak{G}(a' \geq l)$ as $l \to \infty$.

c) Let $\mathbb{P}(j) := \{p \in \mathbb{P} : p \equiv j \pmod{m}\}$, then Dirichlet's PNT for primes in residue classes (e.g. Theorem 4.4.4 of [14]) asserts that for each j relatively prime to m,

$$\# \left(\mathbb{P}(j) \cap \{2, \ldots, N\} \right) \sim \frac{1}{\phi(m)} \frac{N}{\log N} \quad \text{as } N \to \infty.$$

Via an easy argument parallel to the proof of (4.2), this shows that

$$\mu_\mathfrak{G}\left(A'_l \cap \{a' \equiv j \pmod{m}\} \right) \sim \frac{1}{\phi(m) \log 2} \cdot \frac{1}{l \log l} \quad \text{as } l \to \infty,$$

and hence $\mu_\mathfrak{G}(A'_l(j)) \sim \mu_\mathfrak{G}(A'_l)/\phi(m)$ with $A'_l(j) := A'_l \cap \{a' \equiv j \pmod{m}\}$ a ξ-measurable set. Another direct application of Theorem 10.2.b) in [25] then completes the proof of our theorem. □

The result thus established essentially contains (3.1).

Proof of Theorem 3.2 Theorem 3.3 a) contains the statement that $\mu_\mathfrak{G}(A'_l) \varphi'_{l,1}$ converges to a standard exponential law. Using the natural duality $\{M'_n < l\} = \{\varphi'_{l,1} \geq n\}$ this is easily seen to imply (3.1). □

Proof of Theorem 3.4 The result follows directly by [16, Theorem 3] by considering the sets $A_n = \{a_n \in \mathbb{P} \cap \Gamma_n\}$. The only thing to check is that

$$\sum_{n=1}^{\infty} \lambda(A_n) \cdot \lambda(A_n^c) = \infty. \tag{6.1}$$

We note that $\sum_{n=1}^{\infty} \lambda(A_n) \cdot \lambda(A_n^c) \geq \sum_{n=1}^{\infty} \lambda(A_n)$ and thus for (A) (6.1) follows from the proof of Theorem 2.1 a). (B) corresponds to [16, Theorem 5A] and for (C) (6.1) follows from the proof of Theorem 2.1 c). □

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

1. Aaronson, J.: Random f-expansions. Ann. Probab. **14**, 1037–1057 (1986). https://doi.org/10.1214/aop/1176992457
2. Aaronson, J., Nakada, H.: Trimmed sums for non-negative, mixing stationary processes. Stoch. Proc. Appl. **104**, 173–192 (2003). https://doi.org/10.1017/S0143385704001051
3. Aaronson, J., Thaler, M., Zweimüller, R.: Occupation times of sets of infinite measure for ergodic transformations. Ergod. Theory Dynam. Sys. **25**, 959–976 (2005). https://doi.org/10.1017/S0143385704001051
4. Baker, R.C., Harman, G., Pintz, J.: The difference between consecutive primes. II. Proc. Lond. Math. Soc. **83**, 532–562 (2001). https://doi.org/10.1112/plms/83.3.532
5. Bernstein, F.: Über eine Anwendung der Mengenlehre auf ein aus der Theorie der säkularen Störungen herrührendes Problem. Math. Ann. **71**, 417–439 (1911). https://doi.org/10.1007/BF01456856
6. Bingham, N.H., Goldie, C.M., Teugels, J.L.: Regular variation. Cambridge University Press, Cambridge (1989)
7. Borel, E.: Les probabilités dénomerables et leurs applications arithmétiques. Rend. Circ. Mat. Palermo **27**, 247–271 (1909). https://doi.org/10.1007/BF03019651
8. Diamond, H.G., Vaaler, J.D.: Estimates for partial sums of continued fraction partial quotients. Pacific J. Math. **122**, 73–82 (1986)
9. Doeblin, W.: Remarques sur la théorie mé trique des fractiones continues. Comp. Math. **7**, 353–371 (1940)
10. Gauss, C.F.: Letter to Laplace. Göttingen, January 30th 1812
11. Haeusler, E.: A nonstandard law of the iterated logarithm for trimmed sums. Ann. Probab. **21**, 831–860 (1993)
12. Hardy, G.H., Wright, E.M.: An introduction to the theory of numbers. Oxford University Press, London (1989)
13. Iosifescu, M., Kraaikamp, C.: Metrical theory of continued fractions. Kluwer (2002). https://doi.org/10.1007/978-94-015-9940-5
14. Jameson, G.J.O.: The prime number theorem. Cambridge University Press, Cambridge (2003)
15. Kesseböhmer, M., Schindler, T.I.: Strong laws of large numbers for intermediately trimmed Birkhoff sums of observables with infinite mean. Stoch. Proc. Appl. **129**, 4163–4207 (2019). https://doi.org/10.1016/j.spa.2018.11.015 Corrigendum in Stoch. Proc. Appl., 130 (2020), 7019. https://doi.org/10.1016/j.spa.2020.07.014
16. Kesseböhmer, M., Schindler, T.I.: Limit theorems for counting large continued fraction digits. Lith. Math. J. **60**, 189–207 (2020). https://doi.org/10.1007/s10986-020-09479-5
17. Kesseböhmer, M., Schindler, T.I.: Mean convergence for intermediately trimmed Birkhoff sums of observables with regularly varying tails. Nonlinearity **33**, 5543–5566 (2020). https://doi.org/10.1088/1361-6544/ab9585
18. Kesten, H.: Convergence in distribution of lightly trimmed and untrimmed sums are equivalent. Math. Proc. Cambridge Philos. Soc. **113**, 615–638 (1993). https://doi.org/10.1017/S0305004100076246

19. Khinchin, A.. Ya..: Metrische Kettenbruchprobleme. Comp. Math. **1**, 361–382 (1935)
20. Philipp, W.: Some metrical theorems in number theory II. Duke Math. J. **38**, 447–458 (1970). https://doi.org/10.1215/S0012-7094-70-03756-7
21. Philipp, W.: A conjecture of Erdös on continued fractions. Acta Arithmetica **28**, 379–386 (1976). https://doi.org/10.4064/aa-28-4-379-386
22. Philipp, W.: Limit theorems for sums of partial quotients of continued fractions. Monatsh. Math **105**, 195–206 (1988). https://doi.org/10.1007/BF01636928
23. Zweimüller, R.: Kuzmin, coupling, cones, and exponential mixing. Forum Math. **16**, 447–457 (2004). https://doi.org/10.1515/form.2004.021
24. Zweimüller, R.: Mixing limit theorems for ergodic transformations. J. Theor. Probab. **20**, 1059–1071 (2007). https://doi.org/10.1007/s10959-007-0085-y
25. Zweimüller, R.: Hitting times and positions in rare events. Ann. H. Lebesgue. **5**, 1361–1415 (2022). https://doi.org/10.5802/ahl.154