Journal of Number Theory 133 (2013) 2153–2181

Contents lists available at SciVerse ScienceDirect

Journal of Number Theory

www.elsevier.com/locate/jnt

A Gauss–Kuzmin-type problem for a family of continued fraction expansions

Dan Lascu

Mircea cel Batran Naval Academy, 1 Fulgerului, 900218 Constanta, Romania

ARTICLE INFO

Article history:
Received 27 September 2011
Revised 17 December 2012
Accepted 17 December 2012
Available online xxxx
Communicated by David Goss

Keywords:
Invariant measure
Perron–Frobenius operator
Random system with complete connections

ABSTRACT

In this paper we study in detail a family of continued fraction expansions of any number in the unit closed interval $[0, 1]$ whose digits are differences of consecutive non-positive integer powers of an integer $m \geqslant 2$. For the transformation which generates this expansion and its invariant measure, the Perron–Frobenius operator is given and studied. For this expansion, we apply the method of random systems with complete connections by Iosifescu and obtained the solution of its Gauss–Kuzmin type problem.
© 2013 Elsevier Inc. All rights reserved.

1. Introduction

The purpose of this paper is to prove a Gauss–Kuzmin type problem for non-regular continued fraction expansions introduced by Chan [5]. In order to solve the problem, we apply the random systems with complete connections by Iosifescu [10]. First we outline the historical framework of this problem. Then, in Section 1.2, we present the current framework. The main theorem will be shown in Section 1.3. In this subsection we will also give a detailed outline of the paper.

1.1. Gauss' Problem

One of the first and still one of the most important results in the metrical theory of continued fractions is so-called Gauss–Kuzmin theorem. Write $x \in [0, 1)$ as a regular continued fraction

E-mail address: lascudan@gmail.com.

0022-314X/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jnt.2012.12.007

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}} := [a_1, a_2, a_3, \ldots],$$

where $a_n \in \mathbb{N}_+ := \{1, 2, 3, \ldots\}$. The metrical theory of continued fractions started on 25th October 1800, with a note by Gauss in his mathematical diary. Gauss wrote that (in modern notation)

$$\lim_{n\to\infty} \lambda\big(\tau^n \leqslant x\big) = \frac{\log(1+x)}{\log 2}, \quad x \in I := [0,1].$$

Here λ is a Lebesgue measure and the map $\tau : [0,1) \to [0,1)$, the so-called *regular continued fraction* (or *Gauss*) transformation, is defined by

$$\tau(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor, \quad x \neq 0; \qquad \tau(0) := 0,$$

where $\lfloor\cdot\rfloor$ denotes the *floor* (or *entire*) function. Gauss' proof (if any) has never been found. A little more than 11 years later, in a letter dated 30 January 1812, Gauss asked Laplace to estimate the error

$$e_n(x) := \lambda\big(\tau^{-n}[0,x]\big) - \frac{\log(1+x)}{\log 2}, \quad n \geqslant 1,\ x \in I.$$

This has been called *Gauss' Problem*. It received a first solution more than a century later, when R.O. Kuzmin (see [16]) showed in 1928 that $e_n(x) = \mathcal{O}(q^{\sqrt{n}})$ as $n \to \infty$, uniformly in x with some (unspecified) $0 < q < 1$. One year later, using a different method, Paul Lévy (see [17]) improved Kuzmin's result by showing that $|e_n(x)| \leqslant q^n$, $n \in \mathbb{N}_+$, $x \in I$, with $q = 3.5 - 2\sqrt{2} = 0.67157\ldots$. The Gauss–Kuzmin–Lévy theorem is the first basic result in the rich metrical theory of continued fractions.

1.2. A non-regular continued fraction expansion

In this paper, we consider a generalization of the Gauss transformation and prove an analogous result.

In [5], Chan shows that any $x \in [0,1)$ can be written in the form

$$x = \cfrac{m^{-a_1(x)}}{1 + \cfrac{(m-1)m^{-a_2(x)}}{1 + \cfrac{(m-1)m^{-a_3(x)}}{1 + \ddots}}} := \big[a_1(x), a_2(x), a_3(x), \ldots\big]_m, \tag{1.1}$$

where $m \in \mathbb{N}_+$, $m \geqslant 2$ and $a_n(x)$s are non-negative integers.

For any $m \in \mathbb{N}_+$ with $m \geqslant 2$, define the transformation τ_m on I by

$$\tau_m(x) = \begin{cases} \frac{m^{\{\frac{\log x^{-1}}{\log m}\}} - 1}{m-1}, & \text{if } x \neq 0, \\ 0, & \text{if } x = 0, \end{cases} \tag{1.2}$$

where $\{\cdot\}$ stands for fractionary part. It is easy to see that τ_m maps the set Ω of irrationals in I into itself. For any $x \in (0,1)$ put

$$a_n = a_n(x) = a_1\left(\tau_m^{n-1}(x)\right), \quad n \in \mathbb{N}_+, \tag{1.3}$$

with $\tau_m^0(x) = x$ and

$$a_1 = a_1(x) = \begin{cases} \lfloor \log x^{-1} / \log m \rfloor, & \text{if } x \neq 0, \\ \infty, & \text{if } x = 0. \end{cases} \tag{1.4}$$

Transformation τ_m which generates the continued fraction expansion (1.1) is ergodic with respect to an invariant probability measure, γ_m, where

$$\gamma_m(A) = k_m \int_A \frac{dx}{((m-1)x+1)((m-1)x+m)}, \quad A \in \mathcal{B}_I,$$

with $k_m = \frac{(m-1)^2}{\log(m^2/(2m-1))}$ and $\mathcal{B}_I$ is the σ-algebra of Borel subsets of I (which, by definition, is the smallest σ-algebra containing intervals).

The ergodicity of τ_m plays a key role in the study of the asymptotic growth rate of the random Fibonacci type sequences $\{f_n\}$ defined by $f_{-1} = 0$, $f_0 = 1$, $c_0 = 0$ and

$$f_n = m^{c_n} f_{n-1} + (m-1)m^{c_{n-1}} f_{n-2}, \tag{1.5}$$

where c_n, $n \geqslant 1$, are the digits from (1.1). As is known, the Fibonacci sequence is defined using the linear recurrence relation

$$F_{n+1} = F_n + F_{n-1}, \quad n \in \mathbb{N}_+, \quad \text{with } F_0 = F_1 = 1,$$

and Binet's formula is

$$F_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{n+1} - \frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^{n+1}, \quad n \in \mathbb{N}.$$

It is known that using Binet's formula we can compute the asymptotic growth rate of the Fibonacci sequence $\{F_n\}$, which is given by

$$\lim_{n\to\infty} \frac{1}{n} \log F_n = \log\left(\frac{1+\sqrt{5}}{2}\right) = 0.4812\ldots.$$

In the case of random Fibonacci type sequences, defined by (with fixed f_1 and f_2)

$$f_n = \alpha(n) f_{n-1} + \beta(n) f_{n-2},$$

where $\alpha(n)$ and $\beta(n)$ are random coefficients, the quest for the asymptotic growth rate is more difficult. Recently, Viswanath (see [25]) proved that the asymptotic growth rate of the random Fibonacci sequences defined by $f_1 = f_2 = 1$ and

$$f_n = \pm f_{n-1} \pm f_{n-2},$$

where the signs are chosen independently and with equal probabilities, is given by

$$\lim_{n\to\infty} \frac{1}{n} \log f_n = \log(1.13198824\ldots) = 0.12397559\ldots$$

with probability 1. But Viswanath's method is not the only way through. So, Chan proved in [5] that for almost all x with respect to the Lebesgue measure, the asymptotic growth rate of $\{f_n\}$ from (1.5) is given by

$$\lim_{n\to\infty}\frac{1}{n}\log f_n = k_m\int_0^1 \frac{\log(1/x)}{((m-1)x+1)((m-1)x+m)}\,dx \leqslant k_m\frac{3m-1}{2m(2m-1)}. \qquad \square$$

1.3. Main theorem

We show our main theorem in this subsection. For this purpose let μ be a non-atomic probability measure on $\mathcal{B}_I$ and define

$$F_n(x)=\mu\left(\tau_m^n<x\right), \quad x\in I,\ n\in\mathbb{N},$$
$$F(x)=\lim_{n\to\infty}F_n(x), \quad x\in I,$$

with $F_0(x)=\mu([0,x))$.

Then the following holds.

Theorem 1.1 *(A Gauss–Kuzmin-type theorem). If μ has a Riemann-integrable density, then*

$$F(x)=\frac{k_m}{(m-1)^2}\log\frac{m((m-1)x+1)}{(m-1)x+m}, \quad x\in I, \tag{1.6}$$

where $k_m=\frac{(m-1)^2}{\log(m^2/(2m-1))}$.

If the density of μ is a Lipschitz function, then there exist two positive constants $q<1$ and k such that for all $x\in I$ and $n\in\mathbb{N}_+$ we have

$$\mu\left(\tau_m^n<x\right)=\frac{k_m}{(m-1)^2}\left(1+\theta q^n\right)\log\frac{m((m-1)x+1)}{(m-1)x+m}, \tag{1.7}$$

where θ is a certain constant determined by μ, n, x such that $|\theta|\leqslant k$.

The paper is organized as follows. In Section 2 we give the basic metric properties of the continued fraction expansion in (1.1). Hence, we give a Legendre-type result and the Brodén–Borel–Lévy formula used to determine the probability structure of $(a_n)_{n\in\mathbb{N}_+}$ under λ. In Section 2.4, we find the invariant measure of τ_m. The proof of this result is given in a different manner from that described by Chan in [5]. In Section 3 we consider the so-called *natural extension* $\overline{\tau}_m$ (see [19]), define extended incomplete quotients $\bar{a}_l$, $l\in\mathbb{Z}$, and we generalize some results presented in Section 2. In Section 4, we derive the associated Perron–Frobenius operator under different probability measures on $\mathcal{B}_I$. We study the Perron–Frobenius operator of τ_m under the invariant measure γ_m induced by the limit distribution function, we derive the asymptotic behavior of this operator and we restrict the Perron–Frobenius operator to the linear space of all complex-valued functions of bounded variation and to the space of all bounded measurable complex-valued functions. Section 5 is divided into three parts. The first subsection has as purpose defining the notion of random system with complete connections. In the second subsection we set up the necessary machinery to prove the main theorem whose proof is contained in the last subsection. To determine where $\mu(\tau_m^n<x)$ tends as $n\to\infty$ and give the rate of this convergence, we use the ergodic behavior of the random system with complete connections

associated with this expansion. For a more detailed study of the theory and applications of dependence with complete connections to the metrical problems and other interesting aspects of number theory we refer the reader to [10,12,13,22–24] and others.

2. Metric properties of the continued fraction expansions in (1.1)

Roughly speaking, the metrical theory of continued fraction expansions is about properties of the sequence $(a_n)_{n\in\mathbb{N}}$ and related sequences (see Section 3.2). The main purpose of this section is to determine the probability structure of $(a_n)_{n\in\mathbb{N}_+}$ under the Lebesgue measure λ. Before that, we shortly present the metrical theory of these continued fraction expansions. Another important result is the *Legendre-type theorem* (see, e.g., [3,11,15]) which is one of the main reasons for studying continued fractions, because it tells us that good approximations of irrational numbers by rational numbers are given by continued fraction convergents.

2.1. Some elementary properties of the continued fraction expansion in (1.1)

Here, we want to prove the convergence of expansion of the type of (1.1). First, note that in the rational case, the continued fraction expansion (1.1) is finite, unlike the irrational case, when we have an infinite number of non-negative digits.

Define $[a_1, a_2, \ldots, a_n]_m$ the convergent of $\omega \in \Omega$ by truncating the expansion on the right-hand side of (1.1). We want to show

$$\omega = \lim_{n\to\infty} [a_1, a_2, \ldots, a_n]_m, \quad \omega \in \Omega. \tag{2.1}$$

To this end, define integer-valued functions $p_n(\omega)$ and $q_n(\omega)$, for $n \in \mathbb{N}_+$, by

$$p_n(\omega) = m^{a_n} p_{n-1}(\omega) + (m-1)m^{a_{n-1}} p_{n-2}(\omega), \quad n \geqslant 2, \tag{2.2}$$

$$q_n(\omega) = m^{a_n} q_{n-1}(\omega) + (m-1)m^{a_{n-1}} q_{n-2}(\omega), \quad n \geqslant 1, \tag{2.3}$$

with $p_0(\omega) = 0$, $q_0(\omega) = 1$, $p_1(\omega) = 1$, $q_{-1}(\omega) = 0$ and $a_0 \equiv 0$.

Now, it is easy to prove by induction that for any $n \in \mathbb{N}_+$ we have

$$p_n(\omega)q_{n-1}(\omega) - p_{n-1}(\omega)q_n(\omega) = (-1)^{n-1}(m-1)^{n-1} m^{a_1+\cdots+a_{n-1}}, \tag{2.4}$$

and

$$\cfrac{m^{-a_1}}{1 + \cfrac{(m-1)m^{-a_2}}{1 + \ddots + \cfrac{(m-1)m^{-a_n}}{1 + (m-1)t}}} = \frac{p_n(\omega) + (m-1)tm^{a_n} p_{n-1}(\omega)}{q_n(\omega) + (m-1)tm^{a_n} q_{n-1}(\omega)}, \tag{2.5}$$

with $0 \leqslant t \leqslant 1$.

It follows from the definitions of τ_m and a_n that for any $\omega \in \Omega$ we have

$$\tau_m^{n-1}(\omega) = \frac{m^{-a_n}}{1 + (m-1)\tau_m^n(\omega)}, \quad n \in \mathbb{N}_+, \tag{2.6}$$

hence

$$\omega = \cfrac{m^{-a_1}}{1 + \cfrac{(m-1)m^{-a_2}}{1 + \ddots + \cfrac{(m-1)m^{-a_n}}{1 + (m-1)\tau_m^n(\omega)}}}, \quad n \in \mathbb{N}_+. \tag{2.7}$$

By combining (2.7), (2.2) and (2.3) we have

$$\omega = \frac{p_n(\omega) + (m-1)\tau_m^n(\omega)m^{a_n}p_{n-1}(\omega)}{q_n(\omega) + (m-1)\tau_m^n(\omega)m^{a_n}q_{n-1}(\omega)}, \quad \omega \in \Omega,\ n \in \mathbb{N}_+. \tag{2.8}$$

Taking $\tau_m^n(\omega) = 0$ in (2.8) gives

$$[a_1, a_2, \ldots, a_n]_m = \frac{p_n(\omega)}{q_n(\omega)}. \tag{2.9}$$

Now, using (2.4), (2.8) and (2.9), for any $\omega \in \Omega$ we obtain

$$\left|\omega - \frac{p_n(\omega)}{q_n(\omega)}\right| = \frac{(m-1)^n \tau_m^n(\omega) m^{a_1+\cdots+a_n}}{q_n(\omega)(q_n(\omega) + (m-1)\tau_m^n(\omega)m^{a_n}q_{n-1}(\omega))}, \quad n \in \mathbb{N}_+. \tag{2.10}$$

Note that this equation measures the difference between $\omega \in \Omega$ and its convergent and is the key ingredient of the following estimate.

Lemma 2.1. *For any $\omega \in \Omega$ we have*

$$\left|\omega - \frac{p_n(\omega)}{q_n(\omega)}\right| \leqslant \left(\frac{m-1}{m}\right)^n, \quad n \in \mathbb{N}_+. \tag{2.11}$$

Proof. By applying $\tau_m^n(\omega) \leqslant 1$ to (2.10), we have

$$\left|\omega - \frac{p_n(\omega)}{q_n(\omega)}\right| \leqslant \frac{(m-1)^n m^{a_1+\cdots+a_n}}{q_n(\omega)(q_n(\omega) + (m-1)m^{a_n}q_{n-1}(\omega))}, \quad n \in \mathbb{N}_+. \tag{2.12}$$

Let

$$t_n := \frac{(m-1)^n m^{a_1+\cdots+a_n}}{q_n(\omega)(q_n(\omega) + (m-1)m^{a_n}q_{n-1}(\omega))}, \quad n \in \mathbb{N}_+. \tag{2.13}$$

From (2.3), we have that $q_n(\omega) + (m-1)m^{a_n}q_{n-1}(\omega) \geqslant m \cdot m^{a_n}q_{n-1}(\omega)$, i.e., $q_n(\omega) \geqslant m^{a_n}q_{n-1}(\omega)$. Thus, by (2.13) and since $q_n(\omega) \geqslant q_{n-1}(\omega) + (m-1)m^{a_{n-1}}q_{n-2}(\omega)$, we have

$$\begin{aligned} t_n &\leqslant \frac{m-1}{m}\left(\frac{(m-1)^{n-1}m^{a_1+\cdots+a_{n-1}}}{q_n(\omega)q_{n-1}(\omega)}\right) \\ &\leqslant \frac{m-1}{m}\left(\frac{(m-1)^{n-1}m^{a_1+\cdots+a_{n-1}}}{q_{n-1}(\omega)(q_{n-1}(\omega) + (m-1)m^{a_{n-1}}q_{n-2}(\omega))}\right) \\ &= \frac{m-1}{m}t_{n-1}. \end{aligned} \tag{2.14}$$

Now, by direct computation, we have

$$t_1 \leqslant \frac{m-1}{m} m^{-a_1} \leqslant \frac{m-1}{m}$$

and (2.14) shows that $t_n \leqslant (\frac{m-1}{m})^n$, i.e., (2.11). $\square$

Finally, (2.1) follows from (2.11), as $\frac{m-1}{m} < 1$.

2.2. Approximation result

Diophantine approximation (see, e.g., [15]) deals with the approximation of real numbers by rational numbers. Before we give the corresponding approximation result, we define the *cylinder* (or *fundamental interval*) of rank n, $I_m(i^{(n)})$, and show that any $I_m(i^{(n)})$ is the set of irrationals from a certain open interval with rational endpoints.

For any $n \in \mathbb{N}_+$ and $i^{(n)} = (i_1, \ldots, i_n) \in \mathbb{N}^n$ we will say that

$$I_m\big(i^{(n)}\big) = \big\{\omega \in \Omega\colon\ a_k(\omega) = i_k,\ 1 \leqslant k \leqslant n\big\} \tag{2.15}$$

is the *fundamental interval of rank n* and make the convention that $I_m(i^{(0)}) = \Omega$.

For example, for any $i \in \mathbb{N}$ we have

$$I_m(i) = \big\{\omega \in \Omega\colon\ a_1(\omega) = i\big\} = \Omega \cap \big(m^{-(i+1)}, m^{-i}\big). \tag{2.16}$$

We will write $I_m(a_1, \ldots, a_n) = I_m(a^{(n)})$, $n \in \mathbb{N}_+$. If $n \geqslant 2$ and $i_n \in \mathbb{N}$, then we have

$$I_m(a_1, \ldots, a_n) = I_m\big(i^{(n)}\big).$$

From the definition of τ_m and (2.8) we have

$$I_m\big(a^{(n)}\big) = \Omega \cap \big(u\big(a^{(n)}\big), v\big(a^{(n)}\big)\big), \tag{2.17}$$

where

$$u\big(a^{(n)}\big) = \begin{cases} \frac{p_n(\omega)+(m-1)m^{a_n} p_{n-1}(\omega)}{q_n(\omega)+(m-1)m^{a_n} q_{n-1}(\omega)}, & \text{if } n \text{ is odd,} \\ \frac{p_n(\omega)}{q_n(\omega)}, & \text{if } n \text{ is even,} \end{cases} \tag{2.18}$$

and

$$v\big(a^{(n)}\big) = \begin{cases} \frac{p_n(\omega)}{q_n(\omega)}, & \text{if } n \text{ is odd,} \\ \frac{p_n(\omega)+(m-1)m^{a_n} p_{n-1}(\omega)}{q_n(\omega)+(m-1)m^{a_n} q_{n-1}(\omega)}, & \text{if } n \text{ is even.} \end{cases} \tag{2.19}$$

Now, using (2.4), a direct computation shows that

$$\lambda\big(I\big(a^{(n)}\big)\big) = \frac{(m-1)^n m^{a_1+\cdots+a_n}}{q_n(\omega)(q_n(\omega) + (m-1)m^{a_n} q_{n-1}(\omega))} \tag{2.20}$$

and from (2.10) and (2.11) we have that

$$\lambda\big(I\big(a^{(n)}\big)\big) \leqslant \left(\frac{m-1}{m}\right)^n. \tag{2.21}$$

We now give a Legendre-type result for these continued fraction expansions. First we define the *approximation coefficient* $\Theta_m := \Theta_m(\omega)$ by

$$\Theta_m := q^2\left|\omega - \frac{p_n}{q_n}\right|, \quad n \in \mathbb{N}_+,$$

where $\frac{p_n}{q_n}$ is the nth continued fraction convergent of $\omega \in \Omega$. The approximation coefficient gives a numerical indication of the quality of the approximation.

Proposition 2.2. *For $\omega \in \Omega$ and p/q being a rational number with $p < q$, $q > 0$ and* g.c.d.$(p, q) = 1$, *let*

$$\frac{p}{q} = [i_1, \ldots, i_n]_m, \qquad \frac{p_{n-1}}{q_{n-1}} = [i_1, \ldots, i_{n-1}]_m$$

with $p_0 = 0$ and $q_0 = 1$, where the length $n = n(p/q) \in \mathbb{N}_+$ of the continued fraction expansion of p/q is chosen in such a way that it is even if $p/q < \omega$ and odd otherwise. Then

$$\Theta_m < \frac{(m-1)^n m^{i_1+\cdots+i_n} q}{q + (m-1)m^{i_n} q_{n-1}} \quad \text{if and only if} \quad \frac{p}{q} \quad \text{is a convergent of } \omega.$$

Proof. If p/q is a convergent of ω, then by (2.10) we have

$$\Theta_m = q^2\left|\omega - \frac{p}{q}\right| = \frac{(m-1)^n \tau_m^n(\omega) m^{i_1+\cdots+i_n} q}{q + (m-1)\tau_m^n(\omega) m^{i_n} q_{n-1}(\omega)} \leqslant \frac{(m-1)^n m^{i_1+\cdots+i_n} q}{q + (m-1)m^{i_n} q_{n-1}}.$$

Conversely, if $\Theta_m < \frac{(m-1)^n m^{i_1+\cdots+i_n} q}{q+(m-1)m^{i_n} q_{n-1}}$, then

$$q\left|\omega - \frac{p}{q}\right| < \frac{(m-1)^n m^{i_1+\cdots+i_n}}{q + (m-1)m^{i_n} q_{n-1}}.$$

Assuming that n is even, then $\omega > \frac{p}{q}$ and we have $\omega - \frac{p}{q} < \frac{(m-1)^n m^{i_1+\cdots+i_n}}{q(q+(m-1)m^{i_n} q_{n-1})}$. Thus,

$$\frac{p}{q} < \omega < \frac{p}{q} + \frac{(m-1)^n m^{i_1+\cdots+i_n}}{q(q + (m-1)m^{i_n} q_{n-1})} = \frac{p + (m-1)m^{i_n} p_{n-1}}{q + (m-1)m^{i_n} q_{n-1}}.$$

Hence, $\omega \in I_m(i^{(n)})$, i.e., $\frac{p}{q} = [i_1, \ldots, i_n]_m$ is a convergent of ω. The case when n is an odd is treated similarly. $\square$

2.3. *The probability structure of* $(a_n)_{n\in\mathbb{N}_+}$ *under the* λ

We start by deriving the so-called Brodén–Borel–Lévy formula (see, e.g., [10,11]) for these type of expansions. First, define s_n, $n \in \mathbb{N}_+$, by

$$s_n = m^{-a_n}\frac{q_n}{q_{n-1}} - 1, \qquad s_1 = 0, \tag{2.22}$$

where $m \geqslant 2$ and a_n, q_n are defined in (1.3) and (2.3), respectively.

Next, (2.3) implies that

$$s_n = \frac{(m-1)m^{-a_n}}{1+s_{n-1}}, \quad n \geqslant 2, \tag{2.23}$$

hence

$$s_n = \cfrac{(m-1)m^{-a_n}}{1+\cfrac{(m-1)m^{-a_{n-1}}}{1+\ddots+\cfrac{(m-1)m^{-a_3}}{1+(m-1)m^{-a_2}}}} = (m-1)[a_n, a_{n-1}, \ldots, a_2, \infty]_m, \tag{2.24}$$

for $n \geqslant 2$.

Proposition 2.3 *(Brodén–Borel–Lévy formula type). For any $n \in \mathbb{N}_+$ we have*

$$\lambda\left(\tau_m^n < x \,\middle|\, a_1, \ldots, a_n\right) = \frac{(s_n+m)x}{(s_n+(m-1)x+1)}, \quad x \in I, \tag{2.25}$$

where s_n is defined by (2.22) or (2.23).

Proof. As we know, for any $n \in \mathbb{N}_+$ and $x \in I$, we have

$$\lambda\left(\tau_m^n < x \,\middle|\, a_1, \ldots, a_n\right) = \frac{\lambda((\tau_m^n < x) \cap I_m(a_1, \ldots, a_n))}{\lambda(I_m(a_1, \ldots, a_n))}.$$

From (2.8) and (2.17) we have

$$\begin{aligned}\lambda\left(\left(\tau_m^n < x\right) \cap I(a_1, \ldots, a_n)\right) &= \left|\frac{p_n}{q_n} - \frac{p_n + (m-1)xm^{a_n}p_{n-1}}{q_n + (m-1)xm^{a_n}q_{n-1}}\right| \\ &= \frac{(m-1)^n x m^{a_1+\cdots+a_n}}{q_n(q_n + (m-1)xm^{a_n}q_{n-1})}.\end{aligned}$$

Hence, from (2.20) we have

$$\begin{aligned}\lambda\left(\tau_m^n < x \,\middle|\, a_1, \ldots, a_n\right) &= \frac{\lambda((\tau_m^n < x) \cap I_m(a_1, \ldots, a_n))}{\lambda(I_m(a_1, \ldots, a_n))} \\ &= \frac{x(q_n + (m-1)m^{a_n}q_{n-1})}{q_n(q_n + (m-1)xm^{a_n}q_{n-1})} \\ &= \frac{(s_n+m)x}{s_n+(m-1)x+1},\end{aligned}$$

for any $n \in \mathbb{N}_+$ and $x \in I$. $\square$

The Brodén–Borel–Lévy formula allows us to determine the probability structure of $(a_n)_{n\in\mathbb{N}_+}$ under λ.

Proposition 2.4. *For any $i \in \mathbb{N}$ and $n \in \mathbb{N}_+$ we have*

$$\lambda(a_1 = i) = (m-1)m^{-(i+1)} \tag{2.26}$$

and

$$\lambda(a_{n+1} = i \mid a_1, \ldots, a_n) = P_m^i(s_n), \tag{2.27}$$

where

$$P_m^i(x) = \frac{(m-1)m^{-(i+1)}(x+1)(x+m)}{(x+(m-1)m^{-i}+1)(x+(m-1)m^{-(i+1)}+1)}. \tag{2.28}$$

Proof. As shown above, we have

$$\{\omega \in \Omega\colon\, a_1(\omega) = i\} = \Omega \cap \left(m^{-(i+1)}, m^{-i}\right).$$

Thus,

$$\lambda(a_1 = i) = \left|m^{-(i+1)} - m^{-i}\right| = (m-1)m^{-(i+1)}.$$

From (2.6), we have that

$$\tau_m^n(\omega) = [a_{n+1}, a_{n+2}, \ldots]_m, \quad n \in \mathbb{N}_+,\ \omega \in \Omega,$$

and so we have

$$\begin{aligned}
\lambda(a_{n+1} = i \mid a_1, \ldots, a_n) &= \lambda\left(\tau_m^n \in \left(m^{-(i+1)}, m^{-i}\right] \,\middle|\, a_1, \ldots, a_n\right) \\
&= \frac{(s_n+m)m^{-i}}{s_n+(m-1)m^{-i}+1} - \frac{(s_n+m)m^{-(i+1)}}{s_n+(m-1)m^{-(i+1)}+1} \\
&= \frac{(m-1)m^{-(i+1)}(s_n+1)(s_n+m)}{(s_n+(m-1)m^{-i}+1)(s_n+(m-1)m^{-(i+1)}+1)} \\
&= P_m^i(s_n). \qquad \square
\end{aligned}$$

Hence, the sequence $(s_n)_{n\in\mathbb{N}_+}$ with $s_1 = 0$ is a homogeneous I-valued Markov chain on $(I, \mathcal{B}_I, \lambda)$ with the following transition mechanism: from state $s \in I \setminus \Omega$, $s \geqslant 1$ the only possible one-step transitions are those to states $m^{-i}/(1+(m-1)s)$, $i \in \mathbb{N}$, with corresponding probabilities $P_m^i(s)$, $i \in \mathbb{N}$.

2.4. *The invariant measure of τ_m*

In this subsection we will give the explicit form of the invariant probability measure γ_m of the transformation τ_m, i.e., $\gamma_m(A) = \gamma_m(\tau_m^{-1}(A))$, $A \in \mathcal{B}_I$.

Let $\mathcal{B}_I$ denote the σ-algebra of Borel subsets of I. The metric point of view in studying the sequence $(a_n)_{n\in\mathbb{N}_+}$ is to consider that the a_n, $n \in \mathbb{N}_+$, are non-negative integer-valued random variables which are defined almost surely on $(I, \mathcal{B}_I)$ with respect to any probability measure on $\mathcal{B}_I$ that assigns probability 0 to the set $I \setminus \Omega$ of rationals in I. Such a measure is a Lebesgue measure λ.

Another measure on $\mathcal{B}_I$ more important than Lebesgue measure, that assigns probability 0 to the set of rationals in I, is the *invariant probability measure* γ_m of the transformation τ_m.

Proposition 2.5. *The invariant probability density ρ_m of the transformation τ_m is given by*

$$\rho_m(x) = \frac{1}{((m-1)x+1)((m-1)x+m)}, \quad x \in I, \tag{2.29}$$

with the normalizing factor $k_m = \frac{(m-1)^2}{\log(m^2/(2m-1))}$.

Proof. See Appendix A. □

Hence

$$\gamma_m(A) = k_m \int_A \frac{dx}{((m-1)x+1)((m-1)x+m)}, \quad A \in \mathcal{B}_I. \tag{2.30}$$

The normalization constant k_m defined above is chosen so that $\gamma_m([0,1]) = 1$.

3. The natural extension of τ_m and extended random variables

By its very definition, the sequence $(a_n)_{n\in\mathbb{N}_+}$ in (1.3) and (1.4) is strictly stationary under γ_m. As such, there should exist a doubly infinite version of it, say $\bar{a}_l$, $l \in \mathbb{Z} := \{\ldots, -1, 0, 1, \ldots\}$, defined on a richer probability space. It appears that this doubly infinite version can be effectively constructed on $(I^2, \mathcal{B}_I^2, \overline{\gamma}_m)$, where $\overline{\gamma}_m$ is the so-called extended measure which expression is given below.

3.1. *Definition and basic properties*

For τ_m in (1.2), the *natural extension* $\overline{\tau}_m$ of τ_m [19] is the transformation of $[0,1) \times I$ defined by

$$\overline{\tau}_m(x, y) = \left(\tau_m(x), \frac{m^{-a_1(x)}}{(m-1)y+1}\right), \quad (x, y) \in [0,1) \times I. \tag{3.1}$$

This is a one-to-one transformation of Ω^2 with the inverse

$$\overline{\tau}_m^{-1}(\omega, \theta) = \left(\frac{m^{-a_1(\theta)}}{(m-1)\omega+1}, \tau_m(\theta)\right), \quad (\omega, \theta) \in \Omega^2. \tag{3.2}$$

It is easy to check that for $n \geqslant 2$ we have

$$\overline{\tau}_m^n(\omega, \theta) = \left(\tau_m^n(\omega), \left[a_n(\omega), a_{n-1}(\omega), \ldots, a_2(\omega), a_1(\omega) + \frac{\log(1+(m-1)\theta)}{\log m}\right]_m\right), \tag{3.3}$$

and

$$\overline{\tau}_m^{-n}(\omega,\theta)=\left(\left[a_n(\theta),a_{n-1}(\theta),\ldots,a_2(\theta),a_1(\theta)+\frac{\log(1+(m-1)\omega)}{\log m}\right]_m,\tau_m^n(\theta)\right). \tag{3.4}$$

Now, define the *extended measure* $\overline{\gamma}_m$ on $\mathcal{B}_I^2$ as

$$\overline{\gamma}_m(B)=k_m\iint\limits_B\frac{dx\,dy}{((m-1)(x+y)+1)^2},\quad B\in\mathcal{B}_I^2. \tag{3.5}$$

A simple calculus shows us that

$$\overline{\gamma}_m(A\times I)=\overline{\gamma}_m(I\times A)=\gamma_m(A),\quad A\in\mathcal{B}_I. \tag{3.6}$$

The result below shows that $\overline{\gamma}_m$ plays with respect to $\overline{\tau}_m$ the part played by γ_m with respect to τ_m.

Proposition 3.1. *The extended measure $\overline{\gamma}_m$ is preserved by $\overline{\tau}_m$.*

Proof. See Appendix A. □

3.2. Extended random variables

Define extended incomplete quotients $\bar{a}_l$, $l\in\mathbb{Z}$, on Ω^2 by

$$\bar{a}_{l+1}(\omega,\theta)=\bar{a}_1\big(\overline{\tau}_m^l(\omega,\theta)\big),\quad l\in\mathbb{Z},$$

with

$$\bar{a}_1(\omega,\theta)=a_1(\omega),\quad(\omega,\theta)\in\Omega^2.$$

By (3.3) and (3.4) we have

$$\bar{a}_n(\omega,\theta)=a_n(\omega),\qquad\bar{a}_0(\omega,\theta)=a_1(\theta),\qquad\bar{a}_{-n}(\omega,\theta)=a_{n+1}(\theta),\quad n\in\mathbb{N}_+,\ (\omega,\theta)\in\Omega^2.$$

Remark 3.2. Since $\overline{\tau}_m$ preserves $\overline{\gamma}_m$, the doubly infinite sequence $(\bar{a}_l)_{l\in\mathbb{Z}}$ is strictly stationary under $\overline{\gamma}_m$.

Theorem 3.3. *For any $x\in I$ we have*

$$\overline{\gamma}_m\big([0,x]\times I\,\big|\,\bar{a}_0,\bar{a}_{-1},\ldots\big)=\frac{((m-1)a+m)x}{(m-1)(x+a)+1}\quad\overline{\gamma}_m\text{-a.s.}, \tag{3.7}$$

where $a=[\bar{a}_0,\bar{a}_{-1},\ldots]_m$.

Proof. Let $I_{m,n}$ denote the fundamental interval $I_m(\bar{a}_0,\bar{a}_{-1},\ldots,\bar{a}_{-n})$, $n\in\mathbb{N}$. We have

$$\overline{\gamma}_m\big([0,x]\times I\,\big|\,\bar{a}_0,\bar{a}_{-1},\ldots\big)=\lim_{n\to\infty}\overline{\gamma}_m\big([0,x]\times I\,\big|\,\bar{a}_0,\ldots,\bar{a}_{-n}\big)\quad\overline{\gamma}_m\text{-a.s.}$$

and

$$
\begin{aligned}
\overline{\gamma}_m\big([0,x]\times I \,\big|\, \bar{a}_0,\dots,\bar{a}_{-n}\big) &= \frac{\overline{\gamma}_m([0,x]\times I_{m,n})}{\overline{\gamma}_m(I\times I_{m,n})} \\
&= \frac{k_m \int_{I_{m,n}} dy \int_0^x \frac{du}{((m-1)(u+y)+1)^2}}{\gamma_m(I_{m,n})} \\
&= \frac{1}{\gamma_m(I_{m,n})} k_m \int\limits_{I_{m,n}} \frac{x}{((m-1)(x+y)+1)((m-1)y+1)}\,dy \\
&= \frac{1}{\gamma_m(I_{m,n})} \int\limits_{I_{m,n}} \frac{x((m-1)y+m)}{(m-1)(x+y)+1}\gamma_m(dy) \\
&= \frac{x((m-1)y_n+m)}{(m-1)(x+y_n)+1},
\end{aligned}
$$

for some $y_n \in I_{m,n}$. Since

$$
\lim_{n\to\infty} y_n = [\bar{a}_0,\bar{a}_{-1},\dots]_m = a, \tag{3.8}
$$

the proof is complete. □

The stochastic property of $(\bar{a}_l)_{l\in\mathbb{Z}}$ under $\overline{\gamma}_m$ is given by the following corollary of Theorem 3.3.

Corollary 3.4. *For any $i \in \mathbb{N}$ we have*

$$
\overline{\gamma}_m(\bar{a}_1 = i \mid \bar{a}_0,\bar{a}_{-1},\dots) = P_m^i\big((m-1)a\big) \quad \overline{\gamma}_m\text{-a.s.},
$$

where $a = [\bar{a}_0,\bar{a}_{-1},\dots]_m$.

Proof. Let us denote by $I_{m,n}$ the fundamental interval $I_m(\bar{a}_0,\bar{a}_{-1},\dots,\bar{a}_{-n})$, $n \in \mathbb{N}$. We have

$$
(\bar{a}_1 = i) = \big(m^{-(i+1)}, m^{-i}\big] \times [0,1)
$$

and

$$
\overline{\gamma}_m(\bar{a}_1 = i \mid \bar{a}_0,\bar{a}_{-1},\dots) = \lim_{n\to\infty} \overline{\gamma}_m(\bar{a}_1 = i \mid I_{m,n}).
$$

Now

$$
\begin{aligned}
\overline{\gamma}_m\big(\big(m^{-(i+1)}, m^{-i}\big) \times [0,1) \,\big|\, I_{m,n}\big) &= \frac{\overline{\gamma}_m((m^{-(i+1)}, m^{-i}) \times I_{m,n})}{\overline{\gamma}_m(I\times I_{m,n})} \\
&= \frac{1}{\gamma_m(I_{m,n})} \int\limits_{I_n} P_m^i\big((m-1)y\big)\gamma_m(dy) \\
&= P_m^i\big((m-1)y_n\big),
\end{aligned}
$$

for some $y_n \in I_{m,n}$. From (3.8) the proof is complete. □

Remark 3.5. The strict stationarity of $(\bar{a}_l)_{l\in\mathbb{Z}}$, under $\overline{\gamma}_m$ implies that

$$\overline{\gamma}_m(\bar{a}_{l+1} = i \mid \bar{a}_l, \bar{a}_{l-1}, \ldots) = P^i_m\big((m-1)a\big) \quad \overline{\gamma}_m\text{-a.s.}$$

for any $i \in \mathbb{N}$ and $l \in \mathbb{Z}$, where $a = [\bar{a}_l, \bar{a}_{l-1}, \ldots]_m$. The last equation emphasizes that $(\bar{a}_l)_{l\in\mathbb{Z}}$ is a chain of infinite order in the theory of dependence with complete connections (see [10, Section 5.5]).

Motivated by Theorem 3.3 we shall consider the family of (conditional) probability measures $(\gamma^a_m)_a$ on $\mathcal{B}_I$ defined by their distribution functions

$$\gamma^a_m\big([0,x]\big) = \frac{((m-1)a+m)x}{(m-1)(x+a)+1}, \quad x \in I,\ a \geqslant 0. \tag{3.9}$$

Note that the limit case $a = \infty$ is $\gamma^\infty_m = \lambda$.

For any $a \geqslant 0$ put $s^a_0 = a$ and

$$s^a_n = \frac{(m-1)m^{-a_n}}{1+s^a_{n-1}}, \quad n \in \mathbb{N}_+. \tag{3.10}$$

For $a \geqslant 0$ we have

$$s^a_1 = \frac{(m-1)m^{-a_1}}{1+a}$$

and

$$s^a_n = (m-1)\left[a_n, \ldots, a_2, a_1 + \frac{\log(a+1)}{\log m}\right]_m, \quad n \geqslant 2.$$

Then $(s^a_n)_{n\in\mathbf{N}_+}$ is an $I \cup \{a\}$-valued Markov chain on $(I, \mathcal{B}_I, \gamma^a_m)$ which starts from $s^a_0 = a \geqslant 0$ and has the following transition mechanism: from state $s \in I \cup \{a\}$ the possible transitions are to any state $m^{-i}/((m-1)s+1)$ with the corresponding transition probability $P^i_m((m-1)s)$, $i \in \mathbb{N}$.

Now, it is easy to check by induction that

$$s^a_n = m^{-a_n}\frac{(m-1)p_n + (a+1)q_n}{(m-1)p_{n-1} + (a+1)q_{n-1}} - 1, \tag{3.11}$$

for any $n \in \mathbb{N}_+$ and $a \geqslant 0$.

Thus, a simple calculation shows that for any $n \in \mathbb{N}_+$ we have

$$\begin{aligned}\gamma^a_m\big(\tau^n_m < x \,\big|\, a_1, \ldots, a_n\big) &= \frac{\gamma^a_m((\tau^n_m < x) \cap I_m(a^{(n)}))}{\gamma^a_m(I_m(a^{(n)}))} \\ &= \frac{x((m-1)((m-1)p_n + (a+1)q_n) + m^{a_n}((m-1)p_{n-1} + (a+1)q_{n-1}))}{(m-1)((m-1)p_n + (a+1)q_n) + xm^{a_n}((m-1)p_{n-1} + (a+1)q_{n-1})}.\end{aligned}$$

By (3.11) for any $n \in \mathbb{N}_+$ we have

$$\gamma^a_m\big(\tau^n_m < x \,\big|\, a_1, \ldots, a_n\big) = \frac{((m-1)s^a_n + m)x}{(m-1)(x+s^a_n)+1}, \quad a \geqslant 0,\ x \in I. \tag{3.12}$$

The last equation is the generalization of the Brodén–Borel–Lévy formula from Section 2.3.

4. The Perron–Frobenius operator of τ_m under γ_m

In this section we derive and study the associated Perron–Frobenius operator of τ_m under the invariant measure γ_m.

Let μ be a probability measure on $(I, \mathcal{B}_I)$ such that $\mu(\tau_m^{-1}(A)) = 0$ whenever $\mu(A) = 0$, $A \in \mathcal{B}_I$, where the transformation τ_m is defined in (1.2). In particular, this condition is satisfied if τ_m is μ-preserving, that is, $\mu\tau_m^{-1} = \mu$. It is known from previous section, that the Perron–Frobenius operator P_μ of τ_m under μ is defined as the bounded linear operator on $L^1_\mu = \{f : I \to \mathbb{C} \mid \int_I |f|\, d\mu < \infty\}$ which takes $f \in L^1_\mu$ into $P_\mu f \in L^1_\mu$ with

$$\int_A P_\mu f \, d\mu = \int_{\tau_m^{-1}(A)} f \, d\mu, \quad A \in \mathcal{B}_I. \tag{4.1}$$

In particular, the Perron–Frobenius operator P_λ of τ_m under the Lebesgue measure λ is (see [4, p. 86])

$$P_\lambda f(x) = \frac{d}{dx} \int_{\tau_m^{-1}([0,x])} f \, d\lambda = \sum_{t \in \tau_m^{-1}(x)} \frac{f(t)}{|\tau_m'(t)|} \quad \text{a.e. in } I. \tag{4.2}$$

The following results will be proved in Appendix A.

The following proposition gives the expression of the Perron–Frobenius operator of τ_m under the invariant measure γ_m (4.3) and under a probability measure which is absolutely continuous with respect to the Lebesgue measure (4.6). Also, we derive the asymptotic behavior of this operator (4.8).

Proposition 4.1.

(i) *The Perron–Frobenius operator $U_m := P_{\gamma_m}$ of τ_m under γ_m is given a.e. in I by the equation*

$$U_m f(x) = \sum_{i \in \mathbb{N}} P_m^i\big((m-1)x\big) f\big(u_m^i(x)\big), \quad f \in L^1_{\gamma_m}, \tag{4.3}$$

where P_m^i is defined in (2.28) and $u_m^i(x)$ is given by the equation

$$u_m^i(x) = \frac{m^{-i}}{(m-1)x + 1}, \quad x \in I. \tag{4.4}$$

(ii) *Let μ be a probability measure on $\mathcal{B}_I$. Assume that μ is absolutely continuous with respect to λ (and denote $\mu \ll \lambda$, i.e., if $\mu(A) = 0$ for every set A with $\lambda(A) = 0$) and let $h = d\mu/d\lambda$ a.e. in I. Then:*

(a) *The Perron–Frobenius operator P_μ of τ_m under μ is given a.e. in I by the equation*

$$P_\mu f(x) = \frac{1}{h(x)} \sum_{i \in \mathbb{N}} \frac{h(u_m^i(x))}{((m-1)x+1)^2} (m-1)m^{-i} f\big(u_m^i(x)\big) \tag{4.5}$$

$$= \frac{U_m g(x)}{((m-1)x+1)((m-1)x+m)h(x)}, \quad f \in L^1_\mu, \tag{4.6}$$

where $g(x) = ((m-1)x+1)((m-1)x+m) f(x) h(x)$, $x \in I$.
The powers of P_μ are given a.e. in I and for any $f \in L^1_\mu$ and any $n \in \mathbb{N}_+$ by the equation

$$P_\mu^n f(x) = \frac{U_m^n g(x)}{((m-1)x+1)((m-1)x+m)h(x)}. \tag{4.7}$$

(b) *We have*

$$\mu\big(\tau_m^{-n}(A)\big) = \int\limits_A \frac{U_m^n f(x)}{((m-1)x+1)((m-1)x+m)}\,dx, \tag{4.8}$$

for any $n \in \mathbb{N}$ *and* $A \in \mathcal{B}_I$, *where* $f(x) = ((m-1)x+1)((m-1)x+m)h(x)$, $x \in I$.

In the next proposition the domain of U_m will be successively restricted to the following Banach spaces: $BV(I)$ is the linear space of all complex-valued functions of bounded variation and $B(I)$ is the collection of all bounded measurable functions $f : I \to \mathbb{C}$. The variation $\operatorname{var}_A f$ over $A \subset I$ of a function $f : I \to \mathbb{C}$ is defined as

$$\sup \sum_{i=1}^{k-1} \big| f(t_i) - f(t_{i-1})\big|,$$

the supremum being taken over $t_1 < \cdots < t_k$, $t_i \in A$, $1 \leqslant i \leqslant k$, and $k \geqslant 2$. We write simply $\operatorname{var} f$ for $\operatorname{var}_I f$.

Proposition 4.2.

(i) *If* $f \in BV(I)$ *is a real-valued function, then*

$$\operatorname{var} U_m f \leqslant K_m \operatorname{var} f, \tag{4.9}$$

where $K_m = \frac{(m-1)(3m^2-3m+1)}{(2m-1)(m^2+m-1)}$. *The constant cannot be lowered.*

(ii) *The operator* $U_m : B(I) \to B(I)$ *is the transition operator of the Markov chain* $(s_n^a)_{n \in \mathbb{N}_+}$ *on* $(I, \mathcal{B}_I, \gamma_m^a)$, *for any* $a \in I$, *where* $(s_n^a)_{n \in \mathbb{N}_+}$ *and* γ_m^a *are give in* (3.9) *and* (3.10), *respectively.*

5. Proof of the Gauss–Kuzmin-type theorem

In this section we prove our main theorem. The main tool of this section is the random system with complete connections. We will first give a brief introduction to the theory of random systems with complete connections and list some of the main applications and some important properties. The general concepts presented here will be customized in the second subsection for the continued fraction expansion presented in this paper. All these concepts will be applied in Section 5.3 to solve our main theorem.

5.1. Random systems with complete connections

The purpose of this subsection is to recall the definition of random systems with complete connections, and take this opportunity to inform nonspecialists a little about some applications of the theory of random systems with complete connections.

The first explicit formal definition of the concept of dependence with complete connections was given by Onicescu and Mihoc in the 1930s when studying so-called *urn schemes* (see, e.g., [21], or [12] or the Introduction in [10]). The concept of random system with complete connections was defined by Iosifescu [9]. There are many other areas where the theory of RSCC can be applied. Let us just mention a few: *mathematical modeling of learning processes* (see, e.g., [20,12,14]), *chains of infinite order* (see, e.g., [6,7]), *partially observed random chains* (see, e.g., [12]), *image coding* (see [2]), and *continued fraction expansion* (see [10]). Nowadays RSCC are called *iterated functions systems with place-dependent probabilities* or simply *iterated functions systems* (IFS). This terminology was introduced by Barnsley et al. in the middle of the 1980s in [1]. It only became fashionable in the framework of fractals and chaos but, before that, it appeared as the simplest case of a random system with complete connections and, in particular, as the Bush–Mosteller model for learning with experimenter-controlled-events [see, e.g., [2,8]]. An application of IFS to continued fractions can be found in the paper [18].

5.1.1. Definitions and explanations

First, let $(W, \mathcal{W})$ and $(X, \mathcal{X})$ be two measurable spaces. A real valued function P defined on $W \times \mathcal{X}$ is called a *transition probability function* from $(W, \mathcal{W})$ to $(X, \mathcal{X})$ if $P(w, \cdot)$ is a probability on $\mathcal{X}$ for any $w \in W$ and $P(\cdot, A)$ is a $\mathcal{W}$-measurable function for any $A \in \mathcal{X}$.

A quadruple

$$\big\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\big\} \tag{5.1}$$

is named a *random system with complete connections* (RSCC) if

(i) $(W, \mathcal{W})$ and $(X, \mathcal{X})$ are measurable spaces;
(ii) $u : W \times X \to W$ is a $(\mathcal{W} \otimes \mathcal{X}, \mathcal{W})$-measurable function;
(iii) P is a transition probability function from $(W, \mathcal{W})$ to $(X, \mathcal{X})$.

The definition of an RSCC can be extended to the non-homogeneous case in the sense that all the entities constituting it are allowed to depend on $t \in T$, where T is either the set $\mathbb{N}$ of natural numbers or the set $\mathbb{Z}$ of integers.

The set W is usually called the *state space*, the set X is often called the *event space* and the function u is often called the *response-function*. We also call $u(\cdot, x) : W \to W$ a response-function.

The interpretation of this structure is as follows. If X denotes the set of possible observations and W the range of possible states of the system, then P induces for every state $w \in W$ the distribution $P(w, \cdot)$ of the random observation following w. The function u represents the transition function of the system, which transforms a given state w and an actual observation x into a new state $u(w, x)$.

To every RSCC $\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\}$ and every $w \in W$ (an arbitrary fixed element of W) one can generate two stochastic sequences $\{\xi_n\}_{n \in \mathbb{N}}$ and $\{\zeta_n\}_{n \in \mathbb{N}_+}$ as follows: we set $\xi_0 = w$, pick an element $\zeta_1 \in X$ using $P(\xi_0, \cdot)$, define $\xi_1 = u(\xi_0, \zeta_1)$, pick ζ_2 in X using $P(\xi_1, \cdot)$, define $\xi_2 = u(\xi_1, \zeta_2)$, and generally we pick ζ_n in X using $P(\xi_{n-1}, \cdot)$, and define $\xi_n = u(\xi_{n-1}, \zeta_n)$. Thus, the two stochastic sequences can be described as follows:

$$\xi_0 = w, \qquad \xi_{n+1} = u(\xi_n, \zeta_{n+1}), \quad n \geqslant 1,$$

$$P(\zeta_1 \in A) = P(w, A), \quad A \in \mathcal{X},$$

$$P(\zeta_{n+1} \in A \mid \xi_n, \zeta_n, \ldots, \xi_1, \zeta_1, \xi_0) = P(\xi_n, A), \quad A \in \mathcal{X}.$$

We call the sequence $\{\xi_n\}_{n \in \mathbb{N}}$ of W-valued random variables the *state sequence* and the sequence $\{\zeta_n\}_{n \in \mathbb{N}_+}$ of X-valued random variables the *event sequence*. When we want to emphasize the initial point w, we write

$$\xi_n = \xi_n(w) \quad \text{and} \quad \zeta_n = \zeta_n(w).$$

The central issue in the theory of dependence with complete connections is the sequence $\{\zeta_n\}_{n \in \mathbb{N}_+}$ which is a stochastic process that is no longer Markovian, but a chain with complete connections (processes whose transition probabilities depend on the whole past history).

From the definition of ξ_n it is clear that the state sequence $\{\xi_n\}_{n \in \mathbb{N}}$ is a Markov chain (the so-called associated Markov chain) with transition probability function Q, where

$$Q(w, A) = P\big(w, \big\{x \in X \bigm| u(w, x) \in A\big\}\big) \tag{5.2}$$

with $A \in \mathcal{W}$.

The transition operator $U : B(W, \mathcal{W}) \to B(W, \mathcal{W})$ is defined by

$$Uf(w) = \sum_{x \in X} P(w, x) f\big(u(w, x)\big), \quad f \in B(W, \mathcal{W}), \tag{5.3}$$

where $B(W, \mathcal{W})$ is the Banach space of all bounded $\mathcal{W}$-measurable complex-valued functions defined on W.

5.1.2. Examples of RSCCs

In this section we shall give two examples of RSCCs which occur either in various chapters of probability theory or as a result of modeling phenomena in various fields.

Example 5.1. The concept of a random system with complete connections may be regarded as a generalization and formalization of the notion of a stochastic learning model. Learning may be defined as an adaptive modification of behavior in the course of repeated trials. By mathematical learning theory we mean the body of research methods and results concerned with the conceptual representation of learning phenomena, the mathematical formulation of hypotheses about learning, and the derivation of testable theorems. The purpose of mathematical learning theory is to provide simple, quantitative descriptions of processes which are basic to behavioral modifications.

All stochastic models for learning studied so far fit the following general theoretical scheme. The behavior of the subject on trial n is determined by its state S_n (an indicator of the subject's tendencies) at the beginning of the trial. Here S_n is a random variable taking values in a measurable space $(S, \mathcal{S})$. On trial n an event E_{n+1} occurs that results in a change of the state. Here E_{n+1} is a random variable taking values in the measurable space $(E, \mathcal{E})$ and specifies those occurrences on trial n that affect the subsequent behavior. To represent the fact that the occurrence of an event affects a change of state it is necessary to consider a measurable map v from $S \times E$ into S and postulate that $S_{n+1} = v(S_n, E_{n+1})$, $n \in \mathbb{N}$. Finally assume that the probability distribution of E_{n+1} given $S_n, E_n, \ldots, S_1, E_1, S_0$ depends only on the state S_n and denote it by $R(S_n, \cdot)$. By a general learning model we mean the collection $\{(S, \mathcal{S}), (E, \mathcal{E}), v, R\}$ which is trivially an RSCC. Notice that in fact we only changed the notation. Various special learning models are obtained by simply particularizing S, E, v and R (see, e.g., [12,20]). $\square$

Example 5.2. As we mentioned in Section 1.1, any irrational number y in the unit interval $[0, 1]$ has an infinite continued fraction expansion of the form

$$y = \cfrac{1}{a_1(y) + \cfrac{1}{a_2(y) + \cfrac{1}{a_3(y) + \ddots}}},$$

where the $a_n(y)$, $n \in \mathbb{N}_+$, are natural numbers. Define $(s_n)_{n \in \mathbb{N}_+}$ by

$$s_1 = \frac{1}{a_1}, \qquad s_{n+1} = \frac{1}{s_n + a_{n+1}}, \quad n \in \mathbb{N}_+.$$

Let us consider the RSCC $\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\}$, where

$$W = [0, 1], \qquad \mathcal{W} = \mathcal{B}_{[0,1]},$$

$$X = \mathbb{N}_+, \qquad \mathcal{X} = \mathcal{P}_{\mathbb{N}_+},$$

$$u : W \times X \to W, \quad u(w, x) = \frac{1}{w + x},$$
$$P : W \times \mathcal{X} \to W, \quad P(w, x) = \frac{w + 1}{(w + x)(w + x + 1)}.$$

The sequences $(a_n)_{n\in\mathbb{N}_+}$ and $(s_n)_{n\in\mathbb{N}_+}$, $s_0 = 0$, are equivalent to the chain with complete connections $(\zeta_n)_{n\in\mathbb{N}_+}$ and the Markov chain $(\xi_n)_{n\in\mathbb{N}}$ associated with the above RSCC. More precisely, defining the one-to-one map θ from $(\mathbb{N}_+)^{\mathbb{N}_+}$ into $[0, 1]$ by

$$\theta(a_1, a_2, a_3, \ldots) = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}}, \quad a_i \in \mathbb{N}_+,\ i \in \mathbb{N}_+,$$

we have $\zeta_n(\sigma) = a_n(\theta(\sigma))$, $\xi_n(\sigma) = s_n(\theta(\sigma))$, $n \in \mathbb{N}_+$, $\sigma \in (\mathbb{N}_+)^{\mathbb{N}_+}$. $\square$

5.1.3. Properties of the associated operators

In this subsection we present the asymptotic and ergodic properties of the associated operators. These properties are used to obtain the ergodicity of an RSCC by letting the associated Markov chain satisfy some topological properties. To state these results we need some preliminary definitions.

Let Q_n be the transition probability function defined by

$$Q_n(w, A) = \frac{1}{n}\sum_{k=1}^{n} Q^k(w, A)$$

where Q^k, $k \geqslant 1$, is the k-step transition probability function of the Markov chain associated with RSCC (5.1). Let U_n be the Markov operator associated with Q_n.

Next, let us consider the norm $\|\cdot\|_L$ defined on $L(W) =$ the space of Lipschitz complex-valued functions defined on W by

$$\|f\|_L = \sup_{w\in W}\bigl|f(w)\bigr| + \sup_{w'\neq w''}\frac{|f(w') - f(w'')|}{|w' - w''|}, \quad f \in L(W).$$

As is well known, $(L(W), \|\cdot\|_L)$ is a Banach space.

The following can be found in [10].

If there exists a linear bounded operator U^∞ from $L(W)$ to $L(W)$ such that

$$\lim_{n\to\infty}\bigl\|U_n f - U^\infty f\bigr\|_L = 0,$$

for any $f \in L(W)$ with $\|f\|_L = 1$, we say U is *ordered*.

If

$$\lim_{n\to\infty}\bigl\|U^n f - U^\infty f\bigr\|_L = 0,$$

for any $f \in L(W)$ with $\|f\|_L = 1$, we say U is *aperiodic*, where U^n is the nth iterate of U, $n \in \mathbb{N}$, with U^0 is the identity.

If U is ordered and $U^\infty(L(W))$ is one-dimensional space, it is named *ergodic* with respect to $L(W)$.

If U is ergodic and aperiodic, it is named *regular* with respect to $L(W)$ and the corresponding Markov chain has the same name.

The definition below is due to M.F. Norman [20] and isolates a class of RSCCs, called *RSCCs with contraction*.

An RSCC $\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\}$ is said to be an RSCC *with contraction* if and only if there is a distance d on W and the metric space (W, d) is separable, $r_1 < \infty$, $R_1 < \infty$, and there exists a natural integer k such that $r_k < 1$, where

$$r_k = \sup_{w' \neq w''} \sum_{X^k} P_k\big(w, x^{(k)}\big) \frac{\mathrm{d}(w'x^{(k)}, w''x^{(k)})}{\mathrm{d}(w', w'')}, \quad k \in \mathbb{N}_+,$$

and

$$R_k = \sup_{A \in \mathcal{X}^k} \sup_{w' \neq w''} \frac{P_k(w', A) - P_k(w'', A)}{\mathrm{d}(w', w'')}.$$

The following result can be found in [10].

Theorem 5.3. *Let W be a compact metric space with a distance* d *and $\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\}$ be an RSCC with contraction.*

(i) *The Markov chain associated to the RSCC is regular if and only if there exists a point $w_0 \in W$ such that*

$$\lim_{n \to \infty} \mathrm{d}\big(\sigma_n(w), w_0\big) = 0,$$

for any $w \in W$, where $\sigma_n(w) = \operatorname{supp} Q^n(w, \cdot)$ ($\operatorname{supp} \mu$ denotes the support of the measure μ).

(ii) *The supports of $Q^n(w, \cdot)$, $n \in \mathbb{N}_+$, $w \in W$, can be iteratively computed as follows:*

$$\sigma_{m+n}(w) = \overline{\bigcup_{w' \in \sigma_m(w)} \sigma_n\big(w'\big)},$$

for any $m, n \in \mathbb{N}_+$, $w \in W$, where the overline means the topological closure.

An RSCC $\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\}$, whose associated Markov chain is regular with respect to $B((W, \mathcal{W}))$, is *uniformly ergodic* and $\lim_{n \to \infty} \varepsilon_n = 0$, where

$$\varepsilon_n := \sup_{\substack{w \in W, r \in \mathbb{N}_+ \\ A \in \mathcal{X}^r}} \big| P_r^n(w, A) - \mathrm{P}_r^\infty(A) \big|,$$

while P_r^∞ is the probability on $\mathcal{X}^r$.

Theorem 5.4. *Let W be a compact metric space with a distance* d. *If the RSCC $\{(W, \mathcal{W}), (X, \mathcal{X}), u, P\}$ with contraction has regular associated Markov chain, then it is uniformly ergodic.*

5.2. The RSCC associated with expansion of the type of (1.1)

First, it is easy to check that P_m^i from (2.28) defines a transition probability function from $(I, \mathcal{B}_I)$ to $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, i.e., $\sum_{i \in \mathbb{N}} P_m^i(x) = 1$, $x \in I$.

Let us consider the random system with complete connections

$$\big\{(I, \mathcal{B}_I), \big(\mathbb{N}_+, \mathcal{P}(\mathbb{N}_+)\big), u, P\big\}, \tag{5.4}$$

where $u : I \times \mathbb{N} \to I$, $u(x, i) = u_m^i(x)$ is given in (4.4) and the function $P(x, i) = P_m^i(x)$ given in (2.28).

We denote by U_m the associated Markov operator of RSCC (5.4) with the transition probability function

$$Q_m(x, A) = \sum_{\{i \in \mathbb{N}:\ u_m^i(x) \in A\}} P_m^i(x), \quad x \in I,\ A \in \mathcal{B}_I.$$

Then $Q_m^n(\cdot,\cdot)$ will denote the n-step transition probability function of the same Markov chain.

The ergodic behavior of RSCC (5.4) allows us to find the limiting distribution function F and the invariant measure Q_m^∞ induced by F.

Proposition 5.5. *RSCC (5.4) is uniformly ergodic.*

Proof. We apply Theorem 5.4. Putting $\Delta_i = m^{-i} - m^{-2i}$, $i \in \mathbb{N}$, we get

$$P_m^i(x) = (m-1)\left[m^{-(i+1)} + \frac{\Delta_i}{x + (m-1)m^{-i} + 1} - \frac{\Delta_{i+1}}{x + (m-1)m^{-(i+1)} + 1}\right].$$

We have

$$\begin{aligned}
&\frac{d}{dx} u(x,i) = -\frac{(m-1)m^{-i}}{((m-1)x+1)^2},\\
&\frac{d}{dx} P(x,i) = (m-1)\left[\frac{\Delta_{i+1}}{(x + (m-1)m^{-(i+1)} + 1)^2} - \frac{\Delta_i}{(x + (m-1)m^{-i} + 1)^2}\right],
\end{aligned}$$

for all $x \in I$ and $i \in \mathbb{N}$, so that $\sup_{x \in I} |\frac{d}{dx} u(x,i)| = (m-1)m^{-i}$ and $\sup_{x \in I} |\frac{d}{dx} P(x,i)| < \infty$. Hence the requirements of definition of an RSCC with contraction are fulfilled. To prove the regularity of U with respect to $L(I)$ let us define recursively $x_{n+1} = (x_n + 2)^{-1}$, $n \in \mathbb{N}$, with $x_0 = x$.

A criterion of regularity is expressed in Theorem 5.3(i), in terms of supports $\sigma_n(x)$ of the n-step transition probability functions $Q_m^n(x,\cdot)$, $n \in \mathbb{N}_+$. Clearly $x_{n+1} \in \sigma_1(x_n)$ and therefore Theorem 5.3(ii) and an induction argument lead us to the conclusion that $x_n \in \sigma_n(x)$, $n \in \mathbb{N}_+$. But, $\lim_{n\to\infty} x_n = \sqrt{2} - 1$ for any $x \in I$. Hence

$$\mathrm{d}\big(\sigma_n(x), \sqrt{2} - 1\big) \leqslant |x_n - \sqrt{2} + 1| \to 0 \quad \text{as } n \to \infty,$$

where $\mathrm{d}(x,y) = |x-y|$, for any $x, y \in I$. The regularity of U_m with respect to $L(I)$ follows from Theorem 5.3. Moreover, $Q_m^n(\cdot,\cdot)$ converges uniformly to a probability measure Q_m^∞ and there exist two positive constants $q < 1$ and k such that

$$\left\| U_m^n f - U_m^\infty f \right\|_L \leqslant k q^n \|f\|_L, \quad n \in \mathbb{N}_+,\ f \in L(I), \tag{5.5}$$

where

$$U_m^n f(\cdot) = \int_I f(y) Q_m^n(\cdot, dy), \tag{5.6}$$

$$U_m^\infty f = \int_I f(y) \mathrm{Q}_m^\infty(dy), \tag{5.7}$$

and Q_m^∞ is the invariant probability measure of the transformation τ_m, i.e., Q_m^∞ has the density $\rho_m(x)$ given in (2.29), $x \in I$. □

Now we are able to find the limiting distribution function

$$F(x) = F_\infty(x) = \lim_{n\to\infty} \mu\left(\tau_m^n < x\right)$$

and obtain a convergence rate result.

5.3. Proof of Theorem 1.1

We prove Theorem 1.1 in this subsection.

Proof of Theorem 1.1. By (5.7) we have

$$U_m^\infty f_0 = \int_I f_0(y) \mathrm{Q}_m^\infty(dx) = k_m, \quad f_0 \in L(I).$$

Taking into account (5.5), there exist two constants $q < 1$ and k such that

$$\left\| U_m^n f_0 - U_m^\infty f_0 \right\|_L \leqslant k q^n \| f_0 \|_L, \quad n \in \mathbb{N}_+.$$

Further, consider $C(I)$ the metric space of real continuous functions defined on I with the supremum norm $\| f \| = \sup_{x \in I} |f(x)|$. Since $L(I)$ is a dense subset of $C(I)$ we have

$$\lim_{n\to\infty} \left\| \left(U_m^n - U_m^\infty\right) f_0 \right\| = 0, \tag{5.8}$$

for all $f_0 \in C(I)$. Therefore, (5.8) is valid for a measurable function f_0 which is Q_m^∞-almost surely continuous, that is, for a Riemann-integrable function f_0. Thus, we have

$$\begin{aligned} F(x) = \lim_{n\to\infty} \mu\left(\tau_m^n < x\right) &= \lim_{n\to\infty} \int_0^x U_m^n f_0(u) \rho_m(u)\, du \\ &= k_m \int_0^x \rho_m(u)\, du \\ &= \frac{k_m}{(m-1)^2} \log \frac{m((m-1)x+1)}{(m-1)x+m}. \end{aligned}$$

Hence (1.6) is proved. □

Acknowledgments

The author would like to express their sincere thanks to the referees for their valuable comments.

Appendix A. Proofs of propositions

We prove Propositions 2.5, 3.1, 4.1 and 4.2 in this appendix.

Proof of Proposition 2.5. We briefly give some general properties about the Perron–Frobenius operator (see, e.g., [4,11]) which will be useful both to demonstrate this proposition and in Section 4.

Let $(X, \mathcal{X}, \mu)$ be a probability space. A transformation τ of X is said to be *μ-non-singular* if and only if $\mu(\tau^{-1}(A)) = 0$ for all $A \in \mathcal{X}$ for which $\mu(A) = 0$; it is said to be *measure-preserving* if and only if $\mu\tau^{-1} = \mu$, i.e., $\mu\tau^{-1}(A) = \mu(A)$ for all $A \in \mathcal{X}$. Clearly, any μ-preserving transformation is μ-non-singular.

The Perron–Frobenius operator P_μ associated with a μ-non-singular transformation τ is defined as the linear bounded operator on $L^1_\mu = \{f : I \to \mathbb{C}\colon \int_I |f|\, d\mu < \infty\}$ which takes $f \in L^1_\mu$ into $P_\mu f \in L^1_\mu$ with

$$\int_A P_\mu f\, d\mu = \int_{\tau^{-1}(A)} f\, d\mu, \quad A \in \mathcal{X},$$

or, equivalently

$$\int_X g P_\mu f\, d\mu = \int_X (g \circ \tau) f\, d\mu$$

for all $f \in L^1_\mu$ and $g \in L^\infty_\mu$.

In particular, the Perron–Frobenius operator P_λ of τ under the Lebesgue measure λ is (see [4, p. 86])

$$P_\lambda f(x) = \frac{d}{dx} \int_{\tau^{-1}([0,x])} f\, d\lambda = \sum_{t \in \tau^{-1}(x)} \frac{f(t)}{|\tau'(t)|} \quad \text{a.e. in } I. \tag{A.1}$$

The probabilistic interpretation of P_μ is immediate: if an X-valued random variable ξ on X has μ-density h, that is, $\mu(\xi \in A) = \int_A h\, d\mu$, $A \in X$, with $h \geqslant 0$ and $\int_X h\, d\mu = 1$, then $\tau \circ \xi$ has μ-density $P_\mu h$. The following properties hold:

(i) P_μ is positive, that is, $P_\mu f \geqslant 0$ if $f \geqslant 0$;
(ii) P_μ preserves integrals, that is, $\int_X P_\mu f\, d\mu = \int_X f\, d\mu$, $f \in L^1_\mu$;
(iii) $\|P_\mu\|_{p,\mu} := \sup(\|P_\mu f\|_{p,\mu}\colon\ f \in L^p_\mu,\ \|f\|_{p,\mu} = 1) \leqslant 1$ for any $p \geqslant 1$ and $p = \infty$;
(iv) for any $n \in \mathbb{N}_+$ the nth power P^n_μ of P_μ is the Perron–Frobenius operator associated with the nth iterate τ^n of τ under μ;
(v) $(P_\mu f)^* = P_\mu f^*$ for any $f \in L^1_\mu$, where $z^* =$ complex conjugate of $z \in \mathbb{C}$ ($=$ the set of complex numbers);
(vi) $P_\mu((g \circ \tau) f) = g P_\mu f$ for any $f \in L^1_\mu$ and $g \in L^\infty_\mu$;
(vii) $P_\mu f = f$ if and only if τ is ν-preserving, where ν is defined by $\nu(A) = \int_A f\, d\mu$, $A \in \mathcal{X}$. In particular, $P_\mu 1 = 1$ if and only if τ is μ-preserving.

Proof of Proposition 2.5 From above, it is sufficient to show that the function ρ_m defined in (2.29) is an eigenfunction of the Perron–Frobenius operator of τ_m with the eigenvalue 1:

$$P_{\tau_m} \rho_m(x) = \sum_{t \in \tau_m^{-1}(x)} \frac{\rho_m(t)}{|\tau'_m(t)|}. \tag{A.2}$$

First, we note that

$$\tau_m^{-1}(x) = \left\{ \frac{m^{-i}}{1 + (m-1)x}\colon i \geqslant 1,\ x \in I \right\}. \tag{A.3}$$

Thus

$$\begin{aligned}
P_{\tau_m}\rho_m(x) &= \sum_{i=0}^{\infty}\frac{(m-1)m^{-i}}{(1+(m-1)x)^2}\rho_m\left(\frac{m^{-i}}{1+(m-1)x}\right)\\
&= \sum_{i=0}^{\infty}(m-1)m^{-(i+1)}\frac{1}{((m-1)x+(m-1)m^{-(i+1)}+1)}\frac{1}{((m-1)x+(m-1)m^{-i}+1)}\\
&= \frac{1}{m-1}\sum_{i=0}^{\infty}\left(\frac{1}{(m-1)x+(m-1)m^{-(i+1)}+1}-\frac{1}{(m-1)x+(m-1)m^{-i}+1}\right)\\
&= \frac{1}{m-1}\left(\frac{1}{(m-1)x+1}-\frac{1}{(m-1)x+m}\right)\\
&= \frac{1}{((m-1)x+1)((m-1)x+m)} = \rho_m(x). \qquad \square
\end{aligned}$$

Proof of Proposition 3.1. We should show that $\overline{\gamma}_m(\overline{\tau}_m^{-1}(B)) = \overline{\gamma}_m(B)$ for any $B \in \mathcal{B}_I^2$ or, equivalently, since $\overline{\tau}_m$ is invertible on Ω^2, that

$$\overline{\gamma}_m\big(\overline{\tau}_m(B)\big) = \overline{\gamma}_m(B), \quad \text{for any } B \in \mathcal{B}_I^2. \tag{A.4}$$

We start with $B = (a,b)\times(c,d)$, where

$$a = m^{-(i+1)}, \qquad b = m^{-i}, \quad i \in \mathbb{N},$$

and c and d are the arbitrary numbers from $(0,1)$. Then

$$\overline{\tau}_m(B) = \left\{\left(\tau_m(x), \frac{m^{-a_1(x)}}{(m-1)y+1}\right) \,\middle|\, x \in (a,b),\ y \in (c,d)\right\}. \tag{A.5}$$

Taking $x = m^{-(i+\theta)}$, $0 < \theta < 1$, we have

$$\tau_m(x) = \frac{m^{\theta}-1}{m-1}, \qquad a_1(x) = i$$

such that

$$\overline{\tau}_m(B) = \left((0,1), \left(\frac{m^{-i}}{(m-1)d+1}, \frac{m^{-i}}{(m-1)c+1}\right)\right). \tag{A.6}$$

Let

$$I(m,i,c,d) \equiv \left(\frac{m^{-i}}{(m-1)d+1}, \frac{m^{-i}}{(m-1)c+1}\right).$$

A simple computation yields

$$\overline{\gamma}_m\big(\overline{\tau}_m(B)\big) = k_m \int_0^1 dx \int_{I(m,i,c,d)} \frac{dy}{((m-1)(x+y)+1)^2}$$

$$= k_m \int_{m^{-(i+1)}}^{m^{-i}} dx \int_c^d \frac{dy}{((m-1)(x+y)+1)^2} = \overline{\gamma}_m(B)$$

that is, (A.4) holds.

Next, we consider the case

$$a = \frac{m^{-i}}{(m-1)m^{-j}+1}, \qquad b = \frac{m^{-i}}{(m-1)m^{-(j+1)}+1}, \quad i, j \in \mathbb{N},$$

and (c, d) is an arbitrary interval. Now, with

$$x = \frac{m^{-i}}{(m-1)m^{-(j+\theta)}+1},$$

we have

$$\left\{\frac{\log x^{-1}}{\log m}\right\} = \left\{i + \frac{\log(1+(m-1)m^{-(j+\theta)})}{\log m}\right\} = \frac{\log(1+(m-1)m^{-(j+\theta)})}{\log m}$$

and

$$a_1(x) = \left\lfloor \frac{\log x^{-1}}{\log m} \right\rfloor = i.$$

Thus,

$$(m-1)\tau_m(x) = m^{\frac{\log(1+(m-1)m^{-(j+\theta)})}{\log m}} - 1 = (m-1)m^{-(j+\theta)}.$$

Hence,

$$\overline{\tau}_m(B) = \big(m^{-(j+1)}, m^{-j}\big) \times \left(\frac{m^{-i}}{(m-1)d+1}, \frac{m^{-i}}{(m-1)c+1}\right). \tag{A.7}$$

A straightforward calculation shows us that

$$\overline{\gamma}_m\big(\overline{\tau}_m(B)\big) = k_m \int_{m^{-(j+1)}}^{m^{-j}} dx \int_{I(m,i,c,d)} \frac{dy}{((m-1)(x+y)+1)^2}$$

$$= k_m \int_{I(m,i,m^{-j},m^{-(j+1)})} dx \int_c^d \frac{dy}{((m-1)(x+y)+1)^2} = \overline{\gamma}_m(B)$$

that is, (A.4) holds.

Since any arbitrary interval (a,b) can be written as a reunion of fundamental intervals the proof is complete. $\square$

Proof of Proposition 4.1. (i) Let $\tau_{m,i} : I_i \to I$ denote the restriction of τ_m to the interval $I_i = (m^{-(i+1)}, m^{-i}]$, $i \in \mathbb{N}$, that is,

$$\tau_{m,i}(x) = \frac{1}{m-1}\left(\frac{m^{-i}}{x} - 1\right), \quad x \in I_i. \tag{A.8}$$

For any $f \in L^1_{\gamma_m}$ and any $A \in \mathcal{B}_I$, we have

$$\int\limits_{\tau_m^{-1}(A)} f\, d\gamma_m = \sum_{i\in\mathbb{N}} \int\limits_{\tau_m^{-1}(A\cap I_i)} f\, d\gamma_m = \sum_{i\in\mathbb{N}} \int\limits_{\tau_{m,i}^{-1}(A)} f\, d\gamma_m. \tag{A.9}$$

For any $i \in \mathbb{N}$, by the change of variable

$$x = \tau_{m,i}^{-1}(y) = \frac{m^{-i}}{(m-1)y+1}, \tag{A.10}$$

we successively obtain

$$\begin{aligned}
\int\limits_{\tau_{m,i}^{-1}(A)} f\, d\gamma_m &= k_m \int\limits_{\tau_{m,i}^{-1}(A)} \frac{f(x)}{((m-1)x+1)((m-1)x+m)}\, dx \\
&= k_m \int\limits_A \frac{f(u_m^i(y))}{((m-1)u_m^i(y)+1)((m-1)u_m^i(y)+m)} \frac{(m-1)m^{-i}}{((m-1)y+1)^2}\, dy \\
&= k_m \int\limits_A f\big(u_m^i(y)\big)(m-1)m^{-(i+1)} \frac{1}{((m-1)y+(m-1)m^{-i}+1)} \\
&\quad \times \frac{1}{((m-1)y+(m-1)m^{-(i+1)}+1)}\, dy \\
&= \int\limits_A P_m^i\big((m-1)y\big) f\big(u_m^i(y)\big)\gamma_m(dy).
\end{aligned} \tag{A.11}$$

Now, (4.3) follows from (A.9) and (A.11). $\square$

(ii)(a) From (A.8) and (A.10), for any $f \in L^1_{\gamma_m}$ and any $A \in \mathcal{B}_I$, we have

$$\begin{aligned}
\int\limits_{\tau_m^{-1}(A)} f\, d\mu &= \sum_{i\in\mathbb{N}} \int\limits_{\tau_m^{-1}(A\cap I_i)} f\, d\mu = \sum_{i\in\mathbb{N}} \int\limits_{\tau_{m,i}^{-1}(A)} f\, d\mu \\
&= \sum_{i\in\mathbf{N}} \int\limits_{\tau_{m,i}^{-1}(A)} f(x)h(x)\, dx = \sum_{i\in\mathbf{N}} \int\limits_A \frac{f(u_m^i(y))h(u_m^i(y))(m-1)m^{-i}}{((m-1)y+1)^2}\, dy \\
&= \int\limits_A \sum_{i\in\mathbf{N}} \frac{h(u_m^i(x))}{((m-1)x+1)^2}(m-1)m^{-i} f\big(u_m^i(x)\big)\, dx.
\end{aligned} \tag{A.12}$$

Since $d\mu = h\,d\lambda$, (4.5) follows from (A.12). Now, since $g(x) = ((m-1)x+1)((m-1)x+m)f(x)h(x)$, from (4.3) we have

$$U_m g(x) = \frac{((m-1)x+m)}{(m-1)x+1}(m-1)\sum_{i\in\mathbb{N}} m^{-i} h\big(u_m^i(x)\big) f\big(u_m^i(x)\big). \tag{A.13}$$

Now, (4.6) follows immediately from (4.5) and (A.13). □

(b) We will use mathematical induction. For $n=0$, Eq. (4.8) reduces to

$$\mu(A) = \int_A h(x)\,dx, \quad A \in \mathcal{B}_I,$$

which is obviously true. Assume that (4.8) holds for some $n \in \mathbb{N}$. Then

$$\begin{aligned}
\mu\big(\tau_m^{-(n+1)}(A)\big) &= \mu\big(\tau_m^{-n}\big(\tau_m^{-1}(A)\big)\big) \\
&= \int_{\tau_m^{-1}(A)} \frac{U_m^n f(x)}{((m-1)x+1)((m-1)x+m)}\,dx \\
&= \frac{1}{k_m}\int_{\tau_m^{-1}(A)} U_m^n f(x)\,d\gamma_m(x).
\end{aligned}$$

By the very definition of the Perron–Frobenius operator $U_m = P_{\gamma_m}$ we have

$$\int_{\tau_m^{-1}(A)} U_m^n f\,d\gamma_m = \int_A U_m^{n+1} f\,d\gamma_m.$$

Therefore,

$$\begin{aligned}
\mu\big(\tau_m^{-(n+1)}(A)\big) &= \frac{1}{k_m}\int_A U_m^{n+1} f\,d\gamma_m \\
&= \int_A \frac{U_m^{n+1} f(x)}{((m-1)x+1)((m-1)x+m)}\,dx
\end{aligned}$$

which ends the proof. □

Proof of Proposition 4.2. (i) For $x, y \in I$ we have

$$\begin{aligned}
U_m f(x) - U_m f(y) &= \sum_{i\in\mathbb{N}} \big(P_m^i\big((m-1)x\big) f\big(u_m^i(x)\big) - P_m^i\big((m-1)y\big) f\big(u_m^i(y)\big)\big) \\
&= \sum_{i\in\mathbb{N}} \big(P_m^i\big((m-1)x\big) - P_m^i\big((m-1)y\big)\big)\big(f\big(u_m^i(x)\big) - f\big(u_m^0(x)\big)\big) \\
&\quad + \sum_{i\in\mathbb{N}} P_m^i\big((m-1)y\big)\big(f\big(u_m^i(x)\big) - f\big(u_m^i(y)\big)\big)
\end{aligned}$$

$$= \sum_{i\in\mathbb{N}_+} \big(P_m^i\big((m-1)x\big) - P_m^i\big((m-1)y\big)\big)\big(f\big(u_m^i(x)\big) - f\big(u_m^0(x)\big)\big)$$
$$+ \sum_{i\in\mathbb{N}} P_m^i\big((m-1)y\big)\big(f\big(u_m^i(x)\big) - f\big(u_m^i(y)\big)\big).$$

Note that the function P_m^0 is increasing, while the functions P_m^i, $i \in \mathbb{N}_+$, are all decreasing. Let $x < y$, with $x, y \in I$. It follows from the above equation that

$$\big|U_m f(x) - U_m f(y)\big| \leqslant \left(\sum_{i\in\mathbb{N}_+} \big(P_m^i\big((m-1)x\big) - P_m^i\big((m-1)y\big)\big)\right) \operatorname{var} f$$
$$+ \sup_{y\in I, i\in\mathbb{N}} P_m^i\big((m-1)y\big) \sum_{i\in\mathbb{N}} \operatorname{var}_{[x,y]} f \circ u_i(x)$$
$$= \big(1 - P_m^0\big((m-1)x\big) - 1 + P_m^0\big((m-1)y\big)\big) \operatorname{var} f$$
$$+ P_m^0(m-1) \sum_{i\in\mathbb{N}} \operatorname{var}_{[x,y]} f \circ u_i(x).$$

Hence

$$\operatorname{var} U_m f \leqslant \big(2P_m^0(m-1) - P_m^0(0)\big) \operatorname{var} f = \left(\frac{2m(m-1)}{m^2+m-1} - \frac{m-1}{2m-1}\right) \operatorname{var} f$$
$$= \frac{(m-1)(3m^2-3m+1)}{(2m-1)(m^2+m-1)} \operatorname{var} f.$$

Define f by $f(x) = 0$, $0 \leqslant x \leqslant \frac{1}{m}$, and $f(x) = 1$, $\frac{1}{m} < x \leqslant 1$. Then we have $U_m f(x) = P_m^0(x)$, $0 \leqslant x < 1$ and $U_m f(1) = 0$. Since $\operatorname{var} U_m f = \frac{(m-1)(3m^2-3m+1)}{(2m-1)(m^2+m-1)}$ and $\operatorname{var} f = 1$, it follows that the constant K_m cannot be lowered. □

(ii) The transition operator of $(s_n^a)_{n\in\mathbb{N}_+}$ takes $f \in B(I)$ to the function defined by

$$E_a\big(f\big(s_{n+1}^a\big) \,\big|\, s_n^a = s\big) = \sum_{i\in\mathbb{N}} P_m^i\big((m-1)s\big) f\big(u_m^i(s)\big)$$
$$= U_m f(s), \quad s \in I, \tag{A.14}$$

where E_a stands for the mean value operator with respect to the probability measure γ_m^a. □

References

[1] M. Barnsley, S. Demko, J. Elton, J. Gerinomo, Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities, Ann. Inst. H. Poincaré Probab. Statist. 24 (3) (1988) 367–394.

[2] M. Barnsley, J. Elton, A new class of Markov processes for image encoding, Adv. in Appl. Probab. 20 (1988) 14–32.

[3] W. Bosma, C. Kraaikamp, Metrical theory for optimal continued fractions, J. Number Theory 34 (1990) 251–270.

[4] A. Boyarsky, P. Góra, Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension, Birkhäuser, Boston, 1997.

[5] H.-C. Chan, The asymptotic growth rate of random Fibonacci type sequences. II, Fibonacci Quart. 44 (2006) 73–84.

[6] W. Doeblin, R. Fortet, Sur des chaînes à liaisons complètes, Bull. Soc. Math. France 65 (1937) 132–148.

[7] T.E. Harris, On chains of infinite order, Pacific J. Math. 5 (1955) 707–724.

[8] U. Herkenrath, M. Iosifescu, A. Rudolph, Random systems with complete connections and iterated function systems, Math. Rep. 5 (55) (2003) 127–140.

[9] M. Iosifescu, Random systems with complete connections with an arbitrary set of states, Rev. Roumaine Math. Pures Appl. 8 (1963) 611–645.

[10] M. Iosifescu, S. Grigorescu, Dependence with Complete Connections and Its Applications, Cambridge Tracts in Math., vol. 96, Cambridge Univ. Press, Cambridge, 1990; second printing slightly corrected, 2009.
[11] M. Iosifescu, C. Kraaikamp, Metrical Theory of Continued Fractions, Kluwer Academic Publisher, Dordrecht, 2002.
[12] M. Iosifescu, R. Theodorescu, Random Processes and Learning, Springer-Verlag, Berlin, 1969.
[13] S. Kalpazidou, On a problem of Gauss–Kuzmin type for continued fraction with odd partial quotients, Pacific J. Math. 123 (1) (1986) 103–114.
[14] S. Karlin, Some random walks arising in learning models. I, Pacific J. Math. 3 (4) (1953) 725–756.
[15] A.Ya. Khinchin, Continued Fractions, Univ. Chicago Press, Chicago, 1964 (translation of the 3rd Russian edition, 1961).
[16] R.O. Kuzmin, On a problem of Gauss, Dokl. Akad. Nauk SSSR Ser. A (1928) 375–380 (in Russian); French version in: Atti Congr. Internaz. Mat., vol. VI, Bologna, 1928, Zanichelli, Bologna, 1932, pp. 83–89.
[17] P. Lévy, Sur les lois de probabilité dont dépendent les quotients complets et incomplets d'une fraction continue, Bull. Soc. Math. France 57 (1929) 178–194.
[18] R.D. Mauldin, M. Urbański, The doubling property of conformal measures of infinite iterated function systems, J. Number Theory 102 (1) (2003) 23–40.
[19] H. Nakada, Metrical theory for a class of continued fraction transformations and their natural extensions, Tokyo J. Math. 4 (2) (1981) 399–426.
[20] M.F. Norman, Markov Processes and Learning Models, Academic Press, New York, 1972.
[21] O. Onicescu, Gh. Mihoc, Sur les chaînes de variables statistiques, Bull. Sci. Math. 59 (1935) 174–192.
[22] G.I. Sebe, On convergence rate in the Gauss–Kuzmin problem for the grotesque continued fractions, Monatsh. Math. 133 (3) (2001) 241–254.
[23] G.I. Sebe, A Gauss–Kuzmin theorem for the Rosen fractions, J. Théor. Nombres Bordeaux 14 (2) (2002) 667–682.
[24] G.I. Sebe, On a Gauss–Kuzmin-type problem for a new continued fraction expansion with explicit invariant measure, in: Proc. of the 3rd Int. Coll. "Math. in Engg. and Numerical Physics" (MENP-3), Bucharest, Romania, 7–9 October 2004, in: BSG Proc., vol. 12, Geometry Balkan Press, 2005, pp. 252–258.
[25] D. Viswanath, Random Fibonacci sequences and the number 1.13198824 . . . , Math. Comp. 69 (231) (2000) 1131–1155.